International Series in
Operations Research & Management Science

Elaine Chew

Mathematical and Computational Modeling of Tonality

Theory and Applications

Springer

International Series in Operations Research & Management Science

Volume 204

Series Editor

Frederick S. Hillier
Stanford University, CA, USA

Special Editorial Consultant

Camille C. Price
Stephen F. Austin State University, TX, USA

This book was recommended by Dr. Price

For further volumes:
http://www.springer.com/series/6161

Elaine Chew

Mathematical and Computational Modeling of Tonality

Theory and Applications

Springer

Elaine Chew
Centre for Digital Music
Queen Mary University of London
London
UK

ISSN 0884-8289 ISSN 2214-7934 (electronic)
ISBN 978-1-4614-9474-4 ISBN 978-1-4614-9475-1 (eBook)
DOI 10.1007/978-1-4614-9475-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013953210

© Springer Science+Business Media New York 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This book is dedicated to

my father, Chew Kim Lin, who instilled in me
a love for mathematics,

George Dantzig, in memoriam, who gave me
my first glimpse into research, and

Jeanne Bamberger, who showed me I could
combine it all with music

Preface

Blending ideas from operations research, music psychology, music theory, and cognitive science, this book aims to tell a coherent story of how tonality pervades our experience, and hence our models, of music.

The story is told through the developmental stages of the Spiral Array model for tonality, a geometric model designed to incorporate and represent principles of tonal cognition, thereby lending itself to practical applications of tonal recognition, segmentation, and visualization. Mathematically speaking, the coils that make up the Spiral Array model are in effect helices, a spiral referring to a curve emanating from a central point. The use of "spiral" here is inspired by spiral staircases, intertwined spiral staircases: nested double helices within an outer spiral.

The book serves as a compilation of knowledge about the Spiral Array model and its applications, and is written for a broad audience ranging from the layperson interested in music, mathematics, and computing to the music scientist–engineer interested in computational approaches to music representation and analysis, from the music–mathematical and computational sciences student interested in learning about tonality from a formal modeling standpoint to the computer musician interested in applying these technologies in interactive composition and performance. Some chapters assume no musical or technical knowledge, and some are more musically or computationally involved.

I am extremely pleased that this book is to appear fifteen years after the eureka moment that gave rise to the Spiral Array model, and five years—and five house moves, including one cross-country and one cross-Atlantic—after the book proposal, formulated and accepted while I held the Edward, Frances, and Shirley B. Daniels Fellowship at the Radcliffe Institute for Advanced Study. The collaborators who have contributed to this volume include Alexandre R. J. François, Ching-Hua Chuan, and Yun-Ching Chen; our joint work forms the basis of the chapters on visualization, audio key finding, and pitch spelling. Alex is additionally author of the MuSA_RT Mac App, an interactive visualization software based on the Spiral Array that is a part of the supplemental material for this book.

This compendium would not have been possible without the support of Jeanne Bamberger, who has been a mentor well beyond my doctoral research, Kim Lin Chew, my father and the only person I know willing to proofread equations, and other long-suffering members of my family. I thank Jordan Smith for his last-minute voluntary proofreading, Doug Keislar for his (in)voluntary edits to the book draft, the late Lindy Hess for her generous advice on the book proposal, Matthew Amboy for speedy feedback on the book drafts, Camille Price, incoming series editor, for her steadfast encouragement over the years, and Fred Hillier, a former teacher, whose optimism and impending departure as Series Editor provided the catalyst to finally complete the book.

Last but not least, I thank the anonymous student who asked the seemingly innocuous question, "What do you mean by key?," that started this whole undertaking.

London, Singapore, Los Angeles, Boston, August 2013 Elaine Chew

Contents

Part I Introduction

Part II The Model

Part III Key Finding

Part IV Segmentation

Part V Pitch Spelling

Part VI Visualization

Part VII Extensions

Part I
Introduction

Chapter 1
Tonality

Abstract Tonality, the underlying principles of tonal music, is both an elusive and pervasive property of the music with which we are familiar. Elusive because its effects can be felt without the mind being consciously aware that it is actively construing the pitch relations that define tonality; and pervasive because it underpins most of the music that we hear. Its effects can be quickly demonstrated by the ability of the listener to sense when a piece has ended: try humming only the first three phrases of "Happy Birthday." The chapter begins by motivating the study of tonality from the practical standpoint of the listener and of the music practitioner, namely the performer and the composer. It then proceeds to describe the genesis of the project in the pianolab of MIT, when a student asked, "What do you mean by key?," reflecting on what it is that allows a listener to ascertain the most stable pitch in a sequence. In the spirit of Bugliarello's new trivium and quadrivium, in which "no domain can any longer be considered and learned in isolation," the chapter describes how the book bridges the disparate disciplines of music theory and operations research. Spanning C.P. Snow's two cultures, the book mixes mathematical formalisms with qualitative descriptions, mingling intimate and subjective case studies with impartial large-scale and quantitative testing of algorithms. The chapter proceeds to trace the development of the Spiral Array model from its inception through the applications that have followed, thereby providing a narrative of the remaining chapters of the book and the way in which they are interlinked.

Martha Argerich crashes into the rumbling arpeggios that augur the beginning of the quasi cadenza of Strauss' *Burleske*. The orchestra led by Claudio Abbado fades to silence. A pair of doubled octaves strike with a resounding crash. The audience awaits the soloist's virtuosic display in concentrated stillness. It is the Berlin Philharmonic New Year's Eve Concert in 1992.

This chapter incorporates material from the Introduction (Chapter 1) of "Towards a Mathematical Modeling of Tonality" by Elaine Chew, an MIT PhD dissertation, Cambridge, Massachusetts (2000) https://dspace.mit.edu/handle/1721.1/9139

E. Chew, *Mathematical and Computational Modeling of Tonality*,
International Series in Operations Research & Management Science 204,
DOI: 10.1007/978-1-4614-9475-1_1,

Thunderous trills announce blistering streams of arpeggios, each outdoing its precursor by further elaboration or by reaching new heights. The excitement mounts, pushing the cadenza to its apex. The peak is sustained, but not for long. The music falls from the climax; only to be hit by another surge of energy, and another drop, this time to melt away.

A lone melodic line enters, stepping inexorably downward. The soloist pauses. The audience waits, knowing that the cadenza is not done. The orchestra musicians are poised in readiness, awaiting their cue. The soloist holds on to the note, drawing out the suspense, pushing to the limit the audience's focused participation. Then, she turns to the conductor, and smiles to signal her acquiescence. Together, the orchestra and soloist gracefully fall in step to usher in the lyrical theme, and the listener lets out a sigh of contentment ... aah.

Underneath the obvious technical displays of the cadenza lies a complex network of pitch relations that gives inner logic and coherence to the music. Composers—by choosing which note to write and where (i.e. when the note is played in relation to others)—and performers—by choosing which notes or silences to emphasize and how—alike work this system to choreograph and manipulate the listener's expectations. Re-consider the previous anecdote, now embellished with a running subtext describing the pitch relations and the expectations they engender.

In the preceding passages, D minor has been established as the main key, meaning that the pitch D is the most significant pitch, and the note material of the piece, all drawn from the D minor scale, are heard in relation to this reference pitch. Martha Argerich crashes into the rumbling arpeggios that *embellish the G minor triad, a chord (iv) which is composed of the simultaneous combination of three notes based on the fourth degree ($\hat{4}$) of the D minor scale, and* augur the beginning of the quasi cadenza of Strauss' *Burleske*. The orchestra led by Claudio Abbado fades to silence. A pair of doubled octaves strike *the unison A's, the fifth degree ($\hat{5}$) of the D minor scale,* with a resounding crash, *the G and A outlining the first two chords of one of the most prototypical cadential sequence, iv-V-i, setting up the expectation that the chord based on the first degree of the scale, the tonic, also the pitch of greatest stability ($\hat{1}$), is to return.* The audience awaits the soloist's virtuosic display in concentrated stillness, *expecting the technical pyrotechnics that typically accompanies the prolongation of the V chord in the cadenza of a concerto.* It is the Berlin Philharmonic New Year's Eve Concert in 1992.

Thunderous trills announce blistering streams of arpeggioss, *anchored by re-iterations of the A octaves, re-enforcing the root of the V chord in the iv-V-i sequence*, each outdoing its precursor by further elaboration or by reaching new heights *and depths, the A octave eventually hitting the lowest note on the keyboard.* The excitement mounts, pushing the cadenza to its apex, *an emphatic A octave followed by an enigmatic half-diminished chord.* The peak is sustained, but not for long, *the half-diminished chord resolves to an equally unstable fully diminished chord.* The music falls from the climax *following the natural voice leading to the B*♭; only to be hit by another surge of energy, and another drop *returning to B*♭, this time to melt away *into a simple A major triad, A being the natural consequence of B*♭.

A lone melodic line enters, stepping inexorably downward, *starting with the now familiar pitch B♭ followed by A*. The soloist pauses *on the A, emphasizing its connection to the earlier octaves and the fact that we again hear the fifth degree of the scale*, *($\hat{5}$)* The audience waits *for the resolution back to the tonic D, not knowing when the suspense will be over, only* knowing that the cadenza is not done. The orchestra musicians are poised in readiness, awaiting their cue. The soloist holds on to the note *A, lingering on it, deliberately prolonging the wait and* drawing out the suspense, *the bar is stretched to over three times the average length of a bar in the first half of the cadenza* pushing to the limit the audience's focused participation. Then, she turns to the conductor, and smiles to signal her acquiescence. Together, the orchestra and soloist gracefully fall in step to usher in the lyrical theme, *which not only begins with the chord of D minor (i) in the key of D minor, but also with the melody note D ($\hat{1}$) increasing the satisfaction of and pleasure at the expectation fulfilled*, and the listener lets out a sigh of contentment ... aah.

The system of pitch relations that underlie this ebb and flow of expectations is called *tonality*. In the following chapters, I shall describe the theoretical underpinnings of the Spiral Array model, a spatial representation of the relations embodied in tonality, and present applications of the model to a variety of problems in automatic music analysis.

1.1 What is Tonality?

Tonality refers to the underlying principles of tonal music, and is one of the principal ways by which listeners intuit form and structure in music; it is also one of the primary means by which music evokes psychological feelings in listeners [24]. According to Bamberger [3], "tonality and its internal logic frame the coherence among pitch relations in the music with which [we] are most familiar."

The study of tonality has a long and illustrious history dating back to Rameau's 1722 "Treatise on Harmony" [39]. Music theorists, composers, mathematicians, philosophers, and psychologists have sought to uncover the nature of, and formalize the concept of, tonality from a variety of disciplinary perspectives. Monographs written on the subject that are most directly related to this book, and that have influenced its development, include: Krumhansl's "Cognitive Foundations of Musical Pitch" [26], Lewin's "Generalized Musical Intervals and Transformations" [31], and Temperley's "The Cognition of Basic Musical Structures" [47]. Another important contribution that is closely related to this book is Lerdahl's "Tonal Pitch Space" [30]. Scholars who have specifically proposed mathematical formulations for aspects of tonality include Mazzola [34] and Tymoczko [49]. Well-known composers who have added their theories and thoughts include Hindemith [37] and Schoenberg [41].

Dahlhaus [17], in his "Studies in the Origin of Harmonic Tonality," wrote, "In common usage the term [tonality] 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." He further quotes Fétis' 1844 definition of tonality as including the "necessary successive or simultaneous relationships between the notes of a scale." It is worth explicitly noting that tonality impacts relationships between not only simultaneous and successive pitches, but also between pitches or sets of pitches across small and large expanses of time. Indeed, time is a critical element in the experience of tonality, as music itself unfolds and is heard in time.

The term tonality is sometimes synonymized with key, which in turn is often further said to refer to adherence to the pitch set of a major or minor scale. Dahlaus [17] offers a broader definition of tonality:

> ... 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.

Extending beyond the definition of tonality as an expanded key, Dahlhaus states, "Tonality [is] the underlying element of a tonal structure, the effective principle at its heart."

1.2 Elusive Yet Pervasive

Tonality is both an elusive and pervasive property of music.

Elusive because its effects can be felt without the mind being consciously aware that it is construing the relationships amongst the pitches, without the listener deliberately deciphering what is the tonic now, and how are the other tones related to the tonic. Less visceral than the sensing of time structures such as pulse, rhythm, and meter, the understanding of scale degrees (a basic part of tonality) has been noted by Huron [24] to be a cognitive rather than a perceptual phenomenon, as it is an act of the mind in "interpret[ing] physically sounding tones, rather than how the tones are in the world." Equating this unconscious cognitive activity with arithmetic, Leibniz [40] said in 1712, "The pleasure we obtain from music comes from counting, but counting unconsciously. Music is nothing but unconscious arithmetic."

Elusive because while it is a principal music feature that evokes senses both intellectual and hedonistic, the feelings it engenders and how they arise are hard to pin down and to describe in words. In fact, one may argue that time structures such as rhythm and meter give rise to much more visceral sensations that lend themselves more readily to description and communication.

Pervasive because it underpins most of the music that we hear. Listeners all know it, a fact that can be verified with a simple experiment: hum "Happy Birthday" and stop 3/4 of the way through the song—i.e., hum the notes corresponding to the words "happy birthday to you, happy birthday to you, happy birthday to Lisa." The sense of incompleteness is palpable. Stronger than the unfinished prose of the unsung text, or the asymmetry of the phrases, is the unfulfilled longing to hear the tonic on a strong beat at the end of the song, of any song. The unresolved expectation leaves a void in the gut and a strong urge to hum the next phrase. But not just the next phrase in the well known song; any phrase that ends with the tonic on a strong beat can finish the song, some better than others. In fact, any phrase that ends with the tonic can also finish it on the third phrase, the satisfaction of a stable ending overriding the asymmetry of the phrases.

Nearly all music is tonal. One exception to tonality's pervasiveness would be music that is unpitched, for example, music that is based entirely on noise or purely on rhythmic patterns, in which case other organizing principles replace that of tonality. It is difficult to contrive music that is entirely devoid of pitch, for even percussive sounds are often inherently pitched, such as the high and low pitches of palms coming together in Steve Reich's *Clapping Music*.

Because tonality pervades our experience of music, some composers have viewed it as oppressive, and espoused what is sometimes call atonal music, music that deliberately avoids creating the sense of a tonal context. Schoenberg [42], the inventor of twelve-tone music—compositions based on an impartial ordering of the twelve pitches of the chromatic scale in order to avoid the trappings of traditional tonal music—argues eloquently against the existence of atonal music:

> Permit me to point out that I regard the expression atonal as meaningless, and shall quote from what I have already expounded in detail in my *Harmonielehre*. 'Atonal can only signify something that does not correspond to the nature of tone.' And further: 'A piece of music will necessarily always be tonal in so far as a relation exists from tone to tone, whereby tones, placed next to or above one another, result in a perceptible succession. The tonality might then be neither felt nor possible of proof, these relations might be obscure and difficult to comprehend, yes, even incomprehensible. But to call any relation of tones atonal is as little justified as to designate a relation of colors spectral or complementary. Such an antithesis does not exist.'

A major obstacle to creating atonal music is that listeners cannot help but generate mental associations amongst pitches, and in particular, "relationships between pitches having a "tonic" or central pitch as its most important element." Even when confronted with a random sequence of pitches, the mind will construct a tonal context. For example, based on the pitches that have come before, a listener might imagine two consecutive pitches to be the subdominant ($\hat{4}$) and leading-tone ($\hat{7}$) implying a tonic ($\hat{1}$).

Just as tonality can refer to other kinds of tonality in western music, like Hindemith's non-diatonic tonal system, tonality can also refer to similar systems in music of non-western cultures, where the specifics of the inter-pitch associations may differ from that of western classical music, but nevertheless generate interpretations of varying stability amongst the pitches. The thaat system in North Indian classical music is a case in point, where a tonal hierarchy emerges over the course of a raag, with the *vadi* as the most stressed tone, and *samvadi* (typically the fourth or fifth degree of the scale) the second most stressed tone, see [12].

1.3 Inception

This project began as an attempt to formally describe the generating of a tonal center. The quest led to a mathematical model, on which this book is based, that incorporates music knowledge and that can be used to mimic the human decision-making process in apprehending tonality. Nevertheless, the underlying questions that provided the impulse for the extended project remain as important motivating issues.

In my first semester as a pianolab instructor—pianolab was a keyboard skills class for students enrolled in Music Fundamentals and Composition courses—at MIT in 1998, I encountered a few students who had no prior musical training. I asked one such student, after he carefully traced out the melody of "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 at the beginning of the piece and accidentals 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 human cognition of key is an aural experience, and for musicians, a physical one as well. Accidentals, although helpful, are not the primary cues for the key of the passage. One could easily find an example with no accidentals that does not generate the feel of C major or A minor, which are the two keys with no sharps or flats.

For example, in Fig. 1.1, a segment from the 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 (or A minor) as suggested by the key signature and the absence of accidentals. The mind does not need many notes to judge the key of a melody. In the mind's ear, so to speak, the first four notes of "Nobody Knows" already suggest strongly F major. I shall shortly 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 who plays an instrument, each key has its own distinctive terrain on their 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 to get a sense of the tonality is to play a short excerpt of the piece, and based on the way the piece feels in the hands, determine the key; not only the key, the tonal center, the tonic, but also how the various notes they are playing or are about to play relate to that tonic. This approach relies on the physical experience of

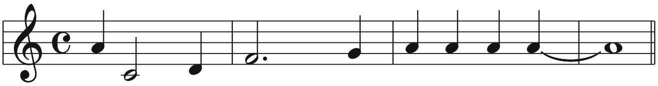

Fig. 1.1 Excerpt from "Nobody Knows the Trouble I've Seen." A musical example with no accidentals, that is actually in the key of F

making music, drawing upon embodied knowledge through the physicality of playing an instrument. But what of the aural experience? A listener who is not playing an instrument, or who does not have the experience of playing an instrument, too, can sense tonality.

I jumped at the next idea that came to mind. I hummed the 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. 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? A closer examination of "Nobody Knows" might shed some light on this matter.

The melody "Nobody Knows the Trouble I've Seen" serves to 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' set up a most stable pitch, and already give 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 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 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 traditional Scottish folksong "My Bonnie Lies Over The Ocean" and Chopin's Nocturne Op. 9 No. 2 in E♭, see Fig. 1.2. In both cases, the two pitches separated by an interval of a major sixth surround and point to the pitch that is a major third interval below the upper pitch as the stable pitch, the likely tonic.

Thus, the mind immediately determines that the first two notes in "Nobody Knows" very likely can be assigned the (movable *do*) solfège syllables *mi* and *sol*, or scale 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 Fig. 1.3. Together, the first, second and fourth notes outline the F major triad, implying an affinity to F major.

In "Nobody Knows," the rhythm in the melody also reinforces the tonic implied by the interval relations. The note of longest duration in the first half of Fig. 1.1 is the fourth note in the melody, and its pitch is F. In addition, the indicated meter places a

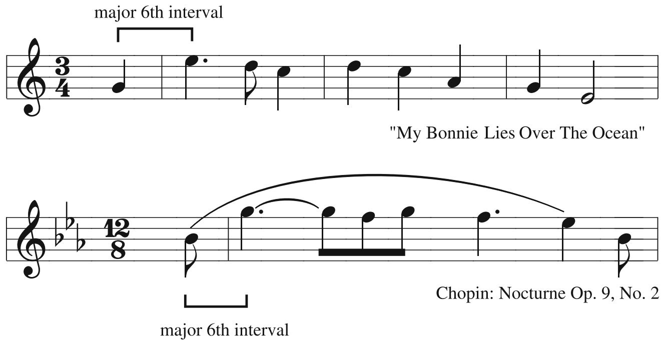

Fig. 1.2 Excerpts from "My Bonnie Lies Over the Ocean" and Chopin's Nocturne Op. 9 No. 2, both beginning with a rising major sixth interval

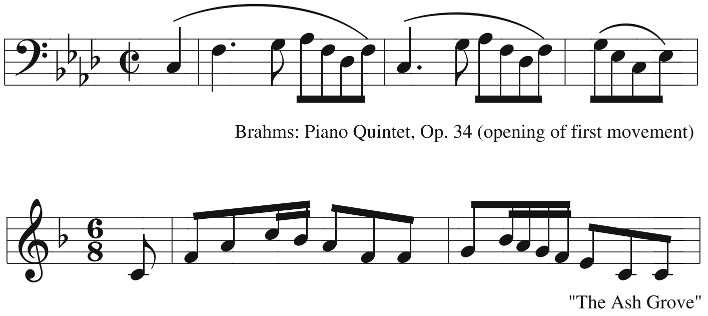

Fig. 1.3 Excerpts from Brahms' Piano Quintet and "The Ash Grove," both beginning with a rising perfect fourth interval

downbeat on this F. Although the last note in Fig. 1.1 has a longer duration than the F, its onset begins on a weak beat (the fourth beat of the third bar).

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 determines the key? Thus one student's seemingly innocent question of how does one find the key led to my quest for a concise and effective model for the generating of tonal centers.

1.4 Design Principles

In the following chapters, I shall describe a model for tonality called the Spiral Array. It behoves me to mention here some of the characteristics and design principles of the model.

The model is generative, and designed to provide a concise description of the characteristics of tonality. The hierarchical structure of tonality is embedded in the model: key representations are generated from chords, and chord representations from component pitches. The model is related to Riemann's theory of tonality [17] in that significant tonal relationships are established by means of the chord functions of the tonic, the dominant and the subdominant. The model can also be used to generate note and chord sequences, although that aspect is not a focus of this book.

Distinct from other geometric and network models for tonal relations, the Spiral Array represents pitches, intervals, chords and keys in the same spatial framework. Derived from the *tonnetz* (tone network), first attributed to Euler [16] and used extensively in neo-Riemannian theory (for example [6, 11, 14–16, 20, 22, 31]), the model fuses both the original network model with properties of continuous space. In this space, any collection of pitches can generate a center of effect, that is essentially a mathematical sum of its parts. Thus, while hierarchical structure is incorporated into the model, at the same time, the model flattens this hierarchy by representing elements from all levels in the same space.

Distance in the model corresponds to perceived closeness in tonal music. Two pitches may be physically close on the piano keyboard, for example F and G♭, but perceptually quite distant. Whereas, two pitches that are farther apart on the piano, for example F and C, are perceived to be closely related. This has something to do with the simple fractional relationships between the frequencies of the pitches, but suffice it to say that the model aims to encapsulate the auditory sense of closeness rather than the configuration of distance on an instrument.

Because of the egalitarian representation of elements of all hierarchical levels in the same space, literal (and conceptual) distance can be measured between any two elements, from any hierarchical level, in the Spiral Array. The model thus offers a way to re-conceptualize tonal relationships. Because keys are represented in the model, the model also offers a way to envision the generating of tonal centers.

A computational model, the Spiral Array raises pertinent questions regarding, and produces insights that illuminate, some basic issues in traditional music theory. I shall demonstrate the versatility of the model by applying it to a number of fundamental problems in the cognition and analysis of western tonal music: that of finding keys (from MIDI, event-based, and digital audio information), pitch spelling (ascribing tonally consistent letter names and accidentals to pitches), and determining key boundaries (searching for modulations, shifts in tonality). The majority of these algorithms are designed for real-time processing. Furthermore, I present the model as a visualization tool, aided by the real-time aspect of the methods as well as the three-dimensional geometry of the Spiral Array space.

Being able to study the nature of these fundamental problems and their solutions is critical to the understanding of human cognition and analysis of tonal music, and also to pedagogical issues pertaining to these problems. Solving these basic tonal analysis problems computationally is a precursor to any computer analysis of western tonal music, and automated systems that interface computer-generated music with real-time performance. Since the model was invented, the field of music information retrieval has burgeoned, fueled by the proliferation of digital music and music streamed over the Internet. Thus, mathematical representations that lend themselves readily to efficient computational implementation have become ever more important for retrieval tasks such as music similarity assessment, segmentation and summarization.

Being able to characterize the relationships that generate a tonal center is crucial to the understanding and the making of composition as well as performance decisions. In composition and improvisation, it affects the choice of notes and note sequences; in performance it impacts the choice of prosody in expressive musical communication.

Knowledge of tonality provides information as to which notes or chords are structurally more stable, and will give the sense of solidity when further emphasized, or are not as stable, and in need of greater stress (or de-emphasis of competing notes) in order to afford the same presence. Knowing the tonality also provides a map of the listener's tonal expectations, revealing which notes can be prolonged (or elaborated upon) to delay the onset of a known and inevitable outcome (typically of closure); which notes can be emphasized to underscore the thwarting of the listener's expectations; and, which notes can be glossed over because they concur with the listener's expectations, or because the primary functions which they support are undertaken by other structurally more significant notes.

1.5 Straddling Cultures

One of the main contributions of this book is the bridging of two disparate disciplines, namely that of music theory and operations research. Operations research is the science of decision-making using mathematical models which integrate the operating criteria of the system in question. According to George Dantzig[1], 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 cognition and analysis of tonal music compositions. As quoted earlier, Leibniz (1646–1716), the German philosopher, physicist, and mathematician, has gone so far as to say: *"Music is nothing but unconscious arithmetic."* 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. While a subset of music theorists have invoked in their scholarly work mathematical techniques, typically from the more abstract branch of pure mathematics, the methods laid out in this book have a practical engineering flavor characteristic of operations research approaches to problem solving, see [36]. An early precursor of this kind of engineering

---

[1] Personal communication, 7 November 1999

approach to fundamental musical practice is Xenakis' thesis [51] on stochastic and other mathematical applications of mathematics to music composition. While core signal processing techniques applied to music audio employ selected optimization algorithms, they rarely accommodate more than a cursory nod to music theoretic knowledge. The model at the heart of this book owes its genesis to music theoretic models, and its development to operations research techniques. This interdisciplinary effort forms a core contribution of this book.

This book project itself has been an exercise in bridging C.P. Snow's two cultures [44]—on the one side the arts and humanities, and the other science—in open acknowledgement of the deep connections that cut across disciplines. A major synthesis of over 10 years' work, this book blends carefully curated selections from earlier writings with newer insights. The styles have been tailored to a mixed audience. Mathematical formalisms commingle with qualitative descriptions. Writing in the first vs. third person varies across some chapters. Up close and subjective case studies exist alongside impartial large-scale and quantitative testing of algorithms. Sometimes, a compromise is struck between humanities and scientific practices. The divide that continues to exist between the two was an important reason for the extended interval between the publication of my dissertation in 2000 and this current book; while book publishing is considered de rigueur in the humanities, a non-peer reviewed publication is of considerably lower value in the sciences. Nevertheless, it is with great pleasure that I find myself in the position to work on this book at this stage.

As Bugliarello [5] wrote in his treatise on a new trivium and quadrivium, "no domain can any longer be considered and learned in isolation." Computational music analysis is an excellent example of one such blended domain of study; it is, by definition, an interdisciplinary study linking human perception and cognition, mathematical and computational modeling, 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. [18] documents several successful attempts to bridge pairs of these disciplines in the past decades. I have summarized in Fig. 1.4 a small sample of some interdisciplinary research in computational music analysis as an example. The development and construction of my model draws upon all three disciplines.

Music is not an easy domain within which to design effective computational models that describe changing tonalities and harmonies. The human cognition of tonality utilizes both top-down and bottom-up analyses [18]. When assessing tonal (and rhythmic) structure, the human mind contemporaneously considers several different structural levels in the music. The mental ability to simultaneously scale up and down allows the listener to gather information at all levels.

Furthermore, music can inherently be described structurally in multiple and equally valid ways [2, 3, 29]. For example, Bamberger [2] 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 [29] demonstration of the fundamental difference between the musical elements of grouping and meter.

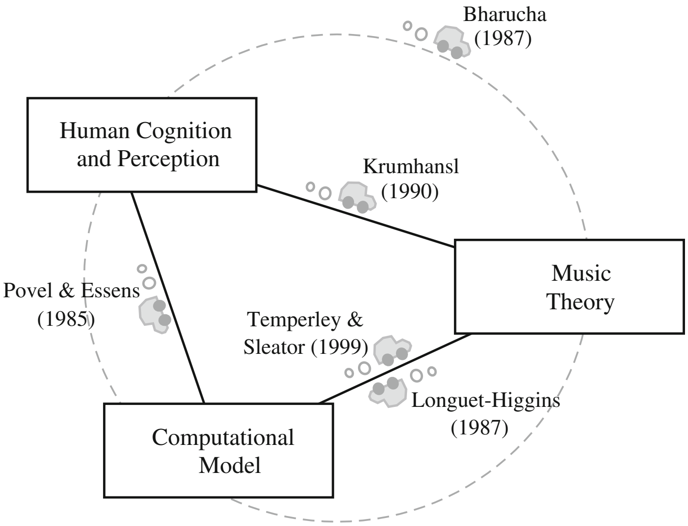

Fig. 1.4 Bridging the disciplines: a sampling of some interdisciplinary research in computational music analysis

At the time of the Spiral Array's invention, several theses had been devoted to the modeling of tonal perception, including Krumhansl's 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" [27] using connectionist methods, and Temperley's thesis on "The Perception of Harmony and Tonality: An Algorithmic Perspective" [45] grounded in cognitive science. Soon after the publication of my dissertation, a special issue on tonality induction edited by Leman and Vos appeared in Music Perception [28]. Dissertations that have followed since include Aarden's "Dynamic Melodic Expectancy" [1], Honingh's "The Origin and Well-Formedness of Tonal Pitch Structures" [21], and Milne's "A Computational Model of the Cognition of Tonality" [35]. Where relevant, newer publications have been incorporated into the literature review of various chapters.

1.6 Overview

The contents of this book represent the culmination of research amassed over fifteen years. The first six chapters, including parts of the present one, are based on material from my dissertation [7]. The next five chapters are based on extensions of the model to post-tonal music, to non-score based digital music information, and data visualization, and derive from a handful of articles selected to illustrate the model's core applications.

Chapter 2 provides some relevant background and survey of models and methodologies that have influenced the design of the Spiral Array and its computational approach. The chapter begins with an intuitive and illustrated overview of the Spiral Array model and the Center of Effect Generator (CEG) key-finding algorithm as a basis for comparison. I review some spatial models of musical pitch that have impacted the model's configuration of pitch representations; next, I describe von Neumann's Center of Gravity algorithm and Dantzig, G. B.Dantzig's bracketing technique that inspired the CEG method. No attempt is made to provide a comprehensive summary of the literature.

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. Later, in Appendix A, I present an example of how the model parameters can be calibrated so that the model represents the cognition of inter-pitch, inter-chord, pitch-chord, and pitch-key distances. The model is sufficiently well defined in Chap. 3 that the calibration details can be skipped without detriment to understanding of the model's applications in later chapters.

Chapter 4 introduces the first computational application that uses the Spiral Array: the problem of key-finding in melodies. I formally propose the CEG key-finding algorithm, explaining how it works by applying it to an example, "Simple Gifts." Step-by-step scientific visualizations accompany the ranked key outputs and distances. I compare the CEG algorithm to those by two other researchers in Chap. 5, giving detailed analyses of the comparison results when applied to the 24 fugue subjects in Book 1 of Bach's *Well-Tempered Clavier*. I show that the CEG algorithm surpasses previous ones in its average performance, and is close to optimal. The MATLAB code for the CEG algorithm, including programs to generate the Spiral Array model, and detailed key and distance output for each fugue subject can be found in Appendix B.

In Chap. 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 March in D, both from his *"A Little Notebook for Anna Magdalena"*. The conclusion suggests more sophisticated variations on this basic algorithm. Chap. 7 presents another segmentation algorithm based on concepts borrowed from Control Theory; the method is named Argus, for the all-seeing giant of Greek mythology, because it scans backward and forward in time to determine if the degree of change has exceeded a threshold, giving rise to a boundary.

In order to apply the tonal analysis algorithms to music data such as MIDI or audio, one must first turn numeric pitch information to tonally consistent letter names, the subject of the next chapter. The key determines pitch spelling, and the spelling of the pitches reveals the key; a chicken and egg problem. Chapter 8 describes a bootstrapping algorithm based on the Spiral Array, work done with Yun-Ching Chen, that presents a solution to this conundrum.

Chapter 9 presents MuSA.RT, "Music on the Spiral Array. Real-Time," an interactive tonal analysis and visualization system based on the Spiral Array model, and incorporating the pitch-spelling and CEG key-finding algorithms, as well as an algorithm for chord recognition similar to that for key-finding. This work was conducted in collaboration with Alexandre R. J. François, whose software architecture style enabled the concurrent processing of multiple data streams—music, video, camera control, and computational algorithms. François has since created a MuSA_RT App [19]—freely available from the Mac App Store—that is the focus of supplemental material to this volume at http://musa-rt.blogspot.com. Inspired by David Huron's study on music-engendered laughter [23], Chap. 10 uses MuSA.RT to conduct an analysis of tonally grounded devices employed by "P. D. Q. Bach" (Peter Schickele) to generate musical humor.

While all of the above techniques have been presented and illustrated using score-based and MIDI information, each can be extended and applied to music audio signals. A score-based tonal analysis algorithm can be extended to MIDI by applying pitch spelling to the note numbers; and a MIDI-based algorithm can be expanded to music audio signals by converting the signals to spectral or pitch-class data. Chapter 11 applies the CEG algorithm to audio key-finding, describing aspects of the system design—including signal processing techniques for transforming audio signals to pitch-class numbers, adaptations of the Spiral Array weights to audio data, and different key determination policies—and presenting an analysis of the system's sensitivity to these model choices.

There is some degree of repetition, especially with regard to summary descriptions of the Spiral Array model in Chap. 6 through 11. I have deliberately left these sections intact so as to improve readability of these later chapters as standalone sources. The general reader, who may not be familiar with some of the musical terms, may find the glossary at the end of the book helpful.

For focus and unity of the volume, this book does not cover a number of applications of the Spiral Array algorithms. The Spiral Array is compared to Lerdahl's Tonal Pitch Space and Krumhansl's and Krumhansl and Kessler's spatial representations of pitch class and key relations in [8]. The model is used to reveal symmetries in Webern's "Sehr Schnell" [10], and pedaling strategies that reinforce tonality [9]. Key and chord distributions and sequences determined with center of effect algorithms can be used to identify musical variations [33, 50] in music information retrieval. Key information forms a first step to automatic learning and generating of melody harmonizations [13]. Tonal novelty quantified using the Spiral Array is one parameter explored in the explaining of the cognition of boundaries in music [43].

Finally, while this volume synthesizes research on the Spiral Array and its applications to date, I view this volume not as the culmination of a decade and a half of work on the Spiral Array, but as an introduction to the model and the beginning of the breadth of computational possibilities offered by the model. The model itself can be further optimized for specific applications, the parameters themselves can be learned from data, they can also be time varying to track our changing cognition of musical structures. Chapter 11 hints at this: the policies for deciding key are subject to debate. As in Temperley's Bayesian reinterpretation of Krumhansl and Schmuckler's probe tone profile method for key-finding [48], the Spiral Array lends itself to extension via spatial probability models.

References

1. Aarden, B.: Dynamic melodic expectancy. Ph.D. Dissertation, Ohio State University, Columbus (2003)
2. Bamberger, J.S.: The Mind Behind the Musical Ear. Harvard University Press, Cambridge, MA (1991)
3. Bamberger, J.S.: Developing Musical Intuition. Oxford University Press, New York (2000)
4. Bharucha, J.: Music cognition and perceptual facilitation: a connectionist framework. Music Percept. **5**(1), 1–30 (1987)
5. Bugliarello, G.: A new trivium and quadrivium. Bull. Sci. Tech. Soc. **23**(2), 106–113 (2003)
6. Callender, C.: Voice-leading parsimony in the music of Alexander Scriabin. J. Music Theory **42**, 219–233 (1998)
7. Chew, E.: Towards a Mathematical Model of Tonality. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA (2000)
8. Chew, E.: Out of the grid and into the spiral: geometric interpretations of and comparisons with the Spiral-Array model. Comput. Musicol. **15**, 51–72 (2007)
9. Chew, E., François, A. R. J.: MuSA.RT and the pedal: the role of the sustain pedal in clarifying tonal structure. In: Proceedings of the 10h International Conference on Music Perception and Cognition, Sapporo, Japan (2008)
10. Chew, E.: Pitch symmetry and invariants in Webern's Sehr Schnell from variations Op.27. In: Klouche, T., Noll, T. (eds.) Mathematics and Computation in Music, CCIS, vol. 37, pp. 240–246. (2009)
11. Childs, A.: Moving beyond Neo-Riemannian triads: exploring a transformational model for seventh chords. J. Music Theory **42**, 181–193 (1998)
12. Chordia, P., Rae, A.: Raag recognition using pitch-class and pitch-class dyad distributions. In: Proceedings of the International Conference on Music Information Retrieval, Vienna (2007)
13. Chuan, C.-H., Chew, E.: Generating and evaluating musical harmonizations that emulate style. Comp. Music J. **35**(4), 64–82 (2011)
14. Cohn, R.: Maximally smooth cycles, hexatonic systems, and the analysis of late-romantic triadic progressions. Mus. Anal. **15**, 9–40 (1996)
15. Cohn, R.: Neo-Riemannian operations, parsimonious trichords, and their tonnetz representations. J. Mus. Theory **41**, 1–66 (1997)
16. Cohn, R.: Introduction to Neo-Riemannian theory: a survey and historical perspective. J. Mus. Theory **42**(2), 167–180 (1998)
17. Dahlhaus, C.: Studies in the Origin of Harmonic Tonality. Robert O. Gjerdingen (Trans.) Princeton University Press, Princeton, New Jersey (1990)
18. Desain, P., Honing, H., Vanthienen, H., Windsor, L.: Computational modeling of music cognition: problem or solution? Music Percept. **16**(1), 151–166 (1998)
19. François, A.R.J.: MuSA_RT (2012) https://itunes.apple.com/ca/app/musa-rt/id506866959?mt=12. Accessed 30 Aug 2013
20. Gollin, E.: Some aspects of three-dimensional tonnetze. J. Mus. Theory **42**, 195–206 (1998)
21. Honingh, A.: The origin and well-formedness of tonal pitch structures. Ph.D. Dissertation, University of Amsterdam, Amsterdam (2006)
22. Hook, J.: Uniform triadic transformations. Ph.D. Dissertation, Indiana University, Bloomington (2002)
23. Huron, D.: Music-engendered laughter: an analysis of humor devices in PDQ bach. In: Lipscomb, S., Ashley, R., Gjerdingen, R.O., Webster, P. (eds.) Proceedings of the 8th International Conference on Music Perception and Cognition, pp. 700–704 (2004)
24. Huron, D.: Sweet Anticipation: Music and the Psychology of Expectation. MIT Press, Cambridge (2006)
25. Krumhansl, C.L.: The psychological representation of musical pitch in a tonal context. Ph.D. Dissertation, Stanford University, Stanford, CA (1978)
26. Krumhansl, C.L.: Cognitive Foundations of Musical Pitch. Oxford University Press, New York (1990)

27. Laden, B.: A model of tonality cognition which incorporates pitch and rhythm. Ph.D. Dissertation, University of Washington, Seattle, WA (1989)
28. Leman, M., Vos, P. (eds.): Special issue on tonality induction. Mus. Perc. **17**(4) (2000)
29. Lerdahl, F., Jackendoff, R.: A Generative Theory of Tonal Music. MIT Press, Cambridge, MA (1983)
30. Lerdahl, F.: Tonal Pitch Space. Oxford University Press, New York (2001)
31. Lewin, D.: Generalized Musical Intervals and Transformations. Yale University Press, New York (1987)
32. Longuet-Higgins, H.C.: Mental Processes. MIT Press, Cambridge, MA (1987)
33. Mardirossian, A., Chew, E.: Music summarization via key distributions: analyses of similarity assessment across variations. In: Proceedings of the International Conference on Music Information Retrieval (2006)
34. Mazzola, G.: The Topos of Music. Birkhäuser, Basel (2002)
35. Milne, A.: A computational model of the cognition of tonality. Ph.D. Dissertation, Open University, Milton Keynes (2013)
36. Morse, P.M.: Mathematical problems in operations research. Bull. Amer. Math. Soc. **54**(7), 602–621 (1948)
37. Ortmann, O.: An analysis of Paul Hindemith's "Unterweisung im Tonsatz". Bull. Am. Mus. Soc. **4**, 26–28 (1940)
38. Povel, D.J., Essens, P.: Perception of temporal patterns. Mus. Perc. **2**(4), 411–440 (1985)
39. Rameau, J.-P.: Treatise on Harmony. Philip Gossett (Trans.) Dover Publications, New York (1722)
40. Sacks, O.: The Man who Mistook his Wife for a Hat. Touchstone, New York (1985)
41. Schoenberg, A.: Structural Functions of Harmony. Norton & Co. Inc., New York (1969)
42. Schoenberg, A.: Style and Idea: Selected Writings of Arnold Schoenberg. Leonard Stein (ed.), Leo Black (Trans.) St. Martins Press, New York (1975)
43. Smith, J.B.L., Chuan, C.-H., Chew, E.: Audio Properties of Perceived Boundaries in Music. In: Li, T., Ogihara, M., Tzanetakis, G. (eds.) Special issue on music data mining, IEEE Transactions on Multimedia (2014)
44. Snow, C.P.: The Two Cultures—C. P. Snow's epochal essay published online for the first time. New Statesman (2013) Available via http://www.newstatesman.com/cultural-capital/2013/01/c-p-snow-two-cultures. Accessed 12 Aug 2013
45. Temperley, D.: The perception of harmony and tonality: an algorithmic perspective. Ph.D. Dissertation, Columbia University, New York (1996)
46. Temperley, D., Sleator, D.: Modeling meter and harmony: a preference-rule approach. Comp. Mus. J. **23**(1), 10–27 (1999)
47. Temperley, D.: The Cognition of Basic Musical Structures. MIT Press, Cambridge (2001)
48. Temperley, D.: Music and Probability. MIT Press, Cambridge (2010)
49. Tymoczko, D.: A Geometry of Music-Harmony and Counterpoint in the Extended Common Practice. Oxford University Press, New York (2011)
50. Unal, E., Chew, E., Georgiou, P., Narayanan, S.: A preplexity based cover song matching system for short length queries. In: Proceedings of the International Conference on Music Information Retrieval (2011)
51. Xenakis, I.: Formalized Music: Thought and Mathematics in Composition. Indiana University Press, Bloomington (1971)