TOWARDS A MATHEMATICAL MODEL OF TONALITY

by

Elaine Chew

S.M., Operations Research, 1998
Massachusetts Institute of Technology
B.A.S., Mathematical and Computational Sciences,
and Music, 1992, Stanford University

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2000

Signature of Author ....
Operations Research Center
January 7, 2000

Certified by ....
Jeanne S. Bamberger
Professor, Music and Urban Education
Thesis Supervisor

Certified by ....
Georgia Perakis
Assistant Professor, Operations Research
Thesis Advisor

Accepted by ....
Cynthia Barnhart
Co-director, Operations Research Center

Towards a Mathematical Model of Tonality

by
Elaine Chew

Submitted to the Sloan School of Management
on January 7, 2000, in partial fulfillment of the degree of
Doctor of Philosophy

Abstract

This dissertation addresses the question of how musical pitches generate a tonal center. Being able to characterize the relationships that generate a tonal center is crucial to the computer analysis and the generating of western tonal music. It also can inform issues of compositional styles, structural boundaries, and performance decisions.

The proposed Spiral Array model offers a parsimonious description of the inter-relations among tonal elements, and suggests new ways to re-conceptualize and reorganize musical information. The Spiral Array generates representations for pitches, intervals, chords and keys within a single spatial framework, allowing comparisons among elements from different hierarchical levels. Structurally, this spatial representation is a helical realization of the harmonic network (*tonnetz*). The basic idea behind the Spiral Array is the representation of higher level tonal elements as composites of their lower level parts.

The Spiral Array assigns greatest prominence to perfect fifth and major/minor third interval relations, placing elements related by these intervals in proximity to each other. As a result, distances between tonal entities as represented spatially in the model correspond to perceived distances among sounding entities. The parameter values that affect proximity relations are prescribed based on a few perceived relations among pitches, intervals, chords and keys. This process of interfacing between the model and actual perception creates the opportunity to research some basic, but till now unanswered questions about the relationships that generate tonality.

A generative model, the Spiral Array also provides a framework on which to design viable and efficient algorithms for problems in music cognition. I demonstrate its versatility by applying the model to three different problems: I develop an algorithm to determine the key of musical passages that, on average, performs better than existing ones when applied to the 24 fugue subjects in Book 1 of Bach's WTC; I propose the first computationally viable method for determining modulations (the change of key); and, I design a basic algorithm for finding the roots of chords, comparing its results to those of algorithms by other researchers. All three algorithms were implemented in Matlab.

Thesis Supervisor: Jeanne S. Bamberger, Professor of Music and Urban Education
Thesis Advisor: Georgia Perakis, Assistant Professor of Operations Research

Biographical Note

Born 1970 in Buffalo, NY, Elaine Chew grew up in Singapore. There, she attended Hwa Chong Junior College (1987-1988), where she completed her 'A' Levels; and, the Singapore Chinese Girls' School (1977-1986), where she completed her 'O' Levels. She studied piano privately with Goh Lee Choo and with Ong Lip Tat. Her studies with Ong prepared her for several national competitions, and for the FTCL (Fellow of Trinity College, London) and LTCL (Licentiate of Trinity College, London) performance diplomas.

She returned to the United States to attend Stanford University, where she graduated with a B.A.S. in Mathematical and Computational Sciences (honors) and Music (distinction) in 1992. At Stanford, she obtained a Summer Undergraduate Research Fellowship (SURF) and worked on the Von-Neumann Center of Gravity Algorithm with Professor George Dantzig. She prepared her senior recital under the tutelage of James Goldsworthy, presenting piano music by Haydn, Rachmaninoff, Debussy and Barber.

At MIT since 1992, she completed an S.M. in Computational Finance (1998) with Professor Dimitris Bertsimas; and has since worked with Professor Jeanne Bamberger on research pertaining to this dissertation. An ONR Graduate Fellowship funded her first few years at MIT, and a Josephine de Karman Dissertation Fellowship made the writing of this thesis possible. Coached by pianist David Deveau, violist Marcus Thompson, composer John Harbison and others, she has presented numerous solo and chamber music concerts at Killian Hall under the auspices of MIT's Chamber Music Society and Advanced Music Performance program. She has performed at the Rockport Chamber Music Festival (1997, 1999), the Embassy Series in Washington D.C. (1999) and at the President's Charity Concert (1998) in Singapore, amongst others. A proponent of new music, she has frequently championed the works of contemporary Chinese composers.

Acknowledgements

The reader is forewarned that seven and a half years in one place generates a lot of debt. Thus, these acknowledgements comprise a significant part of the thesis.

The work in this thesis was supported by a variety of funding sources. An ONR Graduate Fellowship funded me through the first four years of graduate school. A Research Assistantship provided by Richard Larson at MIT's Center for Advanced Educational Services allowed me to develop the ideas for this thesis; and, a Josephine de Karman Dissertation Fellowship afforded me the freedom to devote my time entirely to thesis-writing during my last semester.

In the interim, a variety of Teaching Assistantships have supported me through my years of exploration at MIT. These include: Statistical Analysis for Technology Managers with Arnold Barnett; Urban Operations Research with Richard Larson, Amadeo Odoni and Arnold Barnett; Engineering Risk Benefit Analysis with Amadeo Odoni and George Apostolakis; Data, Models and Decisions with Nitin Patel; and a Trading Room TA with Andrew Lo.

External to academic employment, a number of patrons have provided a creative outlet by allowing me (almost) free reign in designing their websites: the MIT-WHOI Joint Program through Ronni Schwartz; the Operations Research Center through Paulette Mosley and Laura Rose; Andrew Lo, and Peter Child. Andrew Lo provided the machine (a PowerMac 7600, this was top-of-the-line at the time I acquired it) that generated most of the graphics in this thesis.

Many faculty members have influenced my academic development at MIT. Before I settled on this Ph.D. topic, I worked with three professors at the Operations Research Center. In chronological order, they include: Robert Freund, who guided my research in Interior Point Methods during my first year, and continues to encourage my teaching endeavors; James Orlin, whose attempt to bridge the gap between Operations Research and Biology through the Genome Project continues to serve as an inspiration; Dimitris Bertsimas, who believed in me and took me on as a student in Computational Finance. I gained valuable research experience with all three faculty members, and a Masters degree through my work with Professor Bertsimas. In addition, Dimitris and Georgia have always extended their hearty hospitality to their students during the Thanksgiving holidays, where a scrumptious meal at their home always ended with a game of charades.

This doctoral dissertation would not have been possible without the sustained support and encouragement of my thesis supervisor, Jeanne Bamberger. Before my initial conversations with Jeanne, I had a vague idea that I wanted to connect musical logic to mathematical models in some manner. Jeanne's expansive experience quickly pointed me to some important references and got me started on the right track. She believed that education and learning extends beyond the school environment, welcoming me to her summer hideout in Rockport, where picking mussels for lunch and academic critique mingled freely. Her acute intuitions and indomitable energy have guided and inspired me throughout my doctoral research. At the back of my mind, I have always suspected that the pure gumption she demonstrated in taking me on as a student in this math-music research was due in no small part to her humanitarian spirit, as evidenced by her activities in many academic advisory committees.

This interdisciplinary endeavor would not have been possible without my committee members, who displayed a remarkable willingness to brave, and eventually embrace, such an unusual dissertation topic produced in a rather limited time frame. More than most dissertation committees, they have devoted much time and energies into thinking about and improving my work. Georgia Perakis served as my advisor from the Operations Research Center, and helped to clarify the mathematical discussions in the thesis. In a similar manner, Martin Brody's thoughtful suggestions helped to clarify the musical issues explored in the thesis. Evan Ziporyn provided encouragement and further musical guidance at a few critical junctures. My committee chair, William Pounds, demonstrated a prodigious gift for organization, and fostered collegiality at the numerous committee meetings. His cogent advice on career issues were always cheerfully given, out of his genuine curiosity about, and liking for, people.

Academic work at the ORC would not be possible without the able guidance of the administrators: Paulette Mosley, Laura Rose and Danielle Bonaventura. In particular, Paulette and Laura have been at the ORC throughout my tenure as a student, and have always looked out for my interest.

I have also been blessed with excellent company at the ORC. The senior students who took me under their wings included Robert Shumsky, Edieal Pinker, Kerry Malone and Mitchell Burman. The friendly community that welcomed me to the ORC included Armann Ingolfsson, S. Raghavan and Sungsu Ahn. Late nights at the workplace were made more bearable by the company of friends such as Gloria Lee and Eriko Kitazawa. I also enjoyed the friendship of other colleagues, Sarah Stock Patterson, Thalia Chryssikou and Jeremie Gallien.

I would not have been able to make it through graduate school without the enrichment (and diversion) provided by MIT's vibrant music community. David Deveau was my piano teacher, and has always stalwartly supported this part of my life. Marcus Thompson's whole-hearted devotion to his students and to MIT's chamber music society often brought him into school for late night coaching sessions from which I benefited greatly. I have also had the pleasure of working with two of MIT's composers, Peter Child, John Harbison, Edward Cohen and Eric Sawyer – it is a rare honor to work with living composers of such calibre in presenting their work.

I have been involved in numerous concerts and practice sessions at MIT. Through these, I have met kindred spirits, too many to recount. But I will mention a few who have become close friends. Violist Wilson Hsieh and violinist Donald Yeung were like older brothers to me during my beginning years at MIT, and they have continued to be a source of support to me after becoming professors elsewhere. Clarinettist Eran Egozy's quick humour and playful showmanship has enlivened our many collaborations. Pianist Jee-Hoon Yap's powerful and noble playing, violinist Julia Ogrydziak's elegance and quirky artistry, and soprano Janna Baty's unaffected stylishness have brought much beauty and inspiration to my life. Many hours were happily spent rehearsing with the Aurelius Piano Quartet – Walter Federle, violin; Annette Klein, viola; Michael Bonner, cello – and at Sunday musical brunches. The larger group, the Aurelius Ensemble, provided me with more than my fair share of artistic fulfillment and organizational challenges.

All the concerts would not have been possible without MIT's able cast of staff members. At the music library, Forrest Larson was a walking resource on music literature, and Christina Moore and Peter Munstedt were always willing to augment the library's acquisitiongs to include music for our concerts. At the music office, Clarise Snyder would reminisce about warmer climates while she juggled two dozen and one schedules; John Lyons, Mary Cabral and Matthew Agoglia were also invaluable to the smooth running of each new project. Thanks to Ann Richard, I had ample access to the chestnut brown grand piano in Killian Hall prior to any concert. At the Office of the Arts, Mary Haller and Lynn Heinemann carefully and thoroughly looked over each press release, and showed me the ropes to the art of publicizing concerts. The MISTI China

Program, guided by the vision of Deborah Ullrich and Suzanne Berger, gave me a once-in-a-lifetime chance to seek out contemporary Chinese music at its source. Liz Connors was always ready to lend a friendly ear over a cup of tea, and to cheer me on.

Outside of school, I have always had the pleasure of sharing an apartment with an MIT architect. When Afshan Hamid and I decided to move out of Green Hall and put our fates into the hands of Cambridge housing, I had no idea that I was soon to be initiated into the architects' circle. I have gained much from her gift of friendship: experiencing her Pakistani wedding and budding motherhood; a surrogate family in Chicago; and, a few recipes for staple Pakistani dishes. Living with Afshan and Matthias (Jaffe) was the first time I had a home away from home. The architects who came to live at 158 Webster after Afshan were Robert Clocker, then Nina Chen. Together with Hugo Touchette, Nina and I shared many a meal eating and chatting around the kitchen table. Hugo also lent me his thesis template for this dissertation. Paul Keel, not one of my roommates but also an architect, has always been a source of encouragement and arbiter of good taste.

Afternoon tea at the ninth floor of the Green Building became a regular part of my daily routine. There, Ronni Schartz kindly allowed me to raid her fridge for chocolates and cookies. We became thick as thieves, sharing in each other's triumphs (Wellingtonian or otherwise) and sorrows. Like a big sister, she counselled me on many occasions, helped me shop for some of my best concert dresses at bargain prices and taught me that there is always a funny side to every difficult situation.

Through Ronni, I met Brian Arbic, who has been my foremost supporter in this interdisciplinary endeavor. The road to thesis completion would have been much rougher without his nourishment, both emotional and substantial. Brian's colleagues have patiently put up with my presence in the Joint Program's computer facilities during my last stretch of thesis-writing over the holidays. A few – Avon Russell, Chris Hill and Allistair Adcroft – have even helped me with my LaTeX and file-transfer questions. Brian, himself, spent many hours carefully proofreading the first draft of this thesis, and has championed it not only to the Physical Oceanography community, but also to the Operations Research faculty. In addition, Brian's parents, Bernard and Colleen have accorded me much love and encouragement.

Last but not least, I wish to thank my parents, CHEW Kim Lin and TAN Lay Tin, for their love and support through the years; and, my siblings, Vincent and Effie, for the rambunctious times we shared growing up on the Nantah campus. Numerous members of my parents' e·x·t·e·n·s·i·v·e immediate family have also offered encouragement throughout.

Elaine Chew
31 December 1999

Contents

List of Figures

List of Tables

1 — Introduction

This thesis introduces a Spiral Array model, a spatial representation of the relations embodied in tonality[1]. The model is generative in that key representations are generated from chords, and chord representations from component pitches[2]. Distinct from other geometric and network models for tonal relations, the Spiral Array represents pitches, intervals, chords and keys in the same spatial framework. In this space, any collection of pitches can generate a center of effect, that is essentially a mathematical sum of its parts, whose distance from any other element can then be measured. Distance, in the model corresponds to perceived closeness in tonal music.

The model offers a way to re-conceptualize tonal relationships, and the generating of tonal centers. A computational model, it raises pertinent questions regarding, and produces insights that illuminate, some basic issues in traditional music theory. In this thesis, I demonstrate the versatility of the Spiral Array model by applying it to three fundamental

[1] *Tonality* refers to the underlying principles of tonal music. According to Bamberger [5], tonality and its internal logic frame the coherence among pitch relations in the music with which [we] are most familiar. It is also sometimes synonymous with key. Quoting Dahlhaus,

> "The term *tonality* was coined in French (*"tonalité"*) by Castil-Blaze to signify the fundamental notes of a key: the tonic, the 4th and the 5th (*cordes tonales* as distinct from *cordes mélodiques*). In common usage the term denotes, in the broadest sense, relationships between pitches, and more specifically a system of relationships between pitches having a "tonic" or central pitch as its most important element.
>
> In 1844 Fétis defined tonality as the sum total of the "necessary successive or simultaneous relationships between the notes of a scale". According to Fétis the variety of historical and ethnic preconditions gives rise to a multiplicity of "types of tonality". Riemann disputed this relativist premise, holding the view that it could be proved that all types of tonality derive from a single principle: the establishment of significant tonal relationships by means of the chordal functions of the tonic, the dominant and the subdominant. Riemann's system has been disputed in turn: ethnomusicologists and historians restrict its application to the age of tonal harmony in European music (from the 17th century to the 19th or early 20th) or even to the Classical Period alone.
>
> ... While the work "key" is linked with the idea of a diatonic scale in which the notes, intervals and chords area contained, a tonality reaches further than the note content of a major or minor scale, through chromaticism, passing reference to other key areas, or wholesale modulation: the decisive factor in the tonal effect is the functional association with the tonic chord (emphasized by functional theory), not the link with a scale (which is regarded as the basic determinant of key in the theory of fundamental progressions). A tonality is thus an expanded key.
>
> Tonality [is] the underlying element of a tonal structure, the effective principle at its heart."

[2] A *pitch* is a sound of some frequency. High frequency sounds produce a high pitch, and low frequency sounds produce a low pitch. A *note* is a symbol that represents two properties, pitch and duration.

problems in the perception and analysis of western tonal music: that of finding keys[3], determining chords[4], and searching for modulations[5].

I show that the key-finding algorithm (Center of Effect Generator, CEG) devised using the Spiral Array model surpasses previous ones in its average performance, and is close to optimal. The problem of determining modulations is a difficult one computationally, and is one that has not been solved. Again, using the Spiral Array, I design an algorithm (Boundary Search Algorithm, BSA) that can search from boundaries between keys. In addition, I propose a preliminary algorithm (Center of Effect Generator 2, CEG2) that can determine chords in a musical passage.

Being able to study the nature of these problems and their solutions is critical to the understanding of human perception and analysis of tonal music, and also to pedagogical issues pertaining to these problems. Being able to characterize the relationships that generate a tonal center is crucial to the understanding and the making of performance decisions. Solving these basic problems computationally is also a precursor to any computer analysis of western tonal music, and automated systems that interface computer-generated music with real-time performance.

One of the main contributions of this thesis is the bridging of two disparate disciplines, that of music theory and operations research. Operations research is the science of decision-making using mathematical models which integrate the operating criterion of the system in question. According to George Dantzig[6], the inventor of the Simplex Method and father of Linear Programming, *"OR is the wide, wide world of mathematics applied to anything!"* Music theory describes the underlying principles that govern the system of relations organizing the perception and analysis of tonal music compositions. According to Gottfried Leibniz (1646-1716), the German philosopher, physicist and mathematician, as quoted in Lorenz Mizler's *Musikalische Bibliothek*: *"Music is the hidden arithmetical exercise of a mind unconscious that is calculating."* It would seem natural, then, to utilize the techniques in Operations Research to model effectively the perceptual problem solving inherent in the comprehension of western tonal music.

The two disciplines are bridged by a computational geometric model that is inspired by operations research techniques, and is built upon the framework of tonal relations based in music theory. The model, in turn, will provide insights into the symmetries and other relationships of the tonal system. This interdisciplinary effort forms a core contribution of my thesis.

[3]Excerpted from the Oxford Dictionary of Music: A *key*, as a principle in music composition, implies adherence, in any passage, to the note-material of one of the major or minor scales.

[4]A *chord* is any simultaneous combination of three or more notes. In this instance, I mean the name of the root of a triad or tetrachord based on intervals of fifths and thirds.

[5]Again, from the Oxford Dictionary of Music: A *modulation* is the changing from one key to another in the course of a section of a composition by evolutionary musical means and as a part of the work's formal organization.

[6]Personal communications, 7 November 1999.

1.1 Motivation

This project began as an attempt to formally describe the generating of a tonal center[7]. The result of my quest was a mathematical model that mimics the human decision-making process in comprehending tonality. However, the underlying questions that provided the impulse for this thesis remain as important motivating issues. Understanding how a tonal center is generated is a critical part of determining the key of a musical passage. The determination of tonal centers and their progressions in a piece of music is of critical importance to the analysis and perception of tonal music.

What do you mean by key?

In my first semester as a pianolab[8] instructor at MIT, I encountered a few students who had no prior musical background. I asked one such student, after he carefully traced out the melodic line for Yankee Doodle, "What is the key of this piece?" He responded with a reasonable question: "What do you mean by key?"

The obvious (but not altogether accurate) answer was to look at the key signature[9] at the beginning of the piece and accidentals[10] in the passage. There are several problems with this approach. *Looking* at the number of sharps and flats in the key signature ignores the fact that the perception of key is an aural experience. Accidentals, although often helpful, are not the only clues to the tonality of the passage. One could easily find a counter-example with no accidentals which does not generate the feel of a C major[11] tonality. For example, in Figure 1-1, a segment from the negro spiritual "Nobody Knows the Trouble I've Seen", the notes have no sharps or flats or other accidentals, but this melody sounds distinctly in the key of F major, and not C major as suggested by the key signature and the absence of accidentals. The first four notes of "Nobody Knows" already suggest strongly the F major tonality. In a little while, I will attempt to outline some of the decision process I undergo as a listener when encountering this melody. For now, I continue with the narration of the story of the student who asked, "What do you mean by key?"

To any musician, each key has its own distinctive terrain on their musical instrument. The key of a piece of music confers a physical shape, a unique topography, to the moving hands of the musician. An approach sometimes used by instrumentalists is to play a few bars of the piece, and by the way the piece feels in the hands, determine the key. This approach uses the physical experience of making music. But what about the aural experience?

Almost at my wit's end, I jumped at the next idea that came to mind. I hummed the

[7] The *tonal center* is the pitch that has attained greatest stability in a musical passage. The tonal center is also called the *tonic* of the key.

[8] A keyboard skills class for students enrolled in Music Fundamentals and Composition courses.

[9] A *key signature* is a sign placed at the opening of a composition or of a section of a composition, indicating the key. This sign consists of one or more sharps or flats.

[10] An *accidental* is the sign indicating momentary departure from the key signature by means of a sharp, flat or natural.

[11] C major is the major key with no sharps or flats.

piece, and stopped mid-stream. I asked the student if he could sing me the note on which the piece should end. Without a second thought, he sang the correct pitch, the tonic[12], sometimes called the tonal center. The success of this method raised more questions than it answered. These questions are aptly described by Bamberger in *Developing Musical Intuitions* (p.155):

> "How can we explain this tonic function which seems so immediately intuitive? While theorists have argued about answers to this question, most agree that for listeners who have grown up in Western musical culture, the stable function of the tonic derives primarily from its relation to the other pitches which surround it. Thus, the tonic function that a pitch acquires is entirely an internal affair: a pitch acquires a tonic function through its contact with a specific collection of pitches, the particular ordering and rhythmic orientation of this collection as each melody unfolds through time."

What is it we know that causes us to hear one pitch as being more stable than another? How does the function of the tonic evolve over the unfolding of a piece? Perhaps a quick example using "Nobody Knows" might shed some light on this matter.

Example: the Negro Spiritual "Nobody Knows the Trouble I've Seen"

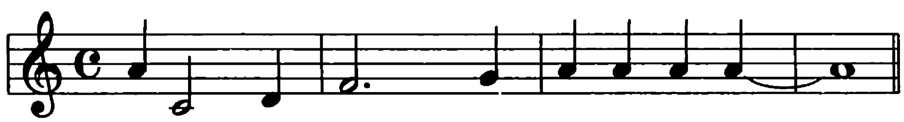


Figure 1-1: Excerpt from "Nobody Knows the Trouble I've Seen". A musical example with no accidentals, that is actually in the key of F.

This example demonstrates that many factors contribute to the listener's perception of tonality. These factors include interval relations, pitch durations and meter. The first four notes of "Nobody Knows" sets up a most stable pitch, and already gives a strong indication of the key. I will outline my own experience in determining this most stable pitch through the first four notes of "Nobody Knows".

The descending major sixth interval[13] between the first and second note (A and C respectively) strongly hints that the most stable pitch is F. This knowledge is the result of

[12]The *tonic* is the pitch of greatest stability. When pitches are ordered in a major or minor scale, this is the first degree of the scale.

[13]Excerpted from the Oxford Dictionary of Music, an *interval* is the distance between any two pitches, expressed by a number. For example, C to G is a 5th, because if we proceed up the major scale of C, the fifth pitch is G. The 4th, 5th and octave are all called Perfect. The other intervals, measured from the first pitch, in the ascending major scale are all called Major. Any Major interval can be chromatically reduced by a semitone (distance of a half step) to become Minor. If any Perfect or Minor interval is so reduced it becomes Diminished; if any Perfect or Major interval be increased by a semitone it becomes Augmented.

experience in listening to western tonal music. Many other tonal melodies begin with two pitches that are a major sixth interval apart. For example, the folksong "My Bonnie Sailed Over The Ocean" and Chopin's Nocturne Op. 9 No. 2 in E♭ (see Figure 1-2). In all three cases, the two pitches separated by an interval of a major sixth surround the stable pitch that is a major third interval below the upper pitch.

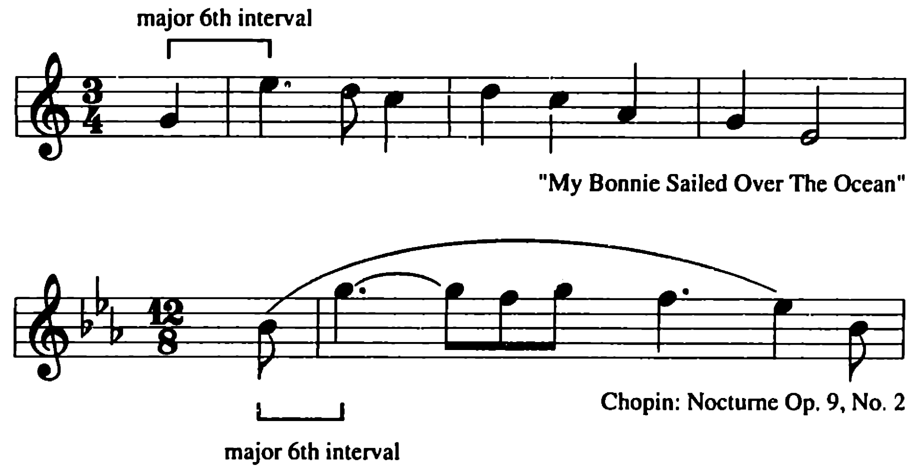


Figure 1-2: Excerpts from "My Bonnie Sailed Over the Ocean" and Chopin's Nocturne Op. 9 No. 2, both beginning with a rising major sixth interval.

Thus, the first two notes in "Nobody Knows" very likely can be assigned the solfege[14] syllables *mi* and *sol*, or scale[15] degrees $\hat{3}$ and $\hat{5}$. The distance between the pitches of the second and fourth note in the melody forms an interval of a perfect fourth. The rising fourth, from C to F, suggests the scale degree assignments $(\hat{5} - \hat{1})$, further reinforcing the F as tonic. Further examples of this rising fourth interval in other melodies are given in Figure 4-14. Together, the first, second and fourth notes outline the F major triad[16], implying an affinity to F major.

In "Nobody Knows", the rhythm in the melody also reinforce the tonic implied by the

[14]According to the Webster Dictionary, *solfege* is the application of the sol-fa syllables to a musical scale or to a melody.

[15]A *scale* is a series of single pitches progressing up or down stepwise. In this thesis, a scale always refers to the diatonic scale, as distinct from the chromatic scale (which uses nothing but semitones), the pentatonic scale (which uses five pitches) or the whole-tone scale (which is free of semitones). When ascending, the diatonic scale degrees are labeled $\hat{1}$, $\hat{2}$, ..., $\hat{7}$, and in have solfege syllables *do*, *re*, *mi*, *fa*, *sol*, *la* and *ti*. The *major scale* has semitone intervals between $(\hat{3} - \hat{4})$ and $(\hat{7} - \hat{1})$, the two halves thus being alike. The *natural minor scale* has semitone intervals between $(\hat{2} - \hat{3})$, $(\hat{5} - \hat{6})$. The *harmonic minor scale* has semitone intervals between $(\hat{2} - \hat{3})$, $(\hat{5} - \hat{6})$ and $(\hat{7} - \hat{1})$; and, the *melodic minor scale* has semitones between $(\hat{2} - \hat{3})$ and $(\hat{7} - \hat{1})$ ascending, and $(\hat{6} - \hat{5})$ and $(\hat{3} - \hat{2})$ descending.

[16]A *triad* is a chord of three notes, basically a "root" and the notes a third and a fifth above it, forming two superimposed thirds, e.g. F-A-C. If lower third is major and the upper minor, the triad is major. If lower third is minor and the upper major, the triad is minor. If both are major the triad is *Augmented*. If both are minor, the triad is *Diminished*.

interval relations. The note of longest duration in the first half of Figure 1-1 is the fourth note, and its pitch is F. In addition, the indicated meter places a downbeat on this F. Although the last note in Figure 1-1 has longer duration than the F, its onset begins on a weak beat (the fourth beat of the third bar).

The Questions

What is it we know that causes us to hear one pitch as being more stable than another? How does the function of the tonic evolve over the unfolding of a piece? Is there a way to describe formally the framework of pitch inter-relations that determine the key? Thus one student's seemingly innocent question of how does one find the key has led to my quest for a concise and effective model for the generating of tonal centers. In the following chapters, I will describe a model for tonality based on Riemann's system whereby significant tonal relationships are established by means of the chord functions of the tonic, the dominant[17] and the subdominant.

1.2 An Interdisciplinary Effort

Computational music analysis, by definition, is an interdisciplinary study linking human perception and cognition, mathematical modeling and computation, and music theory. A confluence of the three could ideally result in fruitful research leading to an enrichment of our understanding of all three disciplines. Desain et al [14] documents several successful attempts to bridge pairs of these disciplines in the past decades. I have summarized in Figure 1-3 a sampling of some interdisciplinary research in computational music analysis. In this thesis, I draw upon all three disciplines in the development and construction of my model.

Music is not an easy domain within which to design effective computational models that describe changing tonalities and harmonies. Being able to determine tonal centers of musical selections and harmonic function of chords is invaluable for better understanding issues related to human perception and performance of music. Being able to determine the tonal centers, and their progressions, in a piece of music is of critical importance to the analysis and perception of tonal and atonal music.

The process of music perception and performance utilizes both top-down and bottom-up analyses [14]. When assessing tonal (and rhythmic) structure, the human mind contemporaneously considers several different structural levels in the music. The ability to simultaneously scale up and down allows the listener to gather information at all levels.

[17] Each scale degree has a name that reflects its function with respect to the first scale degree. The name or key of the scale is the tonic, the first scale degree. The fifth is the "dominant," and the fifth below (the fourth) the "subdominant." Halfway between the tonic and dominant is the third scale degree, the "mediant." On the opposite side, halfway between the tonic and subdominant, the sixth scale degree, is the "submediant." The seventh, "leading note," is thus named for its tendency to move towards the tonic. And the second scale degree, the one above the tonic, is the supertonic.

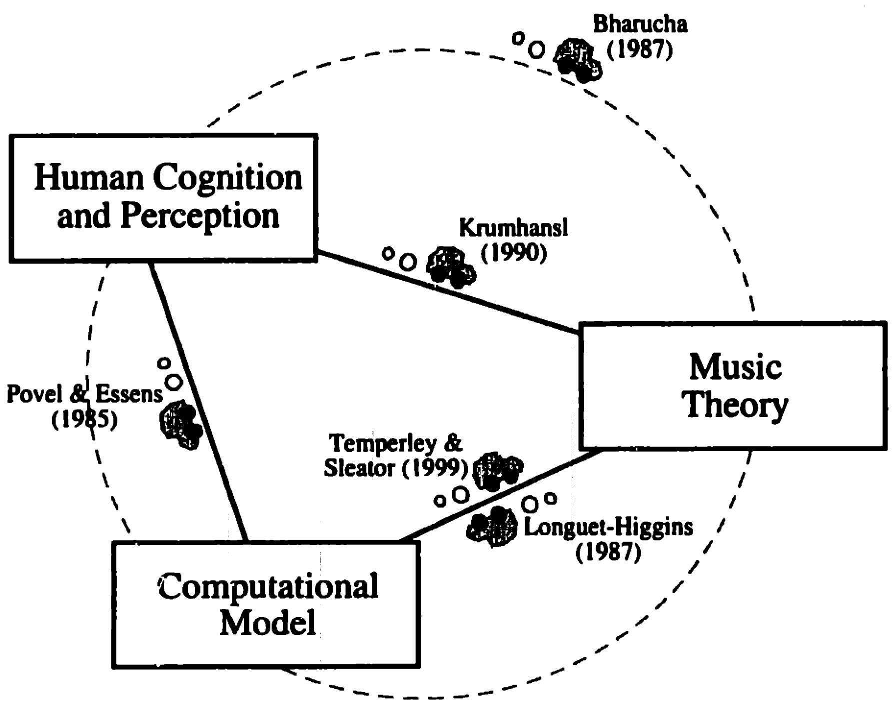


Figure 1-3: Bridging the disciplines: a sampling of some interdisciplinary research in computational music analysis.

Furthermore, music can inherently be described structurally in multiple and equally valid ways [4, 5, 27]. For example, Bamberger [4] showed that even young children are capable of focussing on different but legitimate hearings of the same rhythm which she terms figural and formal. This is directly related to Lerdahl and Jackendoff's [27] demonstration of the fundamental difference between the musical elements of grouping and meter.

To date, several dissertations have been devoted to the modeling of tonal perception, including Krumhansl's 1978 treatise on "The Psychological Representation of Musical Pitch in a Tonal Context" [25] which presents a behavioral approach, Laden's "A Model of Tonality Cognition which Incorporates Pitch and Rhythm" [26] using connectionist methods, and most recently, Temperley's 1996 thesis on "The Perception of Harmony and Tonality: An Algorithmic Perspective" [51] grounded in cognitive science.

I will be following in the tradition of attempting interdisciplinary work in the mathematical sciences and music. One of my goals in the thesis is to characterize and to reveal the latent structure and symmetries of the tonal system. My computational approach, presented in this dissertation, transforms the harmonic network and proposes algorithms that improve upon the performance and expand the scope of its previously known applications.

1.3 Overview of Thesis Content

I provide some relevant background and a literature survey in Chapter 2. Because the Spiral Array model is a spatial one, I review some other spatial model for pitch relations. Also, since the Spiral Array is derived from the harmonic network (*tonnetz*), I provide an overview of the harmonic network's history and applications. The Spiral Array is a versatile model that lends itself easily to computational applications, and I propose a few algorithms for these applications in later chapters. In Chapter 2, I review some approaches proposed by other researchers for these same problems.

Chapter 3 introduces the Spiral Array model, explaining how pitch, chord and key representations are generated in this structure. In addition, some symmetries in the model are highlighted. Chapter 4 presents the arguments for selecting different parameter values; these values affect the positioning of the tonal elements in the Spiral Array, and their proximity in relation to each other. Based on a few of these conditions, I derive the constraints on the parameter values that would satisfy each condition.

Chapter 5 introduces the first computational application that uses the Spiral Array: the problem of key-finding in melodies. I propose a key-finding algorithm, the Center of Effect Generator (CEG), and explain how it works by applying it to an example, "Simple Gifts". In addition, I compare the CEG algorithm to those by two other researchers, analyzing the comparison results when applied to the 24 fugue subjects in Book 1 of Bach's WTC.

In Chapter 6, I propose an algorithm for determining modulations, the Boundary Search Algorithm (BSA). The algorithm is applied to two examples, Bach's *Minuet in G*, and his *Marche in D* (both from "A Little Notebook for Anna Magdalena"). The conclusion suggests more sophisticated variations on this basic algorithm. Chapter 7 presents preliminary work on the use of the Spiral Array to find the roots of chords. A prototypical algorithm (CEG2) is proposed and applied to segments from Beethoven's *Op. 13* (the "Pathetique" Sonata) and Schubert's *Op 33*.

In conclusion, Chapter 8 reviews the contributions of this thesis, and suggests some directions for future research.

2 — Background

Tonality denotes a system of relations among pitches that generates a central pitch as its most important element. Many researchers have proposed models for this system of pitch relations. The principal idea behind this thesis is a spatial model for pitch relations that mimics the process by which pitches generate a center.

The proposed model for tonal relations, the Spiral Array, is a geometric model, and in Section 2.1, I will review some spatial models for pitch relations that have informed my work. The Spiral Array (shown in Figure 3-1) is a three-dimensional realization of the harmonic network (*tonnetz*), and in Section 2.2. The harmonic network, as a model that incorporates perfect fifth and major/minor third interval relations among pitches, has inspired a wide range of research in the Music Theory literature. I will provide an overview of its history and some of its applications. In this thesis, I restrict the application of the Spiral Array to the design of efficient algorithms for a few basic music perception problems.

The Spiral Array is a versatile model that lends itself easily to computational applications. In Chapter 5, I apply the model to key-finding in melodies. In Chapter 6, I use the model to design an algorithm for determining modulations in a musical passage. In Chapter 7, I document preliminary work on using the Spiral Array to determine chords.

Other researchers have addressed the problems of key-finding and determining chords using computational means. In fact, research in the modeling of tonal and harmonic analysis has expanded over the past two decades in artificial intelligence and music perception communities. In Section 2.3, I will survey the literature on key-finding algorithms. In Section 2.4, I will review selected algorithms for chord analysis. In closing, I provide some brief comments on other mathematical models in music.

2.1 Spatial Models for Pitch Relations

This thesis proposes a Spiral Array model, a spatial construct that represents the interrelations among musical pitches. As its name suggests, pitches are represented in the Spiral Array by points in a spiral configuration. The Spiral Array is a *generative* model that captures, spatially and mathematically, the idea that higher-level representations are a composite of their lower-level elements. In distinction to previous models, these different hierarchical elements can reside within the same framework in the same space. In addition, close tonal relations are mirrored by spatial proximity among their corresponding representations. By construction, the Spiral Array gives highest preference to perfect fifth (P5), major third (M3) and minor third (m3) interval relations.

Spatial Analogues as Tools for Solving Problems

Spatial analogues of physical and psychological phenomena are known to be powerful tools for solving abstract intellectual problems [47]. Biologists have long used geometric models for DNA structure and protein folding. Chemists study structural models for chemical bonding. In Mathematics, it was George Dantzig's adroitness at geometry that inspired him to invent the Simplex Method for solving linear optimization models.

Some have argued that problems in music perception can be reduced to that of finding an optimal data representation [50]. The most recent issue (Volume 11) of *Computing in Musicology* [18], covers an assortment of pitch encodings for computer comparisons of melodies. Tonality describes a highly structured system of pitch relations, but few representations have incorporated the functional relations among pitches that generate a tonal center.

Psychologists Shepard and Krumhansl have each sought for representations that would emulate tonal relations in their spatial configurations. I describe their work in the next two sections.

Spiral Models for Pitch Relations

In 1982, Shepard [47], stated that "the cognitive representation of musical pitch must have properties of great regularity, symmetry, and transformational invariance." Recognizing that perfect fifth and major third intervals were perceptually important pitch relations, Shepard aimed to design a spatial model in which these relations have direct counterparts in the geometric structure.

In the tradition of spiral models for pitch dating as far back as 1855, Shepard proposed a model which spaced all twelve chromatic pitches equally over one full turn of a spiral (see Figure 2-1). The equal spacing, distinct from previous models, emphasized the close relationship of pitches related by octave intervals. Further extensions to incorporate perfect fifth interval relations resulted in double helix structures that still did not provide for the major third.

Other Configurations for Pitch Relations

Krumhansl's [25] Ph.D. thesis attempted to uncover the structure of pitch relations in tonality using experimental data. In this, and other later publications [23, 24, 22], Krumhansl uses two main ideas in analyzing data garnered from each experiment. The first is the use of Probe Tone Profiles, judgements with respect to how well each of the twelve pitch classes fit into a given key. The second, is the use of Multidimensional Scaling, a statistical method created by Shepard [46]. This technique maps each profile onto a point in Euclidean space such that high correlation is mirrored by spatial proximity.

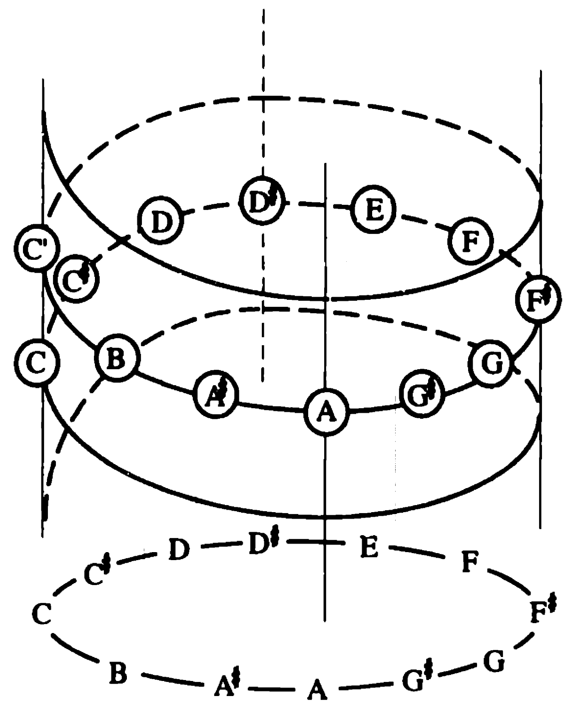


Figure 2-1: Shepard's spiral model of chromatic pitches.

Based on statistical analysis of psychological experiments, Krumhansl proposed a conical structure for pitch relations [25]. The pitches of the major triad were located on a plane closest to the vertex, the remaining diatonic pitches on a second plane, and the remaining pitches in the most distant plane.

This conical structure does not contradict the Spiral Array arrangement. In addition, Krumhansl deduced that an established tonal context changes the pitch relationships, thus suggesting that a geometric representation may have to be altered depending on the key.

Most recently [23, 22], Krumhansl used the Probe Tone Profile method and Multidimensional Scaling to perform studies of listener judgements of triadic proximity, and used empirical data to corroborate the psychological reality of neo-Riemannian transformations.

2.2 The Harmonic Network or Tonnetz

'Tonality describes a highly structured system of pitch relations, a system that has been studied by theorists over the past two centuries. Many theories have been proposed to explain the relations implied in tonal music. One such theory, proposed by the 19th century music theorist Riemann, posits that tonality derives from the establishing of significant tonal relationships through chord functions. This idea has influenced a wide range of research in Music Theory.

Riemann's theory, agrees with Shepard's intuitions that the most significant interval relations are the perfect fifth and the major/minor third. Riemann represented these relations in a harmonic network as shown in Figure 2-2. This harmonic network is also known as the table of tonal relations or *tonnetz*). Cohn [11] has traced the earliest version of this network of pitches related by perfect fifth and major/minor third intervals to the 18th century mathematician Euler. In the 19th century, this representation was appropriated by music theorists such as Oettingen and Riemann.

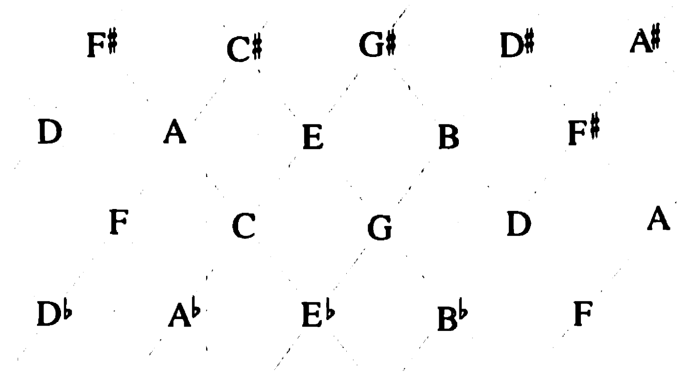


Figure 2-2: The Harmonic Network or *Tonnetz*.

As will be explained in Section 3.1, the proposed Spiral Array is a three-dimensional realization of the harmonic network. In the Spiral Array, adjacent pitches are positioned at each quarter turn of the spiral, shown in Figure 3-1. Neighboring pitches on the spiral are a perfect fifth apart, and pitches vertically above each other are a major third apart. If one considers equal temperament[1], the spiral would close to form a toroid.

The added dimension provided by the spiral arrangement allows computationally efficient and cognitively accurate algorithms to be designed for problems in music perception. These will be described in subsequent chapters. Prior to the Spiral Array, researchers have alluded to the inherent toroid structure of the harmonic network but only used the planar versions in their music analysis applications.

Application I: Key-Finding

In 1962, Longuet-Higgins [32, 33] noted that pitches in a key are positioned in a compact neighborhood on the harmonic network. Following this insight, he proposed a Shape-Matching Algorithm (SMA) for key-finding using this table of harmonic relations. As

[1] *Equal temperament* refers to the practice of tuning a keyboard "out of tune" so that one can play in any key on the same instrument. To demonstrate that this was possible, Bach wrote the "Well-Tempered Clavier." In this system of tuning, enharmonic (pitches that differ from each other in name, but not in any other way when keyboards are concerned) pairs of notes such as A♯ and B♭ are treated as one and the same.

Longuet-Higgins' algorithm is the one most closely related to the Spiral Array, the SMA will be discussed in greater detail in Section 2.3. In Chapter 5, the SMA will be compared to the Spiral Array's Center of Effect Generator (CEG) algorithm.

Longuet-Higgins' work has inspired further mathematical analyses of the harmonic network, such as that of Balzano [3], which favors the (major 3rd, minor 3rd) axes as opposed to the (perfect 5th, major 3rd) arrangement. However, one can demonstrate that the (perfect 5th, major 3rd), (major 3rd, minor 3rd) and the (perfect 5th, minor 3rd) representations are equivalent.

Application IIa: Transformational Theory

In 1982, Lewin [28, 29] revived the use of the ***tonnetz*** in music analysis, thus planting the seed for the emerging field of Transformational Theory for the analysis of triadic post-tonal music [12].

In the harmonic network, similar triads (where similarity is measured in terms of common pitches and parsimonious voice leading) form triangles that have common sides and are consequently near each other. Lewin [28] proposed a set of transformations on the harmonic network that mapped similar triads to each other.

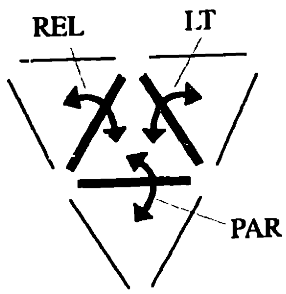


Figure 2-3: Transformations on the Harmonic Network.

The three main types of transformations are parallel (PAR), relative (REL) and leading tone exchange (LT), as shown in Figure 2-3. PAR maps a major (minor) triad to its parallel[2] minor (major). REL maps a major (minor) triad to its relative[3] minor (major). And, LT lowers the root of the triad to its leading tone to form a new triad.

[2] For any major (minor) key, its *parallel* minor (major) is the corresponding minor (major) key with the same tonic.

[3] For a major (minor) key, its *relative* minor (major) is the key with the same sharps or flats. For example, C major is the relative major of A minor because neither has any sharps or flats; and, vice versa, A minor is the relative minor of C major. When referring to triads, this concept extends to the tonic triads of each key.

Application IIb: More Recent Developments

The transformations defined by Lewin can trace paths and cycles through the harmonic network that correspond to triadic movements in tonal music. An entire body of literature has emerged based on these transformations. Researchers have explored transformation cycles and permutations, and sought examples of such transformations in actual musical compositions.

In 1996, Cohn observed [10, 11] that the transformations emphasized parsimonious voice-leading. Based on this criterion, he abstracted transformation cycles on the *tonnetz* that were hexatonic (six-step) and octatonic (eight-step). In his essays, he showed how these patterns corresponded to triadic movements in chromatic music of the nineteenth century.

Recently, these models have been generalized [30], and extended to higher level cycles (such as, cycles of cycles [15]) and to more complex chords (such as, tetrachords [7, 17, 15]). These mathematical approaches are mainly graph-theoretic. Several allude to the inherent toroid structure of the *tonnetz* under equal temperament. However, none use the inherent spiral structure of the harmonic network without equal temperament.

Back to the Spiral Array

The Spiral Array preserves the harmonic network's preference for perfect fifth and major/minor third intervals. Similar triads are still physically close in the Spiral Array. At the same time, the new configuration incorporates the benefits of modeling pitch relations in a higher dimension.

The Spiral Array allows one to calculate mathematical centers for any pitch collection. As a result of this additional benefit, one can define representations for higher level elements such as intervals, triads, tetrachords, and keys. Similar elements (where similarity is measured in terms of common pitches and perfect fifth or major/minor third interval relations) are still physically close in space.

Note that elements that are spatially close in the Spiral Array are not necessarily close in space on a keyboard or finger board. Pitches that are spatially close on a keyboard, such as two pitches a half step apart, are represented by points in space that are relatively far apart. The Spiral Array (and the harmonic network as well) stresses harmonic relations (perfect fifth intervals and major/minor third intervals), and not linear relations such as pitches separated by a half step.

2.3 Key-Finding Algorithms

By "rolling up" the Harmonic Network, as described in Section 3.1 and shown in Figure 3-2, to reflect its spiral structure, I propose that each key be represented by a point within its compact key neighborhood. This formalization of key representations makes the Spiral Array space emminently suited to computational analysis of music, such as the determining of keys.

In the realm of Artificial Intelligence, Longuet-Higgins [35, 34, 32, 33] together with Steedman [37] has used the harmonic network in his key-finding algorithm, utilizing the inbuilt pitch relations in the network. Psychologist Krumhansl [24] together with Kessler [23], on the other hand, uses Multidimensional Scaling on empirical data, thus imputing a spatial structure to key relations. The Spiral Array might well serve as the link between these two seemingly disparate approaches.

The Spiral Array takes to a different level Longuet-Higgins' observation that pitches in a key are located in a compact neighborhood around the tonic. By proposing key representations that emulate the composite effect of the component pitches, the Spiral Array model suggests an explanation for the similarity between Krumhansl's spatial structures and the toroidal harmonic network. It is not a coincidence that, like the Transformational Theorists, Longuet-Higgins has found proximity on the harmonic network a useful criterion for categorizing tonal constructs.

In Chapter 5, I compare my key-finding algorithm with Longuet-Higgins' and Krumhansl's applied to Bach's Well-Tempered Clavier, and I discuss the pros and cons of the different approaches. Here, I briefly outline Longuet-Higgins' and Krumhansl's methods.

Krumhansl's Probe Tone Profile Method

Krumhansl's technique, described in Section 2.1, essentially performs the reverse process of mine. Both address the same phenomenon: pitch information that translates to the generating of a tonal center. I postulate an effect from the Spiral Array, and Krumhansl suggests a spatial structure based on the phenomenon.

As mentioned in Section 2.1, Krumhansl uses Probe Tone Profiles to analyze listeners' judgements of key relations using Multidimensional Scaling. As a result of this analysis, a spatial structure is constructed that reflects these key relations judgements. In Krumhansl's Probe Tone Profile method for key-finding, each musical passage generates a pitch-distribution profile, which is matched against the ideal distribution. Key candidates are ranked by likelihood.

Although this research does not stemming from the symmetries inherent in tonal music, it serves to corroborate the psychological reality of the harmonic network structure.

Longuet-Higgins and Steedman's Shape-Matching Algorithm

Few algorithms for computational tonal analysis have made use of the functional relationships among pitches known to music theorists. A notable exception is Longuet-Higgins' proposed key-finding algorithm which uses Shape Matching on the harmonic network. Due to the discrete nature of this algorithm, it does not allow for ambiguous solutions. As my model is closest in form and spirit to Longuet-Higgins', I summarize his Shape Matching Algorithm (SMA) below.

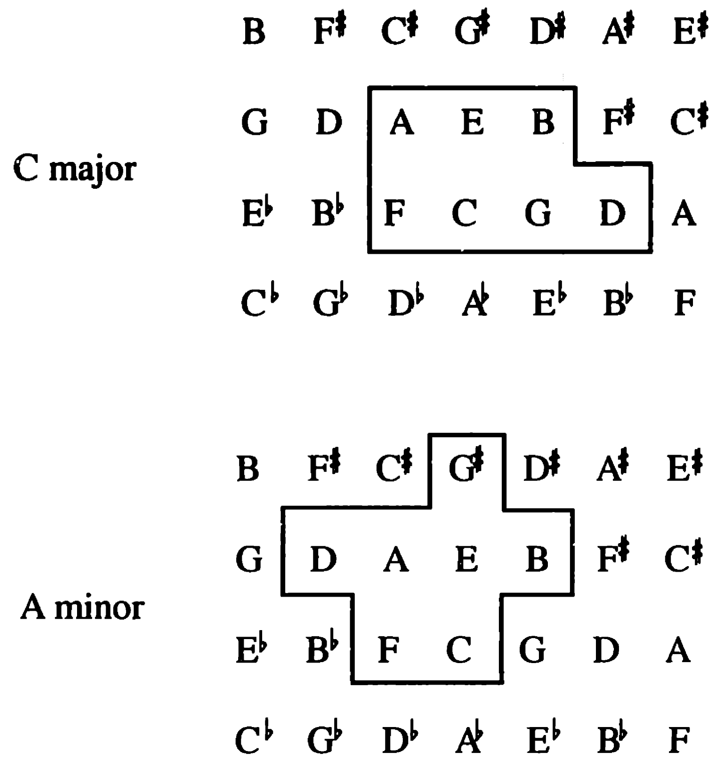


Figure 2-4: Examples of shapes outlining a major key (C major) and a minor key (A minor) in Longuet-Higgins' model of "harmonic space". Other major (minor) keys take on the same shape, but outline different pitch collections.

Longuet-Higgins noted that pitches belonging to a major key were located in such a way as to form a distinct and compact shape (see Figure 2-4). The shape remains the same for all major keys, the only different being the pitch collection inside the shape. Pitches belonging to a minor key formed a different unique and compact shape. Hence each key is represented by one of two shapes, and the key identity is given by the location of the shape.

In effect, the SMA works by a process of elimination. With each pitch event, shapes not covering the pitches sounded are eliminated, together with the keys they represent. This process is repeated until only one key remains.

The SMA could stall in cases of multiple solutions, or if all keys have been eliminated. In the first case, the tonic-dominant rule is invoked. In the latter, the algorithm returns to the point prior to the contradiction, and invokes the tonic-dominant rule. The tonic-dominant rule simply assumes that the first pitch in the theme is most likely to be the tonic, and the second most likely to be the dominant, of the actual key. Hence, if either of these are candidates for selection, they are to be chosen in that order.

The simplicity of the SMA is an appealing attribute. A further advantage of using the harmonic network (although this property was not used) is that intervals can be depicted as vectors. However, few musical passages adhere strictly to any one key, in which case the SMA stalls and invokes the tonic-dominant rule.

2.4 Harmonic Analysis Algorithms

On a large scale, pitch relations determine a key; on a smaller scale, pitch relations determine a chord. The two perceptual problems differ in scale, but not in essence. Since the Spiral Array represents pitches, chords and keys in the same space, by extension, the model can also be used to determine chords. In this section, I outline, amongst others, Winograd's and Temperley's artificial intelligence approaches to the problem of determining chords. In Chapter 7, I compare the Spiral Array's results to that produced by their methods.

Winograd's Systemic Grammar Algorithm

A few approaches have been suggested for the problem of harmonic analysis. An early one by Winograd [54] applied linguistics to harmonic analysis based on Allen Forte's *Tonal Harmony in Concept and Practice.*

More than simply identifying the roots of chords, Winograd's algorithm generates a harmonic analysis of a piece of music, providing chord function details. To avoid the exponential growth of the decision tree with length of the piece, he parses from back to front and introduces tree-pruning techniques.

There are a few drawbacks of Winograd's method. The back-to-front parsing reduces the number of ambiguities because a piece almost always ends on the tonic chord. But listeners assess music sequentially, with little help from the strong cues at the end of the piece unless an opportunity presents itself to listen to the same piece more than once in close succession. The input and results are user-dependent, and the algorithm does not handle incomplete chord information in melodic and contrapuntal lines.

Rule-Based Approach

More recently, several rule-based approaches have been suggested for tackling the problem of harmonic analysis ([38], [52], [53].) Expressing musical knowledge as rules is more intuitive than pure computational methods, and makes the knowledge base modular and extensible. However, completing this knowledge base is a nontrivial task, and the list of rules has to be modified for music from different eras and of dissimilar styles. Maxwell's expert system determines whether a collection of pitches is a chord using a total of 36 rules, then ascertains the chord name using an additional 18 rules.

Temperley and Sleator's algorithm combines both metric and harmonic analyses. Derived from Lerdahl and Jackendoff's Generative Theory of Tonal Music [27], it uses a sequence of preference rules to snap (pitch) event onsets onto hierarchical beat grids, and then sorts the pitches into chord names and ranges. As with Winograd's algorithm, the search space becomes unwieldy and large, and table-pruning must be incorporated into the dynamic programming implementation in order to maintain tractability.

See Meehan [39] for an overview of some other AI approaches. As far as I know, none of the computational methods has yet addressed the determination of key modulations.

2.5 Brief Comments on Other Mathematical Models

Much mathematical analysis of music has been done using Group Theory, Set Theory and Graph Theory. I am not aware of any existing literature on mathematical modeling of music using optimization techniques common in Operations Research.

Several researchers have proposed neural net models for musical applications, and algorithms for pattern matching. Krumhansl's research (see, for example, [25]) employs Multidimensional Scaling, a statistical method, to organize experimental data.

To my knowledge, the closest Mathematics-Music connection that is OR-related would be Xenakis' use of Probability and Game Theory in his compositions of stochastic music [55]. The modeling techniques I use include mathematical constraints that represent preferences and the search for feasible solutions in an optimization framework.

3 — The Spiral Array Model

This thesis proposes a Spiral Array that models the relations embodied in tonality. Each pitch (in this case, pitch class, since I assume octave equivalence[1]), chord and key has a spatial counterpart in this geometric framework.

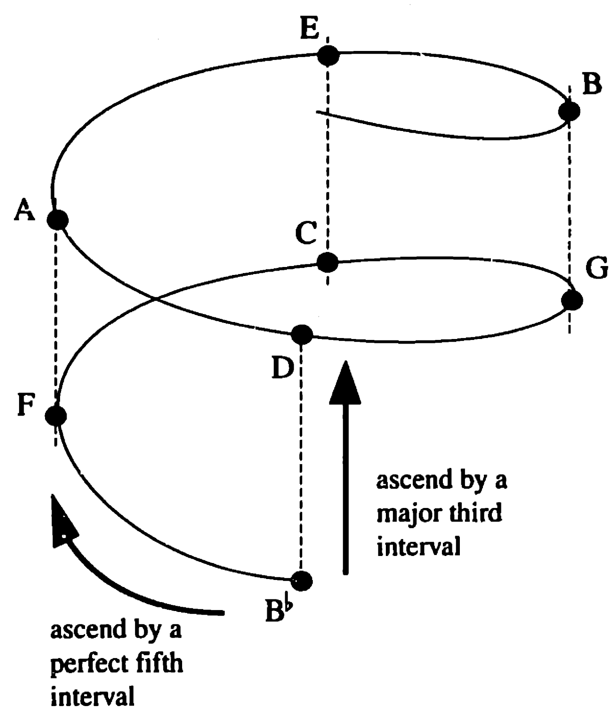


Figure 3-1: The Spiral Array.

The Spiral Array is a *generative* model; key representations are generated from their component chords, and chord representations from their component pitches. These representations are distinct from the root of the chord and the tonal center (tonic) of the key, entities which give chords and keys their names. These chord and key representations attempt to capture spatially the idea that the higher-level representations generated are a composite of their lower-level elements. In distinction to previous models, these different hierarchical elements can reside within the same framework in the same space.

[1] *Octave equivalence* means that pitches of the same letter name are considered to be the same, to belong to the same pitch class.

In this chapter, I introduce the model and show how it successively generates representations for higher level tonal elements as a composite of its lower level componenets. For example, chords representations are the result of their component pitch positions, and key representations are derived from their defining chords. In Chapter 4, I show how the model can be adjusted so that distances between the tonal representations match their perceived closeness. In Chapters 5 through 7, I show how the Spiral Array can be used as a framework in which to design algorithms for finding keys, naming chords and determining modulations.

3.1 The Spiral Array

The Spiral Array, as shown in Figure 3-1, is a three-dimensional realization of the harmonic network. The harmonic network is an array of pitches separated by perfect fifth intervals along the x-axis, and by major third intervals along the y-axis (see Figure 2-2 or 2-4). The pitch names repeat, resulting in periodicity in the network, and redundant representations. Inherent in the harmonic network is a spiral model without redundancies.

Description 1: A Derivation of the Harmonic Network

Take the planar harmonic network and rolling it up as shown in Figure 3-2 so that the A lines up with the A four columns away, and B♭ coincides with the B♭ four columns away, one eliminates these redundancies. The result of this transformation is the Spiral Array, which eliminates the periodicity in the harmonic network.

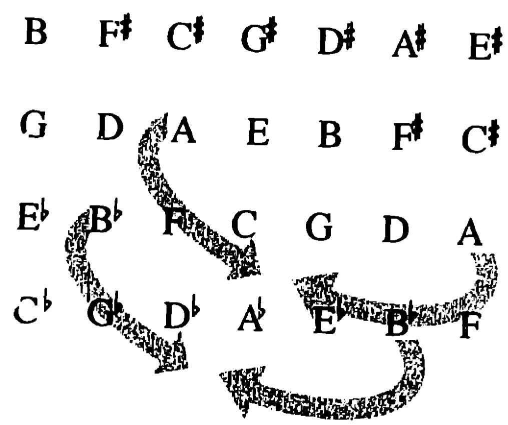


Figure 3-2: "Rolling up" the harmonic network to form the Spiral Array.

Alternatively, one could also generate this structure by wrapping a line pitches ascending by perfect fifth intervals on a vertical cylinder so that pitches a major third apart line up above one another four steps later.

Description 2: A Closer Look

In the Spiral Array, pitches are represented by points on a spiral. Adjacent pitches are positioned at each quarter turn of the spiral as shown in Figure 3-1. Following along the ascending spiral, neighboring pitches are a perfect fifth apart. Along the vertical axis, neighboring pitches are related by major third intervals. Pitches separated by octaves are assumed to be equivalent and share the same position.

In Figure 3-1, follow the winding spiral from the lowest pitch shown, B♭, up a quarter turn to the next pitch. This pitch is a perfect fifth above the previous one, that is, it is F. Continuing along the spiral, one comes to C (a perfect fifth above F) and G (a perfect fifth above C) then D (a perfect fifth above G). By the pitch D one has followed the spiral for four quarter turns after B♭, making one full turn. As a result of this process, D lines up verically above the starting point B♭. D and B♭ are related by an interval of a major third. Similarly, starting from F, four steps up along the spiral brings one to A, which is a major third above F. Continuing along, one can see that pitches vertically above each other are always a major third apart, shown the diagram by dotted lines.

Discussion 1: How do the Intervals Add Up?

Two pitches related by an interval of a perfect fifth are five steps apart along a major (or minor) scale, counting from the lower pitch. However, if one considers all the pitches that lie between the pitches defining a perfect fifth interval, the perfect fifth spans seven half steps (or semitones). Note that this definition of a perfect fifth interval simply gives a method of identifying the interval's span, and is not limited to only the first and fifth degrees of a scale. In fact, in a major scale, ($\hat{1}$,$\hat{5}$), ($\hat{2}$,$\hat{6}$) and ($\hat{3}$ up to $\hat{7}$) are some of the other perfect fifth interval relations among its pitches.

Two pitches related by an interval of a major third are separated by three scale steps along a major scale, which is the same as four half steps. Four perfect fifths piled on top of one another make a total of twenty-eight half steps. Since the model assumes octave equivalence, one can subtract octaves (octave intervals are defined by pitches twelve steps apart) from the twenty-eight half steps. This is equivalent to performing modulo twelve arithmetic.

Hence, the twenty-eight half steps are equivalent to four half steps, which is the distance between pitches defining a major third interval.

$$\begin{aligned} \text{4 perfect fifths} &= 4 \times 7 \text{ half steps} \\ &= 28 \bmod 12 \text{ half steps} \\ &= 4 \text{ half steps } = 1 \text{ major third} \end{aligned}$$

Similarly, a minor third can be represented by three perfect fifths in the opposite direction. Two pitches related by an interval of a minor third are separated by three steps along a minor scale, which is the same as three half steps. Three perfect fifths stacked on top of one another in the opposite direction make a total of twenty-one half steps in the opposite direction. By octave equivalence, twenty-one half steps are equivalent to three half steps, which is the distance between pitches defining a minor third interval.

$$\begin{aligned} -3 \text{ perfect fifths} &= -3 \times 7 \text{ half steps} \\ &= -21 \bmod 12 \text{ half steps} \\ &= 3 \text{ half steps } = 1 \text{ minor third} \end{aligned}$$

Discussion 2: The Inherent Toroid Structure

If one considers equal temperament, the spiral would close to wind itself around a toroid. I choose to keep the Spiral Array in its cylindrical form for reasons both computational and musical. In equal temperament, enharmonic pitches sound the same. However, I argue that the context in which a pitch occurs determines its spelling and vice versa. In addition, for computational purposes, linear distances would not apply in the toroid configuration, resulting in added computational complexity without much gain.

Mathematical Representation

Pitches are indexed by their number of perfect fifths from C. For example, D has index two because C to G is a perfect fifth, and G to D is another. $\mathbf{P}(k)$ denotes the point on the spiral representing a pitch of index k. Since the spiral makes one full turn every four pitches, each pitch can be defined in terms of transformations from its previous neighbor - a rotation, and a translation as shown in Figure 3-3.

Note that C has been arbitrarily chosen as the reference pitch. Any other pitch could have served as the reference, $\mathbf{P}(0)$. Furthermore, the following mathematical definition has arbitrarily fixed the pitch C at the point [0,1,0]. This, too, is an arbitrary choice. It is the relation between the pitch representations that is of utmost importance, and where the spiral begins is of little consequence.

Definition 1. *Each pitch position can be described as a rotation of the previous one by 90 degrees clockwise on the horizontal plane, and elevated by h units in the vertical direction:*

$$\mathbf{P}(k+1) \stackrel{\text{def}}{=} \mathbf{R} \cdot \mathbf{P}(k) + \mathbf{h}, \quad \textit{where } \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \textit{ and } \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix}. \qquad (3.1)$$

Each pitch position $\mathbf{P}(k)$ can also be parametrically defined as a function of the radius of the cylinder encasing the spiral, and the elevation at each step (see Figure 3-3). Section 4.1 discusses how these parameters can be selected to reflect perceived interval relations.

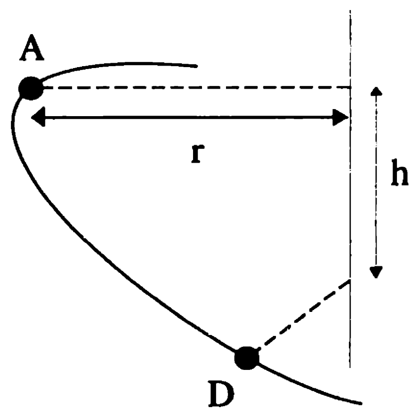


Figure 3-3: The two parameters that uniquely identify pitch position, the radius (r) and vertical step (h).

Definition 2. *Two parameters, the radius of the cylinder, r, and the height gain per quarter rotation, h, uniquely define the position of a pitch representation, which can be described as:*

$$\mathbf{P}(\mathrm{k}) \stackrel{\text{def}}{=} \begin{bmatrix} x_\mathrm{k} \\ y_\mathrm{k} \\ z_\mathrm{k} \end{bmatrix} = \begin{bmatrix} \mathrm{r}\sin\frac{\mathrm{k}\pi}{2} \\ \mathrm{r}\cos\frac{\mathrm{k}\pi}{2} \\ \mathrm{kh} \end{bmatrix}. \tag{3.2}$$

Since the spiral makes one full turn every four pitches to line up vertically above the starting pitch position. Positions representing pitches four indices, or a major third, apart are related by a simple vertical translation. For example, C and E are a major third apart, and E is positioned vertically above C.

Property 1. *Positions representing pitches four indices, or a major third, apart are related by a vertical translation of four elevation steps:*

$$\mathbf{P}(\mathrm{k}+4) = \mathbf{P}(\mathrm{k}) + 4 \cdot \mathbf{h}. \tag{3.3}$$

3.2 Interval Relations

The Spiral Array is, in essense, generated from interval relations. Many of these interval relations are inherited from the harmonic network, and a few are a result of the new configuration. In this section, I will discuss some of the consequences of the Spiral Array pitch arrangement.

Relations Inherited from the Harmonic Network

The model inherits the thirds and fifths interval relationships of the Harmonic Network. In the Spiral Array, the intervals representations undergo a new twist as shown in Figure 3-4.

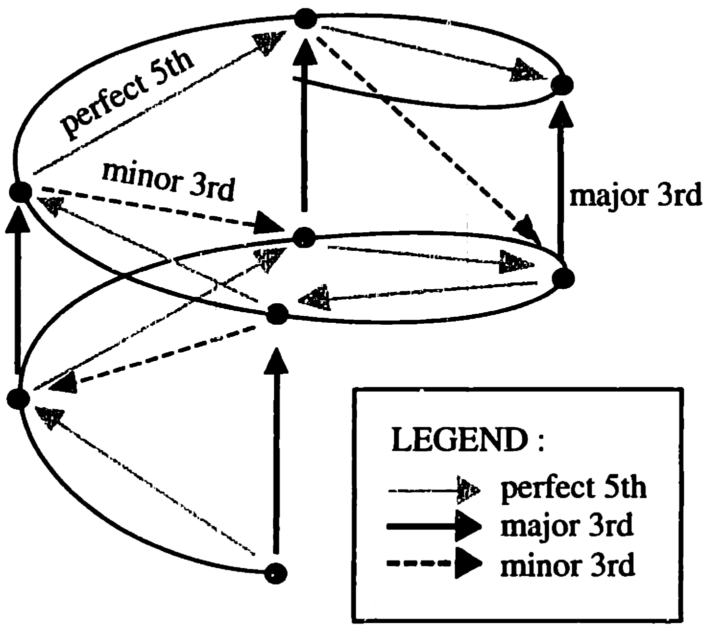


Figure 3-4: Perfect fifth, major third and minor third interval representations in the Spiral Network.

For any given pitch representation, its left neighbor (looking on from the outside) is the pitch a perfect fifth above and its top neighbor a major third above. Because the major third interval is equivalent to four perfect fifth intervals, starting from a given pitch, thepitch a major third above is four vertical steps above this starting pitch. In addition, diagonals represent minor third relations.

This spiral arrangement of adjacent perfect fifths and vertical major thirds affects other interval relations as well. For example, because of the placement of pitches, a major second interval is now represented by pitch positions diametrically opposite each other on the Spiral Array.

Interval Distances

Because the Spiral Array is a spatial model, interval relations correspond to distances between pitch representations. Each pair of pitch positions is separated by some spatial distance. And since pitch pairs define intervals, these spatial distances represent, spatially, different interval distances. One of the design strategies is to select parameters so that interval distances in the model correspond to perceived interval relations.

Note that distance in the model does not correspond to distance on a keyboard. The distance corresponds to the perceived closeness between pitches in tonal music. So, even though two pitches related by a half step interval are close to each other on a keyboard, they are quite distant in the Spiral Array. And, even though on a keyboard a perfect fifth interval spans more pitches than a major third interval, in the Spiral Array two pitches separated by a perfect fifth interval are closer than two pitches separated by a major third interval.

The distances representing the various interval relationships change depending on how much the spiral is "stretched out." Hence, the choice of the aspect ratio $\frac{h}{r}$ has consequences for all interval representations, not just the perfect fifth and major third. The selection of this parameter will be discussed in greater detail in Section 4.1.

In the planar arrangement, intervals could be represented by vectors. In the Spiral Array, intervals are no longer unique vectors. However, they can be represented by two alternative measures: the vertical gain (from one pitch position to the other) or the distance.

3.3 Representing Chords

This section introduces the modeling of a *center of effect* for higher-level tonal constructs. Chords and keys are represented by points in space at the center of their lower level components. This center is literally the mathematical sum and geometric center of the components.

A chord is the composite result, or effect, of its component pitches. A key is the effect of its defining chords. I propose that this effect can be represented spatially. The spatial representation, as well as the effect, cannot be identified only with one of the component elements, but should be a function of their interaction. For example, the C major triad[2] should not be represented by the pitch C alone, but by a combination of its three component pitches, C, E and G.

[2] A *triad* is a chord comprised of three pitches: a root, a third and a fifth. The root and fifth form a perfect fifth. The root and third form a major or minor third for a major or minor triad respectively. In this thesis, the words chord and triad will be used interchangeably.

Description: Center of Triangle

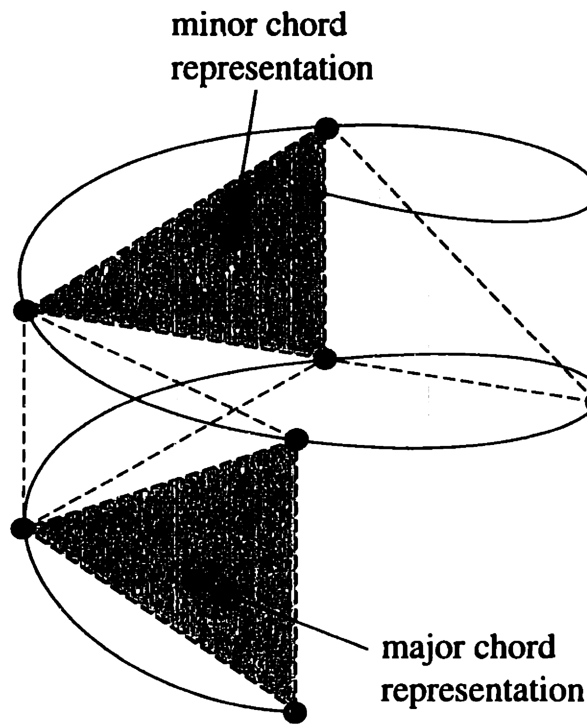


Figure 3-5: Examples of chord representations. Each chord representation is the composite result of its constituent pitches.

As in the harmonic network, chords are represented by triangles. In the Spiral Array geometry, I propose that each chord be represented spatially as a point in the interior of this triangle. Because chords are represented by non-overlapping triangles in the model, any point inside a triangle is sufficient to uniquely identify the designated chord. This point is not the root, the fifth or the third, but a convex combination of the three. It represents the center of effect of the chord pitches.

The three component pitches of the chord form the vertices of the triangle, and the sides of the triangle represent a perfect fifth interval, a major third interval, and a minor third interval. Major and minor triads differ only in their orientation. In the Spiral Array, a major triad can be "flipped" to become a minor triad; in other words, the reflection of a major triad with respect to any one of its sides is a minor triad.

Example 1: The C Major Chord Representation

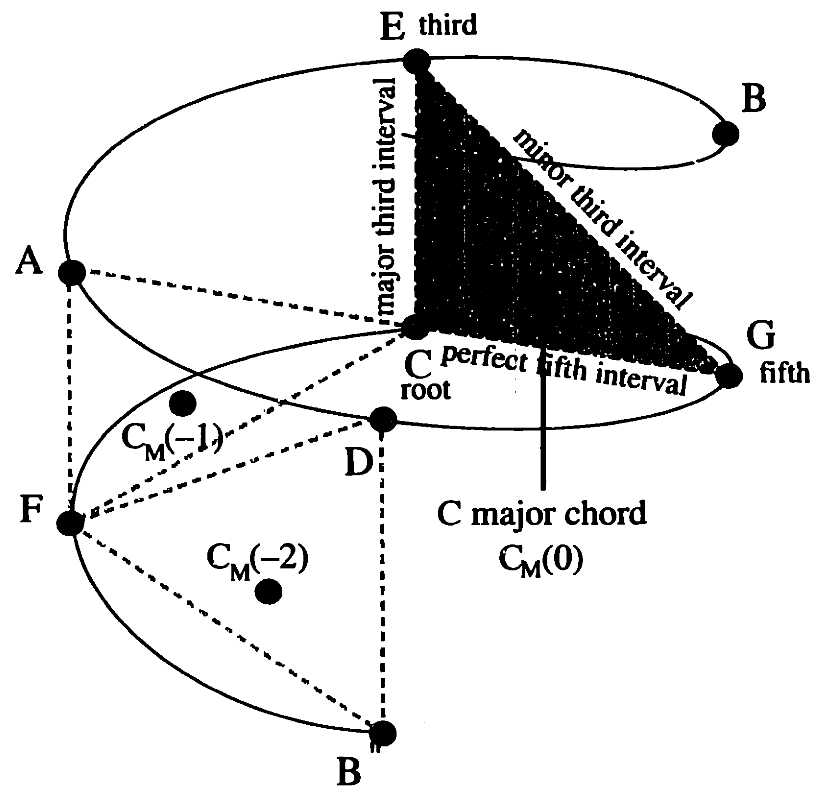


Figure 3-6: Examples of major chord representations.

As shown in Figure 3-6, the C major triad consists of C, E and G. The root, C, is denoted by $\mathbf{P}(0)$; the fifth, G, is a perfect fifth above C, $\mathbf{P}(1)$; and, the third, E, is a major third, or four perfect fifths, above C, $\mathbf{P}(4)$. The root of the chord, C, is connected to the sides representing a perfect fifth interval and a major third interval. The triad is represented by a point on the triangle, $C_M(0)$. The index of the chord corresponds to the index of its root, 0.

Similar representations can be generated for all major triads. These representations also sit on a spiral. This major chord spiral is a little tighter than the outermost pitch spiral, and resides completely inside the outer spiral.

Example 2: The A Minor Chord Representation

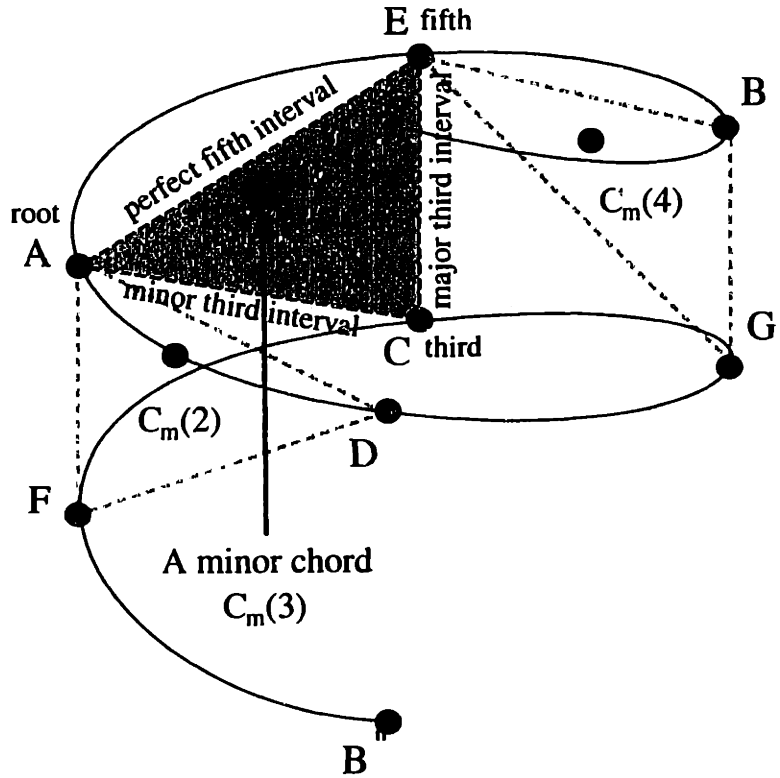


Figure 3-7: Examples of minor chord representations.

As shown in Figure 3-7, the A minor triad consists of A, C and E. The root, A, is denoted by $\mathbf{P}(3)$; the fifth, E, is a perfect fifth above A, $\mathbf{P}(4)$; and, the third, C, is a minor third, or three perfect fifths, below A, $\mathbf{P}(3-3) = \mathbf{P}(0)$. The root of the chord, A, is connected to the sides representing a perfect fifth interval and a minor third interval. The triad is represented by a point on the triangle, $\mathbf{C}_m(3)$. The index of the chord corresponds to the index of its root, 3.

Similar chord representations can be generated for all minor triads. Again, these representations sit on a spiral that is a little tighter than the outermost pitch spiral, and resides completely inside that outer spiral.

Mathematical Representation

Mathematically, the chord's representation is generated by a linear combination of its three component pitch positions. The combination is restricted to be convex. Mathematically, this means that the weights on the pitch positions are non-negative and sum to one. Geometrically, this means that the chord representation resides strictly within the boundaries of the triangle outlined by the triad (see Figure 3-5). Hence, the chord is represented by a weighted average of the pitch positions: the root **P**(k), the fifth **P**(k+1), and the third **P**(k+4) for major triads, and **P**(k-3) for minor triads.

Definition 3. *The representation for a major triad is generated by the convex combination of its root, fifth and third pitch positions:*

$$\mathbf{C}_{\mathrm{M}}(\mathrm{k}) \stackrel{\mathrm{def}}{=} w_1 \cdot \mathbf{P}(\mathrm{k}) + w_2 \cdot \mathbf{P}(\mathrm{k}+1) + w_3 \cdot \mathbf{P}(\mathrm{k}+4), \tag{3.4}$$

$$\textit{where} \quad w_1 \geq w_2 \geq w_3 > 0 \textit{ and } \sum_{i=1}^{3} w_i = 1.$$

The minor triad is generated by a similar combination:

$$\mathbf{C}_{\mathrm{m}}(\mathrm{k}) \stackrel{\mathrm{def}}{=} u_1 \cdot \mathbf{P}(\mathrm{k}) + u_2 \cdot \mathbf{P}(\mathrm{k}+1) + u_3 \cdot \mathbf{P}(\mathrm{k}-3), \tag{3.5}$$

$$\textit{where} \quad u_1 \geq u_2 \geq u_3 > 0 \textit{ and } \sum_{i=1}^{3} u_i = 1.$$

The weights, w_i and u_i, on the pitch positions represent the importance of the pitch to the generated chord. For longstanding psychological, physical and theoretical reasons which are better explained elsewhere (for example [5, 45]), the root is deemed the most important, followed by the fifth, then the third. As such, I limit the weights to be monotonically decreasing from the root, to the fifth, to the third. In order that spatial distance mirrors these relations, there are additional constraints on the aspect ratio h/r which will be discussed in Section 4.2.

Spatial Geometry

As mentioned, the set of points that represent the center of effect for major triads also sit on a spiral, one with a slightly smaller radius than the one for pitches. The chord representations inherit the property of making one full turn in four steps, so that the fourth step lines the current chord vertically above the starting one. Hence, the roots of neighboring major triads are a perfect fifth apart, and the roots of vertical major triad neighbors are a major third apart. For example, for the C major triad, its nearest right neighbor that is also a major triad is the G major triad, and its nearest top major triad neighbor is the E major triad.

Theorem 1. *The representation for each major chord can be described as a 90-degree clockwise rotation of the previous major chord followed by a vertical elevation of h units:*

$$\mathbf{C_M}(\mathrm{k}+1) = \mathbf{R}\cdot\mathbf{C_M}(\mathrm{k}) + \mathbf{h}, \ \textit{where } \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ \textit{and } \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{h} \end{bmatrix}. \tag{3.6}$$

The chord representations whose roots form a major third interval are related by a vertical translation of four elevation steps:

$$\mathbf{C_M}(\mathrm{k}+4) = \mathbf{C_M}(\mathrm{k}) + 4\cdot\mathbf{h}. \tag{3.7}$$

[Proof]: Using definitions for chords as given in Equation 3.4 and for pitch positions as given in Equation 3.1 for

$$\begin{aligned}
\mathbf{C_M}(\mathrm{k}+1) &= w_1\cdot\mathbf{P}(\mathrm{k}+1) + w_2\cdot\mathbf{P}(\mathrm{k}+2) + w_3\cdot\mathbf{P}(\mathrm{k}+5) \\
&= w_1\cdot(\mathbf{RP}(\mathrm{k})+\mathbf{h}) + w_2\cdot(\mathbf{RP}(\mathrm{k}+1)+\mathbf{h}) + w_3\cdot(\mathbf{RP}(\mathrm{k}+4)+\mathbf{h}) \\
&= \mathbf{R}\cdot[w_1\cdot\mathbf{P}(\mathrm{k}) + w_2\cdot\mathbf{P}(\mathrm{k}+1) + w_3\cdot\mathbf{P}(\mathrm{k}+4)] + (w_1+w_2+w_3)\cdot\mathbf{h} \\
&= \mathbf{R}\cdot\mathbf{C_M}(\mathrm{k}) + \mathbf{h}
\end{aligned}$$

The second part of the theorem follows directly from the first. Each step adds a quarter rotation and elevation h, thus, four steps later, the chord is now 4h above where it began.

□

Theorem 2. *Similarly, the representation for each minor chord can be described as a 90-degree clockwise rotation of the previous minor chord followed by a vertical elevated of h units:*

$$\mathbf{C_m}(\mathrm{k}+1) = \mathbf{R}\cdot\mathbf{C_m}(\mathrm{k}) + \mathbf{h}, \ \textit{where } \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ \textit{and } \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{h} \end{bmatrix}. \tag{3.8}$$

The representations for chords whose roots form a major third interval are related by a vertical translation of four elevation steps:

$$\mathbf{C_m}(\mathrm{k}+4) = \mathbf{C_m}(\mathrm{k}) + 4\cdot\mathbf{h}. \tag{3.9}$$

[Proof]: Similar to that of Theorem 1. □

The minor triads neighboring a major chord are known as the relative, and parallel and mediant minors. The major triads neighboring a minor chord are known as the relative, parallel and submediant majors.

3.4 Key Representations

An important property of the Spiral Array is that representations of pitches in a given key still occupy a compact neighborhood. In contrast to the Harmonic Network, only one such neighborhood corresponds to the set of pitches in the key. The added dimension in the Spiral Array allows for the spatial representation of a center of effect that is confined within this neighborhood.

The Major Key Representation

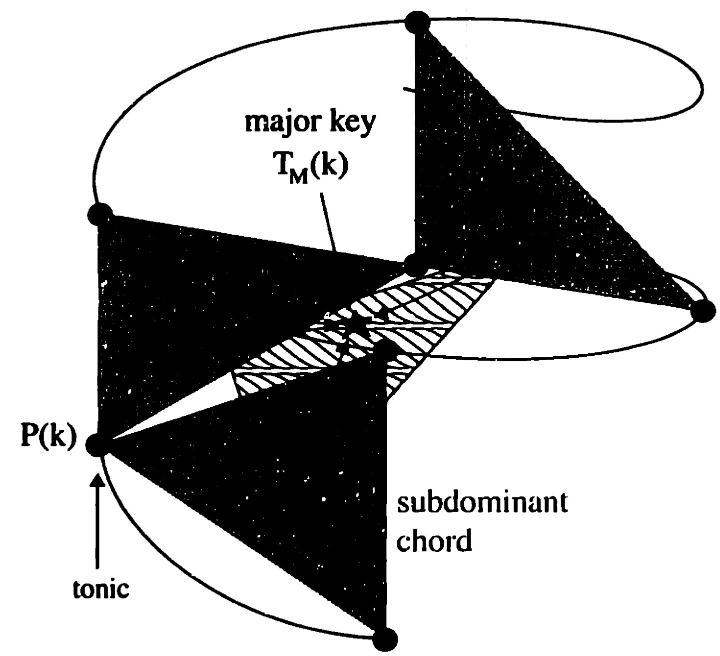


Figure 3-8: Geometric representation of a a major key, a composite of its I, V and IV chords.

Each major chord, together with its right and left neighbor major chords, combine to produce the effect of a major key[3]. In music terminology, these chords are given names, with respect to the key, that reflect their function. The center chord is called the tonic chord (I)[4], the one to its right the dominant (V), and the one to its left the subdominant (IV). Hence, I represent the major key as a combination of its I, V and IV chords. For example, the representation of the C major key is generated by the C major, G major and F major chord representations. See Figure 3-8 for an example of a major key representation.

[3]For a deeper discussion on the connection between chords and keys, see for example [42, 43].

[4]I shall use roman numerals to denote chord function within a key. The number indicates the scale degree of the chord's root. For example, "I" represents the tonic chord. I adopt the convention of denoting major chords by upper case roman numerals, and minor chords by lower case ones. For example, a major chord with the tonic as root is "I" but a minor chord with the same root is "i".

This definition of the center of effect for a key recognizes the bottom-up and top-down nature of the relationship between chords and keys. Chords, depending on their relationships, can generate a key. At the same time, a key gives rise to certain chord relationships with specific functions within the key.

Mathematical Representation

Mathematically, the representation for a major key, $\mathbf{T}_\mathrm{M}(\mathrm{k})$ is the weighted average of its tonic triad ($\mathbf{C}_\mathrm{M}(\mathrm{k})$), dominant triad ($\mathbf{C}_\mathrm{M}(\mathrm{k}+1)$) and subdominant triad ($\mathbf{C}_\mathrm{M}(\mathrm{k}-1)$) representations. As before, the design objective is to have the weights correspond to each chord's significant in the key. Hence, the I chord is given the largest weight, followed by that of the V chord, then the IV chord.

Definition 4. *A major key is represented by a convex combination of its tonic, dominant and subdominant chords. The weights are restricted to be monotonic and non-increasing.*

$$\begin{aligned} \mathbf{T}_\mathrm{M}(\mathrm{k}) &\stackrel{\mathrm{def}}{=} \omega_1 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}) + \omega_2 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}+1) + \omega_3 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}-1), \qquad (3.10) \\ \textit{where} \quad & \omega_1 \geq \omega_2 \geq \omega_3 > 0 \textit{ and } \sum_{i=1}^{3} \omega_i = 1. \end{aligned}$$

The Minor Key Representation

The definition for the minor key is more complicated. This is because there are three types of minor modes: the natural minor, harmonic minor, and melodic minor. In each case, the tonic triad is minor. However, the dominant triad is minor in the case of the natural, major in the case of the harmonic, and could be either major or minor in the case of the melodic. The subdominant triad is minor in the case of the natural, minor in the case of the harmonic, and major or minor in the case of the melodic. Hence I propose that center of effect for the minor key $\mathbf{T}_\mathrm{m}(\mathrm{k})$ be modeled as a combination of the tonic $\mathbf{C}_\mathrm{m}(\mathrm{k})$, the possible dominant triads $\mathbf{C}_\mathrm{M}(\mathrm{k}+1)$ and $\mathbf{C}_\mathrm{m}(\mathrm{k}+1)$, and the possible subdominant triad $\mathbf{C}_\mathrm{m}(\mathrm{k}-1)$ and $\mathbf{C}_\mathrm{M}(\mathrm{k}-1)$. See Figure 3-9 for the spatial representation of a minor key.

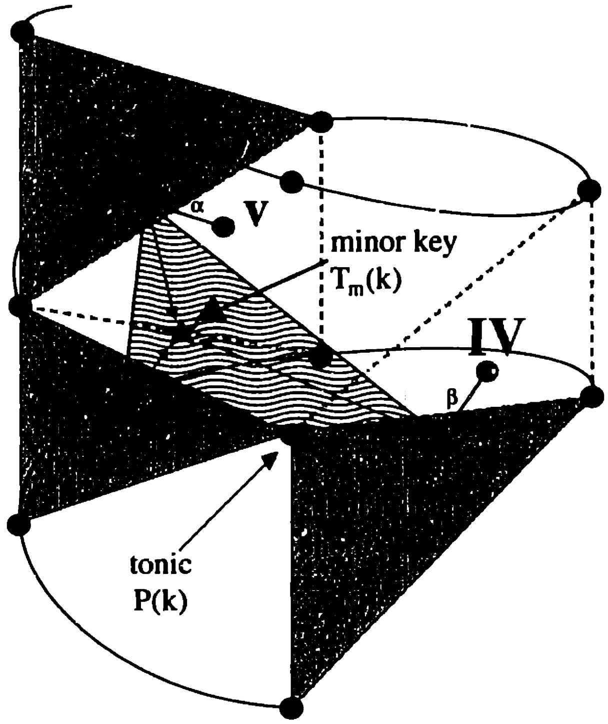


Figure 3-9: Geometric representation of a minor key, a composite of its tonic (i), dominants (V/v) and subdominant (iv/IV) chords.

Definition 5. *The minor key representation is generated by a convex combination of its tonic, dominant (major or minor) and subdominant (major or minor) chords as follows:*

$$\begin{aligned}
\mathbf{T}_m(k) &\stackrel{\text{def}}{=} v_1 \cdot \mathbf{C}_m(k) \\
&\quad + v_2 \cdot [\alpha \cdot \mathbf{C}_M(k+1) + (1-\alpha) \cdot \mathbf{C}_m(k+1)] \\
&\quad + v_3 \cdot [\beta \cdot \mathbf{C}_m(k-1) + (1-\beta) \cdot \mathbf{C}_M(k-1)], \qquad (3.11) \\
\textit{where} &\quad v_1 \geq v_2 \geq v_3 > 0 \textit{ and } v_1 + v_2 + v_3 = 1, \\
\textit{and} &\quad 0 \geq \alpha \geq 1, 0 \geq \beta \geq 1.
\end{aligned}$$

When one focusses on only the dominant major (and not minor) triad, and only on the subdominant minor (and not major) triad, i.e. $\alpha = 1$ and $\beta = 1$, the combination of chords generates the effect of a harmonic minor. When one considers only minor triads, i.e. $\alpha = 0$ and $\beta = 1$, the i, v and iv chords generate the effect of a natural minor. The melodic minor has the composite effect of some combination of all chords represented in the definition.

Spatial Geometry

Again, as a result of the definitions, all points representing the major keys sit on yet another spiral, one with an even smaller radius than the one for major triads. The major key representations inherit the propery of a full rotation in four steps so that the tonics of neighboring major keys are a perfect fifth apart, and the tonics of vertical neighbor major keys are a major third apart. For example, the C major key has, for neighboring major keys, G major to its right, F major to its left, E major above and A♭ major below.

Theorem 3. *Each major key representation can be described as a 90-degree clockwise rotation of the previous major key followed by a vertical elevation of h units:*

$$\mathbf{T}_\mathrm{M}(\mathrm{k}+1) = \mathbf{R}\cdot\mathbf{T}_\mathrm{M}(\mathrm{k}) + \mathbf{h}, \text{ where } \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{h} \end{bmatrix}. \tag{3.12}$$

The representations of keys whose tonics form a major third interval are related by a vertical translation of four elevation steps:

$$\mathbf{T}_\mathrm{M}(\mathrm{k}+4) = \mathbf{T}_\mathrm{M}(\mathrm{k}) + 4\cdot\mathbf{h}. \tag{3.13}$$

[Proof]: Using the definition for major keys as given in Equation 3.10, and invoking Theorem 1,

$$\begin{aligned}
\mathbf{T}_\mathrm{M}(\mathrm{k}+1) &= \omega_1\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}+1) + \omega_2\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}+2) + \omega_3\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}) \\
&= \omega_1\cdot(\mathbf{R}\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}) + \mathbf{h}) + \omega_2\cdot(\mathbf{R}\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}+1) + \mathbf{h}) \\
&\quad +\omega_3\cdot(\mathbf{R}\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}-1) + \mathbf{h}) \\
&= \mathbf{R}\cdot[\omega_1\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}) + \omega_2\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}+1) + \omega_3\cdot\mathbf{C}_\mathrm{M}(\mathrm{k}-1)] + (\omega_1+\omega_2+\omega_3)\cdot\mathbf{h} \\
&= \mathbf{R}\mathbf{T}_\mathrm{M}(\mathrm{k}) + \mathbf{h}
\end{aligned}$$

The second part of the theorem follows directly from the first. □

Theorem 4. *Similarly, each minor key can be described as a 90-degree clockwise rotation of the previous minor key followed by a vertical elevation of h units:*

$$\mathbf{T}_\mathrm{m}(\mathrm{k}+1) = \mathbf{R}\cdot\mathbf{T}_\mathrm{m}(\mathrm{k}) + \mathbf{h}, \text{ where } \mathbf{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{h} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{h} \end{bmatrix}. \tag{3.14}$$

The representations for chords whose roots form a major third interval are related by a vertical translation of four elevation steps:

$$\mathbf{T}_\mathrm{m}(\mathrm{k}+4) = \mathbf{T}_\mathrm{m}(\mathrm{k}) + 4\cdot\mathbf{h}. \tag{3.15}$$

[Proof]: Using the definition for minor keys as given in Equation 3.12, and invoking Theorems 1 and 2,

$$\begin{aligned}
\mathbf{T}_{\mathrm{m}}(\mathrm{k}+1) &= v_1 \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}+1) \\
&\quad + v_2 \cdot [\alpha \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}+2) + (1-\alpha) \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}+2)] \\
&\quad + v_3 \cdot [\beta \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}) + (1-\beta) \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k})] \\
&= v_1 \cdot (\mathbf{R} \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}) + \mathbf{h}) \\
&\quad + v_2 \cdot [\alpha \cdot (\mathbf{R} \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}+1) + \mathbf{h}) + (1-\alpha) \cdot (\mathbf{R} \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}+1) + \mathbf{h})] \\
&\quad + v_3 \cdot [\beta \cdot (\mathbf{R} \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}-1) + \mathbf{h}) + (1-\beta) \cdot (\mathbf{R} \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}-1) + \mathbf{h})] \\
&= \mathbf{R} \cdot \{v_1 \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}) \\
&\quad + v_2 \cdot [\alpha \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}+1) + (1-\alpha) \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}+1)] \\
&\quad + v_3 \cdot [\beta \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}-1) + (1-\beta) \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}-1)]\} + (v_1 + v_2 + v_3) \cdot \mathbf{h} \\
&= \mathbf{R} \cdot \mathbf{T}_{\mathrm{m}}(\mathrm{k}) + \mathbf{h}
\end{aligned}$$

The second part of the theorem follows directly from the first. □

The minor keys neighboring a major key are its relative, parallel and mediant minor keys; and the major keys neighboring a minor are its relative, parallel and submediant major keys.

3.5 The Gamut: A Discussion

The Spiral Array presents a new way to represent and relate musical entities from different hierarchical levels. Its design proposes a new way to generate these entities in the same space. Pitch positions form the basis for both chord and, by inference, key representations. The Spiral Array representations are distinct from those in the Harmonic Network in that the different hierarchical entities are depicted by the same kind of elements - points in space. This similarity offers opportunities for comparison and relation among the different elements. In the tabular arrangement, intervals could be represented by vectors in the Harmonic Network. This is no longer true in the three-dimensional Spiral Array (nor in a toroid configuration). One would need new ways to measure intervals.

A major strength of the Spiral Array lies in its symmetry, regularity and transformational invariance across all hierarchical levels. Figure 3-10 juxtaposes pitch, chord and key positions on the same space. Pitches are represented by positions at every quarter turn of the spiral. As shown by Theorems 1 and 2, major and minor chords are each represented by positions punctuating each quarter turn of their respective spirals. As shown by Theorems 3 and 4, the same is true of major and minor keys. All of these spirals ascend vertically by h units at every quarter turn.

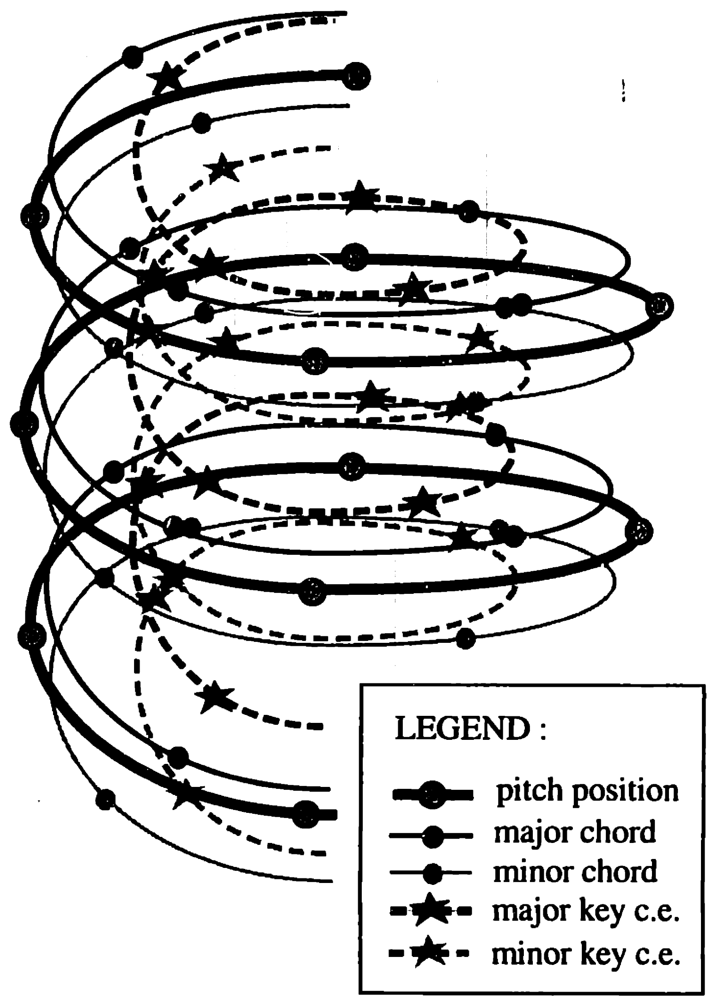


Figure 3-10: Juxtaposing on the same space pitch positions, major and minor chord representations, and positions representing the major and minor keys.

Summary of Definitions

For ease of reference when perusing other chapters, this section lists a compilation of the definitions given in this chapter:

A pitch representation:

$$\mathbf{P}(\mathrm{k}) = \begin{bmatrix} x_\mathrm{k} \\ y_\mathrm{k} \\ z_\mathrm{k} \end{bmatrix} = \begin{bmatrix} \mathrm{r}\sin\frac{\mathrm{k}\pi}{2} \\ \mathrm{r}\cos\frac{\mathrm{k}\pi}{2} \\ \mathrm{kh} \end{bmatrix}.$$

A major chord representation:

$$\begin{aligned} \mathbf{C}_\mathrm{M}(\mathrm{k}) &= w_1 \cdot \mathbf{P}(\mathrm{k}) + w_2 \cdot \mathbf{P}(\mathrm{k}+1) + w_3 \cdot \mathbf{P}(\mathrm{k}+4), \\ \text{where} &\quad w_1 \geq w_2 \geq w_3 > 0 \text{ and } \sum_{i=1}^{3} w_i = 1. \end{aligned}$$

A minor chord representation:

$$\begin{aligned} \mathbf{C}_\mathrm{m}(\mathrm{k}) &= u_1 \cdot \mathbf{P}(\mathrm{k}) + u_2 \cdot \mathbf{P}(\mathrm{k}+1) + u_3 \cdot \mathbf{P}(\mathrm{k}-3), \\ \text{where} &\quad u_1 \geq u_2 \geq u_3 > 0 \text{ and } \sum_{i=1}^{3} u_i = 1. \end{aligned}$$

A major key representation:

$$\begin{aligned} \mathbf{T}_\mathrm{M}(\mathrm{k}) &= \omega_1 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}) + \omega_2 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}+1) + \omega_3 \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}-1), \\ \text{where} &\quad \omega_1 \geq \omega_2 \geq \omega_3 > 0 \text{ and } \sum_{i=1}^{3} \omega_i = 1. \end{aligned}$$

A minor key representation:

$$\begin{aligned} \mathbf{T}_\mathrm{m}(\mathrm{k}) &= v_1 \cdot \mathbf{C}_\mathrm{m}(\mathrm{k}) \\ &\quad + v_2 \cdot [\alpha \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}+1) + (1-\alpha) \cdot \mathbf{C}_\mathrm{m}(\mathrm{k}+1)] \\ &\quad + v_3 \cdot [\beta \cdot \mathbf{C}_\mathrm{m}(\mathrm{k}-1) + (1-\beta) \cdot \mathbf{C}_\mathrm{M}(\mathrm{k}-1)], \\ \text{where} &\quad v_1 \geq v_2 \geq v_3 > 0 \text{ and } v_1 + v_2 + v_3 = 1, \\ \text{and} &\quad 0 \geq \alpha \geq 1, 0 \geq \beta \geq 1. \end{aligned}$$

4 — Model Distances

The goal in this chapter is to select parameters for the spatial representations in the Spiral Array so that relations between represented tonal entities have direct counterparts in the geometric structure. Distance is the measurement for perceived "closeness" between the tonal entities. Note that closeness between pitches in the Spiral Array is different from closeness between pitches sorted by frequency, for example. on a keyboard. Proximity in the Spiral Array indicates some combination of the following: shared pitches, shared intervals, or tonal elements within a perfect fifth, major third or minor third interval of each other.

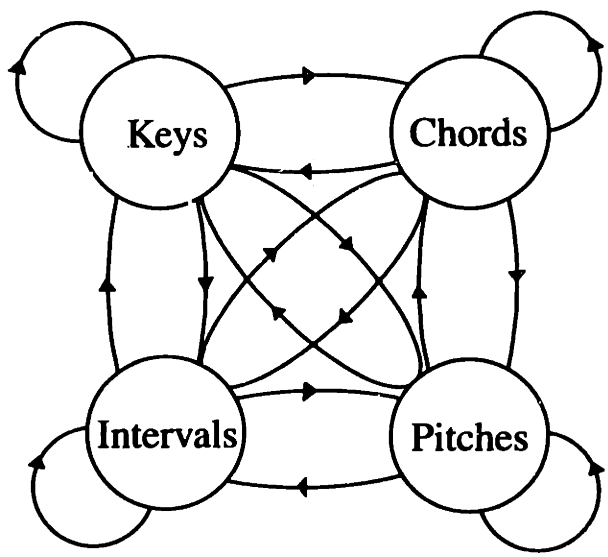


Figure 4-1: Relations that can be assigned to and inferred from the model.

Because pitches, chords and keys are represented in the Spiral Array, the model allows one to relate musical entities from different heirarchies, for example, pitches to pitches, pitches to chords and keys to chords. Figure 4-1 shows the relationships that can be assigned to and inferred from the Spiral Array.

The Objects of Study

In this chapter, the objects of study are the parameters:

$$\begin{aligned}
\mathrm{r} &= \text{radius of the spiral} \\
\mathrm{h} &= \text{height gain per quarter turn of the spiral} \\
w &= [w_1, w_2, w_3], \text{ weights on major chord pitches} \\
u &= [u_1, u_2, u_3], \text{ weights on minor chord pitches} \\
\omega &= [\omega_1, \omega_2, \omega_3], \text{ weights on major key chords} \\
v &= [v_1, v_2, v_3], \text{ weights on minor key chords} \\
\alpha &= \text{preference for V vs. v chord in minor key} \\
\beta &= \text{preference for iv vs. IV chord in minor key}
\end{aligned}$$

The formal definitions of chord and key representations in the Spiral Array space (presented in Chapter 3) constrain, in a useful manner, some of the relations between the tonal elements. However, in order for the model to reflect preceived tonal relations more precisely, a few more constraints will have to be added. In this chapter, I demonstrate the selection of model parameters based on one set of basic criteria. By assigning parameters that conform to these criteria, the Spiral Array aims to model tonal concepts in accordance with perceived pitch relations in western classical music.

4.1 Interval Relations

In this section, I examine the distance relations between pairs of pitch representations, and calibrate the model, so that pitch proximity reflects perceived interval relations in western tonal music.

The Aspect Ratio

The pitch positions are completely defined in the Spiral Array structure by the radius of the spiral and the vertical elevation at each step; these parameters are denoted by r and h respectively. As a result, the aspect ratio $\frac{h}{r}$ determines the distance relationships among the pitch positions. The design goal is that perceptually close intervals should be represented by shorter inter-pitch distances.

Scalar vs. Vector Representation of Intervals

In the Spiral Array, interval pairs are represented by a scalar (distance); this definition is unique. Intervals could be represented by vectors, but then their representation would not be unique; for instance, a vector from $\mathbf{P}(k)$ to $\mathbf{P}(k+1)$ represents a perfect fifth interval, so does a vector from **P**(k−1) to **P**(k). Figure 4-2 shows, geometrically, some interval representations.

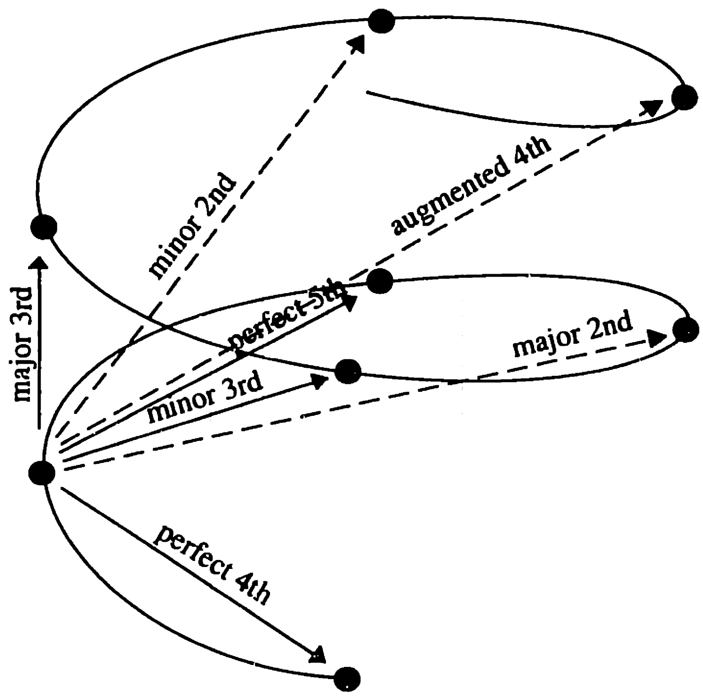


Figure 4-2: Intervals represented on the Spiral Array.

Interval Preferences

In western tonal music, the interval of a fifth is considered the closest harmonic relationship between distinct pitches. Hence, pitches related by a perfect fifth should be nearest each other. Other considerations are: pitches a third apart are perceived to be more closely related than those a second apart; the major is regarded to be closer than the minor for both thirds and seconds; and, the tritone (diminished fifth or augmented fourth) is perceived to be more distant than these other interval classes[1]. These are the conditions against which I will calibrate the model.

Lemma 1. *Consider the interval classes: {(P4,P5), (M3,m6), (m3,M6), (M2,m7), (m2,M7), (d5,A4)}. Inherent in the model are the following interval distance relationships: [1] the (P4,P5) interval class is represented by distances closer than all but the (M3,m6) interval class; [2] (M3/m6) and (m3/M6) are both closer than the (m2/M7) interval class; and, [3] (d5/A4) is the farthest of all.*

[1] By *interval classes* I mean intervals which are equivalent under transformation. For example, a minor sixth interval is the same as a major third interval under inversion; and, a compound interval such as a major tenth, when transposed to the same octave, is equivalent to a major third. The naming convention used here is: P=perfect, M=major, m=minor, d=diminished, A=Augmented.

Interval		Distance^2
Name	Class	
Perfect fifth, perfect fourth	(P4 , P5)	$2r^2 + h^2$
Major third, minor sixth	(M3 , m6)	$16h^2$
Minor third, major sixth	(m3 , M6)	$2r^2 + 9h^2$
Major second, minor seventh	(M2 , m7)	$4r^2 + 4h^2$
Minor second, major seventh	(m2 , M7)	$2r^2 + 25h^2$
Diminished fifth, augmented fourth	(d5 , A4)	$4r^2 + 36h^2$

Table 4.1: Distances associated with interval relations.

[Proof]: The relationships follow naturally from the distances representing each interval class as given in Table 4.1. □

Theorem 5. *In order that: [1] the closest distance between any two pitch positions denote a perfect fifth; [2] pitches a third apart are closer than those a second apart; [3] a major interval is closer than a minor; and [4] a tritone is more distant than any of the above, the aspect ratio $\frac{h}{r}$ should be constrained so that:*

$$\frac{2}{15} < a < \frac{2}{7} \qquad \text{where } a = \frac{h^2}{r^2}. \tag{4.1}$$

[Proof]: Condition [4] is always satisfied, according to Condition [3] in Lemma 1. The following inequality summarizes the remaining conditions:

$$\text{dist(P4,P5)} < \text{dist(M3,m6)} < \text{dist(m3,M6)} < \text{dist(M2,m7)} < \text{dist(m2,M7)}. \tag{4.2}$$

Denoting the distance between any two points (a,b) as d(a,b), the first inequality in Equation 4.2 becomes:

$$\begin{aligned} d((\mathrm{P4,P5})) &< d((\mathrm{M3,m6})) \\ \Rightarrow 2r^2 + h^2 &< (4h)^2 \\ \Rightarrow 2r^2 &< 15h^2, \end{aligned}$$

and, the second inequality implies that:

$$\begin{aligned} d((\mathrm{M3,m6})) &< d((\mathrm{m3,M6})) \\ \Rightarrow (4h)^2 &< 2r^2 + 9h^2 \\ \Rightarrow 7h^2 &< 2r^2. \end{aligned}$$

The next two inequalities follow from the first two:

$$\begin{aligned} d(\mathrm{m3,M6}) &= 2r^2 + 9h^2 = 2r^2 - 7h^2 + 16h^2 \\ &< 2r^2 - 7h^2 + 2r^2 + 9h^2 = 4r^2 + 2h^2, \text{ by the second inequality} \\ &< 4r^2 + 4h^2 = d((\mathrm{M2,m7})), \end{aligned}$$

$$
\begin{aligned}
d(\text{M2,m7}) &= 4\text{r}^2 + 4\text{h}^2 \\
&< 2\text{r}^2 + 15\text{h}^2 + 4\text{h}^2 \text{ by the first inequality} \\
&< 2\text{r}^2 + 25\text{h}^2 = d((\text{m2,M7})).
\end{aligned}
$$

Hence, inequalities combine to give the desired constraint on the aspect ratio:

$$0.3651 = \sqrt{\frac{2}{15}} < \frac{\text{h}}{\text{r}} < \sqrt{\frac{2}{7}} = 0.5345.$$

□

4.2 Major Chord-Chord Pitch Relations

In this section, I derive the conditions on the aspect ratio and the triad weights so that chord pitch to chord distances in the Spiral Array correspond to a set of simple perceived relations.

Desired Distance Relationship

Consider the three pitches that constitute a major triad: its root, fifth and third. In conventional tonal harmony, theorists consider the root the most important pitch in the chord, followed by the fifth, then the third. The goal is to restrict the parameters so that each chord center is closest to its root, followed by the fifth, then the third.

Major triad weights are given priority in the selection process since the behavior of major modes is better understood, and the minor mode is considered by many to be derived from the major.

Definition of Major Chord Weights

By design (see Definition 3), as reflected in the weights, the pitch with greatest influence on each chord representation is its root, followed by the fifth, then the third. Consider the constraints on the weights:

$$
\begin{aligned}
w_1 + w_2 + w_3 &= 1 \\
\text{and} \qquad w_1 > w_2 > w_3 &= 1 - w_1 - w_2 > 0 \text{ (from the above line).}
\end{aligned}
$$

Since w_3 can be written as a function of w_1 and w_2, the above expressions imply the following system of three inequalities on (w_1, w_2) which are shown graphically on Figure 4-3:

$$w_1 > w_2 \tag{4.3}$$

$$w_1 > -2w_2 + 1, \tag{4.4}$$

$$\text{and} \qquad w_1 < -w_2 + 1. \tag{4.5}$$

The shaded area, $\mathbf{F_0}$, shows the intersection of the three constraints.

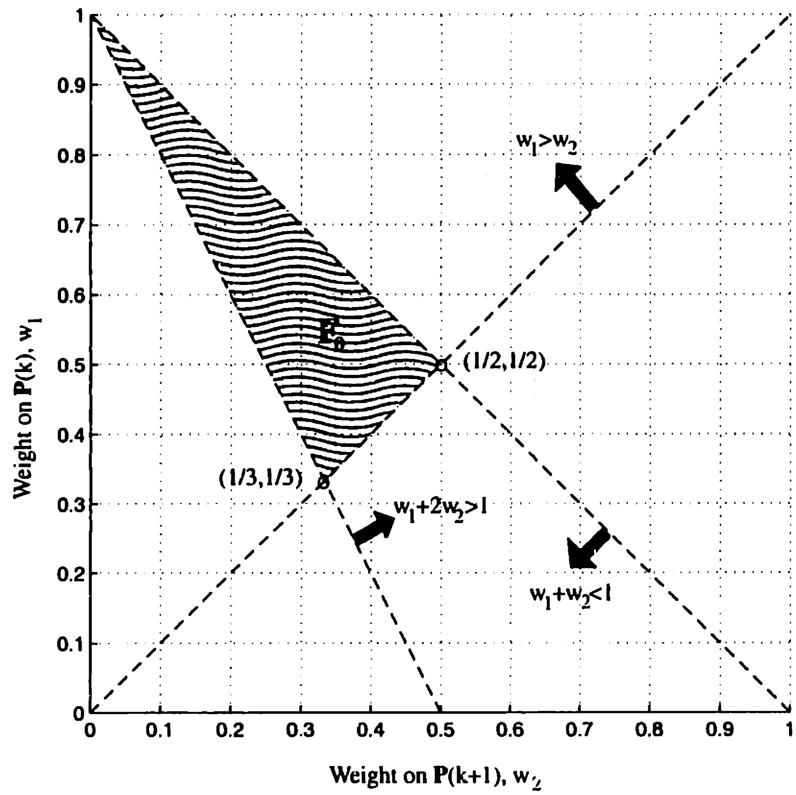


Figure 4-3: System of inequalities defining feasible values of major chord weights as given by the definition in Equation 3.4. A similar set of inequalities applies to the weights on minor chords.

Reason for Further Analysis

Having the heaviest assigned weight does not guarantee that the root will be nearest to the chord representation. See Figure 4-4 for an example in which the point with least weight is actually closest to the center of effect. Thus, new constraints have to be derived so that proximity corresponds to perceived relations.

Constraints on Major Chord Weights

Allowing the major chord weights their full range of values within $\mathbf{F_0}$ (shown in Figure 4-3), I derive the conditions on the aspect ratio so that the representations of the major triad's root, fifth and third are closer than other pitches to the triad's center.

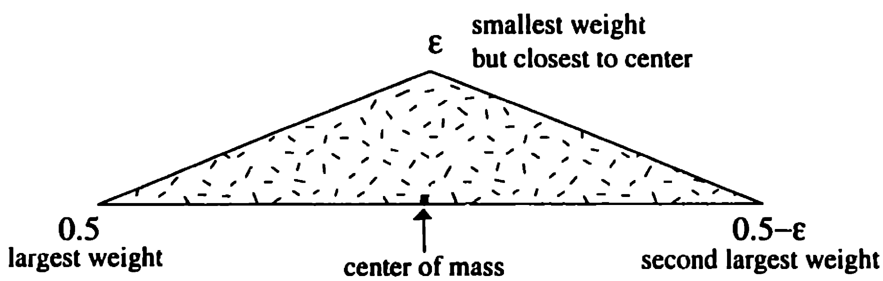


Figure 4-4: An example in which the point with least weight is closest to the center of effect. (Note that in this figure, ε is a small number.)

First, I show that the condition from Theorem 5, $\sqrt{\frac{2}{15}} < \frac{h}{r} < \sqrt{\frac{2}{7}}$, is a necessary and sufficient condition for the major chord's representation to be closest to its root, followed by the fifth, then the third (Theorem 6). Surprisingly, the constraint on the aspect ratio is identical for both Theorem 5 and Theorem 6. So, in the Spiral Array, restricting the perfect fifth to be the closest relation, followed by the major third, minor third, major second and minor second, is tentamount to constraining the major chord representation to be closest to the root, followed by the fifth, then the third.

In the next section, I derive the new constraints on the aspect ratio so that no other pitches are as close to the major chord center as its three constituent pitches.

Theorem 6. *A major chord's representation is closest to its root, followed by its fifth, then its third for all* $(w_1, w_2) \in \mathbf{F_0}$ *if and only if*

$$\frac{2}{15} < a < \frac{2}{7}, \qquad \textit{where} \qquad a = \frac{h^2}{r^2}.$$

[Proof]: The goal is to find the aspect ratio values that satisfy the theorem conditions for all possible choice of (w_1, w_2) in the feasible region, $\mathbf{F_0}$. The theorem states that: for all $(w_1, w_2) \in \mathbf{F_0}$,

$$\begin{aligned} \frac{2}{15} < a < \frac{2}{7} \iff \text{dist(chord, root)} &< \text{dist(chord, fifth)} < \text{dist(chord, third)}, \\ \text{i.e. } \|\mathbf{C_M}(k) - \mathbf{P}(k)\| &< \|\mathbf{C_M}(k) - \mathbf{P}(k+1)\| < \|\mathbf{C_M}(k) - \mathbf{P}(k+4)\|. \end{aligned}$$

Due to the high level of symmetry in the model, it is sufficient to simply consider the case of the C major chord, $\mathbf{C_M}(0)$. The related fifth and third are $\mathbf{P}(1)$ and $\mathbf{P}(4)$ respectively. Their spatial positions are given by:

$$\mathbf{P}(0) = \begin{bmatrix} 0 \\ \mathrm{r} \\ 0 \end{bmatrix}, \mathbf{P}(1) = \begin{bmatrix} \mathrm{r} \\ 0 \\ \mathrm{h} \end{bmatrix} \text{ and } \mathbf{P}(4) = \begin{bmatrix} 0 \\ \mathrm{r} \\ 4\mathrm{h} \end{bmatrix}. \tag{4.6}$$

By definition (see Equation 3.4), the C major chord position is located at:

$$\begin{aligned}\mathbf{C}_{\mathrm{M}}(0) &= w_1 \cdot \mathbf{P}(0) + w_2 \cdot \mathbf{P}(1) + w_3 \cdot \mathbf{P}(4) \\ &= w_1 \begin{bmatrix} 0 \\ \mathrm{r} \\ 0 \end{bmatrix} + w_2 \begin{bmatrix} \mathrm{r} \\ 0 \\ \mathrm{h} \end{bmatrix} + w_3 \begin{bmatrix} 0 \\ \mathrm{r} \\ 4\mathrm{h} \end{bmatrix} = \begin{bmatrix} w_2 \cdot \mathrm{r} \\ (w_1 + w_3) \cdot \mathrm{r} \\ (w_2 + 4w_3) \cdot \mathrm{h} \end{bmatrix}. \end{aligned} \tag{4.7}$$

Define a new variable for the standardized distance between $\mathbf{C}_{\mathrm{M}}(0)$ and a pitch:

$$y_{\mathrm{M}}(s) \stackrel{\mathrm{def}}{=} \frac{1}{r^2} \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(s)\|^2 \tag{4.8}$$

The standardized distances of each chord pitch from the chord are given by:

$$\begin{aligned} y_{\mathrm{M}}(0) &= \frac{1}{\mathrm{r}^2} \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(0)\|^2 \\ &= w_2^2 + (w_1 + w_3 - 1)^2 + (w_2 + 4w_3)^2 \cdot a \\ &= 2 \cdot w_2^2 + (4 - 3w_2 - 4w_1)^2 \cdot a, \end{aligned} \tag{4.9}$$

$$\begin{aligned} y_{\mathrm{M}}(1) &= \frac{1}{\mathrm{r}^2} \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(1)\|^2 \\ &= (w_2 - 1)^2 + (w_1 + w_3)^2 + (w_2 + 4w_3 - 1)^2 \cdot a \\ &= 2 \cdot (1 - w_2)^2 + (3 - 3w_2 - 4w_1)^2 \cdot a, \end{aligned} \tag{4.10}$$

and

$$\begin{aligned} y_{\mathrm{M}}(4) &= \frac{1}{\mathrm{r}^2} \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4)\|^2 \\ &= w_2^2 + (w_1 + w_3 - 1)^2 + (w_2 + 4w_3 - 4)^2 \cdot a \\ &= 2 \cdot w_2^2 + (3w_2 + 4w_1)^2 \cdot a. \end{aligned} \tag{4.11}$$

The desired proximity conditions with respect to the chord center, $\mathbf{C}_{\mathrm{M}}(0)$, are that the root, $\mathbf{P}(0)$, should be closer than the fifth, $\mathbf{P}(1)$, which is in turn closer than the third, $\mathbf{P}(4)$. Mathematically, this means that:

$$\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(0)\|^2 < \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(1)\|^2 \iff y_{\mathrm{M}}(0) < y_{\mathrm{M}}(1) \tag{4.12}$$

and

$$\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(1)\|^2 < \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4)\|^2 \iff y_{\mathrm{M}}(1) < y_{\mathrm{M}}(4). \tag{4.13}$$

These inequalities, simplified, yield the following constraints:

[1] substituting Equations 4.9 and 4.10 into Equation 4.12, one gets

$$\begin{aligned} 2 \cdot w_2^2 - 2 \cdot (1 - w_2)^2 + (4 - 3w_2 - 4w_1)^2 \cdot a - (3 - 3w_2 - 4w_1)^2 \cdot a &< 0 \\ -2 + 7a + (4 - 6a) \cdot w_2 - 8aw_1 &< 0 \end{aligned}$$

$$\begin{aligned} \iff w_1 &> \left(\frac{4 - 6a}{8a}\right) w_2 - \left(\frac{2 - 7a}{8a}\right) \\ &= a_0 \cdot w_2 + \frac{1}{2}(1 - a_0), \qquad \text{where } a_0 = \frac{1}{2a} - \frac{3}{4}. \end{aligned} \tag{4.14}$$

[2] the result of substituting Equations 4.10 and 4.11 into Equation 4.13 is

$$\begin{aligned} 2 \cdot (1 - w_2)^2 - 2 \cdot w_2^2 + (3 - 3w_2 - 4w_1)^2 \cdot a - (3w_2 + 4w_1)^2 \cdot a &< 0 \\ (2 + 9a) - (4 + 18a) \cdot w_2 - 24aw_1 &< 0 \end{aligned}$$

$$\begin{aligned} \iff w_1 &> -\left(\frac{4 + 18a}{24a}\right) w_2 + \left(\frac{2 + 9a}{24a}\right) \\ &= -a_1 w_2 + \frac{1}{2} a_1, \qquad \text{where } a_1 = \frac{3}{4} + \frac{1}{6a}. \end{aligned} \tag{4.15}$$

Hence, the distance constraints given in Equations 4.12 and 4.13 are satisfied if and only if Equations 4.14 and 4.15 are true. The two latter inequalities (4.14 and 4.15) relate the aspect ratio to the chord weights so that the chord representations are closest to the root, followed by the fifth, then the third.

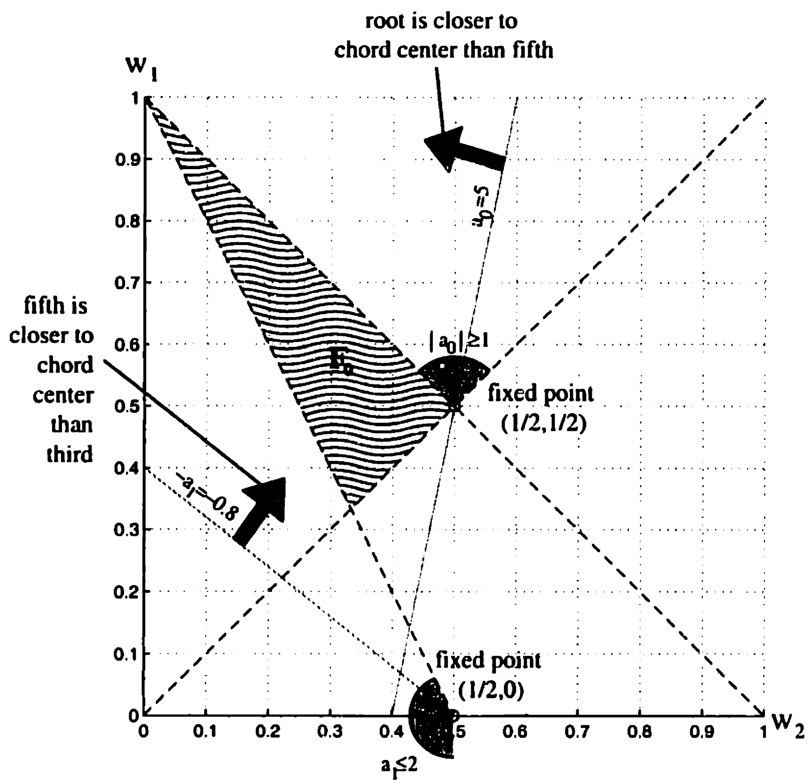


Figure 4-5: Feasible values for major chord weights, (w_1, w_2), based on the desired proximity relations between a major chord and its component pitches.

Graphically, these conditions are satisfied if $\mathbf{F}_1$, the feasible region defined by Equations 4.14 and 4.15, contains $\mathbf{F}_0$. This relationship is illustrated in Figure 4-5.

The line generated by Equation 4.14 has a fixed point at ($w_2 = \frac{1}{2}, w_1 = \frac{1}{2}$), a vertex of the triangular region $\mathbf{F}_0$, as shown in Figure 4-5. The line passes through this point for all values of a_0, the slope of the line. This line forms one of the boundaries for the feasible region, $\mathbf{F}_1$.

The corresponding line from Equation 4.15 has a fixed point at ($w_2 = \frac{1}{2}, w_1 = 0$), passing through this point for all values of a_1. a_1 is the slope of the line that pivots on this fixed point. This line forms the other boundary for the feasible region, $\mathbf{F}_1$.

$\mathbf{F}_0$ will always be contained in $\mathbf{F}_1$ if and only if

$$|a_0| \geq 1 \qquad \text{and} \qquad a_1 \leq 2.$$

Figure 4-5 gives a graphical interpretation. Solving for a:

$$\begin{aligned} a_0 = \frac{1}{2a} - \frac{3}{4} \geq 1 \quad &\text{or} \quad a_0 = \frac{1}{2a} - \frac{3}{4} \leq -1 \\ \frac{1}{2a} \geq \frac{7}{4} \quad &\text{or} \quad \frac{1}{2a} \leq -\frac{1}{4} \\ a \leq \frac{2}{7} \quad &\text{or} \quad a \leq -2 \text{ (an impossibility)}, \end{aligned} \tag{4.16}$$

and

$$\begin{aligned} a_1 = \frac{3}{4} + \frac{1}{6a} &\leq 2 \\ \frac{1}{6a} &\leq \frac{5}{4} \Rightarrow a \geq \frac{2}{15}. \end{aligned} \tag{4.17}$$

Combining Equations 4.16 and 4.17:

$$\mathbf{F}_0 \text{ is a subset of } \mathbf{F}_1 \iff \frac{2}{15} \leq a \leq \frac{2}{7},$$

which is the very same constraint as that given in Theorem 5! □

4.3 Major Chord-All Pitch Relations

The theorem in this section derives the conditions on the aspect ratio so that no other pitches are as close to the major chord center as its three constituent pitches. The theorem is preceded by four lemmas that consider the proximity, to the C major triad $\mathbf{C}_M(0)$, of each of the four types of pitches, $\mathbf{P}(4n)$, $\mathbf{P}(4n+1)$, $\mathbf{P}(4n+2)$ and $\mathbf{P}(4n+3)$, where n is an integer.

Lemma 2. *Consider pitches of the type $\mathbf{P}(4n)$, for example, F♭, A♭, C, E, G♯ and so on. $\mathbf{P}(0)$ is closest to $\mathbf{C}_M(0)$, i.e. C is closest to C major triad, in the Spiral Array, followed by $\mathbf{P}(4)$ (pitch E) for all $(w_1,w_2) \in \mathbf{F}_0$ if and only if $a < \frac{2}{7}$.*

[Proof]: By Equations 3.2 and 4.7:

$$\begin{aligned} y_{\mathrm{M}}(4n) &= \frac{1}{r^2}\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4n)\|^2 \\ &= 2 \cdot w_2^2 + (4w_1 + 3w_2 + 4n - 4)^2 \cdot a \\ &= 2 \cdot w_2^2 + 16a \cdot \left(n - \frac{4 - 4w_1 - 3w_2}{4}\right)^2. \end{aligned} \tag{4.18}$$

Consider the function $(4w_1 + 3w_2)$. For all $(w_1, w_2) \in \mathbf{F}_0$, as shown in Figure 4-3, the function is minimized at $(\frac{1}{3}, \frac{1}{3})$ and maximized at $(1, 0)$ to give the range $\frac{7}{3} < 4w_1 + 3w_2 < 4$.

Relaxing the integer constraint on n, the quadratic function where n is not constrained to be integer, $\tilde{y}(4n)$ has a minimum at $\tilde{n}^* = \frac{1}{4}(4 - 4w_1 - 3w_2)$. But $0 < \frac{1}{4}(4 - 4w_1 - 3w_2) < \frac{5}{12}$.

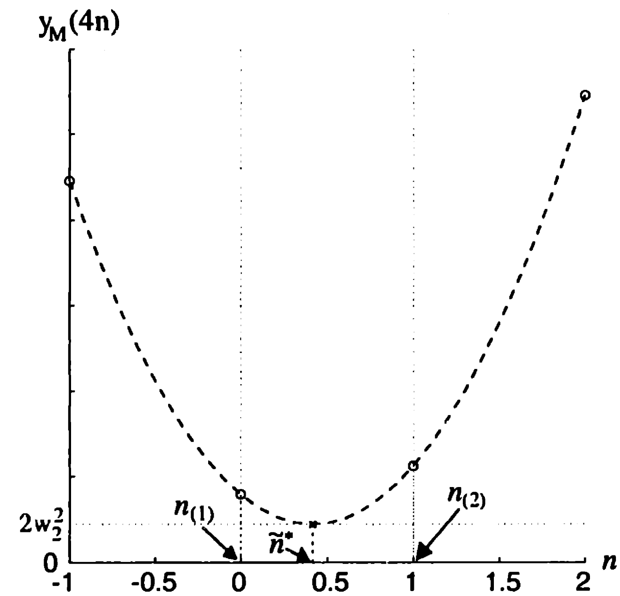


Figure 4-6: A graph of the function $y_{\mathrm{M}}(4n) = \frac{1}{r^2}\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4n)\|^2$ when $r = 1$. This parabola explains the choice of integer $n_{(1)}$ that minimizes the function.

Now, let $n_{(k)}$ denote the kth least value of $y_{\mathrm{M}}(s)$. Observe in Figure 4-6 that because $0 < \frac{1}{4}(4 - 4w_1 - 3w_2) < \frac{5}{12}$, $\tilde{n}^*$ lies between 0 and $\frac{5}{12}$. Hence, the minimum value for $y_{\mathrm{M}}(4n)$ occurs when $n_{(1)} = 0$ at $y_{\mathrm{M}}(0)$, followed by the second least value at $n_{(2)} = 1$ at $y_{\mathrm{M}}(4)$.

In general, $y_{\mathrm{M}}(4n) < y_{\mathrm{M}}(-4n)$ and $y_{\mathrm{M}}(4n) < y_{\mathrm{M}}(4(n+1))$ for non-negative n as shown in Figure 4-6. Therefore,

$$\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(0)\|^2 < \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4)\|^2 < \|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(4n)\|^2, n \neq 0, 1.$$

$\square$

Lemma 3. *Consider pitches of the type* $\mathbf{P}(4n+1)$*, for example, C♭, E♭, G, B, D♯ and so on.* $\mathbf{P}(1)$ *is closest to* $\mathbf{C}_M(0)$*, i.e. G is closest to the C major triad. The second closest pitch is farther than* $\mathbf{P}(4)$ *(E, the third in the triad) for all* $(w_1, w_2) \in \mathbf{F_0}$ *if and only if* $a < \frac{2}{7}$*. When* $4w_1 + 3w_2 > 3$*, this second closest pitch is* $\mathbf{P}(-3)$ *(pitch E♭); when* $4w_1 + 3w_2 < 3$*, the second is* $\mathbf{P}(5)$ *(pitch B), which is always farther than* $\mathbf{P}(4)$*.*

[Proof]: By Equations 3.2 and 4.7:

$$\begin{aligned} y_M(4n+1) &= \frac{1}{r^2}\|\mathbf{C}_M(0) - \mathbf{P}(4n+1)\|^2 \\ &= 2\cdot(1-w_2)^2 + (4w_1 + 3w_2 - 3 + 4n)^2 \cdot a. \end{aligned} \tag{4.19}$$

Getting the expression for $y_M(4)$ from Equation 4.18,

$$\begin{aligned} & y_M(4n+1) - y_M(4) \\ &= 2\cdot(1-w_2)^2 + (4w_1+3w_2-3+4n)^2\cdot a - 2\cdot w_2^2 - (4w_1+3w_2)^2\cdot a \\ &= 2 - 4w_2 + a(-3+4n)(8w_1+6w_2-3+4n) \\ &= 2 - 4w_2 + 16a\left(n - \frac{3}{4}\right)\left(n - \frac{3-8w_1-6w_2}{4}\right) \end{aligned}$$

Because the weights are constrained to be in the feasible region $\mathbf{F_0}$ in Figure 4-3, we know that $-\frac{15}{12} < \frac{1}{4}(3 - 8w_1 - 6w_2) < -\frac{5}{12}$, that is to say, $-\frac{15}{12} < \tilde{n}^* < -\frac{5}{12}$. Consequently, $n_{(1)} = 0$. According to Theorem 6, $y_M(0) < y_M(1) < y_M(4)$.

Depending on the weights (w_1, w_2), the second closest pitch can be indexed either $n_{(2)} = -1$ or $n_{(2)} = 1$, that is to say, the pitch can be either $\mathbf{P}(-3)$ or $\mathbf{P}(5)$. The boundary for this change occurs at $\frac{1}{4}(3 - 8w_1 - 6w_2) = -\frac{3}{4}$. This boundary is shown on Figure 4-7.

When $\frac{1}{4}(3 - 8w_1 - 6w_2) \le -\frac{3}{4}$, $n_{(2)} = -1$:

$$\begin{aligned} y_M(-3) - y_M(4) &= 2 - 4w_2 + 16a\left(-\frac{7}{4}\right)\left(-1 - \frac{3-8w_1-6w_2}{4}\right) \\ &= 2 + 49a - (4+42a)w_2 - 56aw_1 > 0 \end{aligned}$$

$$\begin{aligned} \Longleftrightarrow w_1 &< -\left(\frac{4+42a}{56a}\right)\cdot w_2 + \left(\frac{2+49a}{56a}\right) \\ &= -a_3 w_2 + \frac{1}{2}(a_3+1), \qquad \text{where } a_3 = \frac{1}{14a} + \frac{3}{4}. \end{aligned}$$

Relating this in inequality to the graph in Figure 4-7, the line $w_1 = -a_3 w_2 + \frac{1}{2}(a_3 + 1)$ has a fixed point at $(w_2, w_1) = (\frac{1}{2}, \frac{1}{2})$. The inequality $w_1 < -a_3 w_2 + \frac{1}{2}(a_3 + 1)$ contains the feasible region $\mathbf{F_{0B}}$ if and only if $a_3 > 1$, which means that $a < \frac{2}{7}$.

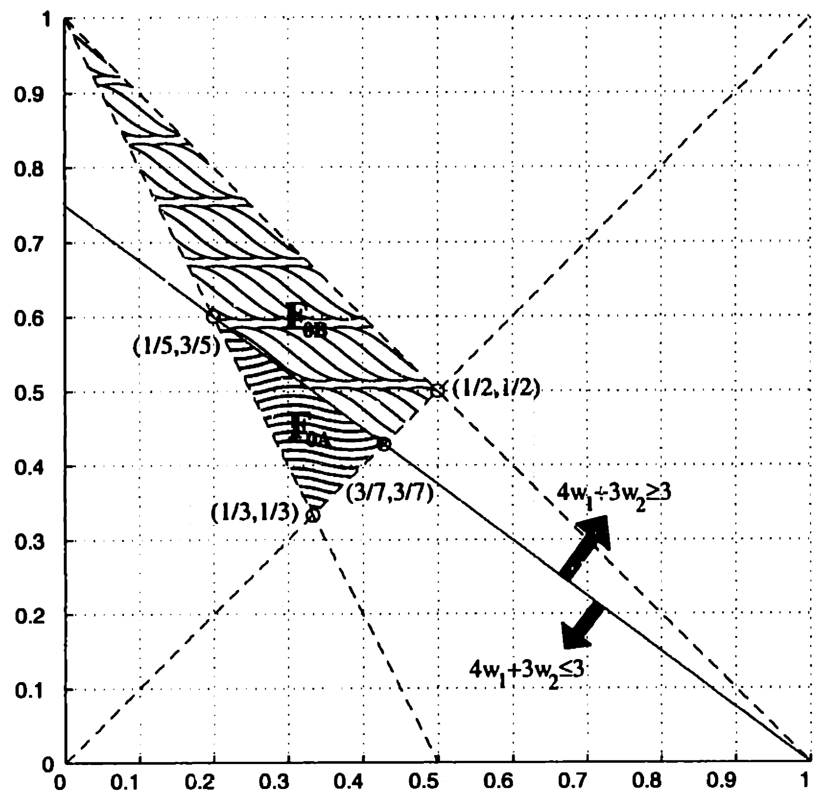


Figure 4-7: Feasible values of (w_2, w_1) and the boundary, $4w_1 + 3w_2 = 3$, between two subsets of weights in the analysis of $y_M(4n+1) - y_M(4)$.

When $\frac{1}{4}(3 - 8w_1 - 6w_2) > -\frac{3}{4}$, $n_{(2)} = 1$.

$$\begin{aligned} y_M(5) - y_M(4) &= 2 - 4w_2 + 16a\left(\frac{1}{4}\right)\left(1 - \frac{3 - 8w_1 - 6w_2}{4}\right) \\ &= 2 + a + 8aw_1 - (4 - 6a)w_2 \end{aligned}$$

$$\begin{aligned} \Longleftrightarrow w_1 &> \left(\frac{4-6a}{8a}\right)\cdot w_2 - \left(\frac{2+a}{8a}\right) \\ &= a_0 \cdot w_2 - \frac{1}{2}(a_0 + 1), \qquad \text{recall that } a_0 = \frac{1}{2a} - \frac{3}{4}. \end{aligned}$$

The fixed point for $w_1 = a_0 \cdot w_2 - \frac{1}{2}(a_0 + 1)$ occurs at $(w_2, w_1) = (\frac{1}{2}, -\frac{1}{2})$. The inequality $w_1 > a_0 \cdot w_2 - \frac{1}{2}(a_0 + 1)$ contains the feasible region $\mathbf{F}_{0A}$ if and only if $a_0 < -11/3$. This means that $a > -\frac{6}{35}$, which is always true. □

Lemma 4. *Consider pitches of the type* $\mathbf{P}(4n+2)$*, for example, G♭, B♭, D, F♯, A♯ and so on.* $\mathbf{P}$*(2) is closest to* $\mathbf{C}_M(0)$*, i.e. D is closest to the C major triad.* $\mathbf{P}$*(2) is farther than* $\mathbf{P}$*(4) (E, the third in the triad) for all* $(w_1, w_2) \in \mathbf{F}_0$ *if and only if* $a < \frac{3}{15}$.

[Proof]: By Equations 3.2 and 4.7:

$$\begin{aligned} y_M(4n+2) &= \frac{1}{r^2}\|\mathbf{C}_M(0) - \mathbf{P}(4n+2)\|^2 \\ &= 2 \cdot w_2^2 - 4w_2 + 4 + (4w_1 + 3w_2 - 2 + 4n)^2 \cdot a, \end{aligned} \tag{4.20}$$

$$\begin{aligned} & y_M(4n+2) - y_M(4) \\ = {} & 2 \cdot w_2^2 - 4w_2 + 4 + (4w_1 + 3w_2 - 2 + 4n)^2 \cdot a - 2w_2^2 - (4w_1 + 3w_2)^2 \cdot a \\ = {} & 4 \cdot (1 - w_2) + 16a \cdot \left(n - \frac{1}{2}\right)\left(n - \frac{2 - 8w_1 - 6w_2}{4}\right). \end{aligned}$$

Since $-\frac{3}{2} < \frac{1}{4}(2 - 8w_1 - 6w_2) < -\frac{2}{3}$, the index of the distance-minimizing pitch, $n_{(1)} = 0$ for all $(w_1, w_2) \in \mathbf{F}_0$.

$$\begin{aligned} y_M(2) - y_M(4) &= 4(1 - w_2) + 16a \cdot \left(-\frac{1}{2}\right)\left(-\frac{2 - 8w_1 - 6w_2}{4}\right) \\ &= 4(1 + a) - (4 + 12a) \cdot w_2 - 16aw_1 > 0, \end{aligned}$$

$$\begin{aligned} \iff w_1 &< -\left(\frac{1 + 3a}{4a}\right) \cdot w_2 + \left(\frac{1 + a}{4a}\right) \\ &= -a_4 \cdot w_2 + (a_4 - \frac{1}{2}), \qquad \text{where } a_4 = \frac{1}{4a} + \frac{3}{4}. \end{aligned}$$

The line $w_1 = -a_4 \cdot w_2 + (a_4 - \frac{1}{2})$ has a fixed point at $(w_2, w_1) = (1, -\frac{1}{2})$. In order that $w_1 < -a_4 \cdot w_2 + (a_4 - \frac{1}{2})$ contains $\mathbf{F}_0$, $a_4 > 2$. This is true if and only if $a < \frac{3}{15}$. □

Lemma 5. *Consider pitches of the type* $\mathbf{P}(4n+3)$*, for example, D♭, F, A, C♯, E♯ and so on. For all* $(w_1, w_2) \in \mathbf{F}_0$*, when* $4w_1 + 3w_2 < 3$*,* $\mathbf{P}(3)$ *(pitch A) is closest to* $\mathbf{C}_M(0)$*, the C major triad; and when* $4w_1 + 3w_2 > 3$*,* $\mathbf{P}(-1)$ *(pitch F) is closest to* $\mathbf{C}_M(0)$*. In either case, the closest pitch is still farther than* $\mathbf{P}(4)$ *(E, the third in the triad) if and only if*

$$3 < 4w_1 + 3w_2 < \frac{1}{5a} + \frac{5}{2} \quad \text{and} \quad 4w_1 + 3w_2 < \frac{1}{2a} + \frac{1}{2}.$$

[Proof]: By Equations 3.2 and 4.7:

$$\begin{aligned} y_M(4n+3) &= \frac{1}{r^2}\|\mathbf{C}_M(0) - \mathbf{P}(4n+3)\|^2 \\ &= 2 + 2 \cdot w_2^2 + (4w_1 + 3w_2 - 1 + 4n)^2 \cdot a. \end{aligned} \tag{4.21}$$

$$\begin{aligned} & y_M(4n+3) - y_M(4) \\ = {} & 2 + 2 \cdot w_2^2 + (4w_1 + 3w_2 - 1 + 4n)^2 \cdot a - 2 \cdot w_2^2 - (4w_1 + 3w_2)^2 \cdot a \\ = {} & 2 + 16a\left(n - \frac{1}{4}\right)\left(n - \frac{1 - 8w_1 - 6w_2}{4}\right). \end{aligned}$$

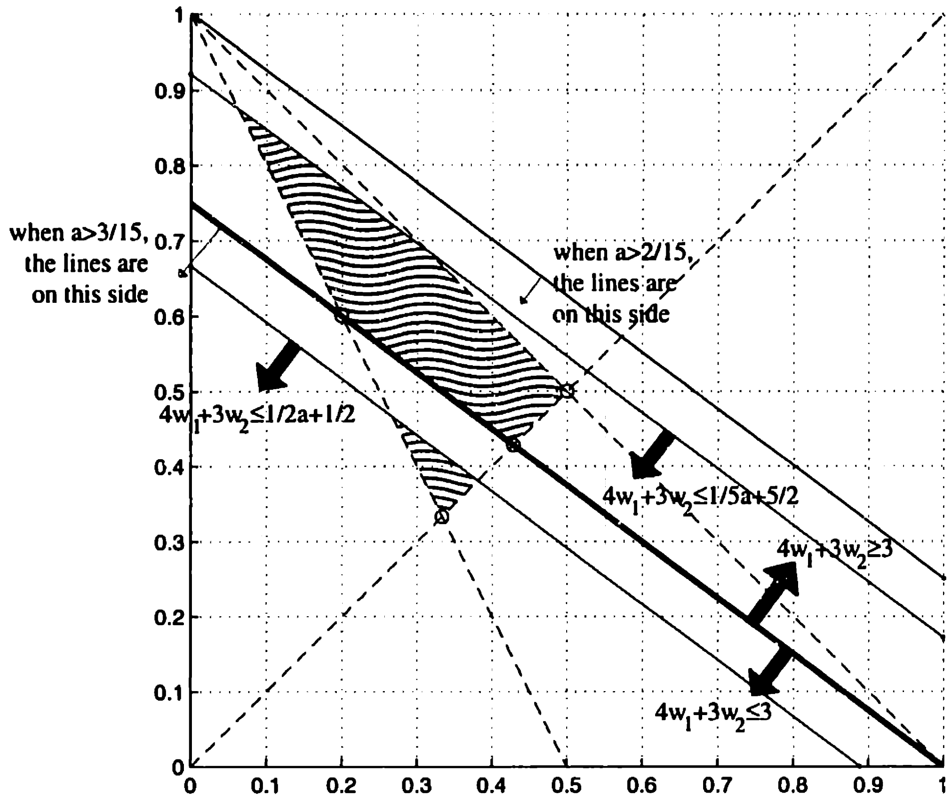


Figure 4-8: Feasible values of (w_1, w_2) based on proximity relations between a major chord and all pitches.

In the possible range, $-\frac{7}{4} < \frac{1}{4}(1-8w_1-6w_2) < \frac{1}{12}$, when $\frac{1}{4}(1-8w_1-6w_2) > -\frac{5}{4}$, $n_{(1)} = 0$; and when $\frac{1}{4}(1-8w_1-6w_2) < -\frac{5}{4}$, $n_{(1)} = -1$.

When $\frac{1}{4}(1-8w_1-6w_2) > -\frac{5}{4}$,

$$\begin{aligned} y_M(3) - y_M(4) &= 2 + 16a\left(-\frac{1}{4}\right)\left(-\frac{1-8w_1-6w_2}{4}\right) \\ &= (2+a) - 8aw_1 - 6aw_2 > 0 \end{aligned}$$

$$\iff 4w_1 + 3w_2 < \frac{2+a}{2a} = \frac{1}{2a} + \frac{1}{2}.$$

Note that when $\frac{2+a}{2a} > 3$ ($\iff a < \frac{3}{15}$), the gap between the two shaded areas in Figure 4-8 close to form one contiguous feasible region. In this case, $4w_1 + 3w_2 < \frac{1}{2a} + \frac{1}{2}$ is not a binding constraint.

When $\frac{1}{4}(1-8w_1-6w_2) < -\frac{5}{4}$,

$$y_M(-2) - y_M(4) = 2 + 16a\left(-1-\frac{1}{4}\right)\left(-1-\frac{1-8w_1-6w_2}{4}\right)$$

$$= 2 + 5a(5 - 8w_1 - 6w_2) > 0,$$

$$\iff 4w_1 + 3w_2 < \frac{1}{5a} + \frac{5}{2}$$

Note that when $a = \frac{2}{15}$, this is not a binding constraint. Figure 4-8 shows the two regions defined by these constraints. □

Theorem 7. *No other pitches are as close to a major chord's representation as its three constituent pitches if* $\frac{2}{15} < a < \frac{3}{15}$ *and* $4w_1 + 3w_2 < \frac{1}{5a} + \frac{5}{2}$.

[Proof]: Note that the chord pitch farthest from the chord representation is the third of the chord (by Theorem 6), so that for all $s \neq 0, 1, 4$:

$$\begin{aligned}&||\mathbf{C_M}(k) - \mathbf{P}(k+s)||\\ &> \max\{||\mathbf{C_M}(k) - \mathbf{P}(k)||, ||\mathbf{C_M}(k) - \mathbf{P}(k+1)||, ||\mathbf{C_M}(k) - \mathbf{P}(k+4)||\}\\ &= ||\mathbf{C_M}(k) - \mathbf{P}(k+4)||\end{aligned}$$

Thus, one need only prove that all non-chord pitches are farther from the chord than its third. By symmetry, one need only show this for C major, i.e. $\mathbf{C_M}(0)$.

By Lemma 2, $a < \frac{2}{7}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n)$ are farther than $\mathbf{P}(4)$ from $\mathbf{C_M}(0)$, for $n \neq 0, 1$. By Lemma 3, $a < \frac{2}{7}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n+1)$ are farther than $\mathbf{P}(4)$ from $\mathbf{C_M}(0)$, for $n \neq 0$. By Lemma 4, $a < \frac{3}{15}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n+2)$ are farther than $\mathbf{P}(4)$ from $\mathbf{C_M}(0)$. Given that $a < \frac{3}{15}$, by Lemma 5, $4w_1 + 3w_2 < \frac{1}{5a} + \frac{5}{2}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n+3)$ are farther than $\mathbf{P}(4)$ from $\mathbf{C_M}(0)$. □

4.4 Minor Chord-Chord Pitch Relations

In this section, I derive the additional constraints on the minor chord weights so that, given the selected aspect ratio range, $\sqrt{\frac{2}{15}} < \frac{h}{r} < \sqrt{\frac{3}{15}}$, the root is closest to the chord center, followed by the fifth, then the third, and that all three are closer than other pitches to the chord center. In the following section, I derive the new constraints so that no other pitches are as close to the minor chord center as its three constituent pitches.

Definition of Minor Chord Weights

As in the case of the weights for the major triads, because of the deinition of the minor triad representation (Definition 3), the constraints on the minor chord weights can be summarized in the three inequalities on (u_1, u_2):

$$u_1 > u_2 \tag{4.22}$$
$$u_1 > -2u_2 + 1, \tag{4.23}$$
$$\text{and} \qquad u_1 < -u_2 + 1. \tag{4.24}$$

These inequalities can be represented in a graph identical (except for the variables' names) to that in Figure 4-3. See Figure 4-9 for a graphical depiction of the feasible region for the minor triad weights, $\mathbf{F}_2$.

Theorem 8. *Given the aspect ratio constraint as given in Theorem 7, that is $\frac{2}{15} < a < \frac{3}{15}$, a minor chord's representation is closest to its root, followed by the fifth, then the third if and only if its weights, $(u_1, u_2) \in \mathbf{F}_2$ satisfy:*

$$a_0 u_1 > u_2 + \frac{1}{2}(a_0 - 1) \qquad \text{where} \qquad a_0 = \frac{1}{2a} - \frac{3}{4}.$$

The sufficient condition, when $a = \frac{2}{15}$ is:

$$u_1 > \frac{1}{3} \cdot u_2 + \frac{1}{3}.$$

[Proof]: The goal is to find the constraint on the weights $(u_1, u_2) \in \mathbf{F}_2$, for the aspect ratio range $\frac{2}{15} < a < \frac{3}{15}$, that satisfy the theorem conditions. The theorem states that: for $\frac{2}{15} < a < \frac{3}{15}$ and $(u_1, u_2) \in \mathbf{F}_2$,

$$\|\mathbf{C}_{\mathrm{m}}(k) - \mathbf{P}(k)\| < \|\mathbf{C}_{\mathrm{m}}(k) - \mathbf{P}(k+1)\| < \|\mathbf{C}_{\mathrm{m}}(k) - \mathbf{P}(k-3)\|$$
$$\iff \qquad a_0 u_1 > u_2 + \frac{1}{2}(a_0 - 1) \qquad \text{where} \qquad a_0 = \frac{1}{2a} - \frac{3}{4}.$$

Again, due to the high level of symmetry in the model, it is sufficient to simply consider the case of the C minor chord, $\mathbf{C}_{\mathrm{m}}(0)$. The related fifth and third are $\mathbf{P}(1)$ and $\mathbf{P}(-3)$ respectively. Their spatial positions are given by:

$$\mathbf{P}(0) = \begin{bmatrix} 0 \\ \mathrm{r} \\ 0 \end{bmatrix}, \mathbf{P}(1) = \begin{bmatrix} \mathrm{r} \\ 0 \\ \mathrm{h} \end{bmatrix} \text{ and } \mathbf{P}(-3) = \begin{bmatrix} \mathrm{r} \\ 0 \\ -3\mathrm{h} \end{bmatrix}. \tag{4.25}$$

By definition (see Equation 3.5), the C minor chord is spatially located at:

$$\begin{aligned} \mathbf{C}_{\mathrm{m}}(0) &= u_1 \cdot \mathbf{P}(0) + u_2 \cdot \mathbf{P}(1) + u_3 \cdot \mathbf{P}(-3) \\ &= u_1 \begin{bmatrix} 0 \\ \mathrm{r} \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} \mathrm{r} \\ 0 \\ \mathrm{h} \end{bmatrix} + u_3 \begin{bmatrix} \mathrm{r} \\ 0 \\ -3\mathrm{h} \end{bmatrix} = \begin{bmatrix} (u_2 + u_3) \cdot \mathrm{r} \\ u_1 \cdot \mathrm{r} \\ (u_2 - 3u_3) \cdot \mathrm{h} \end{bmatrix}. \end{aligned} \tag{4.26}$$

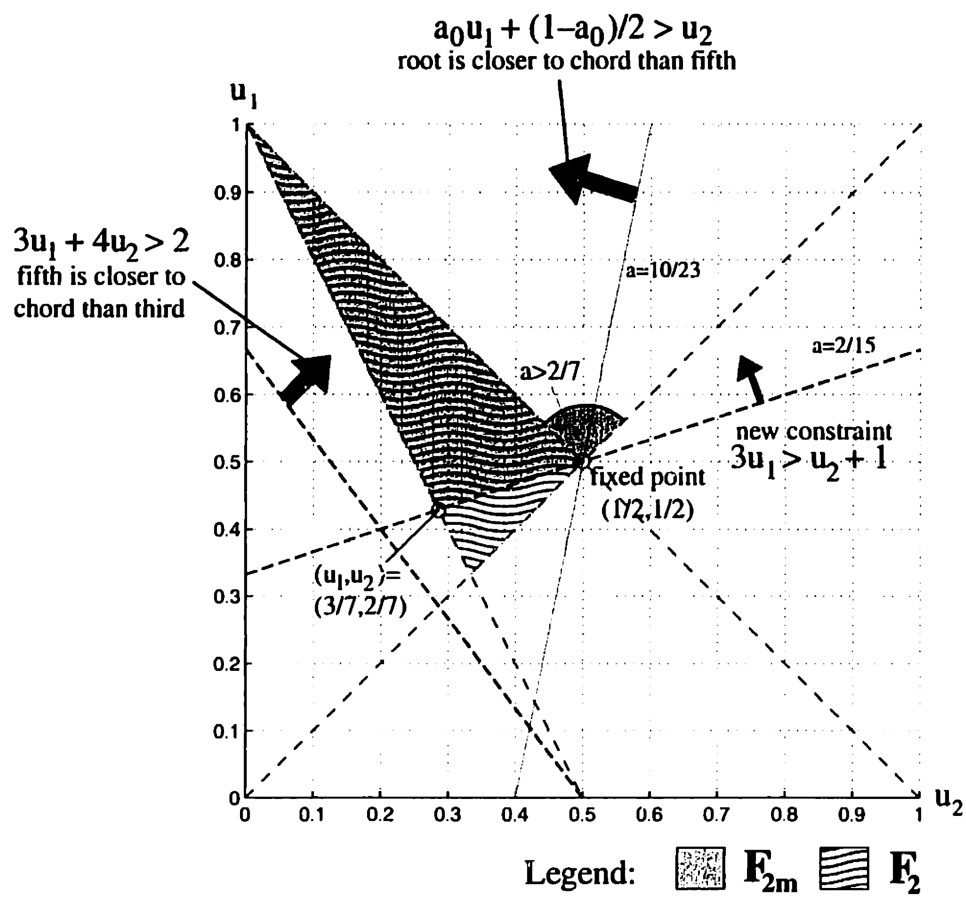


Figure 4-9: Feasible values for minor chord weights, (u_1, u_2), based on the desired proximity relations between a minor chord and its component pitches.

Define a new variable for the standardized distance between $\mathbf{C}_m(0)$ and a pitch:

$$y_m(s) \stackrel{\text{def}}{=} \frac{1}{r^2}||\mathbf{C}_m(0) - \mathbf{P}(s)||^2 \tag{4.27}$$

The standardized distances of each chord pitch from the chord are given by:

$$\begin{aligned} y_m(0) &= ||\mathbf{C}_m(0) - \mathbf{P}(0)||^2 \\ &= (u_2 + u_3)^2 \cdot r^2 + (u_1 - 1)^2 \cdot r^2 + (u_2 - 3u_3)^2 \cdot h^2 \\ &= 2 \cdot (1 - u_1)^2 + (3u_1 + 4u_2 - 3)^2 \cdot a, && (4.28) \\ y_m(1) &= ||\mathbf{C}_m(0) - \mathbf{P}(1)||^2 \\ &= (u_2 + u_3 - 1)^2 \cdot r^2 + u_1^2 \cdot r^2 + (u_2 - 3u_3 - 1)^2 \cdot h^2 \\ &= 2 \cdot u_1^2 + (3u_1 + 4u_2 - 4)^2 \cdot, a && (4.29) \\ \text{and} \quad y_m(-3) &= ||\mathbf{C}_m(0) - \mathbf{P}(-3)||^2 \end{aligned}$$

$$
\begin{aligned}
&= (u_2 + u_3 - 1)^2 \cdot \mathrm{r}^2 + u_1^2 \cdot \mathrm{r}^2 + (u_2 - 3u_3 + 3)^2 \cdot \mathrm{h}^2 \\
&= 2 \cdot u_1^2 + (3u_1 + 4u_2)^2 \cdot a. \qquad (4.30)
\end{aligned}
$$

The desired proximity conditions with respect to the chord center, $\mathbf{C}_\mathrm{m}(0)$, are:

$$
\begin{aligned}
& \|\mathbf{C}_\mathrm{m}(0) - \mathbf{P}(0)\|^2 < \|\mathbf{C}_\mathrm{m}(0) - \mathbf{P}(1)\|^2 \iff y_\mathrm{m}(0) < y_\mathrm{m}(1) \qquad (4.31) \\
\text{and} \quad & \|\mathbf{C}_\mathrm{m}(0) - \mathbf{P}(1)\|^2 < \|\mathbf{C}_\mathrm{m}(0) - \mathbf{P}(-3)\|^2 \iff y_\mathrm{m}(1) < y_\mathrm{m}(-3). \qquad (4.32)
\end{aligned}
$$

These inequalities, simplified, yield the following constraints:

[1] substituting Equations 4.28 and 4.29 into Equation 4.31, one gets

$$
\begin{aligned}
2 \cdot (1 - u_1)^2 - 2 \cdot u_1^2 + (3u_1 + 4u_2 - 3)^2 \cdot a(3u_1 + 4u_2 - 4)^2 \cdot a &< 0 \\
(2 - 7a) - (4 - 6a)u_1 + 8au_2 &< 0 \\
\iff \frac{4 - 6a}{8a}u_1 - \frac{2 - 7a}{8a} &> u_2 \\
\iff a_0 u_1 + \frac{1}{2} \cdot (1 - a_0) &> u_2, \qquad (4.33)
\end{aligned}
$$

[2] the result of substituting Equations 4.29 and 4.30 into Equation 4.32 is

$$
\begin{aligned}
(3u_1 + 4u_2 - 4)^2 \cdot a - (3u_1 + 4u_2)^2 \cdot a &< 0 \\
\iff 16 - 8(3u_1 + 4u_2) &< 0 \\
\iff 3u_1 + 4u_2 &> 2. \qquad (4.34)
\end{aligned}
$$

Hence, the distance constraints given in Equations 4.31 and 4.32 are satisfied if and only if Equations 4.33 and 4.34 are true. The inequalities in Equations 4.33 and 4.34 are shown in Figure 4-9.

The inequality in Equation 4.34 always holds true since the region it defines always includes the values of $(u_1, u_2) \in \mathbf{F}_2$. The inequality in Equation 4.33 would not cut into $\mathbf{F}_2$ if and only if

$$
\begin{aligned}
-1 < \ & \tfrac{4-6a}{8a} < 1 \\
\iff -8a < \ & 4 - 6a < 8a \\
\Rightarrow -2a < 4 \ & \text{and} \ 4 < 14a \\
\Rightarrow \ & a > \frac{2}{7}.
\end{aligned}
$$

Since $\frac{2}{15} < a < \frac{3}{15} < \frac{2}{7}$, the line $a_0 u_1 + \frac{1}{2}(1 - a_0) = u_2$ will always cut into $\mathbf{F}_2$. Hence, the inequality in Equation 4.33 is only satisfied for a subset of $(u_1, u_2) \in \mathbf{F}_2$. The constraint in Equation 4.33 is now required to augment the inequality $u_1 > u_2$.

In Equation 4.33, by setting $a = \frac{2}{15}$, which is to say, $a_0 = 3$, one gets:

$$u_2 < 3u_1 - \frac{1}{2}(1-3) = 3u_1 - 1.$$

Thus, this new constraint on the weights would always place the root closer to the chord's center than the fifth over the range of aspect ratio values $\frac{2}{15} < a < \frac{3}{15}$. □

Note that Theorem 8 effectively excludes the use of the weights $u = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. This means that in the Spiral Array model, the minor triad requires a definite emphasis (as measured by the weight) on the root if it were to be more closely related (as reflected by distance in the model) to the chord than the fifth.

4.5 Minor Chord-All Pitch Relations

The theorem in this section defines the conditions on the minor chord weights so that all other pitches are farther from the minor chord center than its three constituent pitches. The theorem is preceded by four lemmas that consider the proximity, to the C minor triad $\mathbf{C}_M(0)$, of each of the four types of pitches, $\mathbf{P}(4n)$, $\mathbf{P}(4n+1)$, $\mathbf{P}(4n+2)$ and $\mathbf{P}(4n+3)$, where n is an integer.

Lemma 6. *Consider pitches of the type $\mathbf{P}(4n)$, for example, F♭, A♭, C, E, G♯ and so on. $\mathbf{P}(0)$ is closest to $\mathbf{C}_m(0)$, i.e. C is closest to C minor triad, in the Spiral Array, followed by $\mathbf{P}(-4)$ (pitch A♭) for all $(u_1,u_2) \in \mathbf{F}_{2m}$, and $\mathbf{P}(-4)$ is always farther than $\mathbf{P}(-3)$ if and only if $a < \frac{2}{7}$.*

[Proof]: By Equations 3.2 and 4.26:

$$\begin{aligned} y_m(4n) &= \frac{1}{r^2}\|\mathbf{C}_m(0) - \mathbf{P}(4n)\|^2 \\ &= 2\cdot(1-u_1)^2 + (4n - 3u_1 - 4u_2 + 3)^2 \cdot a \qquad (4.35) \\ &= 2\cdot(1-u_1)^2 + 16a\left(n - \frac{3u_1 + 4u_2 - 3}{4}\right)^2. \end{aligned}$$

Consider the function $\frac{1}{4}(3u_1 + 4u_2 - 3)$. For all $(u_1, u_2) \in \mathbf{F}_{2m}$, as shown in Figure 4-9, the function is minimized at $(u_1, u_2) = (\frac{3}{7}, \frac{2}{7})$ and maximized at $(\frac{1}{2}, \frac{1}{2})$. This means that $-\frac{1}{7} < \frac{1}{4}(3u_1 + 4u_2 - 3) < \frac{1}{8}$. Relaxing the integer constraint on n, the function $\tilde{y}_m(4n)$ is minimized at $\tilde{n}^* = \frac{1}{4}(3u_1 + 4u_2 - 3)$, which lies between $-\frac{1}{7}$ and $\frac{1}{8}$.

Using a parabola similar to that in Figure 4-6, it follows that $n_{(1)} = 0$ and $n_{(2)} = -1$. This means that the second farthest pitch of the type $\mathbf{P}(4n)$ is $\mathbf{P}(-4)$. Now, it remains to be proved that $y_m(-4) - y_m(-3) > 0$. Using Equations 4.35 and 4.30,

$$
\begin{aligned}
y_m(-4) - y_m(-3) &= 2\cdot(1-u_1)^2 - 2\cdot u_1^2 + a\cdot(3u_1+4u_2+1)^2 - a\cdot(3u_1+4u_2)^2 \\
&= (2+a) + (6a-4)\cdot u_1 + 8a\cdot u_2 > 0 \\
&\iff \frac{1}{2}(a_0+1) + u_2 > a_0 u_1, \qquad \text{where } a_0 = \frac{1}{2a} - \frac{3}{4}.
\end{aligned}
$$

The equation $a_0 u_1 = \frac{1}{2}(a_0+1)+u_2$ has a fixed point at $(u_1,u_2) = (\frac{1}{2}, -\frac{1}{2})$. This line pivots around its fixed point. The inequality $a_0 u_1 < \frac{1}{2}(a_0+1)+u_2$ and does not interfere with the feasible region $\mathbf{F}_{2m}$ (or for that matter, $\mathbf{F}_2$), shown in Figure 4-9, if and only if $a_0 > 1$.

But $a_0 > 1$ simply means that $a < \frac{2}{7}$. □

Lemma 7. *Consider pitches of the type $\mathbf{P}(4n+1)$, for example, C♭, E♭, G, B, D♯ and so on. $\mathbf{P}(1)$ is closest to $\mathbf{C}_m(0)$, i.e. G is closest to the C minor triad. The second closest pitch $\mathbf{P}(-3)$ (E♭, the third in the triad) for all $(u_1,u_2) \in \mathbf{F}_{2m}$, and all other pitches $\mathbf{P}(4n+1)$ are farther than $\mathbf{P}(-3)$.*

[Proof]: By Equations 3.2 and 4.26

$$
\begin{aligned}
y_m &= \frac{1}{r^2}\|\mathbf{C}_m(0) - \mathbf{P}(4n+1)\|^2 \\
&= 2\cdot u_1^2 + [4n+1-(3u_1+4u_2-3)]^2\cdot a \\
&= 2\cdot u_1^2 + 16a\left(n - \frac{3u_1+4u_2-4}{4}\right)^2. \qquad (4.36)
\end{aligned}
$$

Consider the function $\frac{1}{4}(3u_1+4u_2-4)$. For all $(u_1,u_2) \in \mathbf{F}_{2m}$, as shown in Figure 4-9, Since $-\frac{1}{7} < \frac{1}{4}(3u_1+4u_2-3) < \frac{1}{8}$, this means that $-\frac{11}{28} < \frac{1}{4}(3u_1+4u_2-4) < -\frac{1}{8}$, and $\tilde{n}^*$ lies between $-\frac{11}{28}$ and $-\frac{1}{8}$. It then follows that $n_{(1)} = 0$ and $n_{(2)} = -1$. In general, $y_m(4n) < y_m(-4n+1)$ for non-negative n. Therefore,

$$y_m(0) < y_m(-1) < y_m(0) < y_m(4n+1), n \neq 0, -1.$$

□

Lemma 8. *Consider pitches of the type $\mathbf{P}(4n+2)$, for example, G♭, B♭, D, F♯, A♯ and so on. For all $(u_1,u_2) \in \mathbf{F}_{2m}$, when $3u_1+4u_2 > 3$, $\mathbf{P}(2)$ (pitch D) is closest to $\mathbf{C}_m(0)$, the C minor triad; and when $3u_1+4u_2 < 3$, $\mathbf{P}(-2)$ (pitch B♭) is closest to $\mathbf{C}_m(0)$. In either case, the closest pitch is still farther than $\mathbf{P}(-3)$ (E♭, the third in the triad) if and only if $a < \frac{3}{15}$.*

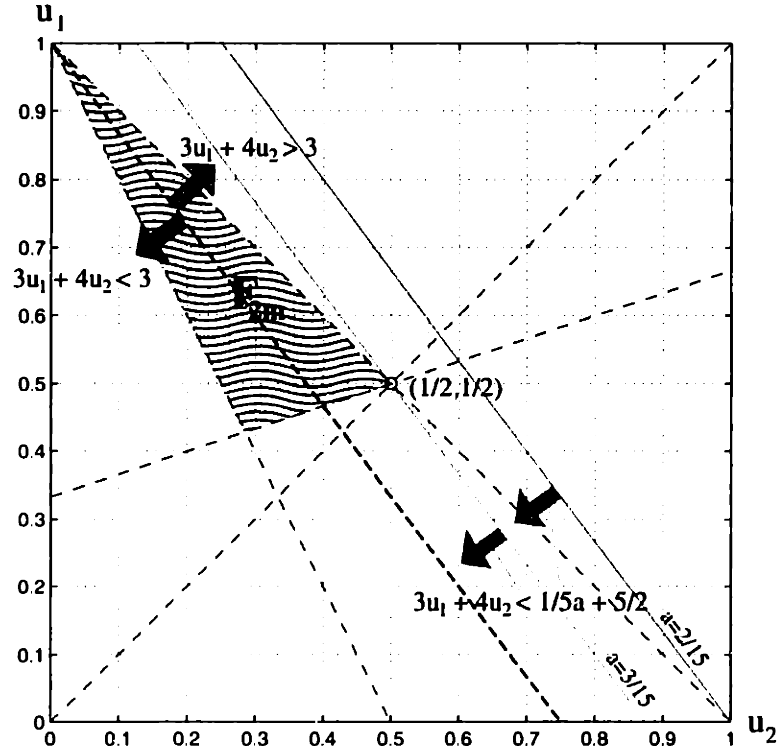


Figure 4-10: Feasible values of (u_1, u_2) and the boundary, $3u_1 + 4u_2 = 3$, between two subsets of weights in the analysis of $y_m(4n+2) - y_m(-3)$.

[Proof]:By Equations 3.2 and 4.7:

$$\begin{aligned} y_m(4n+2) &= \frac{1}{r^2}\|\mathbf{C}_m(0) - \mathbf{P}(4n+2)\|^2 \\ &= 2\cdot(1-u_1)^2 + [4n+2-(3u_1+4u_2-3)]^2\cdot a \qquad (4.37) \\ &= 2\cdot(1-u_1)^2 + 16a\left(n - \frac{3u_1+4u_2-5}{4}\right)^2. \qquad (4.38) \end{aligned}$$

Consider the function $\frac{1}{4}(3u_1+4u_2-5)$. For all $(u_1,u_2) \in \mathbf{F}_{2m}$, as shown in Figure 4-9, Since $-\frac{1}{7} < \frac{1}{4}(3u_1+4u_2-3) < \frac{1}{8}$, this means that $-\frac{9}{14} < \frac{1}{4}(3u_1+4u_2-5) < -\frac{3}{8}$, and $\bar{n}^*$ lies between $-\frac{9}{14}$ and $-\frac{3}{8}$. When $\frac{1}{4}(3u_1+4u_2-5) > -\frac{1}{2}$, $n_{(1)} = 0$; and, when $\frac{1}{4}(3u_1+4u_2-5) < -\frac{1}{2}$, $n_{(1)} = -1$.

When $3u_1 + 4u_2 > 3$, $n_{(1)} = 0$:

$$\begin{aligned} y_m(2) - y_m(-3) &= 2\cdot(1+u_1^2) - 2\cdot u_1^2 + (3u_1+4u_2-5)^2\cdot a - (3u_1+4u_2)^2\cdot a \\ &= (2+25a) + 2(-5)(3u_1+4u_2)\cdot a > 0 \end{aligned}$$

$$\iff 3u_1 + 4u_2 < \frac{2+25a}{10a} = \frac{1}{5a} + \frac{5}{2}.$$

As shown in Figure 4-10, when $a < \frac{3}{15}$, this is not a binding constraint.

When $3u_1 + 4u_2 < 3$, $n_{(1)} = -1$:

$$\begin{aligned}
y_{\mathrm{m}}(-2) - y_{\mathrm{m}}(-3) &= 2 \cdot (1 + u_1^2) - \cdot 2u_1^2 + (3u_1 + 4u_2 - 1)^2 \cdot a - (3u_1 + 4u_2)^2 \cdot a \\
&= (2 + a) + 2(-1)(3u_1 + 4u_2) \cdot a \\
&= (2 + a) + 6a, \qquad \text{since } 3u_1 + 4u_2 < 3 \\
&= 2 - 5a > 0, \qquad \text{because } a < \frac{3}{15} < \frac{2}{5}.
\end{aligned}$$

□

Lemma 9. *Consider pitches of the type* $\mathbf{P}(4n+3)$*, for example, D♭, F, A, C♯, E♯ and so on. For all* $(u_1, u_2) \in \mathbf{F}_{2m}$*,* $\mathbf{P}(-1)$ *(pitch F) is closest to* $\mathbf{C}_{\mathrm{m}}(0)$*, the C minor triad* $\mathbf{P}(-1)$ *(pitch F) is closest to* $\mathbf{C}_{\mathrm{M}}(0)$*. This closest pitch is still farther than* $\mathbf{P}(4)$ *(E, the third in the triad) if and only if*

$$\left(a_5 + \frac{1}{2}\right) \cdot u_1 + u_2 < a_5 \qquad \textit{where} \qquad a_5 = \frac{1}{8a} + \frac{1}{4}.$$

[Proof]:By Equations 3.2 and 4.7:

$$\begin{aligned}
y_{\mathrm{m}}(4n+3) &= \frac{1}{r^2}\|\mathbf{C}_{\mathrm{m}}(0) - \mathbf{P}(4n+3)\|^2 \\
&= 2 \cdot (1 - u_1 + u_1^2) + [4n + 3 - (3u_1 + 4u_2 - 3)]^2 \cdot a && (4.39) \\
&= 2 \cdot (1 - u_1)^2 + 16a\left(n - \frac{3u_1 + 4u_2 - 6}{4}\right)^2. && (4.40)
\end{aligned}$$

Consider the function $\frac{1}{4}(3u_1 + 4u_2 - 6)$. For all $(u_1, u_2) \in \mathbf{F}_{2m}$, as shown in Figure 4-9, Since $-\frac{1}{7} < \frac{1}{4}(3u_1 + 4u_2 - 3) < \frac{1}{8}$, this means that $-\frac{25}{28} < \frac{1}{4}(3u_1 + 4u_2 - 6) < -\frac{5}{8}$, and $\tilde{n}^*$ lies between $-\frac{25}{28}$ and $-\frac{5}{8}$. It follows that $n_{(1)} = -1$.

$$\begin{aligned}
y_{\mathrm{m}}(-1) - y_{\mathrm{m}}(-3) &= 2 \cdot (1 - u_1 + u_1^2) - 2 \cdot u_1^2 + (3u_1 + 4u_2 - 2)^2 \cdot a - (3u_1 + 4u_2)^2 \cdot a \\
&= 2 \cdot [(1 - u_1) + 2a + a(-2)(3u_1 + 4u_2)] \\
&= 2 \cdot [(1 + 2a) - (1 + 6a) \cdot u_1 + 8a \cdot u_2] > 0
\end{aligned}$$

$$\iff \frac{1+2a}{8a} - \left(\frac{1+6a}{8a}\right) \cdot u_1 \quad > \quad u_2$$

$$\iff \left(a_5 + \frac{1}{2}\right) \cdot u_1 + u_2 < a_5, \qquad \text{where } a_5 = \frac{1+2a}{8a} = \frac{1}{8a} + \frac{1}{4}.$$

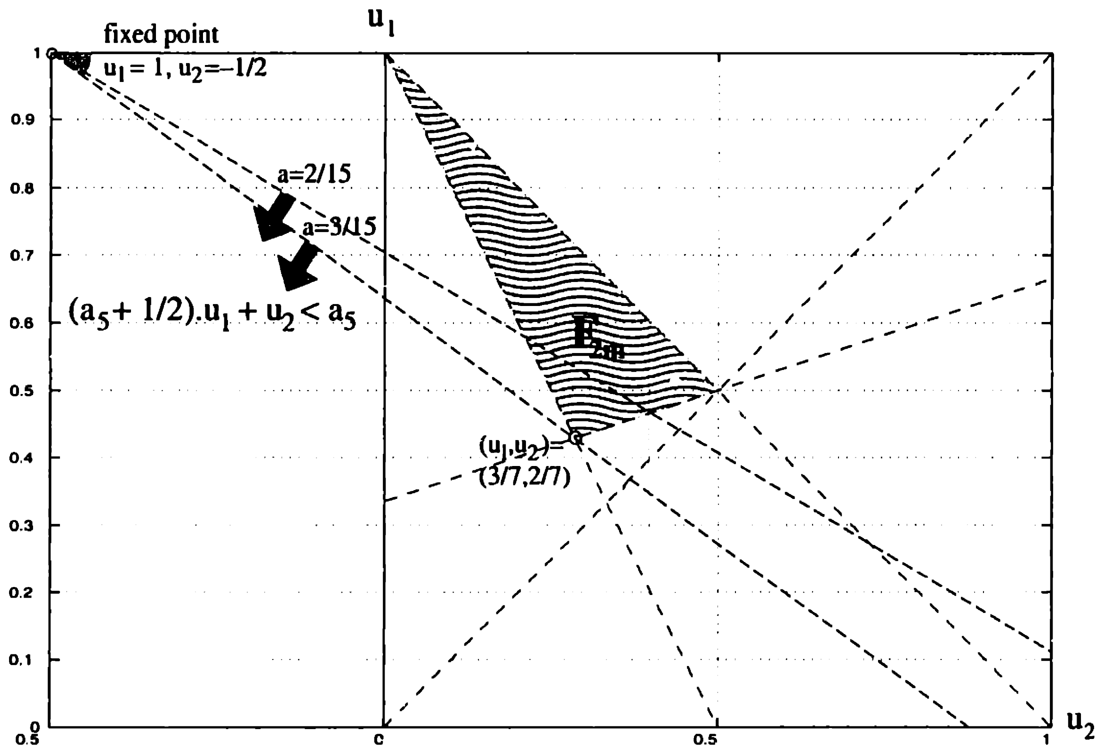


Figure 4-11: Feasible values of (u_1, u_2) based on the analysis of $y_m(4n + 3) - y_m(-3)$.

The equation $(a_5 + \frac{1}{2})u_1 + u_2 = a_5$ has a fixed point at $(u_1, u_2) = (1, -\frac{1}{2})$. Figure 4-11 shows the range of the inequality $(a_5 + \frac{1}{2})u_1 + u_2 < a_5$ over $\frac{2}{15} < a < \frac{3}{15}$.

Theorem 9. *No other pitches are as close to a minor chord center as its three constituent pitches if:*

$$\frac{2}{15} < a < \frac{3}{15}, \qquad 3u_1 > u_2 + 1, \qquad \textit{and} \qquad \left(a_5 + \frac{1}{2}\right) \cdot u_1 + u_2 < a_5,$$

where $a_5 = \frac{1}{8a} + \frac{1}{4}$.

[Proof]: Note that the chord pitch farthest from the chord representation is the third of the chord (by Theorem 8), so that for all $s \neq 0, 1, -3$:

$$\begin{aligned} &||\mathbf{C}_m(k) - \mathbf{P}(k + s)|| \\ &> \max\{||\mathbf{C}_m(k) - \mathbf{P}(k)||, ||\mathbf{C}_m(k) - \mathbf{P}(k + 1)||, ||\mathbf{C}_m(k) - \mathbf{P}(k - 3)||\} \end{aligned}$$

Thus, one needs only prove that all non-chord pitches are farther from the chord than its third. By symmetry, one need only show this for C minor, i.e. $\mathbf{C}_m(0)$.

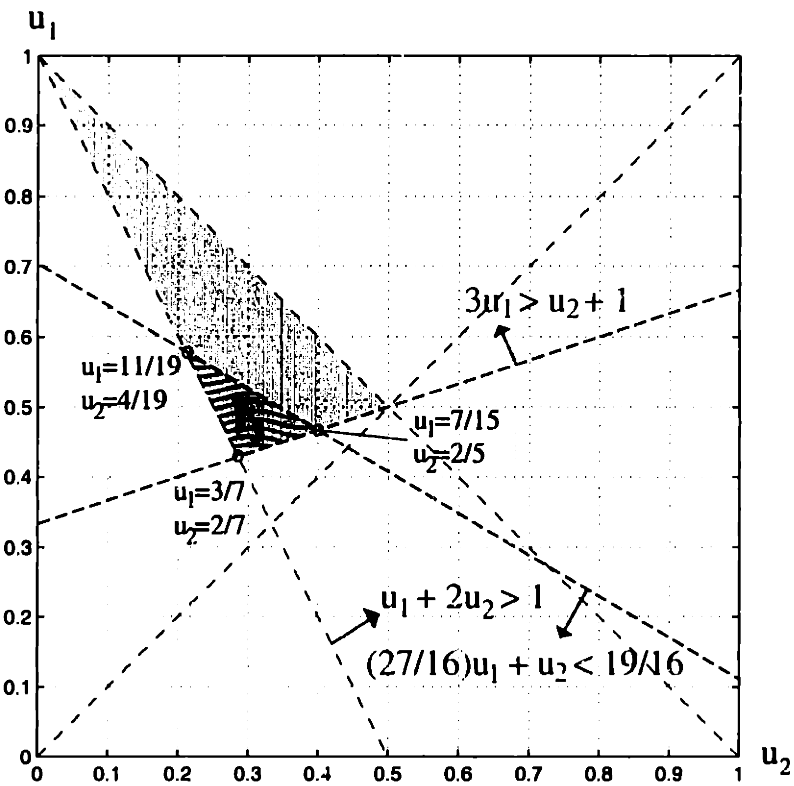


Figure 4-12: Feasible values of (u_1, u_2), when $a = \frac{2}{15}$, based on the proximity relations between a minor chord and all pitches.

By Lemma 6, $a < \frac{2}{7}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n)$ are farther than $\mathbf{P}(-3)$ from $\mathbf{C}_m(0)$, for $n \neq 0$. By Lemma 7, $\mathbf{P}(4n+1)$ are all farther than $\mathbf{P}(-3)$ from $\mathbf{C}_m(0)$, for $n \neq 0, -1$. By Lemma 8, $a < \frac{3}{15}$ is necessary and sufficient to guarantee that all $\mathbf{P}(4n+2)$ are farther than $\mathbf{P}(-3)$ from $\mathbf{C}_m(0)$. Given that $a > \frac{2}{15}$, by Lemma 9, $(a_5 + \frac{1}{2})u_1 + u_2 < a_5$ where $a_5 = \frac{1}{8a} + \frac{1}{4}$, is necessary and sufficient to guarantee that all $\mathbf{P}(4n+3)$ are farther than $\mathbf{P}(-3)$ from $\mathbf{C}_m(0)$. □

4.6 Summary of Constraints on the Aspect Ratio and Chord Weights

With all the math that has transpired, it would be useful to summarize the constraints on the aspect ratio and the chord weights generated by the basic condition: a chord's root should be the pitch closest to the chord representation, followed by its fifth, then the third.

For ease of transcription, instead of referring to the distance between a chord and a pitch, I will use the standardized distances,

$$\begin{aligned} y_{\mathrm{M}}(s) &= \frac{1}{r^2}\|\mathbf{C}_{\mathrm{M}}(0) - \mathbf{P}(s)\|^2, \\ \text{and} \qquad y_{\mathrm{m}}(s) &= \frac{1}{r^2}\|\mathbf{C}_{\mathrm{m}}(0) - \mathbf{P}(s)\|^2. \end{aligned}$$

Summary of Results

The results from the previous sections on major chords can be summarized as: for all $(w_1, w_2) \in \mathbf{F}_0$,

$$\begin{aligned} y_{\mathrm{M}}(0) < y_{\mathrm{M}}(1) < y_{\mathrm{M}}(4) &\iff \frac{2}{15} < a < \frac{2}{7}, \\ y_{\mathrm{M}}(4) < y_{\mathrm{M}}(4n), n \neq 0, 1 &\iff a < \frac{2}{7}, \\ y_{\mathrm{M}}(4) < y_{\mathrm{M}}(4n+1), n \neq 0 &\iff a < \frac{2}{7}, \\ y_{\mathrm{M}}(4) < y_{\mathrm{M}}(4n+2), \forall n &\iff a < \frac{3}{15}, \\ y_{\mathrm{M}}(4) < y_{\mathrm{M}}(4n+3), \forall n &\iff 3 < 4w_1 + 3w_2 < \frac{1}{5a} + \frac{5}{2}, \\ &\qquad \text{and } 4w_1 + 3w_2 < \frac{1}{2a} + \frac{1}{2}. \end{aligned}$$

The results from the previous sections on minor chords can be summarized as: given that $\frac{2}{15} < a < \frac{3}{15}$ and $(u_1, u_2) \in \mathbf{F}_2$,

$$\begin{aligned} y_{\mathrm{m}}(0) < y_{\mathrm{m}}(1) < y_{\mathrm{m}}(-3) &\iff u_1 > \frac{1}{3} \cdot u_2 + \frac{1}{3}, \\ y_{\mathrm{m}}(-3) < y_{\mathrm{m}}(4n), n \neq 0, &\iff a < \frac{2}{7}, \\ y_{\mathrm{m}}(-3) < y_{\mathrm{m}}(4n+1), n \neq -1 &\qquad \text{always true} \\ y_{\mathrm{m}}(-3) < y_{\mathrm{m}}(4n+2), \forall n &\iff a < \frac{3}{15}, \\ y_{\mathrm{m}}(-3) < y_{\mathrm{m}}(4n+3), \forall n &\iff \left(a_5 + \frac{1}{2}\right) \cdot u_1 + u_2 < a_5, \\ &\qquad \text{where } a_5 = \frac{1}{8a} + \frac{1}{4}. \end{aligned}$$

Results for Aspect Ratio $a = \frac{2}{15}$

When $\frac{a=2}{15}$, the distance between pitches a perfect fifth apart and the distance between pitches a major third apart are the same. This is a boundary derived in Theorems 5 and 6.

This value, $\frac{2}{15}$, will be the aspect ratio choice for the examples, illustrations and applications in the remainder of the thesis.

When $a = \frac{2}{15}$, the results from the previous sections on major chords can be summarized as: for all $(w_1, w_2) \in \mathbf{F}_0$,

$$\begin{array}{rrl} & y_M(0) < y_M(1) \leq y_M(4) & \text{always ;} \\ & y_M(4) < y_M(4n), n \neq 0, 1 & \text{always ;} \\ & y_M(4) \leq y_M(4n+1), n \neq 0 & \text{always ;} \\ & y_M(4) < y_M(4n+2), \forall n & \text{always ;} \\ \text{and,} & y_M(4) \leq y_M(4n+3), \forall n & \text{always .} \end{array}$$

The results from the previous sections on minor chords can be summarized as: given that $\frac{2}{15} < a < \frac{3}{15}$ and $(u_1, u_2) \in \mathbf{F}_2$,

$$\begin{array}{rrcl} & y_m(0) < y_m(1) < y_m(-3) & \Longleftrightarrow & u_1 > \frac{1}{3} \cdot u_2 + \frac{1}{3}; \\ & y_m(-3) < y_m(4n), n \neq 0, & & \text{always ;} \\ & y_m(-3) < y_m(4n+1), n \neq -1 & & \text{always ;} \\ & y_m(-3) < y_m(4n+2), \forall n & & \text{always ;} \\ \text{and,} & y_m(-3) < y_m(4n+3), \forall n & \Longleftrightarrow & \frac{27}{16} \cdot u_1 + u_2 < \frac{19}{16}. \end{array}$$

The intersection of $\mathbf{F}_2$ and the first constraint forms $\mathbf{F}_{2m}$ (shown in Figure 4-9). The intersection of $\mathbf{F}_{2m}$ and the second constraint forms $\mathbf{F}_3$ (shown in Figure 4-12).

4.7 Exploring Other Relationships Among Chords and Pitches

In this section, I will explore the implications of the choices of weights on other relationships among major and minor chords and pitches. For example: what is the closest chord to any given pitch? what is the closest pitch to any given chord? what is the closest chord to any given chord?

In Chapter 3, as each chord and key representation were defined, I proved that, because of the definitions, the relationships among major chords, among minor chords, among major keys, and among minor keys, are identical to those among pitches. However, the relationship of across modes has yet to be explored. This section will attempt to illustrate some of these major-minor mode relations for two sets of major triad and minor triad weights.

As shown in the previous section, for any choice of weights for a given chord (w_1, w_2) for major chords and (u_1, u_2) for minor chords, if $a = \frac{2}{15}$, and

$$u_1 - \frac{1}{3} \cdot u_2 > \frac{1}{3} \quad \text{and} \quad \frac{27}{16} \cdot u_1 + u_2 < \frac{19}{16},$$

the chord is always closest to its root, followed by the fifth then the third. In the case of major triads, $\mathbf{C}_M(k)$, the root is $\mathbf{P}(k)$, the fifth $\mathbf{P}(k+1)$, and the third $\mathbf{P}(k+4)$. For minor triads, $\mathbf{C}_m(k)$, the root is $\mathbf{P}(k)$, the fifth $\mathbf{P}(k+1)$, and the third $\mathbf{P}(k-3)$.

This simple set of restrictions allows for the modeling of multiple relationships among chords and pitches, and also between pairs of chord representations.

Parameters Chosen for the Illustrations

As mentioned before, the aspect ratio in all remaining examples are set at $a = \frac{2}{15}$. There are two sets of weights which are used in the illustrations in this section. The major chord weights (w_1, w_2) are restricted to be within $\mathbf{F}_0$ (shown in Figure 4-3); and, the minor triad weights (u_1, u_2) are restricted to be within $\mathbf{F}_3$ (shown in Figure 4-12).

The first set of weights are:

$$w = \left[\frac{1}{3} + \varepsilon, \frac{1}{3}, \frac{1}{3} - \varepsilon\right], \qquad \text{and} \qquad u = \left[\frac{3}{7} + \varepsilon, \frac{2}{7}, \frac{2}{7} - \varepsilon\right].$$

These weights are chosen for their simplicity and because they are close to the vertices of the feasible spaces shown in Figures 4-3 and 4-12 for illustrative purposes. This naive choice of major triad weights reflects close to no preference among the relationships between the different chord pitches and major triad except for the model's own inbuilt preferences. The weights assigned the minor chord pitches shows little preference among the relationships between the minor chord and its fifth or third.

The second set of weights are:

$$w = \left[\frac{1}{2} + \varepsilon, \frac{1}{4}, \frac{1}{4} - \varepsilon\right], \qquad \text{and} \qquad u = \left[\frac{1}{2} + \varepsilon, \frac{1}{4}, \frac{1}{4} - \varepsilon\right].$$

In this case, the two sets of weights are chosen to be equal. The interpretation for these values is that emphasis is placed on the root of the chord, but close to equal preference is assigned to the fifth and the third. In addition, the relationships between minor chord pitches and the minor chord are assumed to be the same way as those between major chord pitches and the major chord.

Implications for Chords wrt given Pitch

Because proximity in the Spiral Array indicates shared pitches and interval relations of fifths or thirds, the chords closest to a given pitch always has this pitch its root, fifth or third. Changes in the weights only shuffle reorders the chords according to the distance gradations.

Tables 4.2 and 4.3 show examples of pitch-chord relations with respect to two different sets of the major and minor chord weights that satisfy the conditions given in Section 4.2. In Table 4.2, $w = [\frac{1}{3} + \varepsilon, \frac{1}{3}, \frac{1}{3} - \varepsilon]$, and $u = [\frac{3}{7} + \varepsilon, \frac{2}{7}, \frac{2}{7} - \varepsilon]$.

h/r	Pitch	Chord positions ranked by closeness					
$\sqrt{\frac{2}{15}}$	**P**(k)	$C_M(k)$	$C_m(k)$	$C_m(k-1)$	$C_M(k-1)$	$C_M(k-4)$	$C_m(k+3)$
	C	C (0.5926)[a]	c (0.6966)	f (0.6966)	F (0.9481)	A♭ (0.9482)	a (1.1538)
$\sqrt{\frac{3}{15}}$	**P**(k)	$C_m(k)$	$C_M(k)$	$C_m(k-1)$	$C_M(k-1)$	$C_M(k-4)$	$C_m(k+3)$
	C	c (0.7183)	C (0.7778)	f (0.8612)	F (0.9778)	A♭ (1.3111)	a (1.5470)

Table 4.2: Example of Pitch-Chord relation when the major chord weights $w = [\frac{1}{3}+\varepsilon,\ \frac{1}{3}-\varepsilon,\ \frac{1}{3}-\varepsilon]$, and the minor chord wegiths $u = [\frac{3}{7}+\varepsilon,\ \frac{2}{7}-\varepsilon,\ \frac{2}{7}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

h/r	Pitch	Chord positions ranked by closeness					
	P(k)	$C_M(k)$	$C_m(k)$	$C_m(k-1)$	$C_M(k-1)$	$C_M(k-4)$	$C_m(k+3)$
$\sqrt{\frac{2}{15}}$	C	C (0.3333)	c (0.5333)	f (0.8000)	F (1.1334)	A♭ (1.1334)	a (1.1334)
$\sqrt{\frac{3}{15}}$	C	C (0.4374)	c (0.5499)	f (0.9500)	F (1.1375)	A♭ (1.6376)	a (1.7501)

Table 4.3: Example of Pitch-Chord relation when the major and minor chord weights are $[\frac{1}{2}+\varepsilon,\ \frac{1}{4},\ \frac{1}{4}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

[a]The number in the bracket denotes spatial distance between each chord representation and the reference pitch. The radius r = 1 and h = the aspect ratio.

Table 4.3 shows the results when the weights are restricted to be both equal to $[\frac{1}{2}+\varepsilon,\ \frac{1}{4},\ \frac{1}{4}-\varepsilon]$.

Implications for Pitches wrt Given Chord

By design, the three closest pitch positions to a given major chord representation are its root, fifth and third. And, because pitches in a key occupy a compact space in the Spiral Array, the closest pitches to chord, $C_M(k)$, will always include the pitches in the key with tonic **P**(k). For example, the pitches closest to the C major chord center will always include those from the key of C major. And, the pitches closest to the C minor chord cenrter will always include those from the key of C minor.

Table 4.5 shows some sample results when the weights are restricted to be $[\frac{1}{3}+\varepsilon,\ \frac{1}{3},\ \frac{1}{3}-\varepsilon]$ for major triads and $[\frac{3}{7}+\varepsilon,\ \frac{2}{7},\ \frac{2}{7}-\varepsilon]$ for minor triads. Table 4.5 shows similar results when the weights are restricted to be both equal to $[\frac{1}{2}+\varepsilon,\ \frac{1}{4},\ \frac{1}{4}-\varepsilon]$.

h/r	Chord	Pitch positions ranked by closeness					
$\sqrt{\frac{2}{15}}$	$C_M(k)$	P(k)	P(k+1)	P(k+4)	P(k+5)	P(k+3)	P(k+2)
	C	C (0.5926)	G (0.9481)	E (0.9482)	B (2.3704)	A (2.4593)	D (2.9037)
$\sqrt{\frac{3}{15}}$	$C_M(k)$	P(k)	P(k+1)	P(k+4)	P(k+3)	P(k+2)	P(k+5)
	C	C (0.7778)	G (0.9778)	E (1.3111)	A (2.5778)	D (2.9111)	B (3.1112)
$\sqrt{\frac{2}{15}}$	$C_m(k)$	P(k)	P(k+1)	P(k−3)	P(k−4)	P(k−2)	P(k−1)
	C	C (0.6966)	G (0.6966)	E♭ (1.1538)	A♭ (2.2204)	B♭ (2.6395)	F (2.6775)
$\sqrt{\frac{3}{15}}$	$C_m(k)$	P(k)	P(k+1)	P(k−3)	P(k−1)	P(k−2)	P(k−4)
	C	C (0.7183)	G (0.8612)	E♭ (1.5470)	F (2.6898)	B♭ (2.7755)	A♭ (3.0041)

Table 4.4: Example of Chord-Pitch relation when the major and minor chord weights are $[\frac{1}{3}+\varepsilon,\ \frac{1}{3},\ \frac{1}{3}-\varepsilon]$ for major triads and $[\frac{3}{7}+\varepsilon,\ \frac{2}{7},\ \frac{2}{7}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

h/r	Chord	Pitch positions ranked by closeness					
$\sqrt{\frac{2}{15}}$	$C_M(k)$	P(k)	P(k+1)	P(k+4)	P(k+3)	P(k−1)	P(k+5)
	C	C (0.3333)	G (1.1333)	E (1.1334)	A (2.5334)	F (2.8000)	B (3.0000)
$\sqrt{\frac{3}{15}}$	$C_M(k)$	P(k)	P(k+1)	P(k+4)	P(k+3)	P(k−1)	P(k+2)
	C	C (0.4375)	G (1.1375)	E (1.6376)	A (2.7375)	F (3.1375)	D (3.2375)
$\sqrt{\frac{2}{15}}$	$C_m(k)$	P(k)	P(k+1)	P(k−3)	P(k−4)	P(k−1)	P(k−2)
	c	C (0.5333)	G (0.8000)	E♭ (1.3334)	A♭ (2.1333)	F (2.5333)	B♭ (2.5333)
$\sqrt{\frac{3}{15}}$	$C_m(k)$	P(k)	P(k+1)	P(k−3)	P(k−1)	P(k−4)	P(k−2)
	c	C (0.5500)	G (0.9500)	E♭ (1.7501)	F (2.5500)	A♭ (2.9500)	B♭ (2.9500)

Table 4.5: Example of Chord-Pitch relation when the major and minor chord weights are $[\frac{1}{2}+\varepsilon,\ \frac{1}{4},\ \frac{1}{4}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

Implications for Chords wrt given Chord

In a similar fashion, the choice of weights on the major and minor chords allow for different distance affiliations between chord centers. As examples, Table 4.6 shows chord-chord relations when the major chord weights are $[\frac{1}{3}+\varepsilon,\ \frac{1}{3},\ \frac{1}{3}-\varepsilon]$ and minor chord weights are $[\frac{3}{7}+\varepsilon,\ \frac{2}{7},\ \frac{2}{7}-\varepsilon]$; and, Table 4.7 shows the same relations when the weights are restricted to be both equal to $[\frac{1}{2}+\varepsilon,\ \frac{1}{4},\ \frac{1}{4}-\varepsilon]$.

h/r	Chord	Chord positions ranked by closeness					
$\sqrt{\frac{2}{15}}$	$C_M(k)$	$C_M(k)$	$C_m(k+4)$	$C_m(k+3)$	$C_m(k)$	$C_M(k+1)$	$C_M(k-1)$
	C	C (0.0000)	e (0.5273)	a (0.6670)	c (0.7812)	G (1.2444)	F (1.2444)
$\sqrt{\frac{3}{15}}$	$C_M(k)$	$C_M(k)$	$C_m(k+3)$	$C_m(k+4)$	$C_m(k)$	$C_M(k+1)$	$C_M(k-1)$
	C	C (0.0000)	a (0.7057)	e (0.7343)	c (1.1151)	G (1.3111)	F (1.3111)
$\sqrt{\frac{2}{15}}$	$C_m(k)$	$C_m(k)$	$C_M(k-4)$	$C_M(k-3)$	$C_M(k)$	$C_m(k-1)$	$C_m(k+1)$
	c	c (0.0000)	A♭ (0.5273)	E♭ (0.6670)	C (0.7812)	f (1.1537)	g (1.1537)
$\sqrt{\frac{3}{15}}$	$C_m(k)$	$C_m(k)$	$C_M(k-3)$	$C_M(k-4)$	$C_M(k)$	$C_m(k-1)$	$C_m(k+1)$
	c	c (0.0000)	E♭ (0.7057)	A♭ (0.7343)	C (1.1151)	f (1.2204)	g (1.2204)

Table 4.6: Example of Chord-Chord relation when the major chord weights are $[\frac{1}{3}+\varepsilon, \frac{1}{3}, \frac{1}{3}-\varepsilon]$, and minor chord weights are $[\frac{3}{7}+\varepsilon, \frac{2}{7}, \frac{2}{7}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

h/r	Chord	Chord positions ranked by closeness					
$\sqrt{\frac{2}{15}}$	$C_M(k)$	$C_M(k)$	$C_m(k-1)$	$C_m(k+4)$	$C_m(k+3)$	$C_M(k+1)$	$C_M(k-1)$
	C	C (0.0000)	c (0.5333)	e (0.8000)	a (0.8334)	G (1.3833)	F (1.3833)
$\sqrt{\frac{3}{15}}$	$C_M(k)$	$C_M(k)$	$C_m(k-1)$	$C_m(k+3)$	$C_m(k+4)$	$C_M(k+1)$	$C_m(k-1)$
	C	C (0.0000)	c (0.7374)	a (0.9376)	e (1.1376)	G (1.4500)	F (1.4500)
$\sqrt{\frac{2}{15}}$	$C_m(k)$	$C_m(k)$	$C_M(k)$	$C_M(k-4)$	$C_M(k-3)$	$C_m(k-1)$	$C_m(k+1)$
	c	c (0.0000)	C (0.5333)	A♭ (0.8000)	E♭ (0.8334)	f (1.1333)	g (1.1333)
$\sqrt{\frac{3}{15}}$	$C_m(k)$	$C_m(k)$	$C_M(k)$	$C_M(k-3)$	$C_M(k-4)$	$C_m(k-1)$	$C_m(k+1)$
	c	c (0.0000)	C (0.7374)	E♭ (0.9376)	A♭ (1.1376)	f (1.2000)	g (1.2000)

Table 4.7: Example of Chord-Chord relation when the major and minor chord weights are $[\frac{1}{2}+\varepsilon, \frac{1}{4}, \frac{1}{4}-\varepsilon]$. The numbers are generated for the case when $\varepsilon = 10^{-5}$.

4.8 Desired Key-Interval-Pitch Relations

In this section, I examine three simple and desirable proximity relations: the nearest key to a given pitch, the nearest key to a half-step interval, and the nearest key to a perfect fourth interval.

The Tonic-Key Relationship

The key representation closest to an isolated pitch should be the major key of the same name, which is to say, the major key with the pitch as its first degree. The interpretation is that: on hearing a single pitch, without further information, the listener perceives the single pitch as the first degree of the major key. For example, the key nearest the pitch C should be C major.

Some other keys close by are: C minor, F major, F minor, A minor and A♭ major. A subtle difference in the weights can lead to any of these other keys being closer to the pitch C than C major.

Mathematically, this condition translates to:

$$\arg\min_{\mathbf{T}} d(\mathbf{T}, \mathbf{P}(\mathrm{k})) = \mathbf{T}_{\mathrm{M}}(\mathrm{k}), \text{ where } \mathbf{T} \in \{T_{\mathrm{M}}(\ell), T_{\mathrm{m}}(\ell) \;\; \forall \text{ integer } \ell\}. \tag{4.41}$$

Implications of the Half Step Interval

The half-step (also called a semitone) is an important element in defining a key. In an ascending diatonic major scale, there are only two half steps: between the mediant and subdominant ($\hat{3} - \hat{4}$), and the leading tone and tonic ($\hat{7} - \hat{1}$). All other intervals between consecutive steps are whole steps. As a result, an informed listener uses half steps as clues to the identity of a key.

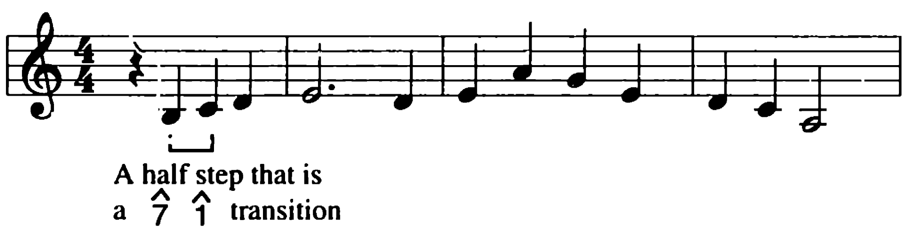


Figure 4-13: Example: "Londonerry Air" begins with a half step interval that forms a ($\hat{7} - \hat{1}$) transition.

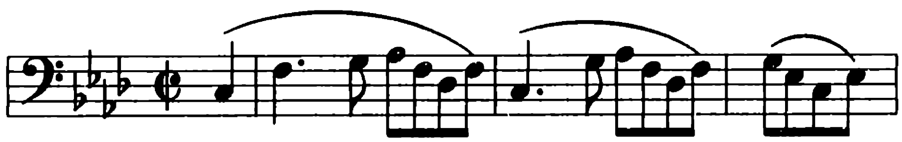
In general, when taken out of context, the sounding of two pitches that forming the interval of a half step most strongly suggests a ($\hat{7} - \hat{1}$) transition, as in the beginning of "Londonerry Air" shown in Figure 4-13. Alternatively, this interval could suggest a ($\hat{3} - \hat{4}$) transition. Given no other information, I choose the former as the likelier interpretation. For example, BC would suggest C major, followed by G major.

In the model, the desired condition is:

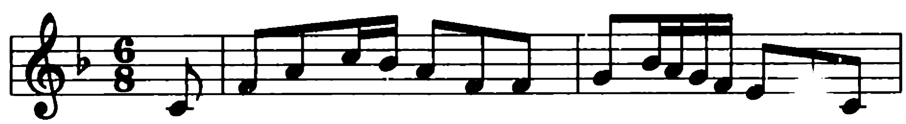
$$\arg\min_{\mathbf{T}} d(\mathbf{T}, \mathrm{ave}(\mathbf{P}(\mathrm{k}+5), \mathbf{P}(\mathrm{k}))) = \mathbf{T}_{\mathrm{M}}(\mathrm{k}), \text{ where } \mathbf{T} \in \{T_{\mathrm{M}}(\ell), T_{\mathrm{m}}(\ell) \;\; \forall \text{ integer } \ell \neq \mathrm{k}+5\}. \tag{4.42}$$

Implications of the Perfect Fourth Interval

The other important element in assessing the key of a passage is the rising perfect fourth interval. This is usually perceived to be the transition from a dominant to the tonic ($\hat{5} - \hat{1}$) for both major and minor scales. An melody beginning with a rising fourth is used as the example in Chapter 5: "Simple Gifts", shown in Figure 5-1. Two other examples (one a minor key, and one in a major key) are given in Figure 4-14.



Brahms: Piano Quintet, Op. 34 (opening of first movement)



"The Ash Grove"

Figure 4-14: Examples: Two melodies that begin with a rising perfect fourth interval that form ($\hat{5} - \hat{1}$) transitions. The Brahms Piano Quintet is in F minor, and "The Ash Grove" in F major.

Less frequently, the rising fourth is a transition from the tonic to the subdominant ($\hat{1} - \hat{4}$). The perfect fourth also exists as a relationship between many other scale degrees. For example, the intervals between the supertonic and dominant ($\hat{2} - \hat{5}$), and the mediant and submediant ($\hat{3} - \hat{6}$) are also perfect fourths. In the melodic minor scale, the relationship between the subdominant and leading tone ($\hat{4} - \hat{7}$) is also a perfect fourth.

Hence, the preference between interpretations is not as strong as that for the half-step. Nevertheless, it is still a strong indicator for key. Giving preference to the major mode, when given two pitches a perfect fourth apart, I select the major key of the upper pitch as the preferred key. The corresponding mathematical relationship is:

$$\arg\min_{\mathbf{T}} d(\mathbf{T}, \mathrm{ave}(\mathbf{P}(\mathrm{k}-1), \mathbf{P}(\mathrm{k}))) = \mathbf{T}_{\mathrm{M}}(\mathrm{k}), \text{ where } \mathbf{T} \in \{T_{\mathrm{M}}(\ell), T_{\mathrm{m}}(\ell) \;\; \forall \text{ integer } \ell \neq \mathrm{k}-1\}. \tag{4.43}$$

4.9 Finding Solutions that Satisfy the Key-Interval-Pitch Relations

In this section, I discuss a selection of precise weights for all key and chord representations using computational methods, bearing in mind the constraints derived earlier in this chapter as summarized in Section 4.6.

The problem of satisfying the conditions stated in Section 4.8, translates to a mathematical one in the Spiral Array. The new problem is one of finding major chord weights w, minor chord weights, u, major key weights, ω, and minor chord weights v that satisfy the conditions stated in Equations 4.41, 4.42, and 4.43.

Even though the definitions for chords and keys (stated in Equations 3.4, 3.5, 3.10 and 3.12, and summarized in Section 3.5) contain simple linear relations, their interaction is nonlinear and not convex. Initial attempts at solving for the weights using the Matlab nonlinear solver turned up no solutions. This does not mean that no solutions exist, just that this particular algorithm failed to find one.

Enumerating all possible values for each set of weights to do an exhaustive search would not be computationally viable. In response, I designed a Flip-Flop Heuristic to generate a set of weights that come close to satisfying all of the three conditions.

The Flip-Flop Heuristic

The Flip-Flop Heuristic is essentially an ad-hoc perturbation technique that flip-flops between infeasible solutions that satisfy a subset of the conditions. Since it is difficult to find a solution to a non-convex, nonlinear problem, this heuristic suggests beginning with an infeasible solution, and attempting to improve upon it by trying to satisfy alternating subsets of conditions. At worst, one could still use the original solution and not any bit worse off; at best, one might actually find a feasible solution. More realistically, one would make incremental improvements on the infeasible solution.

The heuristic first begins with a solution that satisfies two out of the three conditions. Rename the conditions given in Equations 4.41, 4.42 and 4.43 A, B and C respectively. For example, two out of the three could be A and B.

Such a weight is found by restricting the search space so that all weights, $(w, u, \omega,$ and $v)$ are the same, and allowed to take on every possible pair of values in $\mathbf{F}_0$ (shown in Figure 4-3). The equality of all weights is not an unreasonable assumption since, perceptually, the relative importance of the elements in each heirarchy could to be similar.

By varying each set of weights in turn, and alternating between constraints subsets, the heuristic seeks a solution that comes closest to satisfying all three conditions given in Equations 4.41 to 4.43.

Attempting to formalize an ad-hoc algorithm, I list the steps of the Flip-Flop Heuristic:

Step 1: Restrict $w = u = \omega = v$.
Step 2: Find weights (w, u, ω, v) that satisfy a subset of the conditions.
Step 3: Vary one set of weights over its feasible set of values.
(Occasionally, allow infeasible weights.)
Step 4: Find a boundary solution that satisfies another subset of the conditions.
Step 5: Choose another set of weights for Step 3.

Before giving an example of how one set of weights were found using this method, I first address the problem of choosing α and β in the definition of a minor key representation.

A "Democratic" Definition of the Minor Key Representation

I propose the use of a "democratic" definition of the minor key in generating an alternative set of weights. Since the harmonic minor scale is such that the pitches of the dominant chord is major, and the melodic scale is such that the pitches of the dominant dominant chord is major on its way up and minor on its way down, I set $\alpha = 0.75$, i.e. the V chord is used 75% of the time, and the v chord 25% of the time. For similar reasons, I set $\beta = 0.75$ in Definition 3.

What the Flip-Flop Heuristic Found

The Flip-Flop Heuristic was implemented and run with two different definitions of minor keys: the harmonic minor definition which sets $\alpha = \beta = 1$; and the "democratic" definition of minor keys which sets $\alpha = \beta = 0.75$.

A numerical search using the harmonic minor definition for minor keys indicates that for r = 1 and h = $\sqrt{2/15}$, $w = u = \omega = v = [0.5353, 0.2743, 0.1904]$ satisfies Conditions A and B. The Flip-Flop Heuristic was, unfortunately, not able to improve upon this set of initial weights, cycling between the last two infeasible values as shown in Table 4.8.

Note that although this solution only satisfies A and B, it agrees with the constraints on w described in Sections 4.2 and 4.3, so that the major chord pitches, more than any other pitches, are closest to the major chord. And the ordering, by closeness, is root, fifth, third. The solution found, also agrees with the constraints on u described in Sections 4.4 and 4.5, so a similar set of properties apply between the minor chord pitches and the minor chord.

In the "democratic" definition of a minor key, the heuristic begins with the weights $w = u = \omega = v =$ [0.536, 0.274, 0.19]. These weights clearly satisfy Conditions A and B. The ensuing sets of weights generated using the Flip-Flop Heuristic are shown in Table 4.9. These calculations also used h= $\sqrt{2/15}$ (r= 1). This solution satisfies only two of the three

Iteration	Conditions	Major Chord			Minor Chord			Major Key	Minor Key
Initial	A,B	0.5353	0.2743	0.1904	like C_M			like C_M	like C_M
1	B,C	as before			0.212	0.423	0.365	as before	as before
2	A,C	0.593	0.213	0.194	as before			like C_M	like C_M
3	B,C	as before			0.788	0.084	0.128	as before	as before
4	B,C	0.593	0.213	0.194	as before			like C_M	like C_M

Table 4.8: Weights generated by the Flip-Flop Heuristic. Minor keys use harmonic definition.

Iteration	Conditions	Major Chord			Minor Chord			Major Key	Minor Key
Initial	A,B	0.536	0.274	0.19	like C_M			like C_M	like C_M
1	A,C	as before			0.419	0.249	0.332	as before	as before
2	B,C	0.517	0.299	0.184	as before			like C_M	like C_M
3	A	as before			0.5247	0.1627	0.3126	as before	as before
4	B,C	0.5161	0.3155	0.1684	as before			like C_M	like C_M
5	A	as before			0.6011	0.2121	0.1868	as before	as before
6	B,C	0.6025	0.2930	0.1145	as before			as before	as before

Table 4.9: Weights generated by the Flip-Flop Heuristic. Minor keys using "democratic" definition.

conditions, and in addition, u is not in $\mathbf{F}_3$. However, it comes much closer to satisfying all three. By some measures, this solution is less infeasible than the one for the harmonic minor example. I will show and explain how this is true in the remainder part of this section.

Observe the weights in the final row of Table 4.9. The minor chord weights have been assigned the values, $u = [\ 0.6011,\ 0.2121,\ 0.1868\ ]$. The other weights, (w, ω, v), are constrained to be the same while being varied over the range $w \in \mathbf{F}_0$. The venn diagram in Figure 4-15 shows the parts of $\mathbf{F}_0$ that satisfy each of the three conditions, A, B and C.

As can be seen in the magnified plot on the right, there is no intersection between the three sets of weights satisfying each of the conditions. However, they do come minutely close. The final value (w_1, w_2)=(0.6025,0.2930) is chosen because it satisfies Conditions B and C, and is closest to satisfying Condition A. Varying any of the weights do not lead to further improvements. The next chapter uses these sets of values in the applications to key-finding.

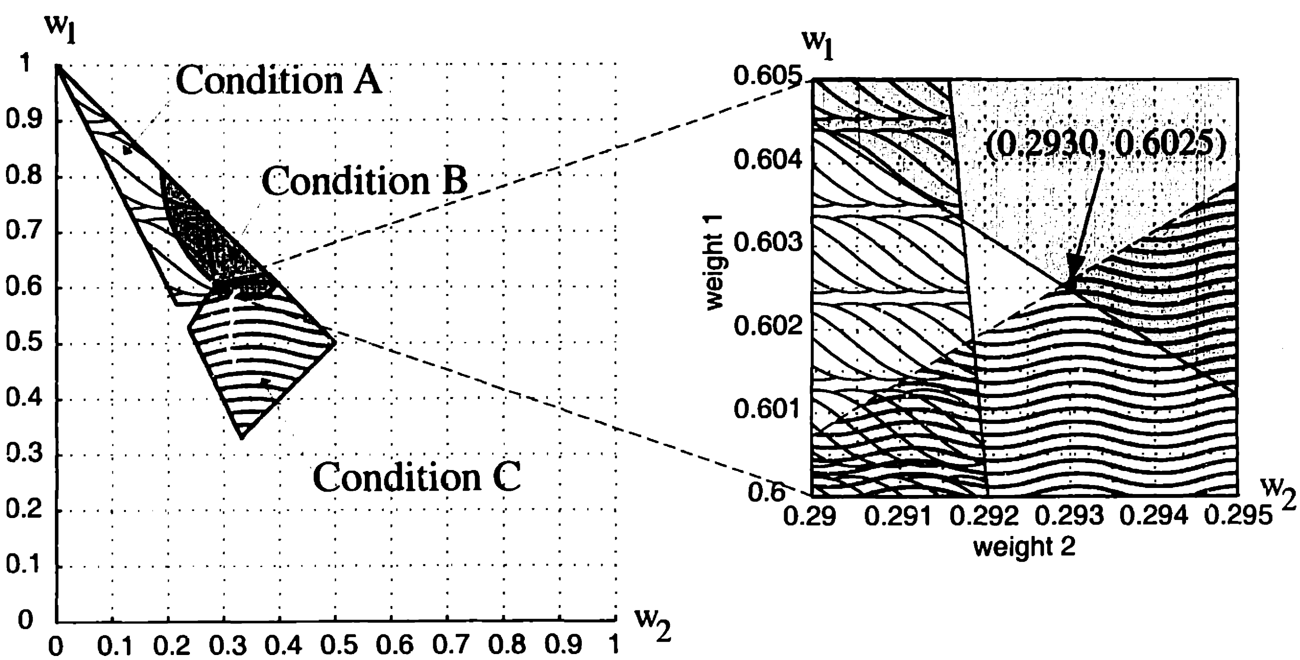


Figure 4-15: Venn diagram of the weights, w, that satisfy different proximity conditions, given that $u = [\ 0.6011, 0.2121, 0.1868\ ]$. The other weights ω and v are restricted to be equal to w.

5 — Finding Keys

Since the entire tonality modeling enterprise was inspired by a question about key, it is only fitting that the first application of the model describes a technique for key-finding. The derivation of key from pitch (and rhythm) information has been widely researched in the fields of music perception, artificial intelligence and cognitive science. The background to some previous research in this area is covered in Section 2.3.

In this chapter I describe my proposed key-finding algorithm which I call the CEG method, and compare its performance to two other algorithms, namely, that of Longuet-Higgins and Krumhansl. The CEG method's average performance surpasses that of previous models, and is close to optimal. I also briefly discuss the issue of model validation.

The application of CEG method to thicker musical textures (which is to say, musical passages with more than one melodic line) will be discussed in Chapter 6 in an application to determining modulations, and in Chapter 7 in an application to determining chords.

5.1 Introduction to the CEG Key-Finding Method

Analyzing the key of a melody poses many challenges. It is a far more difficult problem than attempting to discern the key of a large-scale orchestral piece or even a piece of moderately thick texture such as a two-hand piano piece. In the case of a single melodic line, one must make informed decisions about its key based on little information. Furthermore, there could be more than one equally valid answer, in which case a list for the most likely key candidates for key would be more appropriate than one definite key. This section introduces a key-finding algorithm (CEG) based on the Spiral Array that returns a ranked list of possible keys.

CEG is an acronym for Center of Effect Generator. The letters also coincide with the names of the pitches in the C major triad. CG forms a perfect fifth interval and CE forms a major third interval, important relationships in the Spiral Array model. Because the Spiral Array model uses distance as a measure of perceived closeness, the CEG algorithm uses the Spiral Array to reframe the problem of key recognition as a computationally simple one of finding a distance-minimizing key.

The Center of Effect

The algorithm uses pitch and duration information to generate a center of effect. This c.e. represents, spatially in the model, the tonal space generated by the pitches sounded. In the model, the c.e. is the mathematical sum of the pitches, a composite of their individual positions. In fact, the key representations in the Spiral Array are also c.e.'s, the only difference being that they are generated from a specific combination of pitches.

The c.e. generated by the melody's pitches can then be compared to the framework of key representations in the Spiral Array. The quality of a fit is measured by proximity to the candidate key. All the candidate keys can be ranked by their distance to the c.e., allowing for tonal ambiguities.

Why the Algorithm Works

The collection of pitches in a given key defines a compact space in the Spiral Array. A musical passage in that key employs pitches from this collection. Thus, the pitches from the passage also occupy the same compact space in the Spiral Array. For a melody, as pitches are sounded, one gets a better idea of the pitch collection used by the melody, and by inference, its key. Analogously, in the model, as the music progresses, one gets a better idea of the tonal space occupied by the pitches.

As the number of pitches increases, the geometrical shape defined by the pitch positions, having them as its vertices, becomes more and more complex. Instead of using this complex shape to identify the key, the algorithm collapses the pitch information down to a single point, the center of effect. In this manner, the pitches combine to create an object in space - a point which is the composite sum of the pitch positions.

Since keys are also defined as points in space, it is then simple to compute the distance between the c.e. and the key, and nearby keys, to determine which key is closest to the c.e. Thus the mathematical sum of pitches affords parsimonious descriptions of and comparisons between different pitch collections.

However, the CEG algorithm more than simply compares pitch collections. By definition, the key representations favor triadic pitch configurations, and also tonic-dominant and tonic-subdominant relationships. Thus, not all pitches are weighted equally; and, the key representation is a nonlinear combination of its pitch collection. This is a very specific definition of key representations, that incorporates different levels of heirarchical structure. Thus, by comparing the c.e.'s to these key representations, I am expecting certain pitch relations to prevail.

The algorithm is best explained by an example. The next section applies the CEG algorithm to the melody "Simple Gifts.

5.2 Key-Finding Example Using "Simple Gifts"

An unfolding melody generates a key by the consecutive sounding of pitches in the key, producing interval relations that hint at the tonal space of the melody. As more pitches are heard, the listener forms an idea of the extent of this tonal space and creates a hierarchy based on pitch stability. Correspondingly, in the Spiral Array, as more pitches are sounded and mapped onto their respective spatial representations, one gets a better definition of the analogous space occupied by the key in the model. This information can be summarized by a sum of the pitch positions. I choose to weight each pitch position by its duration since pitch duration is one of the factors that contribute to perceived pitch importance.

The Melody

Figure 5-1: "Simple Gifts".

An example will illustrate how the CEG method spatially models the generation of a key by a melody. The melody used is the Shaker tune, used in Copland's symphonic suite "Appalachian Spring" (1945). The melody is notated in Figure 5-1, for convenience, in the key of F. The model is run with the parameters as defined in Section 4.9. The pitch weights for major chords and the chord weights for the major and minor keys are: [0.6025, 0.2930, 0.1145]; and the pitch weights for minor chords are: [0.6011, 0.2121, 0.1868].

The Center of Effect Generator

At any point in time, the CEG method can generate a c.e. that summarizes the tonal space generated by the pitches sounded. Define a step to be a pitch event. At each step, each of the pitches from the beginning to the present is weighted (multiplied) by its duration, and the center of effect is generated by the sum of these weighted pitch positions. Note that this definition is a general one. If desired, the "beat-in-bar" can also be incorporated into this summation by giving important beats greater emphasis.

However, in this instance of the algorithm, I use pitch and duration information to generate this center. And, a c.e. is generated at each pitch event. If the i-th note is represented in the Spiral Array by pitch position p_i and has duration d_i, then the aggregate center at the i-th pitch event is defined as:

$$c_i \stackrel{\text{def}}{=} \sum_{j=1}^{i} d_i \cdot p_i.$$

The CEG method updates its decision with each note or pitch event. The distance from the key representations to c_i is calculated and ranked. The key that is closest is ranked first, next closest second, and so on:

$$\text{key choices ranked by distance to } c_i \stackrel{\text{def}}{=} \{t^i_{(1)}, t^i_{(2)}, t^i_{(3)}, \ldots\}.$$

The CEG algorithm is applied to "Simple Gifts".

Table 5.2 documents the likely keys ranked by distance. Each row documents the results up to and including the pitch event at that time. In the "Durations" column, the letters represent note values as given in Table 5.1.

Symbol	Derivation	Note value	Assigned weight
b	semi-breve	whole note	1.0
m	minim	half note	0.5
dc	dotted crotchet	dotted quarter note	0.375
c+s	crotchet+semi-quaver	quarter note + sixteenth	0.3125
c	crotchet	quarter note	0.25
dq	dotted-quaver	dotted eigth-note	0.1875
q	quaver	eigth-note	0.125
s	semi-quaver	sixteenth-note	0.0625

Table 5.1: Translation of symbols representing note values.

"Simple Gifts"		Key choice		
Pitches	Durations	First	Second	Third
C	c	f (0.5941)[1]	C (0.5998)	F (0.6152)
F	c	F (0.0742)	f (0.0742)	B♭ (0.8424)
F	q	F (0.0940)	f (0.0982)	B♭ (0.7117)
G	q	F (0.0092)	f (0.0451)	c (0.5263)
A	q	F (0.0118)	f (0.1673)	C (0.6288)
F	q	F (0.0216)	f (0.1603)	B♭ (0.5830)
A	q	F (0.0669)	f (0.2872)	B♭ (0.6861)
B♭	q	F (0.0661)	f (0.2454)	B♭ (0.4499)
C	c	F (0.0102)	f (0.1561)	B♭ (0.5836)
C	q	F (0.0044)	f (0.1374)	c (0.5758)
B♭	q	F (0.0094)	f (0.1194)	B♭ (0.4857)
A	c	F (0.0422)	f (0.2476)	B♭ (0.5941)
G	q	F (0.0455)	f (0.2502)	C (0.5168)
F	q	F (0.0399)	f (0.2344)	B♭ (0.5483)
G	c	F (0.0672)	f (0.2617)	C (0.4024)
C	c	F (0.0508)	f (0.2257)	C (0.3541)
G	c	F (0.1008)	C (0.2521)	c (0.2690)
C	c	F (0.0935)	C (0.2319)	c (0.2491)
G	q	F (0.1201)	C (0.1971)	c (0.2129)
A	q	F (0.1156)	C (0.2137)	c (0.2526)
G	q	F (0.1411)	C (0.1834)	c (0.2202)
E	q	F (0.1597)	C (0.1645)	c (0.2292)
C	c	F (0.1536)	C (0.1571)	c (0.2191)
C	c	F (0.1521)	C (0.1545)	c (0.2140)
F	c	F (0.1076)	C (0.2062)	c (0.2519)
F	q	F (0.0898)	C (0.2327)	f (0.2592)
G	q	F (0.1068)	C (0.2058)	c (0.2436)
A	q	F (0.1049)	C (0.2185)	c (0.2723)
F	q	F (0.0892)	C (0.2434)	f (0.2731)
A	q	F (0.0903)	C (0.2570)	f (0.2910)
B♭	q	F (0.0847)	C (0.2655)	f (0.2760)
C	c	F (0.0828)	C (0.2547)	f (0.2645)
C	q	F (0.0827)	C (0.2505)	f (0.2599)
B♭	q	F (0.0780)	f (0.2472)	C (0.2586)
A	c	F (0.0799)	f (0.2784)	C (0.2808)
G	q	F (0.0915)	C (0.2573)	f (0.2899)
F	q	F (0.0807)	f (0.2757)	C (0.2779)
G	c	F (0.1044)	C (0.2355)	c (0.2814)
C	c	F (0.1015)	C (0.2272)	c (0.2722)
A	c	F (0.1029)	C (0.2479)	f (0.3143)
G	c	F (0.1249)	C (0.2125)	c (0.2765)
F	c	F (0.1031)	C (0.2455)	c (0.3011)
F	c	F (0.0850)	C (0.2788)	f (0.2836)
F	m	F (0.0580)	f (0.2458)	C (0.3452)

Table 5.2: Key selection for "Simple Gifts" at each pitch event.

The First Few Steps

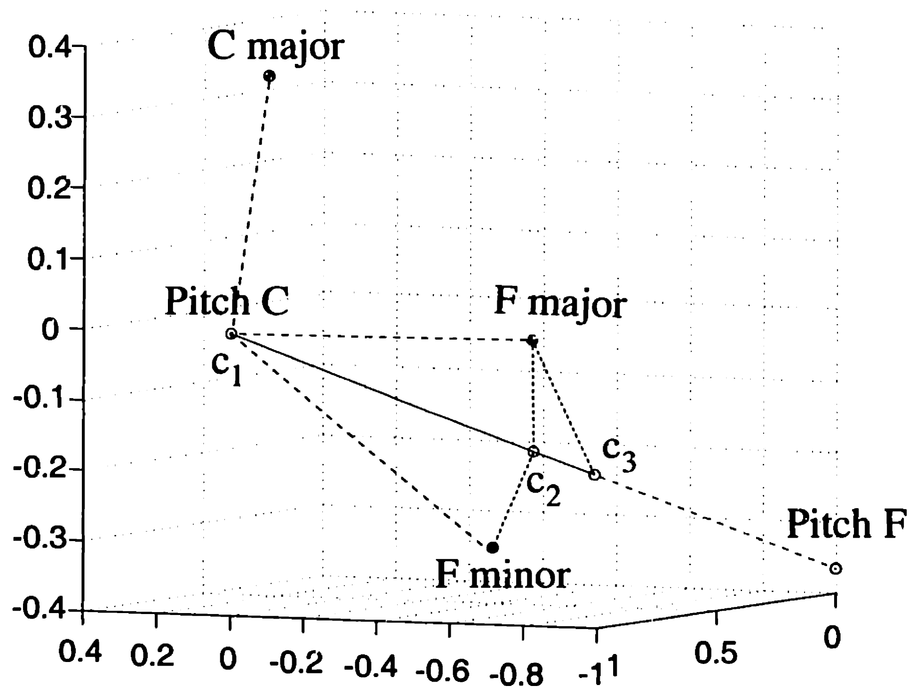


Figure 5-2: Generating centers as "Simple Gifts" unfolds.

At the sounding of the first pitch, C, the first c.e. is generated: $c_1 = \mathbf{P}(0)$. As shown in Figure 5-2, c_1 is close to F minor, C major and F major, and almost equidistant from all three. More precisely, $t^1_{(1)}$ = F minor (0.5941), $t^1_{(2)}$ = C major (0.5998), and $t^1_{(3)}$ = F major (0.6152). The bracketed number after each key represents the distance from c_1 to the key in the model.

At the onset of the second pitch, F, which happens to be of duration equal to the first, the center now shifts to the midpoint of C and F, c_2. Observe, in Figure 5-2, that this new center is now closest to F major and F minor, almost equidistant from either one. This transition from pitch C to pitch F constitutes a rising perfect fourth interval. Section 4.8 explains the implication of such a transition and gives two examples in Figure 4-14 – one in a minor key, and one in a major key. Thus the CEG algorithm's choices of F major and F minor agree with the preceived implications. In this case, $t^2_{(1)}$ = F major (0.0742), $t^2_{(2)}$ = F minor (0.0742).

With the next pitch, F, the center inches closer towards the F major key, prefering this over the F minor. Now, $t^3_{(1)}$ = F major (0.0940), $t^2_{(2)}$ = F minor (0.0982).

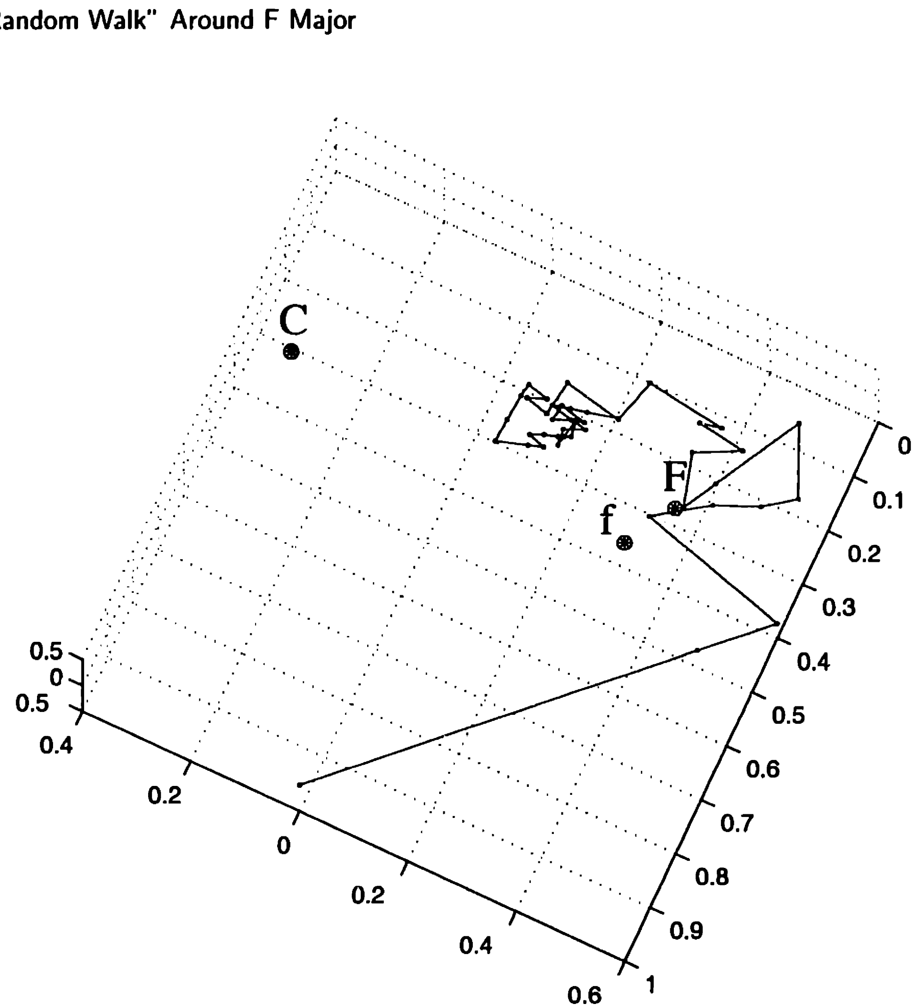


Figure 5-3: An application of the CEG algorithm: A bird's eye view of the path traced by the c.e.'s, $\{c_i\}$, as "Simple Gifts" unfolds, establishing its affiliation to F major.

Examining the contents of Table 5.2, one can see that the rising fourth between the first and second pitch event establishes the melody as being in F major. The generated centers continue to hover around F major, moving farthest away from it when the melody dwells in the dominant area. This happens around the GCGCGAGEC sequence.

Figure 5-3 shows a bird's eye view of how the centers generated by the melody's notes, $\{c_i\}$, approaches F major, then performs a "random walk" around it. The minimum distance is achieved at pitch event $i = 10$, although it comes extremely close at pitch event $i = 4$. Relating this to the numers in Table 5.2, the distance from c_i to F major is 0.0092 when $i = 4$, and 0.0044 when $i = 10$. Compare this to the initial distances of approximately 0.6 each from c_i to F minor, C major and F major.

Figure 5-4 shows the exact distances from the four closest keys (F major, C major, F minor and C minor), at each successive pitch event. Observe, in the graph, that F major quickly establishes itself as the closest key. However, between pitch events $i = 22$ to 24, C major (the dominant of F) vies with F major for preeminence. The melody dwells on the dominant key area at $i = 19$ to 24, outlining the C major triad from $i = 21$ to 24. Thus, this behavior in the model is in accordance with listener perception, and is a desirable trait.

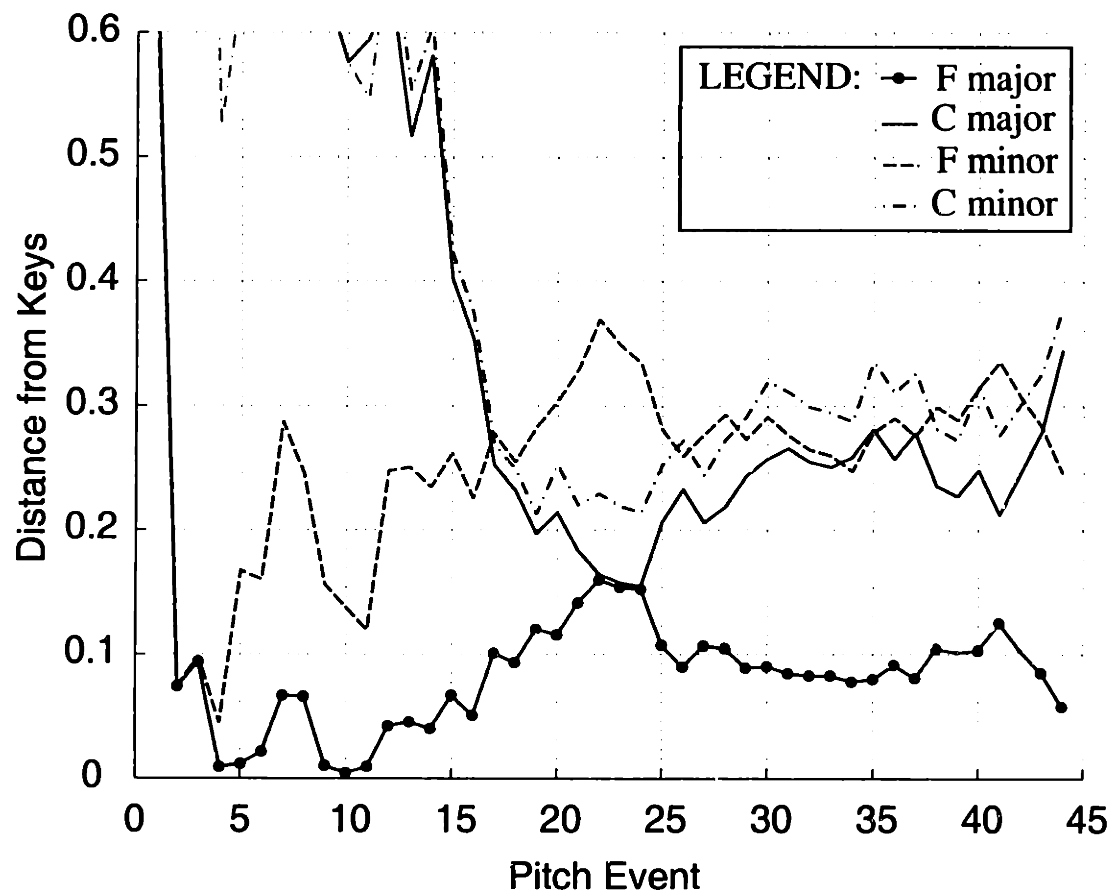


Figure 5-4: Distance to various keys as "Simple Gifts" unfolds.

5.3 Model Validation

How does one know when the algorithm generates the correct answer, or agrees with listener perceptions? In other words, for a given piece of music, how does one decide what is the correct key or when the musical passage deviates from this established key? In Chapter 7, for a collection of pitches, how does one decide if it defines a chord and how does one find the root? In Chapter 6, for modulations, how does one know when a new key has indeed been established?

One systematic way to identify and verify the answer would be to take a vote. There is, however, considerable agreement amongst informed listeners concerning the key and the chords employed in a piece of music. The precise point of modulation is subject to more debate and this will be discussed in Chapter 6.

In this example, and in others in the following chapters, I will be referring to my own musical experience for these benchmark answers. These answers are based on musical study and practice. This method is a time honoured tradition in music theory employed by Lewin to devise his "Generalized Musical Intervals and Transformations" [29], by Lerdahl and Jackendoff in their "Generative Theory of Tonal Music" [27], by Shenker in his theory of a Fundamental Structure [20] underlying musical pieces, and more recently, by Temperley in his 1996 doctoral dissertation [51]. Outside of music theory, in the field of Theoretical Linguistics the researcher refers to his/her own experience to determine if a sentence is syntactically correct.

The method has its flaws. Mine is a subjective assessment, one with which someone else may disagree. However, I proceed on the assumption that there is sufficient agreement when judging the validity of answers concerning simple questions of tonality so that a computational description of my intuitions would serve to model reasonably that of others' as well.

5.4 Comparing Key-Finding Algorithms

In this section, I will focus on comparing the CEG algorithm with two of the most established key-finding methods in the music cognition literature. These are Longuet-Higgins' and Krumhansl's algorithms. Both are described in Section 2.3. I shall call Longuet-Higgin's the Shape-Matching Algorithm (SMA) and Krumhansl's the Probe Tone Profile Method (PTPM). The benchmark examples common to the two are the fugue subjects from Bach's forty-eight Preludes and Fugues (Books 1 and 2) in his Well-Tempered Clavier (WTC). Bach wrote a pair of Prelude Fugue for every possible key in each Book of the WTC. Twelve major, and twelve minor, making a total of twenty-four in each book. He wrote two books, resulting in a total of forty-eight preludes and fugues.

I use the CEG algorithm to analyze the keys of the twenty-four fugue subjects from the WTC (Book 1), comparing these results with corresponding ones by the SMA and PTPM. Summaries of the CEG's output are given in the APPENDIX to this chapter and conclusions about the strengths and weaknesses of each algorithm are discussed at the end of this section.

The Challenge of Finding the Key of a Fugue Subject

Analyzing a single melodic line is itself a challenging task because one has to work with limited information. Furthermore, some melodies can be harmonized in several different ways as evidenced by Bach himself in his chorales [2].

Fugue subjects, by the nature of fugal compositions, are chosen for their contrapuntal potential. That is to say, they are melodies that can be tranformed in many different ways (for example, inverted, augmented, compressed time-wise) and yet still harmonize when overlapped with each other. Often, for this reason, fugue subjects are melodies that can be harmonized in many different ways, and are tonally ambiguous.

In addition, many fugue subjects establish a key, then move to the dominant key area in preparation for the second statement in the dominant key. In spite of the key ambiguities, one can always list the most likely candidates for key, recognizing that this list may change over the course of time.

Method 3: the Shape Matching Algorithm

Longuet-Higgins' Shape Matching Algorithm is discussed in detail in Section 2.3. The algorithm maps pitches onto the harmonic network, then attempts to identify the shape of a key as shown in Figure 2-4. In situations where multiple solutions still exist at the end of the passage, or where all keys have been eliminated, the algorithm invokes the tonic-dominant rule. The tonic-dominant rule derives the key from the first pitch which is assumed to be either the tonic or the dominant of the intended key.

Longuet-Higgins tested the algorithm on all forty-eight fugues in Bach's WTC. The intended key was found in twenty-six of the forty-eight fugue themes. Of the remaining twenty-two, the keys in seventeen of the fugues were found by invoking the tonic-dominant rule.

Method 2: the Probe Tone Profile Method

For Krumhansl's Probe Tone Profile Method, the collective durations of each of the twelve pitch classes are recorded in an input vector, and statistically compared to the listener-generated probe tone profiles. The output of the PTPM is a vector containing numerical values rating the strength of each key. Again, the tonic-dominant rule is invoked when there is a tie.

In *Cognitive Foundations of Musical Pitch* [24], Krumhansl compares the PTPM to Longuet-Higgin's SMA. The probe tone profile method analyzes the "melody-to-date" at each pitch event. With the prior knowledge of the intended key, Krumhansl selects the first instance in which the intended key is ranked first in strength as the point at which the key is determined.

The rule used to measure performance of the model was the number of pitches required to determine the intended key. Considering only the fugue subjects (from Books I and II) in which the tonic-dominant rule was not applied, Krumhansl's algorithm required on average 5.11 pitches, and Longuet-Higgin's 9.42 pitches. I compare the CEG's results against the PTPM and SMA in Table 5.3.

Method 1: the CEG Algorithm

In the CEG method, the Spiral Array is first calibrated to reflect a few basic tonal relationships as discussed in Section 4.8 of Chapter 4. In this implementation, weights at all levels (w, u, ω, v) are restricted to be the same, and assigned the values [0.516, 0.315, 0.168]. h= $\sqrt{2/15}$ (r= 1).

According to the discussions in Chapter 4, this particular assignment ensures that:

1 a pitch is closest to the key of the same name (Condition A);
2 a half step interval is closest to the key of the upper pitch (Condition B);
3 a major chord is closest to the pitch of the same name (its root), followed by its fifth, then its third (Theorem 7); and,
4 a minor chord is closest to the pitch of the same name (its root), followed by its fifth, then its third (Theorem 9).

Similar to Krumhansl's PTPM, the CEG considers pitch and duration information. In contrast to the PTPM that takes as input a vector of pitch names and corresponding durations, the CEG's input is a point in space, the c.e. summarizing the pitch and duration information. In this application, the model does not take into account the beat-in-bar.

	Book 1 Fugue subjects	Steps to key CEG[a]	PTPM	SMA
1	C major	2	2	16†[b]
2	C minor	5	5	5
3	C♯ major	6	7	16
4	C♯ minor	3	3	4
5	D major	2	2	15†
6	D minor	3	3	8
7	E♭ major	2	6	11†
8	D♯ minor	2	6	12†
9	E major	14	12†	11
10	E minor	3	2	7†
11	F major	4	10	6
12	F minor	3	15	4†
13	F♯ major	3	2	8
14	F♯ minor	7	18	5†
15	G major	2	2	15
16	G minor	3	3	4
17	A♭ major	3	2	7†
18	G♯ minor	5	5	5
19	A major	2	4	7
20	A minor	5	5	5
21	B♭ major	4	4	14
22	B♭ minor	2	3	6†
23	B major	2	11	11
24	B minor	3	3	7
	Average	3.75 (3.57)[c]	5.25 (4.79)	8.71 (8.21)

[a]Numbers generated using $h=\sqrt{2/15}$ ($r=1$), and weights across all heirarchies set to [0.516, 0.315, 0.168].

[b]A †indicates that the tonic-dominant rule was invoked.

[c]Numbers in brackets denote average when considering only the fugue subjects in which the tonic-dominant rule was not applied by any of the three methods.

Table 5.3: Applying key-finding algorithm to Bach's fugue subjects in the WTC.

Performance Measure

To align the CEG against the SMA and PTPM in the ensuing tests, for each incremental pitch event, the model gives as output a vector of possible keys ranked by likelihood, and a vector of distances between each key representation and the aggregate position. Note that because the model has been designed so that close relationships are indicated by spatial proximity, a low distance value represents a close match. For more meaningful comparisons, I will determine the number of pitches to key discernment in the same way as Krumhansl. The first occurence of the intended key is recorded as the pitch event at which the key is correctly determined. The key selected at the first note, even if it is the correct key, is not considered a meaningful correct choice.

5.5 Analysis of Key-Finding Algorithms' Comparison Results

The following conclusions are gathered from observations of model performance in the first twelve Fugues of the WTC Book I as given be Table 5.3, and documented in the APPENDIX to this chapter. I examine, in this section, the strengths and weaknesses of each algorithm.

Summary of Comparison Results

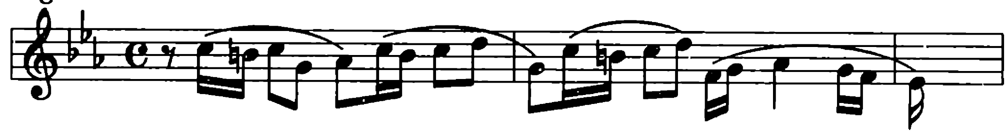
For the fugue subjects in Book I of the WTC, the CEG required on average 3.75 pitches, the PTPM 5.25, and the SMA 8.71. If one considered only the fugue subjects in which the tonic-dominant rule was not applied by any of the three algorithms, the CEG required 3.57 pitches, the PTPM 4.79 and the SMA 8.21. Details of each test run with discussion of performance for the first twelve are given at the end of the report.

Given a melody, a hypothesis of its key based on its first pitch is not a meaningful one. The reliability of a hypothesis based on two pitch events is still questionable. Hence, on average, the absolute minimum number of pitches required to form an opinion of the key is 3. The CEG algorithm required, on average, 3.75 steps to determine the key of the 24 fugue subjects. Based on the reasons stated, I claim that this is an average performance that is close to optimal.

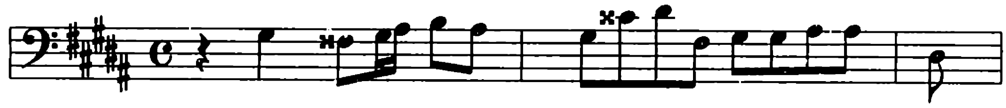
Case 1: Same Number of Steps by All Three Algorithms

The following fugue subjects required the same number of steps to determine their keys in all three algorithms. Each algorithm took five steps for each fugue subject. Fugue Nos. 18 and 20 have essentially the same pitch relations in the first five notes - $\hat{1}$, $\hat{7}$, $\hat{1}$, $\hat{2}$, $\hat{3}$ of a minor scale.

Fugue No. 2 in C minor



Fugue No. 18 in G♯ minor



Fugue No. 20 in A minor

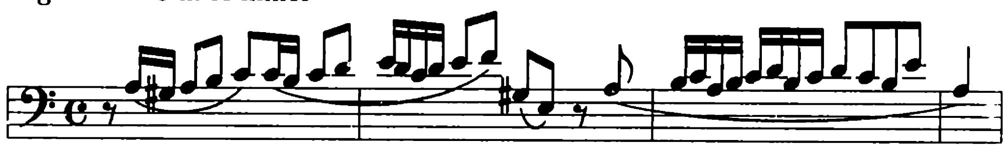


Figure 5-5: Fugue subjects requiring same number of steps to determine key in all three algorithms.

In each example, the fifth pitch is the one that informs the listener that the melody is unequivocally in the minor mode. Listener would not be able to determine if the melody was in a major or minor key given just the first four pitches - $\hat{1}$, $\hat{7}$, $\hat{1}$, $\hat{5}$ in the case of Fugue No. 2; and, $\hat{1}$, $\hat{7}$, $\hat{1}$, $\hat{2}$ in the case of Fugue Nos. 18 and 20. It is the fifth pitch that tips the balance in favor of the minor mode. Thus, all three algorithms agree with the listener's perception.

Case 1a: Almost Same Number of Steps in All Three Algorithms

For fugue subjects from Fugue Nos. 4 and 16, the CEG and PTPM took 3 steps, and the SMA took 4 steps, to determine the intended keys. In these examples, the scale degrees in the first three notes are: $\hat{1}$, $\hat{7}$, $\hat{3}$ in Fugue No. 4; and, $\hat{5}$, $\hat{6}$, $\hat{1}$ in Fugue No. 16.

In this case, it appears that both the CEG (by design) and PTPM (by inference from listener decisions) took into account interval relations important to a triad - perfect fifths and minor thirds. These intervals are formed by $\hat{1}$ – $\hat{3}$ in Fugue No. 4, and $\hat{1}$ – $\hat{5}$ in Fugue No. 16.

If one considered only the pitch collection up to step 3, Fugue No. 16 could be in C minor or G minor. And allowing for enharmonic spellings of the same pitch, Fugue No. 4 could be in E♯ minor or C♯ minor. Thus, the SMA is unable to make a decision until the second half-step (between $\hat{3}$ – $\hat{2}$ in Fugue No. 4, and $\hat{7}$ – $\hat{1}$ in Fugue No. 16) chooses the appropriate minor key.

Figure 5-6: Fugue subjects requiring almost the same number of steps to determine key in all three algorithms.

Case 2: SMA Markedly Worse than CEG and PTPM

I divide the fugue subjects in which the SMA performed markedly worse than the CEG and PTPM into two sets. In the first set, the SMA invoked the tonic-dominant rule, and in the second set, the SMA eventually chose the intended key.

Figure 5-7: Fugue subjects in which the SMA required the tonic-dominant rule to break the tie, but the CEG and PTPM performed well.

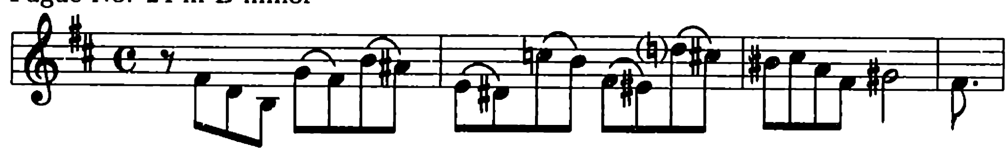
In the first set (as shown in Figure 5-7), the two fugue subjects begin with a lower tetrachord (first four scale degrees) of the intended major scale. The CEG and PTPM immediately chose the appropriate key based on the first two pitches, $\hat{1} - \hat{2}$. This implies that, given two pitches related by a whole-step interval in isolation, a listener assigns the scale degrees $(\hat{1},\hat{2})$ to the two pitches.

However, since the SMA considers only pitch collection, the pitches in fugue subjects No. 1 could be in C major or F major; and, in fugue subject No. 5, could be in D major or G major. The SMA is unable to choose between the two candidate keys for the entire duration of the subjects since the leading-tone is missing in either example.

In the first part of the second set (as shown in Figure 5-8), the SMA required 7 and 8 steps respectively for the two minor key examples, compared to the CEG and PTPM's

3. Again, this difference results from the SMA considering only the pitch collection. Fugue subject No. 24 could be in D major or B minor until the A♯ appears at step 7; and, fugue subject No. 6 could be in C major or F major or D minor or A minor until the C♯ appears at step 8.

Fugue No. 24 in B minor



Fugue No. 6 in D minor

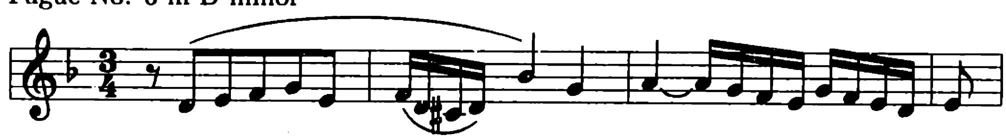


Figure 5-8: Fugue subjects in which the SMA eventually found the key, but the CEG and PTPM performed well (Part 1).

Meanwhile, the CEG recognizes and prioritizes the B minor triad (Fugue No. 24) and the ascending scale figure (Fugue 6) as significant clues to the key. The PTPM's comparisons to listener judgements comes to the same conclusion.

In the second part of the second set (shown in Figure 5-9), the SMA required 14 steps vs. the CEG's and PTPM's 4 in Fugue No. 21; and, 15 steps vs. 2 in Fugue No. 15. The SMA is unable to resolve the ambiguities until step 14 chose B♭ major in favor of F major in Fugue No. 21; and, at step 15 in Fugue No. 15, chose G major in favor of D major. Both the CEG and PTPM decided at step two, for Fugue No. 15, that the whole step interval between the first two pitches implied G major. And the two algorithms again agreed at step 4 of Fugue No. 21, that the key was B♭ and not F.

Fugue No. 21 in B♭ major

Fugue No. 15 in G major

Figure 5-9: Fugue subjects in which the SMA eventually found the key, but the CEG and PTPM performed well (Part 2).

Case 3a: Worst Case for the CEG Algorithm

In one instance, the CEG algorithm performed worse than the PTPM and SMA. This was in Fugue No. 9 (shown in Figure 5-10). The CEG algorithm took 14 steps, the SMA only 11, and the PTPM invoked the tonic-dominant rule at step 12.

Fugue No. 9 in E major

Figure 5-10: Fugue subject in which the CEG performed worst.

Because each pitch is weighted by its duration, the emphasis on F♯ (the longest note, occuring at step 2), tips the balance towards a B major/minor tonality. Ignoring the first note, the next four could be a variation on the opening to Bach's *Minuet in G* (transposed to B major). Thus, B major is not an unreasonable choice. The influence of the B major tonality is so strong, and the ensuing notes of such short duration (thus, having much less influence) that it took 14 steps to cancel out its effect.

The PTPM was probably also affected by its considering pitch duration, which is a hinderance in this case. The SMA took almost as long for different reasons. Step 11 is pitch A, that finally allows the SMA to choose E major in favor of B major.

Case 3b: Best Cases for the CEG Algorithm

The CEG algorithm outperformed the SMA and PTPM in several occasions. These examples have been divided into two sets. The first set gives the fugue subject in which the CEG algorithm chose the intended key in two steps, and both the SMA and PTPM found the key in 11. The second set is a collection of fugue subjects in which the CEG outperformed the PTPM, which in turn performed better than the SMA. Some other examples in which the CEG performed best have been categorized under other scenarios where the SMA or PTPM performed worst.

Fugue No. 23 in B major

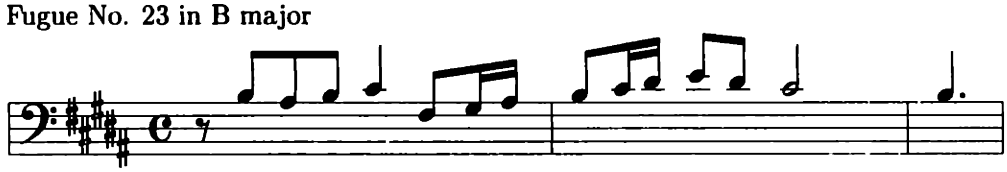


Figure 5-11: Fugue subjects in which the CEG performed best (Part 1).

The first (shown in Figure 5-11) is the Fugue in B major, and it begins with the scale degrees $\hat{1}, \hat{7}$. By design, the CEG algorithm assigns these two pitches the correct scale degrees. Both the SMA and PTPM wait until step 11 to decide on B major in favor of F♯ major.

Figure 5-12: Fugue subjects in which the CEG performed best (Part 2).

In the first two examples in the second set (shown in Figure 5-12), the CEG algorithm chose the correct key after 2 steps. As with Fugue No. 23 in the first set, the first two pitches in fugue subject No. 19 are interpreted to be $\hat{1}, \hat{7}$, and correctly so. The PTPM required step 4 ($\hat{3}$) to make the correct decision, and the SMA needed step 7 to choose A major in favor of F♯ minor.

The correct decision in Fugue No. 22, was arrived at in step 2 by the CEG algorithm rather fortuitously. One of the desired conditions when choosing weights for the model was that a perfect fourth in such a situation would imply B♭ major. However, this particular set of weights did not satisfy this condition, instead prefering the minor mode. The PTPM made the correct choice at step 3, and the SMA required the tonic-dominant rule since the pitch collection could belong to D♭ major or B♭ minor.

In Fugue No. 3, the CEG required only 6 steps. To a listener, the first five pitches could just as likely be in F♯ major as C♯, but the sixth pitch strongly suggests the C♯ major triad, that leans towards C♯ major as the appropriate key. The PTPM required one additional step to come to the same conclusion. The SMA, however, needed the B♯ at step 16 to choose C♯ major in favor of F♯.

In Fugue Nos. 7 and 8, no such defining moment occured for the SMA, and the tonic-dominant rule had to be invoked. Because the CEG algorithm prioritizes triadic relations, the first two pitches in fugue subject No. 7 were assigned, correctly, scale degrees $\hat{5} - \hat{3}$; and, $\hat{1}, \hat{5}$ in No. 8. The PTPM, in both cases, waited for the third pitch in the triad to be realized at step 6 before choosing the appropriate key.

Case 4a: Worst Cases for the PTPM Algorithm

I analyze the results in two parts. The first part address an example in which both the CEG and SMA performed better than the PTPM. In the second part, the SMA invoked the tonic-dominant rule in these examples.

Fugue No. 11 in F major

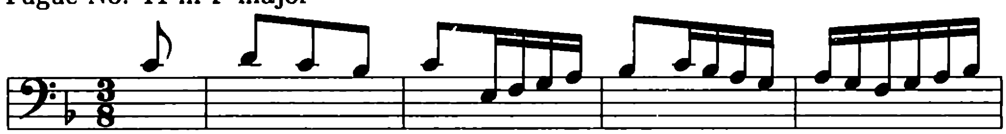
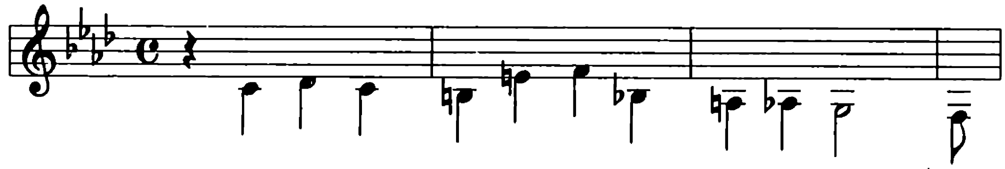


Figure 5-13: Fugue subjects in which the PTPM performed worst (Part 1).

In the first example (shown in Figure 5-13), the CEG algorithm chooses F major at step 4, a reasonable point at which to make this decision since the previous three pitches could as likely be in C major. The SMA waits for the leading tone to choose F major in favor of B♭. For some reason, the PTPM required 10 steps.

In part 2 (shown in Figure 5-14), both the examples employed chromatic pitches not in the scale, and the PTPM required many pitches to choose the correct key. The SMA is stumped early in the melody, and invokes the tonic-dominant rule. Because the CEG algorithm collapses pitch information into a single point, it is a relatively more robust algorithm when applied to melodies with chromatic pitches, especially in the case of Fugue No. 14.

Fugue No. 12 in F minor



Fugue No. 14 in F♯ minor

Figure 5-14: Fugue subjects in which the PTPM performed worst (Part 2).

Case 4b: Best Cases for the PTPM Algorithm

In the next three fugue subjects, the PTPM took 2 steps to ascertain the key while the CEG algorithm took 3 steps. In Fugue No. 13, the SMA took 8 steps; and, in Fugue Nos. 10 and 17, it invoked the tonic-dominant rule.

Fugue No. 13 in F♯ major

Fugue No. 10 in E minor

Fugue No. 17 in A♭ major

Figure 5-15: Fugue subjects in which the PTPM performed best.

In all three cases, the PTPM chose the correct key in 2 steps because it selected the minor mode when given two pitches a minor third apart, and the major mode when given two pitches a perfect fifth apart. The CEG algorithm did the opposite, but this error is corrected in the following step. The SMA required step 8 in Fugue No. 13 to select F♯ major in favor of the other candidate keys; and, required the tonic-dominant rule in the other two cases.

5.6 Commentary on the Three Algorithms

The Shape Matching Algorithm

By considering only the pitch collection, the SMA does not prioritize triadic relations. Because of its emphasis on pitch collection, the SMA is highly dependent on pitches that would eliminate competing key candidates. Often, this pitch is the leading-note (scale degree $\hat{7}$). Also due to this emphasis on pitch collection, the SMA is unable to parse fugue subjects with chromaticism. In these cases, a non-scale pitch eliminates all key candidates, and the tonic-dominant rule has to be invoked.

Only in one case does the SMA perform better than any of the other two algorithms without invoking the tonic-dominant rule. This happened in Fugue No. 11, when it chose the correct key in 6 steps, two more than the CEG algorithm, and 6 less than the PTPM. In all other cases, either the tonic-dominant rule was invoked, or the SMA took more steps than the other two to choose the intended key.

Thus, in a way, the SMA serves as a lower bound for the three key-finding algorithms.

The Probe Tone Profile Method

The PTPM captures listener's judgements about pitch-key relationships using data from live experiments. As such, it models the intuitions of a particular set of listeners.

By inference from its performance, when given two pitches a perfect fourth apart, the PTPM selects the major key of the upper pitch; when given two pitches a minor third apart, the PTPM selects the minor key of the lower pitch; and, when given two pitches a major second apart, the PTPM selects the major key of the lower pitch. These preferences agree with the instance of the Spiral Array model used in this test.

However, the PTPM lacks the appropriate interpretation of pitches defining a half step interval. And it also does not relate pitches a minor third apart to a major triad. In addition, it does not deal well with chromaticism as shown in its worst case scenario discussed in the previous section.

The Center of Effect Generator

In addition to emphasizing the pitch collection, the Spiral Array also accounts for triadic relations (in particular, the perfect fifth and major/minor third interval relations between pitches), fifth relations among chords. As a result, the CEG algorithm prioritizes these relations when assessing pitch information.

The CEG performs better than the PTPM and the SMA in the presence of chromatic pitches. By design, this particular implementation of the CEG method places greatest emphasis on the half-step (which is interpreted to be a $\hat{7} - \hat{1}$ transition), and the tonic of the key.

Different from the PTPM, the CEG interpretes two pitches a minor third apart as being the top of a major triad. Most of the time this works well, except when the minor third is actually the bottom of a minor triad as in the E minor Fugue (No. 10).

An unintentional (and probably, undesirable) trait of this instance of the CEG algorithm selects the minor key of the upper pitch when given two pitches a perfect fourth apart. Also, sometimes, the use of the duration information in generating the center of effect can be a hinderance as happened in Fugue No. 9.

In general, the CEG performed better than the PTPM and SMA, taking an average of 3.75 notes to determine the key, compared to 5.25 (PTPM) and 8.71 (SMA). On average, the absolute minimum number of pitches required to form an opinion of the key in the Spiral Array is three. (One pitch is represented by a point, and can hard hardly approximate a key representation. Two form a line, again not much information on which to guess the space occupied by a key's pitch collection.) Thus, the CEG algorithm's 3.75 average (3.57 when considering the subset where no tonic-dominant rule was invoked by the other two algorithms) is close to optimal.

5.7 Comments About the Comparison Method

Finding a way to measure model performance in the determining of key is a challenging problem. And the use of the required number of pitches as a yardstick is not without its flaws. I have two concerns regarding the use of this statistic to measure algorithm performance.

The first instance in which the intended key becomes most likely is not necessarily the point at which a key has been strongly established. This is especially true for the cases (quite a few) when this number is 2. Certainly, one may form an opinion, even a strong opinion, about the key given two pitch events. But, a key is not found after two pitches. At this point, it is simply a hypothesis to be verified or disqualified in the upcoming pitch events.

A large number of required pitch events does not necessarily portend poor model performance. Often, a melodic line can remain ambiguous, or deliberately non-commital for a while, until a strategically placed note confirms or disconfirms a key. This is particularly true for the minor key Fugues No. 2 (C minor), 18 (G♯ minor) and 20 (A minor). In all three cases, each algorithm determined the intended key after five notes, so the variance is zero. If a similar example occured that required 28 notes, this could seriously skew all the comparison.

The tonic-dominant rule, used by both the SMA and PTPM, is based on the observation, external to the algorithms, that many melodies begin on the tonic or the dominant of its intended key. Hence the first pitch that is sounded is highly likely to be the tonic or dominant of the key. In the CEG, desired conditions such as this can be built into the spiral array. I set up the model to conform to a small number of basic conditions governing relationships between tonal elements. The experiments can also be run with a different set of conditions.

Finally, the goal of the SMA differs from the PTPM and CEG in significant ways. Longuet-Higgins' SMA behaves conservatively, attempting to establish beyond reasonable doubt the identity of the unique key generated by the musical segment. Whereas the PTPM and CEG would venture a guess for the most likely key (albeit an educated guess) from the very first pitch! In addition, both the PTPM and CEG use duration information in their calculations. Pitch, duration and meter work together to establish a tonality. The SMA uses only pitch information to determine the key. Hence, the number of steps to achieve the different goals is not necessarily a fair measure of algorithmic performance.

5.8 APPENDIX: Results of key-finding in Bach's WTC (Book 1)

This section contains data generated by the model when analyzing the twenty-four fugue subjects of Bach's Well-Tempered Clavier. Of the CEG's output, I have chosen to show only the top three selections. Hence the primary result is to be read from the middle column labelled "First". The accompanying numbers in brackets record the distances between each key and the current aggregate position. The smaller the distance, the closer the fit.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C	q	C (0.6117)	c (0.6121)	f (0.6140)
D	q	C (0.1791)	g (0.2129)	F (0.2810)
E	q	C (0.1480)	a (0.2675)	G (0.4288)
F	dq	F (0.1050)	C (0.3154)	d (0.4503)
G	d	F (0.1143)	C (0.2479)	d (0.4449)
F	d	F (0.0860)	C (0.2918)	f (0.4285)
E	q	C (0.2226)	F (0.2541)	a (0.4044)
A	q	a (0.3107)	F (0.3162)	C (0.3624)
D	q	d (0.2109)	a (0.3350)	F (0.3382)
G	dq	C (0.1944)	d (0.2662)	G (0.3590)
A	s	d (0.2323)	C (0.2366)	F (0.3713)
G	s	C (0.1987)	d (0.2572)	G (0.3379)
F	s	C (0.2137)	d (0.2537)	F (0.3211)
E	s	C (0.2021)	d (0.2714)	F (0.3578)

Table 5.4: Key analysis of the fugue subject from the WTC Book I No. 1.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C	s	C (0.6117)	c (0.6121)	f (0.6140)
B	s	C (0.3579)	e (0.4277)	G (0.5384)
C	q	C (0.1515)	c (0.3720)	F (0.8444)
G	q	C (0.1036)	c (0.2008)	G (0.6396)
A♭	q	c (0.1165)	C (0.3010)	f (0.6037)
C	s	c (0.1436)	C (0.3075)	f (0.5769)
B	s	c (0.1009)	C (0.1604)	F (0.7513)
C	q	c (0.1387)	C (0.1882)	f (0.6905)
D	q	c (0.0435)	C (0.0498)	F (0.5010)
G	q	c (0.0362)	C (0.0604)	g (0.5622)
C	s	c (0.0378)	C (0.0606)	F (0.5881)
B	s	C (0.0441)	c (0.0715)	g (0.5612)
C	q	C (0.0453)	c (0.0700)	F (0.6247)
D	q	C (0.0231)	c (0.0685)	g (0.4523)
F	s	C (0.0216)	c (0.0602)	g (0.4425)
G	s	C (0.0268)	c (0.0575)	g (0.4238)
A♭	c	c (0.0180)	C (0.1387)	F (0.4138)
G	s	c (0.0111)	C (0.1329)	F (0.4416)
F	s	c (0.0174)	C (0.1388)	F (0.3846)
E♭	s	c (0.0138)	C (0.1693)	f (0.4077)

Table 5.5: Key analysis of the fugue subject from the WTC Book I No. 2.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
G♯	q	G♯ (0.6117)	g♯ (0.6121)	c♯ (0.6140)
A♯	s	G♯ (0.0714)	g♯ (0.1559)	C♯ (0.1899)
G♯	s	G♯ (0.1120)	g♯ (0.1755)	C♯ (0.2389)
F♯	s	g♯ (0.1327)	C♯ (0.1360)	c♯ (0.2073)
G♯	s	g♯ (0.1238)	C♯ (0.1516)	c♯ (0.1862)
E♯	q	C♯ (0.0685)	G♯ (0.2412)	c♯ (0.3520)
C♯	q	C♯ (0.0118)	c♯ (0.2387)	G♯ (0.4658)
G♯	q	C♯ (0.0230)	c♯ (0.1872)	G♯ (0.3760)
F♯	s	C♯ (0.0041)	c♯ (0.1470)	G♯ (0.4080)
E♯	s	C♯ (0.0077)	c♯ (0.2140)	G♯ (0.4195)
F♯	q	C♯ (0.0244)	c♯ (0.1908)	F♯ (0.3136)
D♯	q	C♯ (0.0633)	c♯ (0.2393)	g♯ (0.3226)
E♯	q	C♯ (0.0531)	c♯ (0.3145)	G♯ (0.3691)
C♯	q	C♯ (0.0340)	c♯ (0.2717)	F♯ (0.3736)
D♯	q	C♯ (0.0613)	c♯ (0.3003)	G♯ (0.3521)
B♯	q	C♯ (0.0770)	G♯ (0.2517)	g♯ (0.3573)
C♯	q	C♯ (0.0469)	c♯ (0.3146)	G♯ (0.3219)

Table 5.6: Key analysis of the fugue subject from the WTC Book I No. 3.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C♯	b	C♯ (0.6117)	c♯ (0.6121)	f♯ (0.6140)
B♯	m	C♯ (0.1462)	c♯ (0.4400)	G♯ (0.7930)
E	m	c♯ (0.0797)	C♯ (0.1542)	g♯ (0.8120)
D♯	b	g♯ (0.1238)	C♯ (0.1516)	c♯ (0.1862)

Table 5.7: Key analysis of the fugue subject from the WTC Book I No. 4.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
D	s	D (0.6117)	d (0.6121)	g (0.6140)
E	s	D (0.1791)	a (0.2129)	G (0.2810)
F♯	s	D (0.1480)	b (0.2675)	A (0.4288)
G	s	G (0.1861)	D (0.2254)	e (0.4207)
F♯	s	D (0.2295)	b (0.2569)	G (0.4417)
E	s	D (0.2483)	e (0.2638)	b (0.2820)
F♯	s	b (0.1643)	D (0.2695)	e (0.3847)
D	s	D (0.1873)	b (0.2492)	G (0.4977)
B	dq	b (0.1349)	e (0.3289)	D (0.4862)
B	s	b (0.1313)	e (0.3092)	B (0.4696)
A	dq	b (0.2262)	D (0.2477)	e (0.2986)
G	s	D (0.2525)	b (0.2772)	e (0.2785)
F♯	dq	b (0.1531)	D (0.2805)	e (0.4150)

Table 5.8: Key analysis of the fugue subject from the WTC Book I No. 5.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
D	q	D (0.6117)	d (0.6121)	g (0.6140)
E	q	D (0.1791)	a (0.2129)	G (0.2810)
F	q	d (0.1218)	F (0.3641)	a (0.4150)
G	q	C (0.1967)	g (0.2259)	G (0.2282)
E	q	C (0.1745)	a (0.2886)	G (0.3022)
F	s	C (0.1727)	d (0.3216)	a (0.3416)
D	s	C (0.2180)	d (0.2377)	G (0.3049)
C♯	s	d (0.1792)	a (0.2267)	D (0.3231)
D	s	d (0.1314)	D (0.2651)	a (0.2803)
B♭	c	d (0.1766)	g (0.1852)	G (0.3349)
G	c	g (0.0709)	G (0.1933)	C (0.3145)
A	c+s	d (0.1318)	g (0.2263)	G (0.2792)
G	s	d (0.1568)	g (0.1955)	G (0.2466)
F	s	d (0.1579)	g (0.2154)	G (0.2915)
E	s	d (0.1573)	g (0.2368)	G (0.2798)
G	s	d (0.1802)	g (0.2098)	G (0.2514)
F	s	d (0.1805)	g (0.2275)	G (0.2920)
E	s	d (0.1819)	g (0.2492)	G (0.2840)
D	s	d (0.1681)	g (0.2335)	G (0.2717)
E	q	d (0.1784)	G (0.2650)	C (0.2761)

Table 5.9: Key analysis of the fugue subject from the WTC Book I No. 6.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
B♭	s	B♭ (0.6117)	b♭ (0.6121)	e♭ (0.6140)
G	s	E♭ (0.3579)	g (0.4277)	B♭ (0.5384)
F	s	B♭ (0.0505)	b♭ (0.3066)	E♭ (0.3467)
G	s	g (0.2740)	B♭ (0.2764)	c (0.3337)
E♭	s	E♭ (0.1573)	c (0.3743)	B♭ (0.3980)
A♭	s	E♭ (0.0844)	c (0.3847)	e♭ (0.4264)
G	s	E♭ (0.1508)	c (0.2742)	B♭ (0.5409)
A♭	s	E♭ (0.1451)	c (0.3427)	A♭ (0.3889)
C	q	c (0.2112)	E♭ (0.3438)	f (0.3710)
B♭	q	E♭ (0.1515)	c (0.3185)	f (0.3934)
A	s	c (0.2410)	E♭ (0.2488)	f (0.2794)
F	s	f (0.2144)	B♭ (0.2697)	c (0.2835)
E♭	q	E♭ (0.1693)	c (0.3210)	f (0.3337)
D	q	E♭ (0.2096)	B♭ (0.2370)	c (0.2883)
C	c	c (0.1636)	f (0.2383)	F (0.3429)
B♭	s	c (0.1979)	f (0.2542)	E♭ (0.3166)

Table 5.10: Key analysis of the fugue subject from the WTC Book I No. 7.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
D♯	c	D♯ (0.6117)	d♯ (0.6121)	g♯ (0.6140)
A♯	dc	d♯ (0.1013)	D♯ (0.1907)	a♯ (0.6867)
B	q	d♯ (0.0772)	D♯ (0.3233)	g♯ (0.7262)
A♯	q	d♯ (0.0704)	D♯ (0.3027)	a♯ (0.8502)
G♯	q	d♯ (0.0167)	D♯ (0.2340)	g♯ (0.5266)
F♯	q	d♯ (0.0583)	D♯ (0.3826)	F♯ (0.5746)
G♯	q	d♯ (0.0589)	D♯ (0.3619)	g♯ (0.3787)
A♯	c	d♯ (0.0292)	D♯ (0.3066)	F♯ (0.5673)
D♯	c	d♯ (0.0213)	D♯ (0.2590)	g♯ (0.4859)
G♯	dc	d♯ (0.0861)	g♯ (0.2100)	G♯ (0.2710)
F♯	q	d♯ (0.0893)	g♯ (0.2389)	G♯ (0.3397)
E♯	c	d♯ (0.0825)	G♯ (0.2878)	g♯ (0.2888)
D♯	dc	d♯ (0.0662)	g♯ (0.2458)	D♯ (0.2548)

Table 5.11: Key analysis of the fugue subject from the WTC Book I No. 8.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
E	q	E (0.6117)	e (0.6121)	a (0.6140)
F♯	c	b (0.0947)	f♯ (0.2101)	B (0.2319)
B	s	b (0.0308)	B (0.1483)	f♯ (0.3391)
C♯	s	b (0.1155)	f♯ (0.1789)	B (0.1867)
D♯	s	B (0.0722)	b (0.1234)	f♯ (0.2241)
E	s	B (0.1057)	b (0.1405)	f♯ (0.2662)
D♯	s	B (0.0495)	b (0.1748)	E (0.3133)
E	s	B (0.0938)	b (0.1993)	E (0.2040)
F♯	s	B (0.0721)	b (0.1579)	E (0.2699)
G♯	s	B (0.0949)	b (0.2402)	E (0.2420)
A	s	B (0.1464)	E (0.2160)	b (0.2301)
B	s	B (0.1194)	E (0.1942)	b (0.1979)
G♯	s	B (0.1403)	E (0.1805)	b (0.2681)
E	s	E (0.1324)	B (0.1822)	b (0.2967)
D♯	s	B (0.1453)	E (0.1623)	b (0.3080)
E	s	E (0.1215)	B (0.1843)	b (0.3333)
F♯	s	E (0.1504)	B (0.1508)	b (0.2855)
E	s	E (0.1151)	B (0.1872)	b (0.3107)
D♯	s	E (0.1410)	B (0.1570)	b (0.3199)
E	s	E (0.1097)	B (0.1908)	b (0.3423)
F♯	s	E (0.1324)	B (0.1603)	b (0.2997)
D♯	s	B (0.1380)	E (0.1580)	b (0.3116)
C♯	s	B (0.1511)	E (0.1723)	b (0.3277)
D♯	s	B (0.1344)	E (0.1984)	b (0.3415)
E	s	B (0.1580)	E (0.1653)	b (0.3541)
F♯	s	B (0.1355)	E (0.1842)	b (0.3201)
G♯	s	B (0.1555)	E (0.1861)	b (0.3637)
A	s	E (0.1687)	B (0.1708)	b (0.3482)
F♯	s	B (0.1511)	E (0.1870)	b (0.3186)
G♯	q	B (0.1864)	E (0.1909)	b (0.3968)

Table 5.12: Key analysis of the fugue subject from the WTC Book I No. 9.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
E	s	E (0.6117)	e (0.6121)	a (0.6140)
G	s	C (0.3579)	e (0.4277)	G (0.5384)
B	s	e (0.1935)	G (0.5801)	E (0.6363)
E	s	e (0.1204)	E (0.4523)	a (0.6805)
D♯	s	e (0.1013)	E (0.1907)	b (0.6867)
E	s	e (0.0797)	E (0.1542)	b (0.8120)
D	s	e (0.0071)	E (0.1820)	b (0.5511)
E	s	e (0.0065)	E (0.1595)	A (0.5772)
C♯	s	e (0.0277)	E (0.0618)	A (0.3761)
E	s	e (0.0394)	E (0.0699)	A (0.3680)
C	s	e (0.0538)	E (0.1752)	a (0.3401)
E	s	e (0.0750)	E (0.1863)	a (0.3377)
B	s	e (0.0400)	E (0.1541)	a (0.4095)
E	s	e (0.0580)	E (0.1640)	a (0.4013)
D♯	s	e (0.0421)	E (0.0823)	A (0.4780)
E	s	e (0.0556)	E (0.0933)	A (0.4738)
A♯	s	E (0.0255)	e (0.0655)	A (0.4494)
C♯	s	E (0.0130)	e (0.1017)	A (0.3849)
G	s	E (0.0316)	e (0.0535)	A (0.3935)
F♯	s	E (0.0277)	e (0.0612)	A (0.3741)
G	s	e (0.0302)	E (0.0545)	A (0.3931)
A♯	s	E (0.0356)	e (0.0707)	b (0.3530)
F♯	s	E (0.0484)	e (0.0930)	b (0.2912)
E	s	E (0.0391)	e (0.0819)	b (0.3329)
D	q	e (0.0818)	E (0.1021)	b (0.2685)
B	q	e (0.0741)	E (0.1036)	b (0.2473)

Table 5.13: Key analysis of the fugue subject from the WTC Book I No. 10.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C	q	C (0.6117)	c (0.6121)	f (0.6140)
D	q	C (0.1791)	g (0.2129)	F (0.2810)
C	q	C (0.0714)	c (0.1559)	F (0.1899)
B♭	q	F (0.1791)	c (0.2129)	B♭ (0.2810)
C	q	c (0.1327)	F (0.1360)	f (0.2073)
E	s	c (0.1263)	C (0.1278)	F (0.1463)
F	s	F (0.0841)	c (0.1880)	C (0.1987)
G	s	c (0.1276)	F (0.1394)	C (0.1491)
A	s	F (0.1063)	C (0.1520)	c (0.1976)
B♭	q	F (0.1140)	c (0.2235)	f (0.2797)
C	s	F (0.1030)	c (0.1972)	f (0.2502)
B♭	s	F (0.1202)	c (0.2223)	f (0.2529)
A	s	F (0.0933)	c (0.2653)	f (0.2733)
G	s	F (0.1272)	c (0.2206)	C (0.2669)
A	s	F (0.1083)	c (0.2654)	C (0.2657)
G	s	F (0.1405)	c (0.2268)	C (0.2340)
F	s	F (0.1120)	c (0.2584)	C (0.2701)
G	s	F (0.1417)	c (0.2237)	C (0.2411)
A	s	F (0.1275)	C (0.2435)	c (0.2634)
B♭	s	F (0.1363)	c (0.2761)	C (0.2869)
C	s	F (0.1230)	c (0.2520)	C (0.2623)

Table 5.14: Key analysis of the fugue subject from the WTC Book I No. 11.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C	c	C (0.6117)	c (0.6121)	f (0.6140)
D♭	c	D♭ (0.3579)	f (0.4277)	A♭ (0.5384)
C	c	f (0.1935)	A♭ (0.5801)	F (0.6363)
B	c	c (0.1625)	f (0.2098)	F (0.2213)
E	c	C (0.1114)	c (0.2653)	F (0.3498)
F	c	F (0.1446)	C (0.2091)	f (0.2838)
B♭	c	F (0.0560)	f (0.1592)	c (0.2390)
A	c	F (0.0347)	f (0.2538)	C (0.2740)
A♭	c	F (0.0383)	f (0.1020)	c (0.3268)
G	m	c (0.1204)	F (0.1388)	C (0.1688)
F	q	F (0.1069)	c (0.1532)	f (0.2007)

Table 5.15: Key analysis of the fugue subject from the WTC Book I No. 12.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
C♯	q	C♯ (0.6117)	c♯ (0.6121)	f♯ (0.6140)
F♯	q	f♯ (0.0797)	F♯ (0.1542)	c♯ (0.8120)
E♯	q	F♯ (0.3153)	C♯ (0.3851)	f♯ (0.5592)
F♯	q	F♯ (0.1227)	f♯ (0.3058)	C♯ (0.5699)
E♯	s	F♯ (0.2063)	f♯ (0.4668)	C♯ (0.4684)
D♯	s	F♯ (0.1418)	C♯ (0.3725)	f♯ (0.4680)
C♯	dq	F♯ (0.1693)	C♯ (0.3117)	f♯ (0.3857)
B	d	F♯ (0.1199)	C♯ (0.2946)	f♯ (0.3242)
C♯	d	F♯ (0.1285)	C♯ (0.2916)	f♯ (0.3202)
D♯	c	F♯ (0.1135)	C♯ (0.2288)	c♯ (0.4242)
C♯	q	F♯ (0.1008)	C♯ (0.1894)	c♯ (0.3653)
B	q	F♯ (0.0689)	C♯ (0.2417)	c♯ (0.3309)
A♯	q	F♯ (0.0809)	C♯ (0.2820)	f♯ (0.4096)
G♯	q	F♯ (0.1170)	C♯ (0.2065)	c♯ (0.3411)
C♯	q	F♯ (0.1040)	C♯ (0.1762)	c♯ (0.3012)
A♯	q	F♯ (0.1108)	C♯ (0.2125)	c♯ (0.3903)

Table 5.16: Key analysis of the fugue subject from the WTC Book I No. 13.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
F♯	c	F♯ (0.6117)	f♯ (0.6121)	b (0.6140)
G♯	c	F♯ (0.1791)	B (0.2810)	f♯ (0.3056)
A	b	A (0.1397)	a (0.4397)	D (0.5415)
G♯	q	A (0.1123)	a (0.4569)	f♯ (0.5300)
A♯	q	A (0.1891)	f♯ (0.3182)	a (0.6482)
B	m	A (0.1676)	f♯ (0.2826)	E (0.3209)
A♯	q	f♯ (0.1851)	A (0.2669)	b (0.3486)
G♯	q	f♯ (0.1870)	A (0.2899)	E (0.3110)
A♯	q	f♯ (0.1259)	b (0.3589)	E (0.3633)
B♯	q	f♯ (0.1250)	F♯ (0.3296)	B (0.3493)
C♯	dc	f♯ (0.0962)	F♯ (0.2942)	B (0.4615)
B	q	f♯ (0.1066)	F♯ (0.3012)	B (0.4013)
A	q	f♯ (0.1070)	F♯ (0.3381)	E (0.4315)
C♯	q	f♯ (0.1005)	F♯ (0.3287)	E (0.4653)
B	q	f♯ (0.1099)	F♯ (0.3341)	E (0.4167)
A	q	f♯ (0.1130)	F♯ (0.3691)	E (0.4212)
G♯	m	f♯ (0.1881)	E (0.3761)	F♯ (0.3844)
F♯	dm	f♯ (0.0853)	F♯ (0.2521)	B (0.3460)

Table 5.17: Key analysis of the fugue subject from the WTC Book I No. 14.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
G	q	G (0.6117)	g (0.6121)	c (0.6140)
A	s	G (0.0714)	g (0.1559)	C (0.1899)
G	s	G (0.1120)	g (0.1755)	C (0.2389)
F♯	s	G (0.0118)	g (0.2387)	D (0.4658)
G	s	G (0.0199)	g (0.2091)	C (0.5158)
A	q	G (0.0738)	D (0.2425)	d (0.2590)
B	s	G (0.0702)	D (0.2772)	g (0.3669)
A	s	G (0.1276)	D (0.1889)	d (0.2789)
G	s	G (0.0720)	D (0.2713)	g (0.3378)
A	s	G (0.1243)	D (0.2029)	d (0.2560)
B	q	G (0.1293)	D (0.2698)	e (0.3552)
A	q	D (0.1696)	G (0.2029)	a (0.3045)
G	q	G (0.1117)	D (0.2576)	d (0.3658)
D	q	G (0.0878)	D (0.2040)	d (0.3014)
C	c	G (0.1215)	C (0.2706)	g (0.3295)
B	q	G (0.1033)	C (0.3347)	g (0.3746)
A	q	G (0.1424)	D (0.3177)	a (0.3254)
G	q	G (0.1062)	C (0.3203)	g (0.3582)
F♯	q	G (0.1054)	D (0.2814)	d (0.3504)
E	c	G (0.1897)	a (0.2834)	D (0.3403)
D	q	G (0.1567)	D (0.2914)	a (0.3147)
E	s	G (0.1759)	a (0.2927)	D (0.3058)
D	s	G (0.1615)	D (0.2844)	a (0.3082)
C	s	G (0.1644)	a (0.2991)	D (0.3091)
B	s	G (0.1625)	a (0.3110)	D (0.3159)
A	q	G (0.1894)	a (0.2674)	D (0.2790)
C	s	G (0.1924)	a (0.2609)	D (0.3025)
B	s	G (0.1887)	a (0.2713)	D (0.3077)
A	s	G (0.2016)	a (0.2526)	D (0.2914)
G	s	G (0.1850)	a (0.2727)	D (0.3058)
F♯	q	G (0.1884)	D (0.2557)	a (0.2812)

Table 5.18: Key analysis of the fugue subject from the WTC Book I No. 15.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
D	q	D (0.6117)	d (0.6121)	g (0.6140)
E♭	q	E♭ (0.3579)	g (0.4277)	B♭ (0.5384)
G	q	g (0.3114)	c (0.5077)	E♭ (0.5783)
F♯	c	G (0.1655)	g (0.2821)	D (0.4929)
G	c	G (0.0927)	g (0.1761)	D (0.7778)
A	s	G (0.0474)	g (0.1420)	D (0.6035)
B♭	s	G (0.0703)	g (0.0853)	d (0.6477)
C	q	G (0.0397)	g (0.0405)	C (0.4798)
B♭	s	g (0.0152)	G (0.0764)	C (0.5030)
A	s	g (0.0182)	G (0.0638)	C (0.4717)
B♭	c	g (0.0150)	G (0.2495)	d (0.5721)
G	c	g (0.0100)	G (0.2110)	c (0.5128)

Table 5.19: Key analysis of the fugue subject from the WTC Book I No. 16.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
A♭	q	A♭ (0.6117)	a♭ (0.6121)	d♭ (0.6140)
E♭	q	a♭ (0.0797)	A♭ (0.1542)	e♭ (0.8120)
C	q	A♭ (0.1462)	a♭ (0.4400)	E♭ (0.7930)
A♭	q	A♭ (0.1515)	a♭ (0.3720)	D♭ (0.8444)
F	q	A♭ (0.1289)	f (0.4762)	a♭ (0.4887)
D♭	q	A♭ (0.1585)	D♭ (0.2718)	a♭ (0.4396)
E♭	c+s	A♭ (0.0012)	a♭ (0.1557)	D♭ (0.4487)

Table 5.20: Key analysis of the fugue subject from the WTC Book I No. 17.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
G♯	c	G♯ (0.6117)	g♯ (0.6121)	c♯ (0.6140)
F♯♯	q	G♯ (0.1462)	g♯ (0.4400)	D♯ (0.7930)
G♯	s	G♯ (0.1401)	g♯ (0.3921)	e♯ (0.8755)
A♯	s	G♯ (0.0392)	g♯ (0.2912)	D♯ (0.5655)
B	q	g♯ (0.0396)	G♯ (0.0738)	d♯ (0.6068)
A♯	q	G♯ (0.0663)	g♯ (0.0799)	d♯ (0.2773)
G♯	q	G♯ (0.0343)	g♯ (0.0459)	d♯ (0.4246)
C♯♯	q	G♯ (0.0687)	D♯ (0.1979)	d♯ (0.2308)
D♯	q	G♯ (0.0870)	D♯ (0.1849)	d♯ (0.2142)
F♯	q	d♯ (0.1227)	G♯ (0.1444)	g♯ (0.1881)
G♯	q	G♯ (0.0858)	g♯ (0.1256)	d♯ (0.1994)
G♯	q	G♯ (0.0524)	g♯ (0.0889)	d♯ (0.2787)
A♯	q	G♯ (0.0892)	g♯ (0.1424)	d♯ (0.1905)
A♯	q	d♯ (0.1301)	G♯ (0.1359)	g♯ (0.2033)
D♯	q	d♯ (0.1238)	G♯ (0.1394)	g♯ (0.1923)

Table 5.21: Key analysis of the fugue subject from the WTC Book I No. 18.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
A	q	A (0.6117)	a (0.6121)	d (0.6140)
G♯	q	A (0.3579)	E (0.5384)	f♯ (0.7211)
C♯	q	f (0.5077)	A (0.5783)	F♯ (0.8664)
A	q	A (0.3200)	f♯ (0.5749)	a (0.7980)
D	q	A (0.1289)	f♯ (0.4762)	a (0.4887)
B	q	A (0.0844)	f♯ (0.3847)	a (0.4264)
E	q	A (0.0381)	a (0.3099)	E (0.3333)
C♯	q	A (0.0833)	f♯ (0.3994)	E (0.4057)
F♯	q	A (0.1480)	f♯ (0.2675)	E (0.4288)
E	q	A (0.0975)	E (0.3351)	f♯ (0.3715)
A	q	A (0.0632)	a (0.3997)	E (0.4227)
D	c	A (0.0893)	D (0.2823)	a (0.3567)
C♯	q	A (0.0975)	D (0.3638)	f♯ (0.4166)
D♯	q	A (0.1473)	f♯ (0.3048)	b (0.4160)
G♯	q	A (0.1604)	f♯ (0.2765)	E (0.3455)
E	q	A (0.1297)	E (0.2950)	f♯ (0.3369)
A	q	A (0.0992)	E (0.3431)	f♯ (0.3633)
G♯	c	A (0.1588)	E (0.2565)	f♯ (0.3460)

Table 5.22: Key analysis of the fugue subject from the WTC Book I No. 19.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
A	s	A (0.6117)	a (0.6121)	d (0.6140)
G♯	s	A (0.3579)	E (0.5384)	f (0.7211)
A	q	A (0.1515)	a (0.3720)	D (0.8444)
B	q	A (0.0609)	E (0.2310)	a (0.2921)
C	q	a (0.0382)	A (0.1597)	e (0.3078)
C	s	a (0.0487)	A (0.2877)	e (0.4066)
B	s	a (0.0935)	e (0.2504)	A (0.2833)
C	q	a (0.1455)	C (0.3809)	e (0.4537)
D	q	a (0.1388)	C (0.3313)	G (0.4051)
E	s	a (0.1218)	C (0.3641)	e (0.4150)
D	s	a (0.1365)	C (0.3595)	G (0.3779)
C	s	a (0.1549)	C (0.2923)	G (0.4026)
D	s	a (0.1723)	C (0.2985)	G (0.3332)
E	q	a (0.1330)	C (0.3404)	G (0.4229)
F	q	a (0.1524)	C (0.2758)	d (0.3953)
G♯	q	a (0.0689)	A (0.3452)	C (0.4410)
E	q	a (0.0639)	A (0.3306)	e (0.4133)
A	q	a (0.0366)	A (0.2841)	e (0.4657)
B	s	a (0.0453)	A (0.2756)	e (0.4027)
C	s	a (0.0542)	A (0.3161)	e (0.4351)
A	s	a (0.0426)	A (0.2960)	e (0.4585)
B	s	a (0.0508)	A (0.2884)	e (0.4021)
C	s	a (0.0591)	A (0.3253)	e (0.4321)
D	s	a (0.0626)	A (0.3243)	e (0.4349)
B	s	a (0.0718)	A (0.3188)	e (0.3862)
C	s	a (0.0788)	A (0.3517)	e (0.4141)
D	q	a (0.0912)	A (0.3557)	e (0.4273)
C	q	a (0.1055)	C (0.3637)	A (0.4157)
B	q	a (0.1194)	C (0.3894)	e (0.3955)
E	q	a (0.1102)	e (0.3737)	A (0.3875)
A	c	a (0.0674)	A (0.3215)	e (0.4311)

Table 5.23: Key analysis of the fugue subject from the WTC Book I No. 20.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
F	q	F (0.6117)	f (0.6121)	b♭ (0.2810)
G	q	F (0.1791)	c (0.2129)	B♭ (0.1899)
F	q	F (0.0714)	f (0.1559)	B♭ (0.3236)
B♭	q	B♭ (0.0176)	b♭ (0.1723)	F (0.4106)
D	q	B♭ (0.0632)	b♭ (0.3931)	F (0.3423)
C	q	B♭ (0.1097)	F (0.1908)	f (0.3684)
A	s	F (0.1494)	B♭ (0.1682)	f (0.3735)
G	s	F (0.1729)	B♭ (0.1941)	g (0.3864)
B♭	s	B♭ (0.1369)	F (0.2224)	g (0.3872)
A	s	B♭ (0.1848)	F (0.1888)	g (0.3156)
G	s	B♭ (0.2105)	F (0.2113)	g (0.3798)
F	s	F (0.1804)	B♭ (0.1881)	g (0.3120)
C	q	F (0.1133)	B♭ (0.2453)	f (0.2838)
E♭	q	F (0.1756)	B♭ (0.2033)	c (0.3302)
D	q	B♭ (0.1995)	F (0.2168)	g (0.3400)
B♭	s	B♭ (0.1652)	F (0.2450)	g (0.3254)
C	s	B♭ (0.1864)	F (0.2151)	c (0.3464)
A	s	F (0.1873)	B♭ (0.2048)	c (0.3565)
B♭	s	B♭ (0.1722)	F (0.2117)	g (0.3382)
C	s	F (0.1871)	B♭ (0.1920)	c (0.3370)
D	s	B♭ (0.1941)	F (0.2064)	g (0.3319)
E♭	s	B♭ (0.1833)	F (0.2280)	c (0.3135)
D	s	B♭ (0.1864)	F (0.2457)	g (0.3212)
C	s	B♭ (0.2025)	F (0.2221)	c (0.3003)
E♭	s	B♭ (0.1953)	F (0.2442)	c (0.3085)
D	s	B♭ (0.1968)	F (0.2589)	g (0.2920)
C	s	B♭ (0.2120)	F (0.2376)	c (0.3149)
B♭	s	B♭ (0.1880)	F (0.2556)	c (0.2897)
C	s	B♭ (0.2030)	F (0.2360)	c (0.3025)
A	s	B♭ (0.2125)	F (0.2134)	c (0.3226)
B♭	s	B♭ (0.1895)	F (0.2299)	c (0.2996)
C	s	B♭ (0.2038)	F (0.2128)	c (0.3147)
D	s	B♭ (0.2043)	F (0.2252)	c (0.2985)
E♭	s	B♭ (0.1987)	F (0.2427)	c (0.3112)
D	s	B♭ (0.1997)	F (0.2544)	g (0.2920)
C	s	B♭ (0.2120)	F (0.2376)	c (0.2787)
E♭	s	B♭ (0.2079)	F (0.2551)	c (0.2876)

Table 5.24: Key analysis of the fugue subject from the WTC Book I No. 21.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
B♭	m	B♭ (0.6117)	b♭ (0.6121)	e♭ (0.6140)
E♭	m	b♭ (0.0797)	B♭ (0.1542)	f (0.8120)
G♭	c	b♭ (0.1334)	B♭ (0.3989)	e♭ (0.5819)
F	c	b♭ (0.0772)	B♭ (0.3233)	e♭ (0.7262)
E♭	c	b♭ (0.0456)	B♭ (0.2727)	e♭ (0.3925)
D♭	c	b♭ (0.0870)	B♭ (0.4331)	e♭ (0.4636)
C	c	b♭ (0.0377)	B♭ (0.3173)	e♭ (0.4521)
D♭	c	b♭ (0.0885)	D♭ (0.3866)	B♭ (0.4581)
E♭	c	b♭ (0.1024)	e♭ (0.3613)	D♭ (0.3674)
F	c	b♭ (0.0748)	D♭ (0.3960)	B♭ (0.4046)

Table 5.25: Key analysis of the fugue subject from the WTC Book I No. 22.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
B	q	B (0.6117)	b (0.6121)	e (0.6140)
A♯	q	B (0.3579)	d♯ (0.4277)	f♯ (0.5384)
B	q	B (0.1462)	b (0.4400)	f♯ (0.7930)
C♯	c	F (0.0843)	f♯ (0.1856)	B (0.1961)
F♯	q	F (0.0714)	f♯ (0.1559)	B (0.1899)
G♯	s	F (0.0777)	f♯ (0.1751)	B (0.1911)
A♯	s	F (0.0515)	f♯ (0.2156)	B (0.2438)
B	q	B (0.1132)	F♯ (0.1423)	f♯ (0.2718)
C♯	s	F (0.1049)	B (0.1509)	f♯ (0.2179)
D♯	s	F (0.1294)	B (0.1445)	f♯ (0.2872)
E	q	B (0.1134)	F♯ (0.2014)	f♯ (0.2656)
D	q	B (0.1130)	F♯ (0.2396)	f♯ (0.3813)
C♯	m	F (0.1036)	f♯ (0.1678)	c♯ (0.2089)
B	dc	B (0.1452)	F♯ (0.2080)	f♯ (0.2427)

Table 5.26: Key analysis of the fugue subject from the WTC Book I No. 23.

Fugue Subject		Key choice		
Pitches	Durations	First	Second	Third
F♯	q	F (0.6117)	f♯ (0.6121)	b (0.6140)
D	q	D (0.5487)	b (0.6953)	d (1.0642)
B	q	b (0.1935)	D (0.5801)	B (0.6363)
G	q	G (0.3579)	b (0.4277)	D (0.5384)
F♯	q	b (0.2516)	D (0.5279)	G (0.6272)
B	q	b (0.1932)	B (0.6110)	G (0.6629)
A♯	q	b (0.0704)	B (0.3027)	f♯ (0.8502)
E	q	b (0.0167)	B (0.2340)	e (0.5266)
D♯	q	b (0.0253)	B (0.1040)	E (0.5514)
C	q	b (0.0477)	B (0.2328)	e (0.3092)
B	q	b (0.0655)	B (0.2338)	e (0.3034)
F♯	q	b (0.0301)	B (0.1968)	e (0.3858)
E♯	q	b (0.0205)	B (0.0757)	E (0.3757)
D	q	b (0.0038)	B (0.1393)	E (0.4310)
C♯	q	b (0.0181)	B (0.1277)	E (0.4034)
B♯	q	B (0.0570)	b (0.0760)	f♯ (0.3260)
C♯	q	B (0.0794)	b (0.1120)	f♯ (0.2502)
A	q	b (0.1201)	B (0.1324)	f♯ (0.2180)
F♯	q	b (0.1103)	B (0.1298)	f♯ (0.2030)
G♯	m	B (0.1480)	f♯ (0.2433)	E (0.2753)
F♯	dq	B (0.1281)	f♯ (0.2083)	b (0.2536)

Table 5.27: Key analysis of the fugue subject from the WTC Book I No. 24.

6 — Determining Modulations

One of the primary tasks performed by the human mind when listening to a piece of music is to organize the information into coherent segments. A means by which to organize musical information is to categorize sections by their keys. The point of modulation marks the end of the influence of one tonal center, and the beginning of a new center. Designing a computer algorithm that can find modulations is a difficult problem which has not yet been solved.

In this chapter, I propose the first computationally viable algorithm (the Boundary Search Algorithm) to determine modulation boundaries using the Spiral Array model. The algorithm uses computer search to find the best boundary between any two keys, the point of modulation in a musical passage. By turning to a geometric model, and using spatial distance to measure closeness, determining the points of key transitions becomes a tractable problem.

6.1 The Boundary Search Algorithm

Given the number of boundaries, the Boundary Search Algorithm scans a musical passage for the best placement of these boundaries between key changes. Section 6.5 suggests ways in which this basic module can be generalized to finding multiple boundaries in more complex pieces.

Fundamental Idea Underlying the Algorithm

The idea behind the Boundary Search Algorithm is a simple one. Like the CEG key-finding algorithm, the BSA uses the center of effect idea to collapse each set of pitch and duration information down to a single point in space.

If a collection of pitches generate a specific tonal center, their center of effect (c.e.) will be close to the key representation in the Spiral Array. The distance from this c.e. to the key is an indication of how well the pitch collection compares to the pitch relations of that key. Because the distance between a c.e. and the closest key indicates how well the pitch collection compares to the pitch relations in that key, a good fit would be one in which this distance is small.

When two or more key areas exist in a musical passage, one can insert boundaries that define these key areas. The problem of finding the point of modulation can then be posed as a problem of determining the placement of boundaries so that the notes in each section generate centers that are close to their respective keys in the Spiral Array.

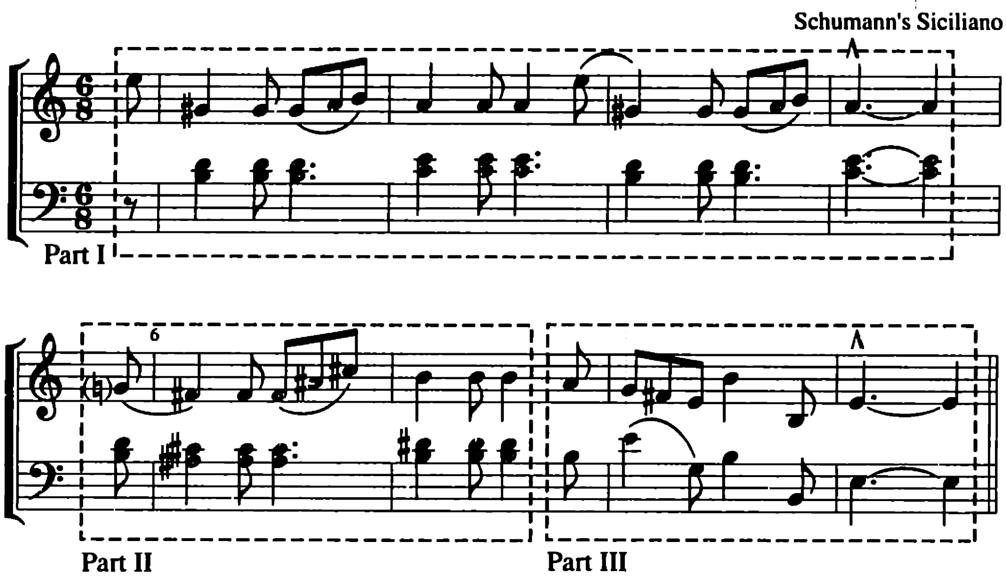


Figure 6-1: Example: Siciliano by Schumann. Grouped (by EC) into three parts, each in a different key.

To get sense of how this idea works, see, for example, Schumann's Siciliano in Figure 6-1. It begins in A minor, shifts to B major in bar 6 as a dominant preparation (B major is the dominant major of E minor) for the next part in E minor. Consider the pitch collections and their respective durations in each of these parts. The set of notes from each part can generate a center in the Spiral Array as described in Section 5.1. The respective center generated by each part will then be close to the key in that part of the piece. For this example, the key representations in the Spiral Array were generated using the parameters $w = u = \omega = v = [\ 0.516, 0.315, 0.168\ ]$, h=$\sqrt{2/15}$ (r=1).

Table 6.1 documents the resulting distances from each part's c.e. to the three closest keys. Observe, from the numbers in the table, that there isn't as clear a preference between the three choices in Part I as there are in Parts II and III. In Part I, E minor is slightly preferred above A minor, which is in turn slightly preferred over A major. In Part II, B major is strongly preferred over the other two keys. In Part III, E minor is strongly preferred over the other two keys. These choices demonstrate the shifting tonal centers in this musical passage.

	Part I	Part II	Part III
Choice 1	e (0.1182)[a]	B (0.0453)	e (0.0132)
Choice 2	a (0.1772)	b (0.2958)	E (0.1762)
Choice 3	A (0.2357)	F (0.3750)	a (0.6132)

Table 6.1: Key choices as determined by CEG (see Chapter 5) for each of the three parts of Siciliano delineated in Figure 6-1.

[a] As before, numbers in brackets indicate the distance of the c.e., generated by the notes in this part, to the key representation in the Spiral Array.

Description of the Algorithm

Having demonstrated that the different key areas in a passage can generate c.e.'s that are close to different key representations, I now describe the Boundary Search Algorithm.

I will restrict the description of the BSA to applications with three parts, and where the key of Part III is the same as the key of Part I. The algorithm's task is to find two boundaries, B_1 separating the intended key of the passage and the secondary key area, and B_2 marking the return to the original key. This is a musical form often employed by composers.

If c_i represents the center generated by the notes between B_{i-1} and B_i (where B_0 and B_3 denote the beginning and end of the piece); and, T_i represents the key closest to c_i, then the boundaries are chosen so that they minimize the sum of the distances between the c.e.'s and the respective key representation:

$$[B_1, B_2] = \arg \min_{B_1, B_2} \sum_{i=1}^{3} \|c_i - T_i\|^2$$

For passges in which the third part returns to the intended key, one would constrain the first and third key areas, T_1 and T_3 to be the same ($T_1 = T_3$).

Boundary Search Algorithm

For all boundaries B_1, B_2,

$$\begin{aligned} c_1 &= \text{c.e. of notes between } B_0 \text{ and } B_1; \\ c_2 &= \text{c.e. of notes between } B_1 \text{ and } B_2; \\ c_3 &= \text{c.e. of notes between } B_2 \text{ and } B_3; \\ T_1 &= \text{nearest key to } c_1; \\ T_2 &= \text{nearest key to } c_2; \\ T_3 &= \text{nearest key to } c_3; \end{aligned}$$

$$d(B_1, B_2) = \|c_1 - T_1\|^2 + \|c_2 - T_2\|^2 + \|c_3 - T_3\|^2$$

end;

Find (B_1^*, B_2^*) that minimize $d(B_1,B_2)$ s.t.

$$\begin{aligned} & T_1 \neq T_2, && \text{(the first and second keys are different)} \\ & T_2 \neq T_3, && \text{(the second and third keys are different)} \\ \text{and} \quad & T_1 = T_3 && \text{(the first and third keys are the same).} \end{aligned}$$

I will apply this basic algorithm to two examples in the following sections. But before applying the BSA to the examples, I first discuss the model validation technique.

6.2 Model Validation: On Modulations and Boundaries

This chapter references two types of modulation boundaries. The first follows from the conventional textbook definition of modulations; and, the second is one that takes into account other factors, such as, figural groupings and retrospective judgements, for example about the dominant preparation of the new key. Brody[1] refers to these two categories as analytic and synthetic, respectively.

In the examples of Section 6.3 and 6.4, I select my boundary choices, the BSA chooses its boundaries independent of my decision, then I compared the two answers. Differences between the two sets of answers exist for a variety of reasons. One main reason for this discrepancy is due to the difference between analytic and synthetic boundaries.

The textbook definition of modulation is based on a strict set of principles that guide analytic judgements about where a modulation occurs. This definition depends significantly on the pitch collection, selecting the boundary for a new key when the current key framework no longer contains the sounding pitches in its pitch collection. In a way, the BSA's approach generates answers that are closer to this conventional definition.

[1] Martin Brody, Professor of Music, Wellesley College. Personal communication, 27 December, 1999.

My boundary choices depend on several factors external to the pitch collection. Sometimes, these choices take into account figural groupings, waiting for a sequence pattern to end before placing a boundary. If the pitch collection prior to a new key area acts as a dominant preparation to the new key, I may choose to think of these pitches as being already in the new key. Such a decision requires one to reconsider the notes that were sounded earlier in light of new key-area information.

As a performer, these subjective decisions influence my choice of interpretations, which could be vary from one performance to the next. However, it should be noted that my boundary choices may not coincide with the textbook definition for modulations.

6.3 Application 1: Minuet in G by Bach

Figure 6-2: Modulation boundaries. [EC = my selections; A = algorithm's choices]

The first example is Bach's *Minuet in G* from "A Little Notebook for Anna Magdalena". For the two examples in this and in the next section, the key representations in the Spiral Array were generated by setting the weights (w, u, ω, v) to be the same and equal to [0.516, 0.315, 0.168], h=$\sqrt{2/15}$ (r=1). Recall, from Section 4.9, that this set of assignments satisfy two conditions: two pitches an interval of a half-step apart generates a center that is closest to the key of the upper pitch; and, an isolated pitch is closest to the key of the same name.

The Minuet is intended to be in the key of G, with a middle section in D major. Figure 6-2 shows a segment of the piece that contains the transitions away from and back to the intended key G. On the figure, are markings denoting my choice of boundaries between the keys of G and D. Also marked, are the BSA's choices for these boundaries.

Note that my choices do not coincide exactly with the BSA's. The difference between A1 and EC1 is one between a synthetic and analytic choice as mentioned in Section 6.2. I chose a boundary, EC1, before the confirmation of D major in bar 20. What I did was, having established in my mind that the piece moves to D major, I listened to the bars before that transition again, and in retrospect, decided that these notes in bar 19 sound to me like they are already in the key of D major. The BSA, on the other hand, lacks the ability to retroactively influence its decision. It chooses the beginning of bar 20, a very reasonable choice, to mark the beginning of the D major section. This choice, in fact, agrees with the textbook definition of modulation.

For the second boundary, B2, both EC2 and A2 fall on either side of the textbook boundary, which would sit just before the C ♮. My decision, EC2, was influenced by the four-bar phrasing that was established from the beginning of the piece. I waited for the latest phrase to end (at the end of the bar) before placing EC2 at the barline. This choice groups the D and C♮ with the D major of the middle section. The BSA chose to group this same D and C♮ with the G major section, a choice that is closer in spirit to the textbook boundary than mine. In this case, the D is a $\hat{5}$ in G major, and not a $\hat{1}$ in D major.

Both choices for modulation boundaries are valid for different reasons. Mine was more subjective and involved retroactive decisions. The BSA took the pitches much more literally, and chose the boundaries accordingly.

6.4 Application 2: Marche in D by Bach

For the second example, I use Bach's *Marche in D*, also from "A Little Notebook for Anna Magdalena". This piece was chosen because it is a little more complex than the previous Minuet example. Its transition from A major back to D major is peppered with brief moments in a number of different keys. A portion of the piece is shown in Figure 6-3.

The A major of the middle section shifts to D major in the second beat of bar 12, portending the return to D major. The D acts as a V to the G major tonality at the end of bar 13; a G♯ in the bass at the end of bar 14 acts as a leading note to A major; there is a momentary hint of B major between bars 15 and 16 that acts as the V of E minor; E minor acts as a V to A, which works as a V to D major which returns in bar 18.

Clearly, this is a less than perfect example on which to test an algorithm that looks for boundaries between only three parts. In spite of this incompatibility, I proceed to feed this piece into the computer, and Figure 6-3 shows the boundaries I chose (EC1 and EC2), and the boundaries the BSA selected (A1 and A2).

The first boundaries, EC1 and A1 were close to identical. The BSA breaks up the phrase structure in bar 6 to include the last beat in the new A major section, while I waited

Figure 6-3: Modulation boundaries. [EC = my selections; A = algorithm's choices]

for the completion of the phrase to place the boundary EC1. This is because, it is unaware of the phrasings, groupings suggested by the composer of the editor. Otherwise, the two choices are not very different. However, it is important to note that the G♯ in the first beat of bar 6 already belongs to the new key area, A major. So, technically speaking, the A major section should begin at the G♯.

For the set of second boundaries, marked EC2 and A2 respectively in Figure 6-3, the choices differed quite a bit. I took into account the sequential pattern starting from the middle of bar 13 and ending at the climax in bar 16, and chose to think of the final D major section as beginning in beat 1 of bar 18. The BSA, on the other hand, did not account for figural groupings and sequential patterns in the notes. It chose, instead, to begin the last D major section as soon as the pitch collection of the third part agreed with the pitch relations in D major.

6.5 Conclusions and Future Directions

Being able to determine modulations is invaluable not only to the computer analysis of music, but also to the interpretation of music for performance. For example, the Bach Minuet example in Figure 6-2 shows two ways to draw the boundaries between the two keys. These different solutions suggest alternative ways to group the figures in the piece.

The BSA selects key boundaries that incorporate elements of the conventional textbook definition of modulations. Its solutions are influenced most strongly by the pitch collection in each part, and triadic relationships among the pitches.

As an averaging technique, the BSA prioritizes triadic relations among pitches, and minimizes non-triadic relations. Its choices do not always correspond to the conventional definitions. Nor does not account for figural groupings and sequential patterns or "know about" dominant preparations to new key areas. In reality, these factors interact with pitch information to generate perceptions of boundaries between key areas, as evidenced by the example of the Marche in Figure 6-3. The human's melody pattern-recognition capability is one which the machine algorithm lacks.

Future Variations

The BSA is a simple prototype for more complicated algorithms that can determine key boundaries using the Spiral Array. More sophisticated variations of this basic module can be designed and applied to music with more complex modulation patterns.

In this chapter, I have limited the computer's task to finding only two boundaries. In the *Marche in D*, this assumption of having only two key changes is violated. In such cases, one could design an algorithm that searched for more than two boundaries. To increase its level of sophistication, one could allow the algorithm the freedom to decide the number of boundaries. For example, a new algorithm, BSA2 that uses the BSA as a module, could iteratively cut the passage into smaller segments: first into three parts, then each of these parts into smaller pieces, and so on.

Alternatively, one could design and algorithm that might be closer to the human process of listening to music, analyzing a piece of music sequentially to find out when a new key enters the picture. In this fashion, the computer model can be used as a tool for examining the conditions for key establishment. Some questions could be: How many different pitches are necessary to establish a tonal center? What are the significant pitches for a tonality? Do their durations make a difference? Such a model provides one means of objectively testing hypotheses on such issues.

On a smaller scale, Chord changes also generate boundaries between different harmonies. Determining when a new chord has occurred is akin to finding when a new key has been established. Thus, one could apply the same concept to the design of a similar procedure to find when a chord change occurs. In the next chapter, I show a different way of determining the roots of chords, using the Spiral Array, in a piece of music.

As a machine model for determining key boundaries, the BSA could become one module in a sophisticated system incorporating multiple factors for computer analysis or computer generating of music.

7 — Determining Chords

This chapter contains preliminary work on the use of the Spiral Array to design an algorithm for finding the roots of chords. The goal is not in finding an optimal algorithm for determining chords, but to demonstrate that the model can be used to design such algorithms. Given a musical passage that has been divided into notes-per-beat or notes-per-bar, the CEG2 (Center of Effect Generator 2) algorithm generates a list of chord names for each group of notes. In order to compare the performance of the Spiral Array to two other known models (Winograd's and Temperley's), the algorithm designed is tested on passages from Beethoven's Op. 13 ("Pathetique" Sonata) and Schubert's Op. 33.

7.1 An Early Example: Minuet in G

This application of the Spiral Array was discovered in my process of trying to find an algorithm to determine keys in Bach's *Minuet in G*, an example which is given in Chapter 5. In the initial test, I first grouped the pitches bar by bar as input for the computer program. The program then attempted to find the nearest key to the center generated by the pitches in each bar.

At that time, the tests were conducted with the simple weights [0.5, 0.3, 0.2], assigned to all levels (major chords and keys). In this early experiment, I only considered major keys. The results have been generated again, in Figure 7-1, using $w = u = \omega = v =$ [0.5353, 0.2743, 0.1904]. The choice of these values are explained in Section 4.9.

Not surprisingly, this method produced a bar-by-bar analysis of the chords employed. And since, for this example, the chords mostly changed bar-by-bar, this technique produced a good approximation to the list of chords traversed in the piece as shown in Figure 7-1. In addition, because the tonic (I) chord contributes the greatest weight to the key representation. As a result, a point which is close to the key, is necessarily also close to its tonic chord.

7.2 Incorporating Seventh Chords

A quick implementation applied to the Beethoven and Schubert pieces finds the definition of chords as given in Section 3.3 lacking because these definitions are limited to triads. Both Beethoven and Schubert used seventh chords frequently, dominant triads plus an

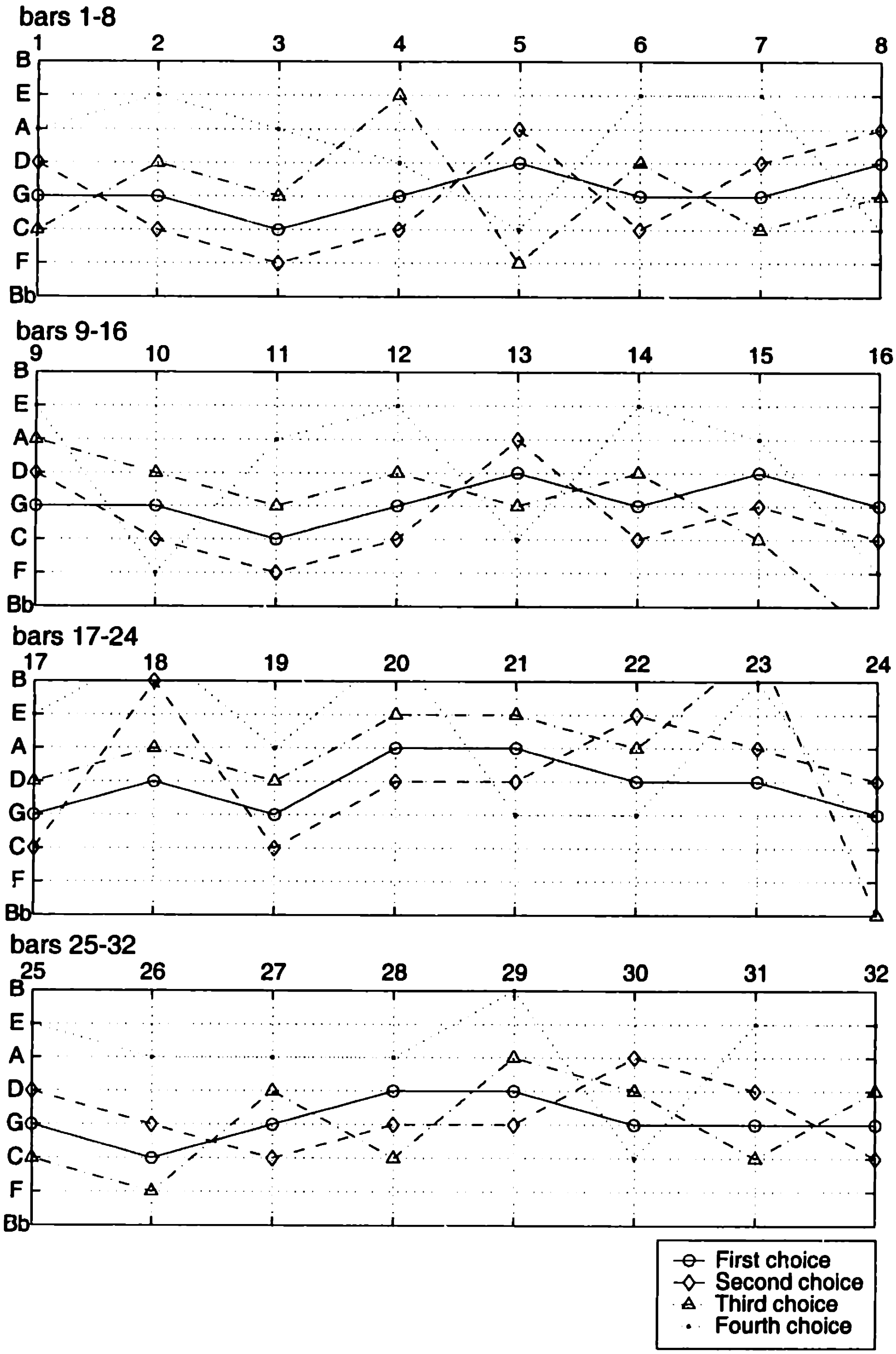


Figure 7-1: Bar-by-bar Analysis of Bach's Minuet in G.

additional fourth pitch a minor seventh from the root, leading to different effects and relative importance of the interval configurations. The function and effect of particular chords differ in different contexts and between different composers and periods. In the works of earlier composers such as Bach, the dominant seventh chord occured less frequently. As a result, it was not necessary to incorporate seventh chords into the definitions.

In order that my model should be able to recognize other chords that occur naturally in the major and minor scales, I have make additional definitions for diminished triads and dominant and other seventh chords. The goal, now, is to design an algorithm that could not only identify the roots of triads, but those of seventh chords as well. Chord identification is carried out by searching for the chord representation that is closest. Again, proximity is the indicator for a good match.

In Section 3.3, a chord is defined to be the weighted average of its root, fifth and third, where the weights sum to one, are non-zero a monotonically decreasing. I extend this definition to seventh chords. For notational ease, I will write chords as a series of thirds from the root up, 1 = major third, and 0 = minor third. So a major chord is $\mathbf{C}_{10}$, and a minor chord is $\mathbf{C}_{01}$.

Definition 6. *A quick and naive approach to defining diminished triads and seventh chords is given as follows:*

$$\begin{aligned}
\mathbf{C}_{00}(\mathrm{k}) &= \frac{1}{3}\cdot\mathbf{P}(\mathrm{k}) + \frac{1}{3}\cdot\mathbf{P}(\mathrm{k}-3) + \frac{1}{3}\cdot\mathbf{P}(\mathrm{k}-6)\\
\mathbf{C}_{010}(\mathrm{k}) &= \frac{3}{8}\cdot\mathbf{P}(\mathrm{k}) + \frac{1}{8}\cdot\mathbf{P}(\mathrm{k}+1) + \frac{2}{8}\cdot\mathbf{P}(\mathrm{k}-3) + \frac{2}{8}\cdot\mathbf{P}(\mathrm{k}-2)\\
\mathbf{C}_{100}(\mathrm{k}) &= \frac{3}{8}\cdot\mathbf{P}(\mathrm{k}) + \frac{1}{8}\cdot\mathbf{P}(\mathrm{k}+1) + \frac{2}{8}\cdot\mathbf{P}(\mathrm{k}+4) + \frac{2}{8}\cdot\mathbf{P}(\mathrm{k}-2)
\end{aligned}$$

The choice of the weights are based on the following rationale: the root is relatively important in the seventh chords ($\mathbf{C}_{010}$ and $\mathbf{C}_{100}$) so I have chosen to assign the root the largest weight. The dominant is implied by the seventh and the third, hence it has the least weight. The diminished triad is inherently unstable and ambiguous so I assign all pitches equal weight for this chord. Assigning equal weights to any two pitches is akin to the idea in information theory the probabilistic distribution with least information is one which assigns the same probability to each outcome. Here, the problem is not one of statistical outcome, but assigning equal weight essentially means that not much is known about the relative importance of the pitches.

In the future, one may derive better weight assignments in a systematic fashion similar to that shown in Chapter 4. To conduct such an analysis, a deeper understanding of the different contexts for seventh chord usage is required. For now, I choose the innocuous weights as given in Definition 6.

7.3 The Algorithm

At present, the algorithm, which I shall call the CEG2 (Center of Effect Generator 2), takes as input only pitch information without their durations. It then generates a center of effect based on this information. Comparing this center to the chord representations in the Spiral Array, the algorithm selects the closest chord as the most likely.

I found, in these experiments for determining chord identity, that the durations are not as crucial to the algorithm. In fact, by allowing durations to affect the position of the center of effect generated, the model has to undergo further fine-tuning. Better weight assignments have to be chosen in order that the model can be robust enough to give valid answers under different sets of pitch durations.

The tests are conducted with the weights [0.6025, 0.2930, 0.1145] for major triads and major and minor keys, and [0.6011, 0.2121, 0.1868] for minor triads. The conditions that these parameter values satisfy are outlined in Sections 4.8 and 4.9. As usual, h= $\sqrt{2/15}$ (r= 1). The weights for the seventh chords are as given in Definition 6.

7.4 Application I: Beethoven Op. 13

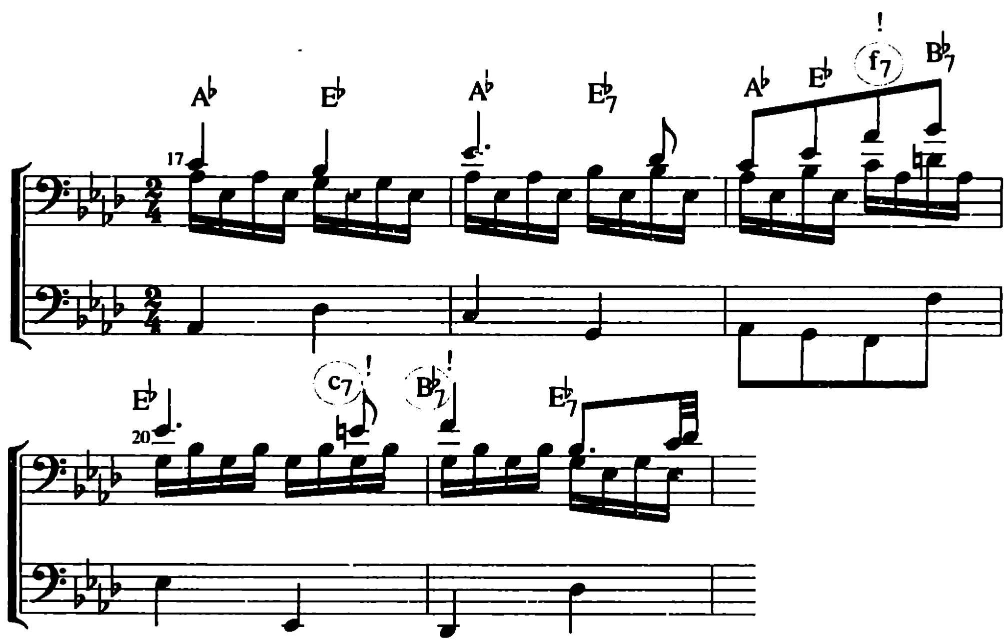


Figure 7-2: Excerpt from Beethoven's Sonata Op. 13 "Pathetique". Chord assignments are chosen by the CEG2 Algorithm.

The CEG2 algorithm was tested on the Beethoven passage shown in Figure 7-2. This passage was chosen so that a comparison can be made to Temperley and Sleator's algorithm [51], which I will call the Preference Rule Approach (PRA). The input were pitches grouped by every eighth-note beat. The output of my algorithm is given in Table 7.1.

In Figure 7-2, the chord names appearing above the music denotes the chord root chosen by the CEG2 algorithm for the group of notes in that eighth-note beat. If no chord name shows above that eigth-note beat, the chord remains the same as the last one designated. Erroneous answers have been circled.

The answers mostly concurr except for eighth-note indices 18 and 20. The CEG2 was unable to recognize the half-diminished E♭ chord in eighth-note 18. This is because the half-diminished seventh chord has not be defined in its structure. In eighth-note 20, the CEG2 agreed with Jeanne Bamberger's hearing of the C as a passing tone between two E♭ seventh chords. Temperley's algorithm placed greater emphasis on the tactus, and interpreted the C as an important chord pitch.

At eighth-note number 16, the CEG2 mistakenly labels an A♭ (I) chord a C_7 because of the E natural. This is because the machine does not know about voice-leading and passing tones. The error in eighth-note number 11, labeling an F chord a F_7 is due to the choice of weights. The weights can be adjusted to reflect this basic preference, that is, the aggregate position of the three triad pitches should be closer to the triad than to the seventh chord with these three pitches.

Eighth-beat Index	Chord Function	Chord Name	
	JBamb	CEG2	PRA (roots only)
1	I	A♭	A♭
2	I	A♭	A♭
3	V_2	$E\flat_7$	E♭
4	V_2	$E\flat_7$	E♭
5	I_6	A♭	A♭
6	I_6	A♭	A♭
7	V_6	$E\flat_7$	E♭
8	V_{65}	$E\flat_7$	E♭
9	I	A♭	A♭
10	V	E♭	E♭
11	VI	f_7	F
12	V_{65} of V	$B\flat_7$	B♭
13	I	E♭	E♭
14	I	E♭	E♭
15	I	E♭	E♭
16	I	c_7	E♭
17	V_2	$B\flat_7$	G
18	V_2	$B\flat_7$	G-B♭
19	V_2	$E\flat_7$	E♭
20	V_2	$E\flat_7$	E♭-A♭

Table 7.1: Comparison of CEG2 and PRA, both applied to the 5 bars of Beethoven's Sonata Op. 13 shown in Figure 7-2. Chord function solutions provided by Jeanne Bamberger.

7.5 Application II: Schubert Op. 33

To compare my algorithm with Winograd's [54], I apply it to Schubert's Op. 33. The passage is shown in Figure 7-3. Winograd applies decision-tree techniques used in linguistics to music, heuristically assigining weights to each branch of the tree. The weights that emerge from the algorithms are important in that they provide information about the importance the listener places on different musical information.

Winograd's algorithm, which I will call the Systemic Grammar Algorithm (SGA), returns as output assignments of functions to each chord. The function of a chord is closely linked to the chord name of the pitch cluster being considered, and the key of the passage. Knowing any two of the three gives one enough information to deduce the third. This, is of course, not a model of how humans recognize chord function, but a crude approximation of the process. In the following sections, I attempt to assess both the key and the chord names of the given Schubert passage.

Determining Chords for Schubert's Op. 33

First, I use the CEG2 algorithm to identify the chords in the Schubert passage. The notes are grouped by hand, to coincide with Winograd's choices for groupings for analysis. The results are tabulated on Table 7.2.

The chords which the CEG2 was not able to to identify correctly were some of the ones that were labeled "auxilliary (aux)", "antecedent (ant)" and "passing (P)". The CEG2 labeled the auxilliary chord in bar 6 a G diminished chord, but the pitch collection does not belong to any one triad or seventh chord. The CEG2 labeled the antecedent chord in bar 11 an $E\flat_7$ chord, and again, the pitch collection does not belong to any known chord. The same for the two passing chords in bars 15 and 16.

As commented upon by Martin Brody, part of what is interesting about the Schubert example is that it contains suspended harmonies. For example, in the opening chords of this passage, and in the 11th bar of the example. The tonic harmony of the upbeat can be heard as a suspended 6/4 chord resolving to V_7. The first beat of the 11th bar uses the now dissonant B♭ chord (functioning as the dominant of E♭, which is IV of B♭) as a suspension over the E♭ bass. These characteristics all affect chord clarity. Currently, the CEG2 is unable to distinguish between chord tones, suspension chords and passing tones. Thus, in a sense, the SGA is more "intelligent" in that it is able to distinguish between real chords and passing ones. In the future, this could perhaps be incorporated into the distances in the Spiral Array using a few basic rules.

Bar	Chord Name		Key
	CEG2	SGA (chord function)	given by SGA
	B♭ (0.1695)[a]	Ic[b]	B♭
1	B♭ (0.1695)	Ic	
	F_7 (0.0333)	V_7	
2	F (0.1695)	V	
	B♭ (0.0116)	P[c]	
	f_7 (0.0287)	P	
3	B♭ (0.1695)	I	
	B♭ (0.1695)	Ic	
4	B♭ (0.1695)	I	
	B♭ (0.1695)	Ic	
5	B♭ (0.1695)	Ic	
	F_7 (0.0333)	Ic	
6	F (0.1695)	V	
	g^o (0.1444)	aux	
	F_7 (0.0333)	V_7	
7	B♭ (0.1695)	I	
	B♭ (0.1695)	Ic	
8	B♭ (0.1695)	I	
	B♭$_7$ (0.0523)	V_7d	E♭
9	B♭$_7$ (0.0333)	V_7	
10	B♭$_7$ (0.0333)	V_7	
11	E♭$_7$ (0.1333)	ant	
	E♭ (0.1695)	I	
12	c (0.1005)	I	c
	G_7 (0.0333)	V_7c	
	c (0.1005)	Ib	
13	F (0.0801)	s	B♭
	B♭ (0.1695)	Ic	
14	d (0.1959)	IIIb	
	B♭ (0.1695)	Ic	
	d^o (0.1444)	P	
15	B♭ (0.0801)	P	
	F_7 (0.0333)	V_7	
16	B♭ (0.1695)	I	

Table 7.2: Comparison of CEG2 and SGA, both applied to the 16 bars of Schubert's Op. 33 shown in Figure 7-3.

[a]Numbers in brackets indicate the distance to nearest chord representation.
[b]b=second inversion, c=third, and d=fourth (for seventh chords).
[c]P=passing tone, aux=auxilliary chord, ant=antecedent.

Figure 7-3: Excerpt from Schubert's Op.33, with modulation boundaries as determined by the BSA described in Chapter 6.

Determining Modulations in Schubert Op. 33

Having named the chord, the next task to be performed before chord functions can be determined is to find the key(s). Applying my Boundary Search Algorithm (described in Chapter 6) to determine the points of modulation, the boundaries are given as shown in Figure 7-3.

Note that the BSA was not able to confirm the E♭ major until the 11th bar even though the A♭, a strong hint for a shift to E♭ major, appears on the third beat of the eighth bar. The texbook definition for modulation would have placed a boundary when the A♭ appeared in bar 11. I attribute this inability to detect the shift to the prominence of the dominent seventh chord in the composition. These seventh chords abound throughout the passage. In addition, as mentioned in the previous section, this passage contains frequent use of suspension chords, that create mild tension in the harmony.

One way to account for the influence of seventh chords in the model is to define keys with the seventh chords in the Spiral Array. At present, the choice of weights for chords and keys in the Spiral Array do not take into account seventh chords. The original chords defined are triads (only three pitches) and the keys are defined from these triads. The function and effect of the interval configurations differ in different contexts and for pieces written by different composers in different periods. Thus, it is not surprising that the model would need to be calibrated with different tonal structures in mind.

7.6 Brief Comments about the CEG2 Algorithm

The CEG2 algorithm has shown that the Spiral Array, by carefully positioning pitches and chords so that distances reflect their perceived relationships, is a structure that can be used to determine chord roots.

As a prototype of how the Spiral Array may be used to find chord roots in clusters of pitches, the CEG2 has been tested only on tonal pieces that utilize a set of well-defined chords. Future work could entend the CEG2 to recognizing chord roots in more complex pitch clusters.

Currently, the CEG2 algorithm requires the groupings of notes to be user-defined. Ideally, the process of determining these groupings would be automated. Suggestions for how this might be done, using a variant of the BSA, are given at the end of the previous chapter on determining modulations in Section 6.5.

8 — Conclusions

My thesis has introduced a Spiral Array that models geometrically the relations embodied in tonality. This model presents a way to re-conceptualize tonal relationships, and offers new insights to some fundamental problems in the perception and analysis of tonal music. The basic idea behind the model is the representation of a center of effect for higher level tonal elements, this center represents the composite effect of the parts.

One of the main contributions of this thesis is the bridging of two disparate disciplines, that of music theory and operations research. I utilized the mathematical modeling techniques in Operations Research to model the perceptual problem solving inherent in the comprehension of tonal music. The two disciplines are bridged by the Spiral Array, a computational geometric model that is based on the harmonic network, a framework for tonal relations.

8.1 A Parsimonious Description for Tonal Relations

This thesis proposes a new and generative way to spatially represent pitches, intervals, chords and keys. This spatial representation incorporates perceived relations among tonal elements, and is well suited to computational implementation and research in music cognition. Bamberger points out in "Developing Musical Intuition" [5] (p.147) that:

> "All notations are partial and they are so in two senses: each notation captures certain features but ignores others, and each is partial to, or favors, certain kinds of features and relations. Thus, different notations are useful for different purposes; the trick is to be able to choose effectively among multiple notations depending on what job you want the notation to do for you."

The Spiral Array representation favors fifth and third interval relations among pitches, important relations in standard music theory. It highlights these relationships by placing pitches that are related by perfect fifth or my major/minor third intervals close to each other. By carefully choosing spatial representations for pitches, chords and keys, the Spiral Array presents a parsimonious description of tonal relations.

The model has a few shortcomings, one being the lack of distinction between intervals and their inversions. Because the model uses octave equivalence, a rising fourth (for example, C *up* to F) and a descending fifth (C *down* to F) are represented in the same way. However, a rising figure creates a different listener response from a descending one, and, a perfect fourth sounds different from a perfect fifth.

As a spatial model incorporating important triadic relations, it is well suited for finding tonality and chord roots in western tonal music. By using spatial proximity to depict closeness, distances between the centers generated by different compositions or melodies can be measured and compared, thus serving as a similarity measure.

One of the model's main advantages is this ability to quantify closeness. In the analyses generated in Chapters 5, 6 and 7, each answer is accompanied by a distance measure. This measure represents the distance from the answer to the center generated by the passage; a small number means a close match, and a large number a distant one. In a way, this number is a "confusion monitor", documenting the algorithm's certainty of the answer. For example, in analyzing a chromatic passage, this distance measure would be high.

8.2 A Research and Pedagogical Tool

In the process of using the Spiral Array, one is constantly faced with questions concerning assumptions about tonality. For example, is a model based on fifths and thirds a good one? Sessions in "Harmonic Practice" [45] isolates the fifth and third from the overtone series and derives the diatonic (major) scale from chords. To what extent is this assumption that western tonal harmony is built on fifths and thirds a reasonable one?

Distance is the measurement for perceived "closeness" in the model. Relations between represented tonal entities have direct counterparts in the geometric structure. As shown in Chapter 4, these relations can be assigned to and inferred from the model. This process of assigning and inferring distance relationships can serve to inform one about important relations in tonality. For example, one could determine which interval-key relations lead to key-finding results that are consistent with human perception. Thus, the Spiral Array is an invaluable resource in analyzing theories in music and in cognitive psychology involving tonality.

The process of designing algorithms, too, can serve to inform about musical perception. For example, in trying to find the keys in Bach's *Minuet in G*, which modulates to D then returns to G, I first grouped the pitches by the bar as input for the computer program. Not surprisingly, this produced a bar-by-bar analysis, and the computer algorithm returned the roots of the chords employed. And since, for this example, the chords mostly changed bar-by-bar, this technique produced a list of chords traversed in the piece. Learning from this experiment, I changed the algorithm to search for the boundaries for a departure from and a return to an established key. A previously unsolved problem because of the computational complexity of other approaches, the Spiral Array model, by numerically quantifying "closeness" in tonality provides a viable method to search for the best choice of boundaries between changing tonalities.

This exercise showed the importance of boundaries in determining modulations. This knowledge, in turn, could provide reinforcement for the hypothesis that people listen in blocks, and that one of the primary tasks performed by the human mind when listening to a piece of music is to organize the information into coherent sections. This knowledge can inform the teaching of music appreciation.

8.3 A Framework on which to Design Efficient Algorithms

The computation approach using the Spiral Array put forth in this thesis proposes algorithms that improve upon the performance and expand the scope of its previously known applications. I have shown how this model can be used to find keys, determine chords, and search for modulations.

Previous models have been designed to find keys or determine chords, but not to do both. The Spiral Array can be used to perform both tasks, with the addition of finding modulations. This is because pitches, intervals, chords and keys are generated in the same spatial framework in the Spiral Array, thus allowing comparisons between elements from different heirarchical levels.

In Chapter 5, I showed that the key-finding algorithm (CEG) devised using the Spiral Array model surpasses previous ones in its average performance, and that its average performance is close to optimal. The problem of determining modulations is a computationally difficult one which has not been solved before. In Chapter 6, I design an algorithm (BSA), the first of its kind, using the Spiral Array model that can search from boundaries between keys. Section 6.5 suggests how this prototypical approach can be extended to address more complex examples. In Chapter 7, I propose a preliminary algorithm (CEG2) that can determine chords in a musical passage; and, Section 7.6 outlines some possible future developments.

In relation to Sections 8.1 and 8.2, being able to study the nature of these problems and their solutions is critical to the understanding of human perception and analysis of tonal music, and also to pedagogical issues pertaining to these problems. As mentioned in Chapter 6, being able to characterize the relationships that generate a tonal center is crucial to understanding and making performance decisions.

Solving these basic problems computationally is also a precursor to computer analysis and generating of western tonal music. These solutions also precede the design of automated systems that interface computer-generated music with real-time performance.

8.4 In Conclusion

Being able to characterize the relationships that generate a tonal center is crucial to the analysis of tonal music. It can potentially inform issues of varying compositional styles that can be distinguished by their rates and types of harmonic changes. Furthermore, knowing the tonality of a segment of music is essential for determining melodic and rhythmic boundaries, which in turn affect performance decisions about musical structure and phrasing.

As demonstrated in this thesis, the Spiral Array is a versatile model for tonal relations. It can be applied to solving many different computational problems in music cognition. It also serves as a tool to research tonal relations. Some future directions include the study of basic, but unanswered, questions about the relationships in tonality, including:

- What makes a melody more complex than another? Why can some melodies be harmonized in more ways than others? Is there a way to quantify its complexity?

- Do people hear a minor third interval as the lower half of a minor triad, or the upper half of a major triad? What are the pitch relations that would generate either perception?

- Why does a piece of music in the minor mode shift easily in and out of its relative major, whereas, a piece in the major mode more often stays within the major realm?

These types of inquiries can be extended to western music of other styles, to some atonal music, and to western music written by non-western composers. For example, in the latter half of this century, there has been a burgeoning of compositions by Chinese for western instruments. As composers from a different culture who have assimilated western tonal concepts and written in western idioms, what have they chosen to absorb, and how have they fashioned western tonality in the light of their own musical culture?

Finally, prior to my work using distance to summarize tonal relationships, attempts at algorithms for determining modulations have proved to be computational quagmires. This technique of using spatial distance to model preferences is a powerful one that can be generalized to many other applications involving decision-making. Perhaps one might be able to apply this knowledge and redirect the research to provide solutions for other highly structured systems of information?

Bibliography

[1] Edward Aldwell. *Harmony and Voice Leading.* Harcourt Brace Jovanovich, San Diego, CA, second edition, 1989.

[2] J. S. Bach. *371 Harmonized Chorales and 69 Chorale Melodies with figured bass.* G. Schirmer, Inc, New York, NY, 1679 edition, 1941. Revised, corrected, edited, and annotated by Albert Riemenschneider.

[3] Gerald J. Balzano. The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, 4(4):66–84, 1980.

[4] Jeanne S. Bamberger. *The Mind Behind the Musical Ear.* Harvard University Press, Cambridge, MA, 1991.

[5] Jeanne S. Bamberger. *Developing Musical Intuition.* Oxford University Press, New York, 2000.

[6] Clifton Callender. Voice-leading parsimony in the music of alexander scriabin. *Journal of Music Theory*, 42(2):219–233, 1998.

[7] Adrian Child. Moving beyond neo-riemannian triads: Exploring a transformational model for seventh chords. *Journal of Music Theory*, 42(2):181–193, 1998.

[8] David Clampitt. Alternative interpretations of some measures from parsifal. *Journal of Music Theory*, 42(2):321–334, 1998.

[9] John Clough. A rudimentary geometric model for contextual transposition and inversion. *Journal of Music Theory*, 42(2):297–306, 1998.

[10] Richard Cohn. Maximally smooth cycles, hexatonic systems, and the analysis of late-romantic triadic progressions. *Music Analysis*, 15(1):9–40, 1996.

[11] Richard Cohn. Neo-riemannian operations, parsimonious trichords, and their tonnetz representations. *Journal of Music Theory*, 41(1):1–66, 1997.

[12] Richard Cohn. Introduction to neo-riemannian theory: A survey and a historical perspective. *Journal of Music Theory*, 42(2):167–180, 1998.

[13] Richard Cohn. Square dances with cubes. *Journal of Music Theory*, 42(2):283–296, 1998.

[14] Peter Desain, Henkjan Honing, Huub Vanthienen, and Luke Windsor. Computational modeling of music cognition: Problem or solution? *Music Perception*, 16(1):151–166, 1998.

[15] Jack Douthett and Peter Steinbach. Parsimonious graphs: A study in parsimony, contextual transformations, and modes of limited transposition. *Journal of Music Theory*, 42(2):241–264, 1998.

[16] Edward Dunne and Mark McConnell. Pianos and continued fractions. *Mathematics Magazine*, 72(2):104–115, 1999.

[17] Edward Gollin. Some aspects of three-dimensional tonnetze. *Journal of Music Theory*, 42(2):195–206, 1998.

[18] Walter B. Hewlett and Eleanor Selfridge-Field, editors. *Computing in Musicology*, volume 11. CCARH, Menlo Park, CA, 1999. Special issue on similarity in music.

[19] Paul Hindemith. *The Craft of Musical Composition*, volume 1. Schott and Company, Ltd., New York, London, 1941-42.

[20] Oswald Jonas. *Introduction to the theory of Heinrich Shenker: the nature of the musical work of art.* Longman, New York, NY, 1982.

[21] Jonathan Kochavi. Some structural features of contextually-defined inversion operators. *Journal of Music Theory*, 42(2):283–296, 1998.

[22] C. L. Krumhansl. Perceived triad distance: Evidence supporting the psychological reality of neo-riemannian transformations. *Journal of Music Theory*, 42(2):265–281, 1998.

[23] C. L. Krumhansl and E. J. Kessler. Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, 89(4):334–368, 1982.

[24] Carol L. Krumhansl. *Cognitive Foundations of Musical Pitch.* Oxford University Press, New York, NY, 1990.

[25] Carol Lynne Krumhansl. *The Psychological Representation of Musical Pitch in a Tonal Context.* PhD thesis, Stanford University, Stanford, CA, 1978.

[26] Bernice Laden. *A Model of Tonality Cognition which Incorporates Pitch and Rhythm.* PhD thesis, University of Washington, Seattle, WA, 1989.

[27] Fred Lerdahl and Ray Jackendoff. *A Generative Theory of Tonal Music.* MIT Press, Cambridge, MA, 1983.

[28] David Lewin. A formal theory of generalized tonal functions. *Journal of Music Theory*, 26(1):32-60, 1982.

[29] David Lewin. *Generalized Musical Intervals and Transformations.* Yale University Press, New Haven, CT, 1987.

[30] David Lewin. Cohn functions. *Journal of Music Theory*, 40(2):181, 1996.

[31] David Lewin. Notes on the opening of the f♯ minor fugue from wtci. *Journal of Music Theory*, 42(2):235–239, 1998.

[32] H. C. Longuet-Higgins. Letter to a musical friend. *Music Rev*, 23:244, 1962.

[33] H. C. Longuet-Higgins. Second letter to a musical friend. *Music Rev*, 23:280, 1962.

[34] H. C. Longuet-Higgins. Perception of melodies. *Nature*, 263:646, 1976.

[35] H. C. Longuet-Higgins. Perception of music. *Proc. R. Soc. Lond. B.*, 205:307, 1979. Review Lecture.

[36] H. C. Longuet-Higgins. *Mental Processes.* MIT Press, Cambridge, MA, 1987.

[37] H. C. Longuet-Higgins and M. J. Steedman. On interpreting bach. *Machine Intelligence*, 6:221, 1971.

[38] H. John Maxwell. An expert system for harmonizing analysis of tonal music. In M. Balaban, K. Ebeioglu, and O. Laske, editors, *Understanding Music with AI*, pages 335–353. MIT Press, Cambridge, MA, 1992.

[39] James R. Meehan. An artificial intelligence approach to tonal music theory. *Computer Music Journal*, 4(2):60–65, 1980.

[40] Walter Piston. *Harmony.* Norton, New York, NY, fifth edition, 1987. Revised and expanded by Mark DeVoto.

[41] D. J. Povel and P. Essens. Perception of temporal patterns. *Music Perception*, 2(4):411–440, 1985.

[42] Arnold Schoenberg. *Style and Idea.* Philosophical Library, New York, NY, 1950. Several essays were originally in German and translated by Dika Newlin.

[43] Arnold Schoenberg. *Structural functions of harmony.* W. W. Norton and Company, New York, NY, revised edition, 1954.

[44] Arnold Schoenberg. *Theory of Harmony.* University of California Press, Berkeley and Los Angeles, 1978. Originally publised in 1911 by Universal Edition. This is a translation by Roy E. Carter based on the third edition.

[45] Roger Sessions. *Harmonic Practice.* Harcourt, Brace and Company, New York, NY, 1951.

[46] Roger N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function (parts i and ii). *Psychometrika*, 27(2-3):125–140, 219–246, 1962.

[47] Roger N. Shepard. Structural representations of musical pitch. In D. Deutsch, editor, *The Psychology of Music*, pages 335–353. Academic Press, New York, NY., 1982.

[48] Roger N. Shepard and Lynn A. Cooper. *Mental Images and their Transformations.* MIT Press, Cambridge, MA, 1982.

[49] Stephen Soderberg. The t-hex constellation. *Journal of Music Theory*, 42(2):207–218, 1998.

[50] Andranick S. Tanguiane. *Artificial Perception and Music Recognition.* Lecture Notes in Artificial Intelligence. Springer-Verlag, New York, NY, 1993.

[51] David Temperley. *The Perception of Harmony and Tonality: An Algorithmic Perspective*. PhD thesis, Columbia University, New York, NY, 1996.

[52] David Temperley. An algorithm for harmonic analysis. *Music Perception*, 15(1):31–68, 1997.

[53] David Temperley and Daniel Sleator. Modeling meter and harmony: A preference-rule approach. *Computer Music Journal*, 23(1):10–27, 1999.

[54] Terry Winograd. Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory*, 12(1):2–49, 1968.

[55] Iannis Xenakis. *Formalized music; thought and mathematics in composition*. Indiana University Press, Bloomington, IN, 1971.