SPATIALIZATION IN MUSIC: THE ANALYSIS AND INTERPRETATION OF SPATIAL GESTURES

by

JASON WYATT SOLOMON

(Under the Direction of Adrian P. Childs)

ABSTRACT

With so much attention given to pitch space in contemporary analytic discourse, theorists have largely ignored the important relationships unfolding within the physical space of musical performance. I propose an analytic methodology for identifying, classifying, and interpreting "spatial gestures." A spatial gesture emerges from the consecutive activity of multiple performers within an ensemble. Various gestures are differentiated by the specific orderings, in time, of sonic events occurring at separate points in ensemble space. Employing integer notation derived from contour theory, gestures may be assigned SG-labels. These labels assist in identifying structural relations between different gestures and in assigning a gesture to one of several "gestural categories." The classification of spatial gestures is based on theories of form perception taken from Gestalt psychology. As a unified whole comprising ordered parts (gestalt structure), a spatial gesture is directed motion through an ensemble that often serves a motivic function: a gesture may be developed and transformed, and the profusion of related gestures imparts spatial coherence to a work of music. Furthermore, the specific directionality and kinetic shape of a spatial gesture is rich in interpretive potential. Image-schema theory and the theory of conceptual metaphor are evoked to construct a hermeneutic account of spatio-gestural activity.

Many composers—particularly those active in the twentieth century—deliberately spatialize their musical works. However, irrespective of a composer's intentionality (documented or surmised), spatial gestures are often readily perceived during performance, and a spatial analysis can unveil the gestures' full capacity for structural unification and the conveyance of meaning. Gestural motions and shapes vary depending upon the performing ensemble's onstage configuration. Since ensemble seating plans are variable and often inconsistent, spatial analysis holds significance for performance practice: the manner in which an ensemble elects to arrange itself (if left unspecified by the composer) becomes a matter of interpretation. Ensemble members may organize themselves in an effort to enhance the perceptual salience of spatial gestures and enliven the spatial construct of a musical performance.

 INDEX WORDS: Spatialization, Spatial gesture, Spatial analysis, Space, Sound, Event, Motion, Movement, Shape, Form, Pan, Panning, Pitch space, Contour, Musical space, Musical gesture, Gesture, Gestural, Polystreaming, Transformation, Gestalt, Prägnanz, Continuity, Continuation, Closure, Similarity, Proximity, Symbol, Semiotics, Categorization, Image schema, Force, Path, Dialogue, Cycle, Container, Balance, Symmetry, Tension, Metaphor, Cross-domain mapping, Psalmody, Polychoral, Antiphony, Antiphonal, Cori spezzati, String quartet, Guitar quartet, Transcription, Arrangement, Electronic music, Auditory stream, Auditory scene analysis, Localization, Acoustics, Op. 131, Beethoven, Ives, Stockhausen, Xenakis, Brant, Chowning, Ginastera

SPATIALIZATION IN MUSIC: THE ANALYSIS AND INTERPRETATION OF SPATIAL GESTURES

by

JASON WYATT SOLOMON

B.Mus., Music Composition and Music Performance, The University of Georgia, 2000M.M., Music Performance, The University of Georgia, 2002

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2007

© 2007

Jason Wyatt Solomon

All Rights Reserved

SPATIALIZATION IN MUSIC: THE ANALYSIS AND INTERPRETATION OF SPATIAL GESTURES

by

JASON WYATT SOLOMON

Major Professor:

Adrian P. Childs

Committee:

John Turci-Escobar Leonard V. Ball, Jr. David Haas Robert Burton

Electronic Version Approved:

Maureen Grasso Dean of the Graduate School The University of Georgia May 2007

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr. Adrian P. Childs for his critical guidance during the writing of this dissertation. Serving as Chair of my doctoral committee, and consistently going above and beyond his prescribed duties as major professor, Dr. Childs sacrificed countless hours reading through various chapter drafts and providing invaluable feedback and commentary. His creative insights and thoughtful suggestions for revision helped dramatically to mold this dissertation into its present form.

I also extend my gratitude to the other members of my doctoral committee, all of whom provided constructive assistance during the conception, research, and writing of this project. Dr. John Turci-Escobar provided great moral support and encouragement, and his extensive knowledge of numerous strands of music theory proved greatly beneficial to me during the dissertating process. As I was completing my doctoral coursework, he and Dr. Childs together created several directed independent study courses that allowed me to engage various fields of music theory under their expert tutelage. Input from Dr. Leonard Ball guided my comparative approach to spatialization in acoustic and electronic music. In fact, my interest in spatialization surged during an electronic music course that I took with Dr. Ball in 2001. My research paper for that course was entitled "An Overview of the Concepts and Development of Spatialization in Music." This dissertation might be viewed as the inflated apogee (hopefully the first of several to come) of a line of research born in that ten-page paper. One of Dr. Ball's compositions (*Wheels*, for guitar quartet) provides several examples of spatial gestures in the third chapter of this dissertation. Dr. David Haas supplied important input from the musicological side of the

iv

fence and provided indispensable writing tips. Finally, Dr. Robert Burton was gracious enough to represent my dual minor area of philosophy and psychology. Dr. Burton's vast knowledge of all things philosophical (and many things musical) ensured that I did not misappropriate concepts from philosophy while applying them to musical discourse. I was fortunate enough to have the opportunity to take an independent study with Dr. Burton, and this provided the means to explore thoroughly many of the philosophical texts drawn from in this dissertation.

Although not an official member of my committee, Dr. David Schiller helped me greatly in completing my Ph.D. at the University of Georgia. I am most grateful for his time and able assistance.

I was fortunate to receive the Graduate School Dissertation Completion Award for the 2006–2007 academic year. This award enabled me to focus on research and writing for an entire year, and it is because of the award that I was able to complete this dissertation in a timely manner. I am grateful to the UGA Graduate School for granting me this honor, and I likewise thank my committee members and graduate coordinator, Dr. Kenneth Fischer, for supporting me in my application for the award.

The musical examples from Beethoven's *String Quartet no. 14 in C# Minor*, Op. 131 appearing in Chapter 6 were produced by Greg Brown. The speed at which he produced these examples helped to guarantee the completion of this dissertation in time for graduation, and his neatness in assembling them is greatly appreciated. I'm almost sorry to have scribbled all over them while tracing out spatial gestures.

Many of my conceptions of sound, space, and spatialized sounds are predicated upon philosopher Casey O'Callaghan's theory of sounds as events. Casey was kind enough to share with me preliminary manuscript drafts of his book entitled *Sounds: A Philosophical Theory*

v

(expected 2007). He even went so far as to email me the final draft of the book as he was sending it to Oxford University Press for publication. Casey likewise reviewed several of my own chapter sections and provided helpful commentary and suggestions. His generosity in sharing his innovative theory, and the constructive criticism that he has lent me, is deeply appreciated.

Although beginning its interminable process of maturation in Dr. Ball's electronic music course, my initial interest in the spatial features of musical performance began with my experiences performing with the Georgia Guitar Quartet. This ensemble was established in 1996 at the behest of John Sutherland, the instructor of guitar studies at the University of Georgia, and is still active with all four original founding members. During our eleven-year tenure, we have employed spatial procedures in many of our original compositions and transcriptions. We have likewise experimented with alternate seating configurations—often treating the onstage spatial layout of the ensemble as a vital element of musical interpretation and presentation. Such performance experiences and experimentations have served dually as both the impetus and the proving ground for many of the spatial theories and concepts outlined in this dissertation. I look forward to many more years of performing, recording, and "spatializing" with the quartet.

Phil Snyder, one of my fellow quartet cohorts, created an arrangement for four guitars of Alberto Ginastera's *Danzas Argentinas*, op. 2. The second of these three dances is subjected to spatial analysis in Chapter 7. I wish to thank Phil for allowing me to reproduce sections of his arrangement in this dissertation.

Last, but certainly not least, I would like to thank my wife, Qiao Chen Solomon, my parents, Larry and Becky Solomon, and my dog, Spooner Chen Solomon. My wife and I were both writing our dissertations at the same time, and she was wise enough to insist that we

vi

purchase a second computer. This spark of genius (coupled with my own suggestion that we put the computers in separate rooms) is likely what preserved our marriage during the writing process. She has always been a profound source of inspiration and encouragement, and I thank her dearly for her patience and tolerance. My parents are the two most supportive people on the planet—emotionally and financially. They never even blinked when, years ago, I informed them that I was going to college to major in guitar, and their unconditional love and ever present kindness has served as a mainstay enabling my modest successes in various musical arenas. And Spooner, our little Boston Terrier, on innumerable occasions waited patiently for me to complete a frantic brainstorm or taxing revision at the computer before finally, finally taking him outside to play ball.

TABLE OF CONTENTS

		Page
ACKNOW	VLEDGMENTS	iv
LIST OF I	FIGURES	xi
INTRODU	JCTION	1
CHAPTE	R	
1	AN OVERVIEW OF SPATIALIZATION PRACTICES AND AESTHETICS	25
	1.1 From Psalmody to Concertato Motets	25
	1.2 Spatial Practices in the Classical and Romantic Periods	32
	1.3 Spatialization in the Twentieth Century	42
	1.4 Electronic Music	54
2	SOUNDS IN SPACE: WHAT SOUNDS ARE, WHAT SOUNDS DO, AND	
	HOW WE HEAR THEM	63
	2.1 Sounds	63
	2.2 Sounds in the Environment	70
	2.3 Auditory Localization	76
	2.4 Ensemble Space	88
	2.5 Auditory Streams	94
3	THE IDENTIFICATION AND CLASSIFICATION OF SPATIAL GESTURE	S:
	AN ANALYTIC METHODOLOGY	105
	3.1 Gestalt Theory: A Summary	105

	3.2 Spatial Gestures	126
	3.3 Migratory and Accumulative Gestures	144
	3.4 Gestural Categories	148
	3.5 Compound Spatial Gestures	169
	3.6 Ensemble Spaces: General Observations	175
4	RELATING MULTIPLE SPATIAL GESTURES AND EXPLORING	
	ALTERNATIVE SPACES	179
	4.1 Alterations of Spatial Gestures	180
	4.2 Polystreaming: Relations Among Concurrent Spatial Gestures	198
	4.3 Gestures in Multi-Dimensional Ensemble Spaces	217
	4.4 Spatial Gestures in Electronic Music	225
5	INTERPRETING THE SPACE: A HERMENEUTIC APPROACH	237
	5.1 Symbols in Space: The Semiotics of Spatialization	238
	5.2 Categorization	245
	5.3 Image Schemata	254
	5.4 Metaphor	325
	5.5 The Interaction of Spatial Gestures with Other Musical Events	336
6	A SPATIAL ANALYSIS OF BEETHOVEN'S STRING QUARTET NO. 14	
	<i>IN C# MINOR</i> , OP. 131	340
	6.1 Movement 1	342
	6.2 Movement 2	353
	6.3 Movement 3	362
	6.4 Movement 4	365

6.5 Movement 5	
6.6 Movement 6	402
6.7 Movement 7	405
6.8 Overview: The "Big Picture"	414
7 IMPLICATIONS, APPLICATIONS, AND CONCLUSIONS	420
7.1 Spatial Analysis and Performance Practice	420
7.2 Spatialization in Transcription	430
7.3 Recommendations for Extension in the Field of Spatial Analysis	
7.4 Conclusions	441
BIBLIOGRAPHY	445
APPENDIX: INDEX OF SPATIAL GESTURES	460
A.1 The Trio: Spatial Gestures in Three-Point Ensemble Space	461
A.2 The Quartet: Spatial Gestures in Four-Point Ensemble Space	463
A.3 The Quintet: Spatial Gestures in Five-Point Ensemble Space	467
GLOSSARY	477

LIST OF FIGURES

Page
Figure 0-1: The circle of fifths
Figure 0-2: The Tonnetz
Figure 0-3: Tymoczko's "scale lattice"
Figure 0-4: Bartók, String Quartet No. 4, Mvt. 2, mm. 142–145
Figure 1-1: Mozart, Serenade for Four Orchestras, Mvt. 1, mm. 1–10
Figure 1-2: Tchaikovsky, <i>Sixth Symphony</i> , Finale theme, mm. 1–4 (violins only)
Figure 2-1: Typical spatial relations among ensemble space and audience space within the
performance space90
Figure 2-2: Spatial relations among ensemble space and audience space in a surround-sound
environment
Figure 2-3: Stationary streams and gestural sequences
Figure 2-4: Stationary sequences and gestural streams
Figure 3-1: Emergence
Figure 3-2: Reification
Figure 3-3: a) Necker Cube and b) Rubin's Figure117
Figure 3-4: Stockhausen, <i>Gruppen</i> , section 119, mm. 1–4
Figure 3-5: CSEG <0132>
Figure 3-6: Common seating plan for a string quartet
Figure 3-7: Integer diagram of four-point ensemble space

Figure 3-8: Normalized integer diagram of four-point ensemble space	135
Figure 3-9: Beethoven, <i>Op. 131</i> , Mvt. 4, m. 89	136
Figure 3-10: Schnittke, String Quartet no. 4, Mvt. 1, mm. 79–82, accumulative SG- 3210	145
Figure 3-11: Berg, <i>Lyric Suite</i> , Mvt. 6, mm. 41–46, dissipative gesture	146
Figure 3-12: Schoenberg, String Quartet No. 3, Mvt. 4, mm. 139-141, adjacent-point	
oscillation gesture(s)	149
Figure 3-13: Ball, Wheels, mm. 314–317, distant-point oscillation gesture	151
Figure 3-14: Duarte, <i>Diptych No. 1</i> , Mvt. 2, mm. 128–131, fused oscillation gesture	154
Figure 3-15: Carter, String Quartet No. 4, mm. 114–116, weak open gesture	158
Figure 3-16: Ball, Wheels, mm. 81–84, strong closed gesture	161
Figure 3-17, Bartók, String Quartet No. 1, Mvt. 1, mm. 1–9, weak closed gesture	162
Figure 3-18: Spiral diagram of SG- 04132	163
Figure 3-19: Schoenberg, Quintet for Wind Instruments, Mvt. 3, m. 30, clockwise inward	
spiral gesture	165
Figure 3-20: A common woodwind quintet seating plan	165
Figure 3-21: Schoenberg, Quintet, Mvt. 4, mm. 309–313, clockwise outward spiral gesture.	166
Figure 3-22: Koshkin, <i>Changing the Guard</i> , mm. 121–131, cyclic panning gesture	171
Figure 3-23: a) Quartet seating plan with minimal arc, b) Seating plan with pronounced arc.	172
Figure 3-24: Schoenberg, String Trio, mm. 155–156, LEFT SWEEP	175
Figure 4-1: Bartók, String Quartet No. 1, Mvt. 1, mm. 32–38	188
Figure 4-2: Ligeti, Wenn aus der Ferne, mm. 1–5	202
Figure 4-3: Suggested seating arrangement for Ellen Taaffe Zwilich's Double Quartet	208
Figure 4-4: Integer diagram of the <i>Double Quartet</i> ensemble space	208

Figure 4-5: Zwilich, Double Quartet, Mvt. 1, mm. 46–50	209
Figure 4-6: Ambiguous streaming	211
Figure 4-7: Crossed streaming	212
Figure 4-8: Mirrored streaming (convergent + divergent)	213
Figure 4-9: Complex Gesture $SG - 031425 = Echo streaming of SG - 012 + SG - 345 $	216
Figure 4-10: Xenakis, <i>Empreintes</i> , recommended layout for the ensemble	218
Figure 4-11: Brant, Desert Forests, seating plan for onstage instruments	223
Figure 4-12: Ascending motion of a perfect fifth in pitch space	226
Figure 4-13: Horizontal passing gestures	232
Figure 4-14: Diagram of 7.1 surround sound system	234
Figure 5-1: Bach, <i>Fugue in C# Minor</i> , mm. 1–3	239
Figure 5-2: Hypothetical gestural category-container with SG- 0123 as the principal	
gestural form	
Figure 5-3: PATH schema	256
Figure 5-4: TENSION-AND-RELEASE schema I	262
Figure 5-5: FORCE schema	263
Figure 5-6: FORCE schema as sine wave	264
Figure 5-7: Elaborated FORCE schema	265
Figure 5-8: T&R schema II	265
Figure 5-9: VERTICALITY schema	
Figure 5-10: Lewin's Figure 0.1 (GMIT)	275
Figure 5-11: DEPARTURE-AND-RETURN (D&R) schema	281
Figure 5-12: CYCLE schema	284

Figure 5-13: Placement of performers and instruments in Persephassa	286
Figure 5-14: Xenakis, Persephassa, mm. 352–356	287
Figure 5-15: Xenakis, Persephassa, mm. 99–103, open counterclockwise circular gesture	289
Figure 5-16: CONTAINER schema	292
Figure 5-17: Nested containers/schema for hierarchy	293
Figure 5-18: Ensemble-space as container (five-point)	294
Figure 5-19: CENTER-PERIPHERY schema	297
Figure 5-20: Partial network for principal gesture SG- 0132	305
Figure 5-21: a) BALANCE schema and b) IMBALANCE schema	310
Figure 5-22: <i>Taijitu</i> , the traditional symbol of yin and yang	313
Figure 5-23: Ligeti, String Quartet No. 1 (Métamorphoses nocturnes), mm. 1207-1210	338
Figure 6-1: Beethoven, Op. 131, Mvt. 1, mm. 1–14	343
Figure 6-2: Mvt. 1, mm. 66–82	346
Figure 6-3: Mvt. 1, mm. 91–100	347
Figure 6-4: Mvt. 1, mm. 104–113	350
Figure 6-5: Mvt. 2, mm. 8–17	354
Figure 6-6: Mvt 2, mm. 84–92	356
Figure 6-7: Mvt. 2, mm. 123–137	359
Figure 6-8: Mvt. 2, mm. 161–168	360
Figure 6-9: Mvt. 2, mm. 169–175	361
Figure 6-10: Mvt. 3 (entire)	363
Figure 6-11: Mvt. 4, mm. 65–76	366
Figure 6-12: Mvt. 4, m. 89	367

Figure 6-13: Mvt. 4, mm. 98–111	
Figure 6-14: Mvt. 4, mm. 114–125	
Figure 6-15: Mvt. 4, m. 131	
Figure 6-16: Mvt. 4, mm. 142–145	
Figure 6-17: Mvt. 4, mm. 153–157	
Figure 6-18: Diagram of multilevel sweeping activity	
Figure 6-19: Mvt. 4, mm. 162–165	
Figure 6-20: Mvt. 4, mm. 180–182	
Figure 6-21: Mvt. 4, m. 195–196	
Figure 6-22: Mvt. 4, mm. 203–204	
Figure 6-23: Mvt. 4, mm. 227–234	
Figure 6-24: Mvt. 4, m. 245	
Figure 6-25: Mvt. 4, mm. 250–254	
Figure 6-26: Mvt. 4, mm. 264–265	
Figure 6-27: Mvt. 5, mm. 1–2	
Figure 6-28: Mvt. 5, mm. 25–39	
Figure 6-29: Mvt. 5, mm. 67–71	
Figure 6-30: Mvt. 5, mm. 88–91	
Figure 6-31: Mvt. 5, mm. 147–153	
Figure 6-32: The spatial process of an oscillation gesture	
Figure 6-33: Mvt. 5, mm. 162–169	
Figure 6-34: Mvt. 5, mm. 328–335	
Figure 6-35: Mvt. 5, mm. 466–470	

Figure 6-36: Mvt. 6, mm. 12–1440)3
Figure 6-37: Graphical representation of two panning gestures and their constituent	
contour relationships40)4
Figure 6-38: Mvt. 7, mm. 52–5540)6
Figure 6-39: Mvt. 7, mm. 124–12640)7
Figure 6-40: Mvt. 7, mm. 160–16840)8
Figure 6-41: Mvt. 7, mm. 264–27140)9
Figure 6-42: Right-to-left supercomplex (deconstruction of the right-to-left	
spatial complex)41	1
Figure 6-43: Mvt. 7, mm. 369–388 (closing)41	2
Figure 7-1: Ginastera, Danza de la moza donosa, arr. Phil Snyder, reduction of mm. 1–2543	2
Figure 7-2: Ginastera, Danza de la moza donosa, arr. Phil Snyder, reduction of mm. 59–8043	3
Figure 7-3: Ginastera, Danza de la moza donosa, arr. Phil Snyder, mm. 80–8143	34
Figure A-1: Integer diagram of three-point ensemble space	1
Figure A-2: Integer diagram of four-point ensemble space46	53
Figure A-3: Integer diagram of five-point ensemble space46	7
Figure G-1: Perfect score space-to-ensemble space translation	7

INTRODUCTION

Our perspective of musical space is utterly frozen and has led to a music in which the movement and direction of sound in space has no function. But the moment we have the means to move sound with any given speed in a given auditorium, or even in a given space outdoors, there is no longer any reason for a fixed spatial perspective of music. [The] movement in space of music becomes as important as the composition of its melodic lines...and its rhythmic characteristics. Whether a sound moves clockwise or counterclockwise, is at the left back or at the front, or any other combination—these are all configurations in space which are as *meaningful* as intervals in melody or harmony.¹

-Karlheinz Stockhausen

It is common in contemporary analytic discourse to portray abstract pitch relations with various models of pitch space. In such models, pitches are metaphorically projected as—and conceptualized in terms of—objects in physical space. Theorists speak of the "space" between two given pitches (or pitch complexes), even though no actual physical space exists between them (except within the geometric diagram of the model). In depicting musical relationships and processes, the "distance" between pitch-objects assumes significance, as does the "motion" between pitch-objects. Different models are employed to depict particular kinds of musical spaces, and the same two musical entities might exhibit contrastive relationships in separate spatial models. For example, in a chromatic space, the pitch A-flat is closer to G than A-natural is to G. However, in a diatonic space with G as the tonic pitch, A-natural is closer to G than is A-flat. The same concept applies to harmony: in a diatonic space, the dominant chord (V) is

¹ Karlheinz Stockhausen, *Stockhausen on Music: Lectures and Interviews Compiled by Robin Maconie* (London: Marion Boyars, 1989), 102–103. Emphasis mine.

closer to the tonic (I) than the supertonic chord (ii) is to I. The distances between pitches and chords in these examples are cognitive, not acoustic.²

During the aural experience of music, musical events are perceived to abide within a purely conceptual, phenomenal musical space. This conceptual space-the form and nature of which is variable from culture to culture as well as from one listener to another within the same culture—might be informed and shaped by exposure to theoretical constructs of musical space and/or conventions of musical notation. Conversely, a metaphorically structured conceptual musical space may direct the formulation and construction of analytic spatial models. The ubiquitous conceptual metaphor PITCH RELATIONSHIPS ARE OBJECT RELATIONSHIPS IN PHYSICAL SPACE guides our reception and conceptualization of musical events. This metaphor provides for the perception of pitches as high or low, which subsequently enables the generation of numerous linguistic metaphors used to describe musical phenomena: a "high" Fsharp, an "ascending" melodic contour, a "distantly" related key, the "leap" of a minor seventh, and A-natural being "closer" than A-flat to G in a specific diatonic space. Such metaphors are typically employed unconsciously; we are not always cognizant of the fact that our core understanding of music and consequent linguistic descriptions of musical events are both highly metaphorical.

Various geometric models of musical space triply emerge from, engender, and reinforce our conceptualization of pitches and/or pitch complexes (such as a chord) as points in space, the relations between musical entities as distances in space, and a musical event or process as the motion between two or more points in space. The measured distance between points, which is often dependent upon their hierarchical structuring within a given space, serves to qualify the nature, function, and affect of the musical relationship or event spanning those points. Spatial

² Fred Lerdahl, *Tonal Pitch Space* (Oxford: Oxford University Press, 2001), 42.

models serve to fortify the notion that we perceive music in a phenomenal, mental space: spatial perception begets spatial modeling, and vice versa.

Consider the familiar geometric model of musical space provided in Figure 0-1. The *circle of fifths* depicts key relationships in tonal pitch space. Traversing this circular model in a clockwise fashion visits keys "up" a fifth in pitch space. Traveling counterclockwise calls on keys "down" a fifth (or up a fourth) in pitch space. The circularity of the model reflects a latent feature of tonal music: if, for example, a work begins in the key of A major and continually modulates up a fifth, it will eventually return to A major. Although seldom found in tonal compositions, this modulation process could repeat incessantly in a cyclic manner.



Figure 0-1: The circle of fifths

The circle of fifths is primarily useful in identifying the relatedness of different key areas. Adjacent keys on the circle are "closely" related, with a difference of only one accidental between their respective key signatures. Non-adjacent keys are more "distantly" related. Keys that are maximally distant in tonal pitch space are also maximally distant in the circle of fifths. For example, a straight line drawn from C major to F-sharp major/G-flat major spans the entire diameter of the circle. The diameter of a circle divides the circle perfectly in half and is the greatest distance between any two points along the circumference of the circle. These two keys have roots related by a tritone (the interval that divides the octave perfectly in half) and key signatures that differ by six sharps (or six flats).

Figure 0-2 depicts another model of musical space, the *Tonnetz*. Early versions of the Tonnetz abound in the eighteenth and nineteenth centuries. The historical Oettingen/Riemann Tonnetz labels its nodes with pitches or pitch classes; the form shown in Figure 0-2 is its dual form, with nodes labeled as triads. This dual form of the Tonnetz shows how various major triads (portrayed by upper case letters) and minor triads (lower case letters) transform into one another via parsimonious voice leading, wherein two chord tones are preserved while the other moves either by half step or whole step.



Figure 0-2: A dual of the Tonnetz³

³ Taken from Clifton Callender, Ian Quinn, and Dmitri Tymoczko, "Generalized Chord Spaces (Draft of 20 June 2005)," available at: http://www.music.princeton.edu/~dmitri/chordspaces.pdf (accessed December 7, 2006), 61.

The neo-Riemannian transformations P (parallel), L (leading-tone exchange), and R (relative) are represented by the lines connecting chords within the model. For instance, by applying "P" to a B-major chord, we progress westward in the Tonnetz to arrive at a B-minor triad—chord tones B and F-sharp are retained while D-sharp moves by semitone to D-natural. "R," represented by motion along the Southeast diagonal, turns B major into a G-sharp-minor chord by shifting Fsharp to G-sharp while B and D-sharp are preserved. Finally, by following the Northeast diagonal (representing the "L" operation, in which B is exchanged with A-sharp), we arrive at Dsharp minor. In the graphic space of the Tonnetz, the chords situated closest together are those that are most similar to one another (sharing two out of three tones). Such a space is useful in depicting non-functional triadic successions, such as those common to late-Romantic music by the likes of Wagner, Mahler, and Strauss.⁴ It is critical to note that the *directionality* of the motion between entities in most graphic representations of music-spatial models is inconsequential: the fact that we traveled leftward in the space from B major to B minor is unimportant. What does matter is the proximity of entities as well as the nature of the "paths" (as transformative motions) that connect those entities.

To demonstrate how varied and complex the geometric modeling of pitch space can become, Figure 0-3 presents Dmitri Tymoczko's "scale lattice." This model depicts semitonal displacement functions that interconnect different transpositions of various pitch collections, including the diatonic, acoustic, harmonic minor, and harmonic major collections. Many spatial models, such as those for the tonal system, express aspects of musical intuition and reflect deeply acculturated music-conceptual spaces. Others are less intuitive—but not necessarily less instructive—or are specifically tailored to the design of a particular work of music.

⁴ For further basic information regarding neo-Riemannian theory, see Richard Cohn, "Introduction to Neo-Riemannian Theory: A Survey and a Historical Perspective," *Journal of Music Theory* 42.2 (1998): 167–180.



Figure 0-3: Tymoczko's "scale lattice"⁵

Apart from referring to theoretical, diagrammatic spatial models or to the cognitive spaces in which musical entities are perceived to inhere, the term "musical space" has yet another metaphorical association. Our basic human understanding of time results from our conceptualization of time in terms of space, the passage of time as motion through space, and temporal relations as spatial relations.⁶ Time is an abstract and intangible concept. To render such a concept sensible, we project features from our concrete knowledge of physical space onto those of time.⁷ Accordingly, we employ spatial terminology when describing temporal events: a certain "point" in time, a "long" time ago, the "distant" past, the years are "flying" by, I'll be

⁵ Callender et al., 63.

⁶ George Lakoff and Mark Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge Western to Western Thought* (New York: Basic Books, 1999), 137–169. Lakoff and Johnson describe two different "metaphorizations" of time: the *moving time metaphor*, in which time is understood to move (or, "flow") past the stationary observer, and the *moving observer metaphor* (or, "time's landscape"), wherein the observer moves through a stationary time-field.

⁷ The process of *cross-domain mapping*, which results in *conceptual metaphor*, shall be detailed in Chapter 5.

with you "shortly," and so on.⁸ Speaking of the "space" in a musical work often refers to the timing (i.e., temporal "spacing") between musical events, which may result from measured silence, fermati, etc.⁹ The type of musical space (as well as the spatial relationships unfolding within this space) elucidated in this dissertation is not metaphorical—it has nothing to do with either metaphorical pitch space or the temporal spacing separating musical events within a given work of music.¹⁰

The description of pitch relationships has long dominated the discipline of music theory and analysis. With so much attention given to *pitch space*, theorists have virtually ignored the important relationships unfolding within the *physical space* of musical performance. Composers, however, have frequently shown great sensitivity to the physical arrangement of performers relative to one another within the performance space. Such awareness of the spatial distribution of performers and the (often) deliberate spatialization of musical events can be traced from the practices of antiphonal psalmody in the ninth-century Mass, through the concertato motets of Adrian Willaert in the sixteenth century, to the most recent electronic and acoustic compositions of the early twenty-first century. Composers such as Henry Brant and Karlheinz Stockhausen have written extensively about their own experiments with the spatialization of musical materials. The lineage and development of spatial practices and aesthetics shall be explored in Chapter 1.

Figure 0-4 presents a passage from the second movement of Béla Bartók's *String Quartet No. 4* (1928).

⁸ The term "flying" is not directly spatial; however, it is an activity that presumes a space in which to occur.

⁹ Orchestration, as well as textural contrast and variety, may also contribute to the sense of spaciousness in music. ¹⁰ However, the interpretation of events within this space *is* often dependent upon metaphorical association, as we shall see in Chapter 5. Likewise, the spatial activity that I explicate *does* interact with concurrent events in pitch space as well as with the temporal organization of music.



Figure 0-4: Bartók, String Quartet No. 4, Mvt. 2, mm. 142-145

In m. 142, the second violin initiates a cycle of ascending glissandi. Each glissando in the cycle occupies two beats of musical time and spans the same range of pitch space with an upward trajectory (an ascending major seventh extending from middle C to B-natural above). It is important in musical discourse to distinguish between musical parts (in the score) and musical voices (in pitch space). Often, the various parts in the score feature individual musical voices, and, particularly in ensemble music, a different musician generally performs each part.¹¹ In fact, the first part (located in the top staff of each system in the score) frequently contains the upper musical voice, and the last part (situated at the bottom of the staff system) often contains the lowest musical voice. Indeed, the score is often (but not exclusively) an accurate visual representation of aural-conceptual pitch space. However, in the passage above, since each instrument is engaged in playing an identical figure, all musical voices are conflated within the same octave range such that there is no distinction of voice. The identical form of the glissando

¹¹ Polyphonic instruments, such as the piano and the guitar, are capable of rendering multiple voices contained in a single part.

figure occludes any differentiation of musical part that might otherwise occur by virtue of registral (voice) and timbral distinctions.¹² However, the "glissando cycle" is distributed among the four parts such that successive glissandi overlap one another in time. The total cycle comprises twelve overlapping glissandi, and each of the four instruments performs three discrete glissandi in direct succession. After the second violin launches the cycle, a spatial pattern emerges as the other instruments—located at separate points in the physical space occupied by the ensemble—enter sequentially with their respective "three-glissandi chains." The spatial pattern emerges from the following series of instrumental entrances: Violin 2, Violin 1, Viola, Cello. Depending upon the manner in which the performing ensemble is configured (which is indeed variable), this pattern may assume a variety of spatial shapes. The spatial pattern, which itself contains four spatially dispersed glissandi, is repeated a total of three times. Thus, the twelve-glissando cycle is apportioned as (and co-occurs with) a cycle of three identical spatial patterns—iterated in immediate succession. Its triple occurrence allows the pattern formed in actual space to function as a recurrent motive.

Thus, in addition to distinguishing between musical parts in the score and musical voices in pitch space, it is likewise imperative to distinguish the separate locations of performers (rendering their respective musical parts) relative to one another within the space inhabited by the performing ensemble. Often, as in the preceding musical example, the distinction of performer locale is the sole (or primary) catalyst for the differentiation of musical parts, the formation of larger motivic units (i.e., spatial patterning), and/or the generation of motion. In many instances, a direct correlation between parts in the score and voices in pitch space (as described above) will likewise share a correspondence to locatedness in actual space.

¹² The four quartet instruments belong to the same instrumental family (the string family). Although the cello playing in its upper register produces a tone somewhat distinct from that produced by the violin performing in its low-to-mid register, the timbral difference is subtle.

The primary purpose of this document is to develop an analytic methodology for identifying, classifying, comparing, and interpreting spatial gestures. A spatial gesture emerges from the consecutive activity of multiple performers within an ensemble.¹³ Various gestures are differentiated by the specific orderings, in time, of sonic events occurring at separate points in ensemble space.¹⁴ The recurring "spatial pattern" identified above in Figure 0-4 is an example of a specific spatial gesture. As a unified whole comprising ordered parts (gestalt structure), a spatial gesture is directed motion through an ensemble. Spatial gestures often serve a motivic function: a gesture may be developed and transformed, and the profusion of related gestures imparts spatial coherence to a work of music. Furthermore, the specific directionality and kinetic shape of a spatial gesture is rich in interpretive potential. The technique of **spatial analysis** proposed and developed in this dissertation (Chapters 3-6) is particularly useful for uncovering and demonstrating cohesion in many modern compositions—works in which pitch relationships are highly contextual and potentially obscured. However, a spatial analysis may reveal interesting features in music from *any* historical period or tradition. Spatial gestures are perhaps most evident (to the perceiver) and exploited (by the composer) within ensembles of identical instruments, such as the guitar quartet. In such an ensemble, the sameness of each instrument negates any inherent differentiation of timbre and range within the ensemble. The emergence of a spatial gesture in such an ensemble is generally more salient due to the timbral and dynamic homogeny of the "parts" within the "whole." Furthermore, the composer or arranger has greater freedom in distributing musical parts to the various instruments of an ensemble when those instruments are identical. In a woodwind quintet, for example, extremely high pitch events would need to be allocated to the flute (doubling piccolo) and low events allotted to the bassoon.

¹³ All terms appearing in bold face (within the body of the text) are defined in the glossary.

¹⁴ The temporal succession of spatially discrete events may involve either contiguous or nonadjacent ensemble members.

Such registral distributions carry the potential to delimit the spatio-gestural *affordances* of an ensemble space.¹⁵ When the sonic potentialities (registration, timbre, articulation, etc.) of various instruments constituting an ensemble are commensurate, musical events may be more freely dispersed. Since music for more than one instrument/performer is inherently spatial (no two objects may occupy the same exact space at the same time), spatial gestures are an omnipresent feature of music, irrespective of the composer's intentional and deliberate spatialization of musical events (or lack thereof). Many composers wittingly use the physical space of performance for contrast and special effect: consider Berlioz's placing of the English horn offstage in the *Symphonie Fantastique*, or the receding, offstage choir at the end of Holst's *The Planets*. I will argue, however, that spatialization is more than mere sonic effect: it is an integral component of musical structure and meaning that is worthy of—and long overdue in receiving—a well-developed analytic methodology.

The profusion of related spatial gestures imparts coherence to a work of music, and specific spatial shapes and patterned motions are rife with interpretive potential. For instance, imagine a string quartet seated in the standard arrangement of instruments (from left to right): Violin I, Violin II, Viola, and Cello.¹⁶ Now, imagine a scoring in which a specific pitch is played consecutively by the first violin, the second violin, the viola, and finally by the cello. The successive activity of the four discrete yet contiguous instruments combines to form a single gesture that spans a specific amount of space (depending upon the onstage spacing of the performers), with a particular directionality, over a certain amount of time. Such an occurrence has a definite gestalt quality: the spatial gesture as a *whole* emerges from the coordinated efforts

¹⁵ The concept of an *affordance*, and its applicability to spatial analysis, shall be outlined formally in Chapter 5. ¹⁶ Under this particular seating configuration, the top-to-bottom ordering of parts in the score corresponds to the leftto-right series of instruments in the ensemble. This particular "score space–to–ensemble space" translation often results in a concordance between activity in registral pitch space and events in the actual space occupied by the ensemble. This point shall be resumed in due course.

of multiple, spatially discrete *parts*.¹⁷ In this case, the gestural result is an obvious left-to-right motion traversing the entire ensemble. Reversing this procedure would result in a right-to-left motion (the retrograde of the original left-to-right directionality). Since these two gestures are related members of the same gestural category (they are closely related variants of one another), their mutual presence in a composition for string quartet would ensure a certain degree of spatial coherence.¹⁸ However, the directionality of each spatial gesture is open to hermeneutic treatment. Western minds would likely be inclined to associate the left-to-right gesture with the notion of progression, while the right-to-left gesture might be interpreted as regressive. An explanation for the contrasting interpretive potentials of these and other spatial gestures is provided in Chapter 5.

Contextualizing "Gesture"

Before progressing, I should qualify my use of the term "gesture." What I refer to as a spatial gesture could easily be called "spatial motive," "spatial theme," or any number of appropriately descriptive appellations. After all, I shall argue in due course that spatial gestures function as unifying motives and may be thematized as subjects of spatio-musical dialogue and narrative. The term "gesture" is optimally suitable and attractive for many reasons. *Gesture* may be defined generally as follows: a movement of part of the body, most commonly a hand or the head, to express an idea or meaning. The hand gestures that accompany speech often clarify or modify the meaning of the spoken words. Sign language is physical gesture that replaces

¹⁷ A spatial gesture in the electronic medium might *not* be composed of discrete, discernible parts. Spatial gestures in electronic music shall be addressed in Chapter 4.

¹⁸ Gestural categories are delineated in Chapter 3.

spoken and audible speech.¹⁹ Though initially and intimately bound to bodily movement, the word gesture is frequently employed to describe sundry musical phenomena. Robert Hatten remarks: "Perhaps no other term has been used in such a bewildering array of contexts as the term gesture in relation to music."²⁰ The fundamental quality of movement (or, motion) is typically preserved in each of the various musical applications of the term gesture. However, the nature of the movement, including the type of space through which the movement occurs, is variable. The movement may occur exclusively in time (a "rhythmic gesture") or in metaphorical pitch space (a rising or falling "melodic gesture"). A "dynamic gesture" might indicate a movement from pianissimo to forte (or vice versa) along the intensity scale. Hatten describes **musical gesture** as the "significant energetic shaping of sound through time."²¹ Generally, the "shaping" involves changes of frequency or timbre, "energetic" refers to intensity, "time" entails the duration of the sound(s), and "significant" indicates that the gesture means something. Hatten notes: "Musical gestures have meaning that is both complex and immediate, and often directly motivated by basic human expressive movements."²² I shall argue in Chapter 5 that the meaning of musical gestures in general, and of spatial gestures in particular, arises from our embodied knowledge-the stored sensorimotor patterns of bodily movement and spatial relations that ultimately serve to inform our understanding of more abstract phenomena such as musical motion.

¹⁹ For a discussion of hand gestures and their interaction with language and meaning, see David McNeill, *Hand and* Mind: What Gestures Reveal About Thought (Chicago and London: The University of Chicago Press, 1992); David McNeill, ed., Language and Gesture, Language, Culture and Cognition, ed. Stephen C. Levinson, no. 2 (Cambridge, UK: Cambridge University Press, 2000); Karen Emmorey and Judy Reilly, eds., Language, Gesture, and Space (Hillsdale, NJ: Lawrence Erlbaum Associates, 1995).

²⁰ Robert S. Hatten, Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, and Schubert (Bloomington and Indianapolis: Indiana University Press, 2004), 93. ²¹ Ibid., 95.

²² Ibid., 94.

My conception of *spatial gesture* conforms to Robert Hatten's more general formulation of musical gesture as the "significant energetic shaping of sound through time." A spatial gesture may be viewed as a specific instantiation of musical gesture—albeit one not expressly identified by Hatten. Hatten states: "The elements synthesized in a musical gesture include specific timbres, articulations, dynamics, tempi, pacing, and their coordination with various syntactic levels (e.g., voice-leading, metric placement, phrase structure)."²³ Hatten does not mention spatially separated sonic events as prospective "elements" to be "synthesized in a gesture." However, much of what Hatten has to say about gesture is directly relevant and applicable to my notion of spatial gesture. Hatten observes that in music, "the coherence of sequential events heard as part of a continuous action, process, or agency...is typically applied to progressions involving a sequence of durations, frequencies (pitch), and intensities, but also applicable to timbral sequences, as in cases of Klangfarbenmelodie."²⁴ I would like to supplement this statement by asserting that a spatial gesture is a sort of "spatial progression" (or, "spatial sequence") likewise heard as a continuous and coherent process.

I shall reference Hatten throughout this dissertation to draw correlates between his "musical gesture" and my "spatial gesture." The reader is cautioned to bear in mind that when Hatten speaks of gesture, he is not referring to spatial gesture (as I conceive of it) but to musical gesture in the more general sense.

²³ Ibid., 94.

²⁴ Ibid., 102. Hatten cites intensity, frequency, timbre, and duration as the "four fundamental properties of sound" (Hatten, *Interpreting Musical Gestures, Topics, and Tropes*, 102). While this conforms to the traditional view of sound, I hold that it is erroneous both to discount a sound's distal locatedness as a fundamental property and to regard the spatial orientation of a sound source as inconsequential to the musical experience. We hear the location of a sound as clearly as we hear its frequency, intensity, etc. As we shall explore in Chapter 2, the location and movement of a sound source relative to the perceiver affects the psychological sensations of intensity (loudness), frequency (pitch), and timbre (tone color). Conversely, the purely sonic attributes of a sound affect our ability to localize that sound accurately.

Spatial Analysis in Context

I adopt a "non-intentionalist" stance in my development of a system for spatial analysis. By non-intentionalist, I mean that any knowledge of a given composer's act of spatializing musical sounds—as a premeditated and calculated compositional endeavor—is not necessarily relevant to analysis, and the lack of such documentation does not discount a spatial analysis. Theory and analysis sometimes serves as the processual and temporal intersection between the past creative activity of the composer and the future perceptive activity of the listener. While theory may function as an intermediary between these two modes of musical experience (the act of listening and the act of composing), I argue that this is not the primary function of music theory. An effective analysis alone can generate a profound musical experience. This analysis may in no way serve as a model of the composer's creative process. It is a fool's errand to endeavor to probe the full extent of any composer's purportedly deliberate or purely intuitive compositional activity.²⁵ If this were the sole function of music theory, the discipline would not have progressed very much since Zarlino. Likewise, an analysis might not reflect the way that a listener would be predisposed to receive and conceptualize music (without having first been prompted by the analysis). Perceptually salient musical events are *not* always the primary objects of analytic discourse. Theory is not confined to modeling the hypothetical listener's perceptual act. The analyst might begin with musical objects that are readily perceivable, or those that are not, and then theorize about them—potentially producing analyses that challenge the musical mind. As noted, there does indeed exist a symbiosis between the spatial perception

²⁵ It is safe to assume that the ratio of *deliberate* activity (that which is intentional, purposive, designed) to *intuitive* activity (that which is unconscious, reactive, unstrategized) can vary drastically from one composer to the next, from one work to another by the same composer, and may even change in proportion during the composition of a single work.

and the spatial modeling of musical events: the former might occasion the latter, and sophisticated forms of the latter might challenge and expand the former.

Finally, a theoretical construct and analytic methodology is not a recipe for composing music. This is not to say, however, that composers do not draw inspiration and compositional strategy from express knowledge of music theory. In addition to engaging existing trends in musical design, a theory of music might be highly speculative—serving to prompt further analytic investigation and functioning as an engine for the inception of new compositional trends. In acknowledgement of the fact that theory does intersect with compositional and perceptual processes, I investigate spatialization as a compositional aesthetic in Chapter 1. In Chapter 2, I describe how we perceive the location of sound sources and how human localization capacity is affected by environmental factors, room characteristics (such as reflective properties), and the positioning of performer(s) relative to the listener. Chapter 3 describes how successive sound events may be perceived as a unified motion within an ensemble. Chapter 5 explores how meaning may be gleaned from perceived musical activity.

Any analysis should be judged based on the quality of musical experience that it affords. If we expect theory and analysis to conclusively model perception and/or the compositional process, we stand to be sorely disappointed. Theory can, and often does, connect up with these two modes of musical experience, and in some cases, theory might function as a connection between these two modes, but an analysis is fundamentally a unique means of experiencing music. An analysis is an interpretation of a musical work of art, and an analysis can itself function as a work of art by opening our ears and minds to deeper levels of understanding and appreciation. Theory is a means unto itself and need not be justified as an instrument for explaining how composers compose or how listeners listen. In the words of David Lewin: "We

16

must conceive the formal space of a GIS [Generalized Interval System] as a space of theoretical potentialities, rather than as a compendium of musical practicalities."²⁶

Ethan Haimo identifies two kinds of analysis that differ in what they purport to accomplish based upon attendant statements made by the analyst. The "type-one statement" indicates that the analyst "tries to reconstruct the composer's techniques or methods of composition, ascribing the cause or origin of specific features of the work directly to the composer's conscious acts." The "type-two statement" is "independent of claims regarding the composer's conscious actions."²⁷ This latter type of statement declares that musical features (as objects of analysis) are simply "there," without attempting to divulge the reason(s) for their sheer presence or their distinct state(s) of being. The following statement is of the type-one variety: "Schoenberg imbued the first of his *Three Piano Pieces*, Op. 11, with [014] trichords." Contrastingly, the following is a type-two statement: "Op. 11, No. 1, contains many [014] trichords." Both statements refer to the pitch complexes inherent in the same musical work. While the first statement tangentially ascribes the existence and recurrence of a certain pitch-class set to the intentional enterprise of the composer, the second makes no such claims—remaining neutral in regard to the "reason" for the set's abundance. Haimo elaborates:

For statements [type one] that rest on assertions of the composer's volitional acts, confirmation of intentions is not only permissible, it is mandatory. If an observer attempts to reconstruct the composer's conscious thought, procedures, or methods, then consideration of intentions is perfectly valid. But if the composer's thought is not invoked in the analysis, then it makes no sense whatsoever to submit the analytical claims to confirmation with respect to the composer's intentions.²⁸

I concur wholly with Haimo's premise: if the analyst indeed lays claims as to the composer's intentions and deliberate activity, those assertions need to be supported by outside documents,

²⁶ David Lewin, *Generalized Musical Intervals and Transformations* (New Haven and London: Yale University Press, 1987), 27.

 ²⁷ Ethan Haimo, "Atonality, Analysis, and the Intentional Fallacy," *Music Theory Spectrum* 18.2 (1996): 178.
²⁸ Ibid., 180.

such as compositional sketches or theoretical treatises. However, it is not crucial to analysis that the composer's intent be either evoked or proven. If intentionality *is* both evoked and proven, then reconstructing the compositional process is not necessarily the analytic end-all: analytic methodologies exterior to the exposed procedures of the composer might provide a telling account of a work's musical design, contributing to a more comprehensive assessment of that work.

Haimo takes Allen Forte to task for the latter's inconsistent statements pertaining to his own system of pitch-class set analysis. At times, Forte deems any correspondence between analytic process and compositional process a moot and dismissible point; elsewhere, he contends that Schoenberg was consciously composing with pitch-class sets (in a seeming effort to justify his own methodology). Haimo argues convincingly that there exists no proof that Schoenberg was mindfully employing permuted pitch-class sets in his early atonal compositions, but he concedes that Forte's analysis would be somewhat less problematic without the type-one statements (in particular, without the contradictory mix of type-one and type-two statements) that surround it. While Haimo goes on to criticize Forte's pc-set analyses on objective grounds, his primary objective is to condone Forte's unfounded and unproven claims regarding Schoenberg's intentionality.²⁹

Whether or not Schoenberg consciously composed with pitch-class sets neither validates nor discredits pc-set analysis as a mode of describing the organization and coherence of his atonal compositions. While Beethoven did not deliberately infuse his *Third Symphony* with the three-line Urlinie, a Schenkerian reading of the Eroica proves to be a valid and potent analytic undertaking—the analysis justifies itself through the musical experience that it affords. Elliot Carter is known to have intentionally loaded some of his works with all-interval tetrachords, but

²⁹ Ibid., 180–191.
an analysis of his music should not solely entail the uncovering of these pitch units. A composer may or may not have entertained the prospect of the ensemble space serving as a venue for spatio-motivic activity. Uncovering the multi-level proliferation of the Ursatz, or identifying the various instantiations of [0137] and [0146], tells us something about how the works in which they reside unfold over time and cohere as unified wholes—irrespective of the composer's intentionality. Likewise, discovering recurring spatial gestures within the mobile architecture of the ensemble space reveals motivic coherence on another musical layer, that of correlated sonic events in physical space.

Several scholars have discussed the importance of space in music. In her dissertation "Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations," Maria Anna Harley surveys the wealth of spatialized music produced in the twentieth century by the likes of Pierre Boulez, Karlheinz Stockhausen, Iannis Xenakis, Henry Brant, and others.³⁰ She deals primarily with composers who deliberately spatialize musical events and treat the spatial complex as a critical—if not primary—aspect of a work's overall compositional design. The purposiveness of these composers is well documented: not only did they precisely diagram how the ensemble members should be arranged relative to one another as well as to the audience (often within a specific performance venue), but they often wrote and lectured about their spatial aesthetics at great length and in considerable detail.

I argue that all music is spatial. Based on this view, I have designed a general system of spatial analysis that is applicable to any work of music—not just "spatialized music." My system is designed to be fluid enough to take into account the variance (and potential discrepancies) of location, distance, and rates of motion within a work's spatial complex that

³⁰ Maria Anna Harley, "Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations" (Ph.D. diss., McGill University, 1994).

might occur from one performance to the next. Furthermore, I have chosen the term *spatial* gesture instead of *spatialized* gesture to underscore the fundamental point that not every composer consciously infuses his or her work with spatial gestures. If we know that Milton Babbitt used a certain matrix to generate the pitch and rhythmic profiles of his *Semi-Simple* Variations (1956), and we embark upon uncovering and identifying the different row forms, we will accomplish little more than uncovering an aspect of Babbitt's compositional process. We will not have come close to exhausting all that can be said about the musical work under investigation. While the recovery and recreation of compositional process is a vital element of analysis that should not be hastily dismissed, it likewise should not function as the sole means of analysis for a given work of music. By applying multiple analytic models (those that are readily applicable) to a work, we might discover relationships and processes that go beyond what the composer "meant to do" in his/her music. An analysis does much more than retrace the composer's steps: it offers a new way of explaining and experiencing music. A spatial analysis might yield a model of spatial events that is more sophisticated and telling than any descriptions potentially provided by the composer. In other instances, a spatial analysis could simplify and/or clarify the composer's account of spatial activity.

While my spatio-analytic system is not intended purely as a model of perception, I do contend that spatial gestures *are* perceived during the experience of music, and that, as objects of perception, they are often more plainly evident than pitch motives—thereby potentially serving a greater unifying function than pitch events.³¹ Localization is an automatic aspect of audition. In addition to sonic attributes such as intensity, frequency, spectral profile, and duration, we hear a

³¹ As we shall discover in Chapter 6, spatial gestures operate on multiple levels of structure. Those arising on the surface of the spatial construct are often markedly perceptible. Many of the deeper levels of spatial structure are penetrable only (or initially) during the analytic process—the gestures unfolding within these levels (frequently over long time spans) often prove difficult to decipher in the absence of analytic guidance.

sound (or the source of that sound) to be located distally at a certain point in space. On the other hand, we must learn to recognize and identify musical characteristics such as interval type and chord quality. This observation conflicts with a statement made by Paul Miller in his dissertation proposal: "It will be presumed that compared with our ability to perceive pitch, rhythm, and timbre relationships, our ability to hear spatial movements in music is underdeveloped."³² I argue that since localization is a natural and inseparable aspect of auditory perception, spatial gestures (along with their dynamic interrelationships) are quite often easier to discern than abstract (and frequently complex) pitch and rhythmic relationships. Listeners simply might not yet be accustomed to hearing and receiving spatial motions and processes as musical. The observation that the four pitches constituting a fully diminished seventh chord are "passed" across a string quartet (as the chord is arpeggiated with one note given to each performer) from the left side of the ensemble to the right is less immediate than the observation that a specific ordering of spatially discrete sound events conspired to form a left-to-right motion through the ensemble. Furthermore, the average listener might find it simpler to interpret a left-to-right movement as a signifier of "progressive" motion than to conceptualize a German augmented sixth chord as "deviant" if it fails to fulfill its usual predominant function and, instead, acts as V^7 in a key one half-step higher than the original key. A degree in music is not required for the aural discernment of the spatial gestures. And again, whether the composer intended for a musical event to dance its way across the space occupied by the ensemble, or whether the spatial motion is incident to other musical processes, or whether it simply arose through intuition or by chance, is utterly irrelevant. If the gesture exists, irrespective of the forces responsible for its

³² Paul Miller, "The Analysis of Spatial Music in Works by Karlheinz Stockhausen" (Ph.D. diss. proposal, Eastman School of Music, 2005), available at: http://www.practiceofpaul.net/research/media/dissertation_proposal.pdf (accessed October 15, 2006), 6.

generation, it may potentially impact the listener and is open to analytic and hermeneutic treatment.

Additionally, as noted above, we metaphorically structure our understanding of abstract musical events as distances or motions within a conceptual musical space. Miller acknowledges this phenomenon, but in making his assertion that hearing spatial movements is "underdeveloped" compared to our ability to perceive other musical relationships, he overlooks the fact that our embodied knowledge (instilled through repeated sensorimotor experience) of concrete relationships in actual, physical space serves as the source domain for structuring our understanding of abstract, musical events. Miller relates specific spatial events in music to certain types of melodic activity. These melodic features, however, are initially conceptualized in terms of concrete spatial events.³³ We, as listeners, are simply not yet entirely acclimated to receiving spatial gestures as important structures of the musical experience—as harbingers of unification and meaning. This is likely what Miller is claiming to be an underdeveloped capacity: attending to spatial relations and motions and judging them to be musically consequential. Hopefully, continued work in the field of spatial analysis will enhance our sensitivity to spatial gestures.

Since my system is applicable to music in which the spatial organization of the ensemble is unspecified, variable performance conditions and psychoacoustic factors of sound that may affect the nature and/or perceptual salience of spatial gestures are described in Chapter 2. Schoenberg's *Woodwind Quintet* may be performed in a variety of venues and by ensembles

³³ Miller cites Fux's "principle of non-coincident extremes in melodic construction" as well as the phenomena of melodic gap-fill and parsimonious voice leading. He also likens hierarchical properties in his model of the spatial activity in Stockhausen's *Lichter-Wasser* to those inhering within models of tonal pitch space, and he compares different types of ensemble spaces to different scale structures. Such comparisons are certainly instructive, but it interesting to note that Miller is using that which is rather abstract (albeit deeply ingrained) to forge an understanding of that which is more concrete and directly experienced—namely, relationships and events in actual space. See Miller, 12, 16, 22.

employing differing seating arrangements. The acoustic properties of the hall could obscure the spatial differentiation of the discrete instruments/performers within the ensemble, which would have an adverse effect on the salience of spatial gestures. The directionality, classification, and "meaning" of a spatial gesture might be fundamentally altered if the ensemble's seating arrangement is changed. Such conditional factors correlate strongly with both performance practice and musical interpretation: ensembles should feel compelled to observe the standard seating arrangements for certain ensemble-types and historical periods. Furthermore, a director or ensemble member cognizant of a work's implicit spatial design may adjust the location of the performers relative to one another, modify the ensemble's positioning relative to the audience, and "tune" the performance hall to optimize the presence of spatial gestures. The physical orientation of ensemble members, when undesignated by the composer or unspecified by practice and tradition, becomes a matter of musical interpretation.

In short, the analytic and interpretive methodology outlined in this dissertation is directly related to the actual, live performance of a musical work. However, one need not have a live quartet at one's disposal, a video recording, or even a "spatially faithful" audio recording of the work being performed to derive and attend to the dynamic spatial design of the work. The composer often specifies the layout of the ensemble space with a pictorial diagram. In cases where the spatial scheme is left unstipulated, it is quite possible to hypothesize the seating arrangement of a particular ensemble based on standardized performance practice. The musical score may be "translated" into a conjectural (or, "imagined") ensemble for analytic purposes. In Hatten's words: "Gestures may be inferred from musical notation [as well as] from a musical performance even when we do not have visual access to the motions of the performer."³⁴ In cases where several viable options for ensemble organization might exist, as with the string

³⁴ Hatten, Interpreting Musical Gestures, Topics, and Tropes, 94.

quartet, a spatial analysis might reveal the most spatially effective seating plan for that particular work, and this, of course, is a matter of interpretation.

CHAPTER 1

AN OVERVIEW OF SPATIALIZATION PRACTICES AND AESTHETICS

The spatial dispersion of performing forces has long been a constituent of musical practice. Beyond the fact that any ensemble of two or more performers must feature a spatial arrangement of some kind among its members, many composers have deliberately treated the space of musical performance as an important compositional element. Spatial practice in music originated in the early psalmodic practices of the Office and Mass. Antiphony reached its first zenith in the concertato motets of the seventeenth century. During the eighteenth and nineteenth centuries, antiphonal practices were unseated by the overwhelming concern for balanced forces and blended sonorities. In the twentieth century, spatialization resurfaced as a critical feature of sonic design—regaining its compositional applicability and vitality. This chapter shall outline the history and development of spatial practice while exploring the varied techniques and aesthetics of spatialization.

<u>1.1 From Psalmody to Concertato Motets</u>

Psalmody, the act of chanting the biblical Psalms in public worship, dates back as far as the late fourth century, although the earliest surviving musical records of psalmodic practice date back only as far as the ninth and tenth centuries. St. Ambrose (d. 397) is credited in some sources as having introduced antiphonal singing into the West in imitation of Eastern Christian

25

models.¹ In its developed and fairly standardized ninth-century state, psalmody takes three primary forms: direct, responsorial, and antiphonal. *Direct psalmody* is simply the chanting of a Psalm by a single, undivided group (a choir or congregation) without the incorporation of a refrain. *Responsorial psalmody* features a soloist (or group of soloists) singing the verses of a Psalm in alternation with a large choir (or the congregation), which generally sings a respond (a refrain that is typically melismatic). The consistent interchange of soloist and choir continues throughout the singing of the Psalm, thus engendering a spatial "oscillation" between a small group and a large group of singers.

In *antiphonal psalmody*, the verses of a Psalm are sung in alternation by the two halves of either the choir or the congregation. The choir is typically divided so that the two halves face one another and are seated on opposite sides of the altar.² The verses of the Psalm are often separated by a refrain-like antiphon, which also appears at the beginning and end of the Psalm.³ Responsorial psalmody and antiphonal psalmody represent the earliest known and recorded practice of spatialization in Western music. All of the antiphonal chants of the Mass are associated with actions, such as the entrance of the priest, the offertory, and the communion; therefore, the spatial separation of chanting forces is ritualized. Just as initially setting the Psalms and other liturgical texts to music was a means of embellishing and intensifying a sacred text, ritualistically spatializing the performing forces further embellishes the musical rendering of the text by "conversationalizing" it.

¹ "Psalmody, Latin," in *The New Harvard Dictionary of Music*, ed. Don Michael Randel (Cambridge, MA, and London: The Belknap Press of Harvard University Press, 1986), 664.

² While responsorial psalmody features an *imbalance* in the spatial proportions of performing (chanting) forces (a small group "versus" a large group), antiphonal psalmody generally exhibits spatial *balance* (in the form of an evenly divided choir/congregation). The issue of spatial balance/imbalance shall be explored in Chapter 5. ³ "Psalmody, Latin," 664–665.

The idea of a "conversation" (or, "dialogue") arising from the interactivity between spatially separated performing forces-nascent in early psalmodic practice-pervades numerous musical styles. The convention of "call and response" transcends geographical and historical lines of stylistic demarcation: it remains a fixture of Western art music and is common to many popular idioms, including folk, blues, gospel, rock, and jazz. The notion of call and response embraces more than just the spatial differentiation of the "caller(s)" and the "responder(s)"—it is commonly manifest in harmonic, melodic, and formal devices. Indeed, call-and-response activity often does not involve spatialized forces at all. For instance, the antecedent-consequent (or, "question-and-answer") construction of a tonal, two-phrase period is a type of call and response encapsulated by melodic and harmonic activity: the attainment of closure in the first phrase is interrupted and subsequently completed in the second phrase (the latter phrase is a definitive response to the open-endedness of the former phrase). In a rock song, the lines of text sung by the lead singer may alternate with "fills" from the electric guitar. As commonly heard through amplification or during playback of a commercial recording, these two distinct (but interlinked) musical actions might not be spatially differentiated. The innumerable musical instantiations of call and response are variations on the early practice of physically separating functionally distinct-yet coordinated-performing forces.

The *polychoral* style features music composed for two or more choirs and is the direct descendant of responsorial and antiphonal psalmody. The Venetian school of the fifteenth and sixteenth centuries, primarily centered around St. Mark's Basilica, was largely responsible for the development and maturation of polychoral music, although the style quickly gained eminence and was cultivated throughout Europe. The terms *coro battente* and *coro spezzato* ("broken" or "split" choir) are often used in association with music of this type. Although

27

Johannes Martini and Johannes Brebis produced "double-choir" Psalm settings in the 1470s for the court of Ferrara, Adrian Willaert was the first Venetian-based composer with whom polychorality is closely associated.⁴ He published his *salmi spezzati*, which consists of polychoral settings of Psalms, in 1550.⁵ Prior to Willaert's musical settings, the two choirs rendering a Psalm would simply alternate singing the verses of that Psalm. In the works of Willaert, the interaction between the two choirs is enhanced: the end of one verse is elided with the beginning of the next (each realized by a different choir), and the choirs perform together at the end—a climactic and conclusive reconciliation of their prior spatial separation. Willaert's polychoral settings of the Psalms are considered to be the "starting point of the double-choir tradition at St. Mark's in Venice."⁶

In the 1550s, music theorists Zarlino and Vicentino both commented on polychoral techniques. While Zarlino primarily discusses the polychoral settings of Psalms, Vicentino widens his discussion to account for polychoral settings of the Mass ordinary, motets, and madrigals. In his *L'antica musica ridotta alla moderna prattica* of 1555, Vincentino observes that the physical separation of sound sources—rather than the sheer presence of a multitude of singers—most effectively fills a hall with sound, and that color contrast between the choirs is achieved by scoring each choir differently. In his *Le istitutione harmonische* of 1558, Zarlino indicates that the multiple choirs should occasionally sing together—particularly at the end of a

⁴ The polychoral tradition had begun to emerge in some secular settings prior to Willaert. Multiple choirs were often employed at festive occasions and ceremonies, such as weddings. Polychorality was incorporated to enhance the "pomp and splendour" surrounding such events. See Anthony F. Carver, *Cori Spezzatti*, vol. 1, *The Development of Sacred Polychoral Music to the Time of Schütz* (Cambridge: Cambridge University Press, 1988), 1–5.

⁵ Allan W. Atlas, *Renaissance Music* (New York and London: W.W. Norton & Company, 1998), 413–416. ⁶ Ibid., 414.

polychoral work. He seems to have prized the effect of oppositional forces coming together and operating in tandem to enhance the finality of a work's ending.⁷

Following Willaert, Giovanni Gabrieli elevated the stature of polychoral music with his *concertato motets*. In these works, Gabrieli often integrates instruments and singers within a single choir, or he separates the vocal choir from the instrumental choir.⁸ The individual choirs are seldom identical to one another: each choir often features a different composite of instruments and/or voice types as well as a different number of performers. Often, the choirs are segregated by register—a group of high voices pitted against a group of low voices. Gabrieli's antiphonal practice thus conforms to Vicentino's advocation to score the separated choirs differently so that contrasts of color and sonority may be easily achieved. Here we see the emergence of the key element of *contrast* that would figure so prominently into the concertato style of the early Baroque period. In Venetian polychoral works, contrast arises from the dynamic, registral, and timbral distinctions between different choirs as well as from the spatial locations of those choirs.

The exact manner in which the discrete choirs were physically separated and oriented relative to one another in St. Mark's has been debated and contested amongst scholars. David Schulenberg notes: "Sometimes the two or more choirs employed in a polychoral work were physically separated, to enhance the 'stereo' effect that results from hearing two or more groups performing in different locations within a church or other large structure; balconies and organ or

⁷ Carver, 5–12.

⁸ One of Gabrieli's most celebrated works, *In ecclesiis* (published posthumously in 1615), is scored for three choirs—two vocal choirs and one instrumental choir—plus organ continuo. It was not uncommon practice in the sixteenth and seventeenth centuries for instruments to perform parts originally written for voice (obviously without delivering the text). Aside from the organ, which generally served as the continuo instrument, the cornetto and the sackbut were the favored instruments for inclusion in polychoral music. See David Schulenberg, *Music of the Baroque* (New York and Oxford: Oxford University Press, 2001), 135–136.

choir lofts were particularly favored for this purpose."⁹ However, Allan Atlas writes: "There is not a shred of evidence to show that, as is often thought, the choirs were dispersed stereo-like in the organ lofts. Rather, they were either situated at floor level next to the altar or squeezed into one of the two pulpits that stood in front of a partition separating altar from nave."¹⁰ Carver echoes Atlas's comments but goes on to remark that "there is no reason to suppose that such [polychoral] psalms were not performed spatially elsewhere."¹¹ Polychoral music invariably features two or more choirs stationed at different locations in space. The point of contention is just how physically separated the choirs actually were in sixteenth and seventeenth-century practice: were they dispersed throughout multiple lofts and balconies—creating a surround-sound atmosphere—or were they partitioned in a centralized location and situated adjacent to one another? Regardless, the choirs rendering polychoral works in St. Mark's Cathedral were spatially segregated to *some* extent: they were not, for example, intermixed in a single choir loft.

Regardless of the actual degree of separation in performance, the spatial differentiation of choirs was a hallmark of late-Renaissance vocal music—the separation of forces providing for a multiplicity of textural combinations and contrastive sonorities.¹² However, the separateness of sounds in these polychoral musical settings was not intended solely as a sonorous and contrast-

⁹ Ibid, 134.

¹⁰ Atlas, 416. Other churches in Northern Italy, such as the Verona Cathedral and St. Anthony's in Padua, feature semi-circular choir lofts behind the alter, such that "double-choir psalms could have been performed without the choirs being unduly separated" (Carver, 11). In spite of the close proximity of the two distinct choirs, however, some degree of stereophony—enabled by the left/right differentiation of performing forces—would undoubtedly infiltrate the musical experience.

¹¹ Carver, 11.

¹² In this chapter, I am primarily discussing musical works in which the composer deliberately conceives of—and actuates—the physical separation of performing forces. This is not to imply that important spatial relations, including what I refer to as *spatial gestures*, do not occur in musics less deliberately spatialized. I shall echo this point repeatedly throughout this dissertation. In a composition for a single mixed choir—as opposed to a polychoral work for multiple choirs—spatial gestures potentially emerge given the fact that the singers rendering each of the four component voice parts (soprano, also, tenor, and bass) are typically (but not exclusively) grouped together within the choir. Thus, a single choir comprises four sections. Although the sections are adjoining and proximate, each section occupies its own spatial area. Any two sections might engage in an antiphonal, call-and-response musical dialogue, and other types of spatial gestures (to be outlined in Chapter 3) may arise as well.

providing special effect: choral antiphony served syntactic and rhetorical functions as well, often being employed to articulate the form, underscore the meaning, and enhance the drama of the text being set. For example, the two choirs would often join together to denote the end of a section of text (such as a verse) or to emphasize important words of the text.¹³ Spatial differentiation and contrast served a critical musical function.

Following Gabrieli's death in 1612, his former pupil Heinrich Schütz extended the development of the polychoral concertato style through the seventeenth century—well into the Baroque period. By this time, the polychoral style had disseminated throughout Europe. With his *Symphonie sacrae*, Schütz popularized the style in Germany, and alongside Michael Praetorius, he published principles governing the instrumentation of polychoral works. In his *Syntagma Musicum* of 1619, Praetorius discusses polychorality in terms of divisions made on the basis of register, resulting in "higher" and "lower" choirs (a practice observed in many of Gabrieli's polychoral works).¹⁴ James Smith Pierce traces the development of the polychoral style in Rome during the seventeenth century, where:

Roman *maestri de cappella* expanded the Venetian concertato to unprecedented proportions, erecting scaffolds where galleries were lacking for their multiple instrumental and vocal choirs. Overwhelming the congregation on all sides four, six, eight, even twelve choirs performed music written in as many as 48 parts. Architecture and music were physically united as a phrase was echoed back and forth over the heads of the congregation, from gallery to gallery, from one choir to another, voices alternating with instruments and high voices contrasted to low.¹⁵

Pierce references a 48-voice mass composed in 1650 by Orazio Benevoli, the maestro de capella at the Vatican. The work involves 150 singers organized into twelve choirs that surround the

¹³ Schulenberg, 133.

¹⁴ Carver, 147.

¹⁵ James Smith Pierce, "Visual and Auditory Space in Baroque Rome," *The Journal of Aesthetics and Art Criticism* 18.1 (1959): 56–57.

audience on all sides.¹⁶ Such a dramatic spatial dispersal of forces exemplifies the apotheosis of polychorality in the mid seventeenth century.

1.2 Spatial Practices in the Classical and Romantic Periods

With the emergence of the pre-Classical style in the first half of the eighteenth century, polychoral techniques largely came to be associated with the *stile antico*. Antiphony remained a fixture of sacred choral music, which was generally less progressive than the secular and instrumental forms coming into fashion in the early Classical period. In reference to this music-historical period, Denis Arnold states that the "cori spezzati…lived on to become something of an anachronism."¹⁷ Carissimi employed the style in his oratorios, as did Bach in his motets and the *St. Matthew Passion*. In such contexts, polychoral techniques served to preserve the sanctity of liturgical music and to heighten the reverence for older musical idioms. As the more progressive musical trends increasingly turned away from counterpoint and began to embrace homophonic textures, the interest in spatial separation as a musical element waned dramatically.

During the Classical and Romantic periods (approx. 1750–1900), a concern for the proper blend and dynamic balance of instrumental forces prompted experimentation with the arrangement of instrumental families within larger ensembles such as the symphony orchestra. It became standard, for example, to situate the brass and percussion sections at the back of the orchestra to compensate for their powerful dynamic range (in comparison to that of the strings and woodwinds). Almost exclusively, the strings were located toward the front, the woodwinds in the middle, and the battery of percussion and brass at the rear of the orchestra. Such practices remain standard in today's symphony orchestras. What has not been rigidly systematized is the

¹⁶ Ibid.

¹⁷ Denis Arnold, "The Significance of 'Cori Spezzati," Music & Letters 40.1 (1959): 14.

exact placement of the different instrumental subsections (generally performing distinct musical parts) relative to one another within their larger familial sections.¹⁸ For instance, it was equally common in the nineteenth century for the first and second violin subsections to be situated either adjacently (side by side) or antiphonally (separated by the viola and cello subsections) within the string section. In contemporary symphony orchestras, the adjacent-violins arrangement preponderates, although some conductors prefer the antiphonal layout.¹⁹ The composers of the Classical and Romantic periods typically did not indicate the preferred spatial positioning of instruments for performances of their works; rather, the orchestra director was responsible for assigning section and subsection locale. Composers frequently conducted their own works and under such circumstances would have had direct control over the onstage spatial arrangement of the musical parts. Additionally, the composer may have (intuitively or deliberately) orchestrated the musical parts (assigned them to specific instruments or groups of instruments) based upon the standard and most common seating for that composer's geographic and historic location.

Christoph Willibald von Gluck applies spatial differentiation to drive the narrative and enhance the drama of his tragic reform opera *Orfée et Euridice* (first performed in Vienna in 1762).²⁰ In addition to the principal orchestra, Gluck incorporates into Act II a second orchestra comprising stings (violin, viola, cello, contrabass) and harp.²¹ In his score, Gluck indicates that this orchestra should be positioned "derrière le théâtre," which is best translated as "offstage" or

¹⁸ The term "section" is commonly used to indicate both the larger instrumental family (such as the entire string section) and the specific members of that family (such as violin, viola, cello, etc.). I shall employ the term "subsection" to describe the smaller groups within the larger instrumental families.

¹⁹ "The Symphony: An Interactive Guide," available at: http://library.thinkquest.org/22673/orchestra.html (accessed January 2, 2007); "The Orchestra: A User's Manual," available at: http://www.mti.dmu.ac.uk/~ahugill/manual/ seating.html (accessed January 7, 2007).

²⁰ Thanks to David Haas for encouraging me to explore this work.

²¹ Gluck employs the second orchestra (minus harp and with added oboe) in the aria "chiamo il mio ben cosi" (Act I, scene 2). Here, the second orchestra is employed sparingly to "echo" musical material presented by the principal orchestra. The intermittent echoing of the second orchestra occurs almost exclusively between lines of the text.

"backstage."²² In the second act, Orfeo travels to the underworld to retrieve his lost Euridice. The chorus and primary orchestra represent the "furies" of the underworld, and they initially deny Orfeo entrance into their domain. He entreats them with his voice and lyre, represented by the harp of the second orchestra, and eventually succeeds in persuading the furies to grant him entrance. As he persistently implores the furies, Orfeo draws out an arpeggiated accompaniment on his lyre (the harp), which is likewise joined by the strings of the second orchestra. Thus, the second orchestra represents not only Orfeo, but also earth-both of which encroach upon the underworld and its demonic inhabitants. The primary, larger, and more "present" orchestra portrays the here and now of the underworld, while the secondary, smaller, and spatially remote orchestra denotes earth. Orfeo's persistent appeals are periodically interrupted by a definitive "non!" from the chorus of furies (which is punctuated by tutti chords from the primary orchestra). This verbal interchange transforms into a more proper dialogue (between Orfeo and the furies) during the latter half of the first scene, as Orfeo begins to sway the furies to admit him into their world. Thus, the literal call and response (question and answer) between Orfeo and the furies is played out in the spatial realm by the physical separation of the two orchestras—with the distant second orchestra depicting the distance of earth. On a deeper level of interpretation, a call and response ensues between two "worlds": the underworld and earth.²³

A notable example of a purely instrumental spatial composition from the Classical period is Mozart's *Serenade No. 8 (Notturno) in D Major for Four Orchestras*, K. 286 (1777). In this highly unconventional work, each of the four individual orchestras consists of a four-part string

²² Regrettably, some musical anthologies omit the offstage instructions for the second orchestra. In the absence of such specifications, the intended spatial separation of the second orchestra from the primary orchestra—which is integral to the unfolding narrative—is indeterminable.

²³ An in-depth assessment of the spatial structure of staged works (including opera and ballet) lies beyond the scope of this dissertation. Such works most certainly merit a comprehensive spatial analysis—one that ties together musical, spatial, textual, and visual (the physical gestures and movement of the performers) attributes of the total performance.

section plus two horns. Mozart does not specify a preferred seating plan for the four ensembles; however, he uses musical devices and processes (such as dynamics and thematic fragmentation) to generate the illusion of echo. Mozart thereby creates a virtual space (with chimerical reflective properties) that becomes an integral element of this particular composition: the space is fabricated from the careful coordination of metric, dynamic, and developmental qualities—as actualized in performance.

A 1964 review of a performance of Mozart's *Serenade* published in *Time Magazine* explains that the work "consists of make-believe echoes, in which a short statement by the first orchestra is repeated in turn by the other three, each abbreviating the phrase until the fourth sounds only a faint fragment of it—just as an echoed shout fades out in the distance."²⁴ The systematic and cumulative fragmenting of a musical phrase represents the distortions that occur to a sound as it travels great distances and reflects from multiple sources: certain components of the sound are filtered and attenuated. Furthermore, late reflected signals may be partially masked by earlier reflections such that only the end of a very late echo is actually perceptible.

Figure 1-1 shows the opening ten bars of Mozart's *Serenade*. After Orchestra I exposits a four bar phrase, Orchestra II repeats the phrase verbatim—this is the "Erstes Echo" (first echo) as marked in the score by the composer (see m. 4). The identical yet spatially dispersed phrases are elided in time: the "first echo" begins before the initial sonic event has concluded. The "Zweites Echo" (second echo) at Orchestra III, appearing in m. 7, consists of only the second half of the original phrase. The "Drittes Echo" (third echo) at Orchestra IV (mm. 9–10) is fragmented further still—it is approximately half of the second echo and consists of only the

²⁴ "A Choice and an Echo," *Time Magazine*, 7 August 1964, available at: http://www.time.com/time/magazine/ article/0,9171,871353,00.html (accessed February 16, 2007), 1.

final cadential figure of the original phrase. The progressive disintegration of this musical phrase reflects the gradual enervation of each subsequent "echo."²⁵



Figure 1-1: Mozart, Serenade for Four Orchestras, Mvt. 1, mm. 1-5 (continued next page)

²⁵ This phrasal deconstruction may be viewed as a special case of *liquidation*, a formal process described by Schoenberg. In Schoenberg's sense of the word, liquidation occurs when the characteristic features of a theme are systematically eliminated to the extent that only a conventional (i.e., cadential) figure exists. Normally, however, this process encompasses a single statement of a theme (which might exist as a period, sentence, or other small form). In the case of Mozart's *Serenade*, the liquidation occurs over four consecutive statements of the same musical theme—the third of which is a reduced version of the first two, and the fourth of which is a further reduction of the third. See Arnold Schoenberg, *The Musical Idea and the Logic, Technique, and Art of its Presentation*, edited, translated, and with a commentary by Patricia Carpenter and Severine Neff (New York: Columbia University Press, 1995), 382.



Figure 1-1 (continued): Mozart, Serenade for Four Orchestras, Mvt. 1, mm. 6-10

For the particular performance of the *Serenade* reviewed in *Time Magazine*, the orchestras were spread out such that "orchestra one was on the stage, two and three were on the...floor at either side of the stage...and four ended up nearly out of sight under a canopy."²⁶

²⁶ "A Choice and an Echo," 1.

The synchronization of the remote fourth orchestra with the other three more proximate orchestras was achieved by employing an additional conductor to direct orchestra four. The second conductor monitored the primary conductor (leading orchestras one, two, and three) over a closed-circuit television and headphones. To further assist in achieving the "distant echo effect," all of the strings in the fourth orchestra were muted, three-quarters were muted in orchestra three, only half in orchestra two, and none in orchestra one. Thus, the progressive reduction in volume as well as the darkening of tone provided by the proportion of muted to unmuted strings in each orchestra aided in the effective conveyance of distance. While not specified by Mozart, the incorporation of mutes is a matter of interpretation that draws out and enhances the inherent spatial illusion of the composition. Obviously, the actual spatial distinction between orchestras one through three and the spatial distance of orchestra four served to facilitate the efficacy of the echo effect.

Since the composers of the Classical and Romantic periods were, in general, less overtly concerned with the spatial distinction of instruments than with achieving a composite, homogenous blend, it seems easy to assume that the variability of orchestral seating plan from one orchestra to the next is relatively insignificant to the performance of a given symphonic work. However, let us consider a passage from the final movement of Tchaikovsky's *Sixth Symphony* (1893).²⁷ Figure 1-2 shows the movement's principal B-minor theme, which descends from the fifth (F#) of a B-minor triad to the root (B) in stepwise fashion before stepping up to the supertonic scale degree (C#): F#-E-D-C#-B-C#. The theme, however, is partitioned note-by-note and dispersed among the first and second violins. Depending upon the elected sub-sectional arrangement of the orchestra's string section, the listener would have a markedly different experience of this section. In either case, the spatial "splitting" of the theme would likely be

²⁷ Thanks to Adrian Childs for bringing this passage to my attention.

quite evident. However, the "antiphonalization" of the theme would be dramatically enhanced if the orchestra's string section were configured with the two violin sections positioned at either end, separated by the cello and viola sections.²⁸



Figure 1-2: Tchaikovsky, Sixth Symphony, Finale theme, mm. 1–4 (violins only)

It is not uncommon for composers to manipulate the dynamic level of a musical work in an effort to impart a sense of either nearness or farness. Loud is often equated with "large"—and soft with "small"—but loud may also signify "close," and soft, "distant."²⁹ The finale of Tchaikovsky's *Sixth Symphony*, for example, recedes into nothingness via textural dissipation and a gradual reduction of dynamic level. An earlier example of changes in virtual distance wrought by dynamic activity is found in Luigi Boccherini's *Guitar Quintet No. 9 in C*, G. 453, likely composed around 1800.³⁰ The fourth and final movement of this work is entitled *La Ritirata di Madrid*, or, "the retreat from Madrid." The movement consists of a march-like theme with twelve variations. It begins at a very low dynamic level and gradually crescendos to the

²⁸ While demonstrating virtuosic and inventive orchestration, perhaps the spatial separation of the theme serves a higher musical function. The thematic bisection could well serve a programmatic purpose—reflecting the fragmented psyche, conflicted persona, and general estrangement of the Symphony's protagonist (perhaps Tchaikovsky himself). This effect would undoubtedly be heightened by the antiphonal orientation of the violins within the orchestra's string section. Interestingly, when the theme is later reprised (beginning at m. 90), it is statically stationed in the first violins. Perhaps this "spatial stabilization" of the theme signifies the protagonist finding peace and resolution prior to death.

²⁹ "Loud," "big," and "close" all share a close correspondence: close objects loom large in our visual field, and the perceived loudness of a sound increases as its source draws nearer. Furthermore, large objects are generally capable of producing louder sounds than are small objects.

³⁰ Boccherini derived his twelve guitar quintets, only eight of which are extant, from his own piano quintets.

exact midpoint of the movement, after which a gradual diminuendo returns the music to its original soft dynamic level by the final bar. Michael Fink remarks that the effect induced by this movement is that "of a military night-watch approaching from a great distance, passing in review, and moving on until it vanishes in the night."³¹ Obviously, the impressions of "distance," "approaching," "passing," "moving," and "vanishing" are purely metaphorical— imparted through the carefully controlled implementation of dynamics. The performing ensemble, being stationary, does not actually approach and recede from the listener. It is our embodied experience that distant sounds are quieter than close ones that enables the analogical association of volume level with distance. More shall be said on this matter in subsequent chapters.

Hector Berlioz relied on spatial effects in several of his works. The *Requiem* (1837), written in a quasi-somber, contrapuntal style harking back to Renaissance and Baroque church music, employs polychoral activity between singers and instrumental groups alike—particularly between the four brass groups meant to be separated from one another within the performance venue. As assessed by Robert Erickson, the "Tuba Mirum" section of the *Requiem* features the "definite and effective use of the positioning of the sound sources for their directionality. Beyond that there is the effect of spatial enlargement, which stems from the ambiguity for the listener of the actual position of any particular groups—because Berlioz composes the music both within and across groups."³² Berlioz's *Te Deum* of 1849, composed twelve years after the *Requiem*, likewise incorporates antiphonal spatial textures, this time toward an explicitly programmatic end: Berlioz employs an organ to contrast the orchestral sonorities, and he states

³¹ Michael Fink, liner notes to *Boccherini: The Guitar Quintets*, Pepe Romero and the Academy of St. Martin in the Fields Chamber Ensemble, Philips 438 769-2, compact disc.

³² Robert Erickson, *Sound Structure in Music* (Berkeley, Los Angeles, and London: The University of California Press, 1975), 146.

that the organ and orchestra should function as "Pope and Emperor, speaking in dialogue from opposite ends of the nave."³³ Berlioz does not restrict his application of spatial contrasts to choral works of a sacred tone. The third movement of the *Symphonie Fantastique*, entitled *scène aus champs* ("scene in the country"), calls for the offstage, or "behind the scene" ("derrière la scène"), placement of the oboe and English horn, which engage one another in a pastoral, call-and-response melodic dialogue meant to evoke the distant piping of shepherds. The use of offstage performers to create the illusion of distance—typically for programmatic, narrative purposes—is perhaps the most pronounced spatial effect utilized during the Romantic period.

Like Berlioz before him, Gustav Mahler experimented with offstage musicians. In the final movement of his *Symphony No. 2* (1893–1894), an offstage military band—consisting of four trumpets, four horns, triangle, cymbal, and bass drum—appears in the development section (mm. 343–379) as well as in the section directly preceding the entrance of the Resurrection chorus (mm. 448–471). While Mahler does not provide a detailed placement scheme for the offstage band in the score, he does indicate that the four trumpets must sound from different directions.³⁴ At times the trumpets sound in alternation, and at times they play simultaneously. Therefore, Mahler is concerned not only with creating the illusion of distance but also with manipulating the azimuth of the sound emerging from that distance. During its initial presentation, the offstage band's military fanfares are superimposed on a layer of music sounding in the string section. The juxtaposition of the two contrasting musical layers is clarified by the spatial separation of the performers. The three distinct entrances of the offstage military band (m. 343, 355, and 376) feature a perpetual increase in dynamic level—as if the apparently mobile

³³ David Cairns, liner notes to *Berloz: Te Deum*, The London Symphony Orchestra & Chorus, Colin Davis, conductor, Philips 839 790 LY, LP.

³⁴ Gustav Mahler, Symphonie II, revised version (Vienna and London: Universal Edition, 1971), 175.

marching band is approaching the stationary orchestra from a distance.³⁵ When the sound of the offstage band returns in bar 448, its dynamic level remains suppressed. It actually fades away—as if the band is receding into the distance. Mahler has effectively varied the perceived distance and directionality of the offstage (and thereby fundamentally remote) instrumental activity. Clearly, Mahler uses the space to enhance the programmatic element of his symphony: the trumpets serve to herald the apocalypse, which immediately precedes the Resurrection.³⁶ However, even if the listener fails to connect the sonic activity of the offstage band to the onset of the apocalypse (realizing on a conscious level the programmatic implications of the offstage music), that listener would likely recognize the offstage music as a contrast (in character as well as in space) to the onstage musical events and a signifier of something "otherworldly" and different.

1.3 Spatialization in the Twentieth Century

Like Mahler in his second symphony, Charles Ives often called for the spatial separation of performing forces in an effort to facilitate the differentiation of two or more simultaneous and relatively independent—layers of sound. Henry Cowell has suggested that Ives was inspired to layer separate groups of sound from having heard "two bands passing each other on the

³⁵ In live performance, the apparent mobility of the offstage marching band is compounded by the fact that it is not visible to the audience. An increasing dynamic level creates the illusion that the sound-source collective is approaching from afar. Yet, its remaining offstage (as a mark of its true immobility) ensures that the military band will never be fully "here" in the sense that the onstage orchestra is "here" (i.e., *in* the performance space proper). Using dynamics alone to give the impression of depth or distance is an effective but imperfect (imprecise) process. As I shall discuss further in Chapter 2, distant sounds are perceived as such because they are quieter, distorted (with high frequency pitches and upper partials attenuated), and often feature a higher proportion of reverberant to direct signals compared to more proximate sounds. Merely adjusting the dynamic level of a sound produced by a static source affects only one of these parameters. Nevertheless, the associability between volume level and perceived distance in musical works is quite pronounced.

³⁶ Donald Mitchell, *Gustav Mahler: The Wunderhorn Years. Chronicles and Commentaries* (London: Faber & Faber, 1975), 183–184.

march, each playing a different piece."³⁷ Consequently, Ives's music frequently consists of two (or more) functionally distinct and physically separated instrumental ensembles-each performing music with self-sufficient harmonic, melodic, and rhythmic features. Were the two ensembles not physically segregated in space, the cacophonic character of the total music would be augmented.³⁸ In *The Unanswered Ouestion* (1908), the string section, representing the "silence of the druids," is to be placed offstage in spatial contrast to the onstage solo trumpet and quartet of flutes.³⁹ The offstage and visibly unobservable strings produce the audible effect of spatial distance-serving as intermittent punctuations of "silence" following the woodwinds' failed "answers" to the persistent queries of the trumpet. The woodwinds, growing ever more frustrated by their inability to answer satisfactorily the trumpet's reiterated question, ultimately fall silent after each of six attempts. The solemn sonic backdrop cast by the remote and muted strings functions as the "awkward silence" between the fruitless exchanges of a foundering musical dialogue between the trumpet and woodwinds. In his foreword to the score, Ives states that after the woodwinds resign in futility, the question is "asked for the [seventh and] last time, and 'The Silences' are heard beyond in 'Undisturbed Solitude.'"⁴⁰ The temporal space of the loosely measured "silence" harmonizes with the spatial dimension of distance.

³⁷ Henry Cowell, "Charles E. Ives," in *American Composers on American Music: A Symposium*, ed. Henry Cowell (Stanford: Stanford University Press, 1933), 134–135.

³⁸ In the music of Ives, the divided ensembles are often *not* separated within the performance space but instead are demarcated by registration, which amounts to separation in pitch space. The second movement of his *Fourth Symphony* is such a case (Harley, "Space and Spatialization," 127–128). Ives likewise supplied only general instructions as to the physical separation of instrumental groups. He viewed the spatial distribution of musical forces in performance as an "aspect of interpretation" and therefore leaves it somewhat open (Erickson, 150).

³⁹ Glenn Watkins, *Soundings: Music in the Twentieth Century* (London: Schirmer Books, 1988), 438.

⁴⁰ Charles E. Ives, *The Unanswered Question for Chamber Orchestra or Chamber Ensemble* (New York: Southern Music Publishing Co. Inc., 1953), 2. Emphasis mine.

Ives sought to free music from being strictly a foreground phenomenon and to give it a background layer of depth through the superposition of dissimilar musical materials.⁴¹ As Peter Burkholder notes, the "background" layer of many works by Ives, which is often dynamically and texturally subordinate to the more prominent "foreground" layer, often evokes either "noises of the environment that one may notice or ignore but are nonetheless always there" or "a background hum...it is the cloud of memory, as each remembered event, person, or thing recalls others aroused involuntarily by their association with or resemblance to the first."⁴² In *The* Unanswered Question, the strings provide the humming background of silence upon which the questioning trumpet and the attempting-to-answer winds-as foreground elements-are superposed in a call-and-response (question-answer) dialogue. Ives's use of spatially separated sound sources and sound-source groups to clarify superimposed layers of contrasting musical material will have a profound impact on later "spatial composers," most notably Henry Brant and Pierre Boulez. Harley states: "According to Ives, the musical coexistence of spatially separated layers stimulates a change in perception by giving the audience the opportunity to focus on individually selected musical strata."43 In the words of Ives himself, "in any music based to some extent on more than one or two rhythmic, melodic, [and/or] harmonic schemes, the hearer has a rather active part to play."⁴⁴ In other words, through active (or, selective) listening, the listener may alternately focus upon different strata of the total music—upon either the

⁴¹ Ives's conception of "foreground" and "background" differs from that of Heinrich Schenker. In Ives's music, the foreground and the background are both realized on the musical surface—it is the surface itself that may be stratified to reveal conflicting layers of musical material. In Schenker's theory, the background does *not* abide at the surface of a musical work, although foreground and background levels are structurally connected via middleground level(s). (Middleground and foreground levels are "composed-out" from the fundamental background structure).

⁴² Peter J. Burkholder, *All Made of Tunes: Charles Ives and the Uses of Musical Borrowing* (New Haven: Yale University Press, 1995), 380.

 ⁴³ Maria Anna Harley, "An American in Space: Henry Brant's 'Spatial Music," *American Music* 15.1 (1997), 75.
 ⁴⁴ Charles Ives, "Music and Its Future," in *Source Readings in Music History*, ed. Oliver Strunk, (New York and London: W. W. Norton & Company, 1998), 1338.

foreground or the background.⁴⁵ The music is thereby akin to mobile architecture: the perspective of the musical work is variable based on the will of the listener, which is perhaps prompted and informed by salient musical cues.

Henry Brant was heavily influenced by Charles Ives's experiments with the spatial dispersal of performers. The spatial separation of sound sources is a distinctive characteristic of nearly all of Brant's compositional output. Brant outlines his spatial aesthetic in an article from 1967 entitled "Space as an Essential Aspect of Musical Composition."⁴⁶ In this article, the composer describes many practical observations made during his many years of experimenting with spatial effects in music. Brant points out that the harnessing of physical space as a compositional resource is considered "optional" and ancillary by most composers, and he emphasizes that in his own music, "the spatial distribution of the performers throughout the hall is a planned, required, and essential element of the music."⁴⁷

One of Brant's primary arguments is that the spatial separation of sound sources may be used to "disentangle" dense textures.⁴⁸ In short, the unwanted or inharmonious blending of tones may be thwarted by the spatial diffusion of performers. He equates the spatial distribution of sound sources to the composing of multiple melodic parts that do not overlap in octave range— essentially conceiving of sectors in actual space as discrete segments of pitch space. However, with the proper spatial separation, the pitch activity between multiple instruments can overlap

⁴⁵ The concepts of active listening and "auditory focus" shall be explored in Chapter 5.

⁴⁶ Henry Brant, "Space as an Essential Aspect of Musical Composition," in *Contemporary Composers on Contemporary Music*, expanded edition, ed. Elliot Schwartz and Barney Childs (New York: Da Capo Press, 1998), 223–242. Henry Brant's concern with the spatial separation of musical forces was first expressed in: Henry Brant, "The Uses of Antiphonal Distribution and Polyphony of Tempi in Composing," *American Composer's Alliance Bulletin* 4.3 (1955): 13–15. Harley notes that the 1955 article predates writings about spatialization by both Karlheinz Stockhausen and John Cage (Harley, "Space and Spatialization," 124).

⁴⁷ Brant, "Space as an Essential Aspect," 223.

⁴⁸ Ibid, 224.

and "collide" in pitch space without occluding textural clarity. In essence, multiple textures are superimposed in time but separated in space.

Brant points out that ensemble difficulties are likely to arise from dispersing members at great distances from one another. He notes that spatial separation "will cause a *deterioration* in the rhythmic co-ordination (keeping together) of the separated groups."⁴⁹ In addition. performers may have trouble intuitively integrating their own parts into the composite texture since it is more difficult to gauge their own dynamic level in relation to distantly oriented, fellow ensemble members. Brant contends, however, that "musicians accustom themselves to this fact [of being unable to hear all the parts] with little difficulty and do not find performance disagreeable under such a condition."⁵⁰ Furthermore, Brant stresses that "spatial arrangements may not be considered optional, but must be carried out in accordance with an exactly specified plan that still allows for some practical adjustment in detail."⁵¹ Throughout the twentieth century, composers increasingly began to indicate in their scores the precise layout of instruments within a performing ensemble, particularly in works featuring an unorthodox combination of instruments (i.e., a nonstandard ensemble).⁵² This practice might still reflect a concern with ensemble balance and blend, but it often betrays the composer's preoccupation with the spatial design of his/her music. When the composer specifies a seating arrangement in the score, such indications should be treated as integral to the work being preformed and should be observed by the performers just as they would observe notated pitches and rhythms.

⁴⁹ Ibid., 233. Emphasis Original. One of Brant's solutions to the problem of ensemble coordination is to write music in which "no exact rhythmic correspondence is intended between the separated groups" (Brant, 234). ⁵⁰ Ibid., 230.

⁵¹ Ibid., 224.

⁵² The act of explicitly indicating the exact distribution of instruments for a performance of his musical works sets Brant apart from Ives. As noted, Ives's instructions for spatial separation are very general and open to interpretation. Brant leaves little room for "spatial interpretation" of his music-at least on the part of the director and/or ensemble members. How listeners and theorists might interpret his (or any other composer's) spatial designs is the underlying subject of this dissertation.

In addition to enhancing textural clarity, Brant observes other desirable consequences of spatially segregating performing forces: an overall increase in volume and resonance occurs, the heights of pitches (their locations and relations in pitch space) come "vividly" into relief, parts generally sound more balanced, and "contrapuntal amalgams, even in the most complex places, [become] easily clear, [with] individual parts easily identifiable by direction."⁵³ In short, the distinct musical parts are more effectively individuated when their instrumental sources are assigned their own spatial sector of the performance venue.

Brant describes a phenomenon that he refers to as "spill," wherein qualitatively similar sounds that are separated in space seem to extend toward one another—perhaps "meeting" at an intermediate point. He describes his own musical experience of the "Tuba Mirum" of Berlioz's *Requiem*, in which each of the four brass ensembles were placed in the four corners of a balcony.⁵⁴ By virtue of the four ensembles "present[ing] a common tone-quality" and "participating in a common harmonic texture, they seem to reach out to each other when all are sounding, to extend the brass-harmonic texture continuously over the *entire* balcony area, not merely confining it to the corners of the balcony where the sounds originate."⁵⁵ Brant's conception of "spill" corresponds to the "effect of spatial enlargement" described by Erickson in regard to the "Tuba Mirum." It is easy to imagine the illusion of sounds in motion arising from the "spill" between instruments of like timbre and similar musical activity. By contrast, according to Brant, no spill whatsoever will result from spatially scattered instruments or instrumental groups of contrasting tone color. Brant refers to the resultant spatial texture in such situations as "non-spill counterpoint," and judging from his own compositional practice, he

⁵³ Brant, "Space as an Essential Aspect," 228.

 $^{^{54}}$ In some performances of the *Requiem*, the four brass ensembles are simply placed in the four corners of the concert stage.

⁵⁵ Brant, "Space as an Essential Aspect," 232. Emphasis Original.

prefers this type of spatial ensemble configuration to one in which kindred or identical instruments are separated from one another in space.

While principally discussing the spatial dispersal of sound sources as a means of ensuring textural clarity and promoting spatial counterpoint, Brant makes some mention of *moving* sonic events. He states: "When instruments placed in fixed positions [adjacent to one another] begin playing one at a time and *accumulating* (staying in)...there is a compelling impression of the hall tangibly *filling* up with sound [and also] of the sound *traveling* gradually."⁵⁶ Brant does not indicate that the instruments engaged in this consecutive and accumulative activity need be identical, nor does he state that they need be producing similar sounds to induce the impression of motion, or "travel." Brant observes further that the "impression of moving direction becomes less well defined as the further entrances and accumulations occur."⁵⁷ Since Brant credits the progressive accumulation of sounds in space with compromising the perception of motion, we may conjecture that were the instruments to play in ordered succession without "staying in," the sense of motion would remain quite vivid. Brant goes on to describe a number of "experiments" conducted in December 1964 to explore sonic motion occasioned by moving performers.⁵⁸ Brant catalogs his findings—which essentially confirm the obvious fact that a listener can aurally track a mobile sound source—but neither applies them to the observation of trends in musical practice nor suggests their prospective implementation in compositional practice.

Up to this point, the discussion of spatialized musical forces has been centered primarily upon ensembles spatially segregated so as to enhance a work's overall sonority and distinguish different parts of the composite musical texture. From the polychoral concertato motets of the sixteenth and seventeenth century to the experimental spatial compositions of Charles Ives and

 ⁵⁶ Ibid., 238. Emphasis Original.
 ⁵⁷ Ibid.

⁵⁸ Ibid., 239–241.

Henry Brant, the primary concern has been the differentiation of musical parts achievable through the spatial separation of musical forces. Although I shall argue that it has always been an emergent element of music for more than one performer, it is not until the mid twentieth century that composers actively and deliberately began to instill their musical works with spatial *motion*. Thus far, the portrayal of spatial motion has been purely metaphorical—as offstage instruments create a sense of remoteness and fluctuating dynamics engender the illusory perception of variable sound-source distance. Many twentieth-century composers began to concern themselves with the actual motion of sound sources or, more commonly, with the composite motions and spatial shapes arising from the consecutive activity of multiple sound sources.

One of Karlheinz Stockhausen's major contributions to contemporary music is giving sounds directed motion within the sound field, whether that field be an acoustic or electronic (or electro-acoustic) sonic environment. Henry Brant makes note of the phenomenon of "travel" (as described above), but he does not exploit it dramatically in his compositional output. Stockhausen, on the other hand, carefully coordinates both the location and the trajectory of sounds. In the 1950s, as Stockhausen was beginning to codify his aesthetic of spatialization, he explicitly discounted early spatial experiments by the likes of Gabrieli, Mozart, and Berlioz—all previously discussed in this chapter—as irrelevant to the future of spatial composition.⁵⁹ The spatial compositions of these earlier composers, as well as those of Charles Ives, were the primary influences upon Henry Brant's conception and style of spatialization. The lineage of spatial aesthetic spanning (and developing) from Gabrieli to Ives, and ultimately to Brant, maintains a concern with using the spatial dispersal of performers to clarify the musical texture. Although Stockhausen is perhaps mistaken to dismiss so hastily the formative influence of such

⁵⁹ Harley, "Space and Spatialization," 152.

ideas and practices, he does indeed break from the static dissemination of sound sources as he embarks upon purposefully conferring motion to sounds-with and without mobile sound sources. This is not to imply, however, that spatial gestures—dynamic spatial shapes and directed motion through an ensemble-do not occur in music composed prior to Stockhausen's spatial compositions. While Stockhausen ultimately embraces imparting directed motion to sounds, he does share Brant's concern with the use of spatial differentiation to establish and maintain textural clarity. Early in his career, Stockhausen was heavily engaged with serial composition, and he viewed spatialization as a means of articulating the easily cluttered surface of serial works. Spatialization was a means of stratifying a complex sonic design into distinguishable layers, making it "possible to articulate longer pointillistic structures by having them wander in space, by moving them from one place to another."60

In addition to using space to clarify a serial texture, Stockhausen experimented with serializing the location of sounds. He describes assigning each of the twelve chromatic pitches to a region of musical space-thus intertwining the serialization of pitch with that of spatial location, creating "a kind of 'space-melody' which evolves in pitch [space] and in [actual] space simultaneously."⁶¹ By adding spatial location as a parameter to the technique of integral serialism, Stockhausen is able to "relate proportions of pitch, duration, timbre, and loudness with those of tone-locality."⁶² Stockhausen applies this concept to one-dimensional ensembles (i.e., a single "line" of performers) as well as to ensembles encircling an audience. In this latter context, the multi-dimensional circularity of space differs from the other serialized parameters of music (pitch, duration, loudness, and timbre)—the contrasts of which persist on a one-dimensional

⁶⁰ Karlheinz Stockhausen, "Music in Space," trans. Ruth Koenig, *Die Reihe* 5, English ed. (1961): 70.
⁶¹ Harley, "Space and Spatialization," 154.
⁶² Stockhausen, "Music in Space," 79.

scale (low to high pitch, short to long duration, soft to loud dynamic, and dark to bright tone color).⁶³

At another point early in his career, Stockhausen devised a circle divided into segments (or, intervals) that form a "scale of localities."⁶⁴ Harley rejects this spatial scale as "a theoretical construct without a perceptual basis," on the grounds that "the acuity of the perception of sound direction differs depending on the orientation of the listener."⁶⁵ This observation is quite accurate: as shall be expounded upon in Chapter 2, localization is dramatically more accurate for sounds that occur directly in front of the listener. However, Harley's criticism is not without its own problems. Most listeners do not hear equally well in all frequency ranges: there is often a band of frequencies in which a listener's auditory response is considerably weakened (this impaired frequency range will vary from one listener to the next). Therefore, under such conditions, one or more components of a musical scale or tone row might be less salient, or altogether imperceptible, depending upon registration. The potentially reduced audibility of certain pitches does not constrain their musical function. Furthermore, localization accuracy can vary drastically depending upon the frequency and spectrum of the sound being heard: sounds in the midrange of human hearing-the portion of the frequency spectrum to which humans are ordinarily most sensitive-are actually the most difficult to localize precisely, even if their sources are located directly in front of the listener. Perhaps Stockhausen's circular "scale of localities" should not be dismissed as musically inconsequential on the sole basis of human localization capacity.

Within his spatial structures, Stockhausen explores the serialization of location as well as the directed motion of sounds. For instance, in his circular constructs, he "concludes that it is

⁶³ Ibid. ⁶⁴ Ibid., 82.

⁶⁵ Harley, "Space and Spatialization," 155.

possible to use continuous sound motion along the circumference of the circle."⁶⁶ He often combines the concepts of spatial differentiation (to clarify components within a dense texture), the serialization of sonic location, and the movement (actual or virtual) of sounds through a space. One such work to achieve this synthesis of spatial techniques is Gruppen für Drei Orchester (1955–1957). The composer describes his approach to this work: "The similarity of the scoring of the three orchestras resulted from the requirement that sound-groups should be made to wander in space from one sounding body to another and at the same time split up similar sound-structures: each orchestra was supposed to call to the others and to give answer or echo."⁶⁷ Interestingly, Harley observes that Stockhausen's dismissal of Gabrieli and Mozart's spatial structures is based on his assessment of "Gabrieli's polychorality as a mere expansion of the principle of dialogue into space" and his criticism of Mozart's Serenades derives from "their exclusive use of the baroque echo principle."68 Not only does Stockhausen employ the term "echo" in his description of Gruppen, but the notion of "call" and "answer" (read: "call and response") implies an antiphonal musical dialogue. Despite this inconsistency between the composer's words and his practice, Stockhausen's allowance of "sound-groups" to "wander in space from one sounding body to another" marks his fresh take on, and original contribution to, spatialization as a compositional procedure. Interestingly, and as noted above, Henry Brant contends that instruments must be physically separated on the basis of timbre: like instruments should not be located at different points throughout the concert hall. Such practice results in "spill," which obstructs spatio-textural differentiation. Brant is critical of Stockhausen's Gruppen für Drei Orchester because each of the three orchestras contain brass, woodwinds, and

⁶⁶ Ibid.

⁶⁷ Stockhausen, "Music in Space," 70.

⁶⁸ Harley, "Space and Spatialization," 152.

percussion, such that instrumental families are not discretely segregated in space.⁶⁹ Apparently,

however, Stockhausen relies on Brant's notion of "spill" to produce the illusory, continuous

movement of sound. A specific example of virtual sound movement in Gruppen shall be

explored in Chapter 3.

Erickson summarizes some of the aesthetic similarities and distinctions among the composers discussed thus far:

Ives, Brant, Stockhausen, and Berlioz all make the point that by the proper positioning of instruments and instrumental groups the composer can delineate, separate, and clarify relationships between textures and their component sound elements. Brant's work was composed for a specific performing space; Berlioz and Stockhausen composed somewhat independently of any particular space; and Ives, while keenly aware of its effects, considered the management of the performing space part of the interpreter's task. Brant composed the vertical spatial dimension of pitch, Stockhausen asked for movement (wandering) of the sound between groups, and Ives drew attention to the effect of distance.⁷⁰

Other composers in the twentieth century have profoundly impacted the developing

practice of spatialization, among them Edgard Varèse, Pierre Boulez, Iannis Xenakis, and

Luciano Berio, to name just a few. I have focused primarily upon the spatial aesthetics of Henry

Brant and Karlheinz Stockhausen due to the originality of their contributions-as novel

extensions of preceding spatial practices and as formulated explicitly in both sound and prose.

These two composers have likewise exerted an appreciable influence upon both their

contemporaries and succeeding generations of composers.

Thus far, the survey of spatial practice has revolved solely around acoustic music. In the final section of this chapter, we shall explore the paramount influence of electronic music on the evolution and expansion of spatialization techniques and concepts.

⁶⁹ Ibid., 133–134.

⁷⁰ Erickson, 150–151.

<u>1.4 Electronic Music</u>

The employment of spatialization as a compositional device was extended by the midcentury advent and proliferation of *musique concrète* (tape music) and *elektronische Musik* (electronic music). Although early experimentation with recording technology was restricted to monophonic sound reproduction, two-channel stereophony became widespread in the 1940s.⁷¹ The mainstream popularization of stereo sound in the early-to-mid twentieth century was preceded by decades of research into multi-channel audio. Clement Adler experimented with multi-channel sound as early as 1881 when he used "multiple sets of telephone transmitters and receivers to relay the sound from remote events."⁷² In the early 1920s, Bell Telephone Laboratories began to work with multi-channel recording techniques. An early experiment at Bell Labs involved placing a "curtain" of microphones before a sound source, such that each microphone would record on a separate audio track. The separate tracks were then synchronized to play the recorded audio over separate loudspeakers. Such techniques paved the way for the earliest commercially available stereo recordings.⁷³

With major improvements in audio reproduction, composers began to see such technology as the means to a creative and artistic end. Several factors lured composers to the electronic medium: one was the contemporary infatuation with timbre (tone color) and texture, already harnessed in numerous acoustic works and systematically exploited in Schoenberg's approach to *Klangfarbenmelodie*.⁷⁴ The serialists were likewise attracted to the electronic medium: it granted them the ability to project highly nuanced, serialized structures that could

⁷¹ Joel Chadabe, *Electronic Sound: The Past and Promise of Electronic Sound* (Upper Saddle River, NJ: Prentice Hall, Inc., 1997), 15.

⁷² David G. Malham and Anthony Myatt, "3-D Sound Spatialization Using Ambisonic Techniques," *Computer Music Journal* 19.4 (1995): 58.

⁷³ Ibid.

⁷⁴ Chadabe, 21–62.
exist in concrete form and thereby be realized "correctly"—without the potential for human error and imperfection that invariably accompanies live performance. Electronic music quickly became a hotbed for practitioners of integral serialism, including Milton Babbitt, Stockhausen, and Pierre Boulez.⁷⁵ M. J. Grant states: "In exploring the boundaries of each [musical] dimension, Boulez promotes spatialisation not as a possibility, but [as] a structural necessity of a music which is not transmitted via a static performer."⁷⁶ Boulez felt that the composer was obligated to use the spatial diffusion of sounds afforded by the electronic medium as a compositional element.

Grant goes on to observe that "spatial transmission is one of the great innovations of electronic music."⁷⁷ Indeed, spatializing musical events within the sonic field has been a vital aspect of electronic music since its infancy, when its practitioners almost immediately "became interested in the possibility of using sound positioning and movement compositionally."⁷⁸ Chadabe states: "At its most sophisticated, the idea of spatialization was to allow a composer to pinpoint the location from which a sound seemed to originate in a listening space and to trace the way it seemed to move."⁷⁹ The technical and artistic means by which this idea is reified and implemented vary widely throughout the latter half of the twentieth century.

Pierre Henry was among the first to spatialize electronic sounds in the 1950s. He experimented with a device built in 1951 by Jacques Poullin (but suggested by Pierre Schaeffer) called the *potentiomètre d'espace*, which utilized induction coils to manipulate the spatial distribution of sounds.⁸⁰ Poullin likewise describes two types of "sound projection" feasible with

⁷⁵ Ibid., 36–39.

⁷⁶ M. J. Grant, *Serial Music, Serial Aesthetics: Compositional Theory in Post-War Europe* (Cambridge and New York: Cambridge University Press, 2001), 99.

⁷⁷ Ibid.

⁷⁸ Malham and Myatt, 59.

⁷⁹ Chadabe, 133.

⁸⁰ Some sources refer to this device as the *pupitre d' espace*. See Chadabe, 31–32.

loudspeakers (during the broadcast of musique concrète): "statisches Relief" (static relief) and "kinematisches Relief" (kinematic relief). Static relief entails "the projection of distinct, simultaneous parts of a composition from different points in space, made possible by multi-track recordings and multi-channel sound projection systems."⁸¹ This form of sound projection is akin to Brant's utilization of spatial separation to distinguish the parts and clarify the texture of his acoustic works. Kinematic relief involves manipulating the trajectory of sounds between loudspeakers and corresponds to Stockhausen's practice of mobilizing both acoustic and electronic sounds. Harley notes: "The experience of hearing spectacular movements of sound images in space during the concerts of *musique concrète* in the early 1950s played a formative role in the development of Stockhausen's and Boulez's theories of spatialization."⁸² Like Poullin, Boulez differentiates between what he terms "fixed distribution" (or, "static relief") and "mobile distribution" (or, "dynamic relief").⁸³ Boulez proposes that mobile distribution (in both acoustic and electronic music) could involve either "conjunct" or "disjunct" movement, the distinction between which Harley states is "not dependent on distance, but on the temporal overlapping of sounds with common features in the domain of pitch, timbre, dynamics and duration."⁸⁴ Conjunct intervals occur when two spatially discrete (yet qualitatively similar) sounds overlap one another in time, such that the second sound begins before the first one ends. Disjunct intervals occur when a short pause separates two sounds ensuing from different locations. If a large time delay separates two sounds, they are perceived as two distinct events and not as a "spatial interval" at all.⁸⁵

⁸¹ Harley, "Space and Spatialization," 145.

⁸² Ibid.

⁸³ Pierre Boulez, *Boulez on Music Today*, trans. Susan Bradshaw and Richard Rodney Bennett (Cambridge: Harvard University Press, 1971), 68.

⁸⁴ Maria Anna Harley, "Spatiality of Sound and Stream Segregation in Twentieth-Century Music," *Organised Sound* 3.2 (1999): 155.

⁸⁵ Boulez, 68–69.

One of the first highly successful multichannel works, Stockhausen's *Gesang der Jünglinge* (1956) made effective use of the space "surrounding" the audience—both immediate and distant.⁸⁶ According to Stockhausen, this work was his first attempt to make "the direction and movement of sound in space" evident as a "new dimension for musical experience."⁸⁷ Other important early electronic works include: *Poème Electronique* (1958) by Edgard Varèse, which was transmitted over approximately 400 loudspeakers in the Philips Pavilion at the Brussels World's Fair in 1958, and Stockhausen's *Spiral* (1969), which necessitated the construction of a spherical auditorium at the Osaka World Fair in 1970. During performances of *Spiral*, the audience sat upon a transparent platform in the middle of the sphere, and Stockhausen controlled the projection of sound over 50 loudspeakers dispersed throughout the sphere.⁸⁸

In the 1970s, John Chowning worked diligently to design computer algorithms that accurately control and modify the azimuth (horizontal and vertical angle of displacement), distance, and velocity (speed and directionality of motion) of a sound within a quadraphonic environment. While endeavoring to simulate a wide array of sounds in a variety of replicated acoustical spaces, Chowning likewise concerned himself with simulating the actual movement of those sounds within the illusory spaces. Chowning states:

A special program has been written which allows the user to specify an arbitrary *path* in a two-dimensional space [one exclusive of elevation] by means of a light pen or a computed geometry. The program evaluates the trajectory and then derives the time functions which control distance and angle for simulation...Using interactive graphic display techniques, a user can specify the location and movement (trajectory) of a

⁸⁶ Chadabe, 39. Stockhausen originally conceived *Gesang* in five tracks: four speakers to be arranged in quadraphonic fashion (encircling the audience) and one speaker to be suspended above the audience. Unfortunately, five-track playback machines did not yet exist at the time of the work's premiere in 1956. After attempting to coordinate two separate tape machines (one to play tracks one through four and another to play the fifth track), and encountering insurmountable difficulties with synchronization, Stockhausen eventually mixed the fifth track with the fourth track, and this quadraphonic version of *Gesang* remains the most frequently performed version.

⁸⁷ Stockhausen, "Music in Space," 68.

⁸⁸ Watkins, 591–594.

sound...[the] program computes the control functions for azimuth, distance, and velocity which are [then] used to modulate the signal to be applied to the loudspeakers.⁸⁹

Chowning's program likewise factors in cues to mimic the *Doppler effect*, which is the phenomenal shift in frequency experienced as a moving sound source passes the listener. When a sound-emitting source approaches and passes the listener, the subject experiences a decrease in the sound's pitch. The amount of decrease is dependent upon (and directly proportional to) the speed of the traveling source: the higher the speed, the greater the change in frequency. Thus, the extent of a frequency shift provides a cue for gauging the speed of a moving sound source. Chowning states: "This frequency shift has been found to be an essential cue in the simulation of moving sources."90 Chowning's algorithms can independently control cues for room size and signal location/movement within that room. Thus, not only is the location of a sound adjustable, but so is the size and shape of the virtual room in which the sound occurs. The artificial space can be made to "sound" larger than the actual room in which the speakers are placed, and sounds can seemingly occur beyond the confines of an acoustically established real or virtual space. Regarding Chowning's accomplishments, Erickson writes: "The possibilities in composing motion are enormous, and go far beyond the 'wandering' of Stockhausen's Gruppen...it is possible for a composer to create in the studio a moving sound object and the illusory space in which it moves."91

Chowning's work is important in part because he achieves remarkable spatial effects with only four speakers. Many other composers and sound engineers have developed extensive diffusion systems involving multiple loudspeakers. Jonty Harrison, famous for his work with

 ⁸⁹ John M. Chowning, John M. Grey, Loren Rush, and James A. Moorer, "Computer Simulation of Music Instrument Tones in Reverberant Environments (grant proposal for CCRMA)" (Stanford: Stanford University, 1974), 60, 6. Emphasis mine. The concept of a "sonic path" shall be explored in subsequent chapters.
 ⁹⁰ Ibid.. 60.

⁹¹ Erickson, 152.

BEAST (Birmingham Electro-Acoustic Sound Theatre), proclaims that eight speakers (the "main eight") are "the absolute minimum for the playback of stereo tapes."⁹² In 1974, François Bayle created a "loudspeaker orchestra" called the *Acousmonium*. This construction "consisted of eighty loudspeakers of various sizes placed across a stage at different heights and distances from the proscenium."⁹³ In the late 1980s, Peter Otto and Nicola Bernardini (and others) developed a system called TRAILS (Tempo Reale Audio Interactive Location System), a "multi-loudspeaker network [consisting of] a twenty-four-by-eight spatialization matrix."⁹⁴ This configuration involves 192 speakers! The ostentation of such experiments in the electronic realm conforms— anachronistically—to that of the Roman polychoral composers (such as Benevoli) in the middle seventeenth century.

During the latter half of the twentieth century, electronic music asserted a strong influence on the composition of acoustic music. Not only did many composers employ extended techniques and unorthodox instrumental combinations in an effort to create new worlds of sound (often in conscious imitation of the exotic sound worlds forged in the electronic medium), but they also became increasingly sensitive to the spatial relationships among musical events. This increasing concern with spatiality in the acoustic realm was undoubtedly due, in large part, to the spatialization practices so common to electronic music. Shortly after composing *Gesang*, Stockhausen conceived of works such as *Gruppen* (1955-1957) for three orchestras and *Carré* (1959-1960) for four orchestras and four choruses that surround the audience. Iannis Xenakis's *Persephassa*, scored for an ensemble of six percussionists encircling the audience and featuring the circular movement of sound, seems strikingly influenced by electronic surround-sound. In

⁹² Jonty Harrison, "Sound, Space, Sculpture: Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion," *Organised Sound* 3.2 (1999): 121.

⁹³ Chadabe, 68.

⁹⁴ Ibid., 132.

some acoustical works, the performers are required to move about while producing sound. Although Berio's *Circles* (1960) is "often cited as the earliest work that requires the soloist to move from one stage position to another," Brant's *Hieroglyphics I* of 1957, which likewise calls for performer mobility, predates *Circles* by three years.⁹⁵ Brant's *Windjammer* (1969) includes "specific walking plans for the performers of the wind quintet."⁹⁶ In Thea Musgrave's *Concerto for Clarinet and Orchestra* (1969), the soloist is required to traverse the entire orchestra intermittently activating and participating in small, concertino groups. In *A Klockwrk Diurnal* (1977) for alto saxophone, French horn, and bassoon, composer Claire Polin indicates that the "players use swivel chairs which should move soundlessly."⁹⁷ The score contains performance instructions such as "swivel out in a slow circle, and in again," "rotate seat outwards," and "swivel inwards."⁹⁸ Several composers have also experimented with the circulation of loudspeakers themselves (as opposed to the circulation of sound amongst multiple speakers). Perhaps the most noteworthy attempt is Annea Lockwood's "Soundball," a flying loudspeaker devised in 1984.⁹⁹

As early as 1953, Stockhausen courted the notion of putting "musicians in chairs and swing[ing] them around."¹⁰⁰ Concerned at that time that performers might object vehemently to such stipulations, he curtailed such experimentation. However, Stockhausen would eventually produce what is arguably the most extreme experiment involving the spatial motility of live performers with his *Helikopter-Streichquartett* ("Helicopter String Quartet"), which is featured in the opera *Mittwoch* ("Wednesday") from the opera cycle *Licht*. In this work, the four

⁹⁵ Harley, "An American in Space," 81.

⁹⁶ Ibid.

⁹⁷ Claire Polin, A Klockwrk Diurnal (Dorn Productions, 1977), 1.

⁹⁸ Ibid., 2, 12.

⁹⁹ Chadabe, 131.

¹⁰⁰ Stockhausen, *Stockhausen on Music*, 101–102.

members of the string quartet are each seated in a separate helicopter. The helicopters are instructed to "circle within a radius of circa 6 km above the performance venue, individually varying their flying altitudes."¹⁰¹ The sounds of the performers, as well as those of the helicopters (rotor blades are recorded with an external contact microphone), are mixed and broadcast over loudspeakers in the auditorium where the work is being experienced. How the circulation of the musician-containing helicopters translates to the space in which the audience receives the sounds is a line of inquiry reserved for later research and analysis.

From the antiphonal chanting of sacred Psalms to airborne string quartets, the aesthetics, techniques, and implementations of spatialization in music have undergone considerable changes over the developmental course of Western musical practice. Nevertheless, the sheer concepts of sonic distribution and spatial differentiation have underpinned much compositional practice. Only in the mid twentieth century do composers begin to deliberately control the *movement* of sound, which is achieved either through sonic illusion (with static performers and/or loudspeakers) or with moving performers. As evident from the preceding discussion, the technique of spatialization in musical composition has been developed quite extensively. However, theories for analyzing spatial relationships in music are virtually non-existent. Furthermore, the sensitivity of the listener to unfolding spatial relationships in musical performance is, in general, ill refined. Perhaps there is a symbiosis between the analytic and perceptual poverty surrounding the spatial design of music. The remainder of this dissertation shall be devoted to constructing an analytic methodology for evaluating and interpreting spatial

¹⁰¹ Karlheinz Stockhausen, liner notes to *Karlheinz Stockhausen: Helikopter-Streichquartett*, trans. Suzee Stephens, The Arditti String Quartet, Montaigne MO 782097, compact disc. I was unable to determine Stockhausen's intended (and perchance, prescribed) directionality of motion; i.e., whether the circular formation of helicopters rotates in a clockwise or counterclockwise fashion (from the audience's terrestrial perspective). As I shall argue in Chapter 4, these two opposing directions have distinct associative meanings.

activity in music. In the process, I hope to encourage more active "spatial listening" on the part of the perceiver.

CHAPTER 2

SOUNDS IN SPACE: WHAT SOUNDS ARE, WHAT SOUNDS DO, AND HOW WE HEAR THEM

This chapter explores the nature of sound, which serves as the raw material of anything musical—including spatial gestures. The discussion involves theories of what sounds actually are, describes how sounds behave in the environment, and explains the process of sound localization as a significant function of human audition. Such a discussion serves to highlight the variable factors that can affect the perception of spatial gestures. Certain decisions on the part of the performers (such as the physical spacing of ensemble members on stage), the composer (specific instructions as to the physical orientation and spacing of the performers), and the listener (where he/she is seated in the concert hall relative to the performing ensemble) can affect the perceptual salience of spatial gestures. The chapter concludes with a discussion of "auditory streaming," which is applied to account for the process by which the listener organizes temporally successive and spatially discrete sounds into a unified spatial gesture.

2.1 Sounds

Sounds are generally defined as "the public objects of auditory perception."¹ But exactly what kind of thing is a sound? Or, to reframe the question: to what metaphysical category do sounds belong? Philosophers and physicists have differed in their response to this question. Three competing views of sound exist. The first is called the *property view*, which was held by

¹ Casey O'Callaghan, "Sounds" (Ph.D. diss., Princeton University, 2002), 1.

John Locke, among others. Under the property view, sounds are considered to be the secondary (or, sensible) qualities of an object. According to Locke: "Whatever reality we by mistake attribute to [sensible qualities], [they] are in truth nothing in the objects themselves, but powers to produce various sensations in us."² Secondary qualities thus arise from the "dispositions objects have to affect perceivers' experiences."³ Colors, tastes, and smells are considered secondary qualities. The perception of a color, for example, is driven by how the brain interprets a specific wavelength of light falling incident upon the retina. How light reflects from an object depends upon the reflective/absorptive qualities of that object's surface—the color we perceive is not actually *in* the object. The property view states that sounds are properties, or secondary qualities, of bodies: a material object "possesses" a sound when it vibrates at a particular frequency and amplitude. The problem with this view is that sounds do not exist in a vacuum—they are dependent upon a medium for propagation. So, an object vibrating in a vacuum does not "possess" any sound.

The second view of sound, posited by Aristotle and still favored by physicists, is called the *wave view*, which states that sounds are waves (pressure disturbances propagated through a medium) produced by vibrating bodies.⁴ This seems quite reasonable; in fact, we often equate "sound" and "wave" when we speak of a "sound wave." However, there are significant problems with this view as well; namely, it entails that a sound itself travels, as a wave, from its source to the listener. Yet, we perceive a sound to be located distally when we identify the spatial location (directionality and distance) of that sound. Sounds are heard to "be" where the

² John Locke, *An Essay Concerning Human Understanding*, complete and unabridged, collated and annotated by Alexander Campbell Fraser (New York: Dover Publications, Inc., 1959), I: 173.

³ Casey O'Callaghan, Sounds: A Philosophical Theory (Oxford: Oxford University Press, in press), 21.

⁴ While most people understand Aristotle to be a wave theorist, O'Callaghan positions Aristotle as a potential historical precedent for his "event view" of sound. When Aristotle describes sound as "a particular movement of air," he is likely referring to the act (i.e., "event") of the air being moved rather than the result of the act of moving (i.e., the moving air). See O'Callaghan, *Sounds: A Philosophical Theory*, 84–85.

events that cause them occur; measuring the locatedness of a sound is a phenomenological fact (and automatic process) of auditory perception.⁵ If the wave view is correct, then we are constantly and systematically misperceiving the location of sounds: they should always seem to be located right at or within our ears! Furthermore, under the wave view, "apparent duration perception results from encounters with the spatial boundaries of sounds."⁶ In other words, if a sound is a wave, then the *spatial* boundaries of the sound—as it travels past the listener—are phenomenally experienced as its *temporal* beginning and ending. The physical "head" of the sound would be located at our ears at the point in time when we begin to hear the sound, and the "tail" of the sound would leave our presence just as we stop hearing it. The "body" of the sound, as it passes us by, would translate to our durational experience of the sound occurring over time. Phenomenologically, we experience the temporal commencement, sustainment, and cessation of a sound—not its spatial form.

Philosopher Casey O'Callaghan offers a third view of sound. On what he calls the *event view*, a sound is "the event of an object or interacting bodies disturbing a surrounding medium in a wave-like manner."⁷ In other words, a sound is the actual event of a medium's being disturbed by a vibrating object. The vibrating object is the *cause* of the sound, and the waves that follow from this event are the *effects* of the sound. The waves merely provide information *about* the sound-event, including its duration, amplitude (loudness), frequency (pitch), timbre (spectral profile), and location. An analog from vision—the most studied and understood of all sense modalities—is instructive: when we "see" an object, we are seeing light waves and particles that

⁵ O'Callaghan debunks the notion that what we actually hear is *only* the location of a sound's source. He states: "We experience sounds, in a wide range of central cases, to be located in the neighborhood of their sources" (O'Callaghan, *Sounds: A Philosophical Theory*, 46). We do not directly hear a sound source; we hear instead the sound it produces (which is located in immediate proximity to its source). We hear the location of the source by or in hearing the location of the sound it has produced. See O'Callaghan, *Sounds: A Philosophical Theory*, Chapter 3. ⁶ O'Callaghan, *Sounds: A Philosophical Theory*, 60.

⁷ O'Callaghan, "Sounds," iii.

are reflected by the surface of the object. Light waves bring information about distal objects to our eyes (the sensory apparatuses for vision), and our brain forms perceptions (internal images) of those objects. We do not perceive the objects to be flying through space and into our eyes. So too with audition: the sounds themselves, as the objects of audition, do not travel through space to our ears; rather, the waves produced by the sounds travel through space, enter the ear canal, and are ultimately transduced into neural impulses that the brain interprets as sound images. O'Callaghan's phenomenological account of sound entails hearing sounds themselves—rather than solely the sound-producing sources—as located in the environment at some distance and in a given direction from the listener's perspective. We do not perceive a sound to be traveling past us; instead, we hear it as distally located and persisting over a specific amount of time. Perceiving the duration of a sound is, in fact, perceiving the "duration of the process of sound wave production."⁸ It is the stationary sound (assuming a motionless source) that produces (or causes) the sound waves that travel to our ears.

O'Callaghan argues that sounds are neither properties nor secondary qualities of an object. Instead, sounds are "*particulars* that possess the audible qualities of pitch, timbre, loudness, along with other inaudible properties."⁹ By "particular," O'Callaghan simply means that a sound is an individual item, as opposed to a universal quality. However, although a sound is not a property or quality in and of itself, it does possess qualities. The audible qualities of a sound are grouped together by virtue of belonging to a single particular that persists over time. The durational aspect of a sound is crucial to the conception of sound as an event: sounds endure and change qualitatively over time, and they survive qualitative changes. O'Callaghan remarks: "We intuitively think of objects, as opposed to time-taking particulars, as being wholly present at

⁸ O'Callaghan, Sounds: A Philosophical Theory, 83.

⁹ Ibid., 23. Emphasis Original.

each time at which they exist." Time-taking particulars and events have "parts that exist and take place at different times."¹⁰ The audible qualities of a sound, as well as the qualitative changes that a sound undergoes over time, are the "signatures" that identify and individuate that particular sound. The fact that a sound is persistent and mutable preempts its being characterized as an object or as the property of an object.¹¹

Viki McCabe defines an event, in general, as "a coherent and meaningful unit that has properties that persist and properties that change as the event occurs over time and space. It could be as short as a raindrop or as long as evolution."¹² This denotation dovetails with and supports O'Callaghan's sound-as-event view. In describing a *motion event*, McCabe writes: "When something moves or appears to change in any way (even by a change in the observer's perspective or through slow development over time), *it can be identified by the patterns produced by its own dynamic structure*."¹³ A spatial gesture, as I shall argue, involves motion— a *change* in the location of a sonic event.¹⁴ This does not necessarily mean that it is the location of a single sound-producing source that is transient; instead, a number of events might ensue, successively, from separate sound sources abiding at distinct points in space. The changing location of *some* event produces a dynamic pattern of motion. The identity of a spatial gesture (its *kind*, or *type*), which determines its categorical membership, is defined by the precise nature of its patterning—its composite shape and emergent directionality.

¹⁰ Ibid., 37.

¹¹ The way that sounds occupy time distinguishes them from objects. A time slice of a sound is not the entire sound. However, a time slice of an object *is* the entire object. An object exists entirely at any given moment, but a sound does not. Thus, O'Callaghan classifies sounds as "event-like particulars," as opposed to "object-like particulars" (O'Callaghan, *Sounds: A Philosophical Theory*, 41).

¹² Viki McCabe and Gerald J. Balzano, eds., *Event Cognition: An Ecological Perspective* (Hillsdale, NJ, and London: Lawrence Erlbaum Associates, 1986), 4–5 (footnote 1).

¹³ Ibid., 7. Emphasis Original.

¹⁴ Aristotle considers "motion" to be change of any kind. Thus, the concept of motion is not restricted to locomotion but refers to any changing quality. Aristotle defines motion as the "actuality of a potentiality." See "Aristotle (384– 322 BCE): Motion and its Place in Nature," *The Internet Encyclopedia of Philosophy*, available at: http://www.iep. utm.edu/a/aris-mot.htm (accessed February 15, 2007).

I adopt and apply the event view of sound since it is dependent upon the automatic processes of externalization (recognizing the sounds as "out there") and localization (identifying precisely where the sounds "are"), which enable the perception of spatial gestures in music. Both the property view and the wave view prove unsatisfactory for a theory of sounds and insufficient to serve as ground zero for a theory of spatialization: the property view discounts the necessity of a medium for the transmission of information about sounds, while the wave view entails that sounds themselves travel and does little to account for the perceived location and locomotion of a distal sound source or the impression of motion arising from the collective activity of two or more stable sound sources. The changing properties of a sound (its audible qualities), and of a spatial gesture (the physical location of its constitutive sounds), are easily accounted for by the event view. Since sounds are inherently spatial, music-comprising sounds—is likewise inherently spatial.¹⁵ In acoustic music for more than one performer, multiple sound sources coexist. Since no two discrete sound sources may occupy the same point in space at the same time, the sounds produced by each source will have different locations. The potential exists for various spatial relationships to ensue from the coordinated activity of these distinct sound sources. These relationships, which unfold over time between two or more sound sources, are spatial gestures. Sounds themselves are not phenomenologically in locomotion (although their sources may be); however, the successive activity of two or more discrete and stationary sound sources may yield a composite motion in space, i.e., the change in location of

¹⁵ By necessity, my current definition of *music* is narrow and overly simplified. I must presently ignore the sociocultural component of music. Sounds are the most fundamental component of music. *Any* sound or sound complex may be heard as "musical" if the perceiver elects to receive the sound(s) as music. This view of music is largely informed by the philosophy and aesthetic of composer John Cage, as outlined in John Cage, *Silence: Lectures and Writing by John Cage* (Hanover: Wesleyan University Press, 1961).

some sound event.¹⁶ O'Callaghan observes that "there might be complex sounds comprised of distinct sounds from a number of sources arranged over time...these...are complex events constituted by many distinct sounds."¹⁷ A spatial gesture is just such a complex event—one from which a dynamic spatial shape emerges. In other words, a spatial gesture is an event composed of smaller events.

Clearly separated sound sources are the norm for the vast majority of performed acoustic music. However, in electronic and recorded music, multiple (and distinct) sounds may be channeled to the same speaker. During playback, therefore, such sounds will actually emanate from the same location in space.¹⁸ When listening to recorded music over headphones, the intracranial process of **lateralization** replaces that of extracranial externalization. Through lateralization, we actually perceive a given sound to be located at some point along a lateral plane inside our head, extending from one ear to the other. Of course, this is a perceptual illusion: the sounds are not actually located inside our head. With these intracranial sonic images, no illusion of azimuth (external directionality) or elevation is possible.¹⁹ Nevertheless, we can still perceive different sounds to be located at different points along the lateral plane or moving from one point to another along the plane.

Despite our relatively accurate ability to externalize and localize sounds, two or more sound-producing performers in close proximity might prove difficult to differentiate solely through audition—particularly if a considerable distance exists between the performers and the

¹⁶ Under Aristotle's view of motion as any type of change, a sustained (temporally continuous) sound ensuing from a fixed point in space (a locationally stable sound source) may be said to "be in motion" if it is undergoing a change in frequency (pitch), amplitude (loudness), and/or spectral profile (timbre).

¹⁷ O'Callaghan, Sounds: A Philosophical Theory, 88.

 ¹⁸ Varying the ratio of intensity (channel proportions) of multiple sounds between two loudspeakers can create an illusory (virtual) stereophonic environment with clearly perceivable spatial separation between the distinct sounds.
 ¹⁹ Stanley A. Gelfand, *Hearing: An Introduction to Psychological and Physiological Acoustics*, 3rd ed. (New York:

Marcel Dekker, 1998), 374. The perceptual and aesthetic distinctions between encountering spatial gestures during live performance and detecting them during the playback of a recording (over headphones) lies beyond the scope of this dissertation. (The broadcasting of electronic music over loudspeakers shall be addressed in Chapter 4.)

listener (as in a large performance space). In such cases, visual cues aid in localization. Although I am primarily concerned with how we hear events in space, vision often plays an important role in the perception of spatial gestures, particularly in a live performance where the audience members can see the performer(s) producing sound.²⁰ Jens Blauert observes: "What the subject sees during sound presentation, and where the subject sees it, are factors determining the position of the auditory event."²¹ In fact, there are certain neurons in the brain that respond exclusively to dual stimulus information ("about" the same distal object) received simultaneously through vision *and* audition. Such multisensory interaction is typically referred to as **crossmodal binding**.²² Of course, electronic/recorded music generally lacks visual cues with the relatively trivial exception of the listener's ability to see where the speakers are situated.

2.2 Sounds in the Environment

We shall now explore the manner in which the waves produced by a sound behave in the environment. It shall prove beneficial first to outline the physical properties of sounds and their psychological correlates. When a sound-producing body (or sound source) disturbs an elastic medium, such as air, it causes the particles of that medium to propagate in all directions as longitudinal waves. Sound waves consist of cyclic changes in air pressure—the alternate compression and rarefaction of air molecules. It is not the individual air particles but the actual

²⁰ In his composition entitled *Rain Tree* for 3 percussion players, Toru Takemitsu coordinates the musical activity of spatially separated performers with flashing spotlights. Sonic spatial gestures are highlighted by concomitant lighting effects, and the synthesis of auditory and visual spatial cues enhances the salience of spatial gestures.
²¹ Jens Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, revised edition (Cambridge,

MA, and London: The MIT Press, 1997), 193.

²² For a discussion of crossmodal binding, see: Albert S. Bregman, Auditory Scene Analysis: The Perceptual Organization of Sound (Cambridge, MA, and London: The MIT Press, 1990), 181–184, 290–291; Richard M. Warren, Auditory Perception: A New Analysis and Synthesis (Cambridge, UK, and New York: Cambridge University Press, 1999), 189–195; Charles Spence and Jon Driver, eds., Crossmodal Space and Crossmodal Attention (Oxford: Oxford University Press, 2004); Casey O'Callaghan, "Shared Content Across Perceptual Modalities: Lessons Learned From Cross-modal Illusions," Electroneurobiología 14.2 (2006): 211–224; and O'Callaghan, Sounds: A Philosophical Theory, Chapter 11.

disturbance of air pressure that is propagated.²³ The velocity of sound-wave propagation is determined by the *elasticity* and *density* of the medium in which it travels rather than by some property of the vibrating source.²⁴ Sound waves propagate through air at a rate of approximately 340 m/sec (or 1,115 ft/sec).²⁵ The rate of vibration of a sound source corresponds to the frequency (cycles per second, measured in hertz) of the wave it produces, and frequency is perceived as pitch. The audible frequency range for human beings is approximately 20–20,000 Hz.²⁶ The human ear is most sensitive to frequencies in the 1000–5000 Hz range.²⁷ The force of a vibrating object (or, the degree of its displacement from a centralized position) equates to the amplitude of the resultant wave, and amplitude is perceived as intensity (or loudness).²⁸ A complex tone's spectral profile—the presence and relative intensities of upper partials (or overtones)—corresponds to the perceived timbre (or, "color") of that tone.

The manner in which sound waves travel through and interact with the environment greatly impacts auditory localization. As a sound wave travels through a medium, it undergoes

²³ W. Lawrence Gulick, George A. Gescheider, and Robert D. Frisina, *Hearing: Physiological Acoustics, Neural Coding, and Psychoacoustics* (New York and Oxford: Oxford University Press, 1989), 21.

²⁴ An elastic medium consists of "particles that resist displacement and that, once displaced, tend to return to the location from which they were displaced" (Gulick et al., 19). Elasticity may be thought of as "stiffness." Density is simply the degree of compactness of a substance. The speed of sound increases as elasticity increases and as density decreases. Temperature likewise affects the speed of sound: sounds travel faster in warm air due to the decreased density of the medium. Humidity actually decreases the density of air because the molecular mass of water is less than that of air, so the speed of sound increases as the index of humidity increases. Sound would travel fastest in hot, humid air and slowest in cold, dry air.

²⁵ Sound travels almost four times faster in water, and faster still in metal, due to the increased elasticity of these mediums. In general, sound travels slowest in gases, faster in liquids, and fastest in solids. Although liquid is more dense than gas, and solid more dense than liquid, the markedly increased elasticity of most liquids and solids is responsible for the increased rate of sound transmission in these mediums. In cases where elasticity is constant between varying mediums, sound will travel faster in the medium with the lower density.

²⁶ The audible frequency range varies among species. For example, canines can hear much higher sounds and elephants much lower sounds than can humans.

²⁷ Michael J. Wagner, *Introductory Musical Acoustics*, 3rd ed. (Raleigh: Contemporary Publishing Company of Raleigh, Inc., 1994), 44.

²⁸ Michael W. Levine, *Fundamentals of Sensation and Perception*, 3rd ed. (Oxford and New York: Oxford University Press, 2000), 333.

diffusion, whereby its total energy is progressively weakened.²⁹ As a sound wave travels through the air, friction causes the energy of particle disturbance to dissipate. As the distance from the sound source increases, the amplitude of the wave is reduced (causing a decrease in perceived loudness) and the quality of the sound becomes distorted. The distortion results from several factors: first, the higher frequency components of a complex sound fade faster than its lower frequency components, so the overall spectral profile of a sound is altered. Second, sound waves are subject to scattering when they encounter objects in the transmitting medium. Depending upon the properties of the object that a sound wave encounters, it may reflect from or penetrate that object. When the conducting medium and the encountered object are alike in density and elasticity, most of the sound wave will penetrate the object and continue traveling. When the medium and object are dissimilar in these respects, a greater proportion of the total sound energy will be reflected. When a sound wave penetrates a medium of varying density, such as a pocket of warm air, the wave front will bend-this is known as refraction. Wind (moving air) can also cause sound waves to refract. Such conditions can lead to misperceptions of a sound source's location.

Diffraction is the spreading out of a sound wave into spaces that are not in a direct path from the sound source. Essentially, diffraction allows a sound to "travel around corners, through open windows, and [to] seemingly 'fill [a] room."³⁰ When a sound wave diffracts into a room through a small window, it will diffuse in a spherical manner from its point of entry. Low frequency sounds diffract much more effectively than high frequency sounds; diffraction distorts a sound by producing "sound shadows," which are caused by the incomplete diffraction of a complex sound's frequency components. When a sound wave is diffracted, the listener's ability

²⁹ The term *diffusion* also refers to the dispersal of a sound's energy such that it is less direct and locationally focused. A "diffuse" sound seems to envelop a room and may therefore prove difficult to localize.
³⁰ Wagner, 47.

to precisely localize the sound source is impeded. Musical performances in which instruments are placed offstage rely on diffraction to deliver their sounds to the audience while creating a sense of distance.

When a sound is heard in a room, the waves produced by that sound reach the listener's ears from a multitude of directions due to **reflection**. When a wave front encounters a surface, only a portion of its total energy is reflected while some is absorbed and transmitted through the surface. (The proportion of absorbed to reflected energy is dependent upon the disparity of elasticity and density between the two media.) For reflecting sound waves, the angle of incidence equals the angle of reflection, and a sound's waves may rebound from multiple surfaces before dissipating completely. The frequency of a sound wave is not affected by reflection; however, the more a sound wave reflects, the more its total energy diminishes, with upper partials dissipating fastest. As such, the loudness and spectral profile of reflected sound waves typically vary from those of the initial, unreflected wave. Higher frequencies generally reflect better than lower frequencies; therefore, sounds rich in upper partials tend to sound more diffuse within a given space.

The sound wave approaching the listener directly from the sound source is the **direct signal. Early reflections** are those that arrive within approximately 30 msec of the direct sound.³¹ They may come at the listener from any direction after having reflected from the walls, ceiling, and/or floor. Since these signals arrive in such close temporal proximity to the direct signal, they are not perceived as separate sounds; rather, they provide the listener with a sense of the "space" in which the sound-events are transpiring. Early reflections approaching the listener from his/her side (after having reflected off of a sidewall) are known as **lateral reflections**, and

³¹ Early reflections are also referred to as "discrete" reflections.

these create a spatial "broadening" of the auditory event.³² Blesser and Salter observe: "Even when the total sound energy remains constant, shifting energy to the early sonic reflections enlarges the...apparent aural size of the sound source."³³ Late reflections are the reflected signals that reach the listener quite late, after the sound wave has reflected multiple times and traveled a greater distance. Late reflections are often perceived as echoes or reverberation, and they potentially degrade the intelligibility of a sound.³⁴ Blesser and Salter distinguish an echo as a "distinct sonic reflection" and reverberation as "multiple sonic reflections."³⁵ The intensity of reflected sounds generally decreases as their delay relative to the direct signal increases due to air friction and the partial absorption of the sound's energy by reflective surfaces. Furthermore, some types of surfaces absorb primarily higher frequency components while others favor the absorption of lower frequencies, thus affecting the sense of "warmth" or "brightness" of an acoustical space.³⁶

Following O'Callaghan's categorization of a sound as an event, a reflection (or echo) of a sound is not a new sound—an echo is not a sound separate from the original (or direct) sound. Rather, the sound wave, which merely transmits information about the sound, is reflected within the environment—reaching the listener's ears at different times, from multiple directions, and with varying degrees of distortion. (Sounds themselves do not travel, although their causal sources *may* travel and their resultant waves *do* travel.) So, an echo is a spatio-temporal distortion of the original sound.³⁷

³² Blauert, 282.

³³ Barry Blesser and Linda-Ruth Salter, *Spaces Speak, Are You Listening?: Experiencing Aural Architecture* (Cambridge, MA, and London: The MIT Press, 2007), 53.

³⁴ Ibid.

³⁵ Ibid., 155.

³⁶ David Butler, *The Musician's Guide to Perception and Cognition* (New York: Schirmer Books, 1992), 96–97.

³⁷ O'Callaghan, "Sounds," 21–42.

Reflections provide useful information about the space in which a sound occurs but can potentially hinder localization of the sound source, particularly in extremely "wet" (highly reverberant) halls. The ratio of direct to reflected signals in any room varies depending upon the distance between the listener and the sound source: as the distance increases, the portion of direct signal decreases in relation to that of the reflected signals. In moderately reverberant rooms, and particularly in "dry" spaces, the direct signal will be much greater in amplitude than the early and late reflections. Localization discriminations of a sound seem to be "based on the dichotic properties of the sound that arrives first, [which] has come to be known as the *precedence* effect."³⁸ In a highly reverberant acoustic environment, the proportion of direct to reflected signals is reduced such that the sound source itself seems markedly disseminated. Under such conditions, the precedence effect is challenged and potentially compromised.³⁹ The location of a single sound source may seem quite diffuse, thereby becoming difficult to localize. Since the reflective properties of a room affect the localizing aptitude of the listener, the acoustic properties of the space in which a musical work is performed have a direct bearing on the salience of spatial gestures.⁴⁰ If the acoustical qualities of a performance venue jeopardize the aural individuation of discretely stationed ensemble members, spatial gestures may become undetectable. Indecipherable spatial gestures are moot: the dynamic spatial construct of the musical performance loses its otherwise considerable ability to affect the musical experience.

³⁸ Gulick et al., 340. The precedence effect is sometimes referred to as the "law of the first wavefront" (Blauert, 222–235) or the "Haas phenomenon" (Gelfand, 387).

³⁹ When the time delay between the direct sound and its echo is less than 1 msec, the precedence effect fails and "the sound source [is] perceived to be at a location intermediate between the locations...of the primary sound and its echo" (Gulick et al., 340).

⁴⁰ For a thorough discussion of conditions affecting the "aural architecture" of a space, as well as a history of acoustic design, see Blesser and Salter, *Spaces Speak*.

2.3 Auditory Localization

Human beings are remarkably adept at discerning the location of a sound within the surrounding environment. The ability to localize the source of a sound involves hearing the directionality (azimuth and elevation), as well as the distance of the sound from the egocentric perspective of the perceiver. Localization also involves detecting the movement (velocity) of a sound source. The process by which we locate sounds within the environment is quite complex. The majority of localization discriminations are made from binaural (or interaural) cues. However, there are important monaural cues as well.

Binaural Hearing

The most accurate means of determining the directionality of a sound is through comparing the sound signals received simultaneously by both ears. With binaural (or dichotic) hearing, the disparity of level and/or phase between the signals at each ear provides for accurate localization. In a process called **binaural fusion**, the "similar but nonidentical signals reaching the two ears are fused into a single, coherent image (gestalt)."⁴¹ Due to the neural integration of the signals at each ear, we do not misperceive a single sound as two different sounds, one at each ear.⁴² Binaural fusion strengthens the argument for the event view of sound: we perceive a single sound, the effect (wave) of which reaches both ears in a slightly different manner. If the sound *were* the wave, we should "hear" two different sounds (or perceive two different instances of the same sound) occurring *at* each ear. Our perceptual system does not allow for such

⁴¹ Gelfand, 372.

⁴² Vision provides an obvious analog. Though we see with two eyes (binocular vision), we typically perceive only one of each object in the visual field. This is due to a type of neural interaction similar to binaural fusion: the inputs of the two eyes are "fused" at a higher level of processing. Misalignment of the eyes—as in *strabismus* (a "crosseyed" or "wall-eyed" subject)—and mental impairments (such as intoxication) can result in *diplopia* (doublevision). Audition does not contain such perceptual pathologies.

misperception. Additionally, we do not perceive the level or timing difference between the signals at each ear; instead, we directly perceive a sound's locatedness.

Lord Rayleigh proposed his enduring "duplex" theory of localization in 1907, which essentially states that low frequency sounds are located laterally by the **interaural time (or phase) differences (ITD)** between the two ears, while high frequency sounds are localized via **interaural level (or intensity) differences (ILD)**. Either of these differences arises when a sound occurs to the right or left of a subject's median (or midsagittal) plane.⁴³ ITDs detect the location of low frequency sounds ranging from the lowest audible frequency up to approximately 1300–1500 Hz, with localization accuracy improving as frequency decreases. ILDs function for frequencies spanning from approximately 1000 Hz to the highest audible sounds, with accuracy improving at higher frequencies. Localization accuracy is poorest around 1500–3000 Hz, which, ironically, is within the midrange of frequencies to which humans are most sensitive (1000–5000 Hz).⁴⁴ The reason for this paradox is as follows: ITDs are most effective for low-range frequencies and ILDs are most effective for high ranges, but neither system functions effectively at midrange.⁴⁵ Within this range, ITDs and ILDs may function in tandem, providing ambiguous localization cues.

Low frequency sounds have longer wavelengths than higher frequency sounds. Sounds below approximately 1000 Hz have wavelengths that are longer than the path around the perceiver's head, so the waves *diffract* (or, "bend") around the head to reach the ear farthest from the sound source after having previously reached the nearer ear. This timing difference results in a discrepancy of phase between the signals at each ear. The ITD is greatest when the sound is

⁴³ Warren, 30. The median (or medial) plane passes longitudinally through the middle of the body from front to back, dividing it into right and left halves.

⁴⁴ Levine, 389.

⁴⁵ Warren, 35.

directly in front of one ear (90 degrees to the medial plane) and is zero when the sound is either directly in front of or directly behind the listener's head. When a sound is thus positioned along the medial plane, the ears are equidistant from the sound and the sound's waves reach both ears simultaneously.

Sounds with frequencies above 1000 Hz have wavelengths that are too small to bend around the head and are effectively "blocked" from reaching the far ear. The resultant acoustic shadow (or, "head shadow") causes a reduction in the intensity of the signal at the far ear. The sound reaches the far ear by traveling *through* the head (discounting environmental reflections), so many frequencies within the sound's spectral profile (assuming a complex sound) will be attenuated. Therefore, in addition to a difference in level, the timbral disparity between the signals at each ear provides an important cue for the localization of higher frequency sounds. ILDs are most effective for complex sounds containing components above 1600 Hz.⁴⁶

Localization accuracy is studied by measuring the *minimal audible angle* (MAA), which is the "smallest angle (or difference in azimuth [between two sound sources]) that a listener can discriminate."⁴⁷ The minimal audible angle is smallest (i.e., localization is most accurate) when the sound sources are located almost directly in front of the listener, close to the median plane. As the sources are moved in either direction toward the listener's binaural axis (which extends out from the listener's ears and runs perpendicular to the median plane), the MAA increases dramatically.⁴⁸ The binaural axis is the region of space (relative to the listener) where interaural differences are the greatest but also the most ineffective, typically resulting in ambiguous cues. Studies have also explored the *minimum audible movement angle* (MAMA), which questions

⁴⁶ Blauert, 166.

⁴⁷ Gelfand, 379.

⁴⁸ These observations hold for two stationary sounds presented either sequentially or concurrently. Frequency and spectral differences between the two sounds affect the MAA. See Gelfand, 380–381.

how far sound source B must travel away from sound source A in order for the listener to know that B is moving. The conclusions are virtually identical to those produced in studying the MAA of stationary sounds: movement is detected most effectively when it occurs toward the front of the listener and less accurately when occurring to the listener's side.⁴⁹ Such findings are important for gauging the salience of spatial motion in multi-dimensional ensemble spaces, which shall be addressed in Chapter 4.

Richard Warren states: "The spatial separation of sound images resulting from binaural differences in acoustic input can allow us to hear signals that would otherwise be masked."⁵⁰ Binaural hearing thus allows us to focus on a given sound, such as the voice of a particular speaker, within a noise-filled room.⁵¹ It is also binaural hearing, in part, that allows us to focus on and attend to a particular musical line within a complex, polyphonic texture, or, to zero in on the spatial location of a particular instrument within a larger ensemble.

Monaural Hearing

The *pinna*, which is the outer, visible portion of the ear, essentially functions as a funnel. The ridges and convolutions of the pinna reflect sound waves, funneling them into the ear canal. The manner in which the pinna reflects sound waves into the ear canal affects the spectrum of the auditory stimulus: certain frequencies are amplified or attenuated depending upon the direction from which the sound waves travel to and impact the ear. Additionally, the sounds reflected from the ridges of the pinna cause very miniscule (but detectable) time delays between the direct and reflected signals.⁵² The pinna's effects on time delay and frequency are correlated;

⁴⁹ Gelfand, 181.

⁵⁰ Warren, 42.

⁵¹ This is known as the "cocktail party effect."

⁵² Gelfand, 83-84.

Warren states: "The pinna with its complex corrugations generates many reflections and hence many frequency-specific time delays."⁵³ The resultant changes in the spectral profile of a broadband sound—called "head-related transfer functions" (HRTFs)—have "values that vary with the azimuth and elevation of the sound source relative to the pinna."⁵⁴ In an effort to maximize these monaural cues, humans often move their head either deliberately or spontaneously, which alters the reflection of sounds into the ear canal.⁵⁵ The qualitative changes of the waves being reflected into the ear canal are calculated at higher levels of perception to increase localization accuracy. Thus, while binaural cues aid in the determination of a sound's azimuth, the pinnae are particularly useful in localizing the height of a sound as well as determining whether or not a sound is coming from the front or the back of the perceiver. If a sound occurs directly in front of or behind a listener, it lies along the median plane and delivers equal energy to each ear-rendering ILDs and ITDs ineffective. The potentially resultant "frontback confusion" is typically corrected by the reflections and resonances of the pinnae, as well as by head movements.⁵⁶ The pinna actually casts a sound shadow when sounds approach from the rear. The shadow is more pronounced for high than for low tones, since high frequencies do not diffract as well as low frequencies.⁵⁷ Front-back localization accuracy is greatest for sounds with high frequencies (above 4000 Hz).⁵⁸

⁵³ Warren, 47.

⁵⁴ Ibid.

⁵⁵ Many animals, such as dogs and cats, can actually move their ears to aid in the localization of sound sources. ⁵⁶ Warren, 45. Front-back confusion may also result from multiple sound positions at one side of the listener that create the same ITD or ILD. For example, a low frequency sound occurring 45 degrees to the listener's right side will create the same ITD as it would if it were 135 degrees to the right (with 0 degrees being directly in front of and 180 degrees being directly behind the subject). However, head movements and pinnae function assist the perceiver in properly discriminating the location of the sound. Warren observes that a spectral difference between the same sound at 0 degrees and at 180 degrees is only detectable to listeners who cannot determine if either sound is coming from the front or the back, which is possible when head movements are inhibited. If the directionality of the sounds is perceived, "then timbre differences [can] not be heard and the spectral differences [are] transformed into localization differences for a sound that seem[s] to have an unchanging timbre" (Warren, 48).

⁵⁷ Gulick et al., 338.

⁵⁸ Gelfand, 379.

In short, the head and the external ear serve to alter the sound signals reaching each eardrum through a combination of diffractive, shadowing, and resonating functions. The differences between the signals at each ear enable the listener to localize a sound quite accurately, and the listener may modify interaural signal differences through voluntary head movements (a combination of rotating, tipping, and pivoting the head). Blauert notes: "In all cases in which localization is at first imprecise or anomalous, exploratory head movements take on great importance."⁵⁹ Spontaneous head movements—resulting from the "audiovisual reflex"—serve to direct the head toward the most probable position of a sound's source, which maximizes localization by bringing the sound into the listener's median plane.⁶⁰ It is easier for a listener to localize the source of a familiar sound as opposed to an unfamiliar sound. Furthermore, broadband sound signals facilitate accurate localization more effectively than do narrowband signals.

Summing Localization

Summing localization occurs whenever two spatially separated sound sources emit nearly identical sound signals, such that the level and timing differences between the two signals reaching the listener are below a certain value. Under such conditions, the listener perceives the sound to be located at a point intermediate between the two sources; this process is often referred to as *interspeaker imaging*. Summing localization is critical for the perception of spatial

⁵⁹ Blauert, 44.

⁶⁰ Ibid., 178–179. Blauert states: "The most frequent combination of classes of movements involves rotating [turning the head to the left or right] and tipping [raising or lowering the chin] movements. As to the direction of the movements, it may be stated that the first movement most subjects [undertake is] such that they more nearly face the direction of sound incidence, though usually they [do] not carry this movement so far as to face it squarely" (Blauert, 182). Facing the sound squarely might trigger front-back confusion.

relationships in recorded music, whether it is played back over loudspeakers or headphones.⁶¹ Blauert describes the phenomenon as follows:

Two loudspeakers are positioned in front of the subject, symmetrically with respect to the median plane...If both loudspeakers are driven with identical signals, the subject perceives a single auditory event in the median plane...A delay [creating an interaural time difference] or weakening [creating an interaural level difference] of one of the two loudspeaker signals leads the auditory event [the perceived location of the sound source] to migrate from the middle toward the loudspeaker that is radiating the earlier or stronger signal.⁶²

Such a situation engenders the illusion of a sound source moving continuously through space, and this illusion may be manufactured in sonic environments with more than two sound sources (such as the quadraphonic sound field).⁶³ Electro-acoustic composers frequently exploit the phenomenon of summing localization to infuse their music with spatial motion (this point shall be resumed in Chapter 4). However, if the two speakers are spread too far apart and/or are not angled correctly (approximately 60 degrees between the speakers from the position of the listener is ideal), a "hole in the middle" emerges—meaning that sounds are "drawn into the speakers, leaving virtually no inter-imaging."⁶⁴ This reduces the effectiveness of producing illusory sonic motion, and in the case of a film's audio track, can cause a reduction in dialogue clarity.

Summing localization also occurs when a primary sound and its reflection(s) are presented at roughly the same intensity level, provided that the delay of the reflection is negligible (below 30–50 msec). The location of a sound is perceived to be in-between the primary sound and the point of its reflection. Blesser and Salter note: "When early sonic

⁶¹ Summing localization is responsible for the intracranial lateralization of sounds heard over headphones. ⁶² Blauert, 204–205. Blauert observes elsewhere that "even sounds that reach the listener's ears up to 50 ms[ec] later than the first arriving sound may take part in this 'summing localization,' provided that their level is adjusted adequately" (Blauert, 325).

⁶³ Ibid., 275.

⁶⁴ Richard Elen, "Ambisonics: The Surround Alternative," *Surround* (2001), 1.

reflections from the sidewalls and ceiling reflectors are appropriately combined, musical instruments on the stage of a concert hall sound closer—aurally larger—than they would otherwise."⁶⁵ Thus, while summing localization enables both the spatial differentiation of multiple sounds and the implementation of illusory sonic movement, it can likewise potentially decrease localization accuracy.⁶⁶

The Auditory Pathway and the Mapping of Space

Frequency is coded topographically on the auditory cortex of the brain, which is the peripheral (and terminal) region of the auditory pathway.⁶⁷ This means that specific locations along the cortex respond to specific frequencies. The frequency range decreases along the cortex from the front to the back of the head, with cortical cells toward the front responding to higher frequencies and those toward the back responding to lower frequencies.⁶⁸ Similar topographic maps exist for vision and touch in different cortical regions of the brain. (The map for frequency on the auditory cortex is often referred to as *tonotopic*.⁶⁹) Auditory space, however, is not mapped on the cerebral cortex; instead, it is mapped topographically on the external nucleus of the inferior colliculus.⁷⁰ The inferior colliculus lies in the midbrain and serves as the approximate midpoint of the auditory pathway.⁷¹ This region contains cells referred to as

⁶⁵ Blesser and Salter, 53. The authors likewise note that curved surfaces (such as walls) function to "focus sound such that the source appears aurally closer or farther, larger or smaller" (Blesser and Salter, 54).

⁶⁶ Summing localization seems to account for Henry Brant's conception of "spill," in which similar sound sources separated in space "seem to reach out to each other when all are sounding" (Brant, 232).

⁶⁷ The *auditory pathway* is the neuronal network that transmits a mechanically received sound from the ears to higher processing systems of the brain.

⁶⁸ Levine, 360–361.

⁶⁹ Gulick et al., 193–194

⁷⁰ Masakazu Konishi, "Listening with Two Ears," *Scientific American: Secrets of the Senses* 16.3 (2006): 30–31.

⁷¹ Once a sound has been transduced into a neural impulse, it follows the following ascending auditory pathway: the auditory (or cochlear) nerve exits the inner ear and enters the brainstem just under the cerebellum. Nearly all of the fibers from the auditory nerve synapse on cells in the *cochlear nucleus*. Most of the nerve fibers exiting the cochlear nucleus cross to the contralateral side of the brain. (As with other sense modalities, most auditory input is coded on the portion of the brain contralateral to the side of the head from which the input was received.) Input from both

"space-specific neurons" that "react only to acoustic stimuli originating from specific receptive fields, or restricted areas in space."⁷² (There are cells in auditory cortex that respond exclusively to sounds emanating from a particular location in space; however, these cells are not arranged topographically on the cortex as they are in the inferior colliculus of the midbrain.⁷³) The spacespecific neurons in the left external nucleus of the inferior colliculus "form a map of primarily the right side of auditory space (the broad region of space from which sounds can be detected), and those of the right external nucleus form a map of primarily the left half of auditory space, although there is some overlap."⁷⁴ ITDs and ILDs are processed in separate but parallel neural circuits that ultimately synapse on the external nucleus where the spatial origin of a sound is mapped.

Hearing the Distance of a Sound Source

The most obvious cue for detecting the distance of a sound source is its loudness. Distance affects the perceived intensity of a sound: as distance increases, intensity decreases. The inverse square law states: "The intensity of a sound is inversely proportional to the square of its distance, so that a twofold change in distance corresponds to a fourfold change in intensity."⁷⁵ Thus, if the distance between a sound source and the listener is cut in half, the intensity of the sound produced by that source is perceived to be four times greater. Another important distance cue is the ratio of direct to reflected sound: distant sound sources generate a higher proportion of reflected sounds than do nearby sources. A final distance cue is the attenuation of high

ears is first received at the *superior olivary complex* in the brainstem. Nerve fibers then travel to the *inferior colliculus*, where auditory space is mapped topographically. From this region, fibers travel to the *medial geniculate nucleus* of the thalamus from whence they ultimately project to the *auditory cortex* (Levine, 346–348).⁷² Konishi, 30.

⁷³ Levine, 349, 389.

⁷⁴ Konishi, 30–31.

⁷⁵ W

⁷⁵ Warren, 51.

frequencies over long distances. The medium through which sound waves travel effectively functions as a low-pass filter, and the upper partials of a complex sound progressively dissipate by the time they reach the listener's ears. Under temperate conditions, "components at 8000 Hz may have an attenuation of 3 dB greater than components at 1000 Hz for each 100 feet of travel," and the attenuation of high frequencies over a fixed distance increases as air humidity increases.⁷⁶ Just as the directionality (azimuth and elevation) of a sound is determinable through both binaural and monaural cues, the judgment of a sound's distance is governed by the interaction of three factors: intensity, the ratio of direct to reflected signals, and spectral profile. Knowledge of (familiarity with) the sound being heard, as well as of the space in which it is heard, improves the judgment of distance.

It has been noted that visual cues aid in locating a sound source. Through crossmodal binding (multisensory interaction), some cortical cells in association areas of the brain respond to auditory and/or visual stimuli. The "response of such a cell [is] maximized when appropriate visual and auditory stimuli that originate from the same location in space are presented simultaneously."⁷⁷ The external nucleus of the inferior colliculus—noted above for its topographic mapping of auditory space-contains such bimodal cells, serving to "integrate some auditory information with the visual and somatosensory systems."⁷⁸ Although we can discern the location of a sound solely through audition, the cooperation of vision and audition strengthens the awareness of a sound source's location. However, since sound waves travel through air approximately 893,000 times slower than light, a discrepancy between visual and auditory stimuli may arise at great distances: we might see a sound-producing source before we

⁷⁶ Ibid., 54. ⁷⁷ Levine, 390.

⁷⁸ Gulick et al., 204.

hear the sound it produces.⁷⁹ This temporal disparity can cause the perceiver to segregate modally distinct stimuli that really belong together. An extreme example of this phenomenon is the distant flash of lightning that is followed by a low rumble of thunder.⁸⁰ The greater the distance of the lightning relative to the perceiver, the longer the delay in its sound waves reaching the perceiver. We must consciously and deliberately integrate these temporally detached phenomena to conceive of them as arising from a single event. Under normal performance conditions, the listener has little difficulty fusing a performer's physical motions with the sound he/she produces. However, in exceptionally large concert halls and cathedrals, or in outdoor venues such as amphitheaters, the slight discrepancies between visual and auditory percepts could prove problematic to the discernment of spatial gestures.

Tracking Moving Sound Sources

As noted above, head movements assist in the localization of static sounds by altering the funneling function of the pinnae and changing the ratio of interaural level/timing differences.⁸¹ Conversely, a subject with a stationary head may aurally track a moving sound source via the same means. Imagine a sound-emitting source traveling from left to right in front of a listener, where the produced sound is sustained and qualitatively constant. The interaural differences produced by the sound will decrease as the source approaches the listener's medial plane and subsequently increase as the source continues to the right of the listener—receding from the medial plane. Furthermore, a sound source approaching a listener will increase in intensity and "brighten" in color, while a source receding from the listener will decrease in loudness and

⁷⁹ Wagner, 43. Light travels at approximately 186,000 miles per second.

⁸⁰ The "lowness" of the rumble of thunder becomes more pronounced as the distance of the lightening increases, due to the progressive attenuation of the thunderclap's upper partials.

⁸¹ Deliberate head movements prove auditory localization to be a somewhat active process, but head movements are often reflexive.

"darken." The ratio of direct to reflected signals will also change as a sound source draws nearer to a subject: the direct signal progressively becomes stronger than the reflected signals. Accordingly, as a sound source approaches the subject, its precise location comes into sharper auditory focus as the precedence effect is heightened. An additional cue for gauging a moving sound is the **Doppler effect**, which is the decrease in pitch brought about as a sound source approaches and subsequently passes a listener. As the sound-emitting source advances toward the subject, the wave crests produced by the sound "pile up" and reach the subject at an increased frequency. As the source passes and begins to recede from the subject, the wave crests are "stretched" and reach the subject at a decreased frequency, corresponding to a lowering of pitch. The greater the velocity of the sound source, the more pronounced the perceived drop in pitch.

As evident from the foregoing discussion, our ability to localize the source of a sound is an amalgam of many parallel physiological and psychoacoustic processes. Warren summarizes the inconstant factors impacting auditory localization:

When listening to external sources, extensive complex changes in the acoustic stimulus reaching the eardrums are produced by room acoustics, head shadows, pinnae reflections, the reflection and diffraction of the sound by the neck and torso, as well as by head movements. The nature of these changes varies with the listener's distance and orientation relative to the source.⁸²

The perception of a sound's locatedness is dependent upon the environment in which the sound occurs—environmental conditions directly affect how the waves produced by a sound reach the listener. Although auditory localization is an automatic process over which we have little control—with the exception of purposive activities such as head movements, the environmental conditions affecting localization are variable, and in many respects, manipulable by those either

⁸² Warren, 44.

producing (performers) or receiving (listeners) musical sounds.⁸³ Harley states: "The location, orientation, attention and, sometimes, movement of [the listener] are crucial for the 'spatialisation' to occur. The listener orients him/herself toward the music, [and] the music is directed toward the listener. Their reciprocal relationship is essential."⁸⁴ The perception of spatial gestures in a musical performance depends upon accurate localization, and the salience of those gestures is affected by many variable factors of performance, including the size, shape, and acoustic qualities of the performance hall, the onstage orientation (positioning and spacing) of ensemble members relative to one another, and the location of the listener relative to the ensemble.

2.4 Ensemble Space

It is necessary to distinguish between the different types of "spaces" encountered during a musical performance. I shall reserve the term **musical space** for the phenomenal, conceptual space constructed mentally during the experience of music. The musical space accounts for the perceived relations between musical entities and/or attributes that are metaphorically structured in terms of objects in space. A component space of musical space is **pitch space**, wherein pitches are perceived to abide at varying heights: higher frequencies correspond to higher pitches and lower frequencies to lower pitches along a continuum. Pitch space is thus the vertical dimension of musical space. Time is often conceived as the horizontal dimension of musical space. This two-dimensional conceptual space (comprising pitch space and time) correlates with the staff of a musical score, where high pitches are positioned higher on the staff than are low

⁸³ Obviously, acousticians play a tremendous role in defining how sonic signals will reach the listener in a concert hall. Blesser and Salter draw distinctions between an "aural architect" and an "acoustic architect." See Blesser and Salter, 5.

⁸⁴ Harley, "Spatiality of Sound and Stream Segregation," 148.

pitches, and the left-to-right ordering of notes corresponds to the flow of time. A third dimension of musical space often results from variable dynamic levels, with loud musical events perceived as "close" and soft musical events considered "distant." This corresponds somewhat to our experience of a constant sound's intensity decreasing as its distance increases. This dimension of musical space is somewhat contrived and conventionalized since other distance cues—such as the attenuation of higher frequencies and the changing ratio of direct to reflected sound—are not (ordinarily) present when the dynamic level of a musical work subsides.⁸⁵

The **performance space** is the room in which a musical performance occurs and in which the sounds are, for the most part, entirely contained. In his 1952 Ph.D. dissertation, Edward A. Lippman claims: "As a spatial setting, the place of performance is of actual musical significance. The amount of reverberation of this space is part of the general quality of the music."⁸⁶ Blesser and Salter echo this point of view by claiming that, since a performance venue affects sound by functioning as a "resonant enclosure," the performance space becomes a *metainstrument*, which Blesser and Salter define as "an extension of the musical instruments played within it."⁸⁷ The performance space contains the performing ensemble (as the collective source of sounds) as well as the audience (the intended target of sounds produced by the ensemble), yet the performance space is larger than the summed areas of the ensemble and audience spaces. The **audience space** is the boundary of all the performers within the ensemble.⁸⁸ All of these spaces occur in a multitude of shapes and interrelations, with one constant factor: musical sounds occur

⁸⁵ Boulez codifies and discusses several different types of "musical space." See Boulez, *Boulez on Music Today*, 83–98.

⁸⁶ Edward A. Lippman, "Music and Space: A Study in the Philosophy of Music" (Ph.D. diss., Columbia University, 1952), 77.

⁸⁷ Blesser and Salter, 135.

⁸⁸ The unoccupied portion of the stage in a concert hall would be considered part of the performance space, since it is the portion of space where sound waves travel and not the portion of space where musical sound events occur.

in the ensemble space, and their resultant waves infiltrate the audience space and reverberate (to varying degrees) within the entire performance space.⁸⁹ Figure 2-1 illustrates the most common relations among these spaces.



Figure 2-1: Typical spatial relations among *ensemble space* and *audience space* within the *performance space*.

In the case of a string quartet, the ensemble space consists of the physical space spanning from the first violin to the cello (or viola). This size and shape of the space is variable depending upon how "spread-out" the ensemble members elect to arrange themselves from one performance to another. The ensemble space may be partitioned into different sections, each of which is an **ensemble-space segment**. Each individual member (performer/instrument) within the ensemble represents a distinct **ensemble point** within the total ensemble space. In Figure 2-1, each "P" located inside the ensemble space depicts an ensemble point. Each discrete ensemble point is a sound source with its own unique spatial coordinates, and any sound issuing forth from a point is a **sound event**. Ensemble spaces and ensemble-space segments may be

⁸⁹ Incidental sounds produced in the audience space are generally ignored (i.e., not consciously attended to) and treated as non-musical, extraneous noises. In some performance settings, the audience space and the ensemble space are conflated. Such is the case, for example, when the ensemble members are dispersed among the audience.
differentiated by the number of points they comprise. The string quartet is an example of a *fourpoint ensemble space* (the unspecified ensemble space in Figure 2-1 is likewise a four-point ensemble space). The two violins of a string quartet (when seated adjacent to one another) constitute a *two-point ensemble-space segment*.

Multiple performers playing the same musical part on the same type of instrument and situated proximate to one another within the ensemble space, as in the viola section of a symphony orchestra, constitute a spatial area. An area is essentially an inflated point; however, the area has the potential to fragment into two or more smaller areas or into multiple points (as would be accomplished through divisi indications). Conversely, areas may be constructed via the process of *assemblage*, which occurs when two or more previously independent yet spatially proximate instruments (of the same kind) begin to play in unison. When an ensemble space comprises two or more spatially separated ensembles, and each individual ensemble features multiple musical parts and/or instruments, the distinct ensembles may each function as a hyperpoint in the global context of the total ensemble space.⁹⁰ The sonic space is the portion of the ensemble space actively engaged in the production of sound at any given moment. The sonic space is nonexistent when none of the ensemble points are producing sound and is equivalent to the ensemble space when all ensemble members are producing sound simultaneously. The ensemble space is relatively fixed in a given performance, but the sonic space is highly fluid and variable. The capacity of the sonic space to wax and wane creates dynamic fluctuations within the spatial construct of a musical work. The aesthetic and interpretive implications of such activity shall be explored in Chapter 5.

⁹⁰ Such is the case in Stockhausen's *Gruppen für drei Orchester*: each of the ensemble space's three component orchestras contains a variety of instruments. The total ensemble space therefore comprises three hyperpoints, and each hyperpoint contains multiple points (some of which occasionally assemble into areas).

The performance space contains the audience space as well as the ensemble space, and the ensemble space contains any number of ensemble points. Normally the listener and the ensemble occupy two different spaces within the performance space; as depicted in Figure 2-1, the ensemble space and the audience space are distinct. However, in works featuring the "surround" effect (either electronic or acoustic), the listener is enveloped by the ensemble and made to abide *within* the ensemble space. Figure 2-2 illustrates the positional relations in such a performance space. In this figure, a six-point ensemble space envelops the audience space.



Figure 2-2: Spatial relations among *ensemble space* and *audience space* in a surround-sound environment

In Figures 2-1 and 2-2, the ensemble space is graphically depicted as being circular while the audience space and the performance space are both rectilinear. This is merely a convention that is observed consistently in both figures for the sake of illustrational and explanatory clarity. While many concert halls are indeed rectangular, the audience space and ensemble space are

variable in terms of their shape—whether existing as a mass of seats (as in the audience space) or merely as an outline (as with the ensemble space).⁹¹

The variable relationship between the perceiver and the ensemble space (within the performance space) affects the perceptibility of spatial gestures, as do the acoustic properties of the performance space. For instance, in an extremely "wet" concert hall-one in which sounds reverberate to a great extent—spatial gestures might be difficult to decipher, particularly if the ensemble is relatively small and the listener is seated at the back of the second balcony. Many concert halls feature adjustable curtains or acoustic panels that may be raised/lowered. The curtains and panels serve to absorb sound waves and minimize reflections. Although likely lacking in loudness and sonorous richness, a "dry" concert hall allows for more accurate localization of discrete ensemble points owing to the minimization of sonic reflections. The listener often has a choice in determining his/her location in the performance space relative to the ensemble space. In a highly reverberant room, the sonic reverberation will be relatively consistent throughout the space, but the intensity of the direct sound decreases as distance from the sound source increases. Obviously, sitting closer to the ensemble space will increase the intensity of the direct signal relative to the reflected signals, and localization will be enhanced. As the perceptual differentiation of each point within an ensemble space becomes more pronounced, the potential for apprehending spatial gestures unfolding within that space improves.

⁹¹ The ensemble space in Figure 2-1 is rightfully a line rather than an elliptical shape. Additionally, "connecting the dots (points)" of the ensemble space in Figure 2-2 might yield a hexagonal shape instead of a circular form (in either case, the ensemble space is surrounding the audience space). The boundedness imposed on ensemble and audience spaces is often purely perceptual.

2.5 Auditory Streams

We shall now consider the process by which a sequence of related or unrelated sounds might be grouped perceptually into a unified entity spanning a specific amount of time. Warren explains that the threshold for the correct perception of the ordering of a sequential pair of unrelated sounds is approximately 20 msec, and that this fundamental limit is "independent of the modality employed."⁹² If two consecutive sounds occur less than 20 msec apart, their temporal ordering cannot be detected. This observation shows that dramatically poor ensemble coordination might yield unintended spatial gestures: if two or more instruments are meant to begin a passage at the same time, but their onsets are grossly misaligned and separated by more than 20 msec, a spatial gesture might be perceived. For items presented only 1–2 msec apart, the resultant "micropatterns' [are] not perceived as a succession of discrete items but rather as unitary perceptual events, with different qualities associated with the different orders."93

Warren cautions that the above observation applies specifically to two-item sequences and should not be applied to extended sequences in which "the initial and terminal sounds of sequences, whether consisting of two items or many items, are identified with special ease."⁹⁴ The perception of order for multi-item sequences requires approximately 200 msec for each component item of the sequence, which equates to a maximum of five items per second. It is rare in musical performance for a spatial gesture to unfold at a rate faster than five events per second. Warren explains that for fast auditory sequences, listeners may deduce the sequential ordering of items from the larger pattern formed by the sequence: "Although it may not be

⁹² Warren, 113–114. ⁹³ Ibid., 120.

⁹⁴ Ibid., 114.

possible to identify the orders of items in such sequences, different arrangements of the same items are not perceptually equivalent but they do form distinguishable patterns."⁹⁵

Warren shows that two very different processes enable a listener to distinguish between "permuted orders of [unrelated] sounds within extended [auditory] sequences: (1) direct identification of components and their orders; and (2) holistic pattern recognition."96 The activation of either process depends upon the rate at which the sequence occurs: slow sequences enable direct assessment of item ordering while a fast sequence is received more holisticallyalthough its item ordering may be deduced from the quality of the pattern as a whole. The permuted orderings of unrelated sounds "can be distinguished without the ability to identify the orders within the sequences."⁹⁷ Warren's observations apply to sequences of unrelated sounds presented equally to both ears (ensuing from the same point in space) that form a unique, holistic sonic pattern. His notion of holistic pattern recognition (possible in the absence of the discrimination of order) is applicable to spatial gestures in musical performance, which form shapes and motions (patterns) in space. Although the majority of spatial gestures occur slowly enough to allow for the discrimination of event order, I shall argue in Chapter 3 that they are still received as coherent patterns (or, "gestalt structures") through the emergence of directed motion. The slow temporal succession of events does not negate holistic pattern recognition.⁹⁸

Albert S. Bregman's fruitful theory of *auditory scene analysis* provides a useful framework for postulating the perceptual salience of spatial gestures. Bregman equates the

⁹⁵ Ibid., 117.

⁹⁶ Ibid., 133.

⁹⁷ Ibid. Bregman summarizes Warren's findings: "The astounding success of subjects in experiments in which they are required to recognize the order of events in extremely rapid sequences (having onset-to-onset times as low as 2.5 msec) is achieved by first recognizing some quality of the sequence as a whole that changes when the order of elements changes, and then remembering that this quality is what is experienced when the components are in a certain order" (Bregman, 70).

⁹⁸ However, incredibly slow sequences would challenge holistic perception by requiring a greater amount of retention on the part of the listener.

auditory scene to the visual field: just as the visual field contains multiple objects upon which we may focus our attention, the auditory scene contains any number of "auditory streams." Bregman defines an auditory stream as "our perceptual grouping of the parts of the neural spectrogram that go together."⁹⁹ The bulk of Bregman's theory is devoted to defining what sonic phenomena "go together." In short, the perceptual act of auditory scene analysis involves stream segregation. An auditory scene may comprise multiple streams. A single stream is the grouping (or integration) of multiple sounds that occur either simultaneously or sequentially. The stream itself endures through time, as do its component sounds. A stream generally consists of multiple sounds that are related in some respect: "A stream serves the purpose of clustering related qualities."¹⁰⁰ Bregman uses "a series of footsteps" to exemplify the concept of an auditory stream, claiming the series "can form a single experienced event [, or 'happening'], despite the fact that each footstep is a separate sound."¹⁰¹ Each footstep produces a qualitatively similar sound, and the multiple footsteps are perceptually grouped over time. Curiously, Bregman says nothing of the fact that the sequence of footsteps would most likely have a perceptible directionality of motion. This oversight shall prove problematic in due course.

Bregman distinguishes between two types of auditory stream segregation: sequential and simultaneous integration.¹⁰² Sequential integration is described as "the process that connects events that have arisen at different times from the same source."¹⁰³ It is sequential integration that allows us to conceive of a melody as a single "thing," although it is composed of successive pitches. Bregman's definition of sequential integration presumes a stationary sound source—he

⁹⁹ Bregman, 9. ¹⁰⁰ Ibid., 10.

¹⁰¹ Ibid.

¹⁰² Multiple sounds are *integrated* to form a stream, and multiple streams are *segregated* from one another (within the auditory scene). "Integration" is synonymous with "grouping."

¹⁰³ Bregman, 30.

says nothing of how a moving source (such as a person walking and producing audible footsteps) might affect streaming.

Simultaneous integration is the process that "takes acoustic inputs occur[ing] at the same time but at different places in the [frequency] spectrum *or in space*, and treats them as properties of a single sound."¹⁰⁴ It is this process by which we perceive chords and harmony in music. Using our traditional metaphors for the organization of musical space, we may think of sequential integration as horizontal grouping (over time) and simultaneous integration as vertical grouping (in time).¹⁰⁵ Sequential integration groups sounds coming from the same source at different times, and simultaneous integration groups sounds coming from different sources at the same time. By thus restricting his categories of stream segregation, Bregman implies that sounds coming from different sources at different times are not integrated into a stream, even if those sounds are qualitatively similar or identical.¹⁰⁶

According to Bregman, for a stream to be formed through sequential integration, the sounds constituting the stream must be similar in some respect. Similarity could exist in the realms of pitch (proximity in pitch space), timbre, duration, envelope, loudness, consistent onset-to-onset interval, emergent temporal patterns (rhythm), continuity of change, and spatial location. (The combination of some or all of these sonic attributes would certainly strengthen perceived similarity relations and improve stream segregation.) An *envelope* is the "temporal

¹⁰⁴ Ibid., 31. Emphasis mine.

¹⁰⁵ Simultaneous integration likewise occurs between different sense modalities, such as vision and audition, and in this respect is synonymous with crossmodal binding. See Bregman, 290–291.

¹⁰⁶ As the distance between the perceiver and the ensemble space increases, sounds from discrete points increasingly seem to come from the same point in space. Imagine a quartet of guitars performing on a concert stage and a listener who is seated a great distance from the four-point ensemble space. During a particular passage, the guitarists consecutively produce comparable sonic events. The activity of the ensemble would constitute a stream of qualitatively similar (but temporally discontinuous) sounds issuing from the same location (from the perspective of the remote audience member). As the perceiver moves closer to the ensemble space, and the discrete points become more pronounced, does the stream stop being a stream or does it simply acquire a dynamic spatial form?

shape of a tone consisting of an attack (or onset), initial decay, sustain, and final decay."¹⁰⁷ The *onset-to-onset interval* could indicate the period of time separating the initiation points of two separate events regardless of the duration (sustainment) of each event. This term could be recast as "*offset*-to-onset interval" to account for the temporal "space" (silence) separating the termination of one event and the commencement of a subsequent event.¹⁰⁸ *Isochrony*, or regular spacing in time, results from uniform onset-to-onset intervals separating discrete events and contributes to stream segregation.¹⁰⁹

Continuity of change accounts for the steady process of transformation that some sonic attribute of a sound may undergo, such as a portamento (continuously changing frequency), a timbral "morph" (continuously changing spectral profile), a crescendo (continuously increasing amplitude), etc. Bregman states: "A homogenous perceptual input contains no units. Only when it is broken up by some sort of discontinuity does it organize itself into units. After being formed, units can be grouped by similarity and other factors [such as spatial and/or temporal proximity] to form higher-order organizations."¹¹⁰ A single, sustained sound, such as a sine tone, is a homogenous perceptual input. However, a single, sustained sound undergoing some sort of seamless, continuous change (such as a modulation of frequency, intensity, tone color, spatial location, etc.) is a single input that cannot be segmented into discrete units. This solitary metamorphic unit, although not grouped together with other units, constitutes a single auditory stream due to its continuity of change over time.¹¹¹ Such continuity effectively replaces the homogeneity of an unchanging, invariable sound. Spatial gestures in an acoustic ensemble space

¹⁰⁷ Wagner, 18. The envelope of a sound accounts for an aspect of its changing persistence over time.

¹⁰⁸ Bregman 65–66.

¹⁰⁹ Ibid., 136.

¹¹⁰ Ibid, 70. The Gestalt grouping principles of similarity and proximity shall be explored in Chapter 3.

¹¹¹ This observation conforms with O'Callaghan's conception of a sound as an event that endures qualitative changes as it persists over time.

inevitably feature discontinuities between the spatial location of discrete ensemble points and therefore do not constitute a continuous change of locale.¹¹² It is because of these discontinuities that the discrete events must be perceptually grouped to form a spatial gesture. Robert Hatten states:

When [musical] gestures encompass more than one musical event, they provide a nuanced continuity that binds together otherwise separate musical events into a continuous whole. Continuity in this sense is not equivalent to continuously sounding; *a discontinuous sequence of sounds (e.g., sounds separated by rests) may nonetheless be linked by a continuous thread of intentional and significant movement.*¹¹³

Applying Hatten's thought to spatial gestures, the directed motion (dynamic shape) of a spatial

gesture is the "thread" linking a sequence of spatially (and perhaps temporally) discontinuous

sounds.

Regarding spatial location, Bregman states:

One would think that the most powerful principle of all for the perceptual decomposing of a complex mixture of sounds should be to group, as part of the same stream, all components of sound that come from the same position in space. It would be a good bet that sounds emanating from the same location have been created by the same sound source.¹¹⁴

Bregman goes on to cite numerous experiments that address the manner in which spatial

separation affects the sequential grouping of sounds.¹¹⁵ Many of these experiments involve the

use of headphones to present different sets of auditory stimuli to each ear.¹¹⁶ When different

strings of pitches are presented to each ear, it is often found that frequency-proximity cues cause

¹¹² Spatial gestures in the electronic medium are another issue and shall be addressed in Chapter 4.

¹¹³ Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 94. Emphasis mine.

¹¹⁴ Bregman, 73.

¹¹⁵ See Bregman, 73–83.

¹¹⁶ Warren raises some concerns regarding the use of headphones in auditory experiments: "Headphone sounds, especially when heard diotically and without echoes, are in some ways more similar to self-generated sounds (e.g., our own voice, chewing sounds, coughing) localized within our head than they are to sounds with external origins. Hence, considerable caution must be used in applying results of experiments using headphones to localization of real sources in the environment, since stabilizing and simplifying stimulus configurations may complicate and confuse perception" (Warren, 44–45). Relating the problem of headphones specifically to localization, Warren states: since complex covariance of many stimulus attributes is normally associated with positional changes of an external source, changing only one attribute while keeping others constant can produce a conflict of cues and prevent perception of clear positional images" (Warren, 45).

tones to be grouped into streams—they are *not* consistently grouped by their ear of presentation.¹¹⁷ Bregman ultimately concludes that "although...stream segregation based on spatial location can occur, it seems to lose out when it is placed into conflict with other bases for grouping [such as similarity]. This suggests that segregation by location alone is not very strong."¹¹⁸ What these observations imply is that homogenous sounds coming from different locations may be grouped into a stream. Bregman continues:

The sounds made by one event, say a musical instrument, can pass around another musical instrument that lies between it and the listener, so that both sounds may appear to be coming from the same place. Also, because of echoing, reverberation, and wrapping around obstructions [diffraction], the spatial origins of a sound may be very unclear. Perhaps location information should not count too strongly unless it is supported by other bases for segregation. Maybe we should expect to find that location differences alone will not be powerful influences on grouping, but will have a powerful multiplying effect when they are consistent with other information such as frequency-based ones.¹¹⁹

These remarks strengthen the notion that a series of spatially separated sounds may be grouped into an auditory stream provided that the sounds are somehow qualitatively similar. If thus occurring within the ensemble space of a musical performance, the resultant stream would be a spatial gesture, because the sequence of like events at different ensemble points would form a composite motion. Citing Bregman and relating auditory streaming to spatialization in music, Harley states: "It is known from psychoacoustics that the fusion of distinct sound images...into one auditory stream (here: the image of continuously moving sound) is possible when the sounds are sufficiently similar."¹²⁰ However, should the uniformity of successive events (spatially "distinct sound images") be a necessary and sufficient condition for the formation (and

¹¹⁷ In other words, pitches that are nearby one another in pitch space (proximity of frequency) might be grouped together more readily than pitches that are close together (i.e., sharing identical source locations) in actual space. ¹¹⁸ Bregman, 82.

¹¹⁹ Ibid., 83.

¹²⁰ Harley, "Space and Spatialization," 295.

perception) of a unified spatial gesture? In what remains of this chapter, as well as in succeeding chapters, I argue that it should not.

In comparing Warren's conception of auditory *sequence* with Bregman's notion of auditory *stream*, we find that both of these concepts involve the grouping of temporally successive sounds. The distinction is that an **auditory sequence** may comprise unrelated sounds, while an auditory stream consists exclusively of related (qualitatively similar) sounds. I suggest that either of these concepts may be conflated with the spatial distinction between component events (or items) to yield a spatial gesture. In other words, a spatial gesture may consist of highly related or utterly dissimilar sounds—what matters is the composite shape and directed motion arising from the consecutive activity of spatially discrete ensemble points. Such dynamic shaping through space is a "gesture" in Hatten's sense of the term, except that it occurs in actual space instead of in pitch space.

In synthesizing the results of different (yet similar) experiments conducted by both Bregman and Warren, we may observe that it is possible for auditory streams to fall out of auditory sequences. Temporarily bracketing the potential for any spatial separation, imagine the following sequence of sounds, where each letter represents a unique sound-item with distinctive pitch and timbre: a-b-c-d-a-b-c-d-a-b-c-d. Within this long sequence, we have three repetitions of a four-item sequence: a-b-c-d. However, when the sequence is relayed at increasingly fast speeds, the long sequence will breakdown into four streams—each comprising three repetitions of the same sound-item: (1) a-a-a; (2) b-b-b; (3) c-c-c; and (4) d-d-d. The streams are grouped on the basis of item similarity (pitch and timbre) rather than direct temporal contiguity.¹²¹ We shall discover in Chapter 4 how this phenomenon is applicable to spatial gestures.

¹²¹ Similarity of pitch is better described as proximity in pitch space.

Now, imagine that each letter (sound-item) in the original sequence is performed by a different instrument within a four-point ensemble. Suppose that the leftmost ensemble point performs every "a-item" and the rightmost instrument performers every "d-item," with items "b" and "c" performed by the interior ensemble points. Figure 2-3 provides a graphic depiction of this situation.



Figure 2-3: Stationary streams and gestural sequences

From this figure we see the emergence and repetition of a spatial gesture that "pans" from the left side of the ensemble space to the right side. This gesture is more closely aligned with the auditory sequence (a-b-c-d...) than with the streams occurring at each discrete ensemble point (a-a-a, b-b-b, etc.). The spatial gesture is formed through the temporal contiguity of successive events and is not dependent upon sonic congruity (of pitch, timbre, duration, etc.) among the consecutive sound-items.

Let us now alter the preceding situation. Imagine a trio (three-point ensemble space) of identical instruments that perform the following sequence: a-a-a-b-b-b-c-c-c-d-d. In this sequence, each letter represents a different pitch with a different duration, dynamic marking, and articulation, so the differences between the sound-items are quite pronounced. Each sound-item is repeated three times, but each time at a different ensemble point (proceeding in a left-to-right

direction through the ensemble space), and each of the three ensemble points performs each sound-item once, as depicted in Figure 2-4 below.



Figure 2-4: Stationary sequences and gestural streams

In this situation, the same kind of spatial gesture as that encountered in Figure 2-3 (exhibiting a left-to-right motion) emerges and is intimately bound with the four auditory streams formed through the sequential integration of identical sound-items. The spatial gestures unfolding in the hypothetical situation of Figure 2-3 are arguably less salient than those occurring in Figure 2-4, primarily because the latter are bound to auditory streams and the former are not. This point shall be resumed in Chapters 3 and 4. For now, it suffices to say that both situations produce similar spatial gestures.

This chapter has described sounds and their audible qualities, the behavior of sound waves in the environment, and the process of auditory localization. I have highlighted many of the variable perceptual factors and performance conditions that affect the reception and perceptibility of spatial gestures. I have defined the different types of spaces encountered in musical experience and have prescribed some specific terminology for discursively addressing such spaces. We have begun to encounter the types of musical situations that might give rise to spatial gestures and have described the basic elements necessary for the emergence of spatial shapes and directed motion in the spatial construct of musical performance—namely, a sequence

of spatially discrete sounds, which may or may not be qualitatively similar. In Chapter 3 we shall further explore the nature of spatial gestures, identify and differentiate the primary "gestural categories," and systematize a means of notating gestures in spatial analysis.

CHAPTER 3

THE IDENTIFICATION AND CLASSIFICATION OF SPATIAL GESTURES: AN ANALYTIC METHODOLOGY

Having surveyed various practices and aesthetics of spatialization in Chapter 1 and outlined theories of sound, motion, and auditory localization in Chapter 2, we shall now explore the nature of spatial gestures in greater detail. At the beginning of this chapter, I summarize the Gestalt theory of form perception that both undergirds my argument for the perceptual authenticity of spatial gestures and frames my delineation of gestural categories. I then explicate the technique of integer notation used to label spatial gestures; such a notational system provides a succinct means of identifying specific gestures and comparing the "spatial forms" of multiple gestures. Next, I reveal how different spatial gestures may be grouped into one of three broad gestural categories (each of which contains at least two subcategories), and I define the assets that a gesture must possess for membership in each respective category. Each category comprises a family of spatial gestures that are related through one or more invariant properties. Finally, I discuss the gestural affordances of different ensemble spaces (i.e., the types of spatial gestures that are feasible within spaces of varying ensemble-point cardinality).

3.1 Gestalt Theory: A Summary

The very perception of spatial gestures, as well as my nomenclatural and taxonomic system for labeling and classifying them, hinges upon the principles of Gestalt psychology.

While I do not claim that my analytic system conforms directly and exactly to human perceptual processes (it is chiefly an analytic model rather than a model of perception), I do maintain that there is a close correspondence between the cognitive processes by which we perceive spatial gestures as coherent entities and the means by which we can identify and categorize them for analytic purposes. The term **gestalt** may be translated as "figure," "pattern," "form," or "shape." In the Gestalt view, "the whole figure is imbued with a property as a unit that is not evident when it is analyzed as a collection of features."¹ An ordered collection of dots on a sheet of paper may be seen as a figure, not merely as a summation of numerous dots. In the words of Max Wertheimer: "There are wholes, the behavior of which is not determined by that of their individual elements, but where the part-processes are themselves determined by the intrinsic nature of the whole."² In other words, the whole is greater than the sum of its parts: simply totaling the separate parts overlooks the critical relationships between those parts. The structural interrelationships among the individual elements are ultimately responsible for forming and defining the whole.

Gestalt psychologists claim that we immediately perceive "things" in a holistic manner, irrespective of the sense modality (vision, audition, olfaction, gustation, etc.) receiving sensory input "about" a distal stimulus.³ In other words, we instantly perceive a thing (as the object of sensation) as a unified whole instead of having to attend separately to each of its constituent parts and amalgamate those parts to reconstitute the whole in perception. This is a kind of *preattentive*

¹ Levine, 185.

² Max Wertheimer, "Gestalt Theory," in *A Source Book of Gestalt Psychology*, ed. and trans. by Willis D. Ellis, with an introduction by Kurt Koffka (London: Routledge & Kegan Paul Ltd., 1967), 1–11.

³ Perception is the development of an internal representation of the outside world based on information presented to the mind by the senses. The mind interprets data received from the senses (sensory input) and constructs percepts (perceptions) of that data. Our experience of the world "out there" is indirect—it is mediated by bodily sensation and mental perception.

processing that occurs automatically—prior to reflective awareness and deliberate cogitation.⁴ We immediately recognize a chair as being a chair without wittingly processing the fact that it consists of an elevated, horizontal surface with four legs extending toward the floor and a vertical structure coming up off of one side. We instantly perceive the chair as a unified entity without having to mentally fuse and integrate its various parts. Yet, we can learn more about a whole by studying the structural relationships between its component parts, such as the functional organization of the chair's legs, seat, and back. Viki McCabe states: "The 'parts' [of a whole] may be conceptually distinct but they are not ontologically separate; parts both constitute and get their meaning from their relationship to one another. These relationships are the working structures from which an integral whole develops."⁵ The whole is shaped and rendered meaningful by the interrelationships between its discrete parts as well as the relationship of each part to the whole. In the words of Kurt Koffka:

If we had mastered one part [of a symphony], we should know a great deal about the whole, even if the part which we had mastered never recurred again in the symphony. [This] mode of approach...would not only tell us *what* each [musician] did at any particular moment but *why* he did it. The whole performance would be meaningful and so would be our knowledge of it.⁶

Since my analytic and interpretive framework is dependent upon them, I shall explore the fundamental principles of Gestalt psychology in greater detail.

Gestalt Psychology emerged in the early twentieth century among German theoretical psychologists, developing at a time when Behaviorism was the prevailing theory of learning in America. The Behaviorists believed that behavioral modification, when prompted by positive or negative reinforcement to stimulus response, enables learning. The Gestaltists argued that innate

⁴ Levine, 191.

⁵ McCabe and Balzano, 9.

⁶ Kurt Koffka, *Principles of Gestalt Psychology* (New York: Harcourt, Brace and Company, 1935), 21. Emphasis Original.

cognitive processes guide learning, perception, and the subject's resultant behavior. Gestalt theory was likewise a reaction against nineteenth-century Sensationalist and Associationist theories, which had emerged from the philosophies of the British Empiricists Locke, Berkeley, and Hume. These doctrines held that "certain atomic sensations were 'given' [and that] all else is a superstructure built up by association."⁷ The Gestaltists, on the other hand, maintained that wholes are given as the objects of perception—that perceptions are generated from sensory input by the acquisition of a structural pattern that organizes the multiplicity of sensations. The principle advocates and practitioners of Gestalt theory throughout the twentieth century were Max Wertheimer, Arnold Köhler, and Kurt Koffka, although the fundamental tenets of Gestalt psychology are rooted in the theories of Christian Von Ehrenfels, Johann Wolfgang von Goethe, Immanuel Kant, Ernst Mach, and the Phenomenologists Franz Brentano and Edmund Husserl.

A major criticism of Gestalt Theory is its tendency to downplay (or outright deny) the important roles of experience and learning in the perceptual organization of figures.⁸ The Gestaltists insist that form perception is a purely spontaneous and innate mental capacity. However, recent trends and findings in the fields of philosophy and psychology have revealed that *embodied knowledge*—acquired through accumulative sensorimotor (bodily) experience within our physical environment—as well as culturally conditioned, learned behavior, both have a profound impact on the manner by which we perceive and structure an understanding of the world around us. In semiotics, many symbols are arbitrary—there is often no direct correlation between the "signifier" and the "signified." Yet, we learn to recognize signs and to know what they "mean." A word, written on a page, is a figure (whole) that consists of multiple letters

⁷ D. W. Hamlyn, *The Psychology of Perception: A Philosophical Examination of Gestalt Theory and Derivative Theories of Perception* (New York: The Humanities Press Inc., 1957), 39.

⁸ For a criticism of Gestalt Theory, see Donald O. Hebb, *The Organization of Behavior* (New York: John Wiley & Sons, Inc., 1952).

(parts). If we are familiar with the language to which that word belongs, we will recognize the word instantly without having to ponder the sequence of letters. We must have first learned the language in order to recognize the word as an actual word (instead of as a meaningless grouping of letters) and then to understand what it means.⁹ Whatever innate capacities human beings possess for organizing and perceiving figures, learning clearly plays a role in pattern and form perception. Nevertheless, the Gestalt laws of perception (to be explicated below) remain valid—they work in conjunction with various forms of learning, including those arising from bodily and acculturated experience. Albert Bregman states: "The Gestalt principles of grouping can be interpreted as rules for [auditory] scene analysis."¹⁰ Bregman argues that through evolution and the continuous adaptation of the species to the environment, some perceptual processes are indeed innate, but that they work in accord with principles of learning. As we shall discover in due course, Gestalt perceptual principles are immediately applicable to both the perception of

⁹ If the letters of a word are disordered, we might still recognize the word and may not even notice that it is misspelled. Several theories for word recognition exist in the field of cognitive linguistics, among them: word shape recognition (due to its composite shape, a word is seen as a complete pattern instead of as the sum of letter parts), serial letter recognition (a word is read letter-by-letter from left to right), and parallel letter recognition (letters are recognized simultaneously and "best guesses" are made to construct phonemes and then words). Of these three theories, the second has been debunked and the third appears to be the most promising (although many typographers continue to endorse the first theory). A misspelled word is often easily recognizable provided that the first and last letters of the word remain intact, particularly in shorter words: a three-letter word cannot be altered under these restrictions, and in a four-letter word the interior letters will simply exchange positions. Transpositions of adjacent letters make the word easier to recognize than transpositions involving more distant letters: "acnadejt" is less recognizable than "ajdacnet" as the misspelling of "adjacent." Words containing letters that may form other words (such as "salt" and "slat") are less easily recognized when jumbled, although context may help with correct word identification. In light of these observations, word recognition is still primarily a holistic enterprise that is acquired through learning. Some of the above concepts apply to gauging the salience of similarity relations among multiple gestural forms, to be addressed in Chapter 4. See Kevin Larson, "The Science of Word Recognition: or How I Learned to Stop Worrying and Love the Bouma," available at: http://www.microsoft.com/typography/ctfonts/ WordRecognition.aspx (accessed January 18, 2007); L. X. McCusker, P. B. Gough, and R. G. Bias, "Word Recognition Inside Out and Outside In." Journal of Experimental Psychology: Human Perception and Performance 7.3 (1981): 538–551; K. Mayall, G. W. Humphreys, and A. Olson, "Disruption to Word or Letter Processing? The Origins of Case-mixing Effects," Journal of Experimental Psychology: Learning, Memory, & Cognition 23 (1997): 1275-1286; D. G. Pelli, B. Farell, and D. C. Moore, "The Remarkable Inefficiency of Word Recognition," Nature 423 (2003): 752-756.

¹⁰ Bregman, 24.

spatial gestures and the formulation of a methodology for codifying and analyzing them. Let us first explore previous applications of Gestalt theory to musical thought.

Gestalt and Music

Advocates of Gestalt psychology frequently use musical examples—melodic construction in particular—to support their theories. Wertheimer summarizes an argument by Von Ehrenfels:

I play a familiar melody of six tones and employ six *new* tones, yet you recognize the melody despite the change. There must be something *more* than the sum of six tones, viz. a seventh something, which is the form-quality, the *Gestaltqualität*, of the original six. It is this *seventh* factor or element which enabled you to recognize the melody despite its transposition.¹¹

Thus, every whole has a **Gestaltqualität**, or "form-quality." It is the Gestaltqualität of a melody that enables its recognition when it is played in different keys and on different instruments. In addition to Gestalt theorists writing about music, many twentieth-century musical thinkers infused their theories with ideas and concepts taken from Gestalt psychology. Arnold Schoenberg distinguishes between what he refers to as motive, Gestalt, and Grundgestalt. He characterizes *motive* as "the smallest part of a piece or section of a piece that, despite change and variation, is recognizable as present throughout...gestalten and grundgestalten are usually composed of several motive forms; but the motive is at any one time the smallest part."¹² The motive is the basic formal unit that manifests itself in melody, harmony, textural contours, and other musical parameters. What we typically refer to as a "motive" in contemporary music-theoretical discourse, Schoenberg refers to as a *gestalt*:

A gestalt usually consists of more than one statement of the motive. Often there are various forms of the motive (for example, inversion or augmentation or diminution of the interval, or both, rhythmic broadening or contraction), but it often consists merely of a motive chain. In any event, a gestalt will have to have a characteristic feature to justify

¹¹ Wertheimer, "Gestalt Theory," 4. Emphasis Original.

¹² Schoenberg, The Musical Idea and the Logic, Technique, and Art of its Presentation, 169–170.

its name: a striking interval or interval progression or a striking rhythm or rhythmic progression.¹³

Thus, a motive might consist of two pitches forming a certain interval, and this interval might recur (at different pitch levels) with a specific rhythmic pattern to constitute a gestalt. Schoenberg describes *Grundgestalten* as "such gestalten as (possibly) occur repeatedly within a whole piece and to which derived gestalten can be traced back."¹⁴ This term is often translated as "ground-form" or "fundamental idea." The Grundgestalt is what today might be referred to as the prime form of a motive, which is often transformed to generate other related (i.e., "derived") motive forms. In Schoenberg's estimation, Grundgestalten originate on the surface of a musical work and are germinal—they are musical "seedlings" that ultimately flower into different, but related, musical parts. In his theoretical writings, Schoenberg explores the part-whole relationships of musical entities in great detail.

Wertheimer invites us to imagine "a Beethoven Symphony where it would be possible for one to select one part of the whole and work from that towards an idea of the structural principle motivating and determining the whole."¹⁵ The Gestalt notions of part-whole relations and hierarchical recursion—latent in the quotation from Wertheimer—seem to have profoundly influenced the musical philosophies and theories of Heinrich Schenker: one can easily conflate Wertheimer's postulation with Schenker's conception of the *Ursatz* (fundamental structure) as a musical object that is replicated on multiple levels of structure. Schenker was likewise influenced by the work of Immanuel Kant (the writings of whom played a formative role in the development of Gestalt theory).¹⁶ In his *Critique of Judgment*, Kant claims that only the *genius*

¹³ Ibid., 169.

¹⁴ Ibid.

¹⁵ Wertheimer, "Gestalt Theory," 11.

¹⁶ Volume I of Schenker's *The Masterwork in Music* opens with the following unreferenced quote from Kant: "The danger here lies not in being disproved, but rather in not being understood." See Heinrich Schenker, *The*

can create a work of fine art. According to Kant, "genius is the innate mental predisposition through which nature gives the rule to art."¹⁷ In other words, nature is the higher power that works through the genius-artist to give true beauty to art; the artist neither creates nor rationally understands the "rules" governing his/her own artistic production.¹⁸ Like Kant, Schenker viewed nature as the generative force guiding the production of art. Schenker's theory, which is most effectively applied to tonal music spanning Bach to Brahms (approximately 1725–1900), begins with the *Naturklang*, or "chord of nature." This chord is a naturally occurring acoustical phenomenon—it is nothing more than a single pitch and its overtone (or harmonic) series. The lower portion of the overtone series yields the first, third, and fifth diatonic scale degrees, which collectively constitute a major triad-the principal sonority in tonal music. Schenker surmised that the Naturklang, as the ultimate source of all "true" music, is given as a unitary whole.¹⁹ According to Schenker's theory, every pitch has the "urge to produce unending generations of overtones, [and] one might compare this urge to that of animals, for it appears in fact to be in no way inferior to the procreative urges of human beings."²⁰ This "urge" gives rise to the Ursatz—a skeletal foundation lurking deep in the background of a musical work and expressed on multiple levels of the musical structure.

The Ursatz is conceived as a unified structure with specified relations between its component parts: the structural tones of the three- or five-line *Urlinie* (fundamental line) and the *Bassbrechung* (bass arpeggiation). The Ursatz is essentially a horizontalized embellishment of

Masterwork in Music: A Yearbook, ed. William Drabkin, trans. Ian Bent et al., Cambridge Studies in Music Theory and Analysis, ed. Ian Bent, no. 5 (Cambridge and New York: Cambridge University Press, 1994), 1.

¹⁷ Immanuel Kant, *Critique of Judgment*, translated by Werner S. Pluhar (Indianapolis: Hackett Publishing Company, 1987), 174.

¹⁸ Ibid., 175.

¹⁹ Allen Cadwallader and David Gagné, *Analysis of Tonal Music: A Schenkerian Approach* (New York and Oxford: Oxford University Press, 1998), 362–367, 380.

²⁰ Heinrich Schenker, *Harmony*, translated by Elisabeth Mann Borgese (Chicago: University of Chicago Press, 1954), 37.

the Naturklang: the overtones of a single pitch are linearized (in descending order) and filled-in with passing tones to yield the Urlinie, while the bass arpeggiates the first partial (fundamental) and third partial of the overtone series (the root and fifth of a major triad). According to Schenker, Auskomponierung (composing-out) involves the visceral elaboration of the Ursatz (on the part of the composer), which explains why different pieces of music are distinctive even though they actually evolve from the same Ursatz. Schenker contends that the "genius" composer responds intuitively to the powerful urges of the Naturklang and progressively elaborates the Ursatz until the complete work is eventually born. In the words of music theorist Lawrence Zbikowski: "According to Schenker, creating a musical composition that both conformed to and properly expanded on an Ursatz was the work of genius. The best a nongenius could do was to come to an appreciation of the finished work by retracing the successive elaborations of the Ursatz."²¹ Thus, while Schenkerian analysis entails reducing a musical work to various middleground levels (the levels of structure that abide between the Ursatz and the musical surface), it is ultimately aimed at uncovering the unique manner in which the Ursatz is composed-out and elaborated upon in a given work of music.

Like Kant, Schenker believed that nature (the Naturklang) gives the rule (fundamental structure) to art (music) through the genius (composer). Like Schoenberg, Schenker conceived of music in terms of part-whole relations—a conception steeped in Gestalt psychology. Schenker's linear analysis, based on exposing the Auskomponierung of the ever-present *Ursatz* (as the background element governing middleground and surface-level phenomena), seems fundamentally at odds with Schoenberg's conception of the multilevel proliferation of a distinct germinal, generative motive (a foreground element identified as the "starting point"). However,

²¹ Lawrence M. Zbikowski, *Conceptualizing Music—Cognitive Structure, Theory, and Analysis* (Oxford: Oxford University Press, 2002), 128.

several music theorists have synthesized Schoenberg and Schenker's views: Janet Schmalfeldt has shown that the two seemingly disparate methods may indeed be reconciled and become notso-strange analytic bedfellows.²²

Leonard Meyer explicitly invokes the principles of Gestalt psychology in an effort to account for pattern perception and expectation during the musical experience.²³ His efforts reflect the fact that one need not adopt all the fundamental tenets of Gestalt Theory (particularly the notion that learning obtained through sensorimotor, environmental, and cultural experience does not affect the perception of form) to apply its empirically tested "laws" of perception to musical thought and discourse. Meyer chiefly applies Gestalt laws to describe melodic shape and continuity, rhythmic and metric continuation, and multiple types of musical closure (melodic, rhythmic, harmonic, and formal).

Foundational Principles

The four fundamental principles of Gestalt perception are: emergence, reification, multistability, and invariance. Emergence is the immediate perception of a group of objects or events as a unified whole. This occurs without the conscious amalgamation of an object's component parts to form the whole in perception. The example posited earlier—that of initially perceiving a chair as a chair, instead of as the complex interrelationships among four legs, a horizontal seat, and a vertical back—exemplifies the core concept of emergence. Bregman states: "Emergent features...are global features that arise at a higher level when information at a

²² Janet Schmalfeldt, "Towards a Reconciliation of Schenkerian Concepts with Traditional and Recent Theories of Form" Music Analysis 10.3 (1991): 233-287. For a discussion of the role of motive in Schenkerian analysis, see Richard Cohn, "The Autonomy of Motives in Schenkerian Accounts of Tonal Music," Music Theory Spectrum 14.2 (1992): 150–170. ²³ Leonard Meyer, *Emotion and Meaning in Music* (Chicago: The University of Chicago Press, 1956), 83–156.

lower level is grouped...they are based on the relations among the elements of a pattern."²⁴

Figure 3-1 presents a concrete visual example of emergence. In this picture, we immediately see a dog—most likely a Dalmatian—sniffing the ground, even though the picture consists of little more than black "splotches" on a white background: the form of the dog *emerges* from the formal assemblage of the splotches. The figure of the dog is an emergent property: it is not a property of any of the individual black blotches but appears from the grouping of those blotches. Spatial gestures are gestalt structures (wholes consisting of interrelated parts) that are received as unified wholes via the process of emergence. A gesture emerges over time—assuming a kinetic profile based on the dynamic grouping of spatially discrete events. Directed motion is the emergent feature of temporally successive sonic events grouped together in perception.



Figure 3-1: Emergence

Reification reflects the constructive, generative aspect of perception and often involves the perceptual process of "connecting the dots" or "filling in the missing piece(s)." Reification

²⁴ Bregman, 138, 172.

occurs when the percept contains more spatial information than the distal, sensory stimulus (the object of perception) actually contains. Figure 3-2 provides a visual demonstration of reification.



Figure 3-2: Reification

In Figure 3-2, we automatically perceive a triangle, although no triangle has actually been drawn. The spatial orientation of the "missing slices" of each of the three "pies" relative to one another gives form to the triangle. I shall argue that the perceived emergent motion of a spatial gesture results from the automatic process of reification: the shifting location of sonic activity within an ensemble of stationary performers gives rise to a composite, time-dependent spatial shape—even though no actual locomotion of a sound source occurs within the ensemble space.

Multistability, also referred to as "multistable perception," gives rise to *figure-ground reversals*. Many visual illusions result from a conflict between two or more viable interpretations; multiple candidates for the assessment of what an object is, or how a figure is spatially oriented, coexist and effectively compete with one another for perceptual prominence. An object cannot be perceived as two totally different things at the same time. When one objectinterpretation prevails (that which is perceived as the *figure*), the other candidate is relegated to the *ground* (or, "background"). Two common pictorial illusions demonstrate the phenomenon of multistable perception: the Necker Cube and Rubin's Figure, as shown in Figure 3-3. Due to the ambiguous visual cues supplied by each pictured object, multistability (figure-ground reversal) allows for two different interpretations of each picture. The brain will not accept both interpretations simultaneously—it works instead to determine which of the two options is superlative.



Figure 3-3: a) Necker Cube and b) Rubin's Figure

The issue with the Necker Cube illusion in Figure 3-3(a) concerns not what kind of object is being seen, but rather its orientation in space: are we viewing it from the upper right or from the lower left? (We likewise infer a sense of three-dimensionality where it does not actually exist.) Rubin's Figure, presented in Figure 3-3(b), invites us to see, alternatively, two totally different scenes: *either* two people with identical profiles facing one another (as if poised to kiss) *or* a single vase. Due to the principal of *exclusive allocation*, the line separating white from black space in Rubin's Figure may be allocated to either the shape of the vase or the faces—never to both simultaneously. However, the *belongingness* of the line may shift (in perception) from one

figure to the other.²⁵ Figure-ground reversals occur in perception—the figure itself does not change. Music theorists Carl Schacter, Kofi Agawu, and Peter H. Smith (among others) have applied these concepts to musical thought in an effort to account for, or refute the prospect of, ambiguity in music.²⁶

Invariance is an important property whereby a single object is recognized irrespective of its spatial (and temporal) orientation. An object may be moved (relocated), turned-around backwards, flipped upside-down, viewed in different lighting, enlarged, shrunk, stretched, and deformed (to a limited extent) and still be recognized as the same object. Invariance is critical to defining set-class (or, category) membership. A musical work may be conceived as being coherent and unified when it contains a preponderance of related motive forms, and we often envisage a melodic motive as an object that is, or might be, manipulated in musical space.²⁷ We can transpose/sequence (relocate), retrograde (reverse), invert (flip vertically), temporally augment or diminish, intervallically expand or compress, fragment, and otherwise deform a melodic motive.²⁸ Regardless of the transformative procedure (or combination of procedures) applied to a motive, the resultant form will share one or more invariant properties with the original version (which might be the prime form, or Grundgestalt), and will therefore be recognized either as the same object (albeit in altered form) or as a similar, related object.²⁹

²⁵ Bregman, 12–13.

²⁶ See Kofi Agawu, "Ambiguity in Tonal Music: A Preliminary Study," in *Theory, Analysis, and Meaning in Music*, ed. Anthony Pope (Cambridge: Cambridge University Press, 1994), 86–107; Carl Schacter, "Either/Or," in *Unfoldings*, ed. Joseph N. Strauss (Oxford: Oxford University Press, 1999), 121–133; and Peter H. Smith, "You Reap What You Sow: Some Instances of Rhythmic and Harmonic Ambiguity in Brahms," *Music Theory Spectrum* 28.1 (2006): 57–97.

²⁷ The proliferation of related motive forms is not necessarily the sole criterion for coherency.

²⁸ This list of transformative procedures is not exhaustive. Although my abstract example utilizes a melodic motive, many of these operations may also be applied to rhythmic patterns, timbral elements, and spatial gestures.

²⁹ This raises the question of whether we conceive of related motive forms in terms of their similarity to one another or we perceive the actual transformative procedures that convert one form into another. Does a musical work contain multiple, related motive forms or just one motive that is perpetually permuted and renegotiated? Perhaps the aural, diachronic experience of music enables us to hear processional and transformative activity, while the

Grouping principles

In addition to the four preceding fundamental principles, Gestalt theory provides two laws explaining the manner by which objects are grouped together in perception: the Law of Similarity and the Law of Proximity.

The Law of Similarity states that we tend to group similar or identical elements together. Visually, the similarity between objects could exist in the realm of shape, color, size, texture, brightness, etc. Aurally, similarity might exist in the domain of timbre, dynamic level, chord quality, articulation, etc. Bregman notes: "Sounds of similar timbres will group together so that the successive sounds of the oboe will segregate from those of the harp, even when they are playing in the same register."³⁰ An ensemble *area* (defined in Chapter 2 as an "inflated point"— consisting of like instruments playing identical musical parts), such as the cello section of a symphony orchestra, is easily perceived as a whole due to the similar sonic attributes and musical activity at each discrete point within the area. The points are likewise perceptually grouped (visually and aurally) by virtue of the fact that they are geographically close to one another, which leads to the next principle of grouping.

The **Law of Proximity** holds that we group together objects that are spatially or temporally near to one another. Objects that are close together in space and/or time often appear to belong together. Bregman observes:

If the analogy [from vision] to audition is a valid one, this suggests that the spatial dimension of distance in vision has two analogies in audition. One is separation in time, and the other is separation in frequency. Both, according to this analogy, are distances, and Gestalt principles that involve distance should be valid for them.³¹

synchronic analysis of a musical score invites us to gauge the similarity of apparently co-existing motive forms. Further investigation into this query lies beyond the scope of this dissertation.

³⁰ Bregman, 19.

³¹ Ibid. Bregman does not acknowledge that perceived distance in pitch space is metaphorical (as is conceiving of time in terms of space).

The consecutive tones of a melody are grouped together into a coherent stream due to their proximity in pitch space. This is perhaps why conjunct motion is optimally "melodic": conjunct motion preserves pitch-space grouping, whereas disjunct motion threatens to compromise the coherent streaming of a melody. A compound (or polyphonic) melody creates the illusion of two separate melodies due to the disjunct motion that parses the melody into two distinct groups in pitch space. The (typical) ensemble space is readily perceived as a unified group due to the spatial proximity of performers on the stage (within the performance space). In addition to proximity in pitch space, a melodic motive is easily discerned as a unified entity when its component notes are close together in time. A spatial gesture, as a type of motive, is formed by the succession of temporally proximate sound events emanating from discrete points within the ensemble space (these sound events may or may not be similar in terms of pitch, duration, timbre, dynamic level, etc.).³² The direct succession of sounds is equivalent to *relative* temporal proximity—irrespective of the amount of time separating events. Spatial gestures comprising successive events at adjacent (i.e., spatially proximate) ensemble points are perhaps more salient than those featuring disjunct spatial motion (in which an ensemble point is "skipped") due to the coordination of spatial and temporal proximity. Such gestures that likewise exhibit a marked degree of similarity among constituent and consecutive sonic events are even more salient due to the complementation of the Law of Similarity and the Law of (spatio-temporal) proximity.

Wertheimer states that both of the above grouping principles hold for auditory organization.³³ At times, these two laws may be in conflict with one another: dissimilar objects may be grouped together if they are adjacent or adjoining, and distantly located objects may be

³² Spatial gestures may unfold over longer time spans and/or at deeper levels of the spatial structure and therefore might *not* comprise temporally proximate sonic events.

³³ Max Wertheimer, "Laws of Organization in Perceptual Forms," in *A Source Book of Gestalt Psychology*, ed. and trans. Willis D. Ellis, with an introduction by Kurt Koffka (London: Routledge & Kegan Paul Ltd., 1967), 74.

grouped together if they are somehow similar. Such perceptual conflict can create the type of multistable perception (figure-ground reversals) cited above: we can mentally "switch back and forth" between grouping objects either by similarity or proximity. A musical example of such conflict might arise during the experience of a pointillistic composition. As we structure the incoming musical sounds, are we to group notes that are temporally contiguous (temporal proximity), those that are registrally close together in pitch space (spatial proximity), or those displaying the same dynamic level and tone color (similarity)?³⁴ The interactions (and potential conflicts) of sound-event grouping on the dual bases of similarity and proximity come into play when we recognize consecutive events occurring at different points within an ensemble as a unified spatial gesture. The similarity and spatial proximity of constituent events may affect grouping tendencies; however, I argue that the primary factor influencing the perception of spatial gestures is the temporal proximity of successive events—events that might *not* be qualitatively similar or spatially proximate.

Beyond the foundational concepts of emergence, reification, multistability, and invariance, and the grouping principles of similarity and proximity, Gestalt psychology breaksdown into slightly more abstract laws. The most basic rule of Gestalt perception is that of **Prägnanz**, which, when translated literally, means "pregnant" (pregnant with *meaning*, not pregnant with child). Prägnanz is generally understood to mean "good form." From a perceptual standpoint, our minds seek to experience figures in as "good" a way as possible: our default experiential preference is for well-formed objects of sensation and perception. "Good," or "wellformed," can mean many different things: regularity, order, simplicity, symmetry, etc. The "goodness" of a whole is directly proportional to its parts' tendencies to group together in

³⁴ Pitches close to one another in register might also exhibit a higher degree of timbral homogeny, thereby enhancing their perceived similarity.

perception. As demonstrated in some of the visual illusions presented above, our perception of objects does not always conform to the reality or actual status of those objects. Thus, according to the Gestaltists, our perceptual apparatus seeks to impose "good" formal values on the objects of our sensation and perception. This is particularly true for vision—our chief and most understood sense modality—but is also true of audition. In general, Prägnanz entails *simplicity* (a good figure is less complex than a poor figure) and *predictability* (predictable figures, with low information content, have good form). Good form is directly proportional to perceptual salience: the better the form of a figure, the greater its perceptibility. Following are some specific Gestalt laws and principles of good form, the majority of which have a direct bearing on my delineation of spatio-gestural categories.

Principles of "Good Form"

Law of Closure: We prefer enclosed and completed figures, such as a perfect circle or a square. Koffka claims that "closed areas seem to be self-sustaining, stable organizations."³⁵ Wertheimer states: "If [items] A, B, C, [and] D are given and AB/CD constitute two self-enclosed units, then *this* arrangement rather than AC/BD will be apprehended."³⁶ Closed figures seem complete. Our minds often add missing elements in order to complete a figure.

Law of Continuity: Figures and grouped objects that continue in one direction, without an abrupt shift in directionality, are strong. Examples of figures displaying "good continuation" include straight lines, curves, circles, and spirals. The issue of continuity is important to Bregman's theory of auditory streaming, as outlined in Chapter 2. In this context, continuity may pertain either to the sustainment of a particular sound's sonic attributes or to the continual,

³⁵ Koffka, 151.

³⁶ Wertheimer, "Laws of Organization," 83. Emphasis Original.

uninterrupted change of some property (or properties) of an object or event. According to Bregman: "The Gestalt psychologists [recognized] that the continuity or smoothness of a change promoted the perceptual integration of the changing experience."³⁷ In the words of Koffka:

In the stimulus there is change, and to this stimulus change there corresponds perceptual segregation, just as in spatial organization segregation [is] produced by inhomogeneity of stimulation. In spatial organization homogenously stimulated parts [segregate] themselves from the rest of the field and [become] units.³⁸

If a portion of a straight line is occluded by another figure, the line is perceived to continue behind the object. The same holds for auditory events: if parts of a soft signal are either masked or removed and replaced by loud signals of a differing quality, the "listener hears the softer sound as continuing unbroken *behind* the louder one."³⁹ In this situation, the interrupted signal may be a sustained, uniform sound or a continuously changing sound, such as a glissando (frequency modulation) or crescendo (amplitude modulation). The listener often infers the process of change to continue "behind" interruptive sonic events.⁴⁰

The indefinite continuation of a figure is often implicit and inferred: our minds often continue and extend a pattern, even after it has stopped. A line may appear to extend in either spatial direction, and a glissando or scale may seem to continue in either spatio-temporal direction—appearing as a fragment of a larger whole not actually present. Meyer states: "Among other things this law helps to account for our being able to hear separate, discrete stimuli as continuous motions and shapes."⁴¹ Meyer applies this statement to melody, where the temporal succession of discrete pitches initiates the mental process of perceiving a melodic line.

³⁷ Bregman, 133. Bregman quite thoroughly relates Gestalt principles to auditory streaming. See Bregman, 196–203.

³⁸ Koffka, 434.

³⁹ Bregman, 345. Emphasis mine. Note the spatial distinction that Bregman assigns to the differing signals: the sound with the softer dynamic level is perceived as phenomenally located "behind" (i.e., "farther away") the intermittent loud sounds.

⁴⁰ Ibid., 348.

⁴¹ Meyer, 92.

Hatten echoes Meyer's claim: "It is natural to hear melody as though it were a single force or line traversing a space, rather than a sequence of distinct sound events...indeed, one might argue that this perception of musical motion is crucial to our experiencing of successive sounds as music, rather than as mere acoustic phenomena."⁴² Hatten goes on to characterize musical gestures (which are often composites of multiple musical elements) as "analog, as opposed to digital or discrete, [and thereby] continuous in a productive sense of continuity (i.e., not necessarily continuous sound, but continuity of shape, curve, motion across silence, etc.).⁴³ Thus, Meyer and Hatten conceive of melody and musical gesture as continuous, although these musical structures are composed of discrete items (such as a sequence of musical tones). Presumably, Hatten's "productive sense of continuity" applies to the listener's act of grouping discrete items into a continuous form. This act of grouping, I argue, is based on the Gestalt principles of grouping outlined herein as well as upon Bregman's concept of auditory stream segregation. These observations may also be applied to the perception of a spatial gesture as a "kinetic shape" weaving its way through space. The shape is a composite of multiple, discrete events occurring at separate points in ensemble space. These events are grouped by the listener to constitute a continuous motion. Spatial gestures in a one-dimensional ensemble space may exhibit the form of a unidirectional line, and gestures in multi-dimensional spaces may trace curvilinear, spiral, and circular paths (among others). Some types of gestures may imply spatiotemporal continuation.44

⁴² Hatten, Interpreting Musical Gestures, Topics, and Tropes, 128.

⁴³ Ibid., 124. Emphasis Original.

⁴⁴ Many artists have developed techniques intended to imply continuation (and/or infinity) beyond the borders which confine the work of art in question—whether those borders be the frame of a painting, the page of a poem, or the amount of real time in which a piece of music unfolds. Irrespective of medium, artists have succeeded in blurring the distinction between finite and infinite—expressing the latter within the former. Nonmusical examples include the stream of consciousness technique in the novels of James Joyce, the fragmented syntax of e.e. cummings's poetry, and the placement of lines and figures in the paintings of Piet Mondrian such that they are "cut-off" by the borders of the canvas. A musical example is Karlheinz Stockhausen's conception of "moment form." This is a

Law of Common Fate: This law is related to the Law of Continuity but applies to grouped objects that are changing together in time. (The Law of Continuity applies chiefly to a single object or event that changes over time.) Elements moving in the same direction, or otherwise changing in time concomitantly, are grouped together (in audition, they are grouped into a single stream via *simultaneous integration*, as explained in Chapter 2). As a visual example of common fate, consider the observance of cloud formations in the sky: the slower moving clouds at a high altitude appear distinct from lower clouds moving at a relatively faster pace. Although the clouds at both levels of altitude might be moving in the same direction, the fact that they are moving at different rates of speed sets them apart.⁴⁵ We will find that the differing velocities (along with other differences) of simultaneously occurring spatial gestures provide for their delineation in perception.

Law of Symmetry: Symmetrical, balanced objects exhibit good form. A symmetrical object will map onto itself across some axis of inversion. Humans are most adept at visually perceiving symmetry across a vertical midline. This preference for left-right symmetry is likely due to our bodily experience: our bodies are (roughly) symmetrical across an imaginary, vertical axis. The notion of symmetry in music often arises in discussions of formal structure and frequently denotes the temporal proportions of different musical sections within the whole of a movement or work. We might also attend to the symmetrical arrangement of ensemble members—assuming that the players are spaced roughly equally in either collinear or curvilinear fashion. For instance, a quartet is balanced by two players on either side of an imaginary line of

compositional technique (fully developed in Stockhausen's *Momente*) in which a piece of music contains a single, or several contrasting, "moments." Each moment is like a slice of a larger piece of music—it does not feature a proper beginning, nor does it involve a prepared ending. In this sense, the moment implies continuation in both temporal directions (past and future)—the moment expanding indefinitely into hours, weeks, years, infinity. The result is the implication of eternal continuation and expansion.

⁴⁵ Levine, 195.

symmetry existing *between* the second and third performers (when viewing the ensemble from left to right). This form of symmetry will hold for any even-numbered ensemble that is positioned in a relatively balanced fashion, with performers roughly equidistant from one another. In a quintet setting, the ensemble is still balanced by two players on each side (performers 1 and 2 on the left, performers 4 and 5 on the right), but the imaginary line of symmetry now extends *through* the third performer. This type of symmetry will hold for any odd-numbered ensemble. The types of spatial gestures and gestural interactions afforded by these two different forms of symmetry (axis between performers, axis through central performer) shall be explored below. Furthermore, in Chapter 5, I shall explore the notion that gestures with opposing directionalities "balance" one another.

3.2 Spatial Gestures

Within an acoustic ensemble, and assuming stationary (immobile) performers, spatial gestures emerge from the consecutive activity of performers within the ensemble. Various gestures arise from—and are differentiated by—the specific orderings, in time, of sonic events occurring at separate points in ensemble space. The temporal succession of spatially discrete events may involve contiguous and/or nonadjacent ensemble members. As a unified whole comprising ordered parts (gestalt structure), a spatial gesture is directed motion through an ensemble. Koffka states: "A moving body never *is* at one place, it always *passes through* places. Just so, a melody, until it is finished and comes, so to speak, to a stop, never *is* at one note but passes through it."⁴⁶ In Koffka's estimation, a melody is a single moving body passing through fixed points (notes) within some region (pitch space). A given melody is distinguished from (and/or related to) other melodies by virtue of the specific sequence of pitches through which it

⁴⁶ Koffka, 434–435. Emphasis Original.
passes: the "path" that a melody weaves through pitch space *is* the melody. A spatial gesture is a complex event (or perhaps, an *event complex*) that functions as a "moving body" passing through the "places" (ensemble points) of an ensemble space. The discrete locations of the ordered events forming a spatial gesture define the holistic nature of the gesture—its patterned shape and direction(s) of motion. This emergent motion and kinetic shape *is* the Gestaltqualität of the spatial gesture. Divorcing the parts (ordered event locations) from the whole (spatial gesture) dissolves the gesture's Gestaltqualität, which might be thought of as the glue that binds the parts together. McCabe claims: "If we actually (rather than conceptually) abstracted parts to construct wholes, we would violate their intrinsic reciprocity and change the nature of both the parts and the whole."⁴⁷ Speaking of musical gesture in the more generalized sense, Hatten states: "Gestures may be comprised of any of the elements of music, although they are not reducible to them; they are perceptually *synthetic* gestalts with *emergent* meaning."⁴⁸ So, too, with spatial gestures (as specific types of musical gesture): reducing a spatial gesture to its component events destroys the motion that emerges from the particular ordering of those events.

Koffka explains that motion can arise from the "*cumulative disposition* generated by the ordered sequence of [objects or events]."⁴⁹ He elaborates:

If an object moves through our field of vision so that it stimulates successively the retinal elements a, b, c, d...and is seen as an object in motion we might explain this...by saying that each element if stimulated alone would give us the experience of a point at a certain place. But because of the serial stimulation each point stimulus after the first one would co-operate with the dispositions left over from the preceding ones. At any moment...the experience of the last position should be the only one of which we are directly aware, but it comes before consciousness as part of a specific whole, the path of motion, and thereby endowed with a specific property, its velocity.⁵⁰

⁴⁷ McCabe and Balzano, 9.

⁴⁸ Hatten, Interpreting Musical Gestures, Topics, and Tropes, 94. Emphasis Original.

⁴⁹ Koffka, 432.

⁵⁰ Ibid., 435.

Koffka claims that the ordering of object location allows for the perception of motion. It is from such an ordering that directed motion—along a "path" and with a specific "velocity"—emerges. In his example, the described motion of the object is continuous through the subject's visual field. Let us modify Koffka's account in keeping with Gestalt theory. Suppose we replace "retinal elements" a, b, c, and d with four lights distally situated in a horizontal row and labeled (from left to right): a, b, c, and d. We then cause the lights to flash in the following sequence: a, b, c, d. In so doing, we perceive a motion directed from left to right in our visual field. The strength of the perceived motion would be variable depending upon the relations among the individual flashes: the rate of the ordered flashing, the degree of overlap between consecutive flashes, the spatial distance between adjacent lights, the similarities of color or brightness between the lights, etc. For example, if all of the lights were the same color, the perception of motion would likely be more pronounced. Additionally, if the flashings of each light slightly overlap one another in time (such that the light b begins to brighten as light a begins to fade, and so on), the motion might appear more continuous. However, in spite of these and other variable conditions affecting the salience of perceived motion, the successive flashing of the lights would impart some sense of apparent motion just by virtue of the spatio-temporal ordering of flashes.

To account for the perception of motion induced by a spatial gesture, the lights as objects in the visual field are replaced by sound sources in the auditory scene (the potential visibility of which is presently dismissed). The succession of sound events produced by these discrete sources, and sensed through audition, likewise engenders the perception of motion. The degree of relatedness between the consecutive sounds does affect the salience of the perceived motion, but I shall argue that unrelated or dissimilar sounds can still form a composite motion: spatial

128

motion is fundamentally abstractable from other sonic qualities and musical happenings, both in perception and in analysis.

What types of sonic events may group together over time to form a spatial gesture? More specifically, what kinds of relations between consecutive events (temporal, dynamic, timbral, etc.) serve to promote event grouping and impart a sense of motion to the spatial construct of a musical performance? Harley states: "Spatial sound movement is usually created by means of overlapping dynamic envelopes; the dynamic peaks mark the shift of sound from one spatial location to another."⁵¹ This statement presumes that a timbrally equivalent sound is "shifted" from an instrument in one location to the exact kind of instrument in another location. Indeed, many composers in the twentieth century (Stockhausen and Xenakis in particular) have employed this technique to deliberately spatialize a musical sound.⁵² The procedure is similar to-and in large part stems from-the technique of "panning" a sound from one channel (speaker) to another in the electronic medium. When successful, the resultant sonic illusion is that of a single, sustained sound moving continuously from one point in space to another, although it actually results from the coordinated and overlapping activity of two separate and static performers (or speakers). Figure 3-4 provides an example of illusory continuous motion produced by overlapping dynamic envelopes from Stockhausen's Gruppen für drei Orchester. This excerpt shows the activity of two of the three orchestras making up the total ensemble of Gruppen. For the present illustration, we may view each orchestra as a discrete hyperpoint. In performance, the three orchestras are spatially separated such that Orchestra I is located to the audience's left, Orchestra II is centered in front of the audience, and Orchestra III is situated to

⁵¹ Harley, "Space and Spatialization," 294.

⁵² It is interesting to recall from Chapter 1 that Henry Brant stresses the importance of placing *different* kinds of instruments in different spatial locations. This practice preempts the potential for the type of continuous spatial motion described by Harley.

the right of the audience. Figure 3-4 reveals the musical activity of Orchestra II, located at the top of the score excerpt, and Orchestra III, located at the bottom.



Figure 3-4: Stockhausen, Gruppen, section 119, mm. 1-4

In the second measure of section 119, a chord is sounded in the brass section of Orchestra III (2 horns, 2 trumpets, and 2 trombones). The chord begins triple piano and rapidly crescendos (in less than 3 seconds) to double forte. Prior to the chord reaching its dynamic climax and then abruptly terminating, the exact same chord—with identical orchestral voicing—begins in Orchestra II (m. 3 of 119). The chord at Orchestra II begins triple piano while the chord at

Orchestra III is likely somewhere in the dynamic neighborhood of forte. The Orchestra III chord ceases as its successor in Orchestra II approaches the piano dynamic on its way to double forte. Although not shown in Figure 3-4, the selfsame chord, with identical dynamic markings, appears in Orchestra I in m. 4 of 119. In this passage, the illusion of sonic motion is generated by a uniform sonic event appearing consecutively, with temporally overlapping dynamic envelopes, at three separate hyperpoints in ensemble space. The audience experiences an angular (or "horseshoe" shaped) motion spanning from right to left.

Deliberately spatialized passages, such as that presented in Figure 3-4, can produce a convincing impression of continuous motion. However, my allowance of what may constitute a spatial gesture is quite broad: the primary condition for the formation and perception of a spatial gesture is the relative temporal proximity of successive sound events. As such, a spatial gesture that features a significant discrepancy of timbre, duration, dynamic, and/or pitch among its constituent events will not be an "auditory stream" in Bregman's strict sense of the word (see Chapter 2).⁵³ Yet, the gesture may still form a spatial shape and effectuate motion through an ensemble: the composite motion (Gestaltqualität) is abstracted from all other sonic and musical qualities. A sense of spatial motion may arise without a congruence of sonic profiles and/or "overlapping dynamic envelopes" between successive musical events. However, a spatial gesture in which the component events are somehow similar to one another could be considered a "mobile auditory stream," which is likely more perceptually salient than a "non-stream gesture" due to the coordination of the two grouping principles of Gestalt theory: (temporal) proximity and similarity.⁵⁴ Due to these and many other factors, some spatial gestures are quite transparent and easily perceptible while others are less immediately obvious (but not necessarily

⁵³ Bregman, 9.

⁵⁴ Bregman's conception of *sequential integration*, one of two types of integration that form an auditory stream, adheres to the Gestalt principle of grouping on the basis of similarity.

any less pertinent to spatial analysis). In Chapter 4 I shall address multigestural activity, or "polystreaming," in which two or more spatial gestures occur at the same time. Generally, the component gestures of a polystreaming complex must likewise be auditory streams (differentiated by some shared attribute among events other than temporal proximity) to avoid ambiguity and allow for the proper identification of the distinct gestures—which is requisite to the accurate portrayal of their interactive relationships.

In what follows, I present an integer-notational method for labeling specific spatial gestures in order that they may be easily identified and compared to other gestures. I then define the necessary conditions for assigning a given gesture to a specific gestural category. Throughout this dissertation, the most extensive consideration is given to small ensembles (trios, quartets, and quintets) that are oriented horizontally from left to right across the stage on a single vertical level and depth plane (this is the most common and traditional spatial distribution of musical ensembles). Occasionally, spatial analysis requires the *normalization* (or, "one-dimensionalization") of the horizontal plane, which effectively entails ignoring—for analytic purposes—any discrepancies of depth resulting from the slight arcing of an ensemble. For instance, string quartets and brass quintets typically configure themselves with a slightly convex arc so that all ensemble members can see one another and intra-ensemble "communication" is not impaired. With some pieces, this normalization procedure clearly will not work—particularly with pieces in which two or more depth planes are active.⁵⁵

My method for labeling spatial gestures derives from contour theory and analysis.⁵⁶ According to Schoenberg, melodic contour and rhythmic profile are the two most salient,

⁵⁵ Ensemble spaces featuring multiple depth planes and/or vertical levels shall be explored in Chapter 4.

⁵⁶ My work here is primarily informed by the contour theory for pitch space that was first developed by Robert Morris and was subsequently expanded by Elizabeth Marvin and Paul Laprade. See: Robert D. Morris, *Composition with Pitch-Classes: A Theory of Compositional Design* (New Haven and London: Yale University Press, 1987);

distinguishing, and memorable characteristics of a motive.⁵⁷ Accordingly, contour analysis (which was not proffered by Schoenberg) ignores any relationships (such as intervallic) among pitches other than the relative heights of pitches and their sequential ordering.⁵⁸ Integer notation is utilized to show the relative heights of consecutive pitches within a *contour segment (CSEG)*. Consider the melodic fragment given in Figure 3-5.



Figure 3-5: CSEG <0132>

The lowest note of the fragment is given the value 0. The highest pitch in the fragment is given the value n - 1, where n is the total number of pitches in the fragment. In this case, the fragment contains four pitches, so the highest pitch (E-flat) is assigned the integer value 3. The values are then presented in sequential order. Thus, the label for this CSEG is <0132>.

While contour analysis depicts the vertical relationships among pitches in metaphorical

pitch space as they unfold over time, the analysis of spatial gestures portrays the (chiefly

horizontal) relationships between temporally adjacent events in "actual," physical space.⁵⁹

divorces horizontal pitch collections away from intervallic content and set-class membership.

Elizabeth W. Marvin and Paul Laprade, "Relating Musical Contours: Extensions of a Theory for Contour," *Journal of Music Theory* 31.2 (1987): 225–67; Robert D. Morris, "New Directions in the Theory and Analysis of Musical Contour," *Music Theory Spectrum* 15.2 (1993): 205–228. An alternative model can be found in Michael Friedmann, "A Methodology for the Discussion of Contour: Its Application to Schoenberg's Music," *Journal of Music Theory* 29 (1985): 223–248. See also Friedmann's response to Marvin and Laprade in Michael Friedmann, "A Response: My Contour, Their Contour," *Journal of Music Theory* 31.2 (1987): 268–274.

 ⁵⁷ Arnold Schoenberg, *Fundamentals of Musical Composition* (New York: St. Martin's Press, 1967), 113–115.
⁵⁸ Pitch-class set analysis abstracts collections of pitches away from contour, registration, etc. Contour analysis

⁵⁹ It is important to emphasize the point that I am appropriating an analytic system initially designed to account for events in pitch space in an effort to describe events in actual space. *Pitch space*, a metaphoric construct, is structured (or, grounded) by our concrete, embodied understanding of events in actual space, which is acquired through repeated (and acculturated) sensorimotor experience. One pitch is not actually higher or lower than another pitch, because the pitches are not actual objects (physical entities that are both visible and tangible) abiding in actual space. It is only through the deeply engrained conceptual metaphor PITCH RELATIONSHIPS (IN PITCH SPACE) ARE OBJECT RELATIONSHIPS (IN ACTUAL SPACE) that we mentally construe pitches in terms of height. (Further discussion of the interpretive issues surrounding this metaphor may be found in Chapter 5.) Therefore, I am constructing my methodology to account for gestures in physical space based on a method created to describe

Consider the diagram of a standard twentieth-century seating arrangement for string quartet shown in Figure 3-6.

Vln. II Viola Vln. I Cello

Figure 3-6: Common seating plan for a string quartet

Moving horizontally from left to right (and disregarding the trivial disparity of depth between the ensemble members that is created by the slight semicircular arc of the quartet), we assign integer values to each instrument, beginning with 0 for the leftmost instrument (in this case, the first violin). Since the total number of instruments (n) is 4, the rightmost instrument will receive the value 3 (n – 1). The corresponding integer values for each instrument are: Violin I = 0; Violin II = 1; Viola = 2; Cello = 3. Abstractly, each integer represents a *point* in ensemble space, and each point represents the relative location of a discrete performing member within the ensemble. Thus, the four-point ensemble space of the string quartet may be diagramed as presented in Figure 3-7.

1 2 0 3

Figure 3-7: Integer diagram of four-point ensemble space

Such an abstract diagram could depict *any* type of quartet, irrespective of its instrumental composition. In the case of the string quartet, we may refer to "point 2" when describing some

events and relationships in pitch space, which is, in the first place, metaphorically structured and conceptualized through knowledge of events in physical space. Similarly, Paul Miller evokes pitch-oriented concepts such as parsimonious voice-leading, melodic gap-fill, hierarchy within tonal pitch space, Fux's principle of non-coincident extremes in melodic construction, and scale structure (gapped, ungapped, etc.) when describing spatial events in Stockhausen's *Lichter-Wasser*. See Paul Miller, "The Analysis of Spatial Music in Works by Karlheinz Stockhausen," 6.

sound event issued by the violist. Figure 3-8 shows the normalization (one-dimensionalization) of the ensemble space, whereby the minimal (and relatively trivial) arc of the ensemble is eliminated and ignored for the sole purpose of simplifying gestural analysis.

0 1 2 3

Figure 3-8: Normalized integer diagram of four-point ensemble space

Any ordering of these integers represents a unique spatial gesture unfolding over any length of time. For instance, the notational label SG-|1320| would indicate the following diachronic sequence of sound events within the string quartet: second violin (1), cello (3), viola (2), and first violin (0). Such integer notation shall be called an **SG-label**, where "SG" simply stands for "Spatial Gesture." A sound event occurring at a given ensemble point may be virtually anything: a single note, the entrance of a theme, the sounding of a prominent motive, a sudden change in color, dynamic, or articulation, etc. Concrete examples of sound events shall be explored throughout the remainder of this dissertation.

Figure 3-9 presents an excerpt from Beethoven's *String Quartet no. 14 in C# Minor*, Op. 131 (1826).⁶⁰ Still assuming the ensemble seating plan depicted in Figure 3-6, m. 89 from the fourth movement of Op. 131 presents a gesture that passes from the far left to the far right of the

⁶⁰ This particular string quartet by Beethoven is treated to a thorough spatial analysis in Chapter 6. Many of the musical examples in this chapter are taken from the string quartet literature. In addition to boasting a wealth of substantive musical works and having endured as one of the most popular chamber-ensemble types for well over two centuries, string quartet ensembles typically (although not exclusively) abide by one of several possible seating arrangements. Such is not the case for other types of chamber ensembles, such as the brass quintet and woodwind quintet, where seating plans are more varied and inconsistent from one group to the next. For all string quartet examples in this document, unless stated otherwise, the following seating configuration shall be assumed (from left to right): violin 1 (point 0), violin 2 (point 1), viola (point 2), and cello (point 3). Potential problems arising from such an assumption, and exceptions to this standard seating plan, shall be addressed in Chapter 7. The left-to-right horizontal ordering of the elected seating plan corresponds to the standard top-to-bottom orientation of the staves in a string quartet score.

ensemble space. The SG-label for this gesture is SG-|0123|. In this particular example, each consecutive instrumental entrance consists of a highly similar sound event: a descending melodic octave in eighth notes with a sforzando dynamic marking. The latter three occurrences of this event feature the same pitch class (E). The ensuing left-to-right motion through the ensemble is quite pronounced despite the lack of overlapping dynamic envelopes, and the salience of the gesture is certainly enhanced by the similitude of musical activity at each ensemble point.



Figure 3-9: Beethoven, Op. 131, Mvt. 4, m. 89

Within the SG-label for any spatial gesture, the first and last integers represent the **temporal endpoints** of the gesture. The lowest and highest integer values depict the **spatial extremes** of the gesture: the lowest integer value represents the leftmost spatial extreme; the highest integer value corresponds to the rightmost spatial extreme. The spatial extremes of a gesture may or may not be the spatial extremes of the total ensemble space in which that gesture unfolds (a gesture might occur in an ensemble-space segment). Often, the spatial extremes and temporal endpoints of a gesture will be one in the same—these event points shall be termed the **spatio-temporal extremes** of the gesture. Such is the case with SG-[0123], presented above in

Figure 3-9. Likewise, the spatio-temporal extremes of this gesture are equivalent to the spatial extremes of the total ensemble space.

A spatial gesture consists of temporally successive events, and these consecutive events are grouped into directed "point shifts." A **point shift**, abbreviated "PS," is the virtual motion from one point to another in ensemble space. Two temporally successive events, each at a different point in ensemble space, function as a unit and constitute a point shift.⁶¹ An ensemble point does not actually shift inside the ensemble space (as might occur with mobile performers); it is the spatial location or position of an event that is transferred from one point to another. The point location of *some* sound event "shifts" from one point in ensemble space to another (the qualities of the sound events at each point may be totally dissimilar). A point shift, which is the smallest unit of motion within a gesture, may occur between adjacent and nonadjacent ensemble points. A point shift in a spatial gesture, then, is akin to Schoenberg's conception of motive as the smallest component of a musical gestalt. His "gestalt" corresponds to my "spatial gesture."

In accounting for part-whole relationships, Wertheimer writes: "The cells of an organism are parts of the whole and excitations occurring in them are thus to be viewed as part-processes functionally related to whole-processes of the entire organism."⁶² A point shift is the smallest "kinetic cell" within a spatial gesture; the aggregation of multiple point shifts (as "part-processes") forms a spatial gesture (as "whole-process"). If a point shift is altered (for example, if its component events are exchanged), the gesture as a whole will likewise be altered. This observation supports Wertheimer's assertion that "modifications of a part frequently involve

⁶¹ It might seem compelling to call such a unit an "event pair" instead of a "point shift." However, the sense of agency implicit in the term "shift" underscores my point that a spatial gesture is actually perceived as directed motion through space. Furthermore, a term such as "pair" seems to imply the temporally coincident presence of two objects (or events), which is not always the case with two events that form a point shift within a larger spatial gesture.

⁶² Max Wertheimer, "The General Theoretical Situation," in *A Source Book of Gestalt Psychology*, ed. and trans. Willis D. Ellis, with an introduction by Kurt Koffka (London: Routledge & Kegan Paul Ltd., 1967), 15.

changes elsewhere in the whole itself."⁶³ Gestural modifications—including those resulting from point-shift alterations—shall be explored in Chapter 4.

Of course, there is no *actual* motion between any two points, since the performers (as sound-event sources) remain static.⁶⁴ However, we likely perceive **apparent motion**, as described by the *Phi phenomenon*. The Phi phenomenon accounts for the illusion of motion that occurs without the actual displacement of an object or event. The Phi effect is a form of *stroboscopic movement* that is caused by the successive presentation of stationary stimuli.⁶⁵ The example of the flashing lights presented above (modified from the discussion by Koffka) is an example of apparent motion induced by stroboscopic movement. Such models are generally applied to visual perception, but analogs to audition are perfectly apt.⁶⁶ A point shift is

⁶³ Ibid., 14.

⁶⁴ In *Der Kleine Harlekin* (1975) for solo clarinet, Karlheinz Stockhausen instructs the soloist to rotate when performing certain strata of the work's compound melody. The performer revolves to his/her right (the audience's left) for lower notes and swivels to the left (the audience's right) for higher notes. Stockhausen's score indications are: "die tiefen Töne nach rechts, die hohen nach links" ("the low pitches to the right, the high pitches to the left"). Thus, registral distinction is coupled with some degree of spatial separation. The mobility of a lone performer can create spatialized activity. In this case, the performer's rotation results in a "split point" that enables pseudo point shifts (and oscillation-type gestures) at a single point. Under normal circumstances—in which performers are both stationary and relatively static—a point shift between two discrete performers (either performing as a duo or within an ensemble) is the smallest unit of spatial motion. See Karlheinz Stockhausen, *Der Kleine Harlekin für Klarinette* (Kurten: Stockhausen-Verlag, 1975), VI, 7.

⁶⁵ Levine, 292, 570, 574; Josef Ternus, "The Problem of Phenomenal Identity," in *A Source Book of Gestalt Psychology*, ed. and trans. Willis D. Ellis, with an introduction by Kurt Koffka (London: Routledge & Kegan Paul Ltd., 1967), 149. Sound events presented successively may not always be totally discrete: latter sound events might overlap or join with earlier sound events (resulting in an *accumulative gesture*, to be addressed below). In such cases, a sense of expansive motion or spatial growth may be inferred.

⁶⁶ Bregman draws an analogy between apparent motion in vision and melodic motion in audition (Bregman, 173– 181). This analogy is easily transferable to spatial motion in audition. Imagine the following scenario: with your eyes closed (to discount any localization cues provided by vision), you hear a sound occur to your front-left. After a few seconds, you hear the exact same quality of sound (identical sonic profile) occur again, but this time to your front-right. You would likely infer that a single sound source was responsible for generating each sound, and that this source had simply moved to a new spatial location prior to producing the second sound. When you open your eyes, you realize that two separate and stationary sources—each located at different spatial coordinates—had produced the identical sounds at different points in time. Your original conclusion, that a single sound source had relocated its position, was incorrect. Of course, this inference was undoubtedly strengthened by the fact that the two sounds were virtually identical (excepting their respective spatial and temporal locations). The sense of apparent motion would likely be more difficult to deduce when two qualitatively different sounds ensue from two separate sources and locations.

effectively the displacement of sound event location and shall be viewed as an instance of apparent motion.⁶⁷

Segmenting a larger spatial gesture into component point shifts enables us to quantify the amount of left-to-right motion and right-to-left motion between events, and it assists us in pinpointing changes in directionality. As an example, let us consider SG-[1320]. This gesture consists of four consecutive events (occurring at points 1, 3, 2, and 0) and three successive point shifts (from point 1 to 3, 3 to 2, and 2 to 0). The notation of individual point shifts employs a dash between the ordered point-activity; thus, the three aforementioned point shifts forming SG-|1320| are identified as PS 1-3, PS 3-2, and PS 2-0. A point shift may be conjunct (occurring between contiguous points) or *disjunct* (occurring between nonadjacent points).⁶⁸ Of the three point shifts comprising SG-[1320], PS 3-2 is conjunct, while PS 1-3 and PS 2-0 are both disjunct. We also see that the gesture contains two instances of right-to-left motion (PS 3-2, PS 2-0) and only one instance of left-to-right motion (PS 1-3). PS 3-2 is the point shift in which the gesture's only directionality change occurs. Since a point shift is the smallest unit of motion that a spatial gesture may contain, we might think of the point shifts forming a gesture as synonymous to the syllables that form a multisyllabic word. Just as a syllable may comprise more than one letter, a point shift contains more than one sound event. Since a point shift does have a specific

⁶⁷ Fundamentally, a point shift is a transformation, in space-time, of event location. The quality of the sonic event being translocated is abstract and highly variable. In Chapter 4 I shall discuss how spatial gestures may be altered, in both the temporal and spatial realms. However, it is important to bear in mind that a spatial gesture itself consists of multiple transformations (of location)—each in the form of a point shift.

⁶⁸ My application of the terms "conjunct" and "disjunct" in describing spatial events differs from that of Boulez, as described in Chapter 1. For Boulez, the use of these descriptors has nothing to do with distance—a conjunct spatial interval is one in which the two discrete events overlap one another in time, while a disjunct spatial interval occurs when there is a temporal pause between the offset of event 1 and the onset of event 2. In my analytic system, "conjunct" and "disjunct" has to do with relative distance within the ensemble space. If a point shift occurs between two adjacent ensemble points, regardless of their spatial proximity (actual distance between one another), the spatial motion is conjunct. If a point shift skips one or more ensemble points, the motion is disjunct.

directionality resulting from the ordering of its component event points, each point shift might be considered a small-scale spatial gesture in its own right.

The exact nature of the specific sound events at either point within a point shift is relatively inconsequential, given my claim that the Law of (temporal) Proximity, as a grouping tendency, overrides the Law of Similarity in the identification (perceptual and analytical) of spatial gestures. Wertheimer states: "It is...quite arbitrary what is coupled in simultaneity and succession. For the togetherness itself the 'content'...is really irrelevant."⁶⁹ Wertheimer does not discount the Law of Similarity altogether but highlights the fact that it may often compete with the Law of Proximity for perceptual dominance. The events constituting a spatial gesture need not be registrally, timbrally, or dynamically homogenous. We recognize a familiar melody irrespective of its key or instrumental assignment-the qualitative features of an object or event may change without compromising its *identity* (recall Von Ehrenfels's argument regarding melodic transposition, quoted above). In the case of a familiar melody, the contour, rhythmic profile, and intervallic relationships between consecutive pitches remain invariant, although the exact pitches being played may be varied (transposition), as may be the timbre of the melody (instrumentation). SG-|01234| is the same gesture regardless of the sonic and musical characteristics of its constituent events: it could be realized by a quintet of five clarinets or by a woodwind quintet. Despite timbral differences (between the above two quintet scenarios and within the second scenario), the spatio-temporal succession of events is preserved. That being said, spatial gestures in which like sound events are shifted between multiple points are generally more perceptually salient; a spatial gesture occurring within an ensemble of identical instruments (such as a guitar quartet) is arguably more readily perceived as a unified whole than would be a similar gesture unfolding within a quartet of unlike instruments, where a pronounced discrepancy

⁶⁹ Wertheimer, "The General Theoretical Situation," 13. Emphasis Original.

of timbre assists in the individuation of each instrument. When the Law of Proximity and the Law of Similarity cooperate, the perceptual salience of a gesture is enhanced. Nevertheless, the fundamental framework of my methodology entails abstracting the spatial location of ensemble points and their ensuing events away from all other sonic/acoustical features and musical elements otherwise attached to those events.

The SG-label reveals the ordered ensemble-point activity of a gesture. Each integer contained in the SG-label represents an event at a particular point in ensemble space and is termed an **event point**. Each event point is situated within an **event slot**. The number of event slots in a label will be equal to the total number of events within a gesture and will be numbered beginning with 1 for the first gestural event. For example, the label for a six-event gesture will indicate the points of consecutive events (event points) in event slots 1, 2, 3, 4, 5, and 6. In this case, the contents of event slots 1 and 6 would indicate the *temporal endpoints*—the initiating and concluding events—of the gesture. Consider SG-[021342]. This gesture comprises six events, so its label contains six event slots:

Event slots: 1 2 3 4 5 6 Event points: 0 2 1 3 4 2

Within this gesture, some event occurs twice at point 2—as both the second event (in event slot 2) and the sixth event (in event slot 6) of the gesture. An SG-label likewise contains overlapping **point-shift slots**. For instance, SG-|021342| consists of PS 0-2 (PS slot 1), PS 2-1 (PS slot 2), PS 1-3 (PS slot 3), PS 3-4 (PS slot 4), and PS 4-2 (PS slot 5). The total number of point-shift slots will always be e – 1, where e is the total number of events within a gesture. (SG-|021342| is a six-event gesture, but it contains only five point-shift slots).

It often proves necessary to describe spatial gestures unfolding within only a portion of the total ensemble space. As described in Chapter 2, such a section of ensemble space is called

an ensemble-space segment. When dealing with any ensemble-space segment, we shall use the same point integers for each ensemble point that we would when dealing with the total ensemble space. For instance, imagine that we are analyzing a work for an ensemble of eight instruments. Each performer (point) within the octet would be represented by an integer numbered from 0 to 7. However, let us presume that for a significant portion of the music, only the four rightmost instruments are actively engaged in the formation of spatial gestures—so we are only analyzing the spatial activity of ensemble-space segment 4–7. SG-|4567| would still be labeled as such; it would not be reduced to SG-[0123].⁷⁰ In addition, SG-[013], occurring in four-point ensemble space, would not be reduced to SG-[012]. This is a critical difference between SG-integer notation and contour segment notation. If CSEG <0123> were restated without the third pitch (represented by integer 2), it would be notated: CSEG <012>. Within an ensemble space, the component ensemble points remain present (and visible) even when they are inactive. Although a musical space (such as pitch space) may be said to contain fixed pitch-points, those pitchpoints may not always be conceptually present when not sounding. In short, ensemble-point integers *always* remain fixed (provided the ensemble members remain stationary). This stipulation simplifies SG-integer notation in general—helping to highlight a gesture's structural features as well as to illuminate the interrelationships among different gestures.

A spatio-analytic system must include a means by which we can indicate when multiple (two or more) sound events occur simultaneously within a spatial gesture, which I shall refer to as **point fusion**. For example, suppose that in our string quartet the first violin initiates a gesture and is soon followed by the second violin. Next, the viola and the cello enter the musical and

⁷⁰ Although SG-|0123| and SG-|4567| are different gestures (as indicated by the differing SG-labels), the strong relationship between the two gestures is quite evident: each gesture features strictly left-to-right motion spanning four adjacent ensemble points. The transformations described in Chapter 4 will provide a vocabulary for discussing such intergestural relationships.

spatial scene at the same time. The integer notation for two concurrent sound events (within an SG-label) shall be indicated by parenthetical enclosure and the separation by "+" of the event points. So, the entrance of the viola (2) and cello (3), occurring at the same point in time, would be indicated: 2+3. The SG-label for the above scenario would read: SG-01(2+3). The parenthetically enclosed integers constitute a single, fused event point (or, "superpoint") that fills a single event slot. A superpoint results from the fusion of two adjacent points. Nonadjacent points may be fused when producing coincident events, but they do not result in a superpoint. Since event points 2 and 3 occur at the same time, their coordinated entrance is perceptually "fused" via *simultaneous integration*, as described in Chapter 2.⁷¹ (Consecutive event points 0, 1, and 2+3 are linked by *sequential integration* to constitute the spatial gesture.) In SG-|01(2+3)|, superpoint 2+3 abides within event slot 3. The last point shift of this gesture is PS 1-(2+3). Within the SG-label, the ordering of the integers representing concomitant events at different points shall depend upon the directionality of the larger gesture in which it is embedded. For instance, SG-|01(2+3)| is a gesture spanning from left to right in ensemble space. If the gesture were to be presented in retrograde order, the label SG-|(3+2)10| would be preferable to SG-|(2+3)10| because the former label highlights the prevailing right-to-left motion of the gesture.

⁷¹ My appropriation of Bregman's concept of "simultaneous integration" is somewhat loose and requires qualification. Since the successive sonic events forming a spatial gesture (via sequential integration) need not be similar (i.e., an auditory stream) for spatial motion (a spatial gesture) to emerge, neither do two concomitant events need be acoustically similar to be fused (via simultaneous integration). The temporal coincidence of event onset (between two or more ensemble points) is the sole criterion for point fusion. Of course, two points that are spatially proximate and producing highly similar sounds at the same time are more effectively integrated—just as an acoustical likeness between consecutive events improves sequential integration and enhances the perceptual salience of a gesture. An ensemble *area*—comprising multiple points performing the same musical part on identical musical instruments (producing unisonous events)—is a prime example of simultaneous integration in Bregman's sense of the term.

3.3 Migratory and Accumulative Gestures

Beyond simply being assigned an SG-label based upon the temporal ordering of spatially discrete events, spatial gestures can be qualified in additional ways depending upon the exact nature of the sound events that give them form. For instance, let us revisit the gesture presented above in Figure 3-9. In this particular instantiation of SG-|0123|, each instrument in the sequence does not play until the previous instrument has ceased its sonic activity. This is an example of a **migratory gesture**, since there is no amassing of sound events among the various parts. It is as if a single sound event "migrates" through the ensemble.⁷²

Figure 3-10 presents an alternate spatial scenario. This gesture, SG-|3210|, occurs toward the end of the first movement of Alfred Schnittke's *String Quartet no. 4* (1989). In it, each player is required to sustain long note values until all four instruments have consecutively entered into the mix. Thus, the staggered and sustained entrances of each instrument result in an enlargement of the sonic space (via the "filling-out" of the available ensemble space). This is an example of an **accumulative gesture**, in which the gradual accretion of sound events leads to an expansion of the sonic space.⁷³ In this particular case, the spatial accumulation coincides with an increase in textural density as a [01235689] octachord is gradually constructed in pitch space.⁷⁴

⁷² Concert hall acoustics would affect the nature of this gesture. In a highly reverberant performance space, for instance, the sonic remnants (reverberation) of one instrument's activity might "bleed over" into the entrance of the next instrument. At the fast tempo in which this gesture occurs (Più mosso), the reverberant signals of point 0 could still be audibly present as point 3 joins the spatial activity. The inherently "migratory" nature of this gesture could be somewhat compromised by a highly reverberant sonic environment. However, under typical performance venue conditions, the level and presence of the direct signal produced by each instrument in the "gestural chain" would exceed that of the early and late reflections.

⁷³ The *sonic space* is identified as the portion of ensemble space in which one or more ensemble points are active (producing sound events) at any given time (see Chapter 2). When no points are active, there is no sonic space, but there *is* still an ensemble space since the points (ensemble members) are present. When all points are active, the sonic space is equal to the total ensemble space. The ensemble space is fixed (provided the performers remain stationary), but the sonic space is extremely elastic and variable. The propensity of the sonic space to grow or shrink in size is an important transformational affordance that shall be explored in subsequent chapters. As discussed in Chapter 2, the term *musical space* is best reserved for the phenomenal, virtual space (typically two- or three-dimensional) in which music is experientially received and conceptually structured. The *musical space*



Figure 3-10: Schnittke, *String Quartet no. 4*, Mvt. 1, mm. 79–82, accumulative SG-|3210|

An accumulative gesture frequently precedes an "ALL PLAY" section. We shall refer to sections in which all ensemble points are active (whether or not it follows an accumulative gesture) as **ALL PLAY**. The various instruments at different points need not be playing identical (or similar) musical parts. However, as dictated by context, we may qualify "ALL PLAY" with terms such as "homorhythmic," "unison," and so forth.

The opposite of an accumulative gesture is a **dissipative gesture**. A dissipative gesture may occur when an ALL PLAY scenario gradually disassembles into a smaller segment of the ensemble space or to a single point, which can occur as instruments successively cease their musical activity. A dissipative gesture results in a contraction of the sonic space, often to a finite

comprises the vertical dimension of pitch/texture, the horizontal dimension of passing time, and, potentially, the third dimension (depth) of dynamic level.

⁷⁴ Regarding textural density, if each instrument involved in an accumulative gesture enters with a unison joining of an ongoing musical line, the musical texture is not altered, although the composite sonority and color of the unison line would change with each successive instrumental entrance (particularly in an ensemble composed of different instruments). On the other hand, if each new instrument renders a different musical part, a textural crescendo accompanies the expansion of the sonic space.

point. A striking example of a dissipative gesture occurs in the concluding bars of Alban Berg's *Lyric Suite for String Quartet* (1926), as shown in Figure 3-11.



Figure 3-11: Berg, Lyric Suite, Mvt. 6, mm. 41-46, dissipative gesture

A lengthy ALL PLAY section precedes these final bars. Beginning at m. 41, the quartet members cease their musical activity in turn, withdrawing from the musical texture as well as from the spatial complex. The second violin (point 1) quits its musical line at the end of m. 41 and is followed by the cessation of the cello (point 3) in bar 43. Violin 1 (point 0) persists until m. 45 before dropping out. The lone viola (point 2) continues for two bars before dying away at the end of m. 46. Such a gesture might be notated: SG-|ALL PLAY -1, -3, -0, -2|. In this particular passage, each instrument diminuendos prior to becoming utterly hushed, which enhances the seamlessness of the sonic space's systematic dissolution. This dissipative gesture is

dramatically visible in the score due to the omission of empty measures in the staves of the instruments that have retired from the spatio-musical activity.⁷⁵

The sound events that mold a spatial gesture need not all be of the same qualitative configuration, as they are in the above examples (Figures 3-9, 3-10, and 3-11). In each of these examples, all four instruments perform highly similar events. The sole factor in labeling a gesture either migratory or accumulative is whether or not spatially discrete sound events endure and meld together in time. Imagine a four-voice fugue performed by a string quartet. Due to the nature of fugal process, some type of spatial gesture would emerge during the initial exposition, where voices enter successively with statements of the fugue subject. In such a case, the instrumental entrances are indeed discrete and isolated, but the voices continue with countersubjects as each new voice enters with the subject. Hence, there is an accumulation of sound events (musical activity) among the various ensemble points. Therefore, a fugue performed by multiple instruments affords the opportunity for accumulative gestures. Whether or not the analyst chooses to qualify a gesture as migratory or accumulative depends entirely upon the spatial context and whether or not the analysis of a specific piece calls for such distinctions to be made. Many gestures will appear as hybridizations of migratory and accumulative gestures: some sound events will "bleed over" one another while others will not. It is not uncommon in musical works to uncover a migratory gesture later transformed into an accumulative gesture (each with an identical SG-label), or vice versa.

A spatial gesture need not proceed from a single point surrounded by silence (in the form of inactive ensemble points), as the foregoing examples seem to imply. A gesture may weave through an ensemble in which all ensemble points are incessantly active (ALL PLAY). A

⁷⁵ In essence, the dissipation of staves on the score complements the evanescence of musical activity on the performance stage—an effective translation of score-space to ensemble-space activity.

gesture may occur in the presence of other, non-gestural musical activity (all within the same ensemble space). In such cases, the discrete sound events conspiring to form the gesture will typically be related by some musical feature other than their staggered spatio-temporal presentation, such as a sforzando marking, a leap of register, a melodic or rhythmic motive, etc. Without the sharing of some extra-spatial quality among the successive event points, a gesture unfolding within an ALL PLAY scenario would be undetectable.

3.4 Gestural Categories

In this section, I identify three primary categories (or, kinds) of spatial gestures: oscillation gestures, open gestures, and closed gestures. All possible gestures occurring in a horizontal, one-dimensional ensemble space may be assigned to one of these categories, and gestures within each of the three primary categories may be further subdivided and grouped into "strong" and "weak" types. Whether a gesture is migratory or accumulative does not affect category membership.

Oscillation Gestures

Perhaps the most common of the three gestural kinds, an **oscillation gesture** occurs between two entities that engage in an antiphonal, "call and response" type of spatio-musical dialogue.⁷⁶ This gesture can occur in duos as well as between two points, superpoints, areas, or hyperpoints within a larger ensemble. Oscillation gestures feature the perpetual alternation of events at two different ensemble points. For example, SG-|131313| involves the consistent interchange of event activity between ensemble points 1 and 3. Oscillation activity may be

⁷⁶ Oscillation may be defined as follows: a single cycle of variation between two values, positions, or states, or the act of moving backward and forward between two extreme points.

regarded as two overlapping point shifts of opposing directionality occurring in turn. SG-|131313| consists of the elided alternation of PS 1-3 and PS 3-1: the second event point of the former point shift is likewise the first event point of the latter point shift. While PS 1-3 is a local instance of left-to-right motion, PS 3-1 is a case of right-to-left motion. Two points need not be engaged in imitative musical activity to produce an oscillating gesture: the oscillation occurs between the spatial location of sonic events—not restrictively between similar or identical musical events that ensue from two different ensemble points in alternation.

An **adjacent-point oscillation gesture** occurs between two contiguous points in the ensemble space. Such gestures feature purely conjunct spatial motion. In a quartet setting, an adjacent-point oscillation gesture could occur between points 0 and 1, 1 and 2, or 2 and 3. Figure 3-12 presents an excerpt from the fourth movement of Schoenberg's *String Quartet No. 3* (1927). In these three measures, two different adjacent-point oscillation gestures unfold simultaneously.



Figure 3-12: Schoenberg, *String Quartet No. 3*, Mvt. 4, mm. 139–141, adjacent-point oscillation gesture(s)

Let us begin by examining the oscillation activity between the viola (point 2) and the cello (point 3). Both of these instruments produce pizzicato melodic intervals in an eighth-note rhythmic

configuration at a pianissimo dynamic level: the viola performs two descending perfect fourths followed by two ascending fourths, and the cello performs two ascending perfect fifths followed by two descending fifths. The qualitative similarity (pizzicato articulation, uniform rhythm, identical dynamic level, inversionally related intervallic content) of the musical events at points 2 and 3 enables them to be grouped together over time (via sequential integration). The fact that the events occur in alternation at two different ensemble points engenders an oscillation gesture. Since those two ensemble points neighbor one another in ensemble space, the resultant gesture is adjacent-point oscillation gesture SG-[23232323].

An additional adjacent-point oscillation gesture unfolds concurrently between ensemble points 0 and 1, where a sixteenth-sixteenth-eighth rhythmic figure is temporally displaced (or offset) to become the subject of both musical and spatial dialogue between the two violins. SG-|0101010101010101010| transpires in the left half of the ensemble space. Again, the similarity of the musical events at each spatially discrete point, coupled with the staggered presentation of those events, allows them to be grouped together in active production of an adjacent-point oscillation gesture—one which provides a kind of "spatial counterpoint" to the gesture occurring simultaneously in ensemble-space segment 2–3. Both of these gestures are of the same categorical kind (adjacent-point oscillation gesture), but each comprises different musical events and unfolds at its own pace. Although they both occur in the exact same span of musical time (beat 3 of m. 139 through beat 2 of m. 141), SG-|01010101010101010101010] occurs at exactly twice the rate of SG-|23232323|.⁷⁷

A **distant-point oscillation gesture** occurs between two nonadjacent points within the ensemble and therefore features disjunct spatial motion: the gesture will "sandwich" at least one

⁷⁷ Structural relations and interactive properties among concomitant gestures (in the context of "polystreaming") shall be addressed in Chapter 4.

ensemble point that is not involved in the gesture (this point may be inactive or otherwise engaged in separate spatio-musical activity). In a quartet setting, such gestures could transpire between points 0 and 2, 1 and 3, or 0 and 3. SG-|030303| is a more expansive distant-point oscillation gesture than either SG-|020202| or SG-|131313|: the former gesture encloses two ensemble points (1 and 2) while the latter two gestures each envelop only one (point 1 and point 2, respectively). A distant-point oscillation gesture occurs in Leonard V. Ball's *Wheels* for guitar quartet (2004), as shown in Figure 3-13.



Figure 3-13: Ball, Wheels, mm. 314-317, distant-point oscillation gesture

Ball composed this work for the Georgia Guitar Quartet. Within this particular ensemble, the member who typically occupies point 0 performs on a seven-string guitar, an instrument featuring an additional low string. In Ball's scoring of the work, the part with the extended bass

range is guitar 4, located at the bottom of the staff system.⁷⁸ Thus, musical parts 1, 2, 3, and 4 (from the top to the bottom staff) correspond, respectively, to ensemble points 3, 2, 1, and 0. In this particular situation, the *bottom-to-top* ordering of staves "translates" to the left-to-right sequence of ensemble points—a distinction that critically informs and affects spatial analysis. Therefore, in mm. 314–317 of Wheels, distant-point oscillation gesture SG-30303030 unfolds between the outermost ensemble members. The musical activity at each point is not identical but is relatively similar: point 3 (guitar 1) presents a four-note descending figure while point 0 (guitar 4) "answers" with a six-note ascending line. While both events unfold in strict sixteenth notes, the unequal durations of the distinct events highlights the asymmetrical, composite meter (5/8) in which they transpire. In bars 314–316, this gesture establishes a "spatial frame" by enclosing points 1 and 2, which are inactive (except for the sustained notes of bar 314). In the pickup to m. 317, guitar 2 (at point 2) enters and is joined by guitar 3 (at point 1) in m. 318. These two parts present interlaced thematic material through m. 337 (not shown). Throughout these 21 measures, the distant-point oscillation gesture continues—persisting in its "framing" occupation.

An oscillation gesture must contain a minimum of three event points, where the first and third event points are the same.⁷⁹ The greater the number of consecutive point shifts between two ensemble points, the *stronger* the gesture: the temporal *continuation* of antiphonal activity is an example of "good form." The perceptual salience of the gesture will be enhanced the longer it occurs, and rapid oscillation gestures are generally more salient than those unfolding at a more leisurely pace. Furthermore, an oscillation gesture that terminates at the same point from which

⁷⁸ Presumably, Ball elected to maintain the convention of situating lower-sounding instruments at the bottom of the staff system.

⁷⁹ Any gesture containing fewer that three event points would be reduced either to a point shift (two events) or to a single event.

it begins is "stronger" than one that does not because it demonstrates closure, or completion (another Gestalt facet of good form). Thus, SG-|020202020| is relatively stronger than SG-|0202| because it contains more exchanges between points 0 and 2 and because it begins and ends at point 0.⁸⁰ Both adjacent-point and distant-point oscillation gestures may be further classified as either "strong" or "weak." A strong oscillation gesture concludes at the same point from which it commenced and is effectively *closed*. A weak oscillation gesture does not terminate at the same point from which it began, and is effectively open. (I have elected to use the terms "strong" and "weak" instead of "open" and "closed" to avoid confusion with the other two gestural categories: open and closed gestures.) Oscillation gestures with an even event cardinality (four, six, etc.) are weak, while those containing an odd number of events (three, five, etc.) are strong. In the example from Schoenberg's String Quartet No. 3 (Figure 3-12 above), SG-0101010101010101010 is a strong adjacent-point oscillation gesture since the first violin both begins and concludes the gesture. The contemporaneous SG-[23232323], on the other hand, is a weak adjacent-point oscillation gesture since it begins with the viola and terminates with the cello. A gesture such as SG-[030303] would be an example of a weak distant-point oscillation gesture, while SG-20202 would exemplify a strong distant-point oscillation gesture.

Ensemble points occasionally work in conjunction to generate a "fused" oscillation gesture. An example of a **fused oscillation gesture** would be SG-|(0+1)(2+3)(0+1)...|, where adjacent points 0 and 1, as well as 2 and 3, are "fused" together in their spatio-temporal activity, resulting in two separate "superpoints." Since the two fused superpoints are themselves adjacent to one another, this would constitute a *fused adjacent-point oscillation gesture* (which may be

⁸⁰ However, with a temporally expansive gesture, the listener might find it difficult to gauge whether or not the gesture has ended at the same point from which it began. The longer the gesture, the more difficult for the listener to retain the exact location of the gesture's starting point. Therefore, an extensive oscillation gesture (with numerous point shifts), which is indeed perceptually strong, could potentially cancel-out the strength (in the form of closure) that it might exhibit by both beginning and ending at the same point.

either strong or weak depending upon its superpoint of conclusion). Another example of a fused oscillation gesture would be SG-|(1+3)(0+2)(1+3)...|, where nonadjacent points (1 and 3, 0 and 2) are fused. Due to the "interlaced" fusion of ensemble points, such a gesture is neither adjacent-point nor distant-point. Either of these two examples could also be notated as two separate yet concomitant gestures, the latter as: SG-|101...| and SG-|323...| or, perhaps less intuitively: SG-|121...| and SG-|303...|. However, viewing and labeling the gesture as a fused oscillation gesture highlights the simultaneity of the component events (1 and 3, 0 and 2). Figure 3-14 reveals an excerpt from *Diptych No. 1 for Four Guitars*, Op. 80 (1979) by John W. Duarte.



Figure 3-14: Duarte, *Diptych No. 1*, Mvt. 2, mm. 128–131, fused oscillation gesture

In this passage, guitars 1 and 4 (points 0 and 3) are fused, as are guitars 2 and 3 (points 1 and 2). Since guitars 2 and 3 are proximate in ensemble space, and since they are identical instruments playing identical chords, their temporary fusion creates a fleeting *area* within the ensemble space.⁸¹ The gestural yield of this passage is SG-((1+2)(0+3)(1+2)(0+3)), which is a weak (i.e.,

⁸¹ Guitars 1 and 4, on the other hand, do not play identical material and are non-adjacent in ensemble space. Therefore, their sharing of a uniform rhythm and similar melodic profile results in point fusion but not in superpoint or area formation. Fused points 0 and 3 effectively "frame" superpoint/area 1+2.

"unclosed") oscillation gesture. Due to the nature of the fusion of ensemble points, this gesture is neither adjacent-point nor distant-point. Instead, it involves the sequential "opening" (or, expansion) and "closing" (or, contraction) of the sonic space (within the total ensemble space). The interpretive implications of such spatial activity shall be addressed in Chapter 5.

Open Gestures

An **open gesture** will begin at one end of ensemble space and terminate at the other end—this is the fundamental distinguishing feature of any open gesture. The gesture may traverse all points within the ensemble or may skip-over one or more points. Ensemble points may be active more than once within the gesture. The requisite qualifier is that the spatial extremes of the ensemble space (or ensemble-space segment) serve as the temporal endpoints of the gesture. For example, in a five-point ensemble space (quintet), points 0 and 4 serve as the spatial boundaries of the total ensemble space. Gestures such as SG-[01234], SG-[01324], SG-[41320], SG-[0124], and SG-[0123234] are all broadly categorized as open gestures by virtue of the simple fact that they begin and end at the opposite poles of the ensemble space.⁸² An open gesture is "open" because its spatio-temporal endpoints are maximally disconnected.

A **partial open gesture** is a gesture that is classified as open within a segment of ensemble space but not within the total ensemble space. The five-point ensemble space (quintet) referenced above may be partitioned into smaller segments: it comprises two contiguous fourpoint segments, or quartets (0 to 3 and 1 to 4) as well as three continuous three-point segments,

⁸² While our hypothetical quintet ensemble space has a point cardinality of five, SG-|0124| contains an event cardinality of four, since point 3 is inactive and effectively "skipped-over" during the gesture. Likewise, SG-|0123234| contains an event cardinality of seven, since points 2 and 3 are both active twice during the unfolding of the gesture.

or trios (0 to 2, 1 to 3, and 2 to 4).⁸³ It is possible to encounter a significant portion of a musical work for quintet in which only three of the five ensemble points are active and/or coordinated in the production of spatial gestures, while the other two points remain inactive or are engaged in musical activity that segregates them from the three-point ensemble-space segment. Partial open gestures might emerge in this local section of ensemble space. For example, within ensemble-space segment 2-4 (still within a five-point total ensemble space), the following gestures would be classified as partial open gestures: SG-[234] and SG-[432]. At the global level of the total ensemble space, however, such gestures would not be considered *open* at all, since one boundary (ensemble point 0) of the total ensemble space does not serve as one of the spatio-temporal endpoints of the gesture.

Open gestures are further delineated into *strong* and *weak* types. A **strong open gesture** moves continuously in one direction through the entire ensemble space (or ensemble-space segment). The motion may be *conjunct*—featuring the sequential sounding of adjacent instruments, or *disjunct*—skipping an ensemble point in one or more component point shifts. Some examples of strong open gestures in a five-point ensemble space include: SG-[01234], SG-[43210], and SG-[0124]. The first two examples are purely conjunct, while the latter is an example of a strong disjunct open gesture, since point 3 is skipped-over during the final point shift (2-4) of the gesture.⁸⁴ A strong open gesture (both conjunct and disjunct) may also be referred to as a **panning gesture** and, depending upon its particular directionality, may be labeled as either a left-to-right panning gesture (L-R PAN) or a right-to-left panning gesture (R-L

⁸³ The four possible two-point segments could only yield point shifts and oscillation gestures and are therefore presently dismissed.

⁸⁴ A strong *disjunct* open gesture might be perceived as somewhat *less* strong than a strong *conjunct* open gesture owing to the fact that the disjunct gesture is less complete, and thereby, perhaps slightly less stable. The listener might expect an event to occur at the skipped-over point (which would yield a closed gesture) and perceptually try to "complete" the gesture by "filling-in" the missing event at whatever point (or points) might have been skipped. For instance, SG-|0124| might cause the listener to anticipate a "concluding" event at point 3, the actual occurrence of which would produce SG-|01243|. This gesture is classified as a *weak closed gesture* (to be explicated below).

PAN).⁸⁵ Panning gestures are highly salient due to their pronounced Prägnanz: the good form of a panning gesture results from the unidirectionality (continuity) and implied continuation of the gesture, as well as from the fact that each discrete shift between contiguous points "moves" in the same direction, exhibiting the gestalt principle of "common-fate."

SG-|013| is an *abbreviated* form of SG-|0123|. While both are classified as left-to-right panning gestures (L-R PAN), SG-[013] is achieved by "deleting" an event at ensemble point 2 (in event slot 3 of SG-|0123|). If dealing with a four-point ensemble space, SG-|0123| is the optimal form for a L-R PAN. Therefore, in such a space, a panning gesture that includes every ensemble point once and only once shall be termed a **perfect panning gesture** to distinguish it from other strong open gestures that involve some degree of disjunction. So, in four-point ensemble space, SG-[0123] and SG-[3210] are the only possible perfect panning gestures.⁸⁶ The spatial gesture presented above in Figure 3-9, SG-|0123|, is a strong open gesture (perfect L-R PAN).⁸⁷ A perfect panning gesture is likely the most perceptually salient of all spatial gestures. In addition to displaying good continuation—a trait shared by all panning gestures—the *perfect* panning gesture features uniformity among all of its component point shifts: assuming that the contiguous ensemble points are equidistantly spaced, every point shift in SG-[0123] (PS 0-1, PS 1-2, and PS 2-3) is conjunct and spans the same approximate distance. Furthermore, all of the point shifts (parts) move in the same direction as the gesture as a whole: reducing the gesture to its spatiotemporal endpoints yields PS 0-3-the "background" point shift that reflects the long-range, left-

⁸⁵ Indicating the specific directionality of a panning gesture will prove essential to assessing the interpretative potential of that gesture as well as evaluating the interactivity of multiple gestures (to be explored in Chapter 5).

 $^{^{86}}$ In an ensemble space of *any* cardinality there will only be two possible forms of *perfect panning gesture*, both of which share in a retrograde and inversional relationship: SG-|3210| is both the retrograde and the inversion (but not the retrograde-inversion) of SG-|0123|, and vice versa. Such transformational relationships are described in the subsequent chapter.

⁸⁷ A series of *displaced panning gestures* can arise from a transformative procedure called *rotational slot displacement*. This and other types of alterations shall be addressed in Chapter 4. A displaced panning gesture is actually categorized as a *weak closed gesture*.

to-right motion of the gesture.⁸⁸ Thinking of the point shifts in a spatial gesture as the motives in a gestalt (in Schoenberg's sense of these terms), the uniformity (of directionality, conjunct motion, and distance) among parts in a perfect panning gesture, as well as the strong directional correspondence between part(s) and whole, strengthens the perceptual coherence of the gesture.⁸⁹

A weak open gesture does not feature strictly unidirectional motion. Like all open gestures, weak open gestures will begin at one extreme of ensemble space and terminate at the other end. A weak open gesture will involve either disjunct motion or conjunct motion in which interior ensemble points are active more than once. Figure 3-15 presents weak open gesture SG-[0213] from Elliot Carter's *String Quartet No. 4* (1986).



Figure 3-15: Carter, String Quartet No. 4, mm. 114-116, weak open gesture

⁸⁸ When the interior event points (1 and 2) of SG-|0123| are reduced away, PS 0-3 (comprising outer event points 0 and 3) emerges as the "background" point shift of the gesture. This point shift depicts the structural frame (the spatio-temporal endpoints) of SG-|0123| and exists only at a background level of the spatial structure (i.e., PS 0-3 does not occur on the spatial surface of the work). We might view PS 0-2 and PS 1-3 as "middleground" point shifts and regard PS 0-1, PS 1-2, and PS 2-3 as "foreground" point shifts.

⁸⁹ The perfect panning gesture presented in Figure 3-9 is likewise "well-formed" owning to the equivalent temporal durations of each event forming the gesture.

SG-|0213| is an *open* gesture because it begins and terminates at the opposing extremes of four-point ensemble space. It is *weak* because it features two instances of left-to-right motion (equidistant point shifts 0-2 and 1-3) and one instance of right-to-left motion (PS 2-1). In fact, the prevailing left-to-right motion of the entire gesture is interrupted by the solitary instance of right-to-left motion.⁹⁰ While each point is active only once within the gesture, there are two cases of disjunct motion (PS 0-2 and PS 1-3). In the figure, arrows are drawn from the entrance of one ensemble point to that of the next in the sequence, such that the path of the gesture in ensemble space is traced out in "score space." As a result, the two instances of left-to-right motion are indicated by a Southeast-directed arrow, and the single right-to-left point shift is depicted by a Northeast-directed arrow. Note that the musical activity of each of the four instruments is rather contrastive (save for the persistent presence of double stops).

Now consider weak open gesture SG-|012123234|, in which the overall direction of motion is left-to-right. This gesture features exclusively conjunct motion with six instances of discrete left-to-right point shifts (PS 0-1, PS 1-2, PS 1-2, PS 2-3, PS 2-3, PS 3-4) and two cases of right-to-left point shifts (PS 2-1, PS 3-2). Several ensemble points operate more than once—point 2 is active three times while points 1 and 3 are each active twice. It would be feasible to conceive of this large gesture as three temporally consecutive and spatially overlapping left-to-right panning gestures: SG-|012| + SG-|123| + SG-|234|. This alternate interpretation reduces the larger gesture to its salient (and related) subgestures: identical and successive activity in three overlapping, three-point ensemble-space segments produces the larger-scale gesture that spans the

 $^{^{90}}$ The left-to-right motion prevails only because the gesture begins at point 0 and ends at point 3, *not* because the number of discrete left-to-right events (2) exceeds that of right-to-left events (1). Consider SG-[03214] transpiring in five-point ensemble space. In this gesture, the number of left-to-right point shifts (0-3, 1-4) balances the number of right-to-left (3-2, 2-1) point shifts. Although this gesture features an interior R-L PAN (SG-[321] in ensemble-space segment 1–3), the prevailing (or, large-scale) motion is still left-to-right since the gesture starts at point 0 and ends at point 4.

entirety of the ensemble space.⁹¹ We can easily see that three small partial open gestures (as parts) conspire to form the larger weak open gesture (as whole). The regular recursion of left-to-right motion on different structural levels (parts and whole) strengthens the perceptual salience of the gesture.

Closed gestures

A **closed gesture** may begin and end at any point within the ensemble, as long as it does not begin at one margin of ensemble space (or ensemble-segment space) *and* conclude at the other fringe of the space; however, a closed gesture may *either* begin *or* end at one of the borders of the ensemble space. In addition, closed gestures may begin and end at the same ensemble point, and this point could be one of the spatial extremes of the ensemble space. A closed gesture will always feature motion that is not unidirectional, and unless one or more ensemble points are active more than once, a closed gesture will feature disjunct motion.

A strong closed gesture will commence from and terminate at the same point in ensemble space. Examples include SG-|013420|, SG-|214302| and SG-|301243|. Such gestures demonstrate Prägnanz in the form of closure: just as the figure of a perfect circle can be said to start from any given point and return to that same point, so too does a strong closed gesture depart from and ultimately return to the same point in ensemble space. However, in the horizontal plane currently under investigation, the spatial shape does not generally occur with the symmetry of a circle.⁹² Nonetheless, a sense of spatial closure is evinced. The temporal

⁹¹ Owing to its easily discernible subgestures, SG-|012123234| is rightfully deemed a *compound spatial gesture* (to be expounded upon in section 3.5). The transformational operation of *nudging* (to be described in Chapter 4) accounts for the relatedness of the discrete subgestures (SG-|012|, SG-|123|, and SG-|234|). The problem of segmentation and the potential for alternate readings will be taken up in Chapter 6.

⁹² The circular (and seemingly continuous) orbit of sound is possible in—and quite common to—the surround environment (acoustic or electronic). Joel Chadabe notes that in *Icon* (1967) by Joji Yuasa, "sounds move at the same time clockwise and counterclockwise at different speeds through a pentagonal array of loudspeakers

endpoints of a strong closed gesture-although separated in time-"connect" at the same point in ensemble space. Figure 3-16 presents another excerpt from Wheels by Leonard V. Ball. Considering only the sextuplet (or compound triplet) figures, and discounting the qualitatively dissimilar descending interval (B-flat to E) in m. 84 (part 4), this passage reveals SG-[3102013] (bear in mind the exceptional "score space-to-ensemble space translation" of this particular work, as described above in reference to Figure 3-13).



Figure 3-16: Ball, Wheels, mm. 81-84, strong closed gesture

This gesture begins and ends at ensemble point 3 (guitar part 1) and thus constitutes a strong closed gesture. Note the palindromic quality of this gesture: beginning from the central event (event point 2 in event slot 4) and branching out in either direction, we encounter event points 0 (in event slots 3 and 5), 1 (in event slots 2 and 6), and 3 (in event slots 1 and 7). The gesture likewise displays perfect temporal symmetry.⁹³ As shall be discussed in Chapter 4, the transformational affordances of such spatially symmetrical gestures are somewhat limited.

surrounding the audience" (Chadabe, 67). Circular spatial gestures, as they occur in both acoustic and electronic spaces, shall be explored further in Chapters 4 and 5.

This observation highlights a concordance of symmetries in both time and space.

A weak closed gesture simply does not end at the same point from which it began.

Some examples of weak closed gestures possible within a quintet setting include: SG-|01243|, SG-|42103|, SG-|21304|, and SG-|04231|. Of these four examples, two are perhaps less weak (or more closed) than the others. Any weak closed gesture in which the temporal endpoints (the event points contained in the outer event slots of the SG-label) are adjacent may be perceived as being more closed (or complete) than those gestures in which the temporal endpoints do not fall on adjacent ensemble points. Such gestures are on the brink of being strong closed gestures their temporal endpoints are *almost* connected at the same point in space. SG-|42103| and SG-|04231| both exhibit temporal endpoints that occur at adjacent points in ensemble space, SG-|01243| and SG-|21304| do not. The opening movement of Béla Bartók's *String Quartet No. 1,* Op. 7 begins with weak closed gesture SG-|0132| (see Figure 3-17).



Figure 3-17, Bartók, String Quartet No. 1, Mvt. 1, mm. 1-9, weak closed gesture
In this excerpt, the opening descending sixth of the first violin is immediately echoed by the second violin at a different pitch level. The entrance of points 2 and 3 is postponed for six measures; when the cello and viola enter in m. 8, the accumulative weak closed gesture and the expansion of the sonic space are completed. The spatial activity of this passage is akin to the tonal convention of filling-in a disjunct melodic leap with stepwise motion in the opposite direction: PS 1-3 (a spatial disjunction proceeding from left to right) is promptly succeeded by PS 3-2 (conjunct, right-to-left spatial motion).

Some specific instantiations of a weak closed gesture (a total of four for any ensemble space with a point cardinality of three or more) exhibit the shape of a *spiral*—inward or outward, clockwise or counterclockwise. Such gestures are referred to as **spiral gestures**. The perceptual activity of "connecting the dots" is required for the perception of any kind of motion during the experience of a spatial gesture. A row of Christmas lights that blink one after the other in rapid succession is easily perceived as a unidirectional line of motion. Likewise, the consecutive activity of adjacent performers within an ensemble space will yield a strong open (perfect panning) gesture when all the "dots" (i.e., event points) are "connected."⁹⁴ Other spatial shapes may emerge from perceptually "connecting" consecutive events at different points in ensemble space. Using a curved line, we may trace the ordered events of a spatial gesture in an ensemble-space diagram. Consider the diagram of SG-[04132] given in Figure 3-18.



Figure 3-18: Spiral diagram of SG-|04132|

⁹⁴ Perceptually connecting the dots (visually and/or aurally) exemplifies the Gestalt principle of *reification*.

The resultant shape of the gesture in Figure 3-18 is a *clockwise inward spiral*. In an effort to distinguish between the four different types of spiral gestures that are possible within a given ensemble space (to be detailed below), and for the sake of consistency, the initial, instigating point shift of any spiral gesture shall always be depicted graphically by a line extending *over* the ensemble-point integers in the diagram of the ensemble space. The second point shift shall be traced below the integers, the third above, and so forth. This is, admittedly, an arbitrary notational convention. Furthermore, the perceptibility of a dynamic spiral in this particular context (a horizontally arrayed ensemble space) is questionable. Instead of discerning a motion that spirals inward (as with SG-|04132|), the listener might be more prone to infer a collapsing, zigzag type of gesture. The assigning of the label "spiral" to such gestures is perhaps more reflective of the analytical process of graphic modeling than of the perception and aesthetic affect of the gesture. This is not say, however, that listeners absolutely do not conceive of such gestures as spiral in nature.⁹⁵

Spiral gestures form a unique subset of the weak closed gestural category. A spiral gesture will either be inward or outward, and clockwise or counterclockwise. Thus, the four possible spiral gestures in a five-point ensemble space are SG-|04132| (clockwise inward spiral), SG-|23140| (clockwise outward spiral), SG-|40312| (counterclockwise inward spiral), and SG-|21304| (counterclockwise outward spiral).⁹⁶ All four of these spiral shapes appear in Schoenberg's *Quintet for Wind Instruments*, Op. 26 (1924). Figure 3-19 depicts SG-|04132| (clockwise inward spiral) as it occurs in the third movement of the *Quintet*.

⁹⁵ Under certain dramatic (and extremely rare) performance conditions, such as Stockhausen's experimentation with total surround in his electronic work entitled *Spiral* (1969), actual spiral gestures may prove easily discernable. In *Terretektorh* for large orchestra (1965–66), Xenakis creates the impression of Archimedean, logarithmic, and hyperbolical spirals by accelerating the velocity of circular gestures in a three-dimensional ensemble space. The coordination of spatial and temporal processes fashion the spiral shapes. (See Harley, "Space and Spatialization," 285–290.)

⁹⁶ Any two spiral gestures are related by one of the following three transformational operations: retrograde, inversion, or retrograde-inversion. These operations shall be described in Chapter 4.



Figure 3-19: Schoenberg, *Quintet for Wind Instruments*, Mvt. 3, m. 30, clockwise inward spiral gesture

Woodwind quintet ensembles abide by an assortment of seating plans. The above musical example forms a clockwise inward spiral if we assume the following relatively standard seating plan (from left to right): flute (point 0), oboe (point 1), French horn (point 2), bassoon (point 3), and clarinet (point 4). Figure 3-20 presents a diagram of this ensemble layout.

	Horn
Oboe	Bassoon
Flute	Clarinet

Figure 3-20: A common woodwind quintet seating plan

Beginning with the flute's F-natural on the downbeat (see Figure 3-19), m. 30 consists of the following sequence of instrumental entrances: flute, clarinet, oboe, bassoon, and horn.⁹⁷ This sequence creates SG-|04132| if realized with the seating plan diagrammed in Figure 3-20. Note that the second and third events (event points 4 and 1) are similar to one another, as are the

⁹⁷ The flute's F-natural is slurred from a B-flat (a compound perfect fourth above) in the previous measure—it is not tied to a previously sounded F-natural. Therefore, the flute does actually articulate the pitch on the downbeat of m. 30 in lieu of sustaining the note across the barline. This articulation instigates the spatial gesture under investigation.

fourth and fifth events (event points 3 and 2). However, the gesture as a whole does not consist of five homogenous events. The composite spatial pattern and emergent motion (or, "whole") formed by the succession of spatially discrete events (or, "parts") transcends the dissimilitude of those events.

A clockwise *outward* spiral gesture transpires in the fourth movement of Schoenberg's *Quintet*. Assuming the same seating plan, SG-|23140| is shown in Figure 3-21.



Figure 3-21: Schoenberg, *Quintet*, Mvt. 4, mm. 309–313, clockwise outward spiral gesture

This gesture is formed by the following sequence of instrumental entrances: horn, bassoon, oboe, clarinet, and flute.⁹⁸ Two of these instrumental parts (horn and clarinet) feature appearances of

⁹⁸ The figure traced out on the score does not resemble the shape of a horizontally (rightward) expanding spiral due to the "poor" translation of score space to ensemble space: part 1 (flute) = point 0; part 2 (oboe) = point 1; part 3 (clarinet) = point 4; part 4 (horn) = point 2; part 5 (bassoon) = point 3. In a "good" translation of score to ensemble

the local *Hauptstimme* ("principal part"). However, all five of the lines begin with a similar rhythmic and melodic figure (the rapid, three-time iteration of a single pitch). While the gesture is not dependent upon it for its identity and validity, the likeness of event-onset activity at each point enhances the salience of the spiral gesture.

Although a comprehensive spatial analysis of Schoenberg's *Quintet for Wind Instruments* lies beyond the immediate scope of this chapter, I will note that the four specific types of spiral gestures recur throughout the quintet—serving to unify the spatial construct of the work.⁹⁹ For example, SG-|23140| (clockwise outward spiral) appears in mm. 107–112 of the first movement as well as in mm. 226–229 of the fourth movement. SG-|40312| (counterclockwise inward spiral) unfolds in mm. 120–122 of the first movement and again in mm. 237–240 of the fourth movement. The four different spiral gestures are interrelated shapes, and their similarities allow them to function collectively as "spatial motives" woven throughout the quintet as a whole—motives in space that are independent from the melodic and rhythmic motives that they contain.

Spiral gestures always contain the maximum number of directionality changes for gestures in which each point is active once and only once. For instance, SG-|21304| (counterclockwise outward spiral) contains two instances of left-to-right motion (PS 1-3 and PS 0-4) and two instances of right-to-left motion (PS 2-1 and PS 3-0). L-R point shifts and R-L point shifts occur in alternating succession (PS 2-1, PS 1-3, PS 3-0, PS 0-4), which yields a total of three changes in directionality. Spiral gestures, as defined and exemplified above, will always

space, the point position of an instrument will be n - 1, where n is the part number (as numbered beginning with 1 for the top staff and increasing by 1 as the staves descend through the system/page of the score). This translation type occurs whenever the parts are ordered consecutively onstage from left to right. With such a translation, score part 1 would correspond to ensemble point 0, part 4 would correspond to point 3, etc. The other type of "good" translation occurs when the ensemble-point value and the part number sum to the same number for each instrument in the ensemble. This occurs whenever the parts are ordered onstage from right to left (instead of from left to right). Such is the case in Leonard V. Ball's guitar quartet entitled *Wheels*, as described above: score part 1 is at ensemble point 3 (1 + 3 = 4), part 2 is at point 2 (2 + 2 = 4); part 3 is at point 1 (3 + 1 = 4); part 4 is at point 0 (4 + 0 = 4).

contain (n - 2) directionality changes, where n is the total number of ensemble points.¹⁰⁰ Spiral gestures are not, however, the only instances of weak closed gestures that exhibit the maximum possible number of directionality changes: any gesture in which successive (and overlapping) point shifts consistently alternate in direction, such as SG-|02143| (which is not a spiral gesture), will contain the maximal number of directionality changes. A change in directionality weakens the salience of a gesture because it defies the Gestalt Law of Continuity. In general, the greater the number of directionality changes, the weaker the gesture. However, the continual alternation of directionally opposed point shifts creates an iterated pattern of change that may serve to heighten gestural salience. Such is the case with oscillation gestures and certain forms of open and closed gestures. Furthermore, the gradual contraction or expansion of a spiral gesture (inward or outward) reflects continuity in the form of an uninterrupted spatial process.

Although the four potential forms of spiral gestures (for any ensemble space of threepoint cardinality or higher) are codified as subspecies of the weak closed gesture category, some are perhaps weaker (or, "less-closed") than others. Take, for example, the outward spiral gestures (either clockwise or counterclockwise). These gestures are actually akin to open gestures through the implied continuation of outward spiraling, whereas the inward spiral gestures seem to close in on themselves—perhaps collapsing to a finite point.

Of course, it is important to bear in mind that the idea of any of these gestures exhibiting the form of a spiral is primarily an analytic tool employed for modeling purposes and to differentiate specific (and related) forms of weak closed gesture. The gestures may not actually be perceived as spirals in performance, particularly since the ensemble space currently under

¹⁰⁰ Gestures in which ensemble points are active more than once can have more than (n - 2) changes in direction.

investigation is one-dimensional—essentially comprising a single, horizontal line from left to right through the ensemble space.¹⁰¹

3.5 Compound Spatial Gestures

Conditions may likewise call for the identification of a **compound gesture**, in which two or more small gestures occur in temporal succession to form a larger gesture.¹⁰² We might think of a compound gesture as a "molecule" comprising multiple atoms, where each "atom" is a distinct spatial gesture. The component, smaller gestures that constitute a compound gesture are subgestures. Discrete subgestures may be related to one another (i.e., belong to the same gestural category), or they may be contrasting in kind and/or directionality. Smaller gestures from different categories or subcategories may combine, over time, to form a compound gesture of still a different categorical kind.¹⁰³ Furthermore, a single type of compound gesture may be categorized differently depending upon the exact nature of its design (i.e., the relationship between its starting and ending points, and the relationship of these points to the whole of the ensemble space). Two specific types of compound gestures shall be explored in this section: cyclic panning gestures and sweeping gestures. In both of these gestures, the discrete subgestures are closely related. For cyclic panning gestures, the relationship among subgestures is identicality. For sweeping gestures, the relationship is a serial transformation (which shall be discussed in greater detail in Chapter 4). Compound gestures need not comprise related

 ¹⁰¹ Although we have dismissed it up to this point, the slight semi-circular arc of an ensemble might aid in the perception of spiral gestures actually "spiraling" within the ensemble space.
¹⁰² A compound gesture consists of temporally adjacent gestures: the gestures do *not* occur at the same time.

¹⁰² A compound gesture consists of temporally adjacent gestures: the gestures do *not* occur at the same time. However, the temporal endpoints of the smaller gestures may elide (the final event point of one gesture may also serve as the first event point of another gesture). The coincident occurrence of two or more spatial gestures results in *polystreaming*, to be explored in Chapter 4.

¹⁰³ The strong closed gesture presented above in Figure 3-16 may be analyzed as a compound gesture: two weak closed gestures, SG-|3102| and SG-|2013|, are essentially coupled at point 2 (event slot 4 of SG-|3102| and event slot 1 of SG-|2013|) to produce the compound strong closed gesture (SG-|3102013|).

subgestures; any extensive gesture might be segmented into unlike subgestures. The analytic ramifications of such segmentation shall be taken up in the context of spatial analysis in Chapter 6.

Cyclic Panning Gestures

The compound gesture in Figure 3-22 begins at the leftmost ensemble point and concludes at the rightmost ensemble point, so it is classified as open. However, the gesture as a whole does not feature purely unidirectional motion, so it cannot be a strong open gesture: while the prevailing motion is left-to-right, there are instances of right-to-left motion, as PS 3-0 occurs

170

eight times (in point-shift slots 4, 8, 12, 16, 20, 24, 28, and 32). Therefore, the gesture as a whole is classified as a weak open gesture.



Figure 3-22: Koshkin, Changing the Guard, mm. 121-131, cyclic panning gesture

The cyclic regularity of this gesture might allow it to be perceived as circular in motion. Thus far, I have advocated normalizing (i.e., "ignoring" for analytic purposes) the semi-circular arc of a predominantly horizontally-oriented ensemble, such as the string quartet and the guitar quartet. However, such an arc bears the potential for shaping the perception of certain kinds of spatial gestures—particularly compound gestures such as the cyclic panning gesture currently under inspection. It is possible that PS 3-0 will *not* be perceived as a disjunct span from point 3 to point 0 that skips-over points 2 and 1 along the way: it could be seen as a motion that "wrapsaround" and closes the "circle" potentially suggested by the arc of the ensemble. When received in this light, PS 3-0 is conjunct, and the gesture assumes a two-dimensional character appearing circular in nature. However, a shift between points 0 and 3 (PS 0-3 or PS 3-0) will always span the greatest spatial distance of any possible point shift since points 0 and 3, as the spatial extremes of the ensemble, will generally be the two most separated ensemble points (provided the ensemble is not perfectly "squared-off"). The precise onstage configuration of the ensemble would affect the reception of a cyclic panning gesture: if the quartet is oriented with a minimal arc, as in Figure 3-23(a), the gesture might be received as a cycle of homodirectional panning gestures. On the other hand, if the ensemble is more "rounded off," as in Figure 3-23(b), the gesture might be perceived as a continuous, clockwise circle.



Figure 3-23: a) Quartet seating plan with minimal arc, b) Seating plan with pronounced arc.

The interpretive implications and aesthetic affect of these potentially conflicting readings could be quite pronounced: if PS 3-0 is not viewed as a disjunction, the whole gesture could easily be read as unidirectional—displaying Prägnanz in the form of circular continuation. Since precise ensemble configuration and organization is a performance-practice issue and is generally not explicitly designated by the composer, such compound gestures shall simply be identified in analysis as cyclic panning gestures, with the ancillary understanding that "cyclic" could imply (and be read as) "circular" in certain contexts.

Cyclic panning gestures may be composed of smaller gestures that are *not* open in nature; they need not be built exclusively from strong open (panning) gestures. SG-[2301] is categorized as a weak closed gesture. However, compound gesture SG-[230123012301] counts as a cyclic panning gesture (still classified as weak closed). Although each of its three discrete four-event gestures are SG-[2301] (SG-[2301,2301,2301]), SG-[0123] is embedded twice in this larger gesture—occurring both in event slots 3–6 and in event slots 7–10: SG-[23,0123,0123,01]. The presence of these two L-R PANs, which are temporally adjacent to one another within the gesture, renders the compound gesture as a whole a cyclic panning gesture. Furthermore, any cyclic panning gesture may end at the same point from which it began and thereby be classified as a compound strong closed gesture. Extending SG-[230123012301] to the next event in the cycle produces SG-[2301230123012]. The gesture now begins and ends at ensemble point 2 and is therefore strongly closed. While this particular compound closed gesture cannot be partitioned neatly into identical, discrete subgestures, it is still classified as a cyclic panning gesture.

Sweeping Gestures (SWEEP)

Consider SG-|0123210|. This gesture, which is termed a **sweeping gesture** (or, **SWEEP**), is another specific type of compound gesture and a special kind of closed gesture. Each point shift is purely conjunct, and the bi-directionality of motion is perfectly balancedthere is no dominance of either left-to-right or right-to-left motion. In fact, SG-|0123210| is an amalgam of two related strong open gestures: perfect L-R PAN (SG-|0123|) and its retrograde (and inversion), perfect R-L PAN (SG-|3210|). In this combination, the two open gestures "hinge" on the sound event occurring at ensemble point 3: the final event point of SG-|0123| likewise functions as the initial event point of SG-|3210|. Depending upon context, however, a SWEEP could easily be perceived as a unified whole—not as two consecutive open gestures. Due to the departure-and-return character of this gesture, and the balance of directional motion, a sweeping gesture is best interpreted as a strong closed gesture.¹⁰⁴

SG-|012343210| shall be labeled LEFT SWEEP, since it both begins and ends at the leftmost point in ensemble space. Likewise, SG-|432101234| shall be identified as RIGHT SWEEP. Since such gestures consist of perfectly conjunct motion and extend from one boundary of the total ensemble space to the other, with only one change in directionality, each shall be considered a **perfect sweeping gesture (perfect SWEEP)**. Abbreviations of sweeping gestures (such as SG-|0134210|), or their occurrence in ensemble segment space (such as SG-|43234| in a five-point ensemble space), shall retain the SWEEP label but will not be considered "perfect." Note, however, that the drastic abbreviation of a sweeping gesture such as SG-|012343210| (LEFT SWEEP) would result in a strong oscillation gesture—either adjacent-point (such as SG-|343|) or distant-point (such as SG-|040|). (Abbreviation, and other indices of similarity, shall be detailed in Chapter 4.)

Figure 3-24 presents a sweeping gesture from Schoenberg's *String Trio* Op. 45 (1946), which is scored for violin, viola, and cello. Assuming the following point positions for the ensemble members: violin (point 0), viola (point 1), and cello (point 2), this passage forms SG-|01210| (LEFT SWEEP). While the first portion of this compound gesture, subgesture SG-|012|,

¹⁰⁴ More about the interpretive potential of "departure and return" and balance (in general) is offered in Chapter 5.

is migratory, the second subgesture, SG-|210|, is accumulative. Thus, the composite sweeping gesture's effect on the sonic space changes as it runs it course: the sonic space initially consists of a single, migrating event. It later expands into the totality of available ensemble space.



Figure 3-24: Schoenberg, String Trio, mm. 155-156, LEFT SWEEP

3.6 Ensemble Spaces: General Observations

Different ensemble spaces afford different opportunities for encountering the types of gestures described above.¹⁰⁵ Some gestural types (including entire gestural categories) are not possible in ensembles of a specific cardinality, and the number of potential gestures increases dramatically as ensemble point-cardinality increases. The appendix offers a delimited list of the spatial gestures afforded by each of the ensembles explored below.

The spatial activity afforded by an ensemble of only two instruments (duo) is quite limited. Since only two ensemble points exist, we may only encounter point shifts and adjacentpoint oscillation gestures. A solitary point shift, such as PS 0-1 or PS 1-0 would be considered

¹⁰⁵ The spatio-gestural affordances of specific ensembles affect the listener's expectations. The concepts of "affordances" and expectation, and how they influence the perception of musical events, shall be taken up in Chapter 5.

open. It is possible to view such activity as a reduced and compressed instance of either a L-R or a R-L PAN, although panning gestures are much stronger (i.e., more salient) if they contain an event-point cardinality of greater than 2 (which is not possible in a duo setting). Distant-point oscillation gestures are not possible: the only two points present must be seen as adjacent, irrespective of the actual distance (physical space) between the two ensemble members in performance. The two points are not separated or divided by any other point(s), so they are adjacent. However, as the distance between the two adjacent ensemble points increases, the antiphonal (stereophonic) spatial effect is enhanced. Both strong and weak adjacent-point oscillation gestures are possible in a duo setting.

In considering only gestures in which each ensemble point is active once and only once (exempting oscillation gestures and strong closed gestures), all types of spatial gestures are possible in three-point ensemble space *except* weak open gestures. It is impossible for any point other than 1 to serve as the interior event point of a gesture in which 0 and 2 are the outer event points. Thus, the only open gestures possible are SG-[012] and SG-[210], both of which are classified as strong. The only four members of the weak closed gesture category in three-point ensemble space are all spiral gestures: SG-[021] (clockwise inward spiral), SG-[102] (counterclockwise outward spiral), SG-[120] (clockwise outward spiral), and SG-[201] (counterclockwise inward spiral). The weak closed gesture category consisting solely of spiral gestures occurs exclusively with three-point ensemble space: in an ensemble space with a cardinality of four or higher, this category will contain some members that are non-spiral gestures. Finally, within a three-point ensemble space, a repeated sequence of *any* ordering of the discrete points shall yield a cyclic panning gesture. For example, SG-[012012012], SG-[120120120], SG-[20210210], SG-[10210210], and SG-[021021021] are all

176

different forms of cyclic panning gesture. Based on their starting and ending points, only two of these compound gestures are classified as weak-open (SG-|012012012| and SG-|210210210|), while the rest are categorized as weak-closed. A cyclic panning gesture such as SG-|2102102| would be considered a strong compound closed gesture in three-point ensemble space.

Obviously, the higher the ensemble-space point cardinality, the greater the variety of spatial gestures afforded by the ensemble. As ensemble-point cardinality increases from one ensemble space to the next, the ratio of spiral to non-spiral weak closed gestures changes exponentially in favor of the latter. For an ensemble space of any cardinality greater than two, there will always be only two perfect panning gestures in the category of strong open gesture. Likewise, there will be only two perfect sweeping gestures in the category of strong closed gesture. Either category may contain abbreviated forms of the perfect gestures. A five-point ensemble space may be partitioned (depending upon context) into smaller ensemble space segments. If prevalent gestures occur within points 0 through 2, this 3-point ensemble-space segment could exhibit the gestures identified in the three-point ensemble space section of the appendix.

Multiple gestures within a given category are related by virtue of their shared categorical membership, the qualifications for which have been delineated above. The invariant qualities that adhere between category members allow them to be received either as similar or as related in specific ways (i.e., retrograde and/or inversional relationships, etc.). A preponderance of like (intracategorical) spatial gestures in a work of music would lend it a certain degree of coherence and unification in the spatial realm. Category members may often be linked via one or more transformative procedures or relational factors, which is the focus of the next chapter.

177

A spatial gesture is a gestalt structure: a unified whole comprising multiple parts—the adherence of which depends upon the specified relations among parts as well as the relations of each part to the whole. A spatial gesture is directed motion through an ensemble. With its dynamic shape materializing from the perceptual grouping of sonic events that occur successively at discrete points in ensemble space, a spatial gesture has a unique Gestaltqualität (form-quality) that transcends the mere summation of its individual event-point contents (the whole is greater than the sum of its parts). Although the motion through ensemble space is not continuous, the emergent patterning is received as a unified kinetic entity—as explained by the Gestalt grouping principles and by Albert Bregman's conception of auditory stream segregation. The events that form a spatial gesture may or may not be similar. A qualitative likeness of component sound events results in a spatial gesture that is likewise a mobile auditory stream (forged via sequential integration). However, an auditory sequence of unlike sounds that are spatially dispersed may likewise give rise to a spatial gesture. The directed motion that emerges from the consecutive activity of multiple ensemble points is the Gestaltqualität of a spatial gesture. This emergent motion may be abstracted from all other musical and sonic qualities.

CHAPTER 4

RELATING MULTIPLE SPATIAL GESTURES AND EXPLORING ALTERNATIVE SPACES

Chapter 3 introduced a system for depicting spatial gestures with integer notation. SGlabels allow for easy identification of a gesture's spatio-directional attributes and, as we shall see in the present chapter, facilitate the comparison of multiple gestures. We likewise explored the specific spatial features that enable a given gesture to be assigned to one of three gestural categories (and respective subcategories). The categories are intimately bound with Gestalt theories of form and pattern perception, and varying gestures belonging to the same category share certain characteristics. In this chapter, I discuss the ways in which spatial gestures may be altered through a number of transformative procedures. These procedures, many of which are comparable to those applied to pitch and rhythmic motives, serve to illustrate the interrelationships between different gestures—gestures that may or may not unfold at the same time in a musical work. I present the tools and terminology for describing the interactivity between two or more gestures occurring simultaneously. Next, I explore the gestures (and alterations thereof) potentially encountered within various two and three-dimensional spaces. Finally, the chapter concludes with a discussion of spatial gestures in electronic music.

4.1 Alterations of Spatial Gestures

Having explored the different categories of spatial gestures that a one-dimensional (spanning left to right across a single horizontal plane) ensemble space may afford, we shall now explore the ways in which such gestures may be developed, varied, and interrelated. The same abstract spatial gesture may assume a variety of guises, since the sonic events that conspire to give a gesture form can be virtually any sound issued by an instrument; gestures are defined and delineated by the temporal succession of event location—irrespective of the sonic attributes of the events at each point. The temporal length, spatial size, and/or overall shape of a gesture may be altered in a variety of ways. The ordering of consecutive event points within a gesture, as well as the gesture's directionality, may be changed. The event cardinality of a gesture may be increased or decreased. Some gestural modifications preserve categorial kind while others alter a gesture's category membership.

Variations of spatial gestures are important because all of the varied forms of a gesture are somehow related, and the proliferation of related gestures gives coherence to the spatial construct of a musical work. Consider the many ways in which a melodic motive might be varied: it may undergo transposition, sequencing, fragmentation, retrogression, inversion, intervallic expansion or compression, temporal augmentation or diminution, and any combination of these procedures. Spatial gestures may be transformed along similar lines. When analyzing a work of music, an atemporal analysis may highlight related gestural forms. Such a spatial analysis may consist of either a catalog of related gestures or a synchronic network revealing the exact relationships between various gestural forms (gestural networks shall be explored in Chapter 5). In describing the diachronic progress of gestural forms as they actually unfold throughout a work of music, the order of presentation of any two related gestures shall be considered in determining which form is considered to be the altered version of the other. For example, SG-|013| shall be considered an abbreviated version of SG-|0123| if the former occurs *after* the latter. If, on the other hand, SG-|0123| emerges after the occurrence of SG-|013|, then SG-|0123| shall be counted as an elaborated form of SG-|013|. Abbreviation and elaboration shall be detailed below.¹ What really matters, however, is that the two gestures are categorically related by some transformative procedure, and this likeness—mingled with variation—serves to unify the spatial structure of the work under investigation.

Temporal Modifications

Temporal modifications include **augmentation** and **diminution** of the timeframe in which a spatial gesture unfolds. As an example, let us suppose that the original presentation of SG-|0132| in a four-point ensemble space takes four seconds to unfold. If, later in the piece, another instance of SG-|0132| transpires over eight seconds, this latter-occurring gesture would function as a temporal augmentation of the initial gesture.² If, still further into the piece, SG-|0132| takes only two seconds to occur, this would be a temporal diminution of the two previous statements. Since the SG-label for a gesture remains the same regardless of the time intervals separating component events, a temporal modification preserves gestural kind.

Due to the nature of oscillation gestures (the perpetual alternation of two event points), this category of gesture requires a slightly more complex apparatus to describe temporal relations. Imagine that SG-|02020202| occurs twice in a work, first unfolding over 3 measures

¹ While some specific musical examples are provided in this section, I mostly employ hypothetical situations (as abstract examples) to explain the various types of gestural alteration. Many of the alterations not supported by a concrete musical example in this chapter shall be demonstrated in the spatial analysis provided in Chapter 6. ² Since a spatial gesture may be abstracted from the sonic qualities of the events they comprise, multiple instances of SG-[0132] would be related despite the degree of qualitative dissimilarity between the events of each gesture. However, a shared sonic or musical attribute between two gestures enhances the salience of their relatedness.

A spatial gesture comprising events that transpire in rapid succession is likely more salient than a temporally expansive gesture. Robert Hatten states: "A prototypical gesture is a relatively short temporal gestalt that generally occurs within the temporal frame of the experiential present, or working memory (ca. 2 seconds)."³ Hatten supports this assertion by stating that such gestures "take advantage of two major forms of representation in the brain: the immediate or imagistic, and the sequential or temporal."⁴ Such gestures "are perceptually salient because they involve a fusion of both *immediate* (understood as 'instantaneous' and 'issuing transparently') imagistic gestalts of complex individual events, and *temporally mediated* gestalts (understood as energetic shaping through time)."⁵ A spatial gesture unfolding in two seconds or less unfolds almost entirely at the "now-cursor" (or, experiential present) of the "perceptual

³ Hatten, Interpreting Musical Gestures, Topics, and Tropes, 101.

⁴ Ibid.

⁵ Ibid., 102. Emphasis Original.

frame." In addition to the experiential present, the *perceptual frame* involves the "retention" and "protension" of prior and anticipated events.⁶ Thus, a rapid spatial gesture transpires more fully in the "now" of perception without relying on the retention of past events to construct the whole. Put another way, a small onset-to-onset interval between consecutive events translates to less temporal discontinuity between events—a condition that promotes and improves perceptual grouping (as the temporal proximity of discrete events is heightened). These observations do not hamper the analytic relevance of spatial gestures that unfold over longer time spans. Such gestures simply may be less immediately perceptible and less "prototypical" in Hatten's estimation.⁷ Hatten grants that "gestures may also be hierarchically organized, in that larger gestures can be comprised of smaller gestures."⁸ Presumably, the larger musical gestures may exceed the two-second timeframe of the experiential present. We shall encounter such situations—in which smaller spatial gestures (as surface phenomena) are embedded in more expansive gestures—in Chapter 6.

Intragestural Temporal Processes

Augmentation and diminution, as described above, allow us to describe the temporal relationships between two or more spatial gestures. Yet, we will often find it useful to portray the changes in temporal proportions within a single gesture. The temporal spacing between the onsets of discrete events within a single gesture will not always be uniformly proportionate. When such spacings are equivalent (isochronic), the sameness of onset-to-onset interval facilitates the perceptual grouping (sequential integration) of successive events into a coherent

⁶ These terms are taken from David Lewin, "Music Theory, Phenomenology, and Modes of Perception," *Music Perception* 3.4 (1986): 329.

⁷ The spatial analysis of a musical work may reveal important gestures that are more difficult to perceive in the real time performance of that work. A spatial analysis might facilitate the aural discernment of such gestures.

⁸ Hatten, Interpreting Musical Gestures, Topics, and Tropes, 94.

gesture. Often however, the lack of regularity between the temporal spacing of successive event onsets will contribute to a process of **intragestural acceleration** or **deceleration**. Either of these processes may result from the actual rhythmic configuration of the gesture (within a steady tempo) or from tempo fluctuations (such as accelerandi or ritardandi). Such temporal processes obstruct grouping on the basis of uniform onset-to-onset interval. However, if the interval between consecutive events progressively increases (intragestural deceleration) or decreases (intragestural acceleration), the continuity of change capacitates grouping. Neither a uniform nor a continuously changing onset-to-onset interval is requisite for gestural formation and perception, but either of these qualities may strengthen event grouping and gestural salience.

Spatial Alterations

A spatial alteration will alter the overall shape, spatial size, event cardinality, and/or directionality of a spatial gesture. Spatial alterations are more numerous and varied in kind—although not necessarily more frequent in occurrence—than the temporal modifications discussed above, and they frequently occur in conjunction with some form of temporal adjustment. In the following discussion, some transformative procedures are classified as **intracategorical alterations**—meaning they preserve a gesture's categorial membership. Of these alterations, some are transformational operations that affect the spatial positioning or directional orientation of a gesture by permuting the order of event points within the SG-label while preserving the total event cardinality. Other intracategorical alterations may change the spatial size of a gesture by increasing or decreasing event cardinality and causing a gesture to acquire or shed disjunctions.

Other procedures are codified as **intercategorical alterations**, which indicates that they affect category or subcategory membership. An intercategorical alteration, for example, might change a strong open gesture into either a weak open gesture or a weak closed gesture. Many alterations may be either intra- or intercategorical depending upon the precise nature of their application. Furthermore, multiple transformative procedures (potentially a free mix of intra- and intercategorical types) may be applied jointly to a single gesture. It is quite common for open and closed gestures to be related via some transformative procedure. However, it is generally difficult to gauge relational factors between oscillation gestures and either open or closed gestures; oscillation gestures do not easily morph into either an open or closed gestural form.⁹ We shall begin with an exploration of intracategorical transformational operations, since these generally yield the most noticeable correlations between gestures.

Intracategorical Transformational Operations

The following intracategorical transformations maintain the event cardinality and number of directionality changes of the original gesture, although the gesture's exact spatial orientation (within the total ensemble space) or overall directionality may be changed. All intracategorical transformations preserve category and sub-category membership: a weak closed gesture will transform into another weak closed gesture.

Nudging occurs when an entire gesture is shifted within the ensemble, always from one segment of ensemble space to another. Nudging preserves the overall directionality of the gesture. For example, imagine that SG-|012| is followed by SG-|123| in a four-point ensemble

⁹ In broad terms, weak oscillation gestures are akin to open gestures and strong oscillation gestures are similar to strong closed gestures. Strong distant-point oscillation gesture SG-|030| may be "filled-in" to produce strong closed gesture SG-|0123210|. In the following discussion, most of the examples of intercategorical alterations/relations exist between open and closed gestures.

space.¹⁰ In this case, each discrete event point within the entire original gesture is "nudged" one ensemble point to the right.¹¹ Spatial nudging is similar to transposition or sequencing in the realm of pitch, where a melodic motive might be transferred up or down in pitch space. We shall indicate that a spatial gesture has been *right-nudged* if it is transferred to the right of the ensemble space (as in the above example) or *left-nudged* if displaced to the left. The additional labels (+x) or (-x), where "x" signifies the degree of displacement, shall indicate the number of ensemble points that a gesture is nudged. So, SG-|123| is the right-nudged (+1) version of SG-|012|. Conversely, SG-|012| is the left-nudged (-1) form of SG-|123|. It is important to be able to identify the degree (i.e., "distance") of nudging, as such considerations come into play when examining the interpretive implications of multiple, consecutive nudgings.¹²

Retrogression occurs when the sequential ordering of event points composing a gesture is reversed; the gesture's directional "path" is likewise reversed. SG-|0132| becomes SG-|2310|, and SG-|21043| becomes SG-|34012|. Retrogression is an operation on the SG-label, in which the contents of event slots are reversed.¹³ Retrogression always reverses the dominant (or, large-

¹⁰ Both of these gestures are classified as partial open gestures—each unfolding in a different three-point segment of ensemble space. Depending upon context and the exact manner in which both of these three-event gestures are presented, it might prove optimal to analyze them collectively as SG-|012123|, which is a weak open gesture. However, we could still break this six-event compound gesture into its two discrete, three-event subgestures to reveal how their "nudged" relationship forms the larger gesture.

¹¹ In this particular example, the direction of the nudging corresponds to the directionality of the spatial gesture(s); therefore, there is a correlation between gestural directionality and transformative (relocational) direction.

¹² For example, imagine the following sequence of spatial gestures occurring within a six-point ensemble space and unfolding over any time span (they need not occur contiguously): SG-|021|, SG-|132|, SG-|021|, SG-|243|, SG-|021|, SG-|354|. All of these gestures are categorically related. SG-|021|, appearing three times, is continuously rightnudged, but the degree of the nudging progressively increases: SG-|021| \rightarrow [right-nudged (+<u>1</u>)] \rightarrow SG-|132|; SG-|021| \rightarrow [right-nudged (+<u>2</u>)] \rightarrow SG-|243|; SG-|021| \rightarrow [right-nudged (+<u>3</u>)] \rightarrow SG-|354|. The gradually increasing distance of nudging betrays the long-term process of transferring the gesture from ensemble-space segment 0–2 to segment 3–5. Furthermore, the greater the degree of nudging, the more "tensive" the procedure seems to be (requiring a greater index of effort and increased agency to "move" something a greater distance). The above sequence of gestures may be interpreted as the gradual heightening of tension.

¹³ Retrogression may also be viewed as inversion affecting the SG-label (*not* the ensemble space): when SG-|0132| becomes SG-|2310|, the inversional axis between event slots 2 and 3 causes event points 1 and 3 as well as 0 and 2 to exchange locations (the contents of event slots 2 and 3 are exchanged, as are the contents of event slots 1 and 4). When SG-|21043| becomes SG-|34012|, the inversional axis *is* event point 0 in event slot 3; thereby, event point 0 maps into itself and remains in event slot 3, while event slots 1 and 5, as well as 2 and 4, "swap" event points.

scale) directionality of a spatial gesture. The spatial gesture featured in Figure 3-16 (see p. 161), SG-[3102013], is a strong closed gesture. The gesture may likewise be viewed as a compound gesture comprising two weak closed gestures, SG-[3102] and SG-[2013], that are elided at event point 2. These two component gestures are related by retrogression—hence the palindromic quality (spatial symmetry) of the gesture.¹⁴ Because of this symmetry, the gesture as a whole is "non-retrogradable" in Messiaen's sense of the term: the gesture cannot be reversed to yield a new gesture.¹⁵ The retrograde of SG-[3102013] is SG-[3102013]—the two gestures are identical. Similarly, a perfect sweeping gesture such as SG-[01210] (presented in Figure 3-24) cannot be retrograded.

Inversion is an operation that maps ensemble points into one another across some axis of symmetry within the ensemble space. A **type-1 inversion** entails reflecting a gesture across an axis abiding within the center of the total ensemble space. In an even-cardinality ensemble space, this axis exists *between* the two central ensemble points: in a four-point ensemble space, the axis would fall between points 1 and 2. In an odd-cardinality ensemble space, the axis occurs *through* the central ensemble point: point 2 in a five-point ensemble space. In such a space, the central ensemble point will always map into itself under type-1 inversion. To find the inversion of SG-[04123] (in a five-point ensemble space), simply "swap" the SG-label positions (event-slot locations) of event points 0 and 4 (the spatial extremes of both the gesture and the ensemble

¹⁴ As previously noted, in addition to being spatially symmetrical, this particular gesture is temporally symmetrical as well. Since it consists of perpetual sixteenth notes (in either sextuplet or triplet figures) at each event point, we may quantify its dynamic temporal proportions as: 6-6-9-12-9-6-6, a numeric palindrome in which each integer represents the total number of sixteenth-notes at each consecutive event point within the gestural sequence. Viewing the compound gesture as a whole, event point 2 (the central "hinge" of the gesture) is the most temporally expansive—enduring for 12 sixteenth notes. We may thus describe the intragestural temporal process of SG-[3102013] as first involving intragestural deceleration (up to event point 2) and subsequently entailing intragestural acceleration (from event point 2 to its conclusion). In other words, the compound gesture's unfolding "slows down" during the presentation of its first subgestural component (SG-[3102]) and then "speeds up" during the laying out of its second component (SG-2013, the retrograde of SG-[3102]).

¹⁵ Olivier Messiaen, The Technique of My Musical Language, trans. John Satterfield (Paris: A. Leduc, 1956).

space in which it unfolds), and then swap the position of integers 1 and 3, resulting in SG-|40321|. Event point 2 remains in event slot 4 because it occurs at the central point of the ensemble space (ensemble point 2).

Figure 3-17 (see p. 162) illustrates weak closed gesture SG-|0132| as exposited at the beginning of Bartók's *String Quartet No. 1*. Figure 4-1 shows the gesture that transpires in mm. 33–38 and begins the second subsection of the quartet's first movement.



Figure 4-1: Bartók, String Quartet No. 1, Mvt. 1, mm. 32-38

This gesture, SG-|3201|, is the inversion of the movement's opening gesture, SG-|0132|. The two gestures unfold at different rates (the latter is a temporal diminution of the former), and the musical material constituting each gesture is quite contrasting. In fact, the musical activity at each point in SG-|3201| is fairly distinct (with the exception of that at points 2 and 1). In spite of

these differences, the spatial gestures are related via the transformational operation of inversion.¹⁶

A type-2 inversion involves an axis of symmetry that does not equally divide the total ensemble space. Type-2 inversion can only affect a spatial gesture that does not contain both spatial extremes of the total ensemble space. If SG-023 occurred in five-point ensemble space, we may invert it within the ensemble-space segment spanning points 0-3 since point 4 (the rightmost spatial extreme of the ensemble space) is not active in the gesture. Therefore, the gesture would reflect across an axis falling between ensemble points 1 and 2, yielding SG-[310]. However, such a gesture may likewise undergo type-1 inversion, which would result in SG-[421] after reflection across ensemble point 2. In this case, both inversion types yield gestures that are related to one another via the operation of nudging: by nudging SG-[310] (the result of type-1 inversion) one point to the right, we arrive at SG-[421] (the result of type-2 inversion). In fact, the gestural result of any type-2 inversion (where applicable) will always be a nudged form of the product of type-1 inversion. For any gesture that contains both spatial extremes of the total ensemble space, type-1 and type-2 inversions are identical processes with matching outcomes. In short, type-1 inversion reflects a gesture through the entire ensemble space (and may therefore map a gesture from one ensemble space segment to another), while type-2 inversion (only applicable to a gesture occurring within ensemble-segment space) retains the ensemble-segment space location of that gesture.

A gesture may undergo **double inversion** when subjected to both type-1 and type-2 inversion. This is only applicable to gestures occurring in a segment of the total ensemble space

¹⁶ When the first movement's opening material returns in m. 53, the initial SG-|0132| is recast as SG-|0123| (mm. 53–58): a weak closed gesture is thus converted into a strong open gesture via *point exchange*—a procedure to be outlined below. SG-|2310|, the retrograde of SG-|0132|, appears in mm. 159-174 of the quartet's third and final movement.

and is particularly useful for describing transformational relationships between palindromic, nonretrogradable gestures.¹⁷ For example, SG-|01210| (in five-point ensemble space) may be transformed into SG-|23432| via double inversion. The order in which the two inversion types are conducted does not affect the resultant gesture.¹⁸

Retrograde-Inversion entails reversing a gesture and then inverting it (or vice versa). This procedure may involve either type-1 or type-2 inversion, as described above. The retrograde-inversion of SG-[10342] (in five-point ensemble space) is SG-[20143].

The following weak closed gestures are related to SG-|0132| by virtue of the three preceding transformational operations: SG-|2310| (R), SG-|3201| (I), and SG-|1023| (RI). Each gesture contains the same number of directionality changes (one), although the proportion of left-to-right to right-to-left motion varies from one gesture to another, as does the prevailing directionality of each gesture.¹⁹

Note that the retrograde and the inversion of some spatial gestures are identical, which also means that the original form and its retrograde-inversion will be duplicates. For example, both the retrograde and the inversion of SG-|0123| (L-R PAN) are SG-|3210| (R-L PAN). The retrograde-inversion of SG-|0123| is the identical gesture SG-|0123|. Such invariance reduces the means by which a gesture may be transformed to yield a related gestural form and thereby

¹⁷ Non-palindromic gestures may often be transformed via retrograde-inversion, to be described below. For instance, in five-point ensemble space, SG-|012| may be transformed into SG-|234| via retrograde-inversion (type 1), double inversion, *or* nudging. Outside of any broader transformational context, it is preferable to choose the simplest transformative procedure to describe gestural relations. In this case, nudging only involves one step, whereas the other two transformations each require two steps, so nudging is the optimal candidate for describing the transformational relationship between gestures. However, it might be desirable to choose one of the other transformative modes if the actual operation itself were recurrent (affecting multiple gestures and perhaps figuring prominently in a transformational network) within a particular section of music.

¹⁸ SG- $|01210| \rightarrow [Type-1 inv.] \rightarrow SG-|43234| \rightarrow [Type-2 inv.] \rightarrow SG-|23432|; or, SG-|01210| \rightarrow [Type-2 inv.] \rightarrow SG-|21012| \rightarrow [Type-1 inv.] \rightarrow SG-|23432|. Note that SG-|23432| is also the right-nudged (+2) form of SG-|01210|: SG-|01210| \rightarrow [right-nudged (+2)] \rightarrow SG-|23432|. Nudging is a more perceptually salient alteration and thereby generally more pertinent to analytic discourse.$

¹⁹ SG-|0132| and SG-|1023| both contain two left-to-right point shifts and one right-to-left point shift, although the two gestures do not have a single point shift (the specific ordering of two event points) in common. Conversely, SG-|2310| and SG-|3201| both contain one left-to-right point shift and two right-to-left point shifts.

restricts the number of members of a given gestural category. The delimitation provided by transformational invariance explains why there are only two possible perfect panning gestures for an ensemble space of any cardinality.

For any weak oscillation gesture (such as SG-|010...1|), the retrograde and type-2 inversion will be identical, as will be the original form and the retrograde-inversion (type 2). However, for a strong oscillation gesture (such as SG-|2323...2|, identity prevails between the original and retrograde forms as well as between the type-2 inversion and retrograde-inversion (type 2) forms (both SG-|3232...3|).

Other Spatial Alterations

The following alterations may be either intracategorical or intercategorical, depending upon the gesture to which they are applied and/or the precise manner in which they are applied to that gesture. With intercategorical alterations, two gestures (the original form and the altered version)—while not categorically affiliated—may exhibit like features that enable them to be perceived as similar and related. A **type-A intercategorical alteration** involves a change of subcategory membership while broad gestural categorization remains the same. For example, a weak *closed* gesture changing into a strong *closed* gesture would be a type-A alteration since both are closed gestures. A **type-B intercategorical alteration** involves the maintenance of either the "strong" or "weak" qualifier across category boundaries, as when a *strong* open gesture is modified to become a *strong* closed gesture.²⁰ A **type-C intercategorical alteration** preserves neither category membership nor the strong/weak descriptor. Clearly, a type-C

²⁰ The relevance of the type-B alteration is supported by the Gestalt principle of Prägnanz (good form). A strong open gesture is considered strong due to its "good continuity," in the form of unidirectional motion (the Law of Continuity). The strength of a strong closed gesture results from its demonstration of spatial "closure" or "completeness" (the Law of Closure). Any strong gesture, irrespective of its categorial affiliation, is well formed based on one of the Gestalt principles of Prägnanz.

alteration will be more drastic (i.e., more of an alteration) than either a type-A or type-B alteration, and the relation between the original and the altered gestural forms may be less obvious (i.e., the two forms may not exhibit any pronounced similarity). All intercategorical alterations are considered to be more severe than intracategorical alterations. While the products of intercategorical relations might be less effective than those of intracategorical relations at conveying a sense of unity and coherence within the spatial complex of a musical work, they may still account for "intergestural motion" (or, connectivity between gestures) within that spatial complex. Furthermore, if a specific type of transformative procedure occurs repeatedly throughout a work, that abstract operation functions as a recurrent and unifying motive—one that links multiple gestures to one another.

The following alterations preserve the event-point cardinality of the original gesture.

Expansion occurs when the space traversed by a gesture is increased. The expansion may occur at either or both spatial extremes of the gesture, provided that the expanded form of a gesture is of the same cardinality as the original. For instance, SG-|012| may expand to SG-|013|; in this case, both are three-event gestures. In four-point ensemble space, oscillation gesture SG-|121...| may be expanded to SG-|020...|, SG-|131...|, and SG-|030...|. Expansion often entails the introduction of disjunct motion into previously conjunct motion (as in the above examples) or an increase in the degree of disjunction (the number of points skipped over in a single point shift) already present within the gesture. Expansion is always an intracategorical alteration.

Compression is the opposite of expansion. Compression occurs when the space that a gesture spans is decreased and the cardinality of the original and the compressed gesture are the same. SG-|023| may be compressed into SG-|123| or SG-|012|. Oscillation gesture SG-|030| may be compressed into SG-|020|, SG-|131|, SG-|010|, SG-|232|, and SG-|121|. Compression

192

frequently involves the reduction or elimination of disjunct spatial motion, resulting in a more (or entirely) conjunct gesture. Like its converse, compression is solely an intracategorical alteration.

A **point exchange** alters a portion of the sequence of event points within a gesture. A point exchange occurs when any two event slots "exchange" event-point integers. The two event slots may be adjacent or nonadjacent. When a point exchange occurs between neighboring event slots, we may refer to this specific form of point exchange as a **point-shift reversal**, since the two event points comprising a single point shift (within a single point-shift slot) are essentially retrograded and the directionality of the point shift is invariably reversed. Consider SG-[01324], a weak open gesture. We may alter this gesture to produce SG-[03124] by exchanging event points 3 and 1 (originally contained in event slots 2 and 3, respectively). We have thus effectively "reversed" PS 1-3 (in point-shift slot 2) to yield PS 3-1; left-to-right motion becomes right-to-left motion. These gestures are strongly related because they share the same spatiotemporal endpoints (they each contain event point 0 in event slot 1 and event point 4 in event slot 5) and are members of the same gestural family. Two interior event points undergoing pointshift reversal is similar to the "adjacent transposition" of interior letters in a word, as discussed in Chapter 3 (see footnote 9). "House" may be misspelled as "huose" or "hosue." The minor alterations produced by "swapping" adjacent and interior letters do not detract from the word's recognizability. Similarly, SG-|01324| and SG-|03124| would likely be perceived (or recognized) as highly similar in a performance context.

Nonadjacent event points may also be exchanged. For example, SG-|34210| may be modified to produce SG-|04213| by exchanging the event-point contents of event slots 1 and 5 (such that event points 0 and 3 "swap" positions). The resultant gesture remains a weak closed

193

gesture, but the number of directionality changes has increased from 2 to 3.²¹ Beginning with SG-|04231|, and exchanging the contents of event slots 2 and 5, SG-|01234| is produced—a weak closed gesture has become a strong open gesture (type-C intercategorical alteration). Continuing the analogy with word recognition, "hsuoe," "eoush," and "uohse" are progressively more difficult to recognize as permuted forms of the word "house." The first misspelling exchanges non-adjacent but interior letters (o and s), the second interchanges the first and last letters (h and e), and the third trades the first letter (h) with an interior letter (u). A point exchange among nonadjacent event points in a spatial gesture, particularly when it involves one or both of the temporal endpoints of a gesture, typically results in less salient gestural relations—undoubtedly due to the fact that such an alteration commonly yields intercategorically related gestural forms. The Gestalt-based conditions for category membership, which are based upon perceptual tendencies, typically override the relational factors that may hold between gestures in different categories.

A **point-shift exchange** is an alteration that occurs when two point shifts are switched. This mechanism is similar to a point exchange, except that a pair of temporally contiguous event points (constituting a point shift) is switched with another discrete point shift. This procedure can occur only between discrete point shift slots—it cannot occur between overlapping PS slots (those that share an event point). Two numerically adjacent PS slots (such as PS slot 1 and PS slot 2) cannot be exchanged because they share an event slot (both PS slot 1 and PS slot 2

²¹ It is possible to view the intracategorical transformational operation of retrogression as a series of point exchanges between points equidistant from the center point. So, with a four-point ensemble space in which SG-|0132| becomes SG-|2310|, the event points in event slots 1 and 4 exchange, as do those in event slots 2 and 3 (the latter exchange also counts as an instance of point-shift reversal). In a quintet scenario, where SG-|21043| becomes SG-|34012|, event point 0 (in event slot 3) remains stationary, serving as the central axis (of the SG-label) across which event points are exchanged. A point exchange occurs between event slots 1 and 5 as well as 2 and 4. No point-shift reversals occur in this instance because no adjacent event slots exchange event points. The retrograde-inversion of SG-|0132| may likewise be achieved by swapping the event-point integers of point-shift slot 1 (0 and 1) as well as those of point shift slot 3 (3 and 2), resulting in SG-|1023|.

contain event slot 2). Consider the modification of SG-|0132| into SG-|3201|. These two gestures share in an inversional relationship, as described above. However, we may also arrive at SG-|3201| from SG-|0132| by exchanging the contents of point-shift slot 1 (PS 0-1) with those of point-shift slot 3 (PS 3-2). The event-point ordering within each point shift is preserved while the sequence of the point shifts themselves is changed. PS 1-3 (in point-shift slot 2) from the initial gesture is eliminated and replaced by PS 2-0 (in the same PS slot).²²

Point-shift exchange may be used to describe alterations that cannot be ascribed to other transformative procedures. Beginning with SG-|13420|, and exchanging the point-shift contents of PS slot 1 (PS 1-3) and PS slot 3 (PS 4-2), we arrive at SG-|42130|. We have thus turned a weak closed gesture into a weak open gesture, a type-B intercategorical alteration. No other procedure can account for the relationship between these two gestures.

Rotational slot displacement involves rotating the contents of every event slot *x* degrees to the right. The rightmost event point "wraps around" to become the leftmost event point. SG-|0123|, when displaced one degree to the right, becomes SG-|3012|. What was initially a perfect L-R PAN (strong open gesture) is now a **displaced panning gesture**, which is classified as a weak closed gesture (a type-C intercategorical alteration). If SG-|3012| is again displaced by one degree, it becomes SG-|2301|. This series of two one-degree displacements is the same as displacing the original L-R PAN (SG-|0123|) by two degrees to the right. To simplify relational modeling, the direction of rotation shall always be to the right. We can use a label such as

²² It is likewise feasible to view this alteration as one point exchange between event slots 1 and 3 and another point exchange between event slots 2 and 4. However, exchanging a fewer number of larger and adjacent units (in this case, only two point shifts) is slightly less cumbersome than exchanging a greater number of smaller and nonadjacent units (four event points). As a general rule, the simplest transformative procedure (i.e., the one requiring the fewest steps) should be used to model the relationship between two gestures. If such relationships are actually perceived during the musical experience, it is likely that the simplest operation is the most perceptually salient. As another example, in many cases where the inversion and the retrograde of a gesture are identical, the retrograde relationship is perhaps easier to discern than the inversional one. SG-[3210] is more readily perceived as the retrograde of SG-[0123], since the reversal of directionality is quite clear.

RSD(2) to indicate that a gesture has undergone rotational slot displacement by two degrees, always to the right.²³ Rotational slot displacement is similar to nudging (described above), except that it affects the entire ensemble space and involves the "wrap-around" procedure.²⁴

Partial rotational slot displacement occurs when only a portion of a gesture's SG-label undergoes rotational slot displacement. SG-|21043| (displaced panning gesture) is related to SG-|21304| (counterclockwise outward spiral) via partial rotational slot displacement: the event points in event slots 1 and 2 do not change, while the contents of slots 3 and 4 each rotate one slot to the right (to slots 4 and 5) and the original component of slot 5 (event point 3) "wraps around" to fill slot 3. The same procedure will change SG-|23401| (displaced panning gesture) into SG-|23140| (clockwise outward spiral), SG-|40123| (displaced panning gesture) into SG-|40312| (counterclockwise inward spiral), and SG-|04321| (displaced panning gesture) into SG-|04132| (clockwise inward spiral). In each of these cases, partial rotational slot displacement is an intracategorical alteration.

The following alteration types do *not* preserve the event cardinality of the original gesture. Some (but not all) of these alterations maintain the gesture's basic directionality and number of directionality changes. Many of these procedures may be applied exclusively to open or closed gestures, and none of them may be applied to oscillation gestures.

Abbreviation refers to any process by which the event cardinality of a gesture is reduced by at least one. Abbreviation often involves "shrinking" a gesture from the total ensemble space to a segment of ensemble space or introducing disjunct spatial motion. In a five-point ensemble

²³ It would certainly be feasible for the index of displacement to be negative—representing movement to the left. Beginning with SG-|0123|, RSD(3) and RSD(-1) both yield SG-|1230|. In this case, RSD(-1) indicates a less severe alteration than does RSD(3). It will prove convenient, however, to keep the direction of displacement constant. Rotation to the right (as the direction of displacement) is, admittedly, an arbitrary convention.

²⁴ If an ensemble performing Bartók's *String Quartet No. 4* assumed the "viola-out" seating configuration (from left to right: Violin 1, Violin 2, Cello, Viola), the passage shown in Figure 0-4 (see p. 8) would yield a *displaced cyclic panning gesture*: SG-|103210321032|. This compound weak closed gesture is "displaced" from SG-|321032103210| via RSD(2).

space, SG-|01234| may be abbreviated to produce the following gestural forms: SG-|0234|, SG-|0134|, SG-|0124|, SG-|034|, SG-|014|, and SG-|024| (via the **deletion** of events at interior ensemble points); SG-|0123| and SG-|1234| (via **truncation**: the elimination of one of the gesture's spatial extremes²⁵); SG-|123| (via **deflation**: the omission of both of the gesture's spatial extremes); SG-|(0+1)234|, SG-|0(1+2)34|, SG-|01(2+3)4|, and SG-|012(3+4)| (via the **fusion** of two points). When applied to a perfect panning gesture such as SG-|01234|, these procedures produce either partial open gestures or imperfect panning gestures.

Elaboration refers to an alteration that increases the event cardinality of a gesture by at least one. Elaboration often entails enlarging a gesture from a segment of ensemble space to either a larger ensemble-space segment or to the total ensemble space. Alternately, elaboration may cause a previously disjunct gesture to recur in a more (or completely) conjunct form via the "filling-in" of interior gaps (spatial disjunctions). Elaboration is the converse of abbreviation. SG-[013] may be elaborated to produce SG-[0123] via the **addition** of an event at interior ensemble point 2. **Extension**, the introduction of a new spatial extreme to a gesture, will cause SG-[123] to become either SG-[0123] (*left-extension*) or SG-[1234] (*right-extension*). SG-[123] may undergo **inflation** (in a five-point ensemble space) to produce SG-[01234]. Finally, superpoint 2+3 in SG-[01(2+3)4)] may be **split** to yield SG-[01234].²⁶

 $^{^{25}}$ SG-|0123| is a *right-truncated* form of SG-|01234|, and SG-|1234| is a *left-truncated* abbreviation of the gesture. SG-|234| would be a more pronounced left-truncation of SG-|01234| since two events are excised. The truncation of SG-|43210| to SG-|4321| is likewise an instance of left-truncation because the leftmost spatial extreme, ensemble point 0, is eliminated. The spatial extremes of a gesture may or may not also be the outermost points of the total ensemble space. In the case of SG-|1432| occurring in five-point ensemble space, event point 1 is the leftmost spatial extreme of the gesture. Ensemble point 0, which is the leftmost spatial extremes and the temporal endpoints are the same: event points 0 (in event slot 1) and 4 (in event slot 5). Such is not the case with a weak closed gesture like SG-|10342|. In this gesture, the spatial extremes are interior gestural events (event point 0 in event slot 2 and event point 4 in event slot 4) and the temporal endpoints occur at interior ensemble points (event point 1 in event slot 1, event point 2 in event slot 5).

²⁶ The splitting of superpoint 2+3 could likewise result in SG-[01324]. Under such circumstances, a strong open gesture is altered to become a weak open gesture (a type-A intercategorical alteration).

All of the above procedures alter a gesture such that certain qualities remain invariant while other aspects of the gesture change. Thus, the original gesture and its altered form are related by some quotient of similarity. Some intergestural relations are more evident and perceivable than others. The overall directionality of a gesture may change (retrogression, inversion), as may its total number of directionality changes. The orientation of the gesture within the ensemble space may be altered (nudging), and the spatial size (or, range) of a gesture may be decreased (compression, abbreviation) or increased (expansion, elaboration). Additionally, the gesture's event cardinality may be decreased (abbreviation) or increased (elaboration). Many of the above alterations may occur in conjunction with one another. For instance, in a six-point ensemble space, SG-[542] may be viewed as the nudged, expanded retrograde of SG-[012].²⁷ Both gestures are classified as partial open gestures, and their mutual occurrence would impart a degree of spatio-motivic coherence to the work in which they materialize.

4.2 Polystreaming: Relations Among Concurrent Spatial Gestures

We often encounter multiple spatial gestures unfolding at the same time (or simply overlapping in time). This type of spatial activity differs from a compound gesture (as described in Chapter 3), in which two or more small and distinguishable gestures occur in succession to yield the composite, larger gesture. When two or more distinct gestures occur simultaneously, it is important to describe their dynamic, spatio-temporal interrelationship(s). As we shall see, the specific spatio-musical context in which gestural activity transpires determines whether a

 $^{^{27}}$ The retrograde of SG-|012| is SG-|210|. This gesture may be nudged 3 event slots to the right to yield SG-|543|, which, in turn, may be expanded to SG-|542|. (This is not the only ordering of the different transformative procedures that would yield the same result.) SG-|210| and SG-|543| are structurally and categorically related by the composite of abstract procedures that change them into one another.
multitude of sonic events should be analyzed as a single, complex spatial gesture or as multiple, simpler gestures unfolding contemporaneously. Generally, the higher the number of distinct points within an ensemble space, the easier it is to identify multiple spatial gestures occurring concurrently.²⁸ However, in an ensemble space of any cardinality (greater than two), we might experience a manifold of concomitant spatial gestures. I refer to the simultaneous presentation of two or more spatial gestures as **polystreaming**, where each individual gesture in the aggregation is considered a separate "stream." In this context, I am using the term "stream" in a somewhat more generalized sense than Bregman's conception of "auditory stream" (as outlined in Chapter 2). In my usage, "stream" is simply employed to indicate the concurrence of more than one spatial gesture. The gestures may or may not be "auditory streams" as defined by Bregman—a series of sounds that are perceptually grouped together (via sequential integration) as a result of their qualitative likeness. Each one of the distinct gestures engaged in polystreaming may emerge as a spatially dispersed sequence of qualitatively dissimilar sonic events. For instance, two gestures occurring in separate segments of ensemble space may be spatially dispersed to such an extent that they may be differentiated from one another solely by virtue of the spatio-temporal proximity of their component events-not requiring any qualitative similarity of event-point activity. Often, however, the different gestures *must* likewise be auditory streams (in Bregman's sense of the word) to aid in their individuation and segregation

 $^{^{28}}$ A single section of an orchestra may consist of twenty performers playing the same part; so just having a larger ensemble with many doubled parts would not necessarily enable the identification of multiple gestures occurring at the same time. Imagine a string quartet (four-point ensemble space) being magnified by having 20 performers on each part, for a total of 80 performers. There would still only be 4 unique parts. We would have to consider the 20 performers on a given part as occupying a single spatial *area* (assuming they are all proximate within the ensemble space). We could still label each of the four areas with integers 0, 1, 2, and 3. The ensemble space would be profoundly augmented (perhaps increasing the salience of spatial gestures), but the potential for specific types of spatial gestures would be identical to the original, four-member string quartet. Large ensembles in which a single section's part is suddenly divided (*divisi*) present a unique situation in which a portion of the ensemble space undergoes a kind of *mitotic division*, and the gestural affordances of the total ensemble are increased.

from the auditory scene—particularly when they unfold within the same region of ensemble space.

Generally, in the context of polystreaming, a given event point will belong to one and only one gesture. This conforms to the Gestalt principle of exclusive allocation, which asserts that a part can belong to one and only one whole at a given time. Bregman applies this concept to his conditions for auditory stream segregation, where he discusses multiple streams "competing" for a sound particular. Bregman states: "Whenever we observe a competition between streams for a sound, we are observing the principle of exclusive allocation in action. The belongingness of an element can shift from moment to moment, but at any one time it tends to be definite."²⁹ An element potentially shifting from one stream to another is tied to the concept of figure-ground reversal (or multistability). Similarity conditions often relegate a sound to one stream or another, but different factors of similarity may compete with one another. An instance of multiple gestures sharing an event point occurs when *point split* causes a gesture to "branch off" from another gesture. Imagine SG-[43210] unfolding in five-point ensemble space. When the gesture reaches point 2, SG-|234| occurs concomitant to the completion of the larger panning gesture. Thus, ensemble point 2 serves mutually as the spatio-temporal midpoint of SG-|43210| and as the initiation point of SG-|234|.³⁰ Concrete examples of this phenomenon shall be explored in Chapter 6.

There are three main varieties of polystreaming: consonant streaming, dissonant streaming, and contrary streaming. These categories are further partitioned into different subcategories.

²⁹ Bregman, 166.

³⁰ Conditions might favor the identification of two larger gestures unfolding simultaneously: SG-|43210| and SG-|43234|. In this case, the two gestures share partial open gesture SG-|432| in ensemble-space segment 2–4.

Consonant Streaming

Consonant streaming involves the coincident occurrence of two or more gestures of the same categorical kind *and* directionality—multiple gestures displaying a commensurate spatial shape. The gestures may occur at the same time in different segments of ensemble space or at overlapping times in the same spatial region (either the total ensemble space or a segment of ensemble space). I identify four subcategories of consonant streaming: canonic, parallel, echo, and uneven streaming. Any case of consonant streaming will feature the concurrence of two or more gestures from the same category and of identical directional profile—what distinguishes the four subcategories of consonant streaming from one another is the temporal relationship between the multiple gestures.

Canonic streaming occurs when the same type of gesture occurs more than once in overlapping temporal succession within the same region of ensemble space. Imagine an eight-point ensemble space with each instrument of the octet performing a discrete part. Now, imagine that SG-[01234567] (perfect L-R PAN) begins to unfold. Before this gesture has reached terminal point 7, another perfect L-R PAN is initiated. The initial gesture is the **lead stream**, and the second gesture is termed the **tracking stream**, since it "follows" the lead stream in both time and space.³¹ If multiple tracking streams are present, they may be labeled "tracking stream 1," "tracking stream 2," etc., depending upon their temporal ordering.

Figure 4-2 presents the opening five measures of György Ligeti's *Wenn aus der Ferne,* No. 2 from *Three Fantasies* (1983). Each voice of this work's SATB scoring contains four

 $^{^{31}}$ In the analysis of melodic canons, the leading voice is called the *dux* and the imitating voice the *comes*. In the continued effort to distinguish spatial activity from that of other musical parameters, I shall not apply these concepts to account for gestural relationships. However, there is a profound degree of functional equivalence between the terms *dux* and *lead stream* as well as between *comes* and *tracking stream*. It would not be surprising to find these related concepts commingled in a work of music—as when a melodic canon is spatialized such that the *comes* functions as a *tracking stream* following the *dux-lead stream*.

distinct parts, for a total of sixteen parts. In Figure 4-2, only the four soprano and four alto parts are shown. Assuming the ensemble is arranged in the following order (from left-to-right): soprano, alto, tenor, basses, and assuming that parts 1–4 also span from left to right in each voice-group, a sixteen-point ensemble space emerges that is a "neat" translation of the score space (the top-to-bottom ordering of staves corresponds to the left-to-right ordering of ensemble points).



Figure 4-2: Ligeti, Wenn aus der Ferne, mm. 1-5

The left half of the ensemble (the soprano and alto sections) opens with a melodic canon that is fused with canonic streaming. Lead stream SG-[01234567] (indicated with a solid arrow) commences with the first four syllables of text, "Wenn aus der Fer-," set in each voice to a melodic [0237] tetrachord. When the lead stream reaches point 3 (m. 2), the syllable "-ne" (completing the word "Ferne") initiates tracking stream 1 at point 0 (tracking streams are

indicated with a dashed arrow).³² This tracking stream unfolds at the same rate as the lead stream. When the lead stream terminates in m. 3, tracking stream 1 is at point 4. Prior to the conclusion of the lead stream, however, two more tracking streams commence: tracking stream 2 on the syllable "da" and tracking stream 3 on the syllable "wir" (both beginning in m. 3). Thus, prior to the lead stream's termination, three tracking streams enter the spatial complex in pursuit of the lead stream. The sonic uniformity of pitch and text-syllable (at different points in the ensemble) individuates the separate streams. Accordingly, each "gestural stream" is likewise an auditory stream.

Canonic streaming may feature streams of different rates as long as each stream follows the exact same sequence of ensemble points *or* an embellished (i.e., abbreviated or elaborated) version of the same sequence in the same region of ensemble space. In short, all the gestures engaged in canonic streaming must be of identical gestural kind. However, lead stream SG-|01234567| might be followed by tracking stream SG-|023467|—an abbreviated (via two deletions) form of the original lead stream. Since SG-|023467| begins before SG-|01234567| has ended, and since both streams are strong open gestures, this scenario counts as an instance of canonic streaming. A **tracking stream overtake** occurs when a tracking stream moves faster than a lead stream (or a preceding tracking stream) and effectively "catches-up" to it potentially terminating at the same time (and point) as the lead stream or perhaps surpassing it to conclude prior to the cessation of the initial lead stream.

Parallel streaming occurs when the exact same gesture takes place in two or more regions of ensemble space at the same time. We can easily envision an octet divided

 $^{^{32}}$ In this excerpt, the melodic *dux* occurs at ensemble point 0. Seven *comes* appear successively (and statically) at points 2–7. Both the dux and its subsequent comes are all spatialized relative to one another, but the individual notes forming the dux (as well as the comes) are not spatialized. Therefore, a melodic canon co-occurs with canonic polystreaming, but the two processes are not perfectly amalgamated.

antiphonally into two quartets—the leftward quartet identified by integer values 0, 1, 2, and 3 and the rightward quartet by integers 4, 5, 6, and 7. Each half of the octet consists of a four-point segment of ensemble space. If both segment spaces contain identical gestures unfolding at the same time, parallel streaming occurs. Imagine that weak open gestures SG-|0132| and SG-|4576|occur in perfect point-for-point simultaneity. SG-|(0+4)(1+5)(3+7)(2+6)| is a convenient (if somewhat convoluted) means of notating the temporal relationship between these two commensurate gestures. The SG-label shows that nonadjacent ensemble points are fused by the synchronicity of events occurring at those points. That this gestural complex is a case of parallel streaming is revealed by the fact that the difference between the two integers contained in each of the four parentheses is the same number (4, in this case).³³ In essence, parallel streaming entails two or more identical gestures (with a uniform event cardinality between the gestures) unfolding at the same time in different (and non-overlapping) segments of the total ensemble space.

Echo streaming is similar to parallel streaming with the exception that the two identical gestures occurring in separate regions of ensemble space are temporally staggered in their joint presentation. If SG-|0132| and SG-|4576| are not exactly synchronized but happen close to one another in time in an overlapping fashion—whereby SG-|4576| begins after the initiation of (and prior to the completion of) SG-|0132| (perhaps in a complex gesture such as SG-|04153726|)— we may refer to this gestural interactivity as an instance of echo streaming, and the latter-occurring gesture is the **echo stream**. As described in Chapter 2, a sonic echo is a spatio-temporal distortion of some primary sound. In view of this, an echo stream is effectively a spatially and temporally displaced version of an earlier gesture. If the two gestures were to occur in immediate succession (SG-|4576| beginning directly after the completion of SG-|0132|), this

 $^{^{33}}$ (4 - 0 = 4), (5 - 1 = 4), (7 - 3 = 4), and (6 - 2 = 4).

would simply be a compound gesture with subgestural components related via *nudging*—not an instance of polystreaming. Echo streaming allows for abbreviation or elaboration of the echo stream: SG-|0132| and SG-|476|, unfolding in overlapping succession, would qualify as an instance of echo streaming, since both gestures are weak closed gestures.

Uneven consonant streaming occurs when two like gestures unfold simultaneously but at completely different (i.e., "uneven") rates in separate regions of ensemble space. The rate of each gesture's materialization may be either fixed (constant) or variable (subject to intragestural temporal processes), and the onset and cessation of the polystreaming may be staggered (one gesture may commence after the other gesture has begun and/or terminate before or after the other gesture has ended). Figure 3-12 (see Chapter 3) depicts two adjacent-point oscillation gestures occurring at the same time in the fourth movement of Schoenberg's *String Quartet No. 3*. Since the gestures are of like kind but unfold at separate rates, they constitute a case of uneven consonant streaming.

Some instances of uneven consonant streaming may be more or less "consonant" than others. For example, the two gestures in Figure 3-12 partake in a proportional temporal relationship: SG-|0101010101010101010| unfolds at exactly twice the rate of SG-|23232323|. This 2:1 temporal ratio is comparable to the 2:1 frequency ratio between two pitches that produces the interval of an octave—a perfect consonance. A 7:5 temporal relationship would be a less consonant instance of uneven streaming (the 7:5 ratio yields the dissonant interval of a tritone in the domain of pitch). Nevertheless, in such a case, the polystreaming would still be considered consonant by virtue of the fact that the two gestures are identical (if not temporally proportional).

205

Dissonant Streaming

Dissonant streaming occurs when two or more categorically different spatial gestures unfold either at the same time within different segments of the ensemble space or in overlapping succession within the total ensemble space. The "spatial dissonance" results from the categorical heterogeneity between the concomitant gestures. For example, SG-|3210| and SG-|76457| are different because the former gesture is a partial open gesture (in ensemble-space segment 0–3) and the latter is a strong closed gesture (in ensemble-space segment 4–7); therefore, the cooccurrence of these gestures is a case of dissonant streaming. On the other hand, consider SG-|3201| and SG-|745|. Despite a discrepancy in event cardinality, these two gestures are members of the same gestural family—SG-|745| is a nudged (+4) and abbreviated (via deletion) form of weak closed gesture SG-|3201|. The coincident presence of these two gestures would be a case of echo streaming, a type of consonant streaming described above.

Contrary Streaming

Spatial gestures related by retrograde and/or inversion are of the same kind, so their simultaneous presentation cannot be a case of dissonant streaming. The streaming likewise cannot be an instance of consonant streaming, since the gestures have differing directionalities. Categorically related gestures with opposing directionalities often yield interesting polystream composites. The two subcategories of contrary streaming are mirrored streaming and crossed streaming.

Mirrored streaming occurs when two gestures unfolding in separate regions of the ensemble space are related by either retrograde, inversion, or retrograde-inversion. There are two types of mirrored streaming: convergent and divergent mirrored streaming. **Convergent**

206

mirrored streaming is when two gestures (each in a separate region) approach one another in ensemble space. The gestures may or may not merge (or, "fuse") at a single point. Imagine SG-|0123| and SG-|7654| occurring at the same time, which, if performed in perfect synchronicity, could be notated: SG-(0+7)(1+6)(2+5)(3+4). The evidence that this SG-label indicates a case of mirrored streaming is provided by the fact that the sum of the two integers in each parenthetical enclosure within the label is the same (7, in this case).³⁴ We know that the mirrored streaming is convergent because the difference between the two integers in each parenthesis decreases by the same amount (2) from left to right (the direction of time) within the label.³⁵ However, the two gestures need not be perfectly synchronized to count as a case of convergent mirrored streaming. Furthermore, the component gestures of the polystreaming complex need not be of the same event cardinality. SG-[0123] may be presented alongside SG-[764], which is the nudged and abbreviated (via deletion) retrograde (or the abbreviated type-1 inversion) of SG-[0123]. Such gestural activity would qualify as an instance of convergent mirrored streaming.

Divergent mirrored streaming prevails when two distinct gestures move away from one another, as with the co-occurrence of SG-|3210| and SG-|4567|. In this example, each gesture begins from a somewhat centralized point within the ensemble space and moves outward toward the spatial extremes. (The gestures could likewise begin from a single point that "splits" to yield two separate streams.) SG-|(3+4)(2+5)(1+6)(0+7)| is the notation for the parallel materialization of SG-[3210] and SG-[4567].³⁶ Again, the two gestures need not be synchronized point for point. The "mirrored" presentation of many closed gestures (including sweeping gestures) may result in the intermingling of convergent and divergent mirrored streaming. For example, SG-|0132| and

³⁴ (0 + 7 = 7), (1 + 6 = 7), (2 + 5 = 7), and (3 + 4 = 7). ³⁵ (7 - 0 = 7), (6 - 1 = 5), (5 - 2 = 3), and (4 - 3 = 1).

 $^{^{36}}$ Within this SG-label, mirrored streaming is again evinced by the uniform sum (7) of each parenthetical pair of integers. Divergent mirrored streaming is indicated by the fact that the difference between each integer pair increases (by 2 in this case) from left to right within the label.

SG-|7645| (the nudged inversion of SG-|0132|) are both weak closed gestures. When presented concurrently, they begin as converging streams and conclude as diverging streams.

An instance of divergent mirrored streaming occurs in Ellen Taaffe Zwilich's *Double Quartet for Strings* (1984). Figure 4-3 presents the composer's "suggested seating arrangement."

Violoncello
Viola
Violin 2
Violin 1

Figure 4-3: Suggested seating arrangement for Ellen Taaffe Zwilich's *Double Quartet* ³⁷

We may observe from this diagram that the quartet to the left of the ensemble space is seated in the common twentieth-century layout for string quartet heretofore assumed for all musical examples taken from the string-quartet literature. The quartet on the right side of the space is essentially the mirror image of the quartet on the left. Figure 4-4 presents an integer-notational diagram of the ensemble space.



Figure 4-4: Integer diagram of the Double Quartet ensemble space

Figure 4-5 presents mm. 46–50 of the quartet's first movement. The top four staves of the system are performed by the quartet on the left, and the bottom four staves by the quartet on the right. In this excerpt, the melodic and rhythmic activity of each discrete quartet is virtually

³⁷ Ellen Taaffe Zwilich, *Double Quartet for Strings*, composed 1984 (Bryn Mawr: Merion Music, Inc., 1985), 2.

identical, with the exception of a few differing notes between the two quartets that create a heterophonic musical texture: the two cellos play different double stops in m. 48–49, and the notes of the two first violins (points 0 and 7) in mm. 49–50 are a minor third apart. Despite these slight differences, the homogeny of parts between the two quartets results in a clear case of divergent mirrored streaming.



Figure 4-5: Zwilich, Double Quartet, Mvt. 1, mm. 46-50

In fact, two cases of dissonant polystreaming occur *within* each quartet—all four of which are subsumed by the larger-scale (and more salient) activity of divergent mirrored streaming *between* the two quartets. Consider only the top four staves of figure 4-5. The cello (point 3) begins the following gesture: SG-|32(1+3)1(0+2)|. Embedded in this elaborate gesture is SG-|3210| (R-L PAN). However, after the initial arrival at point 2, the spatial activity *splits* into two separate

gestures—one progressing to point 1 and another falling back to point 3.³⁸ Thus, SG-|321| (a partial open gesture) and SG-|323| (a strong adjacent-point oscillation gesture) occur simultaneously—accounting for the first instance of dissonant streaming. This activity is repeated as event point 1 splits such that one stream continues to point 0 while another regresses to point 2. Thus, SG-|3210| (strong open gesture) and SG-|3212| (weak closed gesture) occur in point-for-point simultaneity.

SG-|3210| is the large-scale, fundamental motion of this gestural composite (SG-|32(1+3)1(0+2)|); its component events feature eighth-note rhythmic activity at each point (with the exception of point 0), and its salience is only somewhat diminished by the two occasions of dissonant streaming. Since almost identical activity occurs in the rightward quartet of the ensemble, the dominant gesture at points 4–7 is SG-|4567|, a L-R PAN that is likewise minimally screened by two smaller instances of dissonant streaming.³⁹ The contemporaneous existence of SG-|3210| and SG-|4567| is portrayed by the gestural label SG-|(3+4)(2+5)(1+6)(0+7)|. This is a case of divergent mirrored streaming from the innermost ensemble points (3 and 4) to the extremes of the ensemble space (points 0 and 7).

Crossed streaming occurs within the whole of the ensemble space (or within a single segment thereof). The simplest form of crossed streaming involves the juxtaposition of a L-R PAN and its retrograde (or inversion), R-L PAN. Imagine SG-|01234567| and SG-|76543210|

³⁸ In the context of polystreaming, *point split* refers to activity at a single point that progresses outward in either direction, resulting in two distinct (yet temporally overlapping) spatial gestures. For example, point 2 may "split" into points 1 and 3. *Point fusion*—again, in the context of polystreaming—refers to simultaneous activity at two disjunct points that eventually "collapses" into a single, interior point—generally reflecting the "merging" of two distinct spatial gestures. For instance, points 1 and 3 may "fuse" into point 2. In this respect, the terms "fuse" and "split" are borrowed from Clifton Callender, who applies these terms to account for parsimonious voice leading. See Clifton Callender, "Voice-leading Parsimony in the Music of Alexander Scriabin," *Journal of Music Theory* 42.2 (1998): 219-233.

³⁹ The gesture unfolding in ensemble-space segment 4–7 is SG-|45(6+4)6(7+5)|. From this notation, it is easy to discern SG-|456| (partial open gesture), SG-|454| (strong oscillation gesture), SG-|4567| (strong open gesture), and SG-|4565| (weak closed gesture)—all of which are analogs to the gestures occurring at points 0–3.

beginning at the same time and unfolding at the same approximate pace. The streams would effectively "cross" at their respective midpoints, and each gesture would end where the other had begun. In analyzing such spatial activity as an instance of crossed streaming—as with the assessment of any type of polystreaming—we would have to look to other musical parameters to ensure the validity of such a reading. For instance, this would prove to be a well-founded interpretation if there were some homogeny of pitch, range, articulation, etc. among each event (at each point) within each gesture. Figure 4-6 presents a case of **ambiguous polystreaming**, in which two or more feasible explanations of multigestural activity co-exist. Imagine this two-measure example performed by an ensemble of six identical instruments, such as a flute sextet. The spatial equivocation of this passage results from the fact that the same pitch occurs at each point without any differentiation of dynamic, timbre, rhythmic configuration, or articulation. Is this a case of contrary streaming or mirrored (convergent + divergent) streaming? With no extra-spatial cues to disambiguate this conundrum, the passage cannot be analyzed definitively.



Figure 4-6: Ambiguous streaming

Figure 4-7 presents the same passage with added cues that greatly facilitate the perceptual differentiation of two distinct streams. SG-|012345| (L-R PAN) is an integrated stream due to the qualitative identity of musical activity—the persistent pitch A, mezzo piano dynamic level, tenuto articulation, and quarter-note value—at each event point. SG-|543210| (R-L PAN) is likewise a unified stream made distinguishable from SG-|012345| through the constant pitch E, forte dynamic level, staccato articulation, and eighth-note (off the beat) value at every event point. The two related gestures "cross" at the midpoint of the ensemble space (points 2 and 3).



Figure 4-7: Crossed streaming

Figure 4-8 presents another disambiguation of Figure 4-6. Through the redistribution of extra-spatial streaming cues, the passage now exhibits mirrored streaming. The first measure of Figure 4-8 is identical to the first measure of Figure 4-7. It is the activity of the second measure that alters the nature of the polystreaming. The two distinct streams of Figure 4-8 are partial sweeping gestures SG-[01210] and SG-[54345]. These gestures, which are related by type-1

inversion, begin their simultaneous presentation as convergent-mirrored streams (m. 1) and conclude as divergent-mirrored streams.



Figure 4-8: Mirrored streaming (convergent + divergent)

Although the component event points of a spatial gesture need not display any degree of extra-spatial homology whatsoever for their temporal succession to yield a spatial gesture, the above examples offer a prime example of how ambiguous and conflicting readings might arise—particularly in some cases of polystreaming—if other musical factors potentially influencing grouping on the basis of similarity are not taken into consideration.

Apparent Polystreaming

In Chapter 3, the illusion of "apparent motion" (the Phi phenomenon) was introduced to certify the point shift as the smallest unit of motion within a spatial gesture. Bregman draws parallels between auditory stream segregation and apparent motion in vision. When apparent

motion is induced visually by two lights blinking in alternation, the strength of the impression of movement is contingent upon three interlinked factors (as formulated by A. Körte in 1915): the spatial separation of the lamps (i.e., the index of distance between adjacent lamps), the brightness of the lights, and the duration of the flash at each lamp.⁴⁰ Körte's third law essentially states that as the spatial separation of the lights increases, the alternations of flashes must be slowed down in order for the phi phenomenon to occur.⁴¹ If widely spaced lights alternate too rapidly, no motion is perceived. Bregman describes a similar experiment in which six lights (numbered 1–6 from left to right and with a large spatial gap between lights 3 and 4) alternate in the following sequence: 142536. At very slow speeds, Bregman claims that no motion is perceived—only the sequential blinking of lights is observable. At somewhat higher speeds, irregular left-and-right motion (spanning the large spatial gulf) emerges from the pattern. At still higher speeds, the complex pattern "splits" into two distinct streams: 1-2-3 and 4-5-6. Bregman explains the phenomenon:

We assume that potential motions between successive and non-successive flashes are competing with one another for dominance...as we speed up the sequence there is an increased tendency for shorter movements to be favored by Körte's law so that the longer between-triplet motions are suppressed in favor of the stronger within-triplet motions.⁴²

In short, as the rate of flashing increases, spatial proximity overrides temporal proximity as the primary factor influencing perceptual grouping. Bregman and Campbell apply Körte's third law to auditory stream segregation.⁴³ Using a repeating cycle of six tones with contour <031425>, they show that, at high speeds, the sequence segments into two distinct <012> contours in

⁴⁰ A. Körte, "Kinomatoscopishe Untersuchungen," *Zeitschrift für Psychologie der Sinnesorgane* 72 (1915): 193–296.

⁴¹ Bregman, 21.

⁴² Ibid., 22.

⁴³ A. S. Bregman and J. Campbell, "Primary Auditory Stream Segregation and Perception of Order in Rapid Sequences of Tones," *Journal of Experimental Psychology* 89 (1971): 244–249. This article is summarized in Bregman, *Auditory Scene Analysis*, 17–21.

separate regions of pitch space.⁴⁴ They show through further experimentation that stream segregation is stronger both when the frequency separation between low (stream 1-2-3) and high (stream 4-5-6) tones is greater and when speeds are faster: "As the high and low groups are moved further away from one another in frequency, the within-group attractions will become much stronger than the between-group attractions. Speeding the sequence up simply has the effect of moving things closer together on the time dimension."⁴⁵

Although Bregman and Campbell apply Körte's third law to the segregation of melodic streams in pitch space, their findings might also be applicable to gestural segmentation in ensemble space. In a six-point ensemble space, weak open gesture SG-|031425| would be quite noticeable at a relatively slow rate of unfolding. However, at faster rates, the listener might perceive two partial open gestures occurring in separate ensemble-space segments: SG-|012| and SG-|345|. As the rate of the gesture's presentation increases, the nonconsecutive succession of spatially proximate events takes perceptual precedence over the grouping of successive, temporally proximate events. Since these two gestures are interwoven, their segregation constitutes a case of **apparent polystreaming** that is dependent upon a rapid rate of unfolding. In short, perceptual grouping on the basis of either temporal proximity or spatial proximity (both subsumed under the Gestalt Law of Proximity) is dependent upon the rate at which gestural events occur.

The above description of apparent polystreaming assumes that the discrete events composing SG-|031425| do not afford any other groupings on the basis of sonic similarity. Consider, however, the gesture presented in Figure 4-9. In this example, two distinct gestures would be identifiable irrespective of the rate at which the gesture unfolds due to the similarity (in

⁴⁴ Bregman and Campbell label the sequence "142536" with 1 as the lowest and 6 as the highest frequency. They identify the two emergent streams as 1-2-3 and 4-5-6.

⁴⁵ Bregman, 20.

the realms of pitch and articulation) of events at ensemble points 0–2 as well as at points 3–5. Although complex gesture SG-|031425| occurs by virtue of grouping on the sole basis of temporal proximity (the direct successiveness of each event point), the Law of Similarity enters and affects the individuation of two distinct and intertwined gestures: SG-|012| and SG-|345|. At faster rates of unfolding, the two gestures would likely seem less interwoven and more simultaneous, and stream segregation would be further enhanced at faster speeds. Through extra-spatial sonic qualities, including the gesture's rate of unfolding, a single weak open gesture may be experienced as twin partial strong open gestures in a bifurcated ensemble space, yielding a salient instance of consonant echo streaming. SG-|345|, in ensemble-space segment 3–5, is the echo stream to SG-|012|, since SG-|012| begins and ends in ensemble-space segment 0–2 prior to the commencement and termination of SG-|345|. These two gestures are related by nudging, and their coincident occurrence is arguably stronger than the single instance of SG-|031425|.



Figure 4-9: Complex Gesture SG-|031425| = Echo streaming of SG-|012| + SG-|345|

In their experimentation with the segregation of streams on the basis of frequency,

Bregman and Campbell observed: "It was impossible for the listeners to focus their attention on both streams at the same time. When they focused on one of the streams, the other was heard as a vague background."⁴⁶ From this observation, we may make one final inference regarding polystreaming: although the simultaneous presence of multiple spatial gestures may be easily discernible (i.e., the fact that multiple gestures are co-occurring may be plainly evident), only one of the gestures within the polystreaming complex will be the perceptual focus (or, "figure") at any given moment, with the other gesture(s) relegated to the "ground." The listener's perceptual act of figure-ground reversal on the gestural multiplicity—which may or may not be motivated by a gesture's composite sonic attributes—provides a means of assessing the nature of the relations between concomitant gestures in the real time of musical performance.

4.3 Gestures in Multi-Dimensional Ensemble Spaces

The analytic methodology outlined in these pages is focused primarily upon the onedimensional ensemble space. Such an ensemble space involves two or more performers arranged linearly within the performance space. The "line" formed by the ensemble members is actually a sequential arrangement of discrete points. Since this line will frequently be somewhat concave (from the listener's perspective), I have advocated discounting the slight semi-circular arc (i.e., "normalizing" the ensemble space) to account for horizontal spatial motion.⁴⁷ Many ensembles, however, assume an overt and nontrivial two-dimensional (2D) or three-dimensional (3D) spatial form. Such ensemble-space designs are generally specified by composers but often result from large performance forces that need to be somewhat consolidated in the performance space. A

⁴⁶ Bregman, 17–18.

⁴⁷ The slight arc results from the performers' need to see and hear one another to ensure ensemble "togetherness."

symphony orchestra, for example, is a two-dimensional construct: depth planes are added to the left/right horizontal span of the orchestra.⁴⁸ I shall briefly explore the kinds of gestures that might be encountered in such multidimensional ensembles and offer some general tools for describing multidimensional spatial activity.

2-D Ensemble Spaces with Multiple Depth Planes

Figure 4-10 shows the diagram for the "recommended layout" of ensemble forces in *Empreintes* by Iannis Xenakis.

Figure 4-10: Xenakis, *Empreintes*, recommended layout for the ensemble⁴⁹

As evident from the diagram (provided in the score by the composer), the ensemble space is partitioned into three distinct depth planes. I shall use capitalized letters to label the different planes, with "A" designating the plane closest to the public, "B" representing the interior plane, and "C" denoting the plane farthest from the public. Each plane is demarcated by instrumental

⁴⁸ Many large ensembles, including the symphony orchestra, typically feature multiple performers to a given part that are generally seated close together in the ensemble space, constituting an *area*. Since an area is essentially an inflated point, the spatio-gestural affordances of large ensembles might *not* be dramatically augmented by the addition of depth planes.

⁴⁹ Iannis Xenakis, *Empreintes*, composed 1975 (Paris: Editions Salabert, 1987).

timbre: plane A comprises strings, plane B consists of woodwinds, and plane C is composed of brass.⁵⁰ Each plane consists of discrete points that may be identified with integers (0 for the leftmost instrument, n - 0 for the rightmost instrument, etc.). Based on their organization, the component points of planes B and C constitute a single line. However, Plane A is itself partitioned into five sub-planes, each of which is differentiated by its type of instrumental occupant. If pertinent to analysis, lowercase letters (a–e) might be employed to distinguish these five sub-planes: A-a (first violin), A-b (second violin), A-c (viola), A-d (cello), and A-e (contrabass). We may refer to a **plane shift** as the smallest unit of motion between two depth planes. A plane shift is conceptually identical to a point shift—the distinction is the directional orientation of the shift: a point shift moves left-to-right (or vice versa) in the horizontal realm, while a plane shift moves back-to-front (or front-to-back) in the horizontal realm.

While the potential exists for each plane (or sub-plane) to exhibit the left/right horizontal gestures identified in Chapter 3 and the multi-gestural relations explicated above, the possibility of front/back horizontal motion likewise exists. During a performance of *Empreintes*, we might encounter advancing (back-to-front) or receding (front-to-back) gestures. An *advancing (B-F) gesture* might begin at some point within plane C, progress through plane B, and terminate at some point within a sub-plane of plane A. A *receding (F-B) gesture* would unfold in the opposite sequence of planar activity. A *diagonal gesture* might traverse from the leftmost point in the farthest plane to the rightmost point in the nearest plane—cutting through the approximate

⁵⁰ The arrangement of parts in the score does not translate directly to the ensemble space. Plane B (the woodwind family) is positioned at the top of the stave system, plane C (the brass family) is situated in the middle, and plane A (the string family) is located at the bottom of the score. This conventional scoring could be modified to better reflect the spatial design and activity of the music: if the woodwinds exchanged positioning with the brass (in the score), the planes would be arranged as such (from top to bottom): C, B, A. Alternatively, if the strings were relocated to the top of the score, the planes would be arranged: A, B, C. Either of these rearrangements would create a visual correlation between the arrangements of the score and the ensemble space. Such a correspondence would serve to facilitate the identification of spatial activity.

center of the ensemble space.⁵¹ Due to the two-dimensionality of the ensemble space, spatial motion is not restricted to straight lines: we might encounter angular, curvilinear, polygonal, or circular gestures. *Angular gestures* (which would comprise two linked linear gestures) could be qualified and compared based on the size (approximate degree) of the angle (right, acute, obtuse).⁵² *Polygonal* and *circular gestures* might be open/incomplete (beginning and ending at different points) or closed/complete (terminating at the same point of commencement) figures and may materialize in either a clockwise or counterclockwise fashion. *Curvilinear gestures* might be deemed incomplete circular gestures. (The interpretive implications of these gestural types, along with a miniature analysis of *Empreintes*, are found in Chapter 5.) Such gestures may be nudged from one plane to another, rotated from a left/right orientation to a back/front positioning (or vice versa), inverted across a number of axes dividing the ensemble, and otherwise altered by the procedures outlined in section 4.1.⁵³

2-D Ensemble Spaces with Multiple Vertical Levels

A relatively rare ensemble-space structure features more than one vertical level. Musical works featuring such spatial designs require the proper venue (performance space) for their realization. A common experience with verticality in an ensemble space occurs when a choir is placed in an elevated choir loft or upon risers, which is frequently the case in performances of

⁵¹ We may identify the following four directionalities for a prospective diagonal gesture, two of which are *advancing* and two that are *receding* in nature: back-left to front-right (B/L \rightarrow F/R), front-right to back-left (F/R \rightarrow B/L), back-right to front-left (B/R \rightarrow F/L), and front-left to back-right (F/L \rightarrow B/R). Any two of these diagonal gestures are related to one another via some transformational operation—either through retrogression or inversion across an axis that divides the ensemble space in half (either from front to back or from left to right).

⁵² The following notation would indicate a right angle comprising an initial advancing (B-F) gesture and subsequent L-R PAN: $C(0) \rightarrow B(0) \rightarrow A-e(0) \rightarrow A-d(0) \rightarrow A-d(1) \rightarrow A-d(2) \rightarrow A-d(3)$. The integers in parentheses indicate the point at which an event occurs within a given plane (depth planes are represented by the capital letter immediately preceding the parenthetically enclosed integer).

⁵³ For an in-depth assessment of two-dimensional spatial activity, and an alternate, more computational approach to spatial analysis, see Paul Miller, "The Analysis of Spatial Music in Works by Karlheinz Stockhausen."

works such as Beethoven's *Ninth Symphony* (finale), Mahler's *Eighth Symphony*, and any number of Requiem-Masses (including those of Mozart, Berlioz, and Verdi). However, many composers in the twentieth century exploit height as a spatio-musical parameter. One such composer is Henry Brant, whose aesthetic of spatialization was outlined in Chapter 1.

Brant notes the pleasing effect that results from the coordination of high pitches with high elevations (achieved by placing high-pitched instruments at elevated vertical levels) and low pitches with a low physical orientation:

There is no mistaking the compelling naturalness of effect when high pitches originate in a high location...or low pitches from a low position. However, the actual *pitch* need not be high or low, as the case may be; if the *register* in which the instrument plays is *proportionately* acute or deep, this will substitute very well for absolute height or depth in pitch.⁵⁴

In his own musical works exploiting such a spatial scheme, Brant essentially constructs a concrete model of metaphoric pitch space, with high-sounding instruments physically oriented "up" and low-sounding instruments situated "down" in actual space. Such locational concordance between concurrent activity in physical space and in pitch space—the metaphoric derivative of actual space—produces an appreciable, if commonsensical, musical effect.

The various vertical levels of an ensemble space may be indicated with uppercase Roman numerals, with "I" representing the lowest level and the numerals increasing sequentially with each successive level. In a space with three vertical levels, I would represent the ground level, II the middle level, and III the highest level. In an ensemble space with more than one vertical level and only one depth plane, we might expect to encounter rising (down-up) and falling (up-down) gestures. A *rising (D-U) gesture* might begin at level I and advance to level III—passing through level II along the way, while a *falling (U-D) gesture* might descend from the uppermost level to the ground level. Assuming multiple points (constituting a horizontal line) at each level,

⁵⁴ Brant, "Space as an Essential Aspect," 232. Emphasis Original.

rising and falling gestures may be either perfectly vertical or slanted to varying degrees. In fact, we might refer to a gesture beginning at the rightmost point of level I and concluding at the leftmost point of level III as a *slanted rising gesture*. As with two-dimensional ensemble spaces that feature multiple depth planes, in spaces with multiple vertical levels we may encounter angular, curvilinear, polygonal, and circular gestures—the distinction is that the gestures would be oriented vertically as opposed to horizontally.⁵⁵

3-D Spaces with Multiple Depth Planes and Vertical Levels

In the performance notes to Desert Forests: Spatial Panoramas for Separated Orchestral

Groups, Henry Brant explicitly details the manner in which the groups should be arranged

relative to one another within the performance space:

(a) The 9 high woodwinds and their conductor are to be placed in the *top* balcony at the *back* of the hall; (b) the 3 trumpets are to be stationed in a box or balcony in the *middle* of one *side* of the hall, the 3 trombones *opposite* on the *other side*. The three players in each trio may be placed either close together, if in the same box, or separated with each player occupying 1 of 3 adjacent boxes...(c)...The instruments on stage need to be deployed in accordance with the following diagram [in order to] produce clearly identified spatial sonorities.⁵⁶

This work is designed for performance in a concert hall containing three distinct vertical levels:

two balcony levels in addition to the stage level. The instruments forming the first balcony level

(vertical level II) must be situated on the sides of the concert hall. As evident from the diagram

⁵⁵ To highlight the similarity of gestural affordances between ensemble spaces with multiple depth planes and those with multiple vertical levels, imagine how the former might transform into the latter (or vice versa): the back depth plane (C) of a three-plane horizontal space could function as a hinge, allowing the front portion of the space to "lift" 90 degrees until it is perfectly vertical. Thus, depth plane C becomes vertical level I, plane B becomes level II, and plane A becomes level III. Gestures previously occurring in the horizontal space are "translated" to the new vertical space: a receding gesture becomes a falling gesture, and a diagonal advancing gesture becomes a slanted rising gesture. This mental exercise is designed solely to illustrate the homologous gestural shapes afforded by both horizontal and vertical two-dimensional ensemble spaces. Due to the fundamental differences in spatial orientation, however, the interpretation of like gestural shapes may vary from one space to the other, as shall be discussed in Chapter 5.

⁵⁶ Henry Brant, *Desert Forests: Spatial Panoramas for Separated Orchestral Groups*, composed 1983 (New York: Carl Fischer, 2000), 2. Emphasis mine.

supplied in Figure 4-11, the instruments on the stage (vertical level I) are arranged in twodimensional fashion. Vertical level II is likewise a two-dimensional horizontal plane (with two distinct "lines") effectively suspended laterally above the audience. Presumably, the nine high woodwinds constituting vertical level III in the upper balcony at the back of the hall are configured in a single line, although Brant does not specify this arrangement. The following depth planes and vertical levels are conflated: level III/plane A; level II/plane B; level I/plane C. Group III/A is located above and behind the audience, II/B directly above and parallel to the audience, and I/C in front of (and on the same approximate level as) the audience. Although depth planes B and C are themselves horizontal planes (each with distinct sub-planes), a threedimensional diagonal plane extends upward and backward from I/C, through II/B, to III/A. Not only are two-dimensional gestures possible within the planes on levels I and II, but threedimensional gestures are likewise possible from level to level.



Figure 4-11: Brant, *Desert Forests*, seating plan for onstage instruments⁵⁷

⁵⁷ Ibid., 3.

We can envision ensemble spaces shaped like pyramids, cubes, spheres, or any other three-dimensional figure. The audience may abide within or outside of these designs. Stockhausen, speaking specifically to the geometric designs forged by loudspeaker placement in several of his electronic compositions, describes his eight-track work *Oktophonie* (1990–1991) as consisting of "a cube-like arrangement of eight loudspeaker groups with all possible vertical and diagonal movements of sound" and his *Dodekphonie* (1992–1994) as comprising "twelve spatial sources of pyramidic construction."⁵⁸ Such inventive ensemble-space constructs are certainly the exception to the norm (especially in purely acoustic environments).

In a space such as a cube, we might describe a 3-D gesture as such: $I/A(3) \rightarrow II/A(2) \rightarrow$ III/B(1) \rightarrow III/C(0). This hypothetical label depicts a slanted, quasi-angular gesture that rises from level I to III, recedes from plane A to C (by way of plane B), and shifts from right to left (component SG-|3210|) within the cube. Speaking to the gestural alterations unique to a threedimensional ensemble space, entire gestures may be translocated from a vertical (within a single depth plane) to a horizontal (occupying a single vertical level) orientation, and vice versa. Such an alteration might be considered an "inter-dimensional nudging."

The purpose of this section has been to explore the gestural potentials of multidimensional ensemble spaces to a limited extent. These types of ensemble spaces are relatively rare but are important to discussions of deliberately spatialized music of the twentieth century. This section is nascent and deserving of extension—I have attempted to provide the elementary tools for confronting and analyzing spatial gestures within such spaces.

⁵⁸ Karlheinz Stockhausen, "Five Revolutions Since 1950," in *Contemporary Composers on Contemporary Music,* expanded edition, ed. Elliot Schwartz and Barney Childs (New York: Da Capo Press, 1998), 383.

4.4 Spatial Gestures in Electronic Music

Spatial gestures within the electronic medium differ fundamentally from those that unfold within an acoustic space occupied by multiple live performers. Furthermore, some types of gestures are essentially exclusive to electronic music and are all but impossible in acoustic works.⁵⁹ Many electronic compositions are *stereophonic*, meaning that a stereo field is placed in front of the listener (typically via two physically separated loudspeakers). In the electronic stereo field, a sound may be "panned"—in a continuous, analog fashion—from one speaker to the other. The resultant illusion of a mobile sound source is rendered by the psychoacoustic phenomenon of summing localization, as discussed in Chapter 2. Herein lies a critical distinction between spatial gestures in the acoustic realm and those within the electronic medium: in electronic works, sounds may be panned in a continuous motion from one speaker (as a sound source) to another.⁶⁰ When a L-R PAN occurs in an acoustic ensemble space, such as that of a string quartet, the "whole" gesture emerges through the summation of four spatially discrete "parts" (sonic events) occurring consecutively over x amount of time.⁶¹ To highlight the qualitative and perceptual distinctions between gestures in electronic music and acoustic music, Figure 4-12 depicts five different types of melodic motion in pitch space that are relatable to spatial gestures in physical space.

⁵⁹ Electro-acoustic music often features unique kinds of spatial gestures as well, given that one or more stationary performers might be embedded in an environment of shifting electronic sounds. Of course, the sounds produced by the live performer are often amplified (possibly after being subjected to electronic manipulation). Electro-acoustic music exists in such a wide variety of settings that its gestural affordances are difficult to generalize and codify outside the context of a specific work.

⁶⁰ Although the original sounds making up the music were either recorded or manufactured in a studio, the music is played over speakers (or headphones). Each speaker employed for playback is, therefore, a "sound source proper." The analog "movement" of a sound from one speaker to another is a sonic illusion—the actual sound sources do not move. O'Callaghan argues that hearing recorded sounds in general is an illusory and impoverished form of auditory perception because the perspectival content of the original sound(s) is compromised. See O'Callaghan, "Hearing Recorded Sounds" (Draft, Bates College, 2005), and O'Callaghan, *Sounds: A Philosophical Theory*, Chapter 10. ⁶¹ Through the application of overlapping dynamic envelopes, a sound may seem to pan from one ensemble point to another within an acoustic ensemble space. As I have argued, however, such conditions (which frequently betray the composer's deliberate act of spatialization) are not requisite to the experience of spatial gestures.



In all five examples of Figure 4-12, the motion through pitch space spans the interval of a perfect fifth in ascending fashion. Figure 4-12(a) shows a glissando beginning on G and "sliding" upward to D. Depending upon the instrument on which this effect is rendered, it may be performed as either a *portamento* (a continuous frequency slide between the two pitches) or a glissando—wherein the discrete diatonic or chromatic scale degrees falling between the two notated pitches are sounded in rapid succession.⁶² Instruments such as the violin and the human voice are capable of producing portamento; however, a perfectly "smooth" slide cannot be realized on instruments with fixed pitches, such as the piano or the guitar. The keys of the piano and the frets of the guitar cause a pseudo-portamento to be segmented into diatonic or chromatic steps-resulting in a glissando. Edward A. Lippman observes: "A continuous change of frequency appears to be the tonal analogue of visually or haptically perceived [continuous] motion, and in actual fact, a continuous change of frequency gives rise to a compelling experience of motion, more striking than the motion experienced in answer to a succession of tones."⁶³ This gliding motion through pitch space is analogous to continuous spatial pans in the electronic medium and is all but unachievable in an acoustic ensemble space (save for the presence of coordinated overlapping dynamic envelopes between two points).⁶⁴ Harley observes: "Spatial sound movement may be *discrete*, that is, it may proceed stepwise-if a

⁶² The purpose of a glissando is to imply a continuous motion between two pitches; therefore, a glissando is essentially an approximated portamento. In musical discourse, the term glissando is often employed to describe both the portamento and the glissando proper.

⁶³ Lippman, "Music and Space," 224.

⁶⁴ Bregman states: "Continuity in frequency change [such as a portamento] is merely an extreme case of frequency proximity" (Bregman, 179). Similarly, continuity of location change (such as a moving sound source or an analog pan) is an extreme case of spatial proximity.

musical phrase is presented successively in one ensemble of performers after another...sound movement in space may also assume a *continuous* form."⁶⁵ Harley goes on to describe the use of "dynamic shading and temporal overlapping of sounds" enacted by stationary groups of performers to produce continuous motion.⁶⁶ She notes that "the idea of using superimposed dynamic envelopes and temporal shifts to cause continuous changes in the apparent position of instrumental sounds was modeled on an electroacoustic technique, that is, on stereo sound projection."⁶⁷ In short (and as previously posited in Chapter 1), Harley is claiming that spatial trends in the tape and electronic mediums influenced the spatialization of subsequent acoustic works, particularly in the attempts to produce a continuous movement of sound. While continuous motion in either the electronic or acoustic environments is arguably the most salient form of sonic movement, the "discrete" movement of sound (likewise feasible in either setting) is equally cogent. Presence of the former type of motion (continuous) generally betrays some deliberation on the part of the composer, while that of the latter might arise more serendipitously under certain performance conditions; however, the discrete movement of sound is no less critical than continuous sonic motion as an eventuality impacting the musical experience and assuming analytic pertinence.

In his discussion of auditory streaming (the grouping of auditory stimuli into a coherent "stream"), Bregman remarks: "Our perceptual systems tend to break an array of sensory input into separate units only at points of *discontinuity* of properties."⁶⁸ Both a portamento in pitch space and an analog pan in the sonic field of an electronic work are easily perceived as a single event persisting over time because of the continuity of changing properties. In Bregman's

⁶⁵ Harley, "Spatiality of Sound and Stream Segregation," 151. Emphasis Original.

⁶⁶ Ibid.

⁶⁷ Ibid., 153.

⁶⁸ Bregman, 112. Emphasis mine.

words: "There are no inhomogeneities on the glide [between two frequencies] that lead one to segregate it into parts."⁶⁹ In the case of the portamento, the continuously changing property is frequency; with the analog pan, it is spatial location.⁷⁰ Owing to these seamless continuities, there are no discernible parts for our perceptual system to group together into a unified whole—just one modulating stimulus to constitute the stream. In other words, although steadily changing, the "whole" is indivisible. In Figures 4-12(b–e), discontinuities in pitch space exist in the form of intervals (seconds, thirds, and a fifth) articulated by fixed pitches. In an acoustic ensemble space, spatial discontinuities exist in the form of stationary ensemble points.⁷¹ The perception of both a melodic line and a spatial gesture is reliant upon grouping and streaming principles of organization to cluster discontinuous events into an integrated whole.⁷²

Figure 4-12(b) shows the ascent from G to B partitioned into semitonal motion. Were the

glissando of Figure 4-12(a) performed on a single guitar string, the frets of the instrument would

⁶⁹ Ibid. Bregman uses the term "glide" to describe a continuous shift between two frequencies (a "portamento" in musical terminology).

⁷⁰ Boulez refers to continuous spatial motion as a "space-glissando" (Boulez, *Boulez on Music Today*, 66). According to Harley, Boulez disdains the use of space-glissandi—considering them superficial and overused (Harley, "Space and Spatialization," 160).

⁷¹ Not all spatial gestures in the electronic medium need be analog. Gestures comprising discontinuous motion between points (as found in acoustic ensemble music) are common.

⁷² O'Callaghan offers a different view on what may constitute a single event. He states: "When a single instrument seamlessly shifts from playing C sharp to playing B, only its state of sounding [the audible quality of pitch] changes. There is still a single sound event of which each note instance is a part, and so each note instance is part of a single continuous sound...[however], a temporally seamless transition from one instrument's playing a C sharp to another instrument's playing a C sharp involves numerically distinct sounds of the same type" (O'Callaghan, Sounds: A Philosophical Theory, 87). For O'Callaghan, temporal continuity is the key factor in determining the number of sound particulars (distinct events) present in a "scene." Spatial continuity in physical space (relative to the sound source) is likewise critical, but continuity in pitch space (as in the form of a glissando) is not. O'Callaghan does not specify that the "seamlessness" of shifting from C sharp to B need entail a glissando in pitch space, i.e., the seamlessness he describes occurs in time-not in pitch space. Due to the lack of spatial continuity in the latter situation described by O'Callaghan, two distinct—yet qualitatively similar—sounds exist. My claim is that such a situation yields a point shift, which is the smallest possible unit of motion between stationary sound sources. The point shift, although comprising two distinct events, is a unit of grouped particulars, or a "complex event." If an instrument playing a C sharp was to move across the stage while sustaining its pitch (or if a recorded sample of the instrument playing C sharp was panned from one speaker to another), it would still count as a single sound (in both Bregman's and O'Callaghan's estimation) even though it is not locationally stable (in relation to the perceiver) in actual space. The sound is, however, spatially persistent in relation to its source: it is the source itself that is in motion.

ensure the production of Figure 4-12(b).⁷³ Such motion unfolds in chromatic pitch space and is comparable to a perfect panning gesture in an eight-point ensemble space. If the motion of Figure 4-12(c) is viewed as occurring in chromatic pitch space, it would be relatable to a gesture such as SG-[02457] (imperfect panning gesture) occurring in eight-point ensemble space. However, the motion of Figure 4-12(c) is a diatonic span that is perhaps most readily perceived in diatonic pitch space, in which case it corresponds to SG-[01234] (perfect panning gesture) unfolding in a five-point ensemble space. If the glissando of Figure 4-12(a) were performed on the white keys of a piano, the actual result would be Figure 4-12(c), although the rapidity of the gesture might hamper the discernment of five discrete pitches. Figure 4-12(d) shows a triadic span—an ascending arpeggiation of a G-major triad—that we may liken to SG-[024] in five-point ensemble space. The disjunction of the interval of a third, wherein a diatonic scale degree is "leapt over," relates to disjunct spatial motion in which an ensemble point is "skipped" during a point shift. Finally, Figure 4-12(e) is simply an ascending, disjunct interval that corresponds to a disjunctive point shift in ensemble space.

From these observations, we may conclude that an analog gesture in the electronic medium is quite easily received as a "motion of sound" because it lacks the discontinuities between parts that require those parts to be grouped together in perception. However, since grouping tendencies are for the most part automatic, apparent (or virtual) motion is readily perceivable within an acoustic ensemble space. These elemental differences notwithstanding, the overall spatial shape and directionality of a L-R PAN in both mediums is identical, and they carry the same interpretive potential. An oscillation gesture can easily occur in either the stereo or quadraphonic environment. Closed gestures and weak open gestures in the stereo field of

⁷³ The rate of the glissando would affect the perceptibility of the distinct pitches. A fast glissando appears more continuous (i.e., more like a portamento) than a relatively slower glissando.

electronic works often prove difficult to decipher due to the inconsistency of inter-speaker imaging.

In the electronic medium, a panning gesture arises from the gradual adjustment of the level (intensity) of a sound at each speaker. Beginning with a sustained sound concentrated exclusively in the left channel (i.e., panned "hard left"), the level of that same sound is gradually increased in the right channel. Once the sound is equally present in each channel, the amplitude of the sound in the left channel is gradually decreased. The sonic image seems to travel from the left to the right speaker.⁷⁴ Stationary sounds may be placed at any point between two speakers by means of a stable proportion of a sound's intensity between two channels. Besides controlling the azimuth of sound events, the illusion of distance (or, depth) may be created by manipulating the amount of artificial reverberation present and by adjusting both the intensity and the timbre of the sound. Imagine the analog L-R PAN described above accompanied by a progressive increase in reverb (beginning completely "dry" and ending extremely "wet"), a steady reduction of intensity level, and a gradual "darkening" of tone color (achieved by gradually filtering the upper frequency components of the sound). The gesture would illusorily appear to recede to the back of the room (away from the listener) while traversing from the left to the right of the sonic field.⁷⁵ Therefore, the stereophonic field is more than just a lateral plane situated in front of the listener-it is, effectively, a sonic "box" placed before the listener. With the illusion of spatial depth attainable through the aforementioned procedures, the "box" is highly elastic and may even be perceived as larger than the actual performance space in which the music is being performed.

⁷⁴ This process describes overlapping dynamic envelopes between two channels/speakers.

⁷⁵ An increasing ratio of reflected to direct sonic signals, attended by a decrease in overall amplitude and a concomitant attenuation of higher frequencies, effectively fabricates the sonic cues (described in Chapter 2) that enable the localization of distant sounds and thereby engenders the virtual recession of the sound.

Quadraphony entails placing a speaker in each of the four corners of a performance space, resulting in a definitive two-dimensional space. In such a sonic environment, the audience abides *within* the quadraphonic field. As opposed to the stereophonic field, in which the sonic "box" is placed *before* the listener, the listener is now situated *within* the box. Thus, the audience is contained within the ensemble space, which, in turn, is contained within the performance space. Phantom images may exist anywhere within the box, along the "proper" edges of the box (as defined by speaker placement), or beyond the actual boundaries of the box, and sounds may move through and/or around the listener by following a variety of spatial paths. Furthermore, the carefully articulated co-application of artificial reverberation, level changes, and filtration can expand the box in any direction relative to the listener.⁷⁶ Therefore, a sound tracing a single directional vector (moving in only one direction) can appear to approach the listener from afar (from one cardinal direction) and subsequently retreat into the distance, continuing along the same directional vector.⁷⁷

In the four-channel quadraphonic environment, four distinct "quadrants" exist: front-left (F/L), front-right (F/R), back-left (B/L), and back-right (B/R). A sound may belong exclusively to one quadrant, exist somewhere "between" quadrants in a stationary fashion (through inter-speaker "phantom" imaging, which, as noted, is often imprecise), or may travel continuously between two (or more) quadrants. When the latter motion occurs (traveling from one quadrant to another), the gesture shall be termed a *passing gesture*. Let us explore how we might designate the different directionalities of passing gestures that we might encounter in a quadraphonic work.

⁷⁶ This type of illusory "expansion" applies to both the apparent borders of the ensemble space and the perceived size of the performance space and differs from the expansion of the sonic space within (but not beyond) the ensemble space, as in the context of accumulative spatial gestures (see Chapter 3).

⁷⁷ As a sound is made to "pass" the listener, a drop in pitch would create the illusion of the Doppler effect (see Chapter 2) and likely enhance the perception of a moving sound. The faster the sound is "moving," the greater the drop in pitch must be to sustain the illusion.

A passing gesture that occurs between speakers that all lie on the same vertical level (i.e., all four speakers are on the floor, or all are elevated, etc.) is a *horizontal passing gesture*, or HORIZONTAL PASS.⁷⁸ Horizontal passing gestures partition further into the following four types: FRONT PASS, BACK PASS, LATERAL PASS, and DIAGONAL PASS.⁷⁹ Figure 4-13 lists the further divisions of each of these four categories.

FRONT PASS: $L \rightarrow R$ FRONT PASS $R \rightarrow L$ FRONT PASS

BACK PASS: L → R BACK PASS R → L BACK PASS

LATERAL PASS: $F/L \rightarrow B/L \text{ LATERAL PASS}$ $B/L \rightarrow F/L \text{ LATERAL PASS}$ $F/R \rightarrow B/R \text{ LATERAL PASS}$ $B/R \rightarrow F/R \text{ LATERAL PASS}$

DIAGONAL PASS:

F/L → B/R DIAGONAL PASS F/R → B/L DIAGONAL PASS B/L → F/R DIAGONAL PASS B/R → F/L DIAGONAL PASS

Figure 4-13: Horizontal passing gestures

Because sound localization is most accurate to the front and back of the listener, the

FRONT PASS and BACK PASS assume separate categories. However, since lateral sound

localization is relatively inaccurate, I have chosen to group the forward and backward

⁷⁸ In other multi-dimensional electronic spaces, a variety of vertical and oblique (in 3-D space) passing gestures are feasible.

⁷⁹ The FRONT PASS in the quadraphonic space is virtually identical to the L-R or R-L PAN in a stereophonic setting. The term PASS (i.e., passing gesture) shall be used to indicate that the gesture occurs within a quadraphonic space (in an effort to accurately differentiate between stereophonic and quadraphonic spaces). Additionally, an arrow (\rightarrow) is placed between quadrants names in the labels for all passing gestures, as evident in Figure 4-13.

LATERAL PASSES into the same category. Blauert notes that the perceived position of a sound source effected through summing localization "can be well controlled when the loudspeakers are placed in front of the listener...but not when [they] are placed at the side of the listener."⁸⁰ This diminished capacity for accurate lateral localization results from the "cone of confusion" (or, "zone of ambiguity") that exists to either side of the listener. The "cone" is an area extending and spreading in cone-like fashion from the listener's ear to the sound-source in which the listener cannot precisely identify the location of the sound source due to the ambiguous cues arising from interaural time and level differences. In areas close to the listener's binaural axis (perpendicular to the medial plane), interaural differences are often negligible.⁸¹ The listener would know that the sound was occurring to his/her right or left but might have difficulty pinpointing the exact location of the source. Head movements, which alter the sound-funneling function of the pinna, can aid in minimizing the cone's size, but a lateral pass might not be perceived as a continuous motion due to the inefficacy of lateral sound localization.⁸²

Blauert notes: "It is not possible to generate precise auditory events in directions to the side [of the listener] using four loudspeakers, as is common in quadraphonic sound systems. At least six loudspeakers are necessary for a reasonably precise representation of all azimuths around the listener."⁸³ Although he does not specify, it is presumed that the fifth and sixth speakers would be placed directly to the left and right of the listener, along the binaural axis. Such a setup exists in the 7.1 surround sound system, which is diagrammed in Figure 4-14. In addition to the standard quadraphonic setup, the 5.1 surround system features a center speaker

⁸⁰ Blauert, 325.

⁸¹ Gelfand, 379–380. Interaural differences are most noticeable and effective when the sound source is located near the listener's medial plane—in front of or behind the listener.

⁸² The questionable perceptual salience of lateral passing gestures should not compromise their potential analytic relevance.

⁸³ Blauert, 325.

and subwoofer (the ".1"). With the 7.1 configuration, a "surround-left" channel is placed between the front-left and back-left channels, and a "surround-right" assumes the space between the front-right and back-right speakers. Essentially, each of these speakers serves the same function in the lateral plane that the center speaker serves along the front plane: "filling-in the hole" between widely spaced speakers and compensating for potentially poor inter-speaker imaging (ineffective summing localization).



Figure 4-14: Diagram of 7.1 surround sound system

A specific sequence of passing gestures might form a circular gesture. For instance, the following "spatio-gestural progression" yields a *closed clockwise circular gesture*: $L \rightarrow R$ FRONT PASS – F/R \rightarrow B/R LATERAL PASS – R \rightarrow L BACK PASS – B/L \rightarrow F/L LATERAL PASS.⁸⁴ A variety of polystreaming relationships may likewise arise in the quadraphonic environment through a concomitance of multiple passing gestures; for example, the concurrence of F/L \rightarrow B/R DIAGONAL PASS and F/R \rightarrow B/L DIAGONAL PASS results in a two-dimensional instance of crossed streaming.

⁸⁴ The circular gesture is "closed" because it begins and ends in the front-left quadrant.
Recall that within the electronic medium, the "space" is flexible: artificial reverberation can effectively increase the perceived size of both the ensemble space and the performance space. This "spatial pliability" creates the potential for the perceived spatial "length" of any passing gesture to be varied. For example, a F/L \rightarrow B/R DIAGONAL PASS can simply transpire between the actual speaker locations of the F/L and the B/R quadrants. Or, reverberation (along with other artificial cues for creating the illusion of depth) may be added to each temporal endpoint of the gesture—making it seem as if the sound approaches from, and recedes into, a great distance. If both of these F/L \rightarrow B/R DIAGONAL PASS gestures were to occur within an electronic composition (regardless of the quality of the sound being "passed" through the space each time), the two related spatial gestures would function as recurrent (albeit abstract) motives serving to unify the music. The latter example could serve as an extended version of the former, or the former could serve as a truncated version of the latter (depending upon which version of the F/L \rightarrow B/R DIAGONAL PASS occurs first in the composition).

The preceding paragraph described how spatial gestures might (illusorily) be *spatially* extended or truncated. Gestures may also be *temporally* augmented or diminished. In the electronic medium, due to the nature of playback devices, it is quite easy to determine the exact timings of spatial gestures.⁸⁵ Suppose the exposition (original appearance) of a F/L \rightarrow B/R DIAGONAL PASS occurs between 2:10 and 2:13 on the counter, taking a total of three seconds to unfold. A few moments later, the same gesture occurs between 4:02 and 4:08—occupying six seconds of musical time. The latter instance is a temporal augmentation of the original gesture; in fact, it has exactly doubled in duration. Directional and temporal correlations between spatial gestures—as discussed at the beginning of this chapter—reveal the unifying potential of related

⁸⁵ We may easily deduce the temporal relations (augmentation, diminution) between acoustic spatial gestures based on tempo indications and the rhythmic profile of the gesture, but the precise timing of live gestures may vary based upon variable tempi and the possible application of rubato.

gestural forms. The exact disposition of multi-gestural relationships (whether the gestures occur successively or simultaneously) frequently carries strong hermeneutic implications. The interpretation of spatial gestures and their interrelationships is the subject of the next chapter.

CHAPTER 5

INTERPRETING THE SPACE: A HERMENEUTIC APPROACH

The potential for recurrent spatial gestures—functioning as "kinetic motives"—to unify the spatial construct of a musical work is one of the central themes of this dissertation. However, accounting for spatial coherence is only one objective of spatial analysis, and important interpretive questions inevitably emerge: what might the specific spatial structure of an ensemble space symbolize? How might we interpret the specific directionality of a given spatial gesture? How might larger networks of meaning arise from transformations of gestures and multi-gestural interaction (polystreaming)? Finally, how do spatial gestures interface with events in other musical parameters, such as melody, texture, time, and dynamics?

Gestures with particular directionalities and trajectories carry with them powerful connotative and associative meanings. A specific directionality in space does not have any concrete meaning in and of itself that exists independent of our human understanding: we impart meaning to direction based on our embodied experience and enculturation. The meaning ascribed to a given directionality is often universal (based on general consistencies of bodily and environmental experience) but will frequently vary from one culture to another (owing to the specificities and idiosyncrasies of distinct cultural embeddedness). Since all of the musical examples analyzed in this dissertation are from the Western canon, I will primarily examine how the Western mind attaches meaning to spatial gestures based on embodied knowledge and cultural understanding.

5.1 Symbols in Space: The Semiotics of Spatialization

It is possible for the sheer design of an ensemble space to be meaningful: the spatial shape of an ensemble's onstage layout can be interpreted symbolically. This shape is visually evident even when ensemble members are producing no sound, and the full shape may be articulated sonically during sections of ALL PLAY. Geometric forms often assume symbolic meaning. The meaning of a symbol is often unique to a specific culture, and symbolic function may change over time within a given culture. The semiotic interpretation of musical activity is an important branch of musical hermeneutics. Semiotic thought has been applied to musical events, particularly those in the realm of pitch. A musical symbol is often a dynamic spatio-temporal translation of a purely spatial object or figure: it generally unfolds over time in metaphoric pitch space.

In order for the meaning of a musical sign (or, "signifier") to be perceived, the listener must have prior knowledge of what the sign "means" (the "signified," or the object to which a sign refers).¹ Eero Tarasti writes: "Music as a sign cannot exist without the competencies necessary for understanding it as such."² Learning is a requisite and critical component of semiotic understanding, and the precise nature of this learning is often culturally and/or historically specific; signs function within a socio-cultural context due to the networks of meaning assigned to them by the members of the culture. Many symbols originate in, and acquire their meaning through, religious practice. For example, the shape of the cross is a Christian sign. Tarasti discusses the fugue subject of Bach's *Fugue No. 4 in C# Minor* from Book I of the *Well-Tempered Clavier*, BWV 849. As the subject of a five-voice fugue, this

¹ Signs do not refer solely to concrete, tangible objects; a sign may also refer to an idea. William Alston identifies and critiques three types of theories of meaning: referential, ideational, and behavioral. The *ideational theory* identifies the meaning of a word or expression (as a sign) with the ideas with which it is associated. William P. Alston, *Philosophy of Language* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1964), 10–31.

² Eero Tarasti, Signs of Music: A Guide to Musical Semiotics (Berlin and New York: Mouron de Gruyter, 2002), 6.

figure fulfills its expected syntactic function: it appears successively in each of the five voices either at its original pitch level or as a "real" answer (a transposition that preserves the intervallic content of the subject as initially exposited). However, the emergent shape of the subject symbolizes the cross and thus represents the Christ. As shown in Figure 5-1, the interior disjunct interval (B-sharp to E) of the subject head "crosses over" the broad ascent of a major second achieved by the framing pitches C-sharp and D-sharp.



Figure 5-1: Bach, Fugue in C# Minor, mm. 1-3

Tarasti claims that listeners of the Baroque period would have recognized this sign and been cognizant of the religious undertones of the fugue. These types of musical signs are dependent upon our metaphoric construal of pitch space. The "crossing" of directed intervals in Bach's subject is apprehensible only because we conceive of pitch relationships in terms of object relationships in physical space (more shall be said on this point in section 5.4). Furthermore, this musical symbol is a sign of a sign: it signifies the cross, which in turn functions as a sign of the Christ. Perceiving the geometric pattern of an ensemble space is more direct and immediate—it is not dependent upon metaphoric understanding. However, "knowing" (through enculturation) what the spatial shape symbolizes persists in being requisite to the acquisition of symbolic meaning.

Edward A. Lippman observes: "In religious antiphony, a particular performance-practice may symbolize cosmic harmony, for the physical space of performance may be used to sub-serve symbolic ends."³ Thus, the physical separation of performing forces in responsorial and antiphonal psalmody, as well as in polychoral sacred concerti, is symbolic in a religious sense; for instance, the separation of the priest from the congregation (or choir) signifies the distinction between the "Shepard" (Christ) and his "flock" (Christian believers). On occasion, composers have deliberately infused the ensemble space with symbolic shape. For instance, in the second movement of *Ultimos Ritos* (1972) by John Tavener, the ensemble is arranged in the form of a cross, with the timpani outlining the four endpoints of the cross. By additionally infusing the work with descending vertical motion (some ensemble members are positioned in balconies), Tavener imbues this work with layers of symbolic religious imagery: the spatial arrangement of the ensemble symbolizing the Christ and the descending motion representing "the descent of the Eucharist."4

Besides the deliberate "symbolization" of an ensemble-space layout on the part of the composer, any ensemble space assumes a shape of some sort. The shape may or may not form a regular, familiar geometric figure. Harley states: "Even if only seen, the geometric designs may play an important role in the musical compositions to which they belong. As signs and symbols, geometric floor plans and performer placement diagrams are integral, though inaudible, elements of the musical structure."⁵ The shape of an ensemble space may form a straight or arced line segment, a squared or rectangular design, a circular or elliptical figure, and so forth. In the case of multidimensional ensemble spaces, the audience may or may not be contained within the space. For example, in Luciano Berio's Circles, the total ensemble comprises two separate circles on the stage (with a mobile singer). In Xenakis's *Persephassa*, on the other hand, a single circle encloses the audience.

³ Lippman, "Music in Space," 347.
⁴ Harley, "Space and Spatialization," 243–247.
⁵ Harley, "Spatiality of Sound and Stream Segregation," 165.

The circle is one of the most universal signs-it often represents similar ideas in different cultures and historical periods. Since a perfect circle has no beginning and no end, it typically represents completion, repeating cycles, wholeness, perfection, union, and infinity/eternity.⁶ Circles function symbolically in many religions: the circle represented Christian divinity prior to the popularization of the cross, the Wheel of Dharma is a Buddhist symbol that symbolizes Buddha's endlessly spreading teaching, and Stonehenge is a circular construction. Circles are likewise the preferred shape for an assembly of equals, such as a council circle or King Arthur's Round Table. The circle often appears in architecture and the visual arts to represent many of the above concepts: circular halos distinguish holy figures, rings represent union, wheels signify mobility, and so on.⁷ Renaissance architects frequently "reverted to the circle as the basis of design for churches."⁸ The shape of an ensemble space is visibly evident irrespective of the distribution (or utter lack) of sonic activity (i.e., the ratio of sonic space to ensemble space); the geometric pattern of an ensemble space is apparent whether or not any or all members of the performing ensemble are actively engaged in the production of musical sound. The form of a circular ensemble space emerges through the perceptual act of "connecting the dots" (as does the holistic spatial shape of *any* ensemble space). Invariably, there will be some degree of space between adjacent performers forming a curvilinear path. Yet, due to the principles of Gestalt perception-emergence and reification in particular-we conceive of such an arrangement as a whole circle. Whether or not spatial gestures unfolding in such an ensemble space are perceived as geometric forms (such as a perfect or imperfect circle), shall be addressed in section 5.3.

⁶ Tim Wallace-Murphy, *Cracking the Symbol Code: Revealing the Secret Heretical Messages within Church and Renaissance Art* (London: Watkins Publishing, 2005), 93–94.

⁷ Paul A. Calter, *Squaring the Circle: Geometry in Art and Architecture* (Dartmouth: Key College Publishing, in press).

⁸ Alva William Steffler, *Symbols of the Christian Faith* (Grand Rapid, MI, and Cambridge, UK: William B. Eerdmans Publishing Company, 2002), 74.

Other geometric forms carry symbolic meaning: the equilateral triangle represents the Holy Trinity of the Christian faith: "The equality of the three distinct sides and angles expresses the equality of the three distinct Persons [united as] the Triune God."⁹ Depending upon its orientation, a triangle may appear to point upward or downward. An upward-pointing triangle may represent ascension toward heaven, fire energy, or the male phallus. A downward-pointing triangle symbolizes water (which flows downward), female divinity, or female genitalia. The fact that a triangle can symbolize either a religious or a sexual object or concept is potentially problematic—some knowledge of the sign-user's intentions (perhaps only evident through the contextual usage of the sign) would be required to interpret the sign properly. If confronted with a triangular ensemble space, conflicting semiotic readings could easily emerge. That being said, overtly triangular ensemble spaces are less frequently encountered in performance than are circular and rectangular designs.¹⁰ The square is a symbol of the earth. Its non-dynamic shape suggests balance and resolution and "indicates the four points on the compass and the four corners of the earth."¹¹ The rectangle, particularly one displaying the Golden Mean (a ratio of 1 to 1.618...), is "considered the geometrical symbol of harmony."¹² Many large ensembles, such as the symphony orchestra, commonly assume the shape of a square or rectangle.

An ensemble space is a discrete and enduring structure, but the gestures that unfold within it are dynamic and transient. Robert Hatten states: "Musical gesture is movement

⁹ Frederick Roth Webber, Church Symbolism: An Explanation of the More Important Symbols of the Old and New Testament, the Primitive, the Mediaeval, and the Modern Church, 2nd ed. (Cleveland: J.H. Jansen, 1938; reprint, Detroit: Omnigraphics, 1992), 40-42.

¹⁰ In Alax (1985) by Iannis Xenakis, three separated instrumental ensembles (hyperpoints) form the shape of a triangle.

¹¹ Steffler, 80. ¹² Ibid., 80–81.

(implied, virtual, actualized) [that is] interpretable as a sign."¹³ Hatten constructs a cogent semiotic theory of musical gesture based upon the Peircean categories of icon, index, and symbol.¹⁴ By applying Hatten's generalized concepts to spatial gestures, we may interpret certain types of spatio-gestural motion as symbolic. Hatten explains that a musical gesture may be "thematized (foregrounded as a motive and used consistently in a work)" and may therefore function as and be treated (from an analytic standpoint) as a type of motive.¹⁵ However, a gesture-as-motive does not solely serve a syntactic function-its varied forms contributing to the coherence of the musical work in which they abide; musical gestures contribute to, and account for, expressive meaning in music. A gesture may *iconically* relay information concerning the "attitude, modality, or emotional state of the gesturer (or presumed agent)," may *indexically* reveal "reactions, goals, and orientations," and may function as symbol by "rely[ing] on conventions or habits of interpretation (in contexts such as artistic styles) to convey a wealth of extra meaning.¹⁶ Regarding the symbolic functioning of musical gesture, Hatten notes: "Conventionalized music-stylistic gestures (which must be learned, often by direct apprenticeship with musicians competent in the given style) are also relevant at all levels of

¹³ Hatten, *Interpreting Musical Gestures, Topics, and Tropes*, 121–126. Hatten uses the work of David Lidov as a springboard for developing his own semiotic account of gesture. See David Lidov, "Mind and Body in Music," *Semiotica* 66.1/3 (1987): 69–97.

¹⁴ An *icon* is a sign that somehow resembles or imitates the thing being signified. For example, a circle is an iconic sign of a wheel. An *index* is a sign that is physically and/or causally linked to the signified: smoke is an indexical sign of fire (the latter is the cause of the former). A *symbol* does not directly resemble or share a connection to the signified. Symbols are arbitrary and purely conventional. Of these three signs, symbols are the most culturally specific. Hatten states: "[Musical] gesture is the synthesis or integration of many musical elements, and (as bodily index and/or icon of movement shape/force) offers an immediate connection to expressive meaning" (Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 177). I concur that the connection of musical events (including gesture) to meaning is immediate, but not on the basis of semiotic functioning, wherein signs serve as *intermediate* connections between the thing signified and the perceiver. Rather, the immediacy is regulated by the automatic fleshing-out or metaphoric projection of *image schemata*. This notion shall be explored below in section 5.3. ¹⁵ Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 123. Emphasis Original.

¹⁶ Ibid., 125.

structure and form. In other words, as gesture enters into the *symbolic* level of musical style, it can have consequences for musical form.¹⁷

The semiotic interpretation of ensemble-space design and spatio-gestural activity is problematized by both the obligatory awareness of symbolic function within a specific culture and by the required knowledge of any intentionality on the part of the composer (as "signer"). Nevertheless, it might prove difficult or impossible to ignore any deeply instilled associations with a common shape, such as a circle. A circular spatial design might naturally produce an entirely different effect than that furnished by either a polygonal ensemble layout (one with three or more virtual sides and angles) or a simple line-segment design (perhaps the most common formation). However, I shall argue that our aesthetic reaction to (and interpretation of) both ensemble-space shapes and gestures unfolding within those spaces is primarily informed and conditioned by our accumulative embodied knowledge (which includes the experience of bodily motion), which, although somewhat socio-culturally delineated, is generally less culturally constrained than the comprehension of symbols. Hatten concedes this point when he states that "musical gestures need not correspond to any particular culturally encoded gesture to be meaningful, based on the tendency we have as biological creatures to interpret any energetic shaping through time as gestural and hence as potentially significant."¹⁸ The experience of musical events, including spatial gestures, is therefore immediate and need not be mediated by a sign (I shall reinforce this argument in subsequent sections of this chapter). This point, however, does not rescind the effectiveness of analytic applications of semiotic theory. The bulk of Hatten's work is devoted to revealing how the "symbolic levels of a style" enrich the "basic form

¹⁷ Ibid. Emphasis Original.

¹⁸ Ibid., 132.

of communication" supplied by embodied musical gestures.¹⁹ Hatten does not, however, furnish a comprehensive means of interpreting the velocities and directionalities of motion for musical gestures.²⁰ I shall attempt to provide such a method—one primarily applicable to spatial gestures but potentially germane to other types of musical gesture as well.

5.2 Categorization

In addition to the fundamental part that it plays in my system for codifying spatial gestures (as proposed in Chapter 3), categorization assumes an important role in my interpretive method. Since my taxonomic system is steeped in Gestalt psychology, labels such as "open" and "closed" betray a fundamental interpretation of a gesture's spatial form. Additional layers of "spatial meaning" may be superimposed on this most basic interpretation, as we shall discover throughout this chapter. The process of categorization is central to human perception and cognition.²¹ To begin with a basic, non-musical example, consider the mental category for "chair." Like most categories, the chair category contains highly prototypical and less prototypical members. The most prototypical form is often the category member most frequently encountered and utilized in experience.²² The mental image conjured when one simply hears the

¹⁹ Ibid.

²⁰ Under Hatten's conception of gesture, *velocity* would derive from sole or combinatorial attributes of musical time (meter, tempo, rhythm, and sudden or gradual changes in any of these parameters), and *directionality of motion* is essentially limited to ascending and descending motion in pitch space, as well as any "composite shapes" (such as melodic "arcs," etc.) wrought by the combination of such motions. (Dynamics and articulations may also affect the overall shaping of the gesture.) Hatten identifies "lift gestures" that correspond to ascending motion and "subsiding gestures" compatible with descending motion (Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 157, 162–163, 202–203). As we have already seen, spatial gestures may assume a wide array of shapes and directionalities based on the nature of the ensemble space in which they unfold. Since I am primarily exploring one-dimensional ensemble spaces, I am chiefly concerned with horizontal motion in actual space (as opposed to vertical motion in pitch space).

²¹ For an exhaustive account of categorization, see: George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind* (Chicago and London: The University of Chicago Press, 1987).

²² Obviously, categorical structure and prototype designation vary from culture to culture as well as from person to person within the same culture. The prototypical member of a category is often employed in a *metonymic model* to represent the entire category. (See Lakoff, 77–90).

word "chair" is often highly representative of the categorical prototype. We can (and do) conceptualize these mental categories in terms of physical containers. Accordingly, we likewise model the relations among category members with container diagrams, with links (or, "chains") connecting the various members to one another. Imagine the most prototypical member(s) of the chair category (most likely some version of an armless kitchen or desk chair) clustered together at the center of the container, with less prototypical members (such as stools, park benches, and beanbags) arrayed toward the fringes of the container—the least prototypical members being farthest from the center and closest to the boundaries of the container. A higher degree of mental processing is generally required to recognize and relate less typical category members.

Assessing an object's suitability for inclusion in a category is akin to the process of *pattern matching* described by social scientist Howard Margolis.²³ According to Margolis, we "match" perceived patterns to those stored in memory. This cognitive process is instigated via the act of *jumping*, which entails selecting the proper pattern to match incoming stimuli. For typical and recurrent items or experiences, pattern matching occurs automatically and instantaneously. For less typical objects or happenings, pattern matching may take longer and require a greater and more imaginative effort. This effort exists in the form of what Margolis refers to as *checking*, wherein we compare patterns and gauge their similarities and differences. Clearly, the conceptual processes of categorization and pattern matching are closely related (if not virtually synonymous). A family of similar or related patterns constitutes a conceptual category, and pattern recognition corresponds to categorization. In fact, we could consider pattern matching to be the means by which a radial category is organized. With the prototype occupying the central position of the category-container, similarity relations determine the

²³ Howard Margolis, *Patterns, Thinking, and Cognition: A Theory of Judgment* (Chicago: University of Chicago Press, 1997). Candace Brower applies Margolis's theories to music cognition; see Candace Brower, "A Cognitive Theory of Musical Meaning," *Journal of Music Theory* 44.2 (2000): 323–379.

"distance" of less typical category members to the central prototype: the distance of peripheral members to the central prototype is inversely proportional to the degree of similarity, such that a large distance indicates a smaller degree of similarity than does a relatively shorter distance. Less similar category members require more effortful pattern matching, as their interrelationships are less immediately apparent.

In addition to recognizing a chair and evaluating its categorical typicality, we are also directly aware of its potential pragmatic function(s): we immediately estimate the chair's "affordance(s)." An **affordance**, or "what [a thing] furnishes,"²⁴ is the "product both of objective properties and the capacities and needs of the organism that encounters them."²⁵ According to ecological psychologist James J. Gibson, the term affordance "implies the complementarity of the animal and the environment."²⁶ However, Gibson stresses that "the affordance of something does *not change* as the need of the observer changes."²⁷ When we meet with an object, such as a chair, we entertain a vast array of manners in which we might engage that object. These opportunities for interaction with an object are based upon past experience, tacit knowledge of our own sensorimotor capacities, and an estimation of the objective qualities of the object (i.e., its size, weight, malleability, and so forth).²⁸ Therefore, upon recognizing it, we know that we could sit on the chair. (The affordance for a person to sit upon a chair exists whether or not that person "wants" or "needs" to sit upon the chair.) However, we also know that we could stand on the chair to reach highly placed objects or pick the chair up to use in self-

²⁴ James J. Gibson, *The Senses Considered as Perceptual Systems* (Boston: Houghton Mifflin, 1966), 285. Gibson coined the term "affordance."

²⁵ Eric F. Clarke, *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning* (Oxford: Oxford University Press, 2005), 37.

²⁶ James J. Gibson, *The Ecological Approach to Visual Perception* (Hillsdale, NJ, and London: Lawrence Erlbaum Associates, 1986), 127.

²⁷ Ibid., 138–139.

²⁸ Thomas J. Lombardo, *The Reciprocity of Perceiver and Environment: The Evolution of James J. Gibson's Ecological Psychology* (Hillsdale, NJ, and London: Lawrence Erlbaum Associates, 1987), 306–308.

defense. Most objects carry multiple affordances—one of which will generally correspond to the object's primary functional role (as understood through bodily and cultural experience). Recognizing the affordances of an object is intimately bound with (and variable depending upon) how well the object "fits" into a category. Less typical items that require greater effort to identify and categorize might not carry easily discernible affordances. In other words, if we cannot discern what an object *is*, nor qualify its objective properties, we cannot fully assess its manifold affordances. However, Gibson insists that "what we perceive when we look at objects are their affordances, not their qualities."²⁹ Whether or not we successfully classify an object based on its "family resemblances" has no bearing on gauging the functional usefulness of that object. In Gibson's words: "You do not have to classify and label things in order to perceive what they afford."³⁰

Lawrence Zbikowski has shown that categorization is central to the perception and conceptualization of music, particularly when a musical work exhibits a profusion of related musical motives.³¹ A *motive*, as a gestalt structure with specified relations between parts, may exist (and recur) in any musical parameter: a motive may be a sequence or collection of pitches, a harmonic progression, a rhythm, a melodic contour, a spatial gesture, or a transformative procedure. Affiliated motives form a category, the prototypical member of which is generally highly contextual and unique to the work in question. Motivic variation and alteration challenge our capacity for pattern matching.

Generally, the term "prime form" indicates the condition of a motive prior to any alteration or transformation. The prime form of a motive is typically exposited quite early in a musical work and is subsequently varied and developed; hence, the prime form of a motive is

²⁹ Gibson, The Ecological Approach to Visual Perception, 134.

³⁰ Ibid.

³¹ Zbikowski, *Conceptualizing Music*, 23–62.

said to be "generative" of other musical material. In Schoenbergian terms, the *Grundgestalt* is essentially what I am calling the motive prime form.³² I shall employ the term **principal gestural form** to indicate the most dominant spatial gesture of a particular work. The prominence of the principal gestural form is determined by its frequency of occurrence and/or perceptual salience.³³ For example, if throughout a given work the presence of left-to-right panning gestures exceeds that of right-to-left panning gestures, then L-R PAN shall function as the principal gestural form. Identification of the principal gestural form likewise indicates the most dominant gestural category within the work under investigation; in the present hypothetical work, open gestures (particularly strong open gestures) would take precedence over oscillation and closed gestures.

Any musical motive carries with it vast potential (or, "affordances") for variation. Drawing on Gibson's conception of "affordance," the *transformational affordances* of a musical motive are based upon both the structural attributes of that motive and the listener's implicit and explicit knowledge of object manipulation (in general) and motivic variation (as a metaphorical type of object manipulation). When we encounter a salient motive at the beginning of a musical work, we likely expect it to be altered *somehow* during the course of the music. The "experienced listener" might consciously entertain the assortment of related motive forms that the original motive affords—whether or not they actually occur in the work. It is not terribly difficult to imagine (and anticipate) a motive being retrograded, inverted, temporally augmented and diminished, or otherwise altered. The consideration of prospective motivic alteration might occur during the listener's first exposure to a musical work, wherein generalized musical

³² Schoenberg's distinction between *Motive, Gestalt,* and *Grundgestalt* is outlined in Chapter 3.

³³ In Chapter 3, I argue that strong open gestures are inherently (and abstractly) the most perceptually salient of all spatial gestures due to their pronounced Prägnanz. However, any type of spatial gesture may serve as the principal form, as salience is dependent upon compositional factors (as well as performance conditions).

processes, such as motivic variation, are forecasted. Candace Brower refers to such processes as "*musical schemas*—patterns abstracted from musical convention."³⁴ Musical schemas recur in many stylistically and historically diverse works of music. Subsequent encounters with a particular musical work might create stronger "links" between related motive forms that do actually occur in the work—what Brower refers to as "*intra-opus patterns*—patterns specific to the work itself."³⁵

Knowledge derived from the experience and analysis of music (that of motivic variation and developmental procedure in particular) might augment the attendant *transformational affordances* of a salient motive. The presence of such affordances (as potentials for variation and transformation) accords with Leonard Meyer's conception of *expectation* as being paramount to the experience of music: we expect (based on past experience) motives to be varied in certain ways, and that expectation may be satisfied or thwarted during the music's unfolding.³⁶ Various alterations to the principal form of a spatial gesture (such as retrograde, inversion, nudging, abbreviation, elaboration, and others outlined in Chapter 4) produce less typical category members, but these altered forms are still received as interrelated category members provided that some characteristic gestural feature remains invariant despite the application of one or more transformative procedures. Zbikowski points out that a composer's "compositional strategy [can often] exploit and challenge our ability to categorize."³⁷ Indeed, transformational procedures often obscure related motive forms. Zbikowski likewise shows that even structurally related motives might be classified into separate categories if they serve different syntactic and/or

³⁴ Brower, 324. Emphasis Original.

³⁵ Ibid. Emphasis Original. Brower distinguishes "musical schemas" and "intra-opus patterns" from *image schemata* (the latter to be explored below). I argue that image schemata are central to (and generative of) both musical schemas and intra-opus patterns.

³⁶ Meyer, 35–38. David Lewin discusses the related concept of *protension* in "Music Theory, Phenomenology, and Modes of Perception," 329.

³⁷ Zbikowski, Conceptualizing Music, 49, 137–200.

rhetorical functions. Such intercategorical relationships often unfold over long spans of time and can confer large-scale meaning to a work of music.

The identification of related motives (category members) is vital to musical perception and understanding.³⁸ Schoenberg argues that the recognition of musical motives is central to the comprehension of musical coherency. When related motive forms prove to be ubiquitous within a musical work, the work may be received as a unified whole. Here we encounter a prime example of the hierarchical organization of Gestalt perception: we perceive small collections of pitches (generally those in close temporal proximity) with specific characteristics (intervallic, rhythmic, contour, etc.) as individual motives. We perceive the motive as a whole without (necessarily) consciously attending to its "parts" (such as its pitches, intervals, durations). On a higher level, we conceptualize multiple (and related) motive forms as parts of the musical work as a whole. The longer the time-span of a particular work, the greater the challenge to perceiving the work as a unified whole.

The principal gestural form functions as the prototypical member of a gestural category at the local level of the specific work in which it resides. Transformed and altered versions of the principal form are less typical category members. Therefore, while basic category membership remains a constant, the specific hierarchical organization and interrelationships among spatial gestures within the same category may change from one work to the next. Likewise, it is possible to have competing candidates for the principal gestural form label. We might discover processes in which an established principal form is ultimately supplanted by another gestural type—either from the same gestural category or else from a different one altogether. Such a

³⁸ The term *identification* may refer to recognition as it occurs either during the aural experience of music or upon analysis of the score.

discovery would reveal a transformation of the functional (syntactic, rhetorical) relationships between different kinds of gestures.

If we conceptualize any given gestural category as a physical container, we can envision the principal gestural form (as the local categorical prototype) located at the center of the category-container. All permutations of the principal form are situated toward the periphery of the container, with the more typical members (those closest in identity to the prototype) arranged nearest to the prototype and the least-typical members (those resulting from the most drastic transformations/alterations) located closest to the boundaries of the container.³⁹ Secondary gestural subcategories may be organized as a "ring" around the primary subcategory. For instance, if SG-[0123] (L-R PAN) is the principal gestural form, then all strong open gestures (perfect and imperfect) shall be located toward the center of the category-container. All weak open gestures shall exist outside this centralized location, toward the periphery of the container. In Figure 5-2, a dashed circle within the open gesture category-container separates strong open gestures (at the center), which are *more* similar to the principal form, from *weak* open gestures (outside of the center), which are less similar to the principal form. The gestural product of an intercategorical alteration is not to be found in this category-container, though it may still exhibit similarity to the principal form/prototype. In fact, all extracategorical gestures abide in a conceptual space outside of the primary gestural category-container.

³⁹ Frequency of occurrence and perceptual salience determines the principal form. Relational factors (such as event cardinality, directionality, and event-point order) determine the proximity (i.e., "likeness") of other intracategorical forms to the prototype. In other words, all nonprincipal gestural forms are arranged within the category based on their similarity to the principal form, *not* based on their rate of occurrence and/or salience. In fact, we could develop a network showing how all potential spatial gestures relate to the principal gesture, whether or not those gestures actually occur in the work of music. Such a network would represent the affordances of the principal gestural form, whether or not they all are actualized during the work's performance. During the progress of the musical work, a conceptual path through the network is navigated as some (or all) of the potential motive forms are realized.



Figure 5-2: Hypothetical gestural category-container with SG-|0123| as the principal gestural form

From one work to another, the locatedness (geometric coordinates) of each category member (container item) relative to other members within the category-container will change depending upon the exact nature of its employment in each work: the topography of the gestural categories (as arranged in reference to the central, principal gestural form) is variable from one work to the next (assuming different principal gestural forms for each work). Furthermore, if the identity of the principal gesture changes during the course of a work, we may observe a topological change in the organization of the gestural category. Of course, all of these categorial "spaces" are purely conceptual and are used to model both intra- and intercategorical gestural relationships.

5.3 Image Schemata

The interpretable meaning(s) of gestural directionality and of multi-gestural interactive processes may be modeled, in part, by image-schema theory. This theory, which is rooted in the philosophy of Immanuel Kant (among others), finds its twentieth-century maturation in the work of philosopher Mark Johnson.⁴⁰ Johnson's theories have challenged one of the primary tenets of the Western philosophical tradition: that meaning is utterly objective and exists separate from human understanding. This objective view of meaning also maintains that the mind exists and operates independent of the human body. Johnson has challenged this Cartesian, mind-body dualism by claiming that all modes of cognition (conscious and unconscious) are rooted in *embodied knowledge*. He has explored the critical role that the human body—particularly its sensory apparatuses and motor programs—plays in directing the activity of the mind. In essence, we would not think the way we think if our bodies did not function the way they function.

Repetitive bodily and environmental experiences give rise to preconceptual mental structures called **image schemata**, which, in turn, serve to structure our continuous experience as well as our understanding of abstract phenomena. Johnson describes an image schema as "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience."⁴¹ These abstract structures are neither context-bound nor specific to any concrete experience; they simply "organize what can be perceived."⁴² Furthermore, a schema is not tied to any solitary perceptual modality—pertaining exclusively to either vision or audition, for example. In fact, many schemata are likely formed, refined, and

⁴⁰ Recall from Chapter 3 that Kant was likewise a notable influence on the development of Gestalt psychology (which, in turn, influenced the theories of Arnold Schoenberg, Heinrich Schenker, and Leonard Meyer, among others).

⁴¹ Mark Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* (Chicago: The University of Chicago Press, 1987), xiv.

⁴² Lakoff, Women, Fire, and Dangerous Things, 453.

sustained through intricate crossmodal binding and multisensory interaction. Image schemata may be "fleshed-out" or metaphorically projected in an infinite number of ways—allowing for a rich and subjectively variant range of experience. Although, as we shall see, this subjectivity is somewhat constrained by the nature and internal structure of the schemata. While Johnson makes no substantial mention of music in his writings, I contend that these image schemata are the very structures that enable our primary response to music. They likewise serve as the conceptual infrastructure upon which theoretical models may be constructed.

In his theory of auditory scene analysis—in which related sounds are integrated into streams, and individual streams are segregated from one another within the total auditory scene (as summarized in Chapter 2)—Albert Bregman considers both "primitive" and "schema-based" processes for grouping and organizing sonic stimuli. The primitive processes of stream segregation (sequential and simultaneous integration) are influenced by Gestalt principles of perceptual organization and are considered to be somewhat innate, having developed during humankind's evolutionary adaptation to the environment. Bregman describes a schema as "a mental representation of some regularity in our experience" and schema-based integration as deriving from "learned rules that affect the perceptual organization of sound."⁴³ Bregman's conception of the schematic structuring of experience is congruent with that of Johnson. Bregman primarily discusses schematic processing as it affects and shapes auditory perception. Various schemata serve to structure our reception and perceptual organization of multitudinous auditory events, including musical ones.

An image schema may be thought of as the abstract skeletal framework that organizes a host of fundamentally similar experiences. However, we must bear in mind that schemata are mental structures guiding our reception of events—they do not exist *in* the events themselves.

⁴³ Bregman, 43.

An example of a pervasive image schema is what Johnson calls the PATH schema, which is shown in Figure 5-3.



Figure 5-3: PATH schema⁴⁴

This schema consists of three major parts: a source point A, a terminal point B, and a vector tracing a path between them. The vector may be a physical distance, a length of time, or both—depending upon the actual experience the schema is structuring.⁴⁵ The PATH schema structures my experience of walking from the kitchen to the dining room as well as flying from New York to Beijing. Both of these experiences boil down to my traversing from point A to point B, although the nature of those two point-locales, as well as the time, distance, and other factors involved in each respective journey, vary considerably.

From this specific example of the PATH schema, which we shall return to shortly, we may garner several important features of image schemata in general. First, they are *gestalt structures*, meaning they have an irreducible, internal structure with specified relations between parts. Second, schemata are *preconceptual:* we are not consciously aware of their presence in structuring our experience, and they may or may not contribute to the formation of specific concepts that may later be called upon to reason with. Third, they are frequently *dynamic*, entailing change over time. Finally, image schemata are *topological:* their basic structures are not affected by changes in size or shape.⁴⁶ In this sense, they are quite fluid and malleable.

⁴⁴ Taken from Johnson, *The Body in the Mind*, p. 28, Figure 3.

⁴⁵ As noted in the introduction, we conceptualize time in terms of space. See George Lakoff and Mark Johnson, *Philosophy in the Flesh*, 137–169.

⁴⁶ One way to conceive of the image schema is to relate it to Schenker's notion of the *Ursatz*. In Schenker's theory, the Ursatz is the fundamental structure that abides deep in the background of representative works from the tonal

Additionally, any given schema features multiple *entailments*, which Johnson defines as "implications of the internal structure of image schemata."⁴⁷ Candace Brower elaborates Johnson's explanation of entailments by describing them as "consequences of [a schema's] structure that give rise to more extended inferences and thus to more elaborate metaphorical interpretations."⁴⁸ Such entailments regulate and constrain the metaphorical projections of image schemata that enable our perception and understanding of abstract events (such as musical phenomena).

In addition to discussing what image schemata *are*, I shall mention a few things that they are *not:* image schemata are not representational symbols that mediate between our understanding and "things in the world." Instead, they are mental structures of understanding that give order and coherence to all kinds of experiences. We perceive *by means of* these image-schematic structures, we do not directly perceive the structures themselves. Therefore, I see little potential for the semiotic application of image schemata in musical analysis.⁴⁹ Additionally, image schemata are *nonpropositional*—they are not subject-predicate structures of any specificity. However, the pictorial representations of the various schemata *are* propositional, and

repertoire. All such works of music, which may be drastically contrasting on the surface, are ultimately reducible to the same fundamental structure. The Ursatz is topological, dynamic, and has an internal structure: the three or fiveline *Urlinie* over the *Bassbrechung* (an eight-line Urlinie is feasible but rarely encountered). Similarly, an image schema is a basic pattern lurking in the preconceptual portion of our minds that gives structure to a wealth of experience. Just as Schenker argues that the Ursatz itself comes from the *Naturklang* (the "chord of nature," or the naturally-occurring overtone series), so too does an image schema come from natural (and repeated) experience with our bodies within our physical environments. However, this metaphor has its limitations: in Schenker's theory, the Ursatz is manifested at various levels of structure *within* the music. Image schemata, on the other hand, are purely mental frameworks that structure our reception of musical events—they do not abide within the music itself.⁴⁷ Johnson, 22.

⁴⁸ Brower, 327.

⁴⁹ Since, within the realm of semiotics, no intrinsic correlation (objective correspondence) exists between a symbol and the object to which it refers, understanding (within a defined conceptual scheme and culturally established network of meaning) must mediate between the sign (signifier) and what it represents (the signified). Due to the necessary mediation of human understanding, semiotic interpretation yields a type of *indirect* meaning. Even signs that are so pervasive within a culture that they appear to transmit meaning automatically must have been "learned" at some point in time. Image schemata, on the other hand, are not "learned" in the traditional sense of the word they are acquired through experience. Schemata mediate between experience and understanding and are therefore *directly* meaningful. Furthermore, the figure used to depict the structure of an image schema is purely representational and is *not* intended to function as a sign.

they may be discussed and described in propositional form. Johnson stresses that figures (such as that presented in Figure 5-3) are intended as aids in explaining the structure of the schemata, but they should not be confused with rich images or specific mental pictures: "The image schema is not an image. It is, instead, a means of structuring particular experiences schematically, so as to give order and connectedness to our perceptions and conceptions."⁵⁰

I argue that our most basic reactions to music—beginning at the preconceptual level result from the automatic fleshing-out of various embodied schemata during the experience of music; abstract musical events are received and unconsciously organized in terms of the listener's accumulated embodied knowledge-thus rendering the musical experience immediately and directly *meaningful*. Image schemata structure our reception of musical events occurring in any parameter of a musical work, including spatial gestures emerging within the ensemble space. Since perception is an active process, our minds continually and unconsciously work to structure incoming sensory stimuli of all kinds into internal perceptions that areideally-sensible and accurate representations of the external state of affairs. Accordingly, when confronted with music, our minds automatically seek to organize and make sense of the auditory stimuli. This is accomplished through the receiving and realizing of musical events *in terms of* our embodied knowledge, which exists in the form of countless image schemata. Bregman claims that schemata "become active when they detect, in the incoming sense data, the particular pattern [of experience] that they deal with."⁵¹ When an embodied schema is *activated* during experience by incoming stimuli, it is then automatically fleshed-out, or unconsciously *realized*, at the preconceptual level. Schemata may also be projected metaphorically to structure an understanding of abstract or unfamiliar phenomena. Finally, image schemata may be employed

⁵⁰ Johnson, 75. ⁵¹ Bregman, 666.

in an analytic context to model musical events. Such an analytic model may not conform exactly to the preconceptual structuring of musical events, although there is sufficient reason to suppose some degree of correspondence between the two.

I wish to draw a distinction between conceptual metaphor, which can result in the formation of a particular concept (and which shall be explored below), and what I call the "fleshing-out" of image schemata, which is a preconceptual process that need not result in a specific concept.⁵² Image schemata do, however, inform the source domain of conceptual metaphors that we use to reason about musical phenomena, which, in turn, generate the linguistic metaphors that we use to describe musical events. We cannot help but experience music in terms of concrete physical and bodily processes, so it follows that we should think and talk about music in such terms.

The same kind of schema may be activated and fleshed-out by concomitant musical events occurring within different parameters of music.⁵³ Furthermore, different *kinds* of schemata—many of which are dynamically interrelated and complementary—may be realized simultaneously. Johnson acknowledges what he calls "the superimposition of schematic structures, which [allows for] a host of complex meaning structures central to our experience and understanding."⁵⁴ Candace Brower notes: "Because image schema[ta] are both simple and abstract, they are easily adapted to one another and [to] a variety of structural features within the target domain [of a conceptual metaphor]."⁵⁵ As we shall discover, a single image schema

⁵² While a conceptual metaphor yields a concept that we use for thinking about music, the fleshing-out of image schemata (in terms of our sensorimotor experience) allows for a basic response to music that can occur in the absence of (or without having to form, via "cross-domain mapping") specific concepts *about* music. Such concepts do exist, and they inform our conscious thoughts about music. (In general, concepts are a part of the cognitive unconscious, but we employ them to reason at a conscious level.) The same schemata that are fleshed-out at the preconceptual level structure these concepts.

⁵³ It is perhaps best to imagine multiple *copies* of the same schema being realized concurrently.

⁵⁴ Johnson, 125.

⁵⁵ Brower, 327.

seldom operates in solitude in structuring our experience. I shall now explore several different image schemata and describe how each schema can be called-upon to model a host of musical processes, including those in the spatial realm. It shall prove valuable to demonstrate how the same schema might structure events unfolding in multiple musical parameters. The first schema to be discussed, the TENSION-AND-RELEASE schema, is often intrinsically linked to many other image schemata.

The TENSION-AND-RELEASE schema

The notion that certain events and processes within a musical work can signify and induce varying degrees of tension, and the subsequent release thereof, is nothing new to musical discourse. Leonard Meyer employs the term in his seminal work Emotion and Meaning in Music of 1956, where he states: "The greater the buildup...of tension, the greater the emotional release upon resolution."⁵⁶ In *Tonal Pitch Space*, Fred Lerdahl writes: "Fundamental to the experience of tonal music is the hearing of patterns of tension and relaxation as the events in a piece unfold."⁵⁷ Lerdahl goes on to note that "there are [different] kinds of musical tension...[such as] the surface tension of a crescendo, a rise in melodic contour, or a suddenly dense texture."58 What is missing from Lerdahl's account of tension and release is an attempt to explain why we perceive certain musical events in terms of tension and others as a release of that tension. Lerdahl does acknowledge that the very terms "tension" and "relaxation" are metaphorical. Image-schema theory offers an explanation as to how and why the metaphor *works*—how it is grounded in embodied experience, and how these (and other) metaphorical labels that we use to

⁵⁶ Meyer, 28. ⁵⁷ Lerdahl, 142.

⁵⁸ Ibid., 149.

describe musical events actually reflect our initial aesthetic, psychological, and (often) physical response to received musical events.

Unlike the PATH schema depicted above (Figure 5-3), the TENSION-AND-RELEASE schema—henceforth the "T&R schema"—is difficult to represent pictorially as a concise visual image. In fact, it might seem desirable to think of tension and release as being two separate experiential schemata. However, the two are intimately bound. The act of producing tension (in any number of physical or psychological manifestations) necessitates the eventual release of that tension.⁵⁹ Conversely, the release of tension sets the stage for the reapplication of tension. Imagine flexing (or, tensing) a muscle within your body. With the flexing comes the tacit knowledge that the muscle will eventually become fatigued (and possibly sore) following the accumulation of lactic acid, making indefinitely sustained flexing a physical impossibility. Thus, embodied experience teaches us that relaxing is inseparably linked with flexing.⁶⁰ The act of tensing and releasing is generally a cyclic and recursive process. A single cycle comprises a build to the point of highest relative tension and its subsequent release—whether gradual or immediate. Therefore, the T&R schema has a very well defined internal structure, consisting of the indefinite alternation of two main parts (tension and release) that are functionally dependent upon one another and therefore inseparable (see Figure 5-4).

⁵⁹ The term *tension* connotes a multitude of universal physiological and psychological experiences that conspire to form the T&R schema. There are many appropriate synonyms for "tension." Some alternative terms include: flex, stress, pressure, strain, weight, etc. I have chosen to use the word *tension* (and the partner term *release*) for several reasons. First, as I have noted above, it has commonly been employed in prior discourse regarding musical experience and meaning. Second, the term seems to capture the widest range of experiential phenomena; many apparent synonyms are actually more specific types of tension.

⁶⁰ *Expectation* is an intrinsic component of the T&R schema. On a general level, musical tension arouses the expectation for release of that tension, and released tension creates the expectation for the reapplication of tension. Denied or unfulfilled expectation for the release of tension can prolong and intensify perceived tension. The informed listener, when cognizant during the musical experience, might experience expectation in the more specific form of *prediction*. For instance, the recognition of a 4-3 suspension in a tonal work might cause such a listener to specifically predict that the dissonant fourth above the bass will resolve down by step to the consonant chord third. In general, expectation exists at the level of consciousness.



Figure 5-4: TENSION-AND-RELEASE schema I

This structural feature of the image schema itself does not entail that every perceived pattern (or cycle) of tension and release need be complete; for example, building in tension *and* relaxing that tension. Bregman notes: "Because many of the patterns that schema[ta] look for extend over time, when part of the evidence is present and the schema is activated, it can prepare the perceptual process for the remainder of the pattern."⁶¹ There are plenty of pieces in which the buildup of tension (often within multiple parameters of the music) is perpetual, sustained, and unreleased. Perhaps the release of tension is "filled-in" by the listener after the double bar, when all musical sounds have ceased; nevertheless, such works produce an aesthetic effect that is very different from that produced by works in which tension subsides and dissipates prior to the double bar. Image schemata can be partially activated and need not be fully realized during the musical experience. The incomplete fleshing-out of an image schema might correspond to the listener's unfulfilled expectations.

Our bodily experience of tension and release gives rise to the T&R schema. This schema, in turn, structures our reception of abstract musical events. Many musical gestures, such as a dynamic or textural crescendo, are perceived as tensing gestures. Our aesthetic response to such gestures is structured by our embodied knowledge of tension and release. This embodied knowledge is formed in large part by the recurrent experience of flexing and relaxing muscles within our bodies. When we flex a muscle, the muscle becomes larger. We often conceive of

⁶¹ Bregman, 666. Bregman refers to this process as "schema-governed attention."

musical events as having spatial properties such as: large or small, high or low, thick or thin. Thus, an increase in dynamic "level" or textural "density" is equated with an increase in size.⁶² Furthermore, as Andrew Mead observes, it is quite common for a listener to flex various muscles unconsciously while experiencing musical tension and to relax physically as the tension in the music is released.⁶³ The very bodily functions that initially serve to inform our experience of, and reaction to, musical events are then triggered and activated during the musical experience—a reciprocity that strengthens the argument that embodied schemata structure our basic response to music.

It is possible to conceive of the T&R schema as an elaborated version of the FORCE schema. Force is an omnipresent experiential reality—human beings constantly use force to accomplish tasks that allow us to survive and flourish. External forces constantly act upon us, compelling us to resist (by applying counterforce) or succumb. Johnson represents the FORCE schema as a simple arrow, indicating that a force of some (presumably constant) intensity is applied in some direction for some amount of time (see Figure 5-5).

Figure 5-5: FORCE schema⁶⁴

When we apply the FORCE schema to music, the directionality of the force in physical space is supplanted by diachronic time. For our purposes, this schema is elaborated in the sense that the intensity of the force is variable over time. There is a close correspondence between

 ⁶² Note that the terms "level" and "density" are metaphors employed from the domain of physical space to describe musical events. I argue, following Johnson, that such linguistic metaphors *result from* our initial reception of such musical events; they are *not* merely applied after the fact to explain musical phenomena. See Johnson, 104–106.
 ⁶³ Andrew Mead, "Bodily Hearing: Physiological Metaphors and Musical Understanding," *Journal of Music Theory* 43.1 (1999): 1–19.

⁶⁴ Taken from Johnson, p. 2, Figure 1.

force and tension: any musical gesture perceived as increasing in tension could also be construed as increasing the intensity of an abstract force. As I add more and more books to my briefcase, I experience an increase in the force of gravity.⁶⁵ Greater muscular tension is then required for me to counter this increased force and lift my briefcase. A FORCE schema with periodic variance of intensity could be depicted by the familiar graphic representation of a sine wave (see Figure 5-6).



Figure 5-6: FORCE schema as sine wave

In this case, the amplitude (y-axis) represents the intensity of the force and the frequency (x-axis) represents time. In Figure 5-6, the rate of variance of the force intensity over time is regular and cyclic; however, within our embodied understanding, the ebb and flow of force intensity is frequently aperiodic, and the magnitude of force is often variable from one cycle to the next. Similarly, when we perceive of force-intensity fluctuations within various musical parameters, they will often not display periodicity. Therefore, Figure 5-7 provides a better illustration of what such an elaborated FORCE schema *might* look like. Tension and Force share in a complementary, reciprocal relationship: either may be the cause *or* the effect of the other. An increase in force intensity can produce an increase in tension, and tension can enable the exertion

⁶⁵ The force of gravity acting on an object is equal to the weight of that object (on earth).

of force.⁶⁶ Therefore, force and tension are really two sides of the same coin; the Elaborated FORCE schema and the T&R schema are virtually identical. Compare Figures 5-7 and 5-8—the only difference between them is the labeling of the y-axis.⁶⁷



Time

Figure 5-7: Elaborated FORCE schema



Figure 5-8: T&R schema II

⁶⁶ Perhaps this particular cause-and-effect duality is somehow active during musical reception. Might the listener conceive of an abstract "musical force" as an agent or as the product of an agent (abiding *within* the music itself) that not only propels the music forward through time but also produces tension and release progressions within various parameters of the music?

⁶⁷ *Tensional force* is the force that is transmitted through a string, rope, or wire when it is pulled tight by forces acting from each end *(in opposite directions)*. The tensional force is directed along the wire and pulls equally on the objects on either end of the wire. This type of tension *is* a specific kind of force; however, perceived forces/tensing gestures in music seem almost always to be unidirectional (in the direction of time).

We must bear in mind that such figures are only useful in portraying the structure and innerworkings of the schemata. While I am discussing and representing this schema propositionally, it is in fact a nonpropositional mental structure of understanding and meaning. It doesn't actually "look like" anything. Many schemata—including some of those to be discussed below—are difficult or impossible to represent with a concise visual figure.

What types of musical events are received and structured by the T&R schema? Above I discussed the means by which a dynamic crescendo and an increase in textural density can each be perceived as a *tensing gesture* through our embodied knowledge of muscular flexion. Generally, an increase of *something* within a parameter of music is received in terms of tension. Another common example is an increase in tempo within the domain of musical time. We know from experience that the kinetic acceleration of our bodies (or of any other physical object) through space requires greater muscular exertion and expenditure of energy. Therefore, when we encounter an accelerando, we might unconsciously understand that an abstract force of increasing intensity is acting upon the music, causing it to "speed-up." Conversely, the ritardando that frequently appears at the end of a musical phrase is readily perceived in terms of a physical object in motion that is gradually slowing down to a state of rest. If I stop peddling my bicycle (ceasing the exertion of muscular force), I will eventually slow to a stop (assuming that I am on a flat surface).⁶⁸ Thus, an accelerando is a tensing gesture and a ritardando is (generally) a releasing gesture. The application of *rubato* is a more subtle manifestation of a recurrent tension and release cycle (or, "progression") within the domain of musical time.

In the realm of pitch, the perpetual alternation of dissonance and consonance provides for the perception of tension and release progressions. When a dissonant tritone "caves-in" to a consonant major third, a certain degree of tension has been released because an unstable ratio of

⁶⁸ It is the *friction force* between my two tires and the pavement that causes me to decelerate.

frequencies has "resolved" to a more stable ratio, and the vibrations of the basilar membrane (within the cochlea of the inner ear) have been simplified. On the surface of tonal works, the presence of a dissonant nonchord tone (such as a passing tone or suspension) establishes a local tension that requires resolution to a consonant chord tone. Adding a seventh to a dominant chord creates dissonance (a tritone between the chord third and seventh), thereby increasing perceptual tension and strengthening the "pull" to the tonic chord, which will provide a release of tension when ultimately achieved. Dissonance in common-practice tonality is more than just a learned convention—it is a psychoacoustic phenomenon of sensation. In fact, some hold the latter to be responsible for the former. Lerdahl reminds us that when Schoenberg succeeded in "emancipating the dissonance," he emancipated it from its conventional and functional role in music—not from its sensory, psychoacoustic effect.⁶⁹

In the domain of melody, an ascending melodic contour elicits tension. Our human experience on the planet earth involves constant (and unconscious) negotiation with the force of gravity. We know that to work against this force—as when we jump up off of the ground and into the air—requires a counterforce in the form of physical exertion, and of course, physical exertion involves muscular flexing. Therefore, an ascending melodic line may be received in terms of a force that is countering gravity.⁷⁰ Conversely, a descending melodic line is akin to an object "falling" downward to a natural state of rest.⁷¹ Hatten notes: "Meter and tonality each afford analogies to gravitation or dynamic vectoral space, making possible the experience of *embodied motion* subject to dynamics and constraints comparable to those affecting the body in a

⁶⁹ Lerdahl, 380.

⁷⁰ In addition, we know that to vocalize in the upper tessitura of our vocal range requires greater physical tension, which may result in "strained" vocal chords.

⁷¹ A rising melodic line, as a tensing motion, correlates with Hatten's "lift gesture." A descending melodic line is compatible with Hatten's "subsiding gesture," which typically follows a lift gesture. See Hatten, *Interpreting Musical Gestures, Topics, and Tropes*, 157–163.

natural environment."⁷² In other words, the conceptual musical space(s) forged by (tonal) pitch space and musical time is not only structured by our bodily experience in actual space, but the conceptual spaces likewise enable our experience and understanding of musical motion (as it occurs within those spaces). Hatten elaborates: "Much of the meaning we experience in our first encounter with a melodic shape relates to its modality of action, which traces the shape of a lived experience as we *move through* the tones."⁷³ Hatten would most likely claim that the melodic shape functions as an *index* to *signify* a lived experience. I argue that the melodic shape is automatically received and structured via the metaphoric projection of image schemata such as the T&R schema. I therefore agree with Hatten's above statement—I simply differ regarding the cognitive process(es) by which musical activities are mapped onto (and structured in terms of) bodily ones.

An additional schema that is responsible for actively structuring our perception of tension in the realm of melodic contour is the VERTICALITY schema (see Figure 5-9). This schema enables the listener to understand musical pitches as having relative heights, via the process of cross-domain mapping (to be explored later in this chapter).



Figure 5-9: VERTICALITY schema⁷⁴

 ⁷² Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 124.
 ⁷³ Ibid., 203. Emphasis Original.

⁷⁴ Taken from Zbikowski, *Conceptualizing Music*, p. 69, Figure 2.1.

Without the VERTICALITY schema and the conceptual metaphor that it "grounds" (PITCH RELATIONSHIPS ARE OBJECT RELATIONSHIPS IN VERTICAL SPACE), we would not perceive an ascending scale as a succession of pitch-objects ascending in "space." Since the VERTICALITY schema allows us to understand musical pitches as having relative heights (by allowing us to conceptualize "height" in the first place), it is functionally requisite to the fleshing-out of the T&R schema in the domain of melodic contour.⁷⁵

Melodic tension may also be construed in terms of the *attractive* tendencies of tones. Johnson acknowledges the ATTRACTION schema as one of the basic image schemata, but he does not expound upon its structure.⁷⁶ Lerdahl states, "The greater the proximity [of two tones], the greater the attraction...the attraction to the stable tone is inversely proportional to its distance."⁷⁷ When a tone (such as a non-chord tone) is strongly attracted to another tone (a chord tone), a certain type of tension results, just as a magnetic attraction between two oppositely charged objects creates a very real physical tension between them. This is why a tonal work will not typically end on an unstable pitch such as the leading tone: the perceived attraction to the tonic pitch is too strong, and this attraction results in tension. The same factors of attraction in a melody apply to chord progressions as well, in which "the voices in a chord tend toward their attractors."⁷⁸ Perhaps further proof of the attractive tendencies of tones lies in the frequent intonation discrepancy between enharmonic pitches such as F-sharp and G-flat. These two tones are enharmonically equivalent but not functionally equivalent within a tonal framework. As rendered by instruments with indeterminate pitch, such as the voice or a fretless bowed string

⁷⁵ The fact that Western listeners conceive of pitches as being high or low is indeed a convention of our culture (there are cultures that do not perceive pitches in terms of height). However, this convention is based on a conceptual metaphor, and the metaphor is grounded by our everyday experience with up and down. The conceptualization of pitches in terms of verticality allows for the activation of the T&R schema in the realm of pitch relations.

⁷⁶ Johnson, 126.

⁷⁷ Lerdahl, 163.

⁷⁸ Ibid., 173.

instrument, an F-sharp tending toward G will likely be slightly sharper than a G-flat tending toward F; it is as if the note is "pulled" sharp or flat by its semitone attractor.⁷⁹

In the realm of activity within the physical space occupied by an ensemble, an accumulative spatial gesture will educe tension. As mentioned above, the increase (or magnification) of a musical quality (such as dynamic level, tempo, or texture) is associated with an increase in tension. With an accumulative gesture, the sonic space (the portion of the ensemble space that is actively producing sound) gradually expands—potentially filling the entire ensemble space. Conversely, a dissipative gesture—in which the sonic space systematically contracts—is relaxive. However, a release of tension might coincide with the completion of an accumulative gesture through the fleshing-out of a different image schema. The expansion of the sonic space, though immediately tensive due to the increasing size of active ensemble points, likewise produces tension through the instigation and actualization of a process. The PROCESS schema structures the "process" of accumulation that enables the enlargement of the sonic space.⁸⁰ Bound to the PROCESS schema is Meyer's conception of expectation. As the sonic space expands, we might expect more and more for it to spread into the total amount of space "afforded" by the ensemble space. If this expectation is indeed satisfied, a sense of release is furnished by the completion and fulfillment of a dynamic process.⁸¹

⁷⁹ Steve Larson applies the concept of magnetic attraction (wherein the attractive force of a physical magnet increases as its propinquity to a metal object increases) to the individual chord tones of the tonic triad (scale degrees 1, 3, and 5) within the diatonic scale. These stable tones exert a "magnetic pull" on their unstable neighbors. Larson describes the attractive force of these tones as being strong or weak in either direction in pitch space, depending upon the distance to the unstable neighbor: a half-step distance (as exists between scale degrees 1 and 7 as well as 3 and 4) is a strong force, while a whole-step distance (between 1 and 2, 3 and 2, 5 and 6, and 5 and 4) is a weak force—but a perceived force nonetheless. (The distribution of strong and weak forces between stable and unstable tones differs in the minor mode.) See: Steve Larson, "Scale-Degree Function: A Theory of Expressive Meaning and Its Application to Aural-Skills Pedagogy," Journal of Music Theory Pedagogy 20 (1993): 69-84; Steve Larson, "Musical Forces and Melodic Patterns," *Theory and Practice* 22 (1997): 55–72. ⁸⁰ Johnson notes the existence of the PROCESS schema but says little about it. See Johnson, 126.

⁸¹ The commingling of the T&R schema and the PROCESS schema exemplifies Johnson's conception of the "superimposition of image-schematic structures" (Johnson, 125).
Similarly, certain transformative procedures that relate two or more spatial gestures may be tensive. Alterations such as *expansion, extension*, and *inflation* (outlined in Chapter 4) that result in a spatially enlarged form of a previously occurring spatial gesture evoke tension in much the same manner as does a single, accumulative gesture. Furthermore, an *accelerated* gestural form (which must follow a related gesture with a slower point-shift rate to be recognized as accelerated) evokes tension, as does a single gesture undergoing *intragestural acceleration*. Divergent mirrored polystreaming might be deemed tensive, since two gestures effectively work against a potential force of attraction (their diverging "fronts" likewise expanding the sonic space), while convergent polystreaming is relaxive for the opposite reason: two gestures succumb to the force of attraction and merge together (the sonic space framed by their fronts contracting during the process).

At this point, it is important that we remind ourselves that all of the above illustrations pertaining to the T&R schema (gravitational pull, attraction, expansion) are more than just metaphorical expressions devised and employed to describe musical activity. Rather, they are but a few examples of countless related, recurrent, and real experiences that we have with our bodies within our environment that conspire to form this particular schema—a schema that ultimately structures how we receive and "understand" music at a basic level. It is not that we intentionally structure our understanding of music in explicit terms of specific bodily experience—this would imply conscious, deductive, and a priori reasoning with/from an innate conceptual metaphor. Metaphor is embodied, acquired, and a posteriori. Repeated bodily experience gives rise to the T&R schema, which structures our continuous sensorimotor experience as well as our understanding of abstract phenomena such as music. The metaphoric projection of the T&R schema occurs *prior* to our thinking and talking of music in bodily terms;

271

the conceptual metaphor is secondary to the primary response.⁸² Flexing our muscles, or slowing-down on a bicycle, or jumping into the air—these are all examples of common experiences that enable our basic response to music; these are *not* just linguistic metaphors that we bring into play to reason about musical phenomena. Furthermore, the primary realization of the T&R schema does not obligate the subsequent construction of conceptual metaphors by which we can then rationalize and converse about music. Someone with no interest in describing the experience of music might never reason or talk about music in terms of bodily experience; and yet, his/her embodied knowledge undoubtedly shaped his/her initial reception of the music. In the subsequent discussion of other relevant schemata, I shall reveal how they often coincide with the T&R schema.

The DIALOGUE schema

This schema could be called alternatively the CALL-AND-RESPONSE schema, the QUESTION-AND-ANSWER schema, or the CONVERSATION schema. The idea of two distinct entities participating in alternating and interactive activity is summed-up by all of the above descriptors. The social act of engaging in conversational exchange with another person is an omnipresent experience. Furthermore, witnessing a dialogue between two different people is an extremely common occurrence. Many musical styles feature call-and-response activity between musical voices and/or instruments. Hatten categorizes such musical activity as "dialogical gestures," which he defines as "gestures between agencies, or within a single agency...that appear to respond to each other...In the Classical style, [this term is] suggestive of a conversation among equals (string quartets incorporating Haydn's new style of thematic

⁸² The formation of an image schema results from *bottom-up*, or "data-driven" processing. The "data" is the incoming sensory stimuli received during bodily experience. The metaphoric projection of a schema is a type of *top-down*, or "concept-driven" processing.

counterpoint), or oppositional ideas (thematic dialectics), or oppositions between individuals and larger groups (concertato effects).³³ The spatial gesture emerging from the musical conversation between two sound sources is the oscillation gesture. Due to the early spatial practices of responsorial and antiphonal psalmody, the oscillation gesture is perhaps the most ancient kind of spatial gesture, and it is structured by one of the most socio-culturally enforced schemata, the DIALOGUE schema. This gesture—arising from antiphonal (call-and-response) musical practices—continues to pervade Western art music even today. In Ives' The Unanswered Question (discussed in Chapter 1), the DIALOGUE schema structures the musical call-and-response between the solo trumpet (the "questioner") and the woodwinds (the would-be "answerers"). As the question-and-answer cycle unfolds, tension builds as the woodwind's responses grow increasingly agitated. The climactic point of tension follows the trumpet's final posing of the question, when the listener expects (on the basis of foregoing musical behavior) another response from the woodwinds. This expectation, however, is foiled as the question is answered only by "silence" (from the remote and muted strings). As this musical dialogue ensues, so too does an oscillation gesture emerge between the winds and the trumpet. Since the dialogue begins and ends with the trumpet's query, the oscillation gesture is strong (i.e., spatially closed).

While ubiquitous in Western art music, call-and-response musical activity is likewise common to many different cultures, as well as to different musical styles within those cultures. Many popular Western styles feature oscillation gestures that arise from a call-and-response texture, including folk, blues, gospel, rhythm & blues, and rock. The universality of this particular gesture may be attributed to the metaphorical projection of the DIALOGUE schema

⁸³ Hatten, Interpreting Musical Gestures, Topics, and Tropes, 136, 164.

onto musical events. The DIALOGUE schema itself is universal, owing to the fact that some form of spoken language is employed as a communicative tool within all world cultures.

The PATH Schema

Above I introduced the PATH schema to exemplify the concept of an image schema.⁸⁴ Johnson states: "The PATH schema is one of the most common structures that emerges from our constant bodily functioning."⁸⁵ Following Lakoff and Johnson's later appellation of this schema, Candace Brower refers to the PATH schema as the "SOURCE-PATH-GOAL schema," which encapsulates Johnson's point that a path "always [consists of] the same [three] parts: (1) a source, or starting point; (2) a goal, or end-point; and (3) a sequence of contiguous locations connecting the source with the goal."⁸⁶ Johnson occasionally refers to the PATH schema as the "FROM-TO schema."⁸⁷ (I shall simply use Johnson's original term, "PATH schema.") Brower expands Johnson's conception of the PATH schema by adding "trajectory of motion" and "the force propelling the motion" to the components of the schema.⁸⁸ These prove to be fitting additions; after all, the traversal of a path involves directed motion (either real or metaphorical) along that path, and a body in motion requires a force (or, agent) of some kind to propel it through space.

The notion of traversing a path applies to many musical processes. A complete work (or movement) of music has a beginning, an ending, and some length of time separating these two temporal endpoints. Thus, a musical work as a whole traces a path through time. Many

⁸⁴ Remember that specific image schemata themselves are *not* concepts: they are preconceptual mental structures. ⁸⁵ Johnson, 116.

⁸⁶ Brower, 331–332; Johnson, 113. Lakoff (*Women, Fire, and Dangerous Things*) and Lakoff and Johnson (*Philosophy in the Flesh*) employ the term SOUCE-PATH-GOAL schema.

⁸⁷ Johnson, 28.

⁸⁸ Brower, 331.

compositional techniques create the perceptual illusion of timelessness and thereby appear to halt the path-progress of the music.⁸⁹ A melody may be construed as a path: an abstract musical force (or agent) propels a tone through "melodic pathways" in pitch space. The temporal component of the melody provides its forward-propulsion (or, momentum), but the vector of the path may descend or ascend through pitch space. A harmonic sequence or progression (in a tonal context) may likewise travel along a path, especially if the progression or sequence leads from one function to another, as in: I-vi-IV-ii-vii^o-V. In this descending-thirds sequence, tonic function (I-vi) leads to predominant function (IV-ii), which, in turn, leads to dominant function (vii^o-V). I is the "source," and V the ultimate "goal," of this "chordal path."

In his introduction to *Generalized Musical Intervals and Transformations* (1987), David Lewin presents a version of the PATH schema as his opening figure (see Figure 5-10).



Figure 5-10: Lewin's Figure 0.1 (GMIT)⁹⁰

Lewin does not explicitly evoke image-schema theory, but he uses this diagram to formulate his broad conception of *interval*, where i is "the interval from s to t."⁹¹ Lewin's second chapter

⁸⁹ Examples of procedures resulting in "musical timelessness" include: harmonic stasis (in which functional harmonic progression is suspended) such as commonly found in Anton Bruckner's Symphonies, minimalist techniques (overtly sustained repetition), and the incorporation of symmetrical pitch collections, rhythms, formal units, etc. Olivier Messiaen's "modes of limited transposition" and "nonretrogradable rhythms" provide good examples of musical symmetry in the domains of both pitch space and time.

⁹⁰ Taken from Lewin, Generalized Musical Intervals and Transformations, p. xi, Figure 0.1

⁹¹ Lewin, Generalized Musical Intervals and Transformations, xi-xii.

explores twelve different examples of conceptual (or, "intuitive") musical spaces to which this figure applies. In six of these examples, points s and t are pitches or pitch classes and i is the pitch-space distance between those points in either a melodic or a harmonic context. In the other six examples, s and t are measured rhythmic points and i is the temporal duration between them. Lewin states that the interval (i) measures our "*intuitions* of 'distance' or 'motion' [and] these intuitions are highly conditioned by cultural factors."⁹² We may use the term interval to refer to the physical distance between any two points within ensemble space, be they contiguous or nonadjacent. While it might prove cumbersome and unwieldy to measure the exact distance between adjacent points, such that within a four-point ensemble space, i(0-1) = i(1-2) = i(2-3), i(0-2) = i(1-3), i(0-2) = [i(0-1) + i(1-2)], etc.⁹³

In many of Lewin's complex transformational networks, he evokes the concept of a path that navigates the network to account for sequences of related musical entities. Lewin states: "We can view the chronological progress of the piece as a path through the network." The network itself is synchronic—revealing all potential and realized transformational relationships. The progress of the music, as it traces a path through the network, is diachronic. In his analysis of Stockhausen's *Klavierstuck III*, Lewin states: "Rather than asserting a network that follows pentachord relations one at a time, according to the chronology of the piece, I shall assert instead a network that displays all the pentachord forms used and all their *potentially* functional interrelationships, in a compactly organized little spatial configuration."⁹⁴ The resultant network

⁹² Ibid., 17. Emphasis Original.

⁹³ The exact distance between points may vary considerably from one performance to the next (even in successive performances by the same ensemble). However, ensembles (chamber groups in particular) are generally arranged in a symmetrically balanced fashion within the performance space.

⁹⁴ David Lewin, *Musical Form and Transformation: Four Analytic Essays* (New Haven: Yale University Press, 1993), 17. Emphasis mine.

of "potentially functional interrelationships" represents—in essence—the *transformational affordances* offered by the [01236] pentachord exposited in the opening bar of this short work. The network, therefore, is a category-container of transformationally related musical entities. It is doubtful that the listener is consciously aware of (or entertaining the prospect of encountering) each and every one of these affordances, but they exist as potentialities nonetheless. Since the listener most likely does expect *some* kind of variation to occur, the network represents explicitly the transformational potentials implicitly assumed (and vaguely expected) by the listener.

We might apply this format to gestural categories, as delineated in the previous section (5.2) dealing with categorization. For a given work of music, we might construct a containermodel in which the principal gestural form is the local categorical prototype positioned at the center of the category-container. All other spatial gestures radiate away from the central gesture and toward the peripheral boundaries of the container with the least typical (i.e., least similar) gestures located farthest from the principal gesture. If multiple related forms of a spatial gesture occur throughout a work of music, we may think of a "gestural progression" as a navigated *path* through a gestural container. The path may take us to the edge of the container as it explores the most remote categorical forms, and it may traverse categorical boundaries to account for intercategorical relations.

There exists an even more obvious application of the PATH schema to spatial gestures (to open gestures in particular): a spatial gesture *is* a path woven through ensemble space.⁹⁵ A sound event is propelled through "gestural pathways" in ensemble space. The path will have

⁹⁵ A comparison of events in actual space to those in metaphoric space is telling: a melody follows a path through pitch space, and a spatial gesture follows a path through ensemble space. Each degree of a scale is similar to a point in ensemble space. The seven pitches of the diatonic scale (from which tonal melody and harmony are generated) are conceptually present, particularly in the minds of experienced listeners. The ensemble space is *actually* present at all times. Multiple paths may be woven through either of these spaces based on the chosen succession of scale degrees *or* ensemble points.

both a beginning point (source) and an ending point (goal), but it will also have points (intervals) between these spatio-temporal boundaries—the "sequence of contiguous locations connecting the source with the goal."96 In a four-point ensemble space, points 0 and 3 are the spatial extremes of the ensemble. With SG-[3210], points 0 and 3 likewise serve as the spatio-temporal endpoints of the gestural path beginning at point 3 and concluding at point 0. Point 3 functions as the "source" and point 0 functions as the "goal." The path traverses interior ensemble points (and, in this case, event points) 2 and 1.97 Under ordinary performance conditions, the ensemble would be arranged such that the physical distance between each adjacent performer (ensemble point) would be approximately equal. The equidistance of neighboring ensemble points ensures a uniform interval between each point shift of SG-[3210]: PS 3-2, PS 2-1, and PS 1-0 each span the same distance. This equivalence of distance between consecutive point shifts contributes to the "well-formedness" of this gesture. SG-|320|, an abbreviation of SG-|3210|, features a disjunction that causes a discrepancy between point-shift interval sizes: PS 2-0 is twice the length of PS 3-2. Thus, SG-[320] is less "well-formed" than SG-[3210]; the former gesture is classified as "imperfect" while the latter is "perfect" (as described in Chapter 3). It will oftenbut not always—be appropriate for the spatial extremes of the ensemble space to serve as the "source" and "goal" of a spatio-gestural path (as is the case with open gestures).

Other spatial gestures likewise traverse a path through the ensemble space. SG-|02134| (in five-point ensemble space) ultimately spans the entirety of the ensemble space, albeit not in a

⁹⁶ Johnson, 113.

⁹⁷ Perfect panning gestures will contribute to the sense of "path traversal" with increasing effectiveness as the total ensemble-space cardinality increases. SG-[76543210] (in eight-point ensemble space) sketches a more pronounced path than SG-[210] (in three-point ensemble space) owing to the fact SG-[76543210] contains more interior ensemble points, so the uniform right-to-left directionality is accentuated by a greater number of successive point shifts. Furthermore, the eight-point ensemble would likely (but not necessarily) occupy more physical space than the three-point ensemble, so the spatial span of the gesture would be greater and thereby more evident. In general, the salience of a spatial gesture increases as the space inhabited by an ensemble increases, since the distinct spatial locatedness of each point is more pronounced (aurally as well as visually).

unidirectional and strictly conjunct fashion. The spatial progress along the ensemble-space path is hesitant and uncertain in this weak open gesture; however, the endpoints of the ensemble space do function as source/goal. SG-|43012| is a weak closed gesture that traverses the total ensemble space with subgesture SG-|430| and then "backtracks" along the path with subgesture SG-|012|. Two of Brower's entailments of the PATH schema are that "motion may or may not follow a path leading to the goal" and "the endpoint of motion may or may not coincide with the goal."⁹⁸ In this case, we might view SG-|430| as completing the right-to-left path by reaching its goal (point 0), while the endpoint of SG-|012| (ensemble point 2) does *not* coincide with the goal (ensemble point 4)—the projected "end" of the path. The failure of the gesture to reach the intended goal results in unresolved spatial tension. Other forms of weak closed gesture, such as spiral gestures, trace orbital (or, "zigzag") paths toward or away from the central ensemble point (or virtual center) of the ensemble space.

Brower's eighth entailment of the PATH schema is stated as follows: "The approach to a goal tends to be accompanied by an increase in tension and arrival at a goal by relaxation and the slowing and/or stopping of motion."⁹⁹ Thus, the T&R schema and the PATH schema are inextricable—so long as the path's projected goal is defined and understood by the perceiver.¹⁰⁰ Often, the ensemble-point goal of a spatial gesture is *not* foreseen and is only determinable ex post facto. However, certain gestural forms (such as strong open gestures) are better structured by the PATH schema than are other gestures. Through our embodied experience with the force of inertia, we likely predict that a panning gesture will indeed proceed to its termination at the opposite end of the ensemble from which it commenced. *Inertia* is a property of matter by which

⁹⁸ Brower, 331.

⁹⁹ Ibid.

¹⁰⁰ The PROCESS schema is likewise bound to the PATH schema: the traversal of a path (actual or metaphorical) is a singular process (or event) that entails change (of spatial location) over time. Setting out along a path initiates the *process* of reaching the goal of that path.

it continues in its existing state of rest or uniform motion in a straight line, unless that motion is changed by an external force. Steve Larson applies the concept of inertia to melodic forces.¹⁰¹ With a strong open gesture, we might predict future events (continued gestural behavior) based on our observance of past events. So, the latter point shifts completing a L-R PAN are foreseen on the basis of the gesture's initial left-to-right point shifts. Following the commencement of a salient L-R PAN, tension mounts and expectation builds as we anticipate its arrival at the rightmost spatial extreme (goal) of the ensemble space.¹⁰² This encroaching tension may or may not be released—and our expectations satisfied—by the gesture's envisaged conclusion at the rightmost ensemble point.

The DEPARTURE-AND-RETURN schema

The PATH schema involves the departure from a source and ultimate arrival at a goal, such that the sense of initial departure is—at least locally—sustained. As applied to spatial gestures, the goal will be an ensemble point other than the point from which the gesture commenced, such that large-scale point-to-point locomotion occurs within the ensemble space. Often, however, a spatial gesture will conclude at the same point from which it initially began: the "source" of the path will be returned to, and the gesture will comprise a spatial conflation of source and goal (although the two will still be temporally distinct). Such spatial activity is structured by the DEPARTURE-AND-RETURN schema (henceforth the D&R schema). This embodied schema is forged by our continuous experience of departing from and ultimately returning to a physical location.

¹⁰¹ Steve Larson, "Scale-Degree Function," 69–84.

¹⁰² Any such expectation would likely be compounded by having encountered left-to-right panning gestures earlier in the musical work being experienced.

The D&R schema will often comprise two distinct PATH schemata that share in an inverse relationship: the latter-occurring path may trace the former path in reverse. These two paths may be experientially separated by a long length of time and by innumerable intervening events. When I travel from Atlanta to Beijing-unless I intend to change residence permanently—I will eventually return to Atlanta. I might stay in Beijing for two months; all the while, the awareness of my inevitable return to Atlanta suspends the fleshing-out of the D&R schema at a point where the goal of one path is identical to the source of a later path. My return flight might involve connecting flights in cities different from those involved in my departure flight. Nevertheless, the goal (point B) of my departure flight serves as the source of my return flight, and the goal of my return flight is the initial source of the departure flight (point A). Thus, the D&R schema is a gestalt structure consisting of two distinct, but related, PATH schemata. In Figure 5-11, the two PATH schemata forming the D&R schema are slightly curved to facilitate the diagramming of this image schema. This curvature does not necessarily depict a difference in path topography between the "departure path" (A \rightarrow B) and the "return path" (B \rightarrow A). Point A serves as the initial point of departure and goal of the return path. Point B serves as the goal of the departure path and as the point of departure for the return path, which ultimately leads back to point A. Either the departure path or the return path may feature intermediate points (intermediary goals) between their respective source- and goal-points.



Figure 5-11: DEPARTURE-AND-RETURN (D&R) schema

Musical events organized by the D&R schema are numerous: a non-chord tone neighboring away from and back to a single pitch, a harmonic sequence that prolongs a single chord (or function), a tonal work that modulates away from and eventually back to the initial key, and forms such as Ternary and Sonata (among others) all involve departure-and-return scenarios within different musical parameters (melodic, harmonic, tonal, and sectional). In considering spatial activity, strong closed gestures are structured by the D&R schema, since, by definition, they begin and end at the same point in ensemble space. With SG-[0214310], ensemble point 0 functions as the starting and ending point of the gesture. Perfect sweeping gestures are clear examples of departure-and-return, owing to the fact that they comprise two "linked" panning gestures related by either retrograde or inversion. Each panning gesture (as a component subgesture) is structured by the PATH schema, and the two gestures together (constituting a compound strong closed gesture) are structured by the D&R schema. SG-|012343210| (perfect LEFT SWEEP) contains SG-|01234| (perfect L-R PAN) and SG-|43210| (perfect R-L PAN). The hinge of this sweeping gesture, point 4, is the goal of L-R PAN and the source of R-L PAN, while point 0 serves as the source of L-R PAN and the goal of R-L PAN (point 0 is both the source and goal of the sweeping gesture as a whole). Brower writes: "From the taking of a single step, to the execution of a series of steps from one location to another, to the completion of an entire journey involving departure and return, each completed motion takes on meaning in relation to the whole."¹⁰³ Analogically, the "taking of a single step" corresponds to a point shift within a larger gesture. The "series of steps from one location to another" maps onto the idea of an open (or weak closed) gesture that is structured by the PATH schema. The "completion of an entire journey involving departure and return" correlates with a strong closed gesture such as a sweeping gesture. In her statement, Brower has effectively highlighted the

¹⁰³ Brower, 332.

gestalt structure of spatial gestures in general: that large motions (such as SWEEPs) contain smaller ones (such as PANs), and small motions contain smaller motions still (subgestures and point shifts). Most significantly, the "meaning" of the largest motion is determined by the structural relationships of its component, smaller motions.

The DEPARTURE-AND-RETURN schema is intimately bound with the TENSION-AND-RELEASE schema. *Departure* establishes a *tension* that is *released* upon *return*. We tend to expect returns and the attendant release of tension following a departure. In tonal works, a departure from the tonic pitch, tonic chord, or initial key area creates a tension that is released upon returning either to scale-degree 1, the tonic chord, or the primary key area. The return following a departure provides a sense of closure. Spatially, strong *closed* gestures most effectively actualize the *closure* provided by the departure-and-return process.

The CYCLE schema

Closely related to the D&R schema is the CYCLE schema. Johnson summarizes our embodied understanding of cycles:

We come into existence as the culmination of a reproductive cycle. Our bodily maintenance depends upon the regular recurrence of complex interacting cycles: heartbeat, breathing, digestion, menstruation, waking and sleeping, circulation, emotional build-up followed by release, etc. We experience our world and everything in it as embedded within cyclic processes: day and night, the seasons...[and] the revolutions of the heavenly bodies.¹⁰⁴

A cycle is essentially a pattern that is repeated. Johnson represents the CYCLE schema as a circle (see Figure 5-12), which indicates the continuous repetition of a pattern and underscores the fact that we naturally conceptualize cycles as being circular in nature.

¹⁰⁴ Johnson, 119.



Figure 5-12: CYCLE schema¹⁰⁵

Our conceptualization of time is cyclic, with multiple nested cycles producing a temporal hierarchy: 60 seconds = 1 minute; 60 minutes = 1 hour; 24 hours = 1 day; 7 days = 1 week; approximately 4 weeks = 1 month; 12 months = 1 year; 10 years = 1 decade; 10 decades = 1 century; 10 centuries = 1 millennium, and so on. Even the structure of an analog clock provides the visual image of time progressing in a circular fashion as the minute hand is trailed by the hour hand in clockwise motion.

The CYCLE schema is related to the D&R schema outlined above, in that many traditional musical forms (including Ternary, Sonata, Rondo) feature the structural repetition of large sections: the structural "return" of a section is a cyclic process. The CYCLE schema is likewise realized in musical works purported to be "cyclic," such as Beethoven's *Symphony no. 9* and Berlioz's *Symphonie Fantastique*. A cyclic work is one in which themes or sections recur in multiple movements. Such works often feature the process of thematic transformation, in which themes are transformed so as to fit into new musical environments (thematic transformation is structured by the OBJECT MANIPULATION schema, to be outlined below). After all, the repetitions encompassed by a recurring cycle are seldom perfect repetitions: although every week

¹⁰⁵ Taken (with modification) from Johnson, p. 120, Figure 22. Johnson's figure features more clockwise-facing arrows within the circle.

of the year will feature a Wednesday, every Wednesday will be unique in terms of the events contained therein. Although the cycle schema features repetition on some level, it is likewise bound to the concept of diachronic progression. The CYCLE schema, therefore, is realized in works of music that feature either exact or nearly exact repetition of previous musical material. Many cyclic musical processes are best described as altered repetition—musical material recurs, but often does so in a transformed (abbreviated, transposed, etc.) state.

One of the compound spatial gestures identified in Chapter 3 is the *cyclic panning gesture*, which is composed of two or more panning gestures (of the same directionality) in direct succession. The compound gesture's salience is directly proportional to its number of component panning gestures—the perception of a cycle increases as the number of cyclic repetitions likewise increases. Likewise, *oscillation gestures* feature the cyclic (regular or irregular) alternation of activity at two different points—the longer the cycle (i.e., the greater the number of successive alternations), the more salient the gesture. Another application of the CYCLE schema to spatial gestures occurs within a surround-sound environment, in which the audience is contained within the ensemble space. As described in Chapter 4, such a space exhibits the potential for the occurrence of *circular gestures*. Depending upon the relationship between a circular gesture's starting and ending points, the gesture may be closed (complete) or open (incomplete). A closed circular gesture exhibits Prägnanz in the form of spatial closure/completion.¹⁰⁶

In *Persephassa* by Iannis Xenakis, an ensemble of six percussionists encircles the audience, forming the outline of a hexagon (approximated as a circle). The ensemble space itself

¹⁰⁶ In this sense, the circular representation of a cycle (as in Johnson's figure for the CYCLE schema) reciprocally allows for the interpretation of circles as cycles. The circle is a spatial translation of a temporal concept. As such, the literal circle of the ensemble space, as well as the gestures that trace the circular shape of that space, *symbolize* the cycle.

consists of a circle, and Xenakis outlines the ensemble space with circular gestures of varying directionality. Figure 5-13 reveals the ensemble diagram provided in the score to *Persephassa*.



Figure 5-13: Placement of performers and instruments in Persephassa¹⁰⁷

By employing overlapping dynamic envelopes between the successive activities of adjacent ensemble points, Xenakis effectively creates the illusion of continuous sound movement. Many semicircles and circular fragments appear in the buildup to the work's climax, reflecting the process of "constructing a circle." The process reaches its goal in m. 352, when a slowly rotating tremolo on the drums articulates a perfect circle in a counterclockwise direction. This circular

¹⁰⁷ Iannis Xenakis, *Persephassa pour six percussionists* (Paris: Musique Contemporaine, 1970). Note the similarity of this diagram to that provided in Figure 2-2.

gesture repeats its circling (in cyclic fashion). Gradually, more circles enter into the mix: a clockwise circle enters in m. 354 (see Figure 5-14).



Figure 5-14: Xenakis, *Persephassa*, mm. 352–356. The arrows drawn over the score trace the paths of two circular gestures. The dashed line follows the initial counterclockwise gesture; the solid line depicts the clockwise gesture entering in m. 354.

Measures 354–356 present a case of *contrary polystreaming* that alternates between

divergent and convergent mirrored streaming. The two circles continually collide (or,

"converge") at points A and D.¹⁰⁸ Between m. 357 and m. 395, five additional circles join in the

¹⁰⁸ Since they are directly opposite one another in ensemble space, points A and D articulate the diameter of the circle, as do points B and E and points C and F. Throughout this work, *passing gestures* (which effectively "cut through" the circle) extend from one point in each "diameter point pair" to another. Although I am presently concerned with the circular gestures that outline the circumference of the ensemble space, *diametral passing gestures* are another type of gesture active in *Persephassa*. (Since they comprise only two points, these "gestures" are actually point shifts.)

spatial activity. Each one of the seven circles is individuated by virtue of a uniform percussive timbre at each of its event points. The clockwise circle entering in m. 354 is "metal." Cymbals render the circle entering at the pickup to m. 357. Thai gongs begin to circulate in m. 361, followed by wooden simantras (m. 369), tam-tams (m. 381), and wood blocks (m. 395)ultimately culminating in a mass of dissonant directionalities and seven-part polystreaming. However, the consistency of timbre at component event points differentiates each circular gesture within the polygestural composite. Some of the circular gestures conspire to yield consonant canonic streaming as they proceed in the same direction and "follow" one another around the circular ensemble space. However, many of the circular gestures enter with different rates of motion, resulting in *tracking stream overtake* (wherein a *tracking stream* catches-up to and "passes" a lead stream.) Such activity results in a reversal of streaming function: the tracking stream becomes the lead stream, and vice versa. Despite the embedded instances of canonic streaming in this frenzy of circular gestures, contrary streaming (divergent and convergent) is the prevailing form of polystreaming.¹⁰⁹ The increase in textural density provided by the introduction of each new circle is accompanied by a gradual accelerando and crescendo embracing the entire gestural multiplicity. Sirens erupt at each ensemble point in m. 410. Following this climax, the spatial activity subsides to a single layer of rapidly rotating drum tremolos (m. 430). This circular gesture is composed of timbrally heterogeneous percussive sounds, does not feature overlapping dynamic envelopes, and changes direction capriciouslypreventing circular closure at the close of the work (m. 430 to the double bar). A sense of

¹⁰⁹ The perpetual alternation of convergent and divergent contrary streaming between two circular gestures is itself a cyclic process, as is the recurrence of tracking stream overtake between two gestures engaged in canonic streaming.

motion (erratic though it may be) is still perceived, however, in spite of the timbral variety of the sound events composing the gesture.¹¹⁰

In addition to the closing gesture, other circular gestures in *Persephassa* lack overlapping dynamic envelopes between adjacent points. For example, in mm. 99–103, a "three-against-five" rhythmic motive is passed around the circle at a uniform dynamic level (see Figure 5-15).



Figure 5-15: Xenakis, Persephassa, mm. 99-103, open counterclockwise circular gesture

This particular gesture is no doubt a "primitive" circular gesture that anticipates the more continuous circular activity of mm. 352–410 (described above). Nonetheless, mm. 99–103 produce an *open (imperfect) counterclockwise circular gesture* (SG-|ABCDEF|). The nature and

¹¹⁰ This description of the spatial activity in *Persephassa* is partially informed by Harley's analysis of the work. (See Harley, "Space and Spatialization," 293–296.) The interpretation of gestural relationships is my own.

placement (within the musical work itself) of this and other "discontinuous" (i.e., *not* involving overlapping dynamic envelopes) and/or "partial" circles (i.e., *open* circular gestures) prior to mm. 352 highlights the process of "circular construction" alluded to above. The process of circular construction is bound to a concurrent process of transforming discontinuous motion into continuous motion, via the introduction of overlapping dynamic envelopes.

Through the CYCLE schema, we may structure an interpretation of the circular spatial activity in *Persephassa*. The initial perfect circle (mm. 352–353) is counterclockwise and uninterrupted.¹¹¹ After two complete revolutions, the entrance of the clockwise circle presents a conflict: the opposing directionalities represent "cycle" and "anti-cycle" (or, "cyclic reversal"). As more circles of either directionality gradually appear, the primacy of the original cycle is expunged. The decimation of the perfect circle (or, "uniform cycle") of mm. 352–353 (which had been systematically constructed in an elaborate generative process) is hastened by the accelerated rotation of all seven circles. The procedural ruination of a solitary perfect circle— which fleetingly represents an uninterrupted cycle with total and reiterated closure—is sustained through the end of the piece by the mutable closing gesture.¹¹² This final gesture is the disfigured product and remnant of the compromised circularity/cyclicity. Thus, the initial process of constructing continuous, unidirectional circular motion is ultimately undone by a process of circular deconstruction and a return to discontinuous spatial motion.

Cycles often contain a climactic structure: a sequence of alternating high- and low-points. Johnson states: "Life patterns do not simply repeat; they exhibit a character of build-up and

¹¹¹ The interpretive implications of clockwise and counterclockwise circular directionalities shall be explored in the following section (it does not have a direct bearing on the present analysis).

¹¹² Motion around a circle might imply more than cyclic activity and closure. The recursive loop of the circular gesture might signify a lack of progress. The linguistic metaphor "I'm been running around in circles" means that the speaker—presumably in spite of his/her efforts—has failed to accomplish anything substantial. Since circular motion translates to "continuous motion along a closed pathway" (Brower, 329), any projected goal (as typically exists at the end of a linear path) is never actually attained.

release."¹¹³ Johnson cites two examples of fluctuating life processes: "the build-up of sexual or emotional tension followed by a release, [and] the course of an illness that gradually drains our energy before we recover and return to our former 'healthy' state."¹¹⁴ Johnson then evokes the figure of a sine wave to depict the ebb and flow of many experienced cycles. This version of the CYCLE schema is virtually identical to the TENSION-AND-RELEASE schema: as described above, the T&R schema is often fleshed-out in a cyclic manner during the experience of music through the enduring interchange between tensive and relaxive events.¹¹⁵ Johnson continues: "Whether in the simple circular form, or in the sine-wave representation with its imposed climactic character, the CYCLE schema manifests a definite recurring internal structure. This structure constitutes one of our most basic patterns for experiencing and understanding temporality."¹¹⁶ Since music is a temporal art in the literal sense and spatio-temporal both in the metaphorical sense (events in pitch space) and in the literal sense (events in ensemble space), the CYCLE schema plays a prominent role in our comprehension of manifold musical events and processes-from rhythmic regularity, to the harmonic and formal periodicities of tonal music, and to certain types of spatial gestures and polygestural activity.

The CONTAINER schema

According to Johnson:

Our encounter with containment and boundedness is one of the most pervasive features of our bodily experience. We are intimately aware of our bodies as three-dimensional containers into which we put certain things (food, water, air) and out of which other things emerge (food and water wastes, air, blood, etc.). From the beginning, we experience constant physical containment in our surroundings (those things that envelop

¹¹³ Johnson, 119.

¹¹⁴ Ibid., 120.

¹¹⁵ Cycles may be periodic, as depicted by a sine wave (Figure 5-6), or aperiodic, as portrayed by the elaborated FORCE schema (Figure 5-7) and the T&R schema II (Figure 5-8).

¹¹⁶ Johnson, 121.

us). We move in and out of rooms, clothes, vehicles, and numerous kinds of bounded spaces. If we look for common structure in our many experiences of being *in* something, or for locating something *within* another thing, we find recurring organization of structures: the experiential basis for in-out orientation is that of spatial boundedness.¹¹⁷

We understand our bodies to be both contained and containing. It is through this embodied knowledge that we often conceptualize abstract events as somehow involving containment. Earlier in this chapter I discussed the notion of conceptualizing categories as containers and the members of a category as the contents of a container. If the necessary and sufficient conditions for categorical membership are satisfied, then a thing may belong to a given category (i.e., exist *within* a category-container). Johnson represents the CONTAINER schema as shown in Figure 5-16. In this diagram, some item X is contained within the circular enclosure. There is a distinct inside and outside to the container—these regions are separated by the container boundary (which happens to be circular in the figure).



Figure 5-16: CONTAINER schema¹¹⁸

Brower describes a series of nested containers as forming a schema for hierarchy, as illustrated in Figure 5-17. Brower describes the topography of this figure as comprising "successive layers radiating outward from a core."¹¹⁹ She further states: "We tend to conceive of its structure as originating at the center, with each added layer mirroring in some way the

¹¹⁷ Ibid., 21. Emphasis Original.

¹¹⁸ Taken from Johnson, p. 23, Figure 2.

¹¹⁹ Brower, 328.

structure of the core."¹²⁰ Johnson highlights the *transitivity* of nested containers: "If B is in A, then whatever is in B is also in A. If I am in bed, and my bed is in my room, then I am in my room."¹²¹



Figure 5-17: Nested containers/schema for hierarchy¹²²

It is quite clear how such a schema structures our understanding of musical relationships such as metric/hypermetric relationships and Schenkerian structural hierarchy. In a 4/4 meter, each measure will contain four beats. The measures themselves might group into 4-bar units, yielding a 4-bar hypermeter in which the strong-weak-strong-weak accentuation pattern of each measure is replicated on a higher level of rhythmic structure. With the correlation to Schenkerian analysis, we might envision the Ursatz as the center of the container with its multi-level proliferation radiating outward—through the middleground levels to the foreground (surface) level, the latter of which is represented by the container's boundary.¹²³

We conceptualize space as a container. Lakoff and Johnson state: "Containers can be viewed as defining a limited space (with a bounding surface, a center, and a periphery) and as

¹²⁰ Ibid.

¹²¹ Johnson, 22.

¹²² Taken from Brower, p. 328, Figure 4.

¹²³ It might prove useful to imagine Figure 5-17 as a three-dimensional cone opening toward the viewer, such that the center (musical background) is the cone's point located farthest from the observer and the boundary of the cone (musical surface) is closest to the observer.

holding a substance (which may vary in amount, and which may have a core located at the center)."¹²⁴ In modeling the spatial activity of a musical performance, we may view the ensemble space as a container. The ensemble space contains any number of discrete points and is itself contained by the performance space—an observation that conforms to Brower's fourth entailment of the container schema: "Smaller containers may be nested within larger ones."¹²⁵ The sonic space (if present) is likewise contained within the ensemble space. Brower's fifth entailment, that "the boundaries of a container may be fixed or flexible," is applicable to the "flexible" sonic space within the "fixed" ensemble space.¹²⁶ Brower states in her sixth entailment that "a flexible container may expand or contract."¹²⁷ The sonic space frequently dilates and contracts within the ensemble-space container—when the sonic space is expanded to its full potential, it is equivalent in magnitude to the ensemble space.¹²⁸ Figure 5-18 depicts a five-point ensemble space as a container.



Figure 5-18: Ensemble-space as container (five-point)

In Figure 5-18, ensemble points 0 and 4 function as the spatial extremes of the ensemble space. They therefore serve as the boundaries of the ensemble-container. Points 1, 2, and 3 are the

¹²⁴ George Lakoff and Mark Johnson, *Metaphors We Live By* (Chicago and London: The University of Chicago Press, 1980), 92.

¹²⁵ Brower, 328.

¹²⁶ Ibid.

¹²⁷ Ibid.

¹²⁸ The expansion of the sonic space is an analog to registral expansion in pitch space. Expansion itself is tensive, while contraction is generally relaxive. Taking a deep breath (inhalation) requires tensing of the diaphragm, which results in expansion of the chest cavity as the lungs fill with air. Deflation of the lungs (exhalation) follows the "release" of diaphragmatic flexion.

ensemble-container contents, since they abide within the boundaries of the container. Since this is a five-point ensemble space, point 2 is the content-point center, or "core," of the container.¹²⁹

In viewing the ensemble space as a container, we may refer to its discrete points as either central, inner, outer, or peripheral based upon their relative locations within the container. The identity of an ensemble point in relation to the entire ensemble-container is variable depending upon the spatial context in which it is active. Abstractly, in five-point ensemble space, points 1, 2, and 3 are the *inner* (or, "contained") points, point 2 is the *central* point, and points 0 and 4 are the *outer* points (or, "container boundaries").¹³⁰ In an analytic context, point 1 may be considered an *inner point* if functioning independently *or* if gesturally linked to (or fused with) point 2 (or points 2 and 3). Point 1 may be considered a *peripheral point*, however, if linked to point 0 (a natural outer point of the ensemble space). Likewise, point 3 may be considered peripheral if linked to point 4. Finally, point 1 may be considered an *outer point* (or, "container boundary") if ensemble-space segment 1–4 is inspected independent of point 0.

The importance of distinguishing between central, inner, outer, and peripheral functions pertains to describing the spatial motion within the ensemble-container. Many weak closed gestures (spiral gestures in particular) feature either an In \rightarrow Out motional process (such as SG-|23140| and SG-|12304|) or an Out \rightarrow In process (as with SG-|40312| and SG-|03412|).¹³¹ Additionally, many gestures will display either an Out \rightarrow In \rightarrow Out motional process or an In \rightarrow Out \rightarrow In process, both of which may correspond well with the fleshing-out of other image schemata, such as the PATH schema (open gestures) and the DEPARTURE-AND-RETURN

¹²⁹ In conceptualizing categories as containers, a degree of precedence is attached to the central content-item (the categorical prototype or principal gestural form). However, unless exploited by the composer, the central ensemble point of the ensemble-container does not necessarily assume a status of primacy.

¹³⁰ By "outer" I do not mean "outside of the container."

¹³¹ On one level of interpretation, the In \rightarrow Out process is "open" compared to the Out \rightarrow In process. Although outward spiral gestures are technically classified as weak closed gestures, their expansive pattern of "opening" implies continuation beyond the boundaries of the ensemble-container.

schema (closed gestures). Strong open gesture SG-|43210| and weak open gesture SG-|43120| both feature motion from one spatial extreme (or, ensemble-container boundary) to the other, as structured by the PATH schema. The inner ensemble points are articulated in the middle portion of each gesture, yielding the Out \rightarrow In \rightarrow Out process. Strong closed gesture SG-|012310| (in five-point ensemble space) likewise exhibits the Out \rightarrow In \rightarrow Out dynamic, but it returns to the same spatial extreme from which it began. On the other hand, strong closed gesture SG-|213432| reveals the In \rightarrow Out \rightarrow In process. Weak closed gesture SG-|4123| exhibits a weaker instantiation of Out \rightarrow In \rightarrow Out, as the gesture *almost* works its way back to point 4. These closed gestures all are structured by the D&R schema.¹³² Without the distinction between inner and outer ensemble points, these kinetic processes would be meaningless. Moving toward or away from the center of a contained space can recur throughout a musical work and serve both a motivic and rhetorical function within the larger motional complex of the work.

The CENTER-PERIPHERY schema

This schema was evoked earlier in the chapter when I described the hierarchical organization of a gestural category specific to a given work of music. In such a category-container, the principal gestural form is located at the center of the category-container with related gestural forms extending away from the central gesture toward the boundaries (periphery) of the container. The gestures least similar to the principal gesture (as the localized categorical prototype) are located farthest from the center of the category-container. Johnson states: "The fact of our physical embodiment gives a very definite character to our perceptual experience. Our world radiates out from our bodies as perceptual centers from which we see, hear, touch,

¹³² SG-|4123| involves only a partial fleshing-out of the D&R schema since the motion does not return to point 4. As such, this gesture is less complete (i.e., "weak") and more tensive than the strong closed gestures due to its irresolution.

taste, and smell our world."¹³³ Thus, we understand ourselves as central relative to our surroundings, which are peripheral. However, we also impose the center/periphery distinction on surrounding objects that we observe with our senses. Johnson writes: "From our central vantage point we can focus our attention on one object or perceptual field after another as we scan our world. What is 'figure' or 'foreground' at one moment may become 'background' at another."¹³⁴ As discussed in Chapter 3, the distinction between "figure" and "ground" is a fundamental concept of Gestalt psychology. Multistability, or figure-ground reversal, is exactly what Johnson is referring to in the above quotation, where he indicates that we may selectively focus our attention on any object within a perceptual field. The primary object of perception, at any given moment, is the perceptual figure. All other objects are relegated to the background. Johnson's diagram of the CENTER-PERIPHERY schema is reproduced in Figure 5-19.



Figure 5-19: CENTER-PERIPHERY schema¹³⁵

When we focus our eyes on visual objects, the light reflecting from those objects falls on the macula lutea at the center of the retina. The entire macula accounts for approximately 5 degrees of the central visual field. In the center of the macula is the *fovea*, in which fine detail

¹³³ Johnson, 124. ¹³⁴ Johnson, 124.

¹³⁵ Taken from Johnson, p. 124, Figure 24.

discrimination is most pronounced. Visual acuity is greatest in the foveal region due to the high concentration of cones, which are the receptor cells responsible for capturing color and fine detail. Outside of the center of the visual field lies our peripheral vision, which is largely perceived via the rods (the other type of visual receptor cell). Simply turning our gaze to a distal object, such that the proximal image of the object falls on the macula, casts that object into our perceptual center.¹³⁶

The CENTER-PERIPHERY schema likewise applies to audition. When we listen to a polyphonic fugue or to an orchestral work with a lush texture, we may selectively attune our hearing to certain lines or musical events within the larger musical context. How is this possible? The fluid-filled *cochlea* of the inner ear contains the *basilar membrane*, which is studded with the hair cells responsible for transducing pressure changes in the cochlear fluid into neural impulses that the auditory cortex of the brain interprets as sound. Specific locations along the basilar membrane vibrate in accord with specific frequencies. "The mechanical properties of the membrane allow the auditory system to distinguish one frequency from another by the location on the membrane that is maximally excited [and physically displaced] by a particular frequency."¹³⁷ The hair cells on the basilar membrane are divided into two kinds: inner and outer. The inner hair cells connect with afferent nerves that transmit neural impulses to the brain. The outer hair cells are linked to efferent nerves: they receive signals from the brain that modify the activity of the inner hair cells in a process referred to as somatic electromotility (or, *fast motility*). When we actively focus our attention on the viola part of a Mahler Symphony (which may or may not feature the principal melodic line), we are focusing our hearing such that the viola part becomes the "figure," and all other musical parts are consigned to the background,

¹³⁶ Levine, 37, 52–53. ¹³⁷ Ibid., 355.

or "sonic periphery." Likewise, when we ignore the persistent coughing of a fellow audience member, we are dismissing what we perceive to be "extramusical" sounds and attending to the "musical sounds." The extraneous noises are perceptually peripheral to the more central, musical sounds. If the coughing becomes too disruptive, our attention is drawn away from the musical sounds and becomes focused (or, "centered") on an inherently nonmusical sound.¹³⁸

The integration of visual and auditory processes may bring a percept strongly into the "perceptual center." As a result of intricate crossmodal binding, neuronal networks exist between the various sense modalities. As noted in Chapter 2, there are cortical cells that only respond to dual stimulation from two or more modalities, such as vision and audition. Thus, we often come to know the spatial location of a distal object/event by jointly seeing *and* hearing it. Since light travels much faster than sound, it is not uncommon for us to see a "thing" prior to hearing it. A common example is seeing a bolt of lightning in the distant horizon and hearing the thunder it produces a few seconds later. Due to the different transmission rates of light and sound, we naturally perceive these two events to be unconnected and thereby unrelated. The only clue indicating that they might indeed be coupled is the fact that their spatial locatedness seems identical, even though the two events are perceived as being temporally disjunct. Since vision is the dominant sense modality, the ears often direct the activity of the eyes; audition frequently chooses the objects of visual perception. The ears will pick-up sounds coming from any direction and elevation, but we can only see a limited portion of our surroundings. Edward A. Lippman states: "Our ears largely play the part of sentinels in calling our attention to activity which needs investigation; the more exhaustive determination of the happening or object is

¹³⁸ We categorize sounds as either musical or nonmusical. When confronted with 4'33" by John Cage, we are invited to receive as musical certain sounds that we might ordinarily classify as nonmusical (and potentially disruptive): the rustling of a program, the hum of fluorescent lights, etc. If we accept the invitation and elect to receive such sounds as musical, we have (at least temporarily) reconfigured and expanded our conceptual category of "musical sounds." Such sounds become the perceptual *center* against a *peripheral* silence.

reserved for the keener perceptive power of the eyes.^{"139} Hearing a peripheral sound often prompts the turning of the head such that the sound-producing object is brought into the center of the visual field.¹⁴⁰ However, the ears are also most effective when the head is centered on the object of audition (the sound source). Lippman continues: "Turning towards a source is calculated to bring our ears into maximum use; they are most sensitive in almost every respect, including that of localization, when they inspect a source that is straight ahead."¹⁴¹ This increased localization accuracy occurs because interaural differences (of time and/or intensity) are minimized (as discussed in Chapter 2). Due to multisensory interaction (crossmodal binding), a strong perceptual center is forged when both the eyes and the ears are "focused" on an object.

When experiencing a spatial gesture in live performance, any inactive ensemble point is perceptually peripheral both visually and aurally: visually because we are likely not watching an inactive performer, and aurally because of the lack of sonic activity (and resultant silence) at that point. As a migratory L-R PAN traverses the ensemble space, the perceptual figure (from an audio-visual perspective) shifts from point to point. As the figure (object of perception) moves from point 0 to point 1, point 0 recedes (perceptually) to the background. However, the preceding activity at point 0 is retained in memory, for the ultimate aggregation of successive ensemble-point activity forms the spatial gesture as a whole. If the L-R PAN were accumulative, the process would be the same. Newly entering instruments (ensemble points) capture our attention and blend together into the background as the next instrument in the gestural sequence

¹³⁹ Edward A. Lippman, "Spatial Perception and Physical Location as Factors in Music," *Acta Musicologica* 35.1 (1963), 28.

¹⁴⁰ This reflexive action likewise integrates proprioceptive, kinesthetic, and vestibular neural functions.

¹⁴¹ Lippman, "Spatial Perception and Physical Location as Factors in Music," 29.

enters. Once the sonic space has expanded and the total ensemble space is active (ALL PLAY), other musical factors (salience conditions) may attract our perceptual center of focus.

In applying the CENTER-PERIPHERY schema to the figure-ground experience of music, we are evoking a conceptual (or experiential) space. The perceptual figure is not necessarily physically closer (or louder, or larger) than objects or sounds at the background. Johnson addresses the concept of an experiential space, acknowledging it as an "abstract interpretation of the CENTER-PERIPHERY schema":

In my "world" some things, events, and persons are more important than others— they loom larger in my experience and are more central to my interactions. Others are relatively peripheral at a given point in time. [The CENTER-PERIPHERY schema] shows itself not only in the structure of my perceptual field but equally important as a structure of my social, economic, political, religious, and philosophical world. Those "objects" that stand forth as significant in my experiential field are both concrete and abstract entities toward which my interest is directed.¹⁴²

As mentioned earlier, the CENTER-PERIPHERY schema is often bound to the CONTAINER schema in a dynamic superimposition of schematic structures.

A more literal application of the CENTER-PERIPHERY schema arises in considering the experience of music in a multi-dimensional, surround-sound environment. If the performing ensemble or loudspeaker array (in the case of an electronic composition) surrounds a listener, that listener is contained within the ensemble space, and the ensemble space, in turn, is nested within the performance space. Under such conditions, the perceiver is (relatively) "central," the ensemble points are "peripheral," and the performance space is "outer." The perceiver no longer observes the ensemble as a distal container with ensemble-point contents; instead, the listener becomes an item contained within the ensemble space.

¹⁴² Johnson, 124–125.

The OBJECT MANIPULATION schema

Our reception and conceptualization of motivic alteration and transformation is structured by our embodied understanding of object manipulation. From the time that we are able to grasp objects as small children, we learn to manipulate those objects to suit our needs and wants.¹⁴³ If we imagine holding a lump of clay in our hands, anything that we could do to manipulate the form/shape *or* the spatial orientation of the clay will conform to general knowledge of motivic alteration/variation, whether the transformative procedures are applied to pitch motives, rhythmic motives, spatial gestures, or any other type of motive.¹⁴⁴ We could relocate the lump of clay to new spatial coordinates (sequence, transposition, nudging), "flip" the lump of clay horizontally (retrogression) or vertically (inversion), stretch the clay horizontally (temporal augmentation) or vertically (intervallic expansion), compress the clay horizontally (temporal diminution) or vertically (intervallic compression), break-off small pieces of the clay (melodic fragmentation, gestural abbreviation), add new pieces of clay (elaboration, expansion), break-off one end of the clay and add it to the other end (rotational slot displacement), and/or place the clay into (or remove it from) a container.

Such embodied experience of manipulating objects (malleable objects in particular) results in a family of OBJECT MANIPULATION image schemata that structure and inform our continued bodily (sensorimotor) experience as well as our experience of abstract phenomena, such as the alteration of musical motives (including spatial gestures). This family of schemata contains multiple members, all of which are related by the fact that they alter the physical form, temporal dimensions, and/or spatial orientation of an object. When applied to (or associated with) a specific object (such as some type of musical motive), the family of OBJECT

¹⁴³ Gibson states: "The child learns what things are manipulable and how they can be manipulated" (Gibson, *The Senses Considered as Perceptual Systems*, 285).

¹⁴⁴ Alterations of spatial gestures are described in Chapter 4.

MANIPULATION schemata provides for and portrays the numerous *transformational affordances* of that object. We may speak of the INVERSION schema as a specific member of the OBJECT MANIPULATION schematic family that affects the spatial orientation of a spatial gesture or melodic motive.¹⁴⁵ Likewise, the AUGMENTATION schema increases the temporal duration of a motive, and the COMPRESSION schema reduces the size of a musical entity: it decreases the interval(s) of a pitch motive or the breadth of a spatial gesture.

In a category of related gestural forms (as discussed in section 5.2), the conceptual category-container is organized such that the principal gestural form (as local categorical prototype) is positioned at the center of the category with related gestural forms fanning out from the center and toward the periphery of the category. The network may reveal all possible (i.e., "afforded") gestural relations—even though all forms might not appear in the work itself.¹⁴⁶ These related forms are linked to the principal form (as well as to each other) via an array of various OBJECT MANIPULATION schemata. In essence, the transformative procedure structured by a schematic member of the OBJECT MANIPULATION family *is* the relationship between any two gestures. Approaching this established category-container as a synchronic network (in the "Lewinian" sense of the word), we may trace the diachronic progress of spatial gestures via the PATH schema—various paths are woven through the container-category/network as the musical work unfolds in real time. From one gestural form to another (as actually emerge during the performance of a musical work), the PATH schema (within the conceptual category-container/network) will be superimposed on and integrated with a specific

¹⁴⁵ The inversion of a melodic motive is a transformation in pitch space—a metaphorical construct.

¹⁴⁶ For the related gestural forms that *do* appear in the work itself (and that occur subsequent to the presentation and establishment of the principal form), any related gestural form functions as an "inner icon" to the principal form—these forms are signs that evoke the principal form. The greater the similarity (i.e., the less alteration involved), the more enhanced the iconic function of the related form. See Tarasti, *Signs of Music*, 11–12. Tarasti's notion of "inner icon" is highly similar to Brower's formulation of "intra-opus patterns."

OBJECT MANIPULATION schema. For instance, consider SG-|0132| to be the principal gestural form of a given work of music. After this gesture is exposited in the opening portion of the work, suppose that SG-|2310| appears. The PATH schema linking SG-|0132| to SG-|2310| co-occurs with (is conflated with) the RETROGRADE schema. Figure 5-20 depicts a hypothetical network for principal gestural form SG-|0132|, in part.

In this network, principal gestural form SG-|0132| is located at the center of the weak closed gesture category-container. The network depicts neither abbreviated gestural forms nor gestures with fused superpoints. All twenty possible forms of weak closed gesture in a four-point ensemble space (where each point occurs once and only once) are represented in the network.¹⁴⁷ A single bi-directional arrow between any two gestural forms indicates a single operation that may transform either gesture into the other.¹⁴⁸ For instance, inversion (type 1) will transform SG-|0132| into SG-|3201|, and vice versa. This relation is depicted by the INVERSION schema, which is represented by "I" in the diagram.¹⁴⁹ Should SG-|3201| unfold in the ensemble space subsequent to SG-|0132| during an actual musical performance, a conceptual path would traverse the INVERSION schema-link connecting the two gestures in the network—the path would move from SG-|0132| to SG-|3201| based on the ordered appearance of the two related gestures.

¹⁴⁷ See the appendix for a listing of all the potential four-event weak closed gestures in four-point ensemble space. The network of Figure 5-20 depicts the total *afforded* gestural forms for weak closed gestures (given the specified restrictions).

¹⁴⁸ The bi-directional arrow is perhaps best described as a "link" rather than a two-way "path." (It is optimal to think of the PATH schema as unidirectional.) Johnson claims that the LINK schema "makes possible our perception of *similarity*" and allows us to "experience *causal connection* between temporally linked events." Johnson further states: "A *network* (causal or otherwise) is nothing more than a set of links" (Johnson, 117–119). Although a gestural network may contain links between its component gestures, a path weaving through the network may depict the actual sequential presentation of spatial gestures. As such, many links in the synchronic network are navigated as paths during the musical work's unfolding (as well as in the analytic portrayal of diachronic events).

¹⁴⁹ R = RETROGRADE schema; I = INVERSION (type 1) schema; RI = RETROGRADE-INVERSION schema; R/I = RETROGRADE or INVERSION schema; RSD = ROTATIONAL-SLOT-DISPLACEMENT schema; PE = POINT-EXCHANGE schema.



Figure 5-20: Partial network for principal gesture SG-|0132|. Arrows represent specific OBJECT MANIPULATION schemata.

A unidirectional arrow in Figure 5-20 indicates that the procedure of alteration only works in one direction between network items; however, there will be a complementary (inverse) procedure in the opposite direction. For instance, SG-[0132] is related to SG-[1320] via the RSD(3) schema, but SG-[1320] relates to SG-[0132] via the RSD(1) schema.¹⁵⁰ Equivalent procedures likewise emerge from the network: SG-[0132] relates to SG-[3201] via both inversion and RSD(2), as indicated by the dashed line connecting the two gestures in the figure. Metaphorically, this is related to the common understanding that many different roads (or, "paths") may lead to the same place. Intercategorical alterations that change a weak closed gesture into an open (strong or weak) gesture are indicated with a dashed arrow. In an effort not to make the network more cluttered than it already is, all gestural relations are not indicated; many of the gestures are related to one another by point exchange or rotational slot displacement. Furthermore, the relative lengths of different arrows do not necessarily correspond to relative degrees of similarity.¹⁵¹

This category-container/network reveals the specific relatedness of different weak closed gestures to the hypothetical principal gestural form, SG-|0132|. Those located farthest from the prototype-center of the category are (generally) least similar to the principal gesture, but all the forms are related by virtue of their shared membership in the weak-closed category. Within an actual work of music, the preponderance of a certain transformational operation or relational procedure may itself function as a unifying motive. Imagine that the following sequence of gestures occurs in the hypothetical work of music for which we have constructed the above

¹⁵⁰ Although abbreviated forms are not shown in Figure 5-20, SG-|0132| may be altered to produce SG-|032| via the ABBREVIATION (DELETION) schema, but SG-|032| may be modified to yield SG-|0132| via the ELABORATION (ADDITION) schema.

¹⁵¹ Due to the high number of network items and even higher number of arrows interconnecting those items, some arrows have been forced to navigate the network in a circuitous fashion to reduce clutter. Generally, however, a greater distance from the central principal form indicates a more remotely related (i.e., less similar) gestural form.
network: SG- $|0132| \rightarrow$ SG- $|2013| \rightarrow$ SG- $|3201| \rightarrow$ SG- $|1320| \rightarrow$ SG-|0132|. The fact that all of the gestures in this "gestural progression" are of the "weak closed" variety (the first and last gestures in the sequence are identical) provides the music with a degree of spatio-gestural coherence. The first and third gestures within the sequence (SG-|0132| and SG-|3201|) are related by inversion, as are the second and fourth gestures (SG-|2013| and SG-|1320|). However, the gestures are sequentially linked by the procedure of rotational slot displacement.¹⁵² Besides indicating the temporal succession of the spatial gestures, each arrow (\rightarrow) in the notation for the gestural progression represents RSD(1). Therefore, RSD(1), as an abstract transformative procedure structured by the OBJECT MANIPULATION schema, functions as a unifying device at the local level of the section in which this gestural progression occurs.¹⁵³

In the sense that—within our embodied understanding—some index of effort or degree of force is required to physically manipulate an object, any alteration of a spatial gesture might be received as tensive. This is, of course, assuming that some abstract agent (or force) is perceived to be the cause of gestural alteration, which may or may not be the case for the listener.¹⁵⁴ If such is indeed the case, some alterations will be more tensive than others. *Nudging* entails relocating an entire gesture from one segment of ensemble space to another. Gestural *expansion* and *extension* increase the spatial range of a gesture. In spite of the fact that they are "force-requiring" procedures, alterations such as *compression, truncation,* and *deflation* might be deemed relaxive since they decrease the spatial size of a gesture. The seeming interpretive conflict apparent in this paragraph (for instance, that a compressed gestural form may be said to

¹⁵² Refer to either Chapter 4 or the glossary for an explanation of rotational slot displacement.

¹⁵³ Note also the "departure-and-return" quality of this gestural progression: it departs from, and eventually returns to, SG-|0132|. The departure would be more pronounced, and the return arguably more emphatic, if a gestural form other than a weak closed gesture appeared in the interior portion of the gestural progression.

¹⁵⁴ Lewin claims that attending to the actual transformations themselves, rather than solely to the objects being transformed, exemplifies the "attitude of someone inside the music" (Lewin, *Generalized Musical Intervals and Transformations*, 159).

evoke either tension or relaxation) highlights the subjective variability of musical reception and interpretation. The listening subject approaches the musical object with his/her own unique knowledge, experience, ability, and expectations. However, as noted above, the subjective response to musical events is constrained by a set of common bodily, environmental, and cultural experiences shared by a group of listeners. These orthodox experiences exist as a compendium of communal image schemata, which are metaphorically projected (often in a conventionally prescribed fashion) to structure the reception and understanding of musical activity. By identifying the schemata that are potentially active during the experience of music, a band of interpretive probabilities may be narrowed from the full range of interpretive possibilities. In short, how the listener is listening will dictate whether a compressed spatial gesture is received or interpreted as either tensive or relaxive. In either case, a different image schema (or portion of the T&R schema) is metaphorically projected—which schematic portion takes precedence in experience is largely indeterminable from an analytic standpoint. However, context might provide some important hermeneutic cues. For example, if a compressed spatial gesture follows the inarguably tensing process of spatial expansion, the compressed gesture might be perceived (and best analyzed) as relaxive.

The BALANCE schema

Balance is a multifaceted concept. It applies to physical objects such as our bodies, to the judiciary system (the balance of justice), to the concept of equality, to psychological states, to various artistic mediums (including the visual arts, music, dance, and theatre), and to many other social constructs. However, all metaphorical extensions of the basic concept of balance stem from the basic experience of balancing our bodies. Johnson states: "It is crucially important to

308

see that balancing is an activity we learn with our bodies and not by grasping a set of rules or concepts. First and foremost, balancing is something we do."¹⁵⁵ A baby is not taught how to balance its body and stand upright but learns to do so out of necessity. There is no prescribed set of propositional rules that enables a baby to balance itself; the baby simply learns to distribute its weight evenly while negotiating (and countering) external forces such as gravity and wind.

In addition to the continuous (and typically unconscious) process of physically balancing our bodies, we understand balance through the experience of bodily equilibrium: "We understand the notion of systemic balance in the most immediate, preconceptual fashion through our bodily experience."¹⁵⁶ If we are too hot or too cold, we seek to "balance" our internal body temperature knowing that a high fever or hypothermia could be fatal. When we are hungry we eat food. When we are dehydrated we drink water. Much systemic balance occurs without any effort: in the autonomic nervous system (the part of the nervous system that controls involuntary bodily functions), the sympathetic nervous system (which triggers the "fight or flight" reflex in times of stress) is balanced by the parasympathetic nervous system, which reduces the elevated heart rate and blood pressure. Bodily homeostasis is the default (and preferred) state for optimal health and comfort.

Johnson claims that we really only notice balance when it is absent, as when we lose our balance and perhaps fall to the ground.¹⁵⁷ The process of having to reestablish balance makes us intimately aware of its exigency. I posit that the loss of balance is more than a mere lack of balance (as structured by the BALANCE schema): the bodily experience of the loss of balance gives rise to an IMBALANCE schema that is the functional and antithetical complement of the BALANCE schema. There are many ways to depict the concepts of balance and imbalance

¹⁵⁵ Johnson, 74. Emphasis Original.¹⁵⁶ Ibid., 75.

¹⁵⁷ Ibid.

graphically. Figure 5-21 presents "twin-pan" balance and imbalance. The downward arrows represent gravitational force, which is balanced in Figure 5-21(a) and imbalanced in Figure 5-21(b). The different lengths of the arrows in 5-21(b) represent the different force intensities entwined with the imbalance: the longer arrow represents a greater degree of force.



Figure 5-21: a) BALANCE schema and b) IMBALANCE schema¹⁵⁸

Because the experience of imbalance often yields adverse or disadvantageous consequences (such as broken limbs, bruises, embarrassment, physical or mental illness, etc.), the experiential fleshing-out or metaphorical projection of the IMBALANCE schema is often attended by negative connotations. In short, balance is considered "good" while imbalance is deemed "bad." Furthermore, imbalance engenders *tension* while balance produces *relaxation*: sudden physical imbalance necessitates the application of tension (or force) to assist in the

¹⁵⁸ Figure 5-21(a) is taken from Johnson, p. 86, Figure 18. Figure 5-21(b) is my own modification of Johnson's Figure 18.

reestablishment of (i.e., "return to") a balanced bodily state. Once balance is reacquired, the tension subsides.

Johnson explores the metaphorical projection of the BALANCE schema as it applies to visual perception, particularly to the aesthetic affect provided by balance (or the lack thereof) while viewing a work of art (such as a painting or sculpture): "In cases of [both] bodily and visual balance...there seems to be one basic schema consisting of a point or axis around which forces and weights must be distributed, so that they 'counteract' or 'balance off' one another."¹⁵⁹ In viewing a painting, we perceive tension in imbalance as metaphorical forces act on one another. These "forces" exist solely in perception and "give rise to [the experience of] tensions, movement, rest, and balance...there are not physical weights or forces in balance; rather, the shapes, lines, forms, and spatial relationships involve visual weights and forces."¹⁶⁰

Symmetry is a special kind of balance but is not the only kind of balance. In a painting, elements of different sizes, shapes, and colors may balance one another on the canvas, depending upon their dispensation. Force and weight relations (which determine balance) may be influenced by factors such as location in the plane, spatial depth, size, isolation, shape, and color. The aggregate of these factors can produce a sense of balance in the absence of symmetry. However, symmetry is perhaps the most perfect form of balance, as one half of the object will have the exact same amalgam of the above factors as the other half. We experience our bodies (the outwardly visible portion) as symmetrical across an imaginary vertical axis. This symmetry surely facilitates the act of physically balancing our bodies. Symmetry has long held value and sway among philosophers and aestheticians. Plotinus states: "Universally, the beautiful thing is

¹⁵⁹ Johnson, 80.

¹⁶⁰ Ibid., 81.

essentially symmetrical, patterned.¹¹⁶¹ The proposition that geometric proportions lend beauty to physical works of art is certainly not a new concept to Greek thought and can be viewed as a reflection of the various symmetries that occur in nature. Kant states: "Critics of taste adduce geometrically regular figures as the simplest and most indubitable examples of beauty."¹⁶² A geometrically regular figure might be a circle, square, rectangle, or triangle. Kant states further that any "violation of symmetry [is disliked] because [it is] contrapurposive."¹⁶³ This statement highlights the negative connotations commonly associated with imbalance.

The concepts of balance and symmetry pervade non-Western thought as well, as evinced by the Asian concept of *yin* and *yang*, which is prevalent in many Asian philosophies and works of art. The concept of yin and yang originates in ancient Chinese cosmology and philosophy. The Chinese conception of *qi*, often translated as "vital energy," is divided into two strands: *yin* and *yang*. Yin and yang are described as two primal, oppositional, yet complementary forces: they are interdependent and effectively "balance" one another. Symbolic traits of yin include: femininity, the moon, the lake, water, cool temperatures, passivity, docility, and gentleness. Symbolic characteristics of yang include: masculinity, the sun, the mountain, fire, warm temperatures, aggression, vigor, and strength.¹⁶⁴ The sign for yin and yang, the *taijitu* (pictured in Figure 5-22), is not perfectly symmetrical but consists of two identically shaped figures that together create the circularity of the symbol. Everything in the world is believed to contain some amount of yin and yang, which together form the totality of *qi*. These two competing forces are

¹⁶¹ Plotinus, *The Enneads:* Sixth Tracate "On Beauty," translated by Stephen Mackenna (Faber & Faber, 1956), 56. ¹⁶² Kant, 92.

¹⁶³ Ibid.

¹⁶⁴ JeeLoo Liu, An Introduction to Chinese Philosophy: From Ancient Philosophy to Chinese Buddhism (Malden and Oxford: Blackwell Publishing, 2006), 7.

constantly in motion, such that when yin expands, yang decreases, and vice versa. Therefore, the amounts of yin and yang are never perfectly balanced, but the totality of qi is constant.¹⁶⁵



Figure 5-22: Taijitu, the traditional symbol of yin and yang

The incessant competition between yin and yang is governed by the principle of harmony, and

the dynamic flux of yin and yang proportionality is balanced such that their mean proportions are

50/50. According to JeeLoo Liu:

Yijing ["I Ching"] describes a natural harmony in the world of Nature, which balances *yin* and *yang* in the state of *equilibrium*—the balance between two extremes. The Chinese word *zhong* is commonly translated either as "equilibrium" or as "the mean," since it connotes both. The mean is not just the strict middle point between two extremes; rather, it is a state of intrinsic harmony. Hence, some elements could be more than others, as long as the inequality does not disturb the harmonious balance.¹⁶⁶

In addition to this sophisticated view of balance, the notions of cyclic repetition and departure-

and-return are wrapped-up in the concept of yin and yang: when one of the two forces reaches its

maximal state, it will then decline (i.e., "return" toward a state of balance) as its antagonistic

partner increases. Liu continues:

We can envisage a cosmic pendulum in the constant movement of *yin* and *yang*. The goal of the movement of qi is not to keep the pendulum resting still at the middle point, since this eliminates movement. With the constant swing between two ends, the pendulum maintains a consistent flow. This consistent movement is harmony itself. When the mean is understood as equilibrium, the allowance for variance is much greater.

¹⁶⁵ Any yin or yang can be further subdivided into yin and yang. This quality reveals a structural concordance between parts and whole. Furthermore, yin and yang can transform into one another: warm into cold; day into night; life into death (these are all examples of yang transforming into yin).

¹⁶⁶ Liu, 37. Emphasis Original.

A development toward one end or the other is not necessarily violating the principle of the mean, as long as any deviation eventually returns to the mean.¹⁶⁷

The notion of a swinging pendulum represents the idea of cyclic repetition—the pendulum is a metaphor for the fluctuation of yin and yang within the totality of qi. Thus, departure from a state of equilibrium (as defined by Liu) necessitates the return to the state of equilibrium, which might then become imbalanced in the opposite direction. Thus, we see the universally embodied schemata of BALANCE, DEPARTURE-AND-RETURN, and CYCLE forming a partial basis of ancient Chinese philosophy.

This non-Western example of balance supports the notion that universal embodied knowledge gives form to schematic structures that shape our conceptualization of abstract phenomena. Many image schemata are indeed universal owing to the fact that human beings from all walks of life share many bodily experiences. However, recall from section 5.1 that the signs for such concepts (including written or spoken words and other types of symbols—such as the *taijitu*, above) are typically specific to a given culture (although their widespread proliferation may generate quasi-universal, partial understanding). Furthermore, as we shall discover in the following section, the metaphorical projection(s) of a given schema is (are) often circumscribed and restrained by an individual's cultural embeddedness.

Let us explore some of the roles that the BALANCE and IMBALANCE schemata play in structuring our reception of music (via the metaphorical projection of the BALANCE schema onto musical events). Applied to events in most musical parameters, the concept of balance assumes a temporal character and is therefore utterly dependent upon the listener's retention of temporal proportions.¹⁶⁸ Balance was highly revered in music of the Classical period (1750–

¹⁶⁷ Ibid., 37–38.

¹⁶⁸ On the other hand, the analysis of a score (in which the musical events are symbolically represented in a visually synchronic state) may also reveal balance and is less time-dependent.

1820). It is extremely common for the two phrases composing a period to be perfectly equal in length: a four-bar consequent phrase "balances" a four-bar antecedent phrase. In the case where one of the two phrases might be extended by one or more measures, the resultant discordance of phrase-length proportion imparts a sense of imbalance to the period as a whole, which might induce tension as the expected arrival at a "goal" (cadence) is delayed. Many composers likewise construct symmetrical (or, palindromic) musical forms. Such "arc" or "mirror" forms appear quite frequently in the music of Béla Bartók. The third movement of *Music for Strings, Percussion, and Celesta* (1936) consists of the following formal structure: Prologue-ABCDCBA-Epilogue, which is symmetrical about section D. The second movement of Anton Webern's *Symphonie, Op. 21* presents a different type of mirrored form: the second half of the movement is the retrograde of the first half.

In addition to the balance of formal proportions, balance and symmetry exist in parameters of pitch and rhythm as well. Melodic motion in one direction is often said to be "balanced" by subsequent motion in the opposite direction: An ascending melodic line followed by a descending melodic line produces balance in the realm of pitch space. It is likewise quite typical for a melodic leap in one direction to be "filled-in" by stepwise motion in the opposite direction. In this case, balanced motion in pitch space complements a balance between disjunct and conjunct motion. Many rhythmic configurations are symmetrical and therefore "nonretrogradable." Many pitch collections are likewise symmetrical: the diatonic and pentatonic collections are symmetrical about scale-degree 2 (which is seldom exploited in tonal works due to the predominance of scale-degree 1), the octatonic collection has four axes of symmetry, and

315

the whole-tone collection has six axes of symmetry.¹⁶⁹ The symmetry of these collections limits their number of distinct forms due to transpositional and inversional invariance: the octatonic collection has three distinct forms and the whole-tone collection has only two.

Dynamic "balance" is an important consideration for performing ensembles. In this sense of the term "balance," I am referring to the parity of intensity (uniformity of loudness) between multiple instrumentalists. A roughly equal number of singers for each part of a choral work, for example, helps to ensure some degree of balance between the parts. The discrete members performing each part, in turn, generally strive to produce the appropriate blend, such that no single voice stands-out from the rest. In this sense, "blend" is synonymous with "balance," and the proper blend results from relative timbral and dynamic homogeny among the different performers. In a symphony orchestra, the different instrumental families (which constitute the individual sections of the orchestra) strive for dynamic balance: the winds must balance the strings, and neither group must be overcome by the brass section. The seating arrangement of many ensembles (chamber as well as orchestral and choral) is often chosen based on the desire to "balance" the musical sound. For example, in the symphony orchestra, the brass and percussion families—which are typically the loudest instruments in the orchestra—are most frequently positioned at the back of the orchestra so as not to overpower the strings and woodwinds, which are generally oriented toward the front of the ensemble space. In an ensemble such as the brass quintet, the seating plan selected by the ensemble is designed to optimize the balance between the "directional" instruments (trumpets and trombone) and the

¹⁶⁹ Less common collections, such as the hexatonic collection (a series of alternating minor seconds and minor thirds) and the enneadic collection (the chromatic collection minus an augmented triad), are likewise highly symmetrical.

"non-directional" instruments (horn and tuba). (A directional instrument is one that primarily projects the sound in the direction that the performer is facing).¹⁷⁰

A more literal application of balance is to be found in the physical arrangement of many ensembles. Even when not producing any sound, a spatially balanced ensemble is visibly noticeable and perhaps aesthetically pleasing. Stockhausen's Carré is composed for four orchestras and four choruses. The organization of the component forces creates a square surrounding the audience. The square is a symmetrical geometrical figure.¹⁷¹ Composers often indicate the requisite layout of an ensemble graphically in the score, and such diagrams frequently demonstrate sensitivity to the sheer spatial balance of the ensemble. In the absence of such diagrams, directors often organize the ensemble based on principles of spatial as well as dynamic balance. Boulez speaks to the issue of spatially balanced ensembles:

Two groups will be symmetrical if they are situated at an equal distance from an axis of some kind; if they possess homogenous or non-homogenous timbres, identical in quality and density, they can be considered as *regularly symmetrical*; they are *irregularly* symmetrical if their homogeneity is not of the same nature (a group of brass against a group of strings, for example) or if their non-homogeneity differs in quality and in density; they will otherwise be asymmetrical.¹⁷²

In Boulez's estimation, the mutual presence of both spatial and timbral balance within the distribution of an ensemble yields "regular symmetry," while spatial balance conjoined with timbral imbalance produces "irregular symmetry." How do concepts of balance and symmetry apply to spatial gestures? As Boulez observes, many ensembles are arranged symmetrically (regularly or irregularly) within the performance space. Centered upon the stage, and with roughly equal distances between adjacent performers, an ensemble displays spatial balance. In odd-cardinality ensembles (such as trios and quintets), the axis of symmetry stretches *through*

¹⁷⁰ Thanks to Heidi Lucas for explaining this convention to me.
¹⁷¹ The English translation of "carré" is "square."
¹⁷² Boulez, *Boulez on Music Today*, 70. Emphasis mine.

the central performer (ensemble point 1 in a trio; ensemble point 2 in a quintet). The central ensemble point is therefore a *real* center of the ensemble space. In even-cardinality ensembles (such as quartets and sextets), the axis of symmetry falls *between* the two centralized performers.¹⁷³ This distinction is quite important when we consider motion from the periphery to the center of the ensemble-container (or vice versa). An even-cardinality ensemble lacks an actual central point, so the periphery-to-center (Out \rightarrow In) motional process will be imperfect in such an ensemble (unless the point fusion of the two centralized performers, if present, should count as a center).

Boulez states: "Whatever the nature of the ensembles [in which they occur], the transitions or leaps taking place between...sound sources...will be symmetrical or asymmetrical."¹⁷⁴ In a five-point ensemble space, an oscillation gesture unfolding between points 0 and 4 as well as between points 1 and 3 will be spatially balanced. Gestures such as SG-[010...] and SG-[242...] do not exhibit balance within the ensemble space. In fact, within an odd-cardinality ensemble space, all adjacent-point oscillation gestures will be spatially imbalanced, while in an even-cardinality ensemble space, only one adjacent-point oscillation gesture (that arising between the two centralized ensemble points) will be spatially balanced. In a five-point ensemble space, however, SG-[010...] followed by SG-[434...] would exhibit balance within the ensemble space at a higher level of spatial activity.¹⁷⁵

Spatial imbalance may likewise arise from the presence of superpoints (fused ensemble points). Imagine SG-|0(1+2+3)0(1+2+3)0| unfolding in four-point ensemble space. This gesture amounts to a "one-against-three" spatial imbalance, in which superpoint 1+2+3 spatially eclipses

¹⁷³ However, point fusion could yield a pseudo-center in an even-cardinality ensemble space. For example, in a sextet, superpoint 2+3, resulting from the fusion of points 2 and 3, would yield a composite, virtual center. ¹⁷⁴ Boulez, *Boulez on Music Today*, 70.

¹⁷⁵ SG-|434...| is the type-1 inversion (across point 2 as the axis of inversion) of SG-|010...|.

point 0, with which it alternates. Such gestural imbalance, resulting from the unequal distribution of ensemble "weight," imparts a sense of tension. Of course, the subsequent presentation of this gesture's inversion, SG-|3(0+1+2)3(0+1+2)3|, would provide a sense of balance on a deeper level of spatial structure (and over a longer time-span)—although locally tensive, it balances the previous gesture on a more global level of spatial activity.

If, within a hypothetical work of music, the occurrence of open gestures was to grossly outnumber that of closed gestures, we might speak to the imbalance of gestural types. Furthermore, we can identify a balance of gestural directionality when a R-L PAN eventually succeeds a L-R PAN: the right-to-left motion of SG-|3210| perfectly balances the left-to-right motion of SG-|0123|.¹⁷⁶ This type of balance is dependent both upon the respective spatial activities (i.e., the specific directionalities) of more than one gesture as well as the temporal relationship between those gestures. Additionally, specific instances of polystreaming may be more or less balanced. For example, parallel streaming and mirrored streaming are generally balanced within the total ensemble space. Uneven consonant streaming, on the other hand, may indeed be spatially balanced (as is the case with the polystreaming activity presented in Figure 3-12), but the disparity between the component gestures' rates of unfolding (within the polygestural complex) results in a dynamic form of imbalance: one portion of ensemble space (that containing the more rapidly occurring gesture) is more tensive than the other.

¹⁷⁶ In addition to instantiating the concept of departure-and-return, sweeping gestures also display an immediate balance of directionality. On a higher level, a RIGHT SWEEP balances a LEFT SWEEP: the two gestures are inversionally related, and the rightward source/goal of SG-|432101234| balances the leftward source/goal of SG-|012343210|. Furthermore, perfect sweeping gestures are palindromic and thereby directionally symmetrical: the "hinge" of a sweeping gesture (point 4 in SG-|012343210|) is likewise the inversional axis of the gesture (but *not* of the ensemble space in which the gesture unfolds). As a palindrome, a perfect sweeping gesture is non-retrogradable.

Other Schemata

I have explored in detail several pervasive image schemata that seem most pertinent to the reception, conceptualization, and interpretation of spatial gestures. Several other schemata are directly relevant to the comprehension of spatial activity. (Johnson identifies nearly all of the following schemata in his "partial list of schemata."¹⁷⁷) The FULL-EMPTY schema may apply to our experience of sonic space expansion and contraction: when the sonic space is totally expanded, the ensemble space is "full." Conversely, when the sonic space is reduced to total silence (through inactivity at all ensemble points), the ensemble space is "empty" of sonic activity. The NEAR-FAR schema structures our experience of spatial gestures unfolding in an ensemble space with multiple depth planes, such as Xenakis's *Empreintes* (discussed in Chapter 4), or in a surround-sound environment.¹⁷⁸ Likewise, the VERTICALITY schema (which grounds our understanding of pitch-height relationships) structures our experience of gestures occurring in an ensemble space comprising multiple vertical levels. The MERGING schema may structure our reception of point fusion, points-to-area assemblage (unison), spatial reconciliation, and convergent polystreaming. Spatial reconciliation is a type of point fusion that occurs when the points contributing to an oscillation gesture or to a migratory open or closed gesture subsequently become fused. It is a reconciliatory process by which spatio-temporally distinct event-point activity ultimately merges into temporal simultaneity. The MERGING schema's binary opposite, the SPLITTING schema, may inform our experience of point-split, area-to-points fragmentation (divisi), and divergent polystreaming. This discussion of image schemata and their functional roles in structuring our experience and conceptualization of musical events-spatial gestures in particular-is far from complete. We have preconceptual

¹⁷⁷ Johnson, 126.

¹⁷⁸ The illusion of depth may be created in electronic compositions (either stereophonic or quadraphonic), as described in Chapter 4.

schemata formulated for any recurrent bodily and environmental experience—many of which contribute to our understanding of abstract musical phenomena via the metaphorical projection of embodied, image-schematic structures.

Transformations of Image Schemata

George Lakoff identifies several important cases in which we perceptually transform one image schema into another. The most common example is the multiplicity-to-mass imageschema transformation.¹⁷⁹ As the perspectival aspect of our perception changes, a multiplicity may become a mass, and vice versa. Lakoff states: "As one moves further away, a group of individuals at a certain point begins to be seen as a mass. Similarly, a sequence of points is seen as a continuous line when viewed from a distance."¹⁸⁰ Lakoff claims further that: "Such relationships are very common in language. Compare cows (multiplex) and cattle (mass)."¹⁸¹ This schematic transformation is applicable to gauging the listener's location relative to the ensemble space (within the performance space). The closer the listener is seated to the performing ensemble, the better the listener may differentiate (aurally and visually) the individual members (points) of the ensemble. As the listener moves farther and farther away from the ensemble space, the spatial discrimination of discrete ensemble points diminishes, and the musical effort of the ensemble increasingly seems to ensue from one "place" within the performance space. As the MULTIPLICITY schema transforms into the MASS schema (via a shifting perspective), the perceptual salience of spatial gestures declines.

¹⁷⁹ Lakoff and Johnson, *Philosophy in the Flesh*, 145. In this most recent source, Lakoff and Johnson use the term "multiplicity." However, in earlier texts this schematic transformation is referred to as "multiplex $\leftarrow \rightarrow$ mass" (Lakoff, 442) and "multiplex to mass" (Johnson, 26). The conception of the transformation does not change from source to source, but the term "multiplicity" has replaced "multiplex" in the later work. Lakoff's 1987 designation of the transformation (multiplex $\leftarrow \rightarrow$ mass) highlights the fact that it can work in either direction (either multiplex to mass *or* mass to multiplex) and is therefore truly an isomorphic transformational operation.

¹⁸⁰ Lakoff, 442.

¹⁸¹ Ibid., 428.

Another important schematic transformation is what Johnson refers to as "Following a Trajectory."¹⁸² Johnson states: "As we perceive a continuously moving object, we can mentally trace the path it has traversed or the trajectory it is about to traverse."¹⁸³ This notion relates to the earlier discussion of expectation and tension as a strong open gesture traces a unidirectional path through an ensemble. It also syncs with David Lewin's discussion of the perceptual frame established by "primal impressions," "retentions," and "protensions." Lewin's concept derives from Husserl's two-dimensional model of perceptual time. Lewin defines primal impressions as "impressions that follow the now-cursor t," retentions as "projections of remembered times past (and past durations) into my present consciousness," and protensions as "projections of future expectations into present consciousness."¹⁸⁴ The concept of a perceptual frame with the above component parts surely applies to the perception of anything that is of a temporal nature and received largely through audition, including language, melody, and spatial gestures. If we experience a perfect migratory R-L PAN in nine-point ensemble space (SG-[876543210]), the "primal impression" will be whatever ensemble point in the gestural sequence is presently active. By the time the gesture has approached the midpoint of the ensemble space (point 4), subgesture SG-|8765| becomes a "retention" and subgesture SG-|3210| a "protension." The unidirectional fleshing-out of the PATH schema and the Prägnanz (good form) of the strong open gesture strengthen the projection ("into present consciousness") of both retentions and protensions relative to the primal impression. As we follow the gesture's trajectory, it appears as a "continuously moving object" and we easily remember the path it has traveled (from sourcepoint 8) and anticipate (or, "expect") its unimpeded continuance to goal-point 0.

¹⁸² Johnson, 26.

¹⁸³ Ibid.

¹⁸⁴ Lewin, "Music Theory, Phenomenology, and Modes of Perception," 329.

Image Schemata and Gestalt

Prior to concluding this section, I would like to highlight the many correspondences and parallels between Gestalt psychology (as detailed in Chapter 3) and image-schema theory that have emerged during this discussion.¹⁸⁵ The first (and most obvious) correlation between the two theories is that an image schema itself is a gestalt structure comprising multiple parts in specified relations to one another. The entailments that fall out of the internal structure (the interrelationships between parts and between parts and the whole) of an image schema constrain our experiences and inferences. Johnson acknowledges a PART-WHOLE schema that structures our understanding of multifaceted objects and events.¹⁸⁶ But the links between the two theories go even deeper: one of the Gestalt principles of good form (Prägnanz) is the Law of Symmetry. Our experience of a symmetrical object's form as "good" is structured by the metaphorical elaboration of the BALANCE schema. Another principle of good form is the Law of Continuity, which states that unidirectional figures are strong. The PATH schema, which structures innumerable experiences, is abstractly understood as the direct and continuous spatio-temporal motion from one point to another (in a single direction). The Gestalt Law of Closure, yet another principle of good form, corresponds to the closure provided by the DEPARTURE-AND-RETURN schema (and potentially by the CYCLE schema). The fundamental Gestalt principle of *invariance* conforms to our embodied experience of object manipulation, as structured by the OBJECT MANIPULATION schema. We recognize a manipulated object due to the invariant (unchanging) qualities of that object that survive any spatio-structural alteration. The Gestalt notion of *multistability* (or, figure-ground reversal) meshes well with our experience (visual, auditory, conceptual) of center and periphery, as organized by the CENTER-PERIPHERY

¹⁸⁵ It is interesting to note that the philosophy of Immanuel Kant contributed significantly to both Gestalt and imageschema theory. ¹⁸⁶ Johnson, 126.

schema. Finally, the basic Gestalt grouping principles (the Law of Similarity and the Law of Proximity) are means of categorization, which dovetails nicely with the metaphorical projection of the CONTAINER schema as a conceptual category (potentially with a hierarchical organization). In many respects, then, image-schema theory is a repackaging of Gestalt theory. The analogs between them serve to strengthen each of the two theories.

Conclusions

By virtue of their experiential origins, image schemata are directly meaningful in and of themselves¹⁸⁷; therefore, their presence in structuring the reception of musical events renders the musical experience *directly* meaningful, prior to conscious reflection, cultural/historical analysis, and semiotic interpretation. Music need not be designative, associative, or explicitly referential to be meaningful. Nor does musical meaning lie chiefly in abstract (or, "absolute") musical processes. Furthermore, culturally ingrained conventions of musical understanding (such as the minor mode denoting sadness) are not arbitrary: they are rooted in embodied experience. Musical meaning lies principally in our response to a musical work; it is neither extramusical nor *intra*musical. It is *intra*personal—based on the listener's own embodied knowledge. However, many of the schemata realized during the musical experience are universal, allowing for degrees of *inter*personal meaning. We can attribute this universality to the wealth of experience that human beings-the makers and receivers of music-share on account of having bodies of like kind and function. Assuming that the same schemata are fleshed-out by multiple listeners during a shared musical experience, there is reason to believe that they will be fleshed-out in a similar manner. Therefore, the subjective response to music is somewhat constrained by the nature and structure of these universal schemata. Meaning is not universal and objective but is public and

¹⁸⁷ Lakoff, 268.

shared owing to the embodiment of the mind *as well as* the environmental and cultural embeddedness of the mind-containing body. The preconceptual fleshing-out of image schemata gives structure, coherence, and meaning to our ongoing bodily, environmental, and cultural experience. The metaphorical projection of image schemata, the means through which we structure an understanding of abstract phenomena (including musical events of all kinds), shall be outlined in greater detail in the following section in an effort to further elucidate the reception and interpretation of spatial gestures.

5.4 Metaphor

The remainder of my hermeneutic methodology rests on the theory of **conceptual metaphor**, which has become a prominent fixture in the fields of philosophy and cognitive linguistics over the past quarter-century. In the early 1980s, George Lakoff and Mark Johnson presented their pioneering and enduring work on metaphor.¹⁸⁸ They argue that metaphor is not merely a rhetorical, poetic device, but is actually central to human understanding and meaning. Metaphor is more than florid, figurative language; it is an indispensable means through which we conceptualize our world. Lakoff and Johnson claim that we structure our understanding of abstract, unfamiliar concepts *in terms of* our knowledge of concrete, familiar concepts. In the preceding section, I showed the "metaphorical projection of image-schematic structures" to be integral to the perception and understanding of music. Whenever musical events are received and organized in terms of bodily experience, such as when we conceive of a melodic line or spatial gesture as being "balanced" in pitch space or ensemble space, one or more schemata have been metaphorically projected onto the domain of abstract experience. The result is a metaphorical *yet very real* experience of music. The experience is "real" because the metaphor

¹⁸⁸ Lakoff and Johnson, *Metaphors We Live By*.

is grounded in embodied knowledge. I shall now expand on this process in greater detail in order to construct a model for interpreting the specific directionality of a spatial gesture.

A conceptual metaphor consists of a source domain and a target domain. The **source domain** is some form of concrete knowledge. This domain is typically grounded by one or more image schemata, which form the basic framework of our embodied knowledge. **Embodied knowledge** is simply the understanding that results from repeated experience with our bodies' sensory apparatuses and motor programs. Essentially, the source domain of conceptual metaphor consists of knowledge obtained directly from bodily, environmental, and cultural experience. The **target domain**, on the other hand, is an abstract, intangible, and/or subjective concept something that we do not learn about directly through sensorimotor activity and environmental interaction. In a process that Lakoff and Johnson call **cross-domain mapping**, we structure our knowledge of the abstract target domain in terms of an inherently compatible and well-defined source domain. The cognitive mapping occurs *from* the source domain *to* the target domain, and the result of cross-domain mapping is conceptual metaphor.

Consider the following example of a pervasive conceptual metaphor: THEORIES ARE BUILDINGS.¹⁸⁹ Here, the abstract mental conception of "theory" is made sensible by the importation and integration of familiar structural aspects of "building." Since a theory is not something we can actually see or touch, we create a mental correlation between a theory and something that we *can* see *and* touch—in this case, a building. So, we map (or, project) aspects of the source domain, building, onto the target domain, theory. This generally occurs

¹⁸⁹ This is the standard notation for a conceptual metaphor. The first word or phrase (here: THEORIES) is the target domain, and the last word or phrase (BUILDINGS) is the source domain. The connective word ARE (or, in some cases, IS) signifies the mapping between the two domains. Note that the mapping is actually backward in the propositional representation of a metaphor: cross-domain mappings always occur *from* the source domain *to* the target domain, just as an image schema is said to project *onto* abstract experience. The use of all capital letters is also a standard feature of metaphor notation.

automatically and unconsciously. The process of cross-domain mapping that produces conceptual metaphor typically occurs at a level below our reflective awareness.

A host of familiar linguistic expressions falls out of this conceptual metaphor: he "built" his theory on a firm "foundation," she "constructed" her theory on a solid "framework," his theory did not have enough "support" and it "fell apart."¹⁹⁰ Terms such as built, foundation, construct, framework, and support are borrowed from the domain of actual, physical buildings. It is critical to remember that these metaphorical linguistic expressions are secondary: they derive from the primary conceptual metaphor.

Lawrence Zbikowski states that "cross-domain mapping plays [a] very important role in musical understanding...it provides a way to ground our descriptions of elusive musical phenomena in concepts derived from everyday experience."¹⁹¹ Zbikowski identifies the metaphor PITCH RELATIONSHIPS ARE OBJECT RELATIONSHIPS IN VERTICAL SPACE, which structures our conception of pitches as located somewhere in a continuum from low to high. The source domain of this metaphor, "object relationships in vertical space," is grounded by the VERTICALITY schema, which arises from our bodily experience of uprightness as well as from activities such as climbing a staircase or a mountain and observing the level of water rise as we fill a drinking glass. The VERTICALITY schema structures such concrete experiences (or, such experiences "flesh-out" the schema) *and* structures our metaphorical understanding of pitch relations. But what is the experiential basis for our conceiving of pitch relationships in terms of object relationships in vertical space—why does the mapping work? There is one primary explanation: our embodied experience of producing pitches through singing and speaking. When we produce a low sound, our chest resonates.

¹⁹⁰ Lakoff and Johnson, *Metaphors We Live By*, 46.

¹⁹¹ Zbikowski, Conceptualizing Music, 64.

When we produce higher tones, particularly in our falsetto or "head voice," the sound seems to resonate from the head.¹⁹² Lippman claims: "Different parts of the body may respond preferentially to different frequencies of vibration...[high-frequency] noises often seem to make our head vibrate, and we feel low-frequency sounds in the middle of our body."¹⁹³ So, whether a sound is being perceived or produced, "low" pitches seem to occur lower in the body while relatively "higher" pitches occur in the upper portion of the body. It is through this specific mapping that pitch space—an entirely conceptual type of space—exists at all. This conceptual metaphor is so pervasive that it has become a cultural convention that is reinforced by practice and extension.¹⁹⁴ It is from this conceptual metaphor that verbal descriptions of musical events take on a spatial form: a "low" B-flat, a "descending" scale, the "leap" of a tritone, a "closely" related key, and so on. Again, these descriptive linguistic metaphors stem from the larger conceptual metaphor and reflect the cognitive mapping of events perceived in actual space (source domain) onto those in pitch space (target domain).

Other examples of pervasive music-conceptual metaphors include: A MUSICAL WORK IS A JOURNEY (the source domain of which is grounded by the PATH schema) and A MUSICAL FORM IS A CONTAINER. Brower's notion of "nested containers" certainly applies to the latter metaphor, since a formal structure typically "contains" sections that likewise contain

¹⁹² Ibid., 69.

¹⁹³ Lippman, "Spatial Perception and Physical Location as Factors in Music," 33.

¹⁹⁴ Zbikowski points out that some non-Western cultures do not conceive of pitch in terms of height. For instance, in Bali and Java pitches are not regarded as "high" or "low," but as "small" and "large," respectively. Would they perceive an ascending scale as the progressive shrinking of a pitch-object? The Suya of the Amazon basin conceive of pitches as "young" and "old" instead of as "high" and "low." Might they equate a descending scale with the process of aging or receive a Western work in the minor mode as being "older" than a major-mode piece? (See Zbikowski, *Conceptualizing Music*, 67-68). While the members of these cultures do not structure their understanding of pitch relations by means of the PITCH RELATIONSHIPS ARE OBJECT RELATIONSHIPS IN VERTICAL SPACE conceptual metaphor, their conception of pitch is the product of different conceptual metaphors that are grounded in embodied knowledge. The correspondences between these different metaphors are quite obvious: a *young* child has a *smaller* resonating cavity and therefore produces a *higher* pitch (young = small = high). An analysis of the conditions and circumstances that enable the generation and proliferation of a certain conceptual metaphor within a specific culture would be intriguing.

subsections. The affect produced by melodic "direction" derives from metaphors such as TENSION IS ASCENDING MOTION and RELEASE IS DESCENDING MOTION. The source domain of each of these metaphors is grounded by the TENSION-AND-RELEASE schema detailed above, as well as by the VERTICALITY schema.¹⁹⁵ Similar conceptual metaphors exist for the experience of tension in other musical parameters: TENSION IS TEXTURAL EXPANSION, TENSION IS ACCELERATION, TENSION IS INCREASING LOUDNESS, TENSION IS DEPARTURE, TENSION IS SPATIAL EXPANSION, etc.

While Zbikowski and other theorists have applied the theory of conceptual metaphor chiefly to pitch relations, I shall apply the theory to the interpretation of spatial-gesture directionality.¹⁹⁶ To begin with, the three broad gestural categories (oscillation, open, and closed) all have profoundly delineated aesthetic implications. As discussed, a strong open gesture (such as SG-[0123]) and a strong closed gesture (such as SG-[013210]) present contrasting interpretive potentialities, simply due to their respective *openness* (incompleteness) and *closedness* (completeness). The incompleteness of an open gesture might be received as tensive (the perceived tension arising from the unfulfilled expectation for closure), while the completeness of a closed gesture might impart a sense relaxation. However, specific gestures within each category may have entirely different connotative meanings based on differences in directionality.

Since specific forms of directed motion cohere with concrete (embodied) or observed (sensational) experience, specific spatial gestures may serve as the source domain in a cross-

¹⁹⁵ These metaphors are dependent upon (and secondary to) the PITCH RELATIONSHIPS (IN PITCH SPACE) ARE OBJECT RELATIONSHIPS (IN VERTICAL SPACE) metaphor. We must first conceive of a melody as ascending or descending to then associate it with tension or relaxation.

¹⁹⁶ In his Ph.D. dissertation, Arnie Cox explores the metaphoric structuring of musical meaning. In particular, he applies the theories of Lakoff and Johnson to account for the listener's perception of motion within a conceptual, musical space. See Arnie Walter Cox, "The Metaphoric Logic of Musical Motion and Space" (Ph.D. diss., University of Oregon, 1999).

domain mapping. The target domain in the mapping will essentially be the interpretation of the source-domain gesture's directionality. For example, consider the following metaphor: PROGRESSION IS LEFT-TO-RIGHT MOTION.¹⁹⁷ We read text and musical scores from left to right, and we tend to otherwise represent diachronic sequences (such as the alphabet or a string of integers) with this specific directionality. Such experiences ground our understanding of left-to-right motion. Therefore, we might interpret a left-to-right panning gesture (L-R PAN) as a sign of progressive motion. Now consider the converse of this metaphor: REGRESSION IS RIGHT-TO-LEFT MOTION, which may be active during the reception of, and applied in the interpretation of, a R-L PAN. Since right-to-left motion seems to operate against sequential progression (as such sequences are typically depicted graphically), it is interpretable as regressive motion.

Although a L-R PAN and a R-L PAN are members of the same gestural category (they are both strong open gestures; either is the retrograde or the inversion of the other), they have dramatically different aesthetic implications; in fact, they have quite opposite "meanings." Furthermore, a left-to-right motion might be received as tensive since "progression" implies effort (the application of force) on the part of some agent. Regression might result from applying force in the opposite direction (which would likewise be tensive), from not applying any force at all, or from ceasing the application of force (and thus succumbing to an outside

¹⁹⁷ This is very much a Western metaphor that is certainly not shared by all cultures. However, this metaphor is likely becoming increasingly pervasive and universal via the globalization of Western socio-cultural institutions. I have conscientiously opted to employ the term "progression" instead of "progress." "Progression" simply signifies spatial or temporal advancement or procession, while "progress" implies the gradual advancement (actual or metaphorical) to a better (improved) state of being or the acquisition of some positive benefit. Thus, "progression" is a slightly less loaded word than "progress." Although it is quite easy for the Western mind to equate progression and progress, left-to-right motion is not *necessarily* qualitatively superior to right-to-left motion. I will not claim that a L-R PAN is "good" while a R-L PAN is "bad."

force), in which case it might be associated with relaxation.¹⁹⁸ Thus, we might embellish our notation for the directional metaphors listed above as such: PROGRESSION (TENSION) IS LEFT-TO-RIGHT MOTION and REGRESSION (RELEASE) IS RIGHT-TO-LEFT MOTION.

It is interesting to recall a common translation of the "score space" to the ensemble space already mentioned in earlier chapters: the top-to-bottom ordering of staves often translates to a left-to-right series of ensemble points (with bottom-to-top translating to right-to-left). It is quite common for the upper staff (first part) of a work for ensemble to be performed by the leftmost point in the ensemble space, the bottom staff by the rightmost point, and so on. Such a "perfect" score space-to-ensemble space translation (a 1:1 correspondence between the vertical ordering of staves in the score and the horizontal series of points in the ensemble space) often yields a correlation between spatial gestures and motion through registral space: if the score is arranged with higher-pitched instruments in the upper staves and lower-pitched ones at the bottom of the staff system (as is typical), then a bottom-to-top shape in the score (spanning multiple parts/staves)—which entails a broad ascent through registral space—becomes a right-toleft panning gesture when actualized in ensemble space.¹⁹⁹ A bottom-up figure in the score does not always produce an ascending motion in pitch space. If it does, then the metaphors TENSION IS ASCENDING MOTION (in pitch space) and REGRESSION (RELEASE) IS RIGHT-TO-LEFT MOTION (in ensemble space) would be in conflict with one another. Such "interdomain interpretive discordance" shall be addressed in the final section of this chapter.

¹⁹⁸ Motion in general (i.e., *any* directionality of spatial gesture) may be considered tensive due to the agency (force) required to cause said motion. For this reason, right-to-left motion may be construed as tensive. However, through metaphoric associations (such as those described above), left-to-right motion may indeed be considered more tensive than right-to-left motion—at least in isolated contexts.

¹⁹⁹ The string quartet and the guitar quartet are common examples of a perfect score space–to–ensemble space translation. However, plenty of exceptions to such a "neat" translation exist—brass quintets and woodwind quintets, for example.

The above observations apply to solitary instances of either a L-R PAN or a R-L PAN. In the case of a sweeping gesture, the first portion of the gesture might be considered progressive, and the latter portion regressive, irrespective of either leg's directionality. Consider a RIGHT SWEEP (SG-|432101234|). The initial, component panning gesture of this compound gesture SG-|43210|, is a R-L PAN. The subsequent L-R PAN (SG-|01234|) returns (or, "regresses") to the gesture's point of initiation (ensemble point 4). Thus, when considered within the larger motional process of the RIGHT SWEEP, R-L PAN is progressive (and tensive) and L-R PAN is regressive (and relaxive); the sense of spatio-gestural progression and regression is contextual.

As previously noted, *circular gestures* may arise in electronic or acoustic surround environments (such as a quadraphonic electronic composition, Xenakis's *Persephassa*, etc.). Metaphor likewise plays an important role in interpreting the directionality of such gestures. Listeners from many cultures would be inclined to associate a clockwise circular gesture with progression, since the hands of a clock (a pervasive element of experience) travel in a clockwise fashion—representing the diachronic "progression" of time. The metaphor PROGRESSION IS CLOCKWISE CIRCULAR MOTION informs our aesthetic reaction to clockwise motions. On the contrary, a counterclockwise gesture might be interpreted as regressive—represented by the metaphor REGRESSION IS COUNTERCLOCKWISE CIRCULAR MOTION. The idea that left-to-right linear motion and clockwise circular motion both signify progression is not surprising and is anchored in experience: when a rolling object such as a wheel (on a car, bicycle, etc.) passes from the left to the right of our visual field, it revolves in clockwise fashion. Additionally, consider Johnson's figures for both the PATH schema (Figure 5-3) and the CYCLE schema (Figure 5-12). The PATH-schema figure is organized such that the source of

332

the path is to the left, the path's goal is on the right, and the line connecting these points proceeds from left to right. The figure for the CYCLE schema indicates that an abstract cyclic process progresses in a clockwise fashion. While Johnson does not qualify the directionalities latent in his figures, it is probable that he organized them intuitively based on the underlying conceptual metaphors PROGRESSION IS LEFT TO RIGHT MOTION and PROGRESSION IS CLOCKWISE CIRCULAR MOTION. In view of these observations, the dissolution of the prominent counterclockwise circular gesture in *Persephassa* may be interpreted as the abolishment of regressive activity.

As formerly noted, the third dimension of a three-dimensional music-conceptual space is often provided by dynamic level, with loud sounds appearing close and soft sounds seeming distant. The equation of intensity and location (loud with near and soft with far) provides the grounds for interpreting the motion of gestures occurring within ensemble spaces comprising multiple depth planes. In such cases, an actual second dimension exists in the ensemble space that is structured via the NEAR-FAR schema. Just as a crescendo is typically received as tensive, an advancing (back-front) gesture likewise evokes tension. A crescendo is a metaphorical advance in musical space, and an advancing gesture is an actual advance in real (ensemble) space. Both advances amount to an increase in spatial proximity. Conversely, a receding (front-back) gesture might imply relaxation—as does a decrescendo. A large-scale advancing gesture occupies the entirety of Xenakis's *Empreintes* (introduced in Chapter 4). The spatial construct of this work comprises three distinct depth planes, each differentiated by timbre: strings (plane A), woodwinds (plane B), and brass (plane C). The opening gesture, plane C \rightarrow plane A, is a disjunct advancing gesture-skipping plane B. Plane B enters in m. 99 and persists through the end of the work, as planes C and A gradually depart from the musical activity (in m.

333

109 and m. 126, respectively).²⁰⁰ Thus, the overall spatial activity of the work is plane C \rightarrow plane B.²⁰¹ The spatial texture of the entire work is accumulative from mm. 1–99 and dissipative from mm. 109–147 (end), with an ALL PLAY spanning mm. 99–109. The spatial texture therefore exhibits tension and release on two levels: as the sonic space expands (accumulates) and contracts (dissipates) within the ensemble space, and as the tensive advancing motion (plane C \rightarrow plane A) is followed by relaxive receding motion (plane A \rightarrow plane B). However, the long-range (background level) planar advance (plane C \rightarrow plane B) is tensive. Despite the spatial relaxation achieved through the dissipation of the sonic space (from three planes to one), each plane undergoes a textural crescendo as divisi indications systematically augment the textural density.²⁰² This joint tensing activity in the realm of texture and in the spatial realm of each of the three depth planes complements the tension induced by the overall advancement of the sonic space.²⁰³

In the case of works exhibiting multiple vertical levels, as in Henry Brant's *Desert Forests* (introduced in Chapter 4), the same image schema structuring our reception of pitch events, the VERTICALITY schema, structures our conceptualization of actual, spatial events. No metaphor is involved when we encounter a *rising* (D-U) or a *falling* (U-D) gesture. In fact,

²⁰⁰ A portion of plane C (horns only) reappears fleetingly in mm. 143–146 but is quickly subsumed by the dynamic ferocity and textural lushness of plane B.

²⁰¹ The spatial activity of the entire work begins at plane C and ends at plane B. However, the first "spatial advance" bypasses plane B as activity at plane C shifts to plane A. Plane A later recedes to plane B. This spatial motion, plane C \rightarrow plane A \rightarrow plane B, is effectively a "spatial appoggiatura." The long-range (in the temporal sense of the word) target of the planar spatial movement is plane B, but this goal is initially *overstepped*. The spatial "gap" is subsequently "filled in" by spatial motion that *steps* back in the opposite direction (conjunct receding gesture A \rightarrow B). Considering that the English translation of the work's title (*Empreintes*) is "footprints," speaking figuratively of the work's spatial process as an agent-in-motion that "oversteps" plane B and subsequently "steps back" to that plane is quite appropriate.

²⁰² The divisi indications result in the fragmentation of each depth plane: small areas (an area is a group of coordinated points) within each plane are "split" into multiple points, such that the point cardinality of each area increases (but the space occupied by each area remains constant). As a given area splits, its component points gain some degree of independence. Such activity is structured by the SPLITTING schema and is tensive.

²⁰³ This brief spatial analysis is not intended to account for all of the spatial activity present in *Empreintes*. A more thoroughgoing analysis would take into account the smaller gestures occurring within each depth plane and correlate these with the larger gestures identified above.

such spatial observations (aural and visual) are what form the VERTICALITY schema that grounds our metaphorical conception of pitch activity (such as ascending and descending melodic motion). The experience of a rising or falling gesture simply fleshes-out an already existent image schema. However, the metaphor TENSION IS ASCENDING MOTION generates the affect of tension when we encounter a rising gesture. Rising gestures are dominant in *Desert Forests* and are left virtually unchecked due to the relative absence of falling gestures. However, the spatial ALL PLAY (between the three vertical levels) toward the end of the work dissipates (in section 7j) such that only the lowest spatial level is active during the final two sections of the work (sections 7k–7L). This lowest level is the first level to be activated in the entire work and is responsible for the instigation of numerous rising gestures. Therefore, in spite of the prevalence of tensive rising gestures, a sense of departure and return encompasses the large-scale spatial activity of the entire work. Due to the ensemble-space structure of *Desert Forests*, rising gestures are correlated with advancing (back-to-front) gestures of a peculiar type: since the audience is for the most part underneath the diagonal sonic plane of the ensemble space, the gestures do not necessarily "advance" as normal back-to-front gestures do. In fact, depending upon the listener's location within the performance space, such gestures might advance and then recede as they "pass" the listener while on the rise to the far back balcony.

In light of the foregoing discussion, we find a degree of interpretive parallelism between different spatial gestures occurring within different types of ensemble spaces. By equating progression with tension and regression with relaxation, different gestures can be interpreted as producing a similar or identical affect. For example, the following spatial gestures may be considered progressive (and thereby tensive): left-to-right panning gestures, clockwise circular gestures, back-to-front (advancing) gestures, and down-up (rising) gestures. Conversely, right-

335

to-left panning gestures, counterclockwise circular gestures, front-to-back (receding) gestures, and up-down (falling) gestures may be interpreted as regressive and relaxive. The similar interpretations of quite different gestures may be attributed to the types of motion that qualify as progressive/regressive (tensive/relaxive) from an embodied perspective. Of course, these construals apply to abstract gestures divorced from context; a surface-level tensing gesture might actually be seen to function as the "return" portion of a larger departure-and-return scenario, which would ultimately yield a sense of relaxation. The manner in which spatial gestures might interact with events in other musical parameters is the subject of the next and final section.

5.5 The Interaction of Spatial Gestures with Other Musical Events

In the brief exploration of *Empreintes* by Iannis Xenakis (above), some mention was made of the congruous activity between textural processes and spatial gestures. The graduated textural densification of each depth plane and the large-scale advancement of the sonic space were interpreted as mutually tensive. The affective qualities of musical events in different parameters of a work may be either complementary (as in *Empreintes*) or contradictory. Complementary events may bolster the affect (and enhance the reception) of either tension or release, while contradictory events may produce a discordant, cancellation effect—whereby a tensive event in one realm "cancels" a relaxive event in another realm (or vice versa). We might think of such inter-parametric relationships as either consonant (complementary) or dissonant (contradictory). In so doing, we devise an important means of connecting and comparing musical events in different parameters of a musical work. While this line of thought shall not be developed in full in the present section, the subsequent analytic chapter shall highlight such relationships.

Of primary consideration is the manner in which spatial gestures might interact with concomitant melodic gestures in pitch space. In fact, we might conceive of an entirely new class of gesture that results from coincident trajectories in both ensemble and pitch space. In her assessment of the spatial activity of *Persephassa*, Harley identifies a descending circular gesture:

It is the timbre of the first layer of drum tremolos that is the most varied: all available skin instruments are ordered in 6 groups according to their relative pitch (higher bongos, lower bongos, high, medium and low tom toms with timpani, low bass drums). The rotating sound gradually descends, with each level presented during at least one complete revolution.²⁰⁴

Obviously, the "rotating sound" occurs in actual space while the "descent" occurs in pitch space. However, the coordination of these two activities yields a **composite gesture** that fuses two distinct musical parameters: the circular path in ensemble space and the descending path in pitch space collaborate to produce a "descending circular gesture."

In addition to forming composite gestures, concurrent gestures in actual space and in metaphorical space might work in accord or might controvert one another.²⁰⁵ For example, imagine the simultaneous occurrence of a closed spatial gesture with a descending melodic line (a melodic shape that typically denotes closure). The received sense of closure would be potentially compounded by the corresponding events in both actual space and pitch space. Alternatively, events in different parameters of a work might be in conflict with one another. Depending upon context and perceptual salience, events in one parameter might override those in another.²⁰⁶ Under certain conditions, the closure provided by a melodic line descending from scale degree 5 to the "stable floor" of the tonic might assume prominence over the implied

²⁰⁴ Harley, "Space and Spatialization," 295.

²⁰⁵ Spatial gestures are often intimately bound with *Klangfarbenmelodie*. The conjoining of spatial and timbral themes or motives heightens the perceptual salience of each.

²⁰⁶ The perceptual salience of events within any musical parameter might be affected by decisions made on the part of the composer (orchestration, the provision of ensemble diagrams), the performers (seating arrangement, spacing, dynamic interpretation), or by concert hall conditions (the listener's location relative to the ensemble space within the performance space).

continuation suggested by a contemporaneous open spatial gesture. Or, the large-scale rhetoric of a work might involve the gradual resolution of inter-parameter tensions, such as those provided by conflicts between events in pitch space and those in physical space. Such inter-parameter relationships—be they complementary or contradictory—certainly require hermeneutic untangling in analysis.

An overt example of complementation between concurrent musical events in multiple musical domains occurs in a passage toward the end of György Ligeti's *String Quartet No. 1* (*Métamorphoses nocturnes*). The passage is provided in Figure 5-23.



Figure 5-23: Ligeti, String Quartet No. 1 (Métamorphoses nocturnes), mm. 1207-1210

In this brief excerpt, which immediately precedes the final five bars of the quartet, tensing activity transpires jointly in the domains of pitch, dynamic, texture, tempo, and space. The dynamic markings indicate the massive crescendo that occupies this musical segment. As the instruments enter successively, the overall pitch range expands upward—ultimately spanning over four octaves. In addition to this broad ascent, each instrument's activity upon entering

consists of an ascending melodic interval—the measurement of which expands with each consecutive entrance: cello (m7), viola (M9), violin 2 (m10), violin 1 (M10). An accelerando is effected both by the reduced temporal "space" between each instrument's statement of the ascending interval (three beats between the cello and viola, one-and-a-half beats between the viola and second violin, and three quarters of a beat between the second and first violin) as well as by the reduced rhythmic values of the first pitch of each interval (a quarter note in the cello and viola, an eighth note in the second violin, and a sixteenth note in the first violin). As each instrument enters, it sustains the second note of the interval-resulting in an increase in textural density. As the texture thickens, a dissonant [0136] tetrachord accrues prior to the sforzando chords that conclude this passage. The passage comprises an accumulative R-L PAN (SG-[3210]), provided that the ensemble is seated according to the string-quartet seating plan assumed thus far (from left to right: Violin 1, Violin 2, Viola, Cello). This accumulative gesture—which undergoes a composed-out, *intragestural acceleration*—entails an expansion of the sonic space into its full spatial potential: the total ensemble space.²⁰⁷ As discussed earlier in this chapter, all of the above musical activities are tensive: increasing dynamic level, increasing (composed-out) tempo, rising pitch range, intervallic expansion, thickening of textural density, and an accumulative spatial gesture that result in the full expansion of the sonic space.

²⁰⁷ The accumulative R-L PAN dominates Ligeti's quartet. The opening gesture of the entire quartet (mm. 1–7) and the work's final gesture (mm. 1211–end) are both right-to-left panning gestures. The R-L PAN is unchecked by any left-to-right motion, so it serves neither as the departure gesture nor as the return gesture of a larger departure-and-return scenario. In light of the REGRESSION IS RIGHT-TO-LEFT MOTION metaphor, the directionality of the gesture alone may be interpreted as relaxive, which disrupts the consonance of the inter-parametric tensing activity. However, since this is effectively the only directionality of panning gesture present in the work, it is likely received as tensive since it progresses along a spatial path and expands the sonic space in the process.

CHAPTER 6

A SPATIAL ANALYSIS OF BEETHOVEN'S STRING QUARTET NO. 14 IN C# MINOR, OP. 131

In this chapter, I apply the spatio-analytic methodology profiled thus far to account for the spatial relationships unfolding over the course of a multi-movement work. A spatial analysis of Beethoven's *String Quartet No. 14 in C# Minor*, Op. 131 (1826) reveals the tremendous unifying capacity of categorically related spatial gestures. I examine the recurrence and variation of specific gestures within each of the seven movements of Beethoven's Quartet and explore the structural interrelationships between gestures. I demonstrate how the profusion of such gestures lends spatial coherence—and thereby a potential for greater comprehensibility—to the various movements in which they reside. During the process of exploring intra-movement spatial unification, *inter*-movement gestural relationships emerge as well. Members from a given category serve specific syntactic functions, while emergent intercategorical relationships contribute to the unfolding rhetoric and "spatial narrative" of the music.

The interpretive framework of Chapter 5 is evoked to illumine the meaning potentialities nascent within the spatial complex of Beethoven's quartet: this analysis embodies interpretive concepts such as departure-and-return, balance, motion into and out of containment, and "directional meaning." The variable directionalities of certain gestures (panning gestures in particular) bear critical interpretive consequences at both local and global levels of the spatial construct. Of considerable importance, this analysis reveals how spatial gestures may unfold

over long spans of music and exist at deep levels of structure—illuminating the fact that spatial gestures are not merely surface phenomena.

Why elect this particular work by Beethoven to serve as the subject of a spatial analysis? Nearly all of the musical examples contained in previous chapters are from twentieth-century works. Of these examples, many come from works in which the composer is known to have purposely spatialized musical materials-often diagramming precisely how the ensemble should be configured within the performance space. It seems fitting to take a step back not only in historical time, but also in certitude as to the intent of the composer. Since we do not know how deliberate Beethoven was in designing and forging the spatial construct of his music, subjecting any of his works to spatial analysis raises fundamental questions regarding the sheer validity of such an analytic enterprise. In addition, we must posit a particular seating plan upon which to base our assessment of spatial activity. The following analysis presupposes the standard "celloout" arrangement, in which the discrete instruments of the ensemble assume the following point positions: Violin 1 (point 0), Violin 2 (point 1), Viola (point 2), and Cello (point 3). I have chosen the cello-out configuration for two reasons: first, a cursory examination of the work indicated to me that this particular seating plan would effectively draw out many of the gestures and gestural relationships outlined in the preceding chapters.¹ Second, the "perfect" score space-to-ensemble space translation facilitates both the identification and the indication of latent spatial gestures manifest as figures and shapes in the score. The analytic implications of alternative seating plans will be considered at a later point.

Composers have long viewed the genre of the string quartet as a venue for experimentation. In particular, the late quartets of Beethoven are cited for their novelty of

¹ Although lying beyond the immediate scope of this chapter, a comparative analysis of the spatio-gestural activity that emerges under different seating plans would be instructive. Such an analysis would be of use to an ensemble acknowledging that its choice of seating plan functions as an integral component of musical interpretation.

invention. Op. 131 proves to be a viable candidate for spatial analysis owing largely to its textural transparency, which lends itself to spatial clarity. While not wanting in contrast or disruptive musical activity, this maverick work boasts remarkable coherence and continuity: the movements elide into one another, motivic material transcends the already tattered movement boundaries, and complex harmonic relationships unfold over the course of the entire work.² It thereby proves difficult to segregate one part (movement) from the whole. I have therefore elected to analyze all seven movements of this quartet in an effort to expose both the spatio-gestural coherency and the large-scale spatial structure that embrace the entire work.

6.1 Movement 1

The principal gestural form is exposited in the opening 14 bars of the first movement, as presented in Figure 6-1. The fugal opening of the quartet presents a temporally expansive perfect L-R PAN (SG-|0123|). The fugue subject is initially presented at point 0.³ The "real" answer to the subject, unorthodoxly at the subdominant level, begins at point 1 in the pickup to bar 5 while the first violin (point 0) continues with the countersubject. The fugal entries carry on with point 2 entering in the pickup to bar 9 and point 3 entering in the pickup to bar 13.⁴ This opening panning gesture is *accumulative:* ensemble points remain active once they have commenced their sonic activity. Following the entrance of the cello, the entire ensemble engages

² Harold Truscott, *Beethoven's Late String Quartets* (London: Dobson Books Ltd., 1968), 100–119.

³ Robert Hatten discusses the tragic quality of this fugue subject: "The structure of the four-note subject head is analogous to the dramatic scheme of tragedy." Hatten qualifies this interpretation by claiming that the initial "aspiration [of the leading tone resolution to tonic] suffers the reversal of a tragic drop [the sforzando emphasis on the lowered submediant scale-degree] as a brutal intrusion of tragic reality." Robert S. Hatten, *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation* (Bloomington and Indianapolis: Indiana University Press, 1994), 152. Hatten discusses the opening movement of Op. 131 in detail; see Hatten, *Musical Meaning in Beethoven*, 145–160.

⁴ Michael Steinberg notes that it is extremely rare for Beethoven to begin a multi-movement work with a slow movement, and that for him to begin with a fugue is "without precedent." See Michael Steinberg, "Notes on the Quartets," in *The Beethoven Quartet Companion*, ed. Robert Winter and Robert Martin (Berkeley: University of California Press, 1994), 246.
in an ALL PLAY scenario until bar 67. Thus, the opening 13 measures witness the sonic space's *expansion* into the total ensemble space, in a motion beginning from the leftmost ensemble point and proceeding, in unidirectional fashion, to the rightmost point. We shall see throughout the quartet that such expansion of the sonic space often directly precedes substantial ALL PLAY sections.



Figure 6-1: Beethoven, Op. 131, Mvt. 1, mm. 1-14

It is not surprising for the emergence of a spatial gesture to accompany and result from imitative activity—particularly that of fugal process. However, the successive registral ordering of the voice entrances *is* noteworthy, as they enter in series from highest to lowest voice:

soprano, alto, tenor, bass. This is not the most historically common ordering of fugal voice entries.⁵ The "spatial translation" of this sequence of voice entries (given the "cello-out" seating plan of the string quartet) is a perfect L-R PAN. The rightward-bound dilation of the sonic space is coextensive with a progressive descent through pitch space; the unidirectional spatial motion is complemented by a unidirectional registral descent at the global level of the fugal exposition, considering the pitch level and basic range of the fugal entries. Throughout the quartet, however, a descent through pitch space does not always accompany a L-R PAN.

In m. 67, the ALL PLAY comes to a close as points 2 and 3 fall silent. The violins imitate one another with a metrically displaced fragment of the fugue subject for 5 bars (mm. 66–73), yielding strong adjacent-point oscillation gesture SG-|010101010| (see Figure 6-2).⁶ Point 2 enters immediately prior to point 3 in bar 72 with a metrically rearranged fragment of the topic of the ongoing imitation (which is itself a fragment extracted from the tail of the fugue subject). Therefore, mm. 71–73 present SG-|0123|, another instance of L-R PAN (indicated with a dashed arrow in Figure 6-2). This panning gesture is a temporally diminished version of the movement's opening gesture. Points 2 and 3 continue the antiphonal activity previously stated at points 0 and 1 (mm. 72–78), and points 0 and 1 cease their musical activity in bar 74 (after only one-and-a-half bars of accumulated activity). SG-|2323232| emerges in mm. 72–78.⁷ This gesture, which is another strong adjacent-point oscillation, is a right-nudged (+2) and temporally

⁵ Thomas Benjamin observes that the two most common orders of voice-entries in the fugues of J. S. Bach are: 3-2-1-4 and 4-3-2-1. In the *Well-Tempered Clavier*, voice 1 never enters first. It is common, however, for voice 4 to enter last. Beethoven's voice-entries in the first movement of Op. 131 (1-2-3-4) are in the retrograde order of one of Bach's favored voice-entrance schemes (4-3-2-1). Beethoven does preserve many of the musical characteristics traditionally associated with a four-voice fugue (as opposed to a three-voice fugue): graveness, a slow-moving subject, a narrow-ranged subject, and an altogether shorter subject. See Benjamin Thomas, *Counterpoint in the Style of J. S. Bach* (New York: Schirmer Books, 1986), 295–296.

⁶ In m. 67, violin 2 echoes the melody stated by violin 1 in m. 66. The nature of the imitation changes in m. 70 as the violins become more varied in their respective rhythmic and melodic activity. Nevertheless, their consistently staggered activity enables the retention of the call-and-response spatial texture through m. 73.

 $^{^{7}}$ It is difficult to pinpoint exactly where the imitation between the viola and cello terminates. By the time the second violin reenters in m. 79, the imitation between points 2 and 3 is subsumed by the thickening texture and expanding sonic space (mm. 79–81).

diminished version of SG-|010101010|.⁸ It is the nudging of SG-|010101010| from ensemble segment space 0–1 to segment space 2–3 that yields the L-R PAN (SG-|0123|) spanning mm. 71–73 (indicated with a dashed arrow in Figure 6-2). The direction of the nudging (rightward) is identical to the directionality of the panning gesture (left-to-right) linking the two related oscillation gestures. The surface-level oscillation activity at either segment of ensemble space is subordinate to this larger panning motion. Interestingly, the cello imitates the viola in a slightly higher register until bar 78, so this left-to-right panning gesture is *not* coordinated with descending pitch motion.

Point 1 returns in m. 79 and is soon joined by point 0 in m. 81; the violins each execute a further fragmentation of the fugue-subject fragment heretofore subjected to imitation. The spatial activity of bars 78–82 constitutes SG-|3210| (R-L PAN). The first event point of this gesture consists of the cello's diatonic cseg <012> in m. 78. The chromatic retrograde of this event, cseg <210>, serves as event point 2 (at ensemble point 2) in m. 79. The diatonic <012> reappears in the violins (mm. 79–82) to serve as the final two event points of the gesture (<012> occurs twice at ensemble point 1 before finally appearing at ensemble point 0). These spatially dispersed contour segments, which conspire to yield SG-|3210|, are indicated in Figure 6-2 by enclosure in a dashed square. This R-L PAN is the directional antithesis to L-R PAN. However, we might conceive of bars 71–82 as a single, broad gesture: SG-|0123210| (LEFT SWEEP). This sweeping gesture is nested within (and uncoupled by) two related and temporally linked oscillation gestures that together span approximately 16 bars of music. This passage engages the progressive fragmentation of the fugue subject—spatial and fugal processes are entwined. The

⁸ SG-|2323232| may also be viewed as the *double inversion* of SG-|010101010|, meaning that it is subjected to both type-1 and type-2 inversion. Type-1 inversion reflects SG-|010101010| across the total ensemble space, resulting in SG-|323232323]. Since this oscillation gesture is strong (i.e., "closed"), it is palindromic and non-retrogradable. However, type-2 inversion, which maps the gesture onto itself within the ensemble-space segment that it occupies, yields SG-|232323232]. This gesture is then diminished by two event points to yield SG-|2323232].

SWEEP itself features a *balance* of directionality: consequent right-to-left motion "balances" antecedent left-to-right motion.



The concluding motion of this gesture, SG-|3210| can be seen to balance the initial SG-|0123| of mm. 1-13 on a broader level of structure. Thus, local directional balance (mm. 71–82) is nested within more global directional balance (mm. 1–81). The DEPARTURE-AND-RETURN

schema is "fleshed-out" by the sweeping gesture—point 0 serving as the point departed from and eventually returned to. The SG-[3210] of mm. 73–81 is accumulative and leads into an 8measure ALL PLAY section (mm. 82–90). Thus far, we have twice witnessed the amassing of spatial activity—in the form of accumulative panning gestures—serving the syntactic function of introducing an ALL PLAY section.

The next prominent L-R PAN occurs in mm. 91–100 following the cessation of the preceding ALL PLAY section in bar 90 (see Figure 6-3).



Figure 6-3: Mvt. 1, mm. 91-100

Violin 1 begins sequencing a diminished fragment of the fugue subject in the pickup to m. 91 (this is the same fragment treated antiphonally in mm. 66–82). The second violin echoes this activity when it enters later in m. 91, and the viola enters with the fugue subject in its original

form (initial pitch profile and rhythmic durations) in m. 94. The cello finally enters in bar 99 with an augmented statement of the fugue subject. It is interesting to note the transformative process affecting the temporal proportions of the fugue subject at each successive instrumental entrance: the diminished subject fragments of the violins (points 0 and 1) lead to the subject prime form in the viola, which then beckons the augmented form in the cello. A "slowing-down" (or, temporal expansion) of the subject accompanies the spatial expansion of the sonic space. These spatial and temporal processes are accompanied by the descent through pitch space likewise exhibited by the movement's opening L-R PAN. As we might expect, once the gesture has amassed to the point that all ensemble points are active, an ALL PLAY section persists from bar 99 through the end of the first movement (m. 121).

The L-R PAN identified in mm. 91–99 emerges from the consecutive statements of *some* sonic event at each point in the sequence. In considering only the entrances of the subject head (traced with a dashed square in Figure 6-3), we may uncover an additional gesture superimposed on SG-[0123]. The viola (point 2) begins the subject in the pickup to m. 93. This entrance, which likewise counts as the third event point of SG-[0123], is quite salient since it follows nearly two full measures of silence (inactivity at point 2). The subject head appears at point 1 beginning in the pickup to m. 94. Although not directly preceded by silence, the distinctive contour and characteristic intervallic content of this melodic activity marks it as an "answer" to the subject head presented at point 2. The next entrance of the subject is at point 0 in the pickup to m. 99. Thus, partial open gesture SG-[210] spans from m. 93 to m. 99, co-occurring with SG-[0123] (mm. 91–99).⁹ The two gestures share event point 2 (in event slot 1 of SG-[210] and in event slot 3 of SG-[0123]). Since these two gestures are of the same kind (open) but of contrasting directionality, their coincident occurrence is a case of *crossed polystreaming* (see

 $^{^9}$ SG-|210| is a *partial* open gesture because it unfolds in ensemble-space segment 0–2.

Chapter 4 for an explanation of crossed streaming). The two gestures "cross" one another at ensemble point 2, which is the approximate midpoint of SG-|0123| and the generative point of SG-|210|. In actuality, event point 2 of SG-|0123| "splits"—it both spawns SG-|210| and continues along its original path to point 3 (the projected goal of the gestural path), completing the perfect panning gesture. In the context of polystreaming, SG-|210| may be considered a "gestural spur" of SG-|0123|.¹⁰

Some oscillation activity occurs within the ALL PLAY that closes the first movement. Imitative rhythmic activity (the passing of the eighth/dotted-quarter rhythmic cell) ensues between point 1 and 2 in m. 104, yielding SG-|1212...1| through m. 107 (see Figure 6-4). This strong adjacent-point oscillation gesture is categorically related to SG-|010101010| (mm. 66–73) and SG-|2323232| (mm. 72–78); it is the right-nudged version of the former and the left-nudged version of the latter. SG-|1212...1| contains a greater number of point shifts in more rapid succession than the preceding oscillation gestures; it is an *accelerated* form of oscillation gesture. The gesture likewise occurs at interior ensemble points framed by musical activity at points 0 and 3, whereas the former gestures were located on either half of the ensemble space while the other half remained silent. SG-|1212...1| is a "spatial compromise" between SG-|010101010| and SG-|32323|, featuring one event point from each erstwhile oscillation gesture.

The oscillation activity increases in complexity in mm. 107–113, where SG-|10101212020202021212| unfolds. This large gesture comprises four smaller gestures: SG-|10101| \rightarrow SG-|1212| \rightarrow SG-|202020202| \rightarrow SG-|21212|. All but one of these smaller gestures

¹⁰ An alternative reading of this section would be to count the cello's entrance with the augmented version of the subject in m. 99 as the final event point of SG-[2103]. Under this interpretation, SG-[0123] and SG-[2103] share two event points (2 and 3). Since SG-[2103] is a weak closed gesture (it is actually a displaced panning gesture), the concurrence of these two unlike gestures would constitute a case of *dissonant polystreaming*. I prefer the original reading: SG-[0123] emerges from the sheer entrance of event activity at each ensemble point, and SG-[210] takes form through the sequential occurrence of the subject head in its original metric configuration.

are strong (SG-|1212| is weak), and only one of the four is distant-point (SG-|202020202|). As I have segmented them from the larger gesture, adjacent subgestures share overlapping (or, "pivotal") event points: the last event point of SG-|1010<u>1</u>|, for example, is also the first event point of SG-|<u>1</u>212| (located in the second violin part of the first two beats of m. 109). Essentially, the shared event at point 1 functions as a "hinge," effectively relocating the ongoing oscillation by "swinging" it into a new segment of ensemble space.



Figure 6-4: Mvt. 1, mm. 104-113

The only distant-point subgesture, SG-|20202020|, occurs *within* the larger gesture: SG-|21212| prevails prior to the cessation of oscillation activity. SG-|10101212020202021212| has actually

spawned off of the aforementioned SG-|1212...1| (mm. 104–107): the initiating event (again at point 1) of SG-|10101| (pickup to m. 108) is the concluding event of the prevenient SG-|1212...1|. Thus, the comprehensive oscillation activity of mm. 104–113 yields SG-|1212...1010121202020 2021212|. This gesture begins and ends with events oscillating between points 1 and 2—a spatially stable (i.e., "balanced") pair of ensemble points (within the total ensemble space). Departure occurs within the large gesture as the adjacent-point oscillation relocates to points 0 and 1 (m. 108) and is later transformed into distant-point oscillation activity in m. 110. Although the subgestural bookends SG-|1212...1| and SG-|21212| are both locally strong (closed), the gesture as a whole is weak, reflected by the large-scale motion from point 1 to point 2, which, interestingly, is a left-to-right motion that complements the predominant L-R PANs.

A less salient gesture abides "behind" the pronounced oscillation activity of mm. 107– 112. In these bars, varied forms of the fugue's subject head trace their way through the ensemble. Beginning with the pickup to m. 107 (and spanning through bar 112), strong closed gesture SG-|020130| unfolds as these altered subject heads are passed from point to point, yielding a case of dissonant polystreaming with the compound oscillation gesture. Since each is composed of distinct and uniform events, the two gestures engaged in polystreaming do not share any overlapping event points; they are, however, interlaced—often "crossing" one another as point shifts in each respective gesture occur simultaneously. The gesture emerging from the succession of modified subject heads is less salient than the oscillation gesture described above for two primary reasons: first, the intervallic content, melodic shape (contour), rhythmic profile, and dynamic markings (namely, the placement of the sforzando marking in the "aspiration" stage of the head as opposed to within its "tragic drop" phase) of the subject head stray from the original (prime form) configuration of the subject—these changes render it somewhat

351

inconspicuous.¹¹ Second, the gesture as a whole is less rhythmically active than the more rapidly ensuing oscillation gesture with its bouncy rhythm. For these reasons, SG-|020130| functions as a gestural backdrop upon which a portion of SG-|1212...10101212020202021212| unfolds. Note that a subgesture of SG-|020130|, namely SG-|020|, is a strong distant point oscillation gesture—suggesting a structural consensus between gestural "figure" and "ground."

In spite of these complexities, oscillation activity remains subsidiary to panning activity: it is part of a larger panning gesture in mm. 66–78, which in turn, is part of an even larger sweeping gesture spanning mm. 66–81. Oscillation gestures otherwise appear within an ALL PLAY scenario in mm. 104–113, rendering them less salient than the panning gestures, all of which have emerged from a reduced sonic space and accumulated to generate ALL PLAY. This intercategorical dynamic—strongly exposited in the opening movement—shall be opposed and tested throughout the quartet.

In the broad context of the entire first movement, we can clearly see that the LEFT SWEEP of mm. 66–81, particularly with its component R-L PAN, served as the "spatio-gestural departure" to the prominent L-R PANs at the beginning and toward the end of the movement. Within the spatial complex of this movement, a sense of departure-and-return is evinced on three levels: the sweeping gesture of mm. 67–81, the shifting directionality of the panning gestures encompassing the entire movement: L-R PAN \rightarrow LEFT SWEEP (L-R PAN + R-L PAN) \rightarrow L-R PAN, and the subsidiary oscillation activity of mm. 104–113.

¹¹ Hatten states: "The subject head can be 'defused' to some degree by transformations [in which] the drop to [the lowered sixth scale degree] is made less intense by the substitution of either [the raised sixth scale degree] or the dominant scale degree" (Hatten, *Musical Meaning in Beethoven*, 156). The latter substitution occurs with the second violin's rendering of the subject head in m. 110, where the "tragic drop" falls back to the dominant scale degree (G-sharp) prior to stepping up to achieve the A-natural. On other occasions, the "drop" in the transposed (sequenced) versions of the subject head falls *past* the local dominant tone altogether, as in mm. 106–108, where the first violin's D-natural apex drops down a tritone to G-sharp. The mitigation of the subject head wrought by replacing the lowered sixth scale degree with the raised sixth scale degree can be found in the first violin's part in mm. 98–100 (Figure 6-3). The "defusing" of the subject head resulting from scale-degree substitutions, rhythmic reconfiguration, and dynamic displacements serves to disguise the fugue subject.

6.2 Movement 2

Movement 2 begins with the same ALL PLAY spatial scenario with which mvt. 1 concluded. However, the ALL PLAY activity soon evolves to include a series of oscillation gestures that challenge the primacy of the panning gestures. In fact, as we progress through our survey of the quartet, we shall discover that the oscillation activity in the interior movements of op. 131 serves as a departure from the predominant panning motion of the outer movements. Hints of oscillation activity occur in mm. 1–160, where two points (one of which might be a fusion of two points, or "superpoint") engage in brief episodes of antiphonal activity-usually within the context of ALL PLAY.¹² One such instance occurs in bars 13–16 (see Figure 6-5). In this excerpt, violin 1 is delivering the movement's bouncy theme. The accompaniment is divided among points 1–3. Points 2 and 3 are "fused" since they are performing in rhythmic unison and the periodic onsets of their spatial activity are coordinated. This superpoint is indicated in the score by the encapsulating square spanning the staves of the viola and cello. Superpoint 2+3 alternates regularly with point 1, which performs the same rhythm as the superpoint on the upbeat of each measure. The result is strong adjacent-point oscillation gesture SG-|(2+3)1(2+3)1(2+3)|. Of course, the salience of this oscillation gesture is lessened somewhat by the ongoing melodic activity at point 0. This fragment (mm. 13-16) is actually part of a larger and more complex oscillation gesture spanning m. 8 through m. 16: SG-|10(1+3)(0+1)(1+3)0(1+3)0(1+3)(2+3)1(2+3)1(2+3)(1+2). The transformative process of *compression* occurs within this complex oscillation gesture. Subgesture SG-|(1+3)0(1+3)0(1+3)|

¹² A *superpoint* consists of two (or more) adjacent ensemble points. Nonadjacent points engaged in synchronized spatial activity may be fused, but they do not form a superpoint. Point fusion involving nonadjacent ensemble points often obscures the distinction between conjunct and disjunct spatial activity and might indicate the presence of polystreaming, wherein two or more spatial gestures occur simultaneously (as outlined in Chapter 4). Throughout this chapter, superpoints are indicated graphically either by a solid square enclosing the musical activity at adjacent staves or by a bracket showing the joint entrance of continuous musical activity at two points. Noncontiguous fused points are indicated in the score by a dashed line connecting the enclosed musical activity in the nonadjacent staves.

unfolds in mm. 10–12 and is immediately succeeded by subgesture SG-|(2+3)1(2+3)1(2+3)1(2+3)1(2+3)| (2+3)| (mm. 13–16). Thus, the large oscillation gesture is compressed *from* the total ensemble space *into* ensemble-space segment 1–3 as fused points 1 and 3 collapse to become superpoint 2+3 and activity at point 1 replaces that at point 0.



Figure 6-5: Mvt. 2, mm. 8-17

In mm. 17–23, points 1 and 2 perform in rhythmic unison, reflecting a reconciliation of the foregoing oscillation activity. The "reconciliation" of a migratory spatial gesture (such as the above oscillation gesture) occurs as previously consecutive (and potentially alternating) events at

different points are recast so as to occur in temporal simultenaeity. As described in Chapter 5, *spatial reconciliation* is structured by the MERGING schema.¹³

The musical nature of the event being spatialized through the ensemble in mm. 8–16 warrants further exploration. The three eighth-notes rhythmic configuration remains a constant. In the majority of instances, the exact same pitch is stated a total of three times at each point. Occasionally, however, the second of the three pitches serves as an upper neighbor to its framing pitches. One of several examples is the figure in the first violin part on the second dotted-quarter beat of m. 9, where pitch B neighbors away from, and subsequently back to, pitch A. The <010> contour, which is scattered at different pitch levels throughout mm. 8-14, is a pitch-space corollary to the oscillation gesture occurring in ensemble space. Inspect the last beat of m. 8 through m. 9: if we consider only the oscillation of the neighboring rhythmic figure and discount the repeated-note figure on the last beat of m. 9 in the second violin part, we observe SG-|0(1+3)0|. If we reduce-away point 3 by splitting it from fused points 1 and 3, we may analyze this brief subgestural portion of music as SG-[010]. The notational similarities between cseg <010> and SG-|010| are guite obvious: they contain the same sequence of integers.¹⁴ Indeed, the two musical entities are inter-parametric analogs (between the pitch-space and the ensemblespace musical parameters). The importance here is not that the contour-segment and spatialgesture labels both contain the integer sequence "010," but rather that this passage (along with

¹³ As noted in Chapter 5, *spatial reconciliation* affects migratory gestures that precede an ALL PLAY section (within the total ensemble space or an ensemble-space segment). This differs from accumulative gestures that usher in ALL PLAY via the progressive expansion of the sonic space.

¹⁴ SG-|010| will not always be a direct pitch space-to-ensemble space translation of cseg <010>. In contour theory, the interval between integer 0 and integer 1 could be much larger than a step (and thereby involve nonadjacent pitches in pitch-space) since the cseg label accounts for relative pitch height only. On the other hand, SG-|010| *always* indicates an oscillation between adjacent ensemble points. In this example, however, the neighboring motion reflected by the cseg label *is* actually stepwise (within the D-major scale), so the correspondence between neighboring pitch motion and adjacent-point oscillation activity is quite pronounced.

others in the quartet) reveals a concordance (in the form of neighboring activity) between objects in pitch space and events in ensemble space.



Figure 6-6: Mvt 2, mm. 84–92

I shall not attempt to divulge every single instance of oscillation activity in this movement; instead, I shall highlight several salient occasions that contribute to the movement's spatial scheme (which, again, functions within the larger scope of the entire quartet's spatial design). SG-|(3+1)13131(3+1)1(3+1)23232(3+2)2| occurs in mm. 84–91, as depicted in Figure 6-6. This elaborate gesture begins with the mutual onset of event activity at points 1 and 3.

These two points converse for approximately four bars, occasionally overlapping one another, as indicated by the three instances of (3+1) in the SG-label. In m. 88, the oscillation gesture contracts as point 1 is replaced by point 2 in the ongoing dialogue with point 3.¹⁵ This process of compression transforms the gesture from a distant-point oscillation gesture (featuring a spatial disjunction) into an adjacent-point oscillation gesture (featuring conjunct spatial motion). Note the similarity between the rightward compression of this complex oscillation gesture and that of the preceding gesture in mm. 8–16. The transformative procedure of contracting a spatial (oscillation) gesture to the right within the ensemble space is recurrent and thereby serves a unifying, motivic function. The left-to-right directionality corresponds to the dominant directionality of the first movement's panning gestures.

Additionally, SG-|(3+1)13131(3+1)1(3+1)23232(3+2)2| concludes at point 2, which is the point initially "enclosed" by the gesture's instigative event: (3+1). Therefore, by applying the ENSEMBLE SPACE-AS-CONTAINER schema to ensemble-space segment 1–3, we see Out \rightarrow In motion *from* the container boundaries (points 1 and 3) *to* the central point (2). The articulation of "arrival" at point 2 is strengthened by the double stops in the viola part.

Strong adjacent-point oscillation gesture SG-|3(2+1)3(2+1)3| unfolds in mm. 123–126 and is echoed by its inversion, SG-|(3+2)1(3+2)1(3+2)|, in mm. 133–137 (see Figure 6-7). The spatial inversion is type-2 (restricted to ensemble-space segment 1–3), as superpoint 2+1 is reflected across point 2, becoming 3+2, and point 3 maps into point 1. We also observe a complementary contour inversion of the oscillating motive between these related spatial gestures. The motive in each discrete part of mm. 123–126 is cseg <010>, featuring the interval of a second (minor or major) between c-pitch 0 and c-pitch 1. The motive is harmonized in thirds

¹⁵ This change in oscillation activity coincides with the transfer of the principal melody from the viola (mm. 85–88) to the violins (mm. 88–93).

between fused points 1 and 2 (superpoint 2+1). In mm. 133–137, the motive is extended to cseg <1010> in its two consecutive occurrences at superpoint 3+2. Due to the metric placement of the motive, we may consider its first three pitches, csubseg <101> as inversionally related to cseg <010> of mm. 123–126, the metric placement of which is identical. Cseg <2130>, an intervallically expanded form of cseg <1010>, occurs at point 1 in mm. 134–137. Nevertheless, the inter-parametric concordance between inverted events in pitch space (contour) and those in actual space is quite striking in this passage.

These inversionally related spatial gestures are actually separated by a somewhat masked L-R PAN as SG-|0(1+2)(1+3)(0+3)| unfolds in mm. 126–130.¹⁶ This gesture is best viewed as a case of polystreaming: two separate streams begin at point 0 and progress to superpoint 1+2. However, in m. 128, the superpoint splits as activity at point 1 is sustained and point 2 progresses to point 3. Then, as point 3 remains active, point 1 returns to point 0. Due to the bidirectional split of superpoint 1+2, two concurrent gestural streams interact: SG-|0(1+2)3| (L-R PAN) and SG-|0(1+2)0| (strong adjacent-point oscillation gesture).¹⁷ The latter is the type-1 inversion (and event-cardinality reduction) of the preceding SG-|3(2+1)3(2+1)3| (mm. 123–126). Thus, the first appearance of panning motion in the second movement is obscured by a simultaneous oscillation gesture within the context of *dissonant streaming*, which occurs when two or more categorically different spatial gestures transpire at the same time. The ENSEMBLE SPACE–AS–CONTAINER schema is fleshed-out by this polystreaming composite: unidirectional motion into

¹⁶ I am considering the ascending melodic interval (cseg <01>) in this section to be the event that is spatialized (the interval is either a minor second or the leap of an octave). Therefore, I am ignoring the cello's entrance in m. 126 (consisting of a sustained B), which occurs coincident to the entrance of point 0—the gesture's initial event. If we were to consider fused points 0 and 3 as the gesture's first event, we would identify SG-|(0+3)(1+2)(1+3)(0+3)|. We could analyze this case of polystreaming as either *crossed streaming* between SG-|0(1+2)3| (L-R PAN) and SG-|3(2+1)10| (R-L PAN) or as *convergent/divergent mirrored streaming* between SG-|010| and SG-|323| (both strong adjacent-point oscillation gestures related by type-1 inversion).

¹⁷ For the sake of simplicity, SG-|0(1+2)0| does not indicate the truncation of superpoint 1+2 to point 1 that immediately precedes the final event point (0). The "de-fusion" of 1+2, which is actually part of the point-split in the context of polystreaming activity, does not affect the gesture's categorical membership.

the ensemble-container (Out \rightarrow In) is followed by bidirectional (and temporally staggered) motion to the ensemble-container boundaries (In \rightarrow Out).



Figure 6-7: Mvt. 2, mm. 123-137

Further oscillation activity occurs in mm. 161-168 (see Figure 6-8). However, the exact nature of this oscillation gesture changes during its course. Bars 161-164 reveal SG-|3(2+1)3(2+1)3(2+1)3(2+1)|. This gesture is the type-2 inversion of the oscillation subgesture presented in mm. 13-16: points 3 and 1 alternate in each gesture, but point 2 is now fused with point 1 (whereas it was previously fused with point 3). However, in mm. 165-166, point 2 splits

from point 1 and fuses with point 3, yielding: SG-|(2+3)1(2+3)1|. Measure 167 contains PS 3-(2+1), and the gesture concludes with PS (3+2)-(2+1) in m. 168. Together, mm. 167–168 yield SG-|3(2+1)(3+2)(2+1)|.



Figure 6-8: Mvt. 2, mm. 161–168

This fragment of the larger oscillation gesture is still an oscillation gesture in and of itself, due to the consistent oscillation of points 1 and 3. It is still an adjacent-point oscillation gesture because intermediary point 2 is always fused to one of the spatial extremes of the gesture (either point 1 or point 3). When point 2 defects from one point to another, it effectively inverts the spatial orientation of the gesture. The entire 8-bar example features SG-(2(2+1))(2+1)(2+1)

(2+1)(2+3)1(2+3)13(2+1)(3+2)(2+1)|, which is a *weak adjacent-point oscillation gesture*. Note that the prevailing motion of this gesture is right-to-left, as indicated by the outer event points in the SG-label: 3 (in event slot 1) and 1 (in event slot 16). The melodic activity at point 0 gradually falls into the same triplet rhythmic figure being reflected within points 1–3, but it is different enough in its perpetual melodic activity not to readily group with any of the three points engaged in the extended oscillation gesture. The spatial interplay between points 1–3 functions as a kinetic backdrop to the spatially static yet agile melody (which is mobile in pitch space) ensuing from point 0.



Figure 6-9: Mvt. 2, mm. 169-175

Figure 6-9 shows SG-|3(2+1)0| (R-L PAN) unfolding in mm. 169–175. Notice that a single pitch (F-sharp) is sustained at point 0 as point 3 instigates the gesture and is soon joined by superpoint 2+1 (shown bracketed in the figure). The first violin's F-sharp skips up to A-natural in m. 173. We might view this event as the completion of SG-|3(2+1)0| were there not a more viable candidate to serve as the proper "goal-point event" of this panning gesture. In m. 175, point 0 assumes the rhythmic and melodic profile of the other three points. The concordance of rhythmic activity and melodic contour between the three point entrances (one of

which is superpoint 2+1) results in a clearly discernable spatio-gestural stream. Following the completion of the R-L PAN in m. 175, the four points engage in a predominantly homorhythmic ALL PLAY from m. 175 through the end of the second movement (m. 198). The right-to-left directionality of this panning gesture tallies with the chiefly right-to-left motion of the weak adjacent-point oscillation gesture of mm. 161-168. There, however, the overall motion is delayed by the persistent vacillation between points 3 and 1 (with either of these points fused to intermediate point 2), and the motion does not span the entirety of ensemble space: the motion from point 3 to point 0 is interrupted, stalling at point 1. This spatial process promptly begins anew in m. 169—this time in the form of a panning gesture. The R-L PAN completes the motion through the entire ensemble space, in a slightly shorter time span (7 bars), and in a unidirectional manner. Although the last gesture of the movement is a panning gesture, it is the only salient instance of panning activity in this movement dominated by oscillation gestures. It is likewise in the opposite directionality of most of the first movement's panning gestures.

6.3 Movement 3

Movement 3 (the entirety of which is provided in Figure 6-10) seeks to re-establish L-R PAN as the quartet's principal gestural form following the ample oscillation activity and lone R-L PAN of movement 2. After the two introductory ALL PLAY chords in the movement's opening measure, a *migratory* L-R PAN unfolds in mm. 2–3. All panning gestures up to this point have been accumulative. Thus, with this gesture, the sonic space does not expand into the total ensemble space; it merely "migrates" from point 0 to point 3, traversing points 1 and 2 along the way.







Figure 6-10: Mvt. 3 (entire)

Following the ALL PLAY punctuation on beats 2 and 3 of m. 4 (mimicking the introductory figure), an entirely new gesture emerges: SG-[10132] (mm. 4-6). This weak closed gesture is likewise migratory in nature.¹⁸ Notice that the five-event gesture embeds two important three-event gestures: SG-101, perhaps a fleeting remnant of the oscillation activity from movement 2, and SG-1013, which is an *imperfect* (due to the disjunction of PS 1-3) L-R PAN. It likewise imbeds SG-|0132|, a weak closed gesture that will appear more prominently in later movements. Bars 6–7, taken together, yield SG-|32(1+0)|. This particular segmentation of spatial activity extends through the tempo change to Adagio in m. 6, and the musical activity of superpoint (0+1) in m. 7 is quite different from that of points 3 and 2 in m. 6. For this reason, the salience of this accumulative R-L PAN is less pronounced than SG-[10132], in which the musical activity at each successive point is unified by a rhythmic motive. Nevertheless, we might view measures 5–7 as constituting an imperfect LEFT SWEEP: SG-[0132(1+0)].¹⁹ As such, SG-[013] (mm. 5-6) functions as an *abbreviation* (via the *deletion* of an interior event point) of SG-|0123| (mm. 2-3). The remaining four measures (mm. 8-11) of this brief, cadenzalike movement bear no significant spatial gestures. However, PS 1-2 in the final bar functions as a reduction of a L-R PAN.

Considering the spatial process of movement 3 as a whole, we see that it begins to replicate, but does not definitively complete, the spatial process of movement 1. The first movement features two predominant L-R PANs that enclose a LEFT SWEEP, the right-to-left leg of which functions as a departure from the prevailing left-to-right directionality. In similar

¹⁸ Although this closed gesture is weak by definition, it is strengthened by the fact that its starting and ending event points are adjacent to one another within the ensemble space.

¹⁹ We can justify segmenting m. 5 from the motivic activity at point 1 in m. 4 because of the subtle rhythmic discrepancy between the measures. The second violin's rendering of the motive in m. 4 lacks the final eighth-note value shared by all of the motive forms in mm. 5–6. Consequently, the eighth rest at the beginning of m. 5 provides a slight temporal segmentation between m. 4 and m. 5.

fashion, movement 3 begins with a prominent L-R PAN, departs with the LEFT SWEEP, but then closes with a left-to-right directed point shift (PS 1-2). Movement 3 may be viewed as a microcosm of movement 1—temporarily serving to reinstate the sovereignty of panning gestures in general and left-to-right motion, specifically.

6.4 Movement 4

The Quartet's central movement, movement 4 (a theme and variations), achieves a nearperfect balance between oscillation and panning gestures, in terms of both salience and frequency of occurrence. Oscillation activity between the violins begins in the first bar, and SG-|010...0| spans mm. 1–16 as the violins imitate one another first in the same (overlapping) registration (mm. 1–8) and then in registers separated by one octave (mm. 9–16). With the separation in pitch space beginning at m. 9, the spatial separation of the violins in the ensemble space becomes all the more evident. At measure 17, the spatial dispersion of the violins is eliminated as they harmonize one another in thirds in homorhythmic fashion through bar 23. Here again we witness the *spatial reconciliation* of an oscillation gesture as points 0 and 1 cease their call-and-response activity and merge into ALL PLAY with points 2 and 3. This process of "merging," which initially appeared in movement 2, is cutting across movement boundaries to assist in the spatial unification of the quartet as a whole.

After 30 measures of relative ALL PLAY, further oscillation activity ensues between the viola and cello in mm. 47–48, where SG-|232(3+2)232(3+2)| unfolds. This gesture, occupying the right half of the ensemble, succinctly balances the oscillation activity of mm. 1–16, which inhabited the left half of the ensemble. The latter oscillation gesture is a temporally diminished

365

yet accelerated (featuring an increase in the rate of oscillation) version of the former, in addition to being right-nudged (+2).





Figure 6-11: Mvt. 4, mm. 65–76

At the meter change from cut-time to common-time and the tempo change to Più mosso in bar 65 (the beginning of Variation 2), points 1 and 2 jointly assume the task of articulating eighth-note chords on each beat. A prominent oscillation gesture occurs between the outer ensemble points in mm. 66–74: SG-|030| appears as the first violin and cello engage in a calland-response type of spatio-musical dialogue. This strong distant-point oscillation gesture is an *expanded* form of the adjacent-point oscillation gestures witnessed thus far in the fourth movement (the adjacent-point oscillation gestures are expanded into a distant-point oscillation gesture). SG-|1321|, a strong closed gesture, unfolds in mm. 74–76, after which the gestural activity dissolves and the ALL PLAY situation resumes (mm. 78–88). Although belonging to separate gestural categories, SG-|030| and SG-|1321| (shown in Figure 6-11) are related by virtue of the fact that they both end at the same point from which they begin—obtaining strong closure. The ALL PLAY is violently interrupted by a migratory L-R PAN (SG-|0123|) encompassing a single measure (m. 89), as shown in Figure 6-12 (this gesture was originally presented in Figure 3-9).



Figure 6-12: Mvt. 4, m. 89

Up to this point in the quartet, the left-to-right panning gestures have progressively "shrunk" in each successive movement of the quartet in the following two respects: the timespan of the panning gestures has decreased, and the spatial size of the gesture has decreased by virtue of the fact that they have transformed from being accumulative to migratory (however, each gesture *does* span the entirety of ensemble space). Compare the opening L-R PAN of the first movement (mm. 1–13, Figure 6-1) with mm. 2–3 of movement 3 (Figure 6-9) and m. 89 of movement 4 (Figure 6-11). Not only does the inter-movement proliferation of the L-R PAN serve to unify the quartet as a whole, but movements 1–4 also entail the progressive sublimation (via temporal diminution and textural reduction) of the L-R PAN—a process that encompasses and further unites the first half of the quartet. The systematic refinement and enhanced salience of the L-R PAN occurs in spite of the challenges to its supremacy put forth by oscillation gestures and R-L PANs.

Note also that the nature of the sound events at each ensemble point in a L-R PAN varies widely from one movement to the next. This observation underscores my fundamental argument that the motion through an ensemble, in the form of an abstract spatial gesture, can serve as a recurring motive with tremendous unifying capacity—irrespective of the specific musical qualities (pitch motives, rhythmic motives, etc.) entwined with it during different statements. Furthermore, spatial gestures can unify a work of music on multiple levels of structure and over broad spans of music: the systematic permutation of a gesture can encompass large extents of music and reflect a large-scale process of gestural transformation, which can serve a rhetorical function and is open to aesthetic interpretation. In this particular case, the gradually increasing salience of the L-R PAN (up through m. 97 of the fourth movement) reflects its emergent status

368

as principal gestural/directional form (over the intervening oscillation gestures and right-to-left panning gestures) in the quartet as a whole.

The interrupted ALL PLAY activity recommences in m. 90, this time with SG-[2121], forged by the triplet figures passed between the second violin and viola, enclosed by contrary arpeggiated motion at the outer ensemble points. The polyrhythms that arise between the inner (points 1 and 2) and the outer (point 0 and 3) ensemble points enhance the salience of this weak adjacent-point oscillation gesture, the prevailing motion of which is right-to-left (beginning at point 2 and terminating at point 1). We can read this oscillation gesture as a compressed and "weakened" form of SG-[131] (mm. 74–76), which has already been interpreted as a compressed form of SG-[030]. Overall, the oscillation gesture (as an abstraction) in this section of movement 4 has transformed from comprising events at the outer ensemble points to comprising events at the interior points, and this large-scale transformational process, consisting of three distinct stages, has occurred over twenty-seven measures of music. This localized "shrinking" of the oscillation gesture complements the more hemispheric (occupying the first half of the quartet) shrinking of the panning gesture, as described in the preceding paragraph. After a homorhythmic ALL PLAY (mm. 94–96), this section of the fourth movement ends with a varied repetition of the L-R PAN (from bar 89) in m. 97. The syntactic function of this one-bar panning gesture has changed: what was initially interruptive (locally) has become (comparatively) conclusive.

The tempo and meter change at bar 98 (the beginning of Variation 3) herald complex gestural activity. Bars 98–103 bear SG-|32323232|, a weak adjacent-point oscillation gesture arising from imitative melodic activity between the cello and viola (see Figure 6-13). This activity is taken-up by the violins at bar 106, producing SG-|01010101| (mm. 106–113). This gesture is the type-1 inversion (reflected across the central axis of the total ensemble space) of

369

the initial SG-[32323232]. SG-[23232323] (viewed either as the retrograde *or* type-2 inversion of SG-[323232323]) occurs contemporaneous to SG-[01010101] (mm. 106–113). The sustained whole-note/half-note composition of this gesture renders it markedly less salient than the dotted-rhythmic/arced-melodic activity occurring at each event point within SG-[01010101]. Nevertheless, we may view the spatial activity of bars 106–109 as a case of *echo streaming*, in which two categorically identical gestures (weak adjacent-point oscillation gestures) unfold in separate segments of ensemble space in a staggered manner. Each event of SG-[2323] (the *lead stream*, indicated with a solid line) slightly precedes that of SG-[0101] (the *echo stream*, indicated with a dashed line): in m. 106, the whole note in the viola part (point 2) begins on the downbeat, whereas the melodic activity in the first violin begins on the second beat of the measure. The same relationship exists between points 3 and 1 in m. 107, and the combined spatial interplay of mm. 106–107 recurs in mm. 108–109.

In bars 110–112, the nature of the echo streaming changes: activity at point 0 (pickup to m. 110) precedes the half-note activity at point 2 (m. 110); likewise, activity at point 1 precedes that at point 3 in the pickup to m. 111. As a result, SG-|0101| now becomes the lead stream (solid line) and SG-|2323| assumes the role of the echo stream (dashed line). This reversal of streaming function is caused by the metric displacement of event activity in the violins (mm. 109–110), from the second beat to the upbeat of a bar. The "echo" nature of echo streaming has to due chiefly with the onset of event activity in different parts and not with the melodic and/or rhythmic character of the event. As such, echo streaming may occur in the absence of antiphony/imitative polyphony. In mm. 106–112, melodic imitation occurs *within* each discrete adjacent-point oscillation gesture unfolding on either half of the ensemble space, but not *between* the two gestures (one as the lead and the other as the echo stream).



Figure 6-13: Mvt. 4, mm. 98–111

After a brief ALL PLAY (mm. 112–113), the cello emerges with a new take on the arc melody exposited in m. 98. The viola quickly echoes this melody, ushering-in SG-|32323232|, which fills mm. 114–122 (see Figure 6-14). The call-and-response activity of this gesture is slightly less pronounced due to the overlapping of the two-bar theme between points 3 and 2.

For instance, the viola's entrance in m. 115 occurs at the beginning of the second measure (and second half) of the cello's statement of the theme. This oscillation gesture is therefore *accumulative,* due to the continuous and coincident activity at each event point. Yet, it is clearly evident that a musical entity with identical rhythmic and contour profiles (the nature of which changes at bar 118) is being "passed" to and fro between points 3 and 2.



Figure 6-14: Mvt. 4, mm. 114–125

In bar 122, points 0 and 1 reiterate the foregoing activity of points 3 and 2, once again yielding SG-|01010101| (mm. 122–129). However, points 3 and 2 continue their musical activity, harmonizing points 1 and 0 (respectively) in thirds. This seven-bar span features another instance of polystreaming; however, this time it is a case of *parallel streaming* as opposed to *echo streaming*. Parallel streaming occurs when two (or more) ensemble-space segments display the exact same gestures at the exact same time, such that the onset of events are synchronized at coordinated points within the different segments. In this case, the parallel streaming technically begins in m. 123, where the viola (point 2) falls instep with the first violin (point 0) and the second violin enters with the cello. SG-|(0+2)(1+3)|(0+2)(1+3)|(0+2)(1+3)| (0+2)(1+3)| spans the remainder of this section of the fourth movement. The parallel streaming occurs within the context of ALL PLAY. Although the identification of parallel streaming is perfectly valid, we might consider the antiphonal activity between points 0 and 1 (SG-|01010101|) as primary, since these two points are effectively "echoing" the preceding activity between points 3 and 2 and the harmonizing activity of the viola and cello remains subsidiary.

In the section of movement 4 stretching from m. 98 to m. 129 (Variation 3), we have identified four primary oscillation gestures, all of which are categorically related, if not identical: SG-|32323232| (mm. 98–103), SG-|01010101| (mm. 106–113), SG-|32323232| (mm. 114–122) and SG-|01010101| (mm. 122–129). The entire third variation consists of an oscillation gesture at a higher level of structure: SG-|(3+2)(1+0)(3+2)(1+0)|. Although oscillation activity occurs locally between points 3 and 2 (mm. 98–103, mm. 114–122), they function together as a fused superpoint on a more global level—as do points 1 and 0 (mm. 106–113, mm. 122–129). The SG-|(3+2)(1+0)(3+2)(1+0)| label seems to ignore the echo streaming of mm. 106–113 as well as the parallel streaming of mm. 122–129. However, in mm. 106–113, the nature of the material

373

oscillating between points 0 and 1 is more closely related (via contour) to the musical material vacillating between points 3 and 2 in mm. 98–103 and mm. 114–122. Therefore, the echo streaming is more of a surface phenomenon—albeit one that interacts with the spatial activity abiding at a deeper level of structure.²⁰ In viewing the activity of points 0 and 1 as primary in mm. 122–129, the parallel stream (oscillating between points 3 and 2) is likewise supportive and relegated to the status of a surface phenomenon. However, we might view the parallel streaming as a metric normalization of the earlier echo streaming. Furthermore, since the parallel streaming occurs within an ALL PLAY context, we might view it as a spatial reconciliation of the large-scale oscillation between superpoints 3+2 and 1+0.

Just like its four component gestures, SG - |(3+2)(1+0)(3+2)(1+0)| is a weak adjacentpoint oscillation gesture. This observation reveals a striking accord between parts and whole. The prevailing directionality of the large-scale gesture is right-to-left, and it leads into the ALL PLAY (consisting of parallel streaming) that closes the third variation.

In addition to its four formative oscillation gestures, SG-|(3+2)(1+0)(3+2)(1+0)| embeds two other types of gestures. In considering only the instrumental entrances of mm. 98–107, we encounter SG-|3201|, a weak closed gesture. SG-|3201| occurs again with the staggered ensemble-point entrances of mm. 114-123. Both gestures are accumulative, but the latter gesture is perhaps more accumulative than the former due to the thickened texture of its resultant ALL PLAY section. We can also discern a L-R PAN (SG-|0123|) in mm. 108–111 (Figure 6-13), embedded in the more predominant echo streaming and overlapping the point where the ensemble location of the lead and echo streams are reversed (pickup to m. 110). This gesture comprises heterogeneous musical activity between its first leg (PS 0-1) and its second leg (PS 2-

²⁰ This interaction between different levels of spatial structure is very similar to the Schenkerian, melodicprolongation technique of *reaching-over*, in which surface-level appoggiatura activity interacts with an ascending melodic line (such as the Anstieg) occurring at the middleground.

3) but emerges nonetheless from the successive activity of the discrete ensemble points. These four bars mark a local subordination of panning gestures to oscillation gestures.



Figure 6-15: Mvt. 4, m. 131

Variation 4, an Adagio in compound duple time, begins in m. 130. In the midst of an ALL PLAY scenario spanning mm. 130–141, five important point shifts arise from staggered pizzicato articulations. PS 1-3 occurs three different times, in m. 131 (see Figure 6-15), m. 133, and m. 135. In the final occurrence of PS 1-3, the pizzicato activity at point 3 is extended to the middle of bar 137. Due to the disjunctive span (skipping over point 2) of the point shift, its apparent motion is quite pronounced; it spans three quarters of the total ensemble space. Although these point shifts are sufficiently separated by time and ALL PLAY activity, we may lump them together as weak oscillation gesture SG-[131313]. The left-to-right directionality of PS 1-3 is counterbalanced by the right-to-left motion of PS 2-0, which transpires in m. 139 and again in m. 141 (where it appears slightly modified as SG-[2(1+0)]), collectively yielding SG-

|202(1+0)|, the quasi-type-1 inversion (and event-cardinality reduction) of SG-|131313|.²¹ Considering only the pizzicato events in measures 135–139, SG-|1320| appears. This weak closed gesture functions as the hinge between the related oscillation gestures within the larger gesture SG-|131313202(1+0)|. SG-|1320| is related to SG-|3201| (twice embedded in the preceding section) via *rotational slot displacement*, wherein all of the event-slot contents of SG-|3201| are shifted one slot to the right, and the last event point (event point 1 in event slot 4) "wraps-around" to fill event slot 1. Thus, SG-|1320| results from the rotational slot displacement of SG-|3201|. Both of these gestures are categorically related as weak closed gestures.

Measures 142–145 contain four different instantiations of right-to-left motion. By considering only the rapid sixteenth note motion (and ignoring the dotted-eighth notes), weak open gesture SG-|3120| makes two consecutive appearances in mm. 142–143 (see Figure 6-16).²² This gesture is related to the preceding SG-|1320| by an *intercategorical point exchange* between the first two event slots of the gesture, whereby PS 1-3 (in point-shift slot 1 of SG-|1320|) becomes PS 3-1 (yielding SG-|3120|).²³ The importance of this gesture, however, is that it features predominantly right-to-left motion; abstractly, it results from a *point exchange* between the interior events of a perfect R-L PAN (SG-|3210|). This gesture begins at the rightmost

²¹ The presence of superpoint 1+0 (more specifically, the fusion of point 1 to point 0) in SG-|202(1+0)| complicates the inversional relationship between it and SG-|131313|.

²² Should we count the dotted-eighth notes occurring on the second beat at point 0 in both m. 142 and m. 143, the gestural yield would be SG-[32010], which is still classified as a weak open gesture and consists of predominantly right-to-left motion. The gesture now contains imperfect panning subgesture SG-[320] and oscillatory subgesture SG-[010]. However, due to the Gestalt Law of Similarity (see Chapter 3), we are justified in grouping only the sixteenth-note configurations ensuing from separate points.

²³ SG-|3120| is likewise related to SG-|3201| (mm. 98–123) via *partial rotational slot displacement*, in which only a portion of the SG-label undergoes rotational slot displacement while one or more event points maintain their event slot locations. Beginning with SG-|3201|, we allow event slot 1 to retain event point 3 while slots 2–4 undergo rotational slot displacement, such that point shift 2-0 (initially filling point-shift slot 2) slides rightward to fill point-shift slot 3, and event point 1 (initially in event slot 4) wraps-around to occupy event slot 2, resulting in SG-|3120|. In this case, the partial rotational slot displacement results in a type-B intercategorical alteration, wherein a weak *closed* gesture becomes a weak *open* gesture (both gestures are weak).

ensemble point and concludes at the leftmost point, but it does not span the ensemble space in a unidirectional manner: PS 1-2 indicates the single instance of left-to-right motion.



Figure 6-16: Mvt. 4, mm. 142-145

PS (3+1)(2+0) occurs in m. 144. We can view this point shift between two fused superpoints as both a compression of SG-|3120| and as a spatial reconciliation of the pizzicato activity in mm. 131–141 that yielded SG-|131313202(1+0)| (described above). Finally, SG-|32(1+0)| (R-L PAN) provides a final localized instance of right-to-left motion (m. 144–145). This section of recurrent right-to-left motion began weak and ended with a strong open (panning) gesture, which, although migratory in nature, still introduces an ALL PLAY passage (mm. 146–151). Within this ALL PLAY passage, SG-|323| (a strong adjacent-point oscillation gesture) appears as the distinct points successively state an ascending scalar line.

As shown in Figure 6-17, a single four-bar L-R PAN appears in mm. 154–157, providing nice directional balance to the four-bar right-to-left activity of mm. 142–145. During these measures, the entire ensemble is still engaged in ALL PLAY, but the ascending conjunct scalar activity in perpetual sixteenth notes (which had initially forged an oscillation gesture in mm. 146–151) weaves its way through the ensemble. Taking a closer look at m. 153 (Figure 6-17),

we discover that it functions as a transitionary "pivot" between the right-to-left motion of mm. 142–145 and the left-to-right motion of mm. 154–157.



Figure 6-17: Mvt. 4, mm. 153–157

Taken as a whole, and considering only ensemble-point entrances, bar 153 contains SG-|21(0+2)3|, a weak closed gesture. However, embedded in this gesture are two smaller gestures: SG-|210| (in the first three quarters of the measure, indicated by solid arrows) and SG-|123| (the last three quarters of the measure, traced with dashed arrows). The relationship between these two gestures may be described either by type-1 inversion or by the twin operations of retrograde
and nudging: the retrograde of SG-|210| is SG-|012|, which becomes SG-|123| when nudged one point to the right. However, the two gestures overlap one another in time as well as in space: they both share the exact same event at point 1 (in event slot 2 of SG-|210| and event slot 1 of SG-|123|). Essentially, point 1 "splits" to continue as SG-|210| as well as to initiate SG-|123|. The concomitant occurrence of these two related gestures is yet another case of polystreaming. This time, however, the spatial activity yields *divergent mirrored streaming*: the two gestures move in opposite directions from a single (and nearly centralized) point (point 1). The right-toleft motion begins and ends prior to the termination of left-to-right motion, which supports the claim that this gesture is pivotal—marking the end of rightward motion and the beginning of leftward motion. Perhaps the SG-|323| of mm. 146–149 serves a similar function, since its initial point shift moves from right-to-left (PS 3-2) and it concludes with a left-to-right point shift (PS 2-3). In this respect then, both SG-|323| and SG-|21(0+2)3| function together to *prolong* the pivot from rightward to leftward motion: SG-|21(0+2)3| is a spatially extended yet temporally diminished annex of SG-|323|.

On a deeper level of structure, mm. 142–157 comprise an expansive RIGHT SWEEP, which is a compound, strong closed gesture. Figure 6-18 provides a diagram of this large-scale gestural activity. The underlying SWEEP is embellished by prolongational spatial activity on the musical surface. The prevailing right-to-left motion of mm. 142–145 is subdivided into four small right-to-left episodes: two weak open gestures, one point shift, and one strong open gesture. We might forge a middleground R-L PAN with point 3 in m. 142 being prolonged to point 3 (at the final three sixteenths) in m. 144, at which point the final foreground SG-[32(1+0)] likewise provides the final two events, 2 and (1+0), of the middleground R-L PAN. As

379

for example, SG-|3210123| pivots on point 0, which functions both as the last event of a R-L PAN (antecedent subgesture) and as the first event of a L-R PAN (consequent subgesture). At the deeper-level sweeping activity of mm. 142–157, however, the pivot consists of a composed-out section spanning eight measures (mm. 146–153). The "background" identified in Figure 6-18 reveals the source- and goal-points of the SWEEP (these points are in boldface at all structural levels identified in the figure): point 3 is the source and goal of the entire sweeping gesture (at the global level of this section's spatial structure). Point 0 functions locally as both the goal of the initial right-to-left motion and the source of the final left-to-right motion.²⁴ The prolongation of this intermediary point (the "hinge proper" of the RIGHT SWEEP) occupies the longest span of time in the entire sweeping gesture: point 0 is prolonged from m. 145 to m. 153 (within the surface-level pivot) and again from m. 153 to m. 154, where the final L-R PAN begins. Arrows drawn between points at the middleground level show spatial prolongations between the underlined numerals.²⁵

(Motion):	$R \rightarrow L$		"PIVOT"	$L \rightarrow R$	
Surface:	3 210 3210 (3+1)(2+0) 32(1+ 0)	//	323 21(0+2)3	//	012 3
Middleground	$: \underline{3} \longrightarrow \underline{32(1+\underline{0})}$	//	<u>→</u> <u>0</u> →		<u>0</u> 12 3
Background:	3 0	//		11	3
(Section):	(mm. 142–145)		(146–153)	(154–157)

Figure 6-18: Diagram of multilevel sweeping activity

This Adagio episode is brought to a close by pizzicato point shift 2-(3+2) in m. 161. This is the compressed form of the pizzicato point shifts featured in mm. 134–141, and it features a

²⁴ The deep background-level reduction of this gesture is actually strong oscillation gesture SG-|303|.

²⁵ I am deliberately evoking the Schenkerian concepts of prolongation and the hierarchical organization of structural levels. However, I only loosely apply the methodological framework of Schenkerian analysis in an effort to show that broad spatial gestures can and do occur at levels of structure removed from the immediate surface of a work.

return to the initial left-to-right motion of PS 1-3. The two occasions of right-to-left motion afforded by PS 2-0 (mm. 139 and 141) may now be seen as the directional "departure" in a largescale departure-and-return scenario—the "return" not being effected until m. 161. Furthermore, the return is weakened threefold by the conjunction of PS 2-(3+2) (now only spanning one-half of the ensemble space), the fusion of points 3 and 2 in the final event of the point shift, and the solitary local occurrence of the point shift. Interestingly, the final event of the fourth movement (m. 277) is an *unspatialized* pizzicato event shared by points 1–3 (with concurrent non-pizzicato activity in the first violin). The spatial location of the pizzicato articulation occupies the original span from points 1 to 3, but it is now filled-in by coincident activity at point 2 and consists of a single event at superpoint 1+2+3. Further observations regarding this closing event shall be given below.

Once again, a joint change in tempo and meter ushers in a new variation (Variation 5), and with it, a prominent spatial gesture: SG-|0132| appears in mm. 162–165, as shown in Figure 6-19 (recall that this gesture was embedded in the imperfect LEFT SWEEP of movement 3).



Figure 6-19: Mvt. 4, mm. 162-165

This weak closed gesture is the type-1 inversion of SG-|3201|, which twice appeared embedded within the larger oscillation framework of mm. 98–129 (as described above). The family of weak closed gesture appears to be vying for gestural prominence. However, SG-|0123| occurs in mm. 180–182 (see Figure 6-20). This gesture might be surmised as the "opening-up" of the previous SG-|0132|, resulting from a *point exchange* between event slots 3 and 4: PS 3-2 becomes PS 2-3, and a weak closed gesture becomes a strong open gesture (L-R PAN). The conversion of SG-|0132| to SG-|0123| is underscored by the reiteration of SG-|0123| brought about by the repeat of mm. 178–185.



Figure 6-20: Mvt. 4, mm. 180–182

ALL PLAY carries us through this variation and into the sixth variation, *Adagio ma non troppo e semplice*, beginning at m. 187. Although the musical texture is predominantly homophonic, and the ALL PLAY spatial texture is for the most part sustained, some interesting ensemble divisions begin to emerge that affect the remainder of the fourth movement. In mm. 195–198, a three-against-one division occurs as ensemble points 0–2 continue their quarter-note pulsation while the cello iterates a sixteenth-note motivic figure (see Figure 6-21).²⁶ The ensemble space is heavily weighted toward the left in this imbalanced partitioning.



Figure 6-21: Mvt. 4, m. 195–196

Throughout mm. 203–213, R-L PAN gives way to imbalanced ensemble-space divisions on four occasions. SG-|3(2+1)0| occurs in bars 203–204 (see Figure 6-22), 205–206, 211, and 213. Each time (with the exception of mm. 205–206, which introduces ALL PLAY), the R-L PAN hinges with the three-against-one segregation of ensemble space exposited in mm. 195–198 (shown partially in Figure 6-21). This subsection of Variation 6 closes with the three-againstone division (spanning mm. 213–219); therefore, on a broader level, we can observe the departure-and-return character of mm. 195–219, in which three-against-one is first presented (mm. 195–198), subsequently "interrupted" by SG-|3(2+1)0| (R-L PAN) on four separate

²⁶ It is possible to view the spatial activity of mm. 195–198 as oscillation gesture SG-|(0+1+2)3(0+1+2)3(0+1+2)3(0+1+2)3|(0+1+2)3|. Under such a reading, the fusion of points 0, 1, and 2 produces the largest superpoint encountered thus far in the quartet. This spatial gesture is weakened by the lack of balance between point 3 and superpoint 0+1+2; oscillation gestures exhibiting spatial balance are stronger (i.e., "better-formed," in the Gestalt sense) than those lacking balance between component points. However, owing to the continuous activity at ensemble-space segment 0–2 (which provides a sonic buttress to the periodic interjections of the cello), viewing the ensemble-space as divided (in an imbalanced fashion) is perhaps the optimal reading.

occasions (mm. 203–213), and then decisively returned-to (mm. 213–219). In this spatial process, the panning gestures serve as the kinetic points of departure to a static spatial imbalance.



Figure 6-22: Mvt. 4, mm. 203-204

An ALL PLAY D-sharp diminished-seventh chord on the downbeat of m. 220 paves the way for an entirely new spatial gesture—our first encounter with a strong closed gesture on the spatial surface. SG-|02310| occurs in mm. 220–225, as a melody winds its way through the ensemble. This gesture is punctuated by activity in other instruments: 1+2+3 in m. 221 and m. 223, 0+1+2 in mm. 223–224, and ALL PLAY on the first two beats of m. 225 and m. 226. Incomplete panning gestures continuously approach the trilling first violin (point 0) in mm. 227–230, where four statements of SG-|321| yield a *cyclic panning gesture* in ensemble-space segment 1–3 (see Figure 6-23). SG-|321321321321321] is technically classified as a compound, weak open gesture (within ensemble-segment space). However, as the first violin ceases its trilling activity at the end of m. 230, it essentially joins the panning activity of its ensemble cohorts—resulting in a perfect R-L PAN (SG-|3210]). Thus, the panning activity left incomplete

(within the total ensemble space) by the three SG-|321|s of mm. 227–229 is finally completed in m. 230.



Figure 6-23: Mvt. 4, mm. 227-234

In an alternate reading, we might view the pitch-changes in the first violin (bars 228 and 229) as twice completing the right-to-left panning motion through the ensemble space, in which case, the panning gestures would overlap, yielding SG-|321(0+3)21(0+3)213210|. Due to the overlap of a gesture's final event point with another gesture's initial event point (indicated by 0+3), this reading would count as an instance of *canonic streaming*, in which identical gestures overlap one another within the total ensemble space (or within the same ensemble-space

segment). However, the trills at point 0 differ drastically from the coordinated arpeggiation of points 1-3, so a division of the ensemble space such that point 0 and ensemble-space segment 1-3 are segregated (on the basis of their respective musical activity) seems most apt. This disproportionate ensemble division remains intact during mm. 231–242 (a portion of which is shown in Figure 6-23), as the musical activity of the first violin remains distinct from the homorhythmic activity (mm. 231–238) and silence (mm. 239–242) at points 1-3.²⁷ As such, we now find a one-against-three ensemble division, which is the type-1 inversion of the threeagainst-one ensemble-space division identified earlier. In either partitioning, a spatial extreme of the ensemble space is the "odd man out." The one-against-three imbalance of mm. 227-242 (where point 0 is segregated from points 1-3) balances the preceding three-against-one imbalance of mm. 195–219 (where point 3 is segregated from points 0–2). Thus, we find local spatial imbalance (as surface phenomena) within the ensemble space on two occasions within the section stretching from m. 195 to m. 242. However, due to the inversional relationship between these two imbalances, a sense of balance is produced on a deeper level of the spatial structure as the three-against-one scenario is transformed (via type-1 inversion) into a one-against-three division.

Rapid oscillation activity occurs in mm. 243–249, as SG-|303...0| emerges from the sixty-fourth-note exchanges between the first violin and cello (m. 245 is given in Figure 6-24).²⁸ The salience of this gesture is questionable given the continuous trilling activity in the first violin part. However, a rapid trill in pitch space (statically located at a single point in ensemble space)

 $^{^{27}}$ In mm. 231–242, the first violin plays an unspatialized version of the theme from the beginning of the movement (where it was originally split between the two violins, resulting in an oscillation gesture between points 0 and 1). This recasting of the theme at a single ensemble point reflects the initial spatial reconciliation of the theme occurring at m. 17 (and described above).

²⁸ The spatial activity of mm. 243–249 features a division of ensemble space such that the outer points (3 and 0) are set against the inner points (1 and 2). The second violin and viola (inner points) present the primary thematic material in octaves while the cello and first violin (outer points) decorate this line with rapid figures in pitch space and oscillation activity in ensemble space. The primary melody is thus spatially contained by decorative filigree.

complements the concurrent oscillation gesture in ensemble space. The trill is essentially an "oscillation" between two "close" pitches (either a half-step or whole-step apart in pitch space). In this passage, the trill is the pitch-space counterpart to the oscillation gesture, and their simultaneous presentation is yet another instance of inter-parametric unification.²⁹



Figure 6-24: Mvt. 4, m. 245

The oscillation and trilling activity ultimately concludes in a cyclic sweeping gesture within ensemble-space segment 1–3 in mm. 250–253 (see Figure 6-25). After SG-[303...0] concludes in m. 249, the violin resumes its sustained trilling enterprise first exposited in mm. 227–230. In fact, mm. 250–253 is a varied repetition of mm. 227–230: the violin trills are transposed up an octave, and the arpeggiated activity of points 1–3 is reworked rhythmically as a triplet configuration and spatially as sweeping instead of panning (still within ensemble segment space 1–3). SG-[2321232123212321] occupies mm. 250–253. This compound gesture contains three consecutive RIGHT SWEEPs (SG-[32123]); it likewise contains three LEFT SWEEPs (SG-[12321]). The prevailing motion of this gesture is difficult to gauge given that it begins at point 2

²⁹ The correspondence between the oscillation gesture and trilling activity would be more absolute if the oscillation gesture were adjacent-point instead of distant-point.

and ends at point 1. We might view the event activity at point 2 as being prolonged by neighboring spatial motion to and from point 3 (as "rightward neighbor") and point 1 (as "leftward neighbor"), alternately, until the spatial articulation provided by ALL PLAY occurs at m. 254.



Figure 6-25: Mvt. 4, mm. 250–254

The spatial neighboring activity meshes with the pitch neighbors produced by the trills in the first violin as well as by the neighboring motion (in pitch space) from B-natural to both C-sharp and C-natural in the violin's melodic line (it is these three principal tones that are ornamented by the trills in mm. 250–253). We may also view the violin activity at the end of m. 253 as

breaking the ensemble-segment sweep and yielding a R-L PAN (SG-|32(1+0)|). This reading parallels nicely with our prior analysis of m. 230, of which m. 253 is a recomposed musical analog. Bars 254–263 function as a transposed and temporally extended version of mm. 231– 239, and they feature the same imbalanced oscillation activity as the former section: SG-|(3+2+1)0(3+2+1)...0|, which, again, complements the one-against-three ensemble segregation of mm. 250–253. Given the repeat of this imbalance, the scales seem to be tipping in favor of the one-against-three scenario, potentially compromising the large-scale balance heretofore provided by the inversionally related spatial imbalances.

A somewhat concealed L-R PAN crops up as the final spatial gesture of the movement. In m. 264, the first violin begins a rapid scalar passage against the sustained V7/A broken-up among points 1–3. In m. 265, the chord is reproduced as staggered double stops that are dispersed through the lower parts in an expansive motion projecting rightward through the ensemble space (see Figure 6-26). In keeping with one of its established syntactic functions, an accumulative SG-[0123] (L-R PAN) introduces the ALL PLAY conclusion to this movement (mm. 266–277).



Figure 6-26: Mvt. 4, mm. 264–265

A large-scale RIGHT SWEEP ranges across mm. 227–265. The first part of the sweeping gesture is the repeated right-to-left panning activity in mm. 227–230 (Figure 6-21), which finally spans the total ensemble space in mm. 230–231. The "hinge" of the RIGHT SWEEP occurs with the cyclic sweeping activity on the spatial surface of mm. 250–253. The left-to-right panning gestures contained within this pivotal gesture forecast the L-R PAN of mm. 264–265, which serves as the final component of the expansive RIGHT SWEEP. This movement has exhibited two RIGHT SWEEPs—each abiding at a deeper level of the spatial structure and comprising smaller surface-level gestures.

The final measure of the movement serves as a reminder of the one-against-three spatial imbalance as superpoint 1+2+3 performs a pizzicato A-major chord against point 0's arco rendering of upper-voice pitch A. This imbalance is not dynamically spatialized—as it was in some preceding passages—since all four instruments execute the chord in rhythmic unison. Instead, the static spatial imbalance is evinced through the differences in articulation (attack types) between simultaneous events at point 0 and superpoint 1+2+3.

In short, within movement 4, oscillation gestures assume a more prominent role roughly equal to panning gestures in terms of prevalence and perceptual salience. This movement has likewise displayed several closed gestures, and the incidence of right-to-left panning gestures has risen to rival that of L-R PANs, the salience of which has waned since m. 97. As the quartet progresses, we shall discover if oscillation activity will ultimately surpass that of panning activity or if the panning gestures will reassert their preeminence. If the strong open gestures (panning gestures in general) *do* come out on top, will the left-to-right directionality continue to prevail or will the R-L PAN supersede the L-R PAN?

6.5 Movement 5

The cello's "false start" in m. 1 of the fifth movement is significant. This outburst, shown in Figure 6-27, actually completes a large-scale oscillation gesture begun in the preceding movement but left unfinished at its end.



Figure 6-27: Mvt. 5, mm. 1-2

The oscillation gesture emerges from the isolation of points in the imbalanced ensemble-space divisions of movement 4 and at the head of movement 5. In the three-against-one imbalance of mm. 195–219, the cello (point 3) is segregated from ensemble-space segment 0–2. In mm. 227–242, the one-against-three imbalance (which functions to balance the former imbalance) entails the setting apart of point 0 from points 1–3. This imbalance is sustained through the end of the fourth movement—it is present in mm. 250–264 and in the final bar (m. 277). The entrance of the cello in m. 1 of movement 5 with the movement's head motive—set against the backdrop of silence provided by the inactivity at points 0–2—completes the large-scale strong oscillation gesture SG-[303], which, again, arises from the marginalization of the outer ensemble points over

a span of approximately 83 measures (across two movements).³⁰ While the beginning of movement 5 is introductory in many respects (motivic, harmonic, formal, etc.), it is conclusive in a broad spatial sense.

The first 23 measures of movement 5 continue the ALL PLAY situation with which movement 4 concluded. Bars 25–36 present a migratory, cyclic panning gesture: SG-|012301230123|, which contains three discrete instances of SG-|0123| (see Figure 6-28). The perceptual salience of each component L-R PAN is enhanced by the fact that the same two pitchclasses (constituting the same ascending melodic interval) are passed from one ensemble point to the next: F-sharp to B/perfect fourth (mm. 25–28), G to B/major third (mm. 29–32), and G-sharp to B/minor third (mm. 33–36). Enwrapped in this cyclic panning gesture is the systematic compression—from one subgestural unit (discrete L-R PAN) to the next—of the interval being panned from point to point, which is effected by raising the lower pitch of the interval by one semitone in each successive SG-|0123|: F-sharp \rightarrow G \rightarrow G-sharp. This three-stage intervallic compression is related to the three-stage process of spatial compression undergone by the oscillation gestures in mm. 66–93 of movement 4 (discussed above). This observance highlights a dynamic processual concordance between events in pitch space and events in actual (ensemble) space, and it further reflects a unity that transcends parametric and temporal boundaries.

Following the third and final L-R PAN of the larger compound gesture, a R-L PAN unfolds in mm. 36–39 (see Figure 6-28). Although this SG-|3210| consists of musical material different from that of the preceding cyclic panning gesture (namely, the canonic treatment of the movement's head motive), bars 33–39, taken together, constitute a perfect LEFT SWEEP (SG-|0123210|). The final R-L PAN of the sweeping gesture balances the thrice-occurring left-to-

³⁰ The surface-level oscillation gesture SG-|303...0| (mm. 243–249) is contained (temporally) within this larger oscillation gesture. Although both are distant-point oscillation gestures involving the outermost points (spatial extremes) of the ensemble space, the surface gesture is classified as weak and the larger gesture as strong.

right motion of the cyclic panning gesture.³¹ Therefore, the L-R PAN of mm. 33–36 is subsumed by two different (and overlapping) compound gestures: the cyclic panning gesture of mm. 25–36 and sweeping gesture of mm. 33–39.



Figure 6-28: Mvt. 5, mm. 25-39

An additional LEFT SWEEP, SG-|0123(2+1)0|, occurs in mm. 67–69 (see Figure 6-29).

This sweeping gesture is a temporally diminished version of the preceding LEFT SWEEP (mm.

33–39). In m. 69, the event activity at superpoint 2+1 "splits": it continues leftward to point 0 to

³¹ The immediate counterbalance of the three successive L-R PANs by a lone R-L PAN is another instantiation of the "three-against-one" imbalance. In movement 4, this spatial dialectic (as well as its converse, one-against-three) is played-out by uneven divisions of the ensemble. In the fifth movement, it is manifest by the frequency of occurrence of panning gestures of a given directionality: three L-R PANs against one R-L PAN.

complete the sweeping gesture, and it also moves rightward to become part of an adjacent-point oscillation gesture that overlaps the sweeping gesture: PS 2-3 (in PS slot 3 of the SWEEP's SGlabel) is the opening point shift of SG-[2323...2], which ranges from m. 68 to m. 77.



Figure 6-29: Mvt. 5, mm. 67-71

Michael Steinberg refers to this musical activity, which functions as accompaniment to the thematic material presented in the violins, as "a rapid Ping Pong of quarter notes for viola and cello."³² Steinberg does not explicitly qualify his description of the musical interplay, but the "ping-pong" metaphor that he employs is quite clear. Interestingly, the "ping ponging" occurs both in pitch space (the viola's descending fourth alternating with the lower descending fifth in the cello) and in ensemble space (the oscillation gesture between points 2 and 3). In fact, the amalgamation of the motions in both conceptual space (pitch space) and actual space (ensemble space) produces a *composite gesture*: a diagonal (or, slanted) motion repeatedly traces an "upperleft-to-bottom-right" vector (the top-to-bottom motion is provided by events in pitch space and the left-to-right motion by events in ensemble space).³³

 ³² Steinberg, 258–259.
³³ A similar integration of spaces occurs in the fugal opening of the entire quartet (mvt. 1), in which descending motion in registral space is conflated with left-to-right motion in ensemble space. The result is blended motion (i.e.,

Following the ALL PLAY activity of mm. 78–85, the oscillation activity between the viola and cello resumes with weak adjacent-point oscillation gesture SG-|32323232| in mm. 86–90. This gesture is transferred to the violins in mm. 90–101, where it is "strengthened" and elaborated to produce SG-|0101010101010101010101010]. The spatial displacement of this gesture, as it is shifted from ensemble-space segment 2–3 to segment 0–1, is tantamount to a large-scale "ping ponging" of the "Ping Pong" figure. SG-|32323232| and SG-|010101010101...0| are both adjacent-point oscillation gestures (the former is weak, while the latter is strong) that are related to one another via type-1 inversion. While large-scale point shift PS (3+2)-(1+0) accounts for right-to-left motion between bars 86–90 (3+2) and 90–93 (1+0), the weak closed gesture SG-|3201| is embedded in the passing of antiphonal activity from points 3 and 2 to points 0 and 1 (mm. 89–90). SG-|3201| is outlined in Figure 6-30.



Recall that SG-|3201| appeared twice within the oscillation context of mm. 98–129 in movement4. Despite the joint entrance of the viola and cello in m. 94 (of mvt. 5), the violins continue their

a *composite gesture*) spanning from northwest to southeast within a hybridized conceptual-actual "meta-space." In effect, within this meta-space, the horizontal (left-right) dimension of time in the default conceptual musical space is replaced by a dimension of space-time: in this example, the gesture's left-to-right direction of motion in space accords with the conceived, left-to-right flow of time).

inherited antiphonal interplay through m. 101. However, on the last beat of m. 101, the violins merge together to play the same figure one octave apart—continuing the descending root-fifth-root arpeggiation of an E-major triad that was formerly antiphonally dispersed. This consolidation of pitch and rhythmic activity effects a point fusion in ensemble space. The spatial reconciliation of oscillation activity (resulting from the fusion of two or more points) was seen in mm. 8–23 of the second movement, where m. 17 marked the resolution of the preceding oscillation gesture, as well as in mm. 1–24 of the fourth movement.

The oscillation figure of mm. 85–97 returns in m. 143—this time materializing at the inner points of the ensemble space (points 1 and 2). Here we observe the opposite of point fusion/spatial reconciliation as a superpoint is "split" into its two component pitches: the predominantly homorhythmic enterprise of the second violin and viola (mm. 126-141) is bifurcated in m. 142, yielding SG-|212...1| through m. 149. We may apply the Out \rightarrow In (Periphery \rightarrow Center) process (enabled by the ENSEMBLE SPACE-AS-CONTAINER schema) to account for the spatial process occupying mm. 68–149. The boundaries and peripheral "outskirts" of the "container" are explored and outlined in mm. 68-83, first with the oscillation gesture at points 2 and 3 (mm. 68–76, 86–90) and following with the transference of that gesture (in altered form) to points 0 and 1 (mm. 90–101). We may now observe the interpretive importance of the point-shift directionality within these two local oscillation gestures: the first gesture ends with PS 3-2 (a small-scale Out \rightarrow In point shift) and the second gesture begins with PS 0-1 (also a surface-level manifestation of the Out \rightarrow In process). Thus, the fulcrum of these two gestures, weak closed gesture SG-|3201| (mm. 89-90; see Figure 6-30), emphasizes the local Out \rightarrow In process and portends its impending occurrence at a higher level of structure. The appearance of SG-[212...1] in m. 142 completes the large Out \rightarrow In journey embarked upon by

396

the adjacent-point oscillation gesture. SG-|(3+2)(0+1)(1+2)| accounts for the large-scale spatial activity of mm. 68–149. (The local oscillation activity between parenthetically enclosed points is reduced-away in this SG-label.) In essence, this gesture is a weak closed gesture that begins by navigating the boundaries (spatial extremes) of the ensemble-container and concludes at the interior ensemble points. However, this higher-level spatial process is not yet finalized...

The oscillation activity at points 1 and 2 continues at m. 150, although it undergoes temporal augmentation (x2) and is made to consist exclusively of ascending intervals (see Figure 6-31). SG-|212121212121| encompasses mm. 150–159. Points 1 and 2 undergo point fusion in mm. 159–160, again marking the reconciliation of oscillation activity. When this gesture begins in m. 150, the SG-|212...1| of mm. 142–149 splits and expands into the outer ensemble points: SG-|303...0| occurs concurrent to SG-|2121212121| in mm. 150–159. The former gesture unfurls at twice the rate of the latter; this temporal differentiation enhances the salience of each gesture, resulting in a clear case of uneven, yet proportional (in a 2:1 ratio), consonant polystreaming.



Figure 6-31: Mvt. 5, mm. 147–153

Despite the polystreaming and the continued oscillation activity between points 1 and 2 in mm. 150–159, the oscillation gesture at the outer ensemble points completes the ongoing,

higher-level spatial process. The In \rightarrow Out motion resulting from the relocation of SG-[2121212121] (mm. 142–149) to SG-[303...0] (mm. 150–159) provides the "return" portion of the wide-reaching departure-and-return process occupying mm. 68-160. SG-|(3+2)(0+1)(1+2)|(0+3), a strong closed gesture at a higher level of spatial structure, reflects the large-scale Out \rightarrow In \rightarrow Out path mapped-out by the itinerant oscillation gesture.³⁴ The initial "Out" occupation of the large closed gesture, SG | (3+2)(0+1) |, is temporally staggered and weakened by the mutual presence of an inner ensemble point (which may be viewed as "peripheral" when functionally attached to an outer ensemble point): 3+2 (in event slot 1 of SG-((3+2)(0+1)(1+2)(0+3))) contains outer point 3 and peripheral point 2; 0+1, in event slot 2, contains outer point 0 and peripheral point 1. Event slot 3 contains interior ensemble points 1 and 2—neither of which is antiphonally engaged with (or otherwise fused to) an outer ensemble point. Therefore, these points are rightly considered "inner" rather than "peripheral."³⁵ Up to m. 150, the oscillation gesture weaving its way through the ensemble space has been purely adjacent-point. However, the oscillation gesture of event slot 4 is distant-point, and it strongly outlines the ensemblecontainer's boundaries. In a sense, then, the concluding oscillation gesture encapsulated in this mammoth closed gesture supplies a purified form of the ensemble-container's outer region, and the last two oscillation gestures (event slots 3 and 4 of SG-|(3+2)(0+1)(1+2)(0+3)|) provide a

³⁴ In this situation, the peripatetic oscillation gesture, navigating its way through ensemble space, is less abstract than some of the large-scale spatial relations elucidated above. The oscillation gesture *at all times* consists of the same rhythmic figure (and similar melodic profile) passed back and forth between two ensemble points, irrespective of the point-identity (location in ensemble space) of the partnered points. Accordingly, the spatio-kinetic process of this gesture should be quite noticeable to the listener.

³⁵ As prescribed in Chapter 5, the identity of an ensemble point in relation to the entire ensemble-as-container may be transformed by the spatial context in which it is immersed. Abstractly, in four-point ensemble space, points 1 and 2 are the *inner* (or, "contained") points, and points 0 and 3 are the *outer* points (or, "container boundaries"). In an analytic context, point 1 may be considered an *inner point* if functioning independently *or* if gesturally linked to (or fused with) point 2, the other inner point. Point 1 may be considered a *peripheral point*, however, if linked to point 0 (a natural outer point of the ensemble space). Finally, point 1 may be considered an *outer point* (or, "container boundary") if ensemble-space segment 1–3 is inspected independent of point 0.

palpable differentiation between "inner" (contain<u>ed</u>) and "outer" (contain<u>er</u>). Figure 6-32 summarizes the spatial process.

SG- (3 +2)(0 +1)		(1+2)		(0+3)	
mm. 68–101		mm. 142–149		mm. 150–160	
Out	\rightarrow	In	\rightarrow	Out	
		Departure	\rightarrow	Return	

Figure 6-32: The spatial process of an oscillation gesture

The uneven consonant streaming in mm. 150–160 emphasizes the inner/outer spatial dialectic in synchronic fashion and at a local level of spatio-gestural activity.

A pizzicato sequence yields SG-|10232102| in mm. 162–169 (see Figure 6-33). This gesture is classified as a weak closed gesture. Considered as a whole, this highly salient migratory gesture (unified by pizzicato articulation and partitioned by measured silence) contributes little to the unfolding spatial rhetoric of the quartet at large—at least in terms of spatio-motivic recurrence. (The arco motive stated by the cello in m. 168 is different enough to be counted as "other" and is excluded from the gesture.) However, the gesture as a whole is unified by the SG-|102|s that bookend it. Furthermore, it embeds an imperfect (abbreviated, via deletion) LEFT SWEEP: by ignoring the spatial endpoints of the gesture (event point 1 in event slot 1 and event point 2 in event slot 8), we may identify SG-|023210|. In fact, if the first two event points of the entire gesture were exchanged, the sweeping gesture would be "perfected" as SG-|0123210|. This is not the case, however, and the sweeping activity is to some extent concealed. The first four events of the entire gesture yield SG-|1023|, the retrograde of SG-|3201|, which has already occurred throughout the quartet. SG-|10232102| effectively conflates

open and closed gestures, suggesting uncertainty as to which categorical kind or directionality of gesture will ultimately prevail.



After the repetition of the preceding section (mm. 170–326 is a repeat of mm. 3–160), SG-|10232102| is replaced by SG-|10121320120121| (mm. 328–335; Figure 6-34). The pitch sequence of the original pizzicato gesture (SG-|10232102) is recast in mm. 328–335 with each pitch doubled (i.e., presented twice in direct succession). However, the duplicate pitches are delegated to different instruments, and thereby, to different ensemble points. Each pair of identical pitches produces a conjunct point shift with the exception of the B-naturals in m. 332, which create disjunct point shift 2-0. The pairing of points (by pitch) produces the more elaborate gesture SG-|10121320120121|; however, the first note of each pair presents SG-|10232102|. As a whole, this gesture is strongly closed since it begins and ends at point 1. It likewise embeds smaller gestures of contrasting kind and directionality: SG-|01213| is a weak open gesture (left-to-right), SG-|320| is an imperfect panning gesture (right-to-left), SG-|3201| is a recurrent weak closed gesture, and SG-|012| (occurring twice in a row) is a partial left-to-right panning gesture (SG-|012012| is a cyclic panning gesture in ensemble-space segment 0–2). Thus, the categorical and directional conflict of the pizzicato figure initially presented in mm. 162–169 is compounded in mm. 328–335.



Measures 336–462 feature more repeated musical activity, and the pizzicato gesture appears once again in a new guise in mm. 466–471, this time as SG-|120120310212| (see Figure 6-35).³⁶ Temporally adjacent unison pitches are still coupled, and each coupled pair is still spatialized, but the sequence of pitches is unrelated to that of the two preceding pizzicato passages. In this passage, however, the disjunct point shifts among identical pitches (PS 2-0, PS 3-1, and PS 0-2) outnumber the conjunct point shifts (PS 1-2 and 0-1) of unison pitches by 1. Embedded gestures are less apparent in this weak closed gesture. SG-|0120310| (event slots 3–9) is related to a LEFT SWEEP, but the would-be initial L-R PAN (SG-|01203|) is occluded by event point 0 (event slot 4) and the final R-L PAN is imperfect, featuring a disjunction between points 3 and 1. The reduced salience of internal spatial gestures (subgestural components)

³⁶ The sectional repetition of this scherzo movement owes to its extended ternary design, outlined as follows: A (Scherzo)-B (Trio)-A-B-A-Coda (B-A; Trio fragments and contracted Scherzo). Amy Carr-Richardson discusses the formal structure, tonal design, and motivic organization of this movement in terms of "rotational symmetry." See Amy Carr-Richardson, "Rotational Symmetry as a Metaphor for the Scherzo of Beethoven's Opus 131," *Indiana Theory Review* 20.1 (1999): 1–23. Thus, symmetry permeates this movement on multiple levels, as it is found in the work's durational proportions of sections, sequence of key areas, and spatial activity (the symmetrical sweeping gesture).

highlights the rhetorical function of the pizzicato gesture as an agent of liquidation and dissolution of any concrete spatial gesture—a role it has fulfilled with increasing effectiveness throughout the fifth movement.³⁷ Movement 5 closes with 28 measures of ALL PLAY.



Figure 6-35: Mvt. 5, mm. 466–470

6.6 Movement 6

Movement 6, a 28-measure adagio, features only two salient and identical spatial gestures: SG-|(3+2)10| initially occupies mm. 12–14 (see Figure 6-36) and recurs in mm. 19–21. The double occurrence of this R-L PAN—sans any intervening gestures—provides this concise movement with a degree of spatial coherence. The gestures occur in relatively close proximity in the interior portion of the movement, surrounded by ALL PLAY. However, more pressing is the right-to-left directionality of the gesture. Locally, the rightward motion is left unchecked due to the lack of any left-to-right motion. However, on the global level of the entire quartet, this evanescent movement displays the emergence of R-L PAN as the new principal gesture of the

³⁷ In each of these pizzicato passages, the parts-distribution of pitches defies the typical registration of the ensemble instruments. For example, in mm. 162–169 (Figure 6-33), the second violin's A-natural (m. 162) is a minor third *higher* than the first violin's F-sharp (m. 163). Material in the upper musical voice is typically allocated to the first violin. Similarly, the viola's D-sharp (m. 164) is *lower* in pitch space than the cello's F-sharp (m. 165). The "jumbling" of registral distribution complements the dissolution of spatio-gestural activity.

quartet, providing a long-term balance to the L-R PANs that, along with the oscillation gestures, have virtually dominated the spatio-gestural activity of the quartet up to this point.



Figure 6-36: Mvt. 6, mm. 12-14

Also noteworthy is the ascension through pitch space that accompanies SG-|(3+2)10|. In m. 12, the interval of a minor sixth is provided by superpoint 3+2 as the gesture's opening event features D-sharp at point 2 undergirded by F-double-sharp at point 3. The C-sharp entrance of point 1 is a minor seventh above point 2's D-sharp, and the first violin enters on B a minor seventh above point 1's C-sharp. Points 2, 1, and 0 each feature local stepwise descents, but the gesture as a whole features a rise in registral space and is therefore tensive, which contradicts the "spatial relaxation" often associated with right-to-left motion—particularly that which follows a left-to-right motion (as described in Chapter 5). A similar relationship between pitch events and spatial events occurs with the reprise of the R-L PAN in mm. 19–21, albeit with altered pitches.

Recall the opening gesture of the entire quartet: measures 1–13 of movement 1. This gesture is an accumulative L-R PAN interwoven with a global descent through pitch-space that comprises local ascension at the beginning of each discrete ensemble point's melodic activity. The two panning gestures of movement 6 are the *exact inverse* of this collective process: a R-L PAN occurs with *global* ascension through pitch-space and *local* descending motion at each

entry point (note that the three gestures under comparison are all *accumulative*). Thus, not only has the directionality of the panning gesture (now renewing its reign as principal gestural type following movements 4 and 5) been reversed, but so too has the interlinked pitch activity. This observation is important because the R-L PAN of movement 6 constitutes the first instance of R-L PAN's localized (intra-movement) independence: it no longer serves as a point of departure to L-R PAN. The opening L-R PAN of movement 1 establishes not only the superiority of the panning gesture (irrespective of directionality) over other gestural categories but also the local primacy of left-to-right motion. Thus, it is strikingly fit for inverted pitch-space activity to accompany the cross-movement directionality reversal of the quartet's expositional panning gesture. Figure 6-37 illustrates this relationship.

nts	0:	<9te>							<et98></et98>
Poi	1:		<678>					<7654>	
mble	2:			<345>			<3210>		
Ense	3:				<012>		Х		
		{SG- 0	1	2	3	//	SG- (3+2)	1	0 }
						TIME (x)	(\rightarrow)		
Movement 1; mm. 1–13			Moven	nent 6; mr	n. 12–14				

Figure 6-37: Graphical representation of two panning gestures and their constituent contour relationships

In Figure 6-37, the x-axis indicates approximately proportionate diachronic time, with // signifying the substantial time-lapse and intervening spatial activity between the first and sixth movements. The y-axis shows the local contour (cseg) at each point as the various instruments enter successively, in imitative fashion, to form each of the two gestures. The local ascents in movement 1 and the local descents in movement 6 are readily apparent from the increasing or

decreasing (respectively) string of integers (representing c-pitches) at each y-point in the graph. All of the csegs in the movement-one column may be *translated* to <012> and those in the movement-six column translated to <3210> to highlight the contour equivalence at each point (within either gesture). However, the untranslated forms at the different y-points within the figure reveal the registral shape of each gesture (spanning all ensemble points) as a whole: the global descent of SG-|0123| and the global ascent of SG-|(3+2)10| are quite evident. Most importantly, Figure 6-37 demonstrates the inverse relationships between the two gestures—in the realm of both pitch-space *and* ensemble-space activity. The X at point 3 in the Movement-6 column indicates the absence of a scalar descent at that point. In light of the relation between these two temporally remote spatial gestures, the fusion of points 3 and 2 in SG-|(3+2)10| (the latter occurring gesture) can be seen to reflect a global process of spatial reconciliation. As we shall see, this process continues in the final movement of the quartet.

<u>6.7 Movement 7</u>

Movement 7 begins with the familiar ALL PLAY accentuation that has begun every movement except the first and fifth. A fused point shift introduces the right-to-left directionality that will pervade the quartet's closing movement: PS (3+2)-(1+0) occurs in mm. 30–31 and is repeated (in slightly varied form) in mm. 34–35. This "superpoint shift" will gradually disperse into a full-blown R-L PAN. SG-|(3+2)10| emerges in mm. 35–38. While points 1 and 0 have disengaged from one another in this panning gesture, the fusion of points 3 and 2 persists. However, superpoint 3+2 fragments in mm. 52–55 at the appearance of SG-|3210| (see Figure 6-38). Thus far, the twin superpoint shifts and two panning gestures have all moved from the right to the left of the ensemble space in an accumulative manner. These four gestures constitute a

405

"right-to-left spatial complex" that spans mm. 30–55. Besides featuring directional unity, the complex entails the step-by-step conversion of a right-to-left point shift into a perfect R-L PAN.



Figure 6-38: Mvt. 7, mm. 52-55

Several closed gestures initiate left-to-right motion but do not complete the motion through the ensemble space. SG-|0210| occurs in mm. 56–67 as descending linear activity is woven through the prevailing ALL PLAY texture. The opening point shift of this gesture, PS 0-2, moves left-to-right in a disjunct manner (skipping point 1); however, point 1 is "filled-in" by the spatially conjunct motion back to point 0. Within the context of a closed gesture, subgesture SG-|210| reaffirms the newly established dominance of right-to-left motion. The retrograde of SG-|0210|, SG-|0120|, unfolds in mm. 91–98 as the principal theme (active at point 0 since m. 81) is fragmented, sequenced, and spatially dispersed through the ensemble. This gesture begins as a L-R PAN, but "closes" by shifting back to point 0 instead of continuing on to point 3.³⁸

³⁸ In actuality, an inferior (on the basis of reduced salience) L-R PAN *does* occur in mm. 91–99. The entrance of point 3 in the pickup to m. 99 does *not* involve the thematic fragment that has been panned among ensemble points 0-2 as the cello was resting in silence (in mm. 94–98). For this reason, SG-[0123] is secondary to SG-[0120].

Interwoven oscillation activity occupies mm. 124–168. A broad SG-|101| (strong adjacent-point oscillation gesture) occurs in mm. 124–141 as an eighth-note scalar passage is passed between the violins. As violin 2 (point 1) first renders this passage (mm. 124–129), it is "sandwiched" by SG-|0(2+3)0(2+3)0(2+3)0(2+3)0(2+3)0(2+3)|, a weak distant-point oscillation gesture produced by the vacillation of an eighth-eighth-quarter rhythmic motive. Figure 6-39 depicts a portion of this gesture, SG-|0(2+3)0(2+3)0(2+3)0(2+3)...|.



Figure 6-39: Mvt. 7, mm. 124-126

As point 0 takes-up the scalar passage (mm. 131-135), the more localized oscillation activity becomes a weak adjacent-point oscillation gesture: SG-|1(2+3)1(2+3)1(2+3)1(2+3)1(2+3)1(2+3)1(2+3)1(2+3)|(2+3)]. As the scalar passage returns to point 1 for mm. 136-141, the local oscillation gesture likewise resumes as SG-|0(2+3)0(2+3)0(2+3)0(2+3)0(2+3)0(2+3)|. In mm. 142, point 2 joins the scalar activity of point 1. Point 0 supplants point 1 in bar 145. The ongoing oscillation activity continues with SG-|30303| (mm. 142–143), which is soon compressed into SG-|3131313| (mm. 145–147). Following this strong distant-point oscillation gesture is an ALL PLAY scenario, which spans mm. 148–159. The oscillation activity draws to a close with SG-|(1+2)3(1+2)3| (3+2)0(3+2)0| (mm. 160–167; see Figure 6-40). This large gesture actually comprises two smaller gestures related by a dual transformational operation: weak adjacent-point oscillation gesture SG-|(1+2)3(1+2)3| is inverted (type-2) and expanded to become weak distant-point oscillation gesture SG-|(3+2)0(3+2)0|.³⁹ This inversional process is indicated by the curved arrow drawn from the top system to the bottom system at the right of Figure 6-40.



Figure 6-40: Mvt. 7, mm. 160–168

³⁹ As SG-|(1+2)3(1+2)3| is inverted across the ensemble-space segment in which it occurs (segment 1–3, with point 2 as axis of inversion), it becomes SG-|(3+2)1(3+2)1|. This gesture is then *left-expanded* to become SG-|(3+2)0(3+2)0|. Alternatively, we could envision SG-|(1+2)3(1+2)3| undergoing type-1 inversion (within the total ensemble space), yielding SG-|(1+2)0(1+2)0|. This gesture could then be *right-expanded* to become SG-|(3+2)0(3+2)0|. The former procedure involves extending a single point (1 to 0), while the latter entails extending a superpoint, wherein (1+2) becomes (2+3).

Further right-to-left motion is evinced in mm. 193–202. PS (3+2)-(1+0) appears in mm. 193–194 and again in mm. 197–198. As occurred earlier in the movement, one of the superpoints of this established point shift splinters to yield SG-|(3+2)10| (mm. 199–203). However, SG-|(3+2)10| is *not* followed by SG-|3210| as it was earlier in the movement, within the right-to-left spatial complex of mm. 30–55. After a substantial ALL PLAY section, SG-|3232(2+1)0(2+1)0| appears in mm. 264–271 (see Figure 6-41) as the movement's prime motive is passed through the ensemble space.



Figure 6-41: Mvt. 7, mm. 264-271

This weak open gesture ultimately spans the entirety of ensemble space from point 3 to point 0, displaying a right-to-left directionality. Embedded in this gesture is SG-[32(2+1)0] (event slots 3-6), a perfect R-L PAN. This subgesture demonstrates the larger gesture's prevailing directionality. SG-|3232(2+1)0(2+1)0| begins with PS 3-2. The rightward motion is interrupted as the gesture slides back to point 3 to begin the traversal process again. After resuming and completing the motion to point 0, superpoint 2+1 "neighbors" away from, and then back to, point 0. The presence of superpoint 2+1 suggests that the gesture might work its way back to point 3, which would yield a RIGHT SWEEP. We might reasonably expect such a gesture, given that sweeping gestures occurred quite saliently earlier in the quartet. However, as 2+1 resolves back to point 0, its prolongational role is made clear. We might also view this large gesture as the pairing of two weak adjacent-point oscillation gestures-the conjoining of which produces the internal right-to-left panning gesture. SG-[3232] (occupying the first four event slots of SG-|3232(2+1)0(2+1)0| and SG-|(2+1)0(2+1)0| (inhabiting the last four event slots of the compound gesture) are related via nudging: the former gesture is left-nudged (+2) and right-extended (via point fusion) to yield the latter gesture.⁴⁰ In mm. 264–267, points 3 and 2 are engaged in imitative activity (both melodic and spatial), and in mm. 268–271, superpoint 2+1 and point 0 imitate one another in the same manner. On a larger scale, the spatio-musical activity of ensemble-space segment 0-3 imitates the preceding activity at segment 2-3.

PS (3+2)-(1+0) recurs twice more in mm. 286–289. It now appears without either SG-|(3+2)10| or SG-|3210|. The regional right-to-left complex presented in measures 30–55 has been systematically deconstructed throughout this movement, as illustrated in Figure 6-42. The three occurrences of the right-to-left spatial complex collectively form a "right-to-left *super* complex"

⁴⁰ The two oscillation gestures are likewise related by *double inversion*—discounting the fusion of point 2 to point 1 in SG-|(2+1)0(2+1)0|.

that spans 259 measures of the seventh movement. The supercomplex itself features the progressive dissolution of right-to-left panning activity: SG-|3210|—the apotheosis of the original right-to-left spatial complex—is discarded in complex 2. SG-|(3+2)10| is subsequently eliminated in complex 3, which consists entirely of the two superpoint shifts. Indeed, SG-|(3+2)10| of complex 2 (mm. 199–202) is the last *perfect* panning gesture of the entire quartet. The reductive procedure pervading the supercomplex irrevocably diminishes the salience of right-to-left motion, which was so strongly established by the initial right-to-left spatial complex (complex 1 of the supercomplex).

Complex 1 (mm. 30–55): PS (3+2)-(1+0) → PS (3+2)-(1+0) → SG-|(3+2)10| → SG-|3210| Complex 2 (mm. 193–202): PS (3+2)-(1+0) → PS (3+2)-(1+0) → SG-|(3+2)10| Complex 3 (mm. 286–289): PS (3+2)-(1+0) → PS (3+2)-(1+0)

Figure 6-42: Right-to-left supercomplex (deconstruction of the right-to-left spatial complex)

Following more ALL PLAY, an incomplete RIGHT SWEEP appears in mm. 313–328: SG-|21(2+1)0123|. This gesture emerges from the musical texture through the passing of a quarter-note–eighth-rest–eighth-note rhythmic motive accompanying a conjunct melodic ascension (save for the entrance of points 2 and 3 in bar 327). This sweeping gesture contains the second (of two) not altogether salient instances of left-to-right motion for this movement.

As the final movement of the quartet draws to a close, oscillation activity and polystreaming abound. Following the 22 measures of ALL PLAY spanning mm. 328–349, SG-|3131...3| zigzags its way through mm. 350–363. This strong (i.e., "closed") distant-point oscillation gesture predicts the large-scale spatial closure that is eminent. In bars 369–382, the two main motives of this movement become antiphonally intertwined, resulting in *dissonant polystreaming* (see Figure 6-43).



Figure 6-43: Mvt. 7, mm. 369–388 (closing)

During these final 13 measures, the passing about of the predominantly conjunct (melodically), quarter-note-eighth-rest-eighth-note rhythmic figure yields SG-[2103010]. In Figure 6-43, the similar musical activity at each event point yielding this gesture is indicated by enclosure in a dashed rectangle (some of which begin in the last measure of a system and continue in the first measure of the following system). Although this gesture does not navigate the total ensemble in a directionally uniform fashion, right-to-left motion prevails. PS 2-1, PS 1-0, PS 3-0, and PS 1-0 all move right-to-left, with only two instances of left-to-right motion: PS 0-3 and PS 0-1. Concurrent to SG-[2103010], the point-to-point bouncing of the mostly disjunct, eighth-eighthquarter-note figure forms SG-[303212121301212] (indicated in the figure by enclosure in a solid rectangle). PS 3-0 occurs twice, disjunctively spanning the total ensemble space in a rightto-left motion. Embedded in this more active stream are two weak adjacent-point oscillation gestures: SG-|21212121| (event slots 4–11) and its truncated retrograde (or inversion), SG-|1212| (event slots 14–17). We may reduce the gesture to SG-[303213012] by eliminating the repetitive oscillation activity. From this reduced SG-label, we can clearly discern two weak closed gestures, appearing here for the first time in the entire quartet: SG-|0321| (event slots 2-5) and SG-3012 (event slots 6–9). These two gestures are related by type-1 inversion, and they both instantiate pronounced Out \rightarrow In motion (in the ensemble space–as–container), as the outer ensemble points are activated prior to the inner ensemble points in each gesture. In the larger gesture, we see that each weak closed gesture (related by inversion) leads into oscillation activity (also related by inversion) as its last point shift is repeated: SG-[0321212121], SG-[301212]. SG-[0321] and SG-[3012] are *displaced panning gestures*: displacing each event point of SG-[3210] (a perfect R-L PAN) one slot to the right (via rotational slot displacement), and allowing event point 0 (in event slot 4) to "wrap-around" to event slot 1, produces SG-|0321|. Performing the

same procedure on SG-|0123|, a perfect L-R PAN, yields SG-|3012|.⁴¹ Both perfect panning gestures were featured prominently throughout the quartet, and we may view them as being transformed (via rotational slot displacement) in the final bars of the quartet. The displaced panning gestures supplement the piecemeal disintegration of panning activity witnessed earlier by the right-to-left supercomplex.

At the tempo change in m. 383, the final spatial gesture of the quartet begins. The eighth–eighth–quarter-note figure disappears altogether, and the melodically ascending quarter-note–eighth-rest–eighth-note figure closes the work. The cello (point 3) begins stating this figure as open perfect fifths (the frame of a C-sharp-major/minor chord) are sustained at points 0 and 2. Point 1 enters with a "response" to point 3's "call" in m. 385. Points 0 and 2 (joined by point 3) enter with the figure in m. 386 after ceasing their sustainment of the open fifth. Thus, mm. 383–386 contain the imperfect panning gesture SG-[31(0+2+3)], which spans the entirety of the ensemble space in a right-to-left motion. The core motion of this gesture is expressed by SG-[310]. However, the entrance of point 0 (event slot 3) is fused with activity at point 2 (which effectively "fills-in the gap" left by the initial PS 3-1) and point 3 (causing the gesture to be fully accumulative). Following the expansion of the sonic space provided by this accumulative, imperfect R-L PAN, three ALL PLAY C-sharp-major chords emphatically conclude the quartet.

6.8 Overview: The "Big Picture"

Let us retrace the spatial design of Beethoven's Op. 131 in an effort to depict the largescale, dynamic interplay between spatial gestures. The most frequent and saliently articulated spatial gestures throughout the entire quartet are panning gestures, but the syntactic function of

⁴¹ A perfect panning gesture (strong open gesture) becoming a displaced panning gesture (weak closed gesture) is a type-C intercategorical alteration.
the strong open gestures varies between outer and inner movements. In the outer movements, panning gestures are primary and other gestures (such as oscillation and closed gestures) are secondary. In the interior movements of the quartet, the local predominance of other gestures (oscillation gestures in particular) serves as a large-scale departure from the prevalent panning activity of the outer movements. In movements 2 and 4, oscillation gestures are principal while panning gestures function as spatio-gestural interruptions (localized points of departure). Movements 4 and 5 feature the most pronounced challenge to the dominance of the panning gestures, panning gestures, sweeping gestures, weak closed gestures, and strong closed gestures all make salient appearances. Movement 5 includes the reiterated cyclic panning gesture/LEFT SWEEP complex as well as the recurrent pizzicato figure that progressively dilutes gestural identity and salience.

A sense of gestural departure-and-return encompasses the entire quartet as panning gestures are firmly established (movements 1–3), confronted and opposed by other gestures (movements 4–5), and ultimately re-established (movements 6–7). The dynamic interplay of the various gestural categories triggers a spatial process that spans the quartet as a whole. Additionally, a strong sense of large-scale directional balance is lent by the specific types (directionality of motion) of panning gestures in the outer movements of the quartet. Movements 1 and 7 each embrace the departure-and-return quality of the entire quartet in microcosmic fashion. Movement 1 features the L-R PAN. The left-to-right directionality is briefly challenged by the R-L PAN contained within the LEFT SWEEP of mm. 67–81, and a solid L-R PAN appears as the movement's final spatial gesture—reinstating the local preeminence of left-to-right motion. The opposite process occurs in the seventh movement. Following the prorogation

415

of panning activity in the quartet's inner movements, a right-to-left directionality is firmly established in movement 6 and sustained in movement 7. Oscillation activity is interspersed throughout the seventh movement (reflective of the prominent oscillation activity in the interior movements), and a RIGHT SWEEP provides the movement's only semi-salient occasion of leftto-right activity. The final gesture of the seventh movement is an imperfect R-L PAN. We may boil down the spatial activity of the entire quartet to discover a LEFT SWEEP operating at the deepest background level of the work. From mm. 1–13 of movement 1 through m. 97 of movement 4, the L-R PAN is refined as it undergoes temporal diminution and textural "pruning" (changing from accumulative to migratory). This transformative process strengthens the primacy of the left-to-right directionality in the first half of the quartet. Movement 5 functions as the hinge for the all-encompassing LEFT SWEEP. In this movement, the L-R PAN retains its pronounced conspiculty—particularly as it occurs three times in direct succession as a migratory cyclic panning gesture. However, the directionality of the L-R PAN is counterbalanced by the LEFT SWEEP that overlaps the third L-R PAN of each statement of the cyclic panning gesture: the final subgesture (SG-|0123|) of the cyclic panning gesture (SG-|012301230123|) is likewise the first subgesture of the perfect LEFT SWEEP (SG-[0123210]). Although outnumbered threeto-one each time it is blended with the cyclic panning activity (three L-R PANs followed by one R-L PAN), the component R-L PAN of this sweeping gesture provides the spatio-directional turning point of the entire quartet.⁴² The perfect sweeping gesture, which occurs a total of six times in the quartet's fifth movement (four of which are fused with the cyclic panning gesture), functions as the fulcrum on which the quartet's large-scale spatial design pivots. In movements 1-4, the *left-to-right* panning gesture predominates. In movements 6 and 7, the *right-to-left*

⁴² The tenacious cyclic panning gesture—comprising the persistent repetition of L-R PAN—fleetingly affirms the retained predominance of left-to-right motion and is perhaps likewise indicative of L-R PAN's "reluctance" to relinquish its title as principal gestural form.

panning gesture supplants SG-[0123] as the principal gestural form. The LEFT SWEEP of movement 5 strikes a balance between these two directionalities. The transformational relationship between the two panning gestures (and the fact that they belong to the same gestural category) enables their mutual prevalence to unify the spatial construct of the ensemble space; however, the distinct syntactic function of the gestures' opposing directionalities proves essential to the spatial rhetoric of the quartet. The sweeping gestures in movement 5 function as smallscale (surface-level) manifestations of the larger LEFT SWEEP that embraces the quartet as a whole, as well as of several "middleground" SWEEPs identified in the foregoing analysis. The hinging activity of the "foreground" LEFT SWEEP of movement 5 is prolonged and emphasized by the compulsive, verbatim repetition of its two concrete forms (first exposited in mm. 33–38 and mm. 67–69).

Movement 6 features the R-L PAN to the exclusion of any other spatial gesture, and this gesture is prolonged through movement 7, where its salience is compromised by the right-to-left supercomplex.⁴³ In fact, SG-|3210| of mm. 52–55 (mvt. 7) is the proper close to the large-scale SWEEP embracing the entire work. This R-L PAN, which is the final component of right-to-left complex 1 and the first gesture to be eliminated within the right-to-left supercomplex, is the last salient perfect panning gesture the event cardinality of which equals the ensemble-point cardinality of the quartet (i.e., no fused superpoints are present in this gesture).⁴⁴ This accumulative R-L PAN (SG-|3210|) is the gestural counterpart to the accumulative L-R PAN (SG-|0123|) that opens movement 1. Although strong open gestures dominate Op. 131, the

⁴³ The structurally significant inverse relationship between the opening L-R PAN of movement 1 and the R-L PAN (mm. 12–14) of movement 6 is detailed in Figure 6-37.

⁴⁴ In general, gestural salience increases as the event cardinality of a gesture increases. Thus, while SG-|3210| and SG-|(3+2)10| are both classified as perfect panning gestures (exhibiting exclusively conjunct motion), the former gesture has a higher event cardinality (4), and is thereby more salient, than the latter (with an event cardinality of 3). The fusion of points 3 and 2 decreases the event cardinality of SG-|(3+2)10|.

spatial form of the entire quartet—as fashioned by the expansive interaction of the two directionally opposed panning gestures—is strongly closed (at point 0). Long-term spatial closure is evinced by the departure-and-return character of the broad LEFT SWEEP: point 0 is firmly established during the first violin's solo rendering of the fugue subject in mm. 1–4 of the first movement. Subsequent left-to-right panning gestures (in the first five movements of the quartet) continuously provide *departure* from point 0. The R-L PANs of movements 5, 6, and 7 stress the *return* to point 0 (from point 3, the other spatial extreme of the ensemble space).

The cumulative disintegration of right-to-left motion follows from the R-L PAN of mm. 52–55 (mvt. 7), as does the gradual shrouding of any clear spatial gesture whatsoever. Although several gestures are identifiable in the section spanning from the end of the right-to-left supercomplex (mm. 289) to the movement's final measure (m. 388), these gestures are opaque and vaguely perceptible at best. We may view the latter portion (and majority) of the movement, from m. 56 to m. 388, as a "spatial coda" to the quartet as a whole. The large-scale spatial design of the work is completed by the R-L PAN of mm. 52-55; the directed return to point 0 in m. 55 marks the resolution of spatio-gestural activity. The interlinked processes of gestural disintegration and spatial reconciliation occur throughout the "spatial coda"-explicable, in part, by the right-to-left supercomplex.⁴⁵ Instead of contributing to the gestural syntax and directional rhetoric of the quartet, the elaborate dissonant polystreaming of mm. 369-382 and the emasculated, imperfect R-L PAN of mm. 383-388 engender and reinforce the large-scale process of spatial reconciliation carried out in the "spatial coda." Throughout the quartet, spatial reconciliation has oftentimes occurred on the surface of the spatial construct; the process now serves to bring the spatio-gestural activity of Op. 131 to a definitive close following its structural

⁴⁵ The process of gestural disintegration at the close of the seventh movement is anticipated by the thrice-occurring pizzicato sequence in movement 5, wherein concrete gestural forms become progressively indiscernible.

resolution in m. 55 of movement 7. The homorhythmic and fortissimo ALL PLAY of the work's final three bars confirms and emphasizes the global consolidation of ensemble-point activity.

CHAPTER 7

IMPLICATIONS, APPLICATIONS, AND CONCLUSIONS

In this final chapter, the spatio-analytic methodology proposed and applied in the preceding chapters is summarized and brought into the larger context of musical practice. Three primary questions shall be addressed: 1) what is the relevance of spatial analysis to performance practice; 2) may the concepts outlined above be applied in any creative contexts outside of analysis; and 3) how might this line of research be extended?

7.1 Spatial Analysis and Performance Practice

A spatial analysis of Beethoven's *Quartet in C# Minor*, Op. 131 reveals the tremendous unifying capacity of spatial gestures. The spatial construct of a musical work is autonomous to a certain extent, but spatial gestures often interact with events in other musical parameters particularly those in pitch space. In fact, the extra-spatial sonic qualities of a gesture affect its perceptual salience: markedly similar or identical events occurring at separate times and ensemble points group together more readily than do highly dissimilar events. Qualitatively similar events constitute what Albert Bregman refers to as an *auditory stream*, which is segregated from the larger auditory scene via the process of *sequential integration*. However, spatial gestures may likewise emerge from a sequence of unrelated (sonically and/or musically) sounds that are dispersed over time and space. Grouping on the basis of temporal proximity may override the relative lack of similarity between consecutive sonic events. The *Gestaltqualität* of a spatial gesture—its emergent shape and directionality of motion—transcends the sonic relatedness (or lack thereof) of its component sound events. Furthermore, the relations among different gestures are not dependent upon the quality or musical function of the sounds composing the gestures: SG-|0123| is a left-to-right panning gesture regardless of its time span and overall sonic profile. In Beethoven's string quartet, many gestures feature a qualitative similitude among constituent events, while others do not. In addition, the numerous panning gestures that unify the spatial construct of the whole work (across movement boundaries) are forged by a multitude of contrastive musical events. In Op. 131, we also find layers of gestural activity (surface-level gestures embedded in more vast, "background" gestures), multi-gestural interaction (polystreaming), and specific directionalities serving particular syntactic functions.

Prior to proceeding with the spatial analysis of Op. 131 in Chapter 6, we had to postulate a specific seating arrangement of the ensemble that might perform this work. How valid is this presumption? The "cello-out" configuration that we selected is not the only seating plan employed by modern string quartets. It is noted by Robin Stowell that "most [quartets] opt for the traditional semi-circular arrangement of (from left to right as one looks 'front-on'): Violin 1, Violin 2, Viola, Cello."¹ However, it is not unusual to find the viola and cello positions exchanged, such that the violist assumes the rightmost position (point 3) and the cellist resides at point 2.² Furthermore, it is not uncommon for the violinists to occasionally swap parts between quartet-works without changing their positions—such that the first violinist at point 0 is actually playing the second violin part while the second violinist at point 1 plays the first part.³ An

¹ Robin Stowell, "Extending the Technical and Expressive Frontiers," in *The Cambridge Companion to the String Quartet*, ed. Robin Stowell (Cambridge: Cambridge University Press, 2003), 165.

² Simon Standage, "Historical Awareness in Quartet Performance," in *The Cambridge Companion to the String Quartet*, ed. Robin Stowell (Cambridge: Cambridge University Press, 2003), 147.

³ Such "part switching" between performers (without "point switching" within ensemble space) is likewise common between the two trumpets of a brass quintet.

entirely new spatial analysis would be needed to account for a performance of Op. 131 in which the ensemble was seated either in the "viola-out" configuration and/or with the parts (but not positions) of the violins switched—the analysis presented in Chapter 6 would be utterly nullified. For example, all of the L-R PANs (strong open gestures) identified in the preceding analysis would become weak closed gestures in the viola-out seating plan: SG-|0123| becomes SG-|0132|. This is a type-C alteration that results from comparing the emergent gestures of one seating plan with those of another. Type-C alterations are the most pronounced intercategorical alterations; in this case, a strong open gesture is both "closed off" and "weakened," and the categorical identity of the quartet's principal gestural form and its "opponents" would be changed. Although the relations among spatial gestures would vary proportionally with a different seating plan, and an argument could still be made for spatial coherence, an alternate seating plan would alter the structural dynamic between the three primary gestural categories.⁴ While many of the relations between different gestures are preserved in varying seating plans, others are not (or else are considerably weakened).⁵ SG-[0123] and SG-[3210] are related by both retrogression and inversion (one or the other, not both together). SG-|0132| and SG-|2310| (the "viola-out" version of SG-[3210]) are related solely by retrograde. The fact that each weak closed gesture is bidirectional subverts the previously identified directional opposition of the work's large-scale

⁴ Of all gestural kinds, oscillation gestures are the least affected by alternative seating designs. An oscillation gesture will remain an oscillation gesture, and retain its "strong" or "weak" modifier, from one seating plan to another. Whether it is an adjacent-point or distant-point oscillation is the only quality that might vary. ⁵ It is interesting to consider the modifications that take place to point pairings as the performing ensemble changes from the cello-out to the viola-out seating plan. In a four-point ensemble space, the fusion of points 0 and 1 (0+1) and points 2 and 3 (2+3) is unaffected from one seating plan to the other: superpoint 2+3 remains intact regardless of the relative positioning of the viola and cello. However, consider the pairings 0+2 and 1+3—disjunctive fusions that are interlaced (overlapping) in the ensemble space. With the viola-out schema, this pairing becomes 0+3, 1+2—outer points are now paired and set against fused inner points. Finally, consider the three-against-one spatial imbalance featured prominently in movement 4 of Op. 131. The sense of bilateral imbalance is heightened by the isolation of point 3 (the far-right spatial extreme of the ensemble space) and the antagonistic activity of superpoint 0+1+2. If the ensemble were seated viola out, point 2 would be set against points 0, 1, and 3: an inner ensemble point (point 2) would function in opposition to the both the outer ensemble points (points 0 and 3) and peripheral point 1 (in this context, point 1 is considered peripheral because it is fused to point 0).

spatial structure (L-R PAN "versus" R-L PAN), as exhibited locally by the sweeping gestures of the fifth movement.⁶ We may observe that some seating arrangements more effectively enliven the spatial construct (by augmenting the salience of spatial gestures and their interrelationships) and contribute to the unfolding of the spatio-musical narrative.

From these observations, an important question surfaces: did Beethoven have a certain seating arrangement in mind when composing his quartet? Did the string-quartet ensembles of Beethoven's day all sit in one (and only one) manner? A definitive answer to these questions might betray a preoccupation on the composer's part with the spatial dispersal of the quartet's musical material. Unfortunately, the answers to these questions do not come easily, partially due to the presence of conflicting information in the existing literature. Simon Standage observes:

There is scant evidence of how the players [of a string quartet] were arranged in an eighteenth-century quartet, but iconography suggests that just about every possible seating permutation seems to have been tried in the nineteenth century. Evidence can be divided into two broad categories, according to whether the two violins are side by side or opposite each other.⁷

It is elsewhere noted by Tully Potter that "the classic nineteenth-century seating arrangement for quartets was to have the violinists facing each other at the front, with the cellist on the leader's left at the rear and the violist on the second violinist's right."⁸ Potter asserts that nineteenth-century quartets assumed the following configuration: Violin 1, Cello, Viola, Violin 2. This

⁶ I maintain that SG-|0132| is less well formed (and thereby less perceptually salient) than SG-|0123|, and that it produces an entirely different affect—one of partial closure instead of implied continuation (this point shall be amplified below). In the spatial analysis of Chapter 6, long-range spatial closure occurs in the last two movements (6 and 7) of Op. 131 as R-L PAN unseats L-R PAN as principal gestural form. In considering the conversion of SG-|0123| to SG-|0132| (and SG-|3210| to SG-|2310|) that occurs if the performing ensemble adopts the viola-out seating plan, we find an increase in localized (small-scale), partial (weak) spatial closure. For example, mm. 1–14 of mvt. 1 (Figure 6-1) and m. 89 of mvt. 4 (Figure 6-12) both feature descending registral motion that encompasses all four parts. Since descending melodic motion is typically associated with closure (as discussed in Chapter 5), it might seem apposite for these passages to be rendered spatially as weak closed gesture SG-|0132| (as would occur in the viola-out plan). However, under the cello-out configuration, the emergent perfect panning (strong open) gestures propel the music forward by denying closure and implying/inviting continuation, and the large-scale spatial structure (the background LEFT SWEEP) provides strong closure toward the end of the entire work.

⁷ Standage, 147.

⁸ Tully Potter, "From Chamber to Concert Hall," in *The Cambridge Companion to the String Quartet*, ed. Robin Stowell (Cambridge: Cambridge University Press, 2003), 42.

seating plan—reflective of the standard orchestral seating of the day—seems to have prevailed in Germany, and was perhaps favored due to the "conversational" quartet styles of Haydn, Mozart, and early Beethoven. The antiphonal separation of the upper parts (which were frequently engaged in musical "dialogue") served to clarify the texture.⁹ Nineteenth-century quartets in London, on the other hand, preferred the "adjacent-violins" plan with either the cello or the viola at the rightmost point.¹⁰ It is this seating plan that persists into the twentieth century and is still standard for contemporary string-quartet ensembles.

Most of Beethoven's late string quartets were composed for the Schuppanzigh Quartet, which is considered to be one of the first professional string quartets. However, Op. 131 was given its first private performance by Josef Böhm's quartet and its first public performance by the Müller Quartet in Halberstadt in 1828—a year after Beethoven's death.¹¹ There is no direct evidence as to how any of these quartets arranged themselves—or even if they consistently abided by a particular seating plan. Beethoven could have been familiar with any number of seating arrangements. We can infer from evidence that he could very well have been most well acquainted with the "violins-out" plan described above. In such a plan, the profuse SG-[0123] identified in Chapter 6 would be experienced as SG-[0321]. Again, the co-principal gestures of the quartet would be weak closed gestures instead of strong open gestures, and their interrelatedness would be minimized.¹²

We must embrace the fact that spatial gestures—the objects of spatial analysis—are often the products of musical performance despite any lack of intentionality (as pertains to

⁹ Standage, 147. We might anticipate distant-point oscillation gestures (between point 0 and 3) arising from such a scenario.

¹⁰ Ibid.

¹¹ Potter, 42–45.

¹² SG-|0123| (in the cello-out configuration) becomes SG-|0321| (in the violins-out seating plan) and SG-|3210| becomes SG-|1230|. SG-|0321| and SG-|1230| are displaced panning gestures related via retrograde but not by inversion. Interestingly, SG-|0321| is "displaced" (abstractly, via rotational slot displacement) from perfect panning gesture SG-|3210|, not from SG-|0123|. Likewise, SG-|1230| is displaced from SG-|0123| (not from SG-|3210|).

spatialization) on the part of the composer. Since we are a twenty-first-century audience, we must consider the spatial gestures that are likely to emerge from a modern quartet's performance of Op. 131. An ensemble, however, is responsible for interpreting the music it presents to the public. Without dwelling on (or discounting) the issue of "historical performance practice," I suggest that an ensemble's choice of seating configuration is a matter of interpretation. Many composers-particularly those in the twentieth century-deliberately spatialize their musical works. Often, composers (such as Henry Brant and Karlheinz Stockhausen) provide detailed diagrams of the obligatory dispositions of performers within the performance space. Other times, composers provide general (and possibly vague) instructions: Erickson states that Ives "views the spatial distribution of instruments as an aspect of the interpretation, and that is probably why his remarks to the conductor [in many of his scores] are general rather than specific."¹³ In the majority of musical scores, no spatial layout (diagram of ensemble configuration) is provided by the composer. This, however, does not mean that the music is not spatialized. I have argued that *all* music (for more than one performer) is spatial—not merely works classified under the rubric of "spatial music." Irrespective of a composer's intentionality (documented or surmised), spatial gestures are often readily perceived during a musical performance, and a spatial analysis can unveil the gestures' full capacity for structural unification and conveyance of meaning. Since spatial motions and shapes are variable depending upon the ensemble's onstage configuration, spatial analysis holds significance for performance practice: the manner in which an ensemble elects to arrange itself-if left unspecified by the composerbecomes a matter of interpretation. An ensemble's onstage disposition involves more than the ordering of instruments relative to one another: the spacing between adjacent ensemble points

¹³ Erickson, 150.

can affect gestural salience.¹⁴ Ensemble members may organize themselves in an effort to enhance the perceptual salience of spatial gestures and enliven the spatial construct of a musical performance. Such decisions may conform to or contradict historical informedness.

I have argued that the different categories of spatial gesture carry contrasting interpretive potentials, as do the varying directionalities of gestures. An open gesture is distinct from a closed gesture, a strong gesture is different from a weak gesture, and the meaning of left-to-right motion contrasts that of right-to-left motion. Some types of gestures are intrinsically more salient than others, irrespective of the musical attributes of the events composing the gesture. Strong open gestures are arguably the most salient of all spatial gestures, for reasons broached at different points throughout this dissertation. Several Gestalt principles contribute to the "good form" (Prägnanz) of perfect panning gestures. First, the Law of Proximity (one of the two Gestalt grouping principles) comes into play as the consecutive event activity (relative temporal proximity) at *adjacent* ensemble points (relative spatial proximity) assists the grouping of spatially and temporally discrete events into a unified gesture.¹⁵ Conjunct spatial motion is more readily perceived than disjunct motion due to the temporal *and* spatial "nearness" of consecutive events. Furthermore, provided that the ensemble points are roughly equidistant, each point shift comprised within a panning gesture will span the same approximate distance.

The Gestalt Law of Continuity aids in perceiving a strong open gesture as a unidirectional line traced through ensemble space. The "openness" of the gesture is supplied by the fact that its

¹⁴ Too much or too little distance between neighboring points may reduce the perceptibility of spatial gestures: a tightly compacted ensemble minimizes the distinctions of location necessary for the perception of apparent motion, while a vastly arrayed ensemble challenges the natural grouping of successive events by magnifying the spatial discontinuities between points.

¹⁵ Adjacent ensemble points are spatially proximate regardless of the distance between the points. The fact that no other ensemble points intervene between point 0 and point 1 of a given ensemble space allows their adjacency to equate to relative spatial proximity. Likewise, the direct succession of events equates to the relative temporal proximity of those events—notwithstanding the amount of time separating consecutive events. Gestures unfolding at faster rates (exhibiting an increase in the *actual* temporal proximity of events) are generally more salient than those occurring at slower rates.

spatio-temporal extremes are maximally disconnected and separated in both time and (ensemble) space, and that it spans the entirety of ensemble space with implied continuation in either spatio-temporal direction. The "strongness" of the gesture arises from its unidirectionality—as structured in perception by the metaphorical projection of the PATH schema. The step-by-step, unimpeded spatial progress along the "path" of ensemble space is strengthened by the expectation (or protension) of continuation and ultimate arrival at the projected "goal," which is the endpoint of ensemble space opposite the "source" of the gesture. With each successive point shift in the same direction, the unidirectionality of the gesture as a whole is confirmed. For this reason, the salience of panning gestures is likely magnified in high-cardinality ensemble spaces: SG-[012345] in six-point ensemble space contains a higher number of left-to-right point shifts (5) than does SG-[012] in a three-point ensemble space (with only 2 left-to-right point shifts). In light of this observation, consider the five examples of ascending-fifth melodic motion presented in Figure 4-12. We may now observe that the spatio-gestural counterparts of examples a–e (discussed in the text of Chapter 4) are ordered from most to least salient.

With open gestures (strong or weak), the spatial extremes and temporal endpoints of the gesture are one in the same, and these spatio-temporal "bookends" of the gesture are likewise the spatial boundaries of the ensemble space. Perfect panning gestures unfolding in an acoustic ensemble space emphasize the spatial discontinuities between events less than other gestures that feature disjunctive motion. An analog pan in the electronic medium contains no discontinuities—and therefore no discernible parts—and is easily received as a whole. Panning gestures in an acoustic ensemble space comprising discrete and static points contain spatial discontinuities between successive events, but these discontinuities are minimized as much as possible through the exclusively conjunct motion of the gesture.

427

For any one-dimensional ensemble space, there are only two potential forms of strong open gesture: L-R PAN and R-L PAN. A degree of transformational invariance limits the number of potential permuted gestural forms for this category: both the retrograde and the inversion of SG-|0123| is SG-|3210| (and vice versa), and under retrograde-inversion, SG-|0123| maps into itself (as does SG-|3210|). Such invariance strengthens the structural relationship between the two gestural forms.

Finally, in cases where the ensemble space is a horizontal translation of the vertical orientation of parts in the score (i.e., a perfect score space-to-ensemble space translation), panning gestures emerge quite effortlessly: whenever a broad gesture through pitch space spans all the parts of the score, it will also span all points in the ensemble. Standard notation for string quartet places the staves in the following order from top to bottom in each staff system: Violin 1, Violin 2, Viola, Cello. As such, the instrument capable of producing the lowest pitches is situated lowest within the system. The first violin, which typically performs the highest musical voice (and which often carries the principal melodic material), is given the uppermost position in the system. With the cello-out ensemble configuration, an ascending line in pitch space that is broken-up and distributed among the four instruments would render a R-L PAN: down \rightarrow up (in pitch space) = right \rightarrow left (in ensemble space). The salience of panning gestures in such situations is likely attributable, in part, to the perfect translation of vertical pitch/registral space (parts in the score) to horizontal ensemble space (points in ensemble space). While it is not uncommon for various types of ensembles to position themselves along a continuum of part ranges, other factors—such as dynamic balance and timbral blend—often supersede this concern (as is typically the case with woodwind quintets and brass quintets.)

The above observations regarding the salience of panning gestures motivated my selection of the cello-out seating plan (in lieu of the viola-out configuration) as the foundation for my spatial analysis of Op. 131. This seating plan most effectively animates the ensemble space by yielding the most salient spatial gestures. Gestures would still emerge from alternative seating arrangements, but their presence and interrelationships would arguably be less pronounced. Of course, gestural salience is affected by (but not dependent upon) other musical factors: a spatial gesture boasting similar sounds at each constituent point (an auditory stream) will likely be more pronounced than a gesture consisting of a sequence of unrelated sounds. Overlapping dynamic envelopes between consecutive events might create a more effective illusion of motion. A fast spatial gesture might be easier to comprehend than a slower gesture (due to the enhanced temporal proximity of events). A gesture with consistent onset-to-onset intervals between events might emerge more readily. While extra-spatial musical factors are not required for the formulation and perception of a spatial gesture, they do affect gestural salience and should be taken into consideration in any thoroughgoing spatial analysis.

This dissertation should encourage performing ensembles to take the issue of spatial orientation into consideration. When the composer stipulates or merely suggests a specific ensemble layout, such indications should be considered vital to the integrity of the composition and should be heeded and honored just as the supplied notes, rhythms, and other performance indications should be faithfully observed. When a seating design is left unspecified by the composer, an examination of the performance practices in effect for a particular ensemble type during the composer's conception of the work in question is appropriate. If such inquiries cannot be answered decisively (as with Beethoven's Op. 131), the ensemble must assume the responsibility of choosing the most effective spatial configuration. This choice entails

429

determining the positioning of ensemble members relative to one another within the ensemble space, as well as the size and location of the ensemble space in relation to the audience space within the performance space.¹⁶ When possible, the ensemble might choose to "tune" the hall. Lowering the dampening curtain(s) or acoustic paneling alters the ratio of direct signal to indirect reflections (in favor of the former). The resultant decrease of indirect reflections (reverberation) would enhance the listener's ability to localize each point of sound within the ensemble accurately. Such choices—electing a seating configuration and altering the acoustical attributes of the performance space (when optional)—may be informed by spatial analysis, general knowledge of acoustics, and experimentation. Taking such issues into consideration is an integral component of interpreting a musical work in performance.

7.2 Spatialization in Transcription

New arrangements of preexisting compositions might exploit the spatial construct. The arranger might take creative license to add spatial gestures into a setting in which they did not previously exist. Imagine a Fugue by J.S. Bach originally composed for solo keyboard and subsequently transcribed for woodwind quartet. The original version—by virtue of being conceived for a solo instrument—would not exhibit any palpable spatiality within the realm of physical space (save for the fact that all musical sounds would issue forth from the same point in

¹⁶ Most professional ensembles sit in the same position for every musical work they perform, unless otherwise instructed by the composer. In contemporary practice, works are often commissioned by specific ensembles that work in close conjunction with the composer during the compositional process. At the very least, the composer is likely to have tangential (if not explicit) knowledge of how the target ensemble organizes itself in performance. The notion of potentially using a different seating plan for different works in a concert performance presents some logistical problems, such as an ensemble having to physically rearrange itself—which might involve shifting sheet music and juggling equipment (mutes, reeds, etc.). Another potential difficulty lies in the performers' having to adjust to hearing different instruments and parts coming from different points in space, which might affect cuing and the maintenance of ensemble togetherness. When an ensemble member anticipates a sonic cue, he/she is likely attuned to the directional location from which the sound will ensue. Some reluctance on the part of ensembles to rearrange themselves is justifiable; nevertheless, the potential obstacles are surmountable. Moreover, these obstacles are worth surmounting in order to rouse and activate the spatial construct and utilize all the means of interpretation at the ensemble's disposal.

space). The ensemble space of any composition for solo instrument consists of a single point, so the dynamic interplay between multiple points that yields a spatial gesture is not feasible. However, when creatively arranged for a quartet of wind instruments, the incorporation of spatial gestures could enkindle and "mobilize" the four-point ensemble space—adding an additional layer of coherence, meaning, and—arguably—interest.

An exemplar of an arranger infusing a preexistent, spatially "static" work with spatial gestures is Philip Snyder's arrangement for four guitars of Alberto Ginastera's *Danzas Argentinas*, Op. 2, originally composed for solo piano.¹⁷ Obviously, Ginastera's original setting of the music lacks the spatial interplay that occurs in ensemble music. In his arrangement of Ginastera's music for guitar quartet, however, Snyder takes full advantage of the ensemble space: by dispersing musical events among the four discrete players of the quartet, he creates a cohesive spatial complex in which recurrent spatial gestures provide an important and effective layer of unity and meaning.

Salient spatial gestures may be found throughout all three movements of *Danzas Argentinas*, as arranged by Snyder. I shall briefly discuss the spatial aspects of the second movement, entitled *Danza de la moza donosa*. In this lyrical slow movement, several related gestures occur over different spans of time, conferring large-scale structural unification to the movement as a whole. Figure 7-1 provides a reduction of mm. 1–25. Guitar 4, seated to the far right (if facing the ensemble), begins with a 3-bar solo introductory figure.¹⁸ Guitar 3 enters with

¹⁷ Philip Snyder, "The Music of Alberto Ginastera Transcribed for Guitar: A Performance Edition of Danzas Argentinas for Guitar Quartet," (DMA diss., University of Georgia, 2003). For an article related to this arrangement, see Philip Snyder, "The Music of Alberto Ginastera: A Fertile Source for Guitar Transcriptions," *Soundboard* 31.4 (2006): 19–23. For a recording of Snyder's arrangement, see the Georgia Guitar Quartet, *Mosaic*, Solponticello SP-015, compact disc.

¹⁸ I portray the positions of the discrete quartet members, as seated from left to right from the audience's perspective, as follows: Guitar 1, Guitar 2, Guitar 3, Guitar 4. This is done to simplify my description of the spatial process of the movement (by establishing a 1:1 correspondence between vertical score space and horizontal ensemble space) but does not conform to Snyder's original notation. His Guitar 2 is my Guitar 4, his Guitar 1 is my

the melody in m. 4. Guitar 2 joins the other two guitars with a countermelody in bar 13. Guitar 1 finally enters in bar 25, and all four guitars engage in an ALL PLAY scenario through m. 52. The first spatial gesture is therefore an accumulative right-to-left panning gesture (SG-|3210|) that occupies over 25 measures of music. With the tempo indication of dotted-quarter = 60 in a 6/8 meter, this R-L PAN takes approximately 50 seconds to unfold. As the gesture traverses the ensemble space with its rightward trajectory, the sonic space slowly expands to its full potential (the total ensemble space).



From bars 53 through 58, guitar 2 is utterly silent and the musical activity of guitar 4 is

sparse—bringing about not only a thinning of the musical texture in vertical (and metaphorical)

Guitar 3, his Guitar 3 is my Guitar 2, and his Guitar 4 is my Guitar 1. However, Snyder and myself are both performing members of the Georgia Guitar Quartet, for whom Snyder made this arrangement. My labeling system for the position of each part is based on our actual practice of performing this piece. Snyder has revealed, through personal correspondence, that he was cognizant of the spatialization in his transcription when he distributed the parts to the quartet members. I mention this discrepancy between Snyder's original notation and my own description of the work's spatiality to avoid any confusion that might arise upon inspection of Snyder's arrangement (as published in his DMA dissertation). The musical examples provided herein are altered from Snyder's original version, with his kind permission. The musical examples conform to the in-text description of spatial gestures (which, in turn, conforms to the ensemble's practice of performing Snyder's arrangement).

pitch space, but also a dilution of the sonic space in horizontal (and actual) physical space. This coordinated attenuation of sonic texture and spatial density immediately succeeds the movement's climax (mm. 48-51). On the downbeat of m. 59, guitar 4 begins its restatement of the solitary, 3-bar introductory figure that opened the movement. We are poised to experience a repeat of the expansive R-L PAN exposited in the opening 25 bars. This time, however, the entrances of the other three parts are somewhat altered, as shown in Figure 7-2.



Figure 7-2: Ginastera, *Danza de la moza donosa*, arr. Phil Snyder, reduction of mm. 59–80

Guitars 3 and 2 enter simultaneously in bar 62, meaning that interior ensemble points 1 and 2 are fused (into superpoint 1+2). These two guitars harmonize in thirds and in rhythmic unison (for the most part) through bar 78. The long-anticipated entrance of guitar 1 is deferred until measure 80. Thus, the second large spatial gesture encountered in this movement is SG-|3(2+1)0|, another R-L PAN. Despite the fact that the musical material and exact nature of this gesture differ somewhat from the movement's opening R-L PAN, the two gestures are categorically related (strong open gestures). They likewise display a uniformity of directed motion (right to left) within the ensemble space. It has taken 21 measures (approximately 42 seconds) for this instantiation of right-to-left panning gesture to unfold. Therefore, the latter-occurring R-L PAN is a slight temporal diminution of the former. Both the shortening, in time, of the recurrent gesture and its abbreviation (reduction of event cardinality achieved via point fusion) coincide with another formal process: the drive to the movement's closing cadential figure and final spatial gesture.

The entrance of guitar 1 in bar 80 serves a pivotal function. It not only completes the R-L PAN of mm. 59-80, but it also initiates the movement's final spatial gesture—a migratory L-R PAN (SG-|0123|) that unfolds entirely in the penultimate bar (m. 80), as shown in Figure 7-3.



Figure 7-3: Ginastera, Danza de la moza donosa, arr. Phil Snyder, mm. 80-81

In m. 80 the tempo changes to *molto lento*, and each guitar—in ordered succession from left to right—plays two consecutive pitches: A and E in ascending order. Not only does the melodic

interval of a perfect fourth ascend in each individual part, but the registration also ascends by octave from one guitar's statement of the interval to the next.¹⁹ In the final bar (m. 81), all four guitars articulate the downbeat with a widely spaced [015] trichord: guitar 4 plays an F-natural while guitars 1–3 perform artificial harmonics on F-sharp (guitars 1 and 3) and C-sharp (guitar 2). This final synchronic event fleetingly recalls the ALL PLAY section of mm. 25–52 and may be viewed as a spatial reconciliation of the preceding migratory gesture.

Danza de la moza donosa features three primary spatial gestures: the first two are R-L PANs, and the third is a L-R PAN. These gestures are all related by virtue of their shared membership in the larger gestural category of strong open gesture. These gestures are likewise related by the transformational operations of both retrogression and inversion: either gesture may be viewed as the retrograde (or inversion) of the other. However, since the lone L-R PAN occurs *after* the two successive statements of R-L PAN, we shall view R-L PAN as the principal gestural form and L-R PAN as a transformation of the principal form. Although the work consists of three open gestures, we can infer some degree of spatial closure: the final L-R PAN nicely balances the opposing directionality of the twice-occurring R-L PAN. In fact, the second R-L PAN and the final L-R PAN constitute a RIGHT SWEEP since point 0 (guitar 1) provides both the final event of R-L PAN and the first event of L-R PAN (eliding the last bar of Figure 7-2 with the first bar of Figure 7-3 reveals the RIGHT SWEEP). The sense of spatial balance is

¹⁹ Note that the overall ascension through pitch space encompassing all four parts is represented in the score as descending motion. This results (in part) from the fact that all four instruments in the ensemble are of identical kind—each with the same range of pitch. In other words, the staff system is *not* organized with higher sounding instruments in the upper staves and lower sounding instruments in the lower staves, which, as noted, is a common notational convention for ensembles composed of different instruments. If, in an effort to expand its composite range in both frequency directions, the guitar quartet contained a bass guitar (the part for which would likely be located at the bottom of the staff system) and a piccolo guitar (its part located in the upper staff), we might reasonably expect to find ascending motion in registral space corresponding to ascending motion in "score space." Observations such as this—as well as any observation highlighting the "translation" of the score space to pitch space and/or ensemble space—are of interest and relevance solely to the analyst and, perhaps, to the performers. In a performance situation, the listener generally lacks access to the score. Accordingly, the hypothetical listener would hear only the broad ascent through pitch space articulated as a left-to-right motion through ensemble space.

also enacted on another level: although the R-L PAN occurs twice consecutively (one instance more than the L-R PAN), it unfolds over a relatively long time span in each instance. The longer a gesture takes to unfold, the harder it is to perceive as a unified and continuous motion. The listener must mentally retain the successive order of entrances among the performers within the ensemble. Although the final gesture occurs in a slower tempo (molto lento), it occupies only one bar of music. The temporal diminution of this gesture enhances its perceptual salience. Furthermore, the conclusive L-R PAN comprises uniform sonic activity at each ensemble point—a feature that likewise enhances its cohesiveness and perceptibility. Thus, the final, concise L-R PAN counterbalances the two similar but temporally expansive (and therefore less salient) R-L PANs. The fact that the final event of the movement involves the combined activity of all four ensemble members reflects the ALL PLAY climax from the middle of the movement and provides a final reconciliation not only to the immediately preceding L-R PAN, but also between the opposing directionalities of L-R PAN and R-L PAN, as both spatial gestures are effectively liquidated in the final bar.²⁰

As mentioned above, the final L-R PAN is integrated with a steady ascent through pitch space, resulting in a *composite gesture* that exhibits a bottom-left-to-upper-right directional shape. Here we witness complementary gestures occurring concomitantly in both actual space *and* pitch space. We tend to associate falling melodic lines with a sense of relaxation and closure—as if gravity is pulling the pitch-objects down to a state of repose. Conversely, rising pitch events often evoke tension—as if the musical figure is struggling upward by working against the force of gravity. As described in Chapter 5, a L-R PAN signifies progression through

²⁰ The spatial process of this movement is similar to that of Beethoven's Op. 131. In Op. 131, R-L PANs serve to balance the more prevalent (and precedent) L-R PANs at multiple levels of structure. In *Danza de la moza donosa*, the closing L-R PAN balances the two preceding R-L PANs. Spatial unification achieved through the recurrence of strong open gestures, and spatial closure engineered through both immediate and long-range directional balance, are common features of both works.

metaphorical association and is thereby tensive. Accordingly, a L-R PAN and an ascending line in pitch space collectively impart a pronounced sense of progressive and continuous motion.²¹ However, in the larger context of the entire movement, the departure-and-return character of the overall spatial design is fulfilled by the final gesture, which traces a spatial path back to the ensemble point from which the spatial motion of the entire movement began (ensemble point 3). We can view the RIGHT SWEEP (SG-|3210123|) as the fundamental spatial structure of the movement. After the initial R-L PAN (mm. 1-25), the steady motion back to point 3 is interrupted by the restatement (or, "starting over") of R-L PAN (mm. 59-80). The spatial activity of the second R-L PAN is condensed in two respects: it is temporally diminished, and it involves the fusion of points 1 and 2 such that the gesture as a whole is abbreviated from its original version (the gesture's event cardinality is reduced from four to three). This spatial process is similar to an interrupted period in tonal music, wherein the harmonic motion back to the tonic is "interrupted" by a half cadence at the end of the first phrase. The second phrase begins the harmonic process anew, the harmonic motion speeds-up, and the harmonic goal is finally achieved with an authentic cadence. The tonic chord at the beginning of the second phrase is *not* the resolution of the dominant chord at the end of the first phrase—the proper resolution does not occur until the end of the second phrase. So too with the spatial process of Ginastera's Danza de la moza donosa: the activity at ensemble point 4 at the beginning of the second R-L PAN is not the proper return to this ensemble point. The actual return is postponed

²¹ Perhaps a greater sense of closure would be evinced by a descending pitch event occurring with a closed spatial gesture (of course, the arranger has no control over the former musical event). Closed gestures are interpretively antithetical to open gestures. We must remember, however, that we are presently analyzing the second movement of a three-movement work. The joint lack of closure in the realms of actual physical space and virtual pitch space invites the continuation (and eventual conclusion) of the collection as a whole, which occurs with the commencement of the third and final movement.

until the final L-R PAN, which progresses to ensemble point 4 in conjunct (point-by-point), unidirectional fashion—providing local spatial closure to the movement upon reaching its goal.²²

Snyder's transcription illustrates the point that the creative arrangement of a preexisting work can harness the spatial construct of musical performance: the arranger may add spatial gestures into a setting in which they did not originally exist. The music thereby stands to gain a new layer of structure and meaning on which to engage the listener in performance. Such is certainly the case with Snyder's arrangement of Ginastera's *Danza de la moza donosa*, in which the effects of motivic cohesion, large-scale directional balance and spatial closure, "directional meaning," and inter-parametric concordance (between events in ensemble space and those in pitch space) result from the careful articulation of ensemble-point activity.

7.3 Recommendations for Extension in the Field of Spatial Analysis

In the interest of concision and for the sake of delimitation, I have elected to focus primarily on the one-dimensional ensemble spaces of works from the Western canon. An obvious extension of this research would be to fully develop the language and analytic tools for confronting spatial gestures occurring in multi-dimensional ensemble spaces. I hope to have provided the impetus for such a foray in section 4.3 of Chapter 4, in which I make ancillary observations and propose some descriptive terminology for describing the gestural activity potentially encountered in such spaces. Throughout this dissertation I have briefly described the multi-dimensional spaces of several twentieth-century compositions, but more complete analyses of these and other representative works are needed.

²² The spatial closure provided by completion of the fundamental RIGHT SWEEP (m. 80) complements the spatial reconciliation occurring in the final bar, which represents the cessation of gestural activity and stabilization of the sonic space.

A more fully developed system for analyzing spatial gestures in the electronic and electro-acoustic mediums could easily grow out of section 4.4. Although exhibiting gestures similar to those found in acoustic works, the electronic medium requires different tools for spatial analysis. Since many gestures are analog—continuously panned from one channel to another-integer notation is ineffectual. The listener might experience a L-R PAN, but the gesture cannot be neatly notated with an SG-label because there exist no discontinuities in the spatial motion to segment the gesture into parts. Furthermore, since most electronic works do not exist in score form, a description of gestural activity would need to be undertaken solely in the form of "aural analysis."²³ Software programs that graphically indicate (via real-time animation) the channel distribution and inter-channel motion of sonic events would greatly facilitate spatial analysis. Computer algorithms, such as those designed by Chowning et al. (described in Chapter 1), that simulate the motion of sounds in a quadraphonic environment could potentially be used to analyze the motion of sounds-whether initially generated by the same program or not—in existent electronic repertoire. The computational results of such an analytic process would, of course, need to be synthesized and interpreted by the human analyst. However, the access to such programs, as well as the equipment on which to run them, can be quite limited.²⁴ Nevertheless, spatialization has been an integral element of electronic music since its infancy. In the absence of traditional pitch and rhythmic relations, the recurrence of (and interrelations among) spatial gestures might prove more vital to the unification of an electronic composition than events in any other musical parameter.

²³ The first published score of an electronic composition was that of Stockhausen's *Elektronische Studie II* (1954). The relative rarity of such scores is primarily due to the fact that they are solely intended to serve as study materials—they are not "recipes" enabling a performer to recreate and actualize the musical work, as are most scores (Watkins, 590).

²⁴ A variety of stereo oscilloscopes, which graphically display the inter-channel disposition of sounds, are available on the consumer market.

The qualitative distinction between experiencing music in live performance and in the recorded format requires further examination. In Chapter 2, it was noted that the localization process of externalization is replaced by that of lateralization when sounds are experienced over headphones. Illusorily experiencing sounds inside our heads rather than in our surrounding environment is a very different way of hearing. Many types of music are experienced solely or most frequently through recording (as opposed to live performance). How effectively does a recording capture the spatial relations of a musical work? A thorough examination of recording techniques and audio reproduction might highlight the potential preservation, enhancement, or degradation of a musical work's inherent spatiality. Extending the analysis of spatial activity to popular music might likewise bear fruitful results. In the process of recording many popular music styles, in which the members of a musical group rarely record their respective parts all at once, spatial gestures are often fabricated in the studio. How often is the stereo separation of different instruments and movement of sound artistically motivated, and how often is it merely a convention of sonic engineering? Furthermore, listening to recorded music either over loudspeakers or headphones eliminates visual cues and their interaction with audition (multisensory interaction). How is the salience of a spatial gesture affected by the inability to see the sound-producing agent?²⁵

An exploration of spatialization in non-Western musics would provide the grounds for comparing the different meanings that various cultures might ascribe to the same spatial formation and/or gestural directionality. Spatial relationships are often ritualized and carefully circumscribed during the performance of musical works from a given culture. The metaphoric underpinning of my interpretive method (Chapter 5) is dependent upon a Western mind

²⁵ A speaker reproduces recorded (or manufactured) sounds—it is not the agent originally responsible for the production of the sound.

perceiving the directionality and shape of a spatial gesture. It is not only our embodied and environmental experience but also our cultural upbringing that dictates how cross-domain mappings occur to yield specific conceptual metaphors. Differing cultural experiences likely produce contrasting interpretations of spatial activity.

An issue that could be addressed in reference to any musical style is how spatial gestures might contribute to tone painting (or, text painting) or otherwise serve a programmatic function. Be it an orchestral lied by Mahler or an alternative rock tune by Radiohead, how might spatial activity reflect, modify, or enhance the actions or feelings described (or induced) by the text— whether that text is set (read or sung) in the music or employed and provided as a para-musical apparatus? In addition, the spatial analysis of staged works (such as opera and ballet) would need to take into account the relationships between musical events, spatio-gestural activity, and the physical orientation, movement, and/or gestures of onstage performers (as components of the unfolding narrative).²⁶ To describe the relationships between text, music, spatial gestures, human movement, and/or other visual elements, the application of Gilles Fauconnier and Mark Turner's system for modeling "conceptual blending" might prove tremendously beneficial.²⁷

7.4 Conclusions

The methodology presented in the foregoing chapters proceeds from the premise that spatial gestures exist in music irrespective of the composer's intentionality (or lack thereof) in

²⁶ The interactivity of events occurring in different musical parameters has been tangentially discussed throughout this dissertation. In Chapter 5, I define a *composite gesture* as the blended motion that results from the integration of events in ensemble space (spatial gestures) with those in pitch space (such as a broad registral ascent or descent). While the elucidation of composite gestures (including the amalgamated, conceptual-actual "space" in which they occur) merits further investigation, this discussion could be widened to include extra-musical and extra-spatial aspects of performance.

²⁷ Gilles Fauconnier and Mark Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities* (New York, Basic Books, 2002). Lawrence Zbikowski applies their "Conceptual Integration Network" to account for the blended meaning that arises from the mixture of text and music in lieder. See Zbikowski, *Conceptualizing Music*, 77–95, 243–286.

manipulating the space. In other words, whether or not a composer is known to have consciously constructed and developed spatial gestures in a specific work has absolutely no bearing on the relevance of spatial analysis—the spatial gestures are still present and perceptible in performance. Certainly, a spatial analysis will not yield interesting and meaningful results with any and every work of music; however, music to be subjected to such an analysis is not limited to works in which the composer has expressly indicated concern with the spatialization of musical events.²⁸ Speaking of "musical gesture" (in the more general sense), Robert Hatten claims: "A movement need not be intentionally produced to be interpreted as significant...in cases where the composer may not have specified a gesture through notation, aspects of style and compositional rhetoric may nevertheless imply a certain gestural realization.²⁹ This statement is readily applicable to spatial gestures: although a composer might not have explicitly indicated the intended ensemble seating plan for a given work, the optimal (or, "correct") seating plan may by inferred by attendant knowledge of historical performance practice (as an "aspect of style"). Spatial gestures may likewise be forged, controlled, and accentuated through creative physical arrangement on the part of the performing ensemble, which may shape its spatial configuration as an aspect of interpretation.³⁰ Furthermore, a spatial analysis should not replace other modes of analysis but should engage them and be applied in conjunction with them (where applicable) to yield a comprehensive musical analysis.

In general, composers' interest in the spatial design of music peaked during the twentieth century and has persisted into the twenty-first century. However, contrary to Stockhausen's

²⁸ This "concern" could be betrayed by little more than a suggestive title, such as Thea Musgrave's *Space Play*, or could be explicitly outlined by a composer, as in the quote by Karlheinz Stockhausen at the head of this dissertation.
²⁹ Hatten, *Interpreting Musical Gestures, Topics, and Tropes,* 112–113.

³⁰ Interpreting a musical work is a multifaceted undertaking. We might refer to the act of physically organizing an ensemble in performance, as well as faithfully capturing and relaying spatial relationships on recording, as "spatial interpretation."

estimation, the "movement and direction of sound in space" has *always* had a function. This function is twofold: structural unification via the recurrence of related spatial gestures and the transmission of meaning. The analyst simply has not hitherto had the means of explicating this dual function via the classification and comparison of spatial gestures, and the hypothetical listener may not have always been sensitive and attentive to the function of sound in space. While spatial analysis is not dependent upon (or justified by) deliberate spatialization on the part of the composer, this document might encourage future composers to think more definitively in spatial terms—prompting them to infuse a given work with the twin functioning of sounds in space as well as to intertwine spatial gestures with events in other musical parameters. It is tempting to think of silence as the composer's "blank canvas" and empty space as the visual artist's canvas—whether that artist be a dancer, painter, sculptor, architect, etc. However, as I hope to have demonstrated by now, the composer may likewise sculpt and shape patterns of sonic activity in the actual space of musical performance.

Although the bulk of my analytic and interpretive tools are predicated on several different perceptual theories (Gestalt, auditory stream segregation, image schema, conceptual metaphor, etc.), the methodology outlined in this dissertation is not solely intended to model what *is* perceived in musical performance. Instead, I hope its application can expose what *can* be perceived in the spatial construct of a musical work. A spatial analysis may serve as an ideal reading of a musical situation and may challenge the listener's perceptual faculties. A spatial analysis may serve as a platform for systematically describing a composer's deliberate act of spatialization but is not restricted to doing so—spatial gestures may arise in musical performance purely as fortuitous phenomena. Their sheer presence disallows their being ignored and discounted. The spatial aspects of a musical work may fluctuate in their degree of perceptual

443

salience from one performance to another based on variable performance conditions, and decisions made by the performing ensemble directly affect the salience of spatial gestures. A spatial analysis can highlight the kinds of spatial gestures that will materialize during the performance of a given work of music. Spatial gestures are ubiquitous in musical works for more than one performer, but the listener might not yet be accustomed to engaging and receiving spatial shapes and motions with a musical mindset. This observation reeks of irony, given the fact that the conceptual musical space in which melodies, harmonies, and other musical events transpire is a metaphoric construct that is grounded in, and structured by, our embodied knowledge of objects and events in actual, physical space. A spatial gesture is a kinetic complex of discrete events in physical (ensemble) space (the composite shape and emergent directionality of which is forged through perceptual grouping tendencies) that is directly perceived without having to be structured by an underlying conceptual metaphor. However, the interpretation of gestural activity is shaped metaphorically: the received meaning and import of a single gesture's directionality, or of multigestural interactivity, arises from our embodied, environmental, and cultural experience with spatial relations and specific directionalities of motion. While all music is inherently spatial, perhaps one of the most significant consequences of this dissertation will be the encouragement and prompting of more deliberate "spatial listening" on the part of the perceiver—potentially enriching the musical experience. Stockhausen is correct: there is no reason for a "fixed spatial perspective of music"—be it on the part of the analyst, composer, arranger, performer, or listener.

444

BIBLIOGRAPHY

BOOKS AND ARTICLES

- Agawu, Kofi. "Ambiguity in Tonal Music: A Preliminary Study." In *Theory, Analysis, and Meaning in Music*, edited by Anthony Pope, 86–107. Cambridge: Cambridge University Press, 1994.
- Alston, William P. Philosophy of Language. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1964.
- Andersen, Richard A. "Multimodal Integration for the Representation of Space in Posterior Parietal Cortex." *Philosophical Transactions: Biological Sciences* 352.1360, What are the Parietal and Hippocampal Contributions to Spatial Cognition? (1997): 1421–1428.
- Aristotle. De Anima. In The Complete Works of Aristotle: The Revised Oxford Translation, edited by J. Barnes. Oxford: Oxford University Press, 1984.
- "Aristotle (384–322 BCE): Motion and its Place in Nature." *The Internet Encyclopedia of Philosophy*. Available at: http://www.iep.utm.edu/a/aris-mot.htm (accessed February 15, 2007).
- Arnold, Denis. *Giovanni Gabrieli and the Music of the Venetian High Renaissance*. London: Oxford University Press, 1979.

------. "The Significance of 'Cori Spezzati." Music & Letters 40.1 (1959): 4-14.

- Atlas, Allan W. Renaissance Music. New York and London: W.W. Norton & Company, 1998.
- Avis, Paul. God and the Creative Imagination: Metaphor, Symbol, and Myth in Religion and Theology. London and New York: Routledge, 1999.

Behrens, Roy R. "Art, Design and Gestalt Theory." Leonardo 31.4 (1988): 299-303.

- Benjamin, Thomas. Counterpoint in the Style of J. S. Bach. New York: Schirmer Books, 1986.
- Blauert, Jens. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA, and London: MIT Press, 1997.
- Blesser, Barry, and Linda-Ruth Salter. *Spaces Speak, Are You Listening?: Experiencing Aural Architecture.* Cambridge, MA, and London: The MIT Press, 2007.

- Boulez, Pierre. *Boulez on Music Today*. Translated by Susan Bradshaw and Richard Rodney Bennett. Cambridge, MA: Harvard University Press, 1971.
- ———. Conversations with Celestin Delierge. London: Eulenburg Books, 1976.
- Bowman, Wayne D. *Philosophical Perspectives on Music*. New York and Oxford: Oxford University Press, 1998.
- Brant, Henry. "Space as an Essential Aspect of Musical Composition." In *Contemporary Composers on Contemporary Music*, edited by Elliot Schwartz and Barney Childs, 221–242. New York: Da Capo Press, 1978.

———. "The Uses of Antiphonal Distribution and Polyphony of Tempi in Composing." *American Composer's Alliance Bulletin* 4.3 (1955): 13–15.

- Bregman, Albert S. *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA, and London: MIT Press, 1990.
- Bregman, Albert S., and J. Campbell. "Primary Auditory Stream Segregation and Perception of Order in Rapid Sequences of Tones." *Journal of Experimental Psychology* 89 (1971): 244–249.
- Brower, Candace. "A Cognitive Theory of Musical Meaning." *Journal of Music Theory* 44.2 (2000): 323–379.
- Burkholder, Peter J. *All Made of Tunes: Charles Ives and the Uses of Musical Borrowing*. New Haven: Yale University Press, 1995.
- Burton, Robert. "A Multilevel, Interdisciplinary Approach to Phenomenal Consciousness." *The Southern Journal of Philosophy* 43.4 (2005): 531–543.
 - ——. "The Human Awareness of Time: An Analysis." *Philosophy and Phenomenological Research* 36.3 (1976): 303–318.
- ------, ed. Natural and Artificial Minds. Albany: State University of New York Press, 1993.
- Busch, Regina. "On the Horizontal and Vertical Presentation of Musical Ideas on Musical Space (I)." *Tempo* 154 (1985): 2–10.
- Butler, David. *The Musician's Guide to Perception and Cognition*. New York: Schirmer Books, 1992.
- Cadwallader, Allen, and David Gagné. *Analysis of Tonal Music: A Schenkerian Approach*. New York and Oxford: Oxford University Press, 1998.

- Cage, John. *Silence: Lectures and Writings by John Cage*. Hanover, CT: Wesleyan University Press, 1961.
- Cairns, David. Liner notes to *Berloz: Te Deum*. The London Symphony Orchestra & Chorus. Colin Davis, conductor. Philips 839 790 LY. LP.
- Callender, Clifton. "Voice-leading Parsimony in the Music of Alexander Scriabin." *Journal of Music Theory* 42.2 (1998): 219-233.
- Callender, Clifton, Ian Quinn, and Dmitri Tymoczko. "Generalized Chord Spaces (Draft of 20 June 2005)." Available at: http://www.music.princeton.edu/~dmitri/chordspaces.pdf (accessed December 7, 2006).
- Calter, Paul A. *Squaring the Circle: Geometry in Art and Architecture*. Dartmouth: Key College Publishing, in press.
- Carr-Richardson, Amy. "Rotational Symmetry as a Metaphor for the Scherzo of Beethoven's Opus 131." *Indiana Theory Review* 20.1 (1999): 1–23.
- Carver, Anthony F. Cori Spezzati. Vol. 1, The Development of Sacred Polychoral Music to the Time of Schütz. Cambridge: Cambridge University Press, 1988.
- ———, ed. *Cori Spezzati*. Vol. 2, *An Anthology of Sacred Polychoral Music*. Cambridge: Cambridge University Press, 1988.
- Chadabe, Joel. *Electronic Sound: The Past and Promise of Electronic Sound*. Upper Saddle River, NJ: Prentice Hall, Inc., 1997.
- Childs, Adrian P. "Moving beyond Neo-Riemannian Triads: Exploring a Transformational Model for Seventh Chords." *Journal of Music Theory* 42.2 (1998): 181–193.
- "A Choice and an Echo." *Time Magazine*, 7 August 1964. Available at: http://www.time.com/ time/magazine/article/0,9171,871353,00.html (accessed February 16, 2007).
- Chowning, John M., John M. Grey, Loren Rush, and James A. Moorer. *Computer Simulation of Music Instruments Tones in Reverberant Environments (grant proposal for CCRMA).* Stanford: Stanford University, 1974.
- Chowning, John. "The Simulation of Moving Sound Sources." *Journal of the Audio Engineering Society* 19 (1971): 2–6.
- Clarke, Erik. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning.* Oxford: Oxford University Press, 2005.
- Cohn, Richard. "The Autonomy of Motives in Schenkerian Accounts of Tonal Music." *Music Theory Spectrum* 14.2 (1992): 150–170.

Cook, Nicholas. "Theorizing Musical Meaning." Music Theory Spectrum 23.2 (2001): 170-195.

- Cowell, Henry. "Charles E. Ives." In *American Composers on American Music: A Symposium*, edited by Henry Cowell, 128–145. Stanford: Stanford University Press, 1933.
- Cox, Arnie Walter. "The Metaphoric Logic of Musical Motion and Space." Ph.D. diss., University of Oregon, 1999.
- Cramer, Alfred. "Schoenberg's 'Klangfarbenmelodie': A Principle of Early Atonal Harmony." *Music Theory Spectrum* 24.1 (2002): 1–34.
- Danesi, Marcel, and Paul Perron. *Analyzing Cultures: An Introduction and Handbook.* Bloomington and Indianapolis: Indiana University Press, 1999.
- Driver, Jon, and Charles Spence. "Cross-Modal Links in Spatial Attention." *Philosophical Transactions: Biological Sciences* 353.1373, Brain Mechanisms of Selective Perception and Action (1998): 1319–1331.
- Eargle, John M. Music, Sound, and Technology. New York: Van Nostrand Reinhold, 1995.
- Elen, Richard. "Ambisonics: The Surround Alternative." Surround (2001): 1-4.
- Ellis, Willis D., ed. and trans. *A Source Book of Gestalt Psychology*. With an introduction by Kurt Koffka. London: Routledge and Kegan Paul Ltd., 1967.
- Emmerson, Simon. "Aural Landscape: Musical Space." Organised Sound 3.2 (1999): 135-140.
- Emmory, Karen, and Judy Reilly, eds. *Language, Gesture, and Space*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1995.
- Erickson, Robert. Sound Structure in Music. Los Angeles: University of California Press, 1975.
- Fauconnier, Gilles, and Mark Turner. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* New York: Basic Books, 2002.
- Fink, Michael. Liner notes to *Boccherini: The Guitar Quintets*. Pepe Romero and the Academy of St. Martin in the Fields Chamber Ensemble. Philips 438 769-2. Compact disc.
- Flock, A., D. Ottoson, and M. Ulfendahl. *Active Hearing*. Oxford and Tarrytown: Pergamon, 1995.
- Forte, Allen. "New Approaches to the Linear Analysis of Music." *Journal of the American Musicological Society* 41.2 (1988): 315–348.

- Friedmann, Michael. "A Methodology for the Discussion of Contour: Its Application to Schoenberg's Music." *Journal of Music Theory* 29 (1985): 223–248.
- ———. "A Response: My Contour, Their Contour." *Journal of Music Theory* 31.2 (1987): 268–274.
- Gelfand, Stanley A. *Hearing: An Introduction to Psychological and Physiological Acoustics*, 3rd ed. New York: Marcel Dekker, 1998.

Gerzon, Michael. "A year of surround-sound." Hi-Fi News 16 (1971): 1412–1420.

———. "What's Wrong with Quadraphonics?" Studio Sound 16.5 (1974): 50–56.

Gibson, James J. *The Ecological Approach to Visual Perception*. Hillsdale, NJ, and London: Lawrence Erlbaum Associates, 1986.

------. The Senses Considered as Perceptual Systems. Boston: Houghton Mifflin Company, 1966.

- Glicksohn, Joseph, and Chanita Goodblatt. "Metaphor and Gestalt: Interaction Theory Revisited." *Poetics Today* 14.1 (1993): 83–97.
- Grant, M. J. Serial Music, Serial Aesthetic. Cambridge: Cambridge University Press, 2001.
- Griffiths, Paul. The String Quartet. New York: Thames and Hudson, 1983.
- Gulick, Lawrence W., George A. Gescheider, and Robert D. Frisina. *Hearing: Physiological Acoustics, Neural Coding, and Psychoacoustics.* Oxford: Oxford University Press, 1989.
- Haimo, Ethan. "Atonality, Analysis, and the Intentional Fallacy." *Music Theory Spectrum* 18.2 (1996): 167–199.
- Hamlyn, D. W. *The Psychology of Perception: A Philosophical Examination of Gestalt Theory and Derivative Theories of Perception.* New York: The Humanities Press Inc., 1957.
- Harley, James. Xenakis: His Life in Music. New York: Routledge, 2004.
- Harley, Maria Anna. "An American in Space: Henry Brant's 'Spatial Music."" American Music 15.1 (1997): 70–92.

------. "Music of Sound and Light: Xenakis's Polytopes." Leonardo 31.1 (1998): 55-65.

———. "Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations." Ph.D. diss., McGill University, 1994.

—. "Spatiality of Sound and Stream Segregation in Twentieth-Century Instrumental Music." Organised Sound 3.2 (1999): 147–166.

- Harrison, Daniel. "Some Group Properties of Triple Counterpoint and Their Influence on Compositions by J. S. Bach." *Journal of Music Theory* 32.1 (1988): 23–49.
 - ——. "Rhetoric and Fugue: An Analytical Application." *Music Theory Spectrum*, 12.1 (1990): 1–42.
- Harrison, Jonty. "Sound, Space, Sculpture: Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion." *Organised Sound* 3.2 (1999): 117–127.
- Hasty, Christopher. "Segmentation and Process in Post-Tonal Music." *Music Theory Spectrum* 3 (1981): 54–73.
- Hatten, Robert S. Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert. Bloomington and Indianapolis: Indiana University Press, 2004.

——. *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation.* Bloomington and Indianapolis: Indiana University Press, 1994.

Hebb, Donald O. The Organization of Behavior. New York: John Wiley & Sons, Inc., 1952.

- Hughes, Bryn. "Rock's Compositional Space: The Stereo Field and Its Relation to Formal Structure." Paper presented at the annual meeting for Music Theory Southeast, Athens, Georgia, March 16–17, 2007.
- Ives, Charles. "Music and Its Future." In Source Readings in Music History, edited by Oliver Strunk, 1335–1339. New York and London: W. W. Norton & Company, 1998. Originally published in American Composers on American Music, edited by Henry Cowell, 191–198. Palo Alto: Stanford University Press, 1933.
- Johnson, Mark. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: The University of Chicago Press, 1987.
- Jourdain, Robert. *Music, the Brain, and Ecstasy: How Music Captures Our Imagination*. New York: Avon Books, 1998. Reprint, New York: Quill, 2002.
- Kahn, Douglas. "John Cage: Silence and Silencing." *The Musical Quarterly* 81.4 (1997): 556–598.
- Kant, Immanuel. *Critique of Judgment*. Translated by Werner S. Pluhar. Indianapolis: Hackett Publishing Company, 1987.
- Kepes, Gyorgy, ed. Sign, Image, Symbol. Vision + Value Series. New York: George Braziller, Inc., 1966.
- Kerman, Joseph. *The Beethoven Quartets*. New York: Knopf, 1967. Reprint, Westport, CT: Greenwood Press, 1982.
- Kinderman, William, ed. *The String Quartets of Beethoven*. Urbana and Chicago: University of Illinois Press, 2006.
- Koffka, Kurt. *Principles of Gestalt Psychology*. New York: Harcourt, Brace and Company, 1935.
- Konishi, Masakazu. "Listening with Two Ears." Scientific American Special Edition: Secrets of the Senses: How the Brain Deciphers the World Around Us 16.3 (2006): 28–35.
- Körte, A. "Kinomatoscopishe Untersuchungen." Zeitschrift für Psychologie der Sinnesorgane 72 (1915): 193–296.
- Krebs, Harald. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. New York and Oxford: Oxford University Press, 1999.
- Lakoff, George. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* Chicago: The University of Chicago Press, 1987.
- Lakoff, George, and Mark Johnson. *Metaphors We Live By*. Chicago: The University of Chicago Press, 1980.
- ———. Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. New York: Basic Books, 1999.
- Larson, Kevin. "The Science of Word Recognition: or How I Learned to Stop Worrying and Love the Bouma." Available at: http://www.microsoft.com/typography/ctfonts/ WordRecognition.aspx (accessed January 18, 2007).

Larson, Steve. "Musical Forces and Melodic Patterns." Theory and Practice 22 (1997): 55-72.

—. "Scale-Degree Function: A Theory of Expressive Meaning and Its Application to Aural-Skills Pedagogy." *Journal of Music Theory Pedagogy* 20 (1993): 69–84.

- Layer, Harold A. "Exploring Stereo Images: A Changing Awareness of Space in the Fine Arts." *Leonardo* 4.3 (1971): 233–238.
- Lerdahl, Fred. Tonal Pitch Space. Oxford: Oxford University Press, 2001.
- Leroi-Gourhan, André. *Gesture and Speech*. Translated by Anna Bostock Berger. With an introduction by Randall White. Cambridge, MA, and London: The MIT Press, 1993.

- Levine, Michael W. *Fundamentals of Sensation and Perception*, 3rd ed. Oxford: Oxford University Press, 2000.
- Lewin, David. *Generalized Musical Intervals and Transformations*. New Haven: Yale University Press, 1987.
 - ——. "Music Theory, Phenomenology, and Modes of Perception." *Music Perception* 3.4 (1986): 327–392.

—. *Musical Form and Transformation: Four Analytic Essays*. New Haven: Yale University Press, 1993.

Lidov, David. *Elements of Semiotics*. New York: St. Martin's Press, 1999.

———. "Mind and Body in Music." *Semiotica* 66.1/3 (1987): 69–97.

- Lippman, Edward A. "Music in Space: A Study in the Philosophy of Music." Ph.D. diss., Columbia University, 1952.
 - ———. "Spatial Perception and Physical Location as Factors in Music." *Acta Musicologica* 35.1 (1963): 24–34.
- Liu, JeeLoo. An Introduction to Chinese Philosophy: From Ancient Philosophy to Chinese Buddhism. Malden and Oxford: Blackwell Publishing, 2006.
- Locke, John. *An Essay Concerning Human Understanding*. Collated and Annotated by Alexander Campbell Fraser. New York: Dover Publications, Inc., 1959.
- Lombardo, Thomas J. *The Reciprocity of Perceiver and Environment: The Evolution of James J. Gibson's Ecological Psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1987.
- Luce, Duncan R. Sound & Hearing: A Conceptual Introduction. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
- Malham, David G. "Approaches to Spatialisation." Organised Sound 3.2 (1999): 167-177.
- Malham, David G., and Anthony Myatt. "3-D Sound Spatialization using Ambisonic Techniques." *Computer Music Journal* 19.2 (1995): 58–70.
- Margolis, Howard. *Patterns, Thinking, and Cognition: A Theory of Judgment*. Chicago: University of Chicago Press, 1987.
- Marvin, Elizabeth W., and Paul Laprade. "Relating Musical Contours: Extensions of a Theory for Contour." *Journal of Music Theory* 31.2 (1987): 225–67.

- Mayall, K., G. W. Humphreys, and A. Olson. "Disruption to Word or Letter Processing? The Origins of Case-mixing Effects." *Journal of Experimental Psychology: Learning, Memory, & Cognition* 23 (1997): 1275–1286.
- McCabe, Viki, and Gerald J. Balzano. *Event Cognition: An Ecological Perspective*. Hillsdale, NJ, and London: Lawrence Erlbaum Associates, 1986.
- McCusker, L. X., P. B. Gough, and R. G. Bias. "Word Recognition Inside Out and Outside In." Journal of Experimental Psychology: Human Perception and Performance 7.3 (1981): 538–551.
- McDermott, Vincent. "A Conceptual Musical Space." *The Journal of Aesthetics and Art Criticism* 30.4 (1972): 489–494.
- McNeill, David. *Hand and Mind: What Gestures Reveal about Thought*. Chicago and London: The University of Chicago Press, 1992.
 - ——, ed. *Language and Gesture*. Language, Culture, and Cognition, ed. Stephen C. Levinson, no 2. Cambridge, UK: Cambridge University Press, 2000.
- Mead, Andrew. "Bodily Hearing: Physiological Metaphors and Musical Understanding." Journal of Music Theory 43.1 (1999): 1–19.
- Messiaen, Olivier. *The Technique of My Musical Language*. Translated by John Satterfield. Paris: A. Leduc, 1956.
- Meyer, Leonard. *Emotion and Meaning in Music*. Chicago: The University of Chicago Press, 1956.
- Miller, Paul. "The Analysis of Spatial Music in Works by Karlheinz Stockhausen." Ph.D. diss. proposal, Eastman School of Music, 2005. Available at: http://www.practiceofpaul.net/research/media/dissertation proposal.pdf (accessed October 15, 2006).
- Miller, Samuel D. "Motion in Musical Texture and Aesthetic Impact." *Journal of Aesthetic Education* 17.1. (1983): 59–67.
- Mitchell, Donald. *Gustav Mahler: The Wunderhorn Years. Chronicles and Commentaries.* London: Faber & Faber, 1975.
- Mitchell, W. J. T. "Spatial Form in Literature: Toward a General Theory." *Critical Inquiry* 6.3 (1980): 539–567.

- Morgan, Robert P. "Spatial Form in Ives." In An Ives Celebration: Papers and Panels of the Charles Ives Centennial Festival-Conference, edited by H. Wiley Hitchcock and Vivian Perlis, 145–158. Music in American Life. Urbana, Chicago, and London: University of Illinois Press, 1977.
- Morris, Robert D. Composition with Pitch-Classes: A Theory of Compositional Design. New Haven and London: Yale University Press, 1987.

———. "New Directions in the Theory and Analysis of Musical Contour." *Music Theory Spectrum* 15.2 (1993), 205–228.

------. "Voice-Leading Spaces." Music Theory Spectrum 20.2 (1998): 175–208.

- Natsoulas, Thomas. "The Ecological Approach to Perception: The Place of Perceptual Content." *The American Journal of Psychology* 102.4 (1989): 443–476.
- Neisser, Ulric. "Implications for Artificial Intelligence." In *Natural and Artificial Minds*, edited by Robert G. Burton. New York: State University of New York Press, 1993, 147–164.
- Nöth, Winfried. *Handbook of Semiotics*. Bloomington and Indianapolis: University of Indiana Press, 1990.

O'Callaghan, Casey. "Hearing Recorded Sounds." Draft, Bates College, 2005.

———. "Shared Content Across Perceptual Modalities: Lessons From Cross-modal Illusions." *Electroneurobiología* 14.2 (2006): 211-224.

. "Sounds." Ph.D. diss., Princeton University, 2002.

——. Sounds: A Philosophical Theory. Oxford: Oxford University Press, in press.

- "The Orchestra: A User's Manual." Available at: http://www.mti.dmu.ac.uk/~ahugill/manual/ seating.html (accessed January 7, 2007).
- Parker, Mara. *The String Quartet, 1750–1797: Four Types of Musical Conversation.* Burlington, VT: Ashgate Publishing Limited, 2002.
- Pelli, D. G., B. Farell, and D. C. Moore. "The Remarkable Inefficiency of Word Recognition." *Nature* 423 (2003): 752–756.
- Pierce, James Smith. "Visual and Auditory Space in Baroque Rome." *The Journal of Aesthetics and Art Criticism* 18.1 (1959): 56–57.
- Plotinus. *The Enneads*: Sixth Tracate "On Beauty." Translated by Stephen Mackenna. London: Faber & Faber, 1956.

- "Psalmody, Latin." In *The New Harvard Dictionary of Music*, edited by Don Michael Randel, 664–665. Cambridge, MA, and London: The Belknap Press of Harvard University Press, 1986.
- Quinn, Ian. "Fuzzy Extensions to the Theory of Contour." *Music Theory Spectrum* 19.2 (1997): 232–263.
- Reed, Edward S. James J. Gibson and the Psychology of Perception. New Haven and London: Yale University Press, 1988.
- Reiser, Oliver L. "Time, Space and Gestalt." Philosophy of Science 1.2 (1934): 197-223.
- Rothfarb, Lee A., ed. and trans. *Ernst Kurth: Selected Writings*. Cambridge Studies in Music Theory and Analysis, ed. Ian Bent, no. 2. Cambridge and New York: Cambridge University Press, 1991.
- Schacter, Carl. "Either/Or." In *Unfoldings*, edited by Joseph N. Strauss, 121–133. Oxford: Oxford University Press, 1999.
- Schenker, Heinrich. *Harmony*. Translated by Elisabeth Mann Borgese. Chicago: University of Chicago Press, 1954.
- ———. *The Masterwork in Music: A Yearbook*. Edited by William Drabkin. Translated by Ian Bent, William Drabkin, Richard Kramer, John Rothgeb, and Heidi Siegel. Cambridge Studies in Music Theory and Analysis, ed. Ian Bent, no. 5. Cambridge and New York: Cambridge University Press, 1994.
- Schmalfeldt, Janet. "Towards a Reconciliation of Schenkerian Concepts with Traditional and Recent Theories of Form." *Music Analysis* 10.3 (1991): 233-287.
- Schoenberg, Arnold. *Fundamentals of Musical Composition*. New York: St. Martin's Press, 1967.
- ———. *The Musical Idea and the Logic, Technique, and Art of its Presentation*. Edited, translated, and with a commentary by Patricia Carpenter and Severine Neff. NewYork: Columbia University Press, 1995.
- Schulenberg, David. *Music of the Baroque*. New York and Oxford: Oxford University Press, 2001.
- Sebeok, Thomas A., ed. *A Perfusion of Signs*. Bloomington and London: Indiana University Press, 1977.

- Shinn-Cunningham, Barbara G., Antje Ihlefeld, Satyavarta, and Eric Larson. "Bottom-up and Top-down Influences on Spatial Unmasking." Acta Acustica United With Acustica 91 (2005): 967–979.
- Shinn-Cunningham, Barbara G., Norbert Kopco, and Tara J. Martin. "Localizing Nearby Sound Sources in a Classroom: Binaural Room Impulse Responses." *Journal of the Acoustical Society of America* 117.5 (2005): 3100–3115.
- Shinn-Cunningham, Barbara G., Norbert Kopco, and Scott Santarelli. "Tori of Confusion: Binaural Localization Cues for Sources Within Reach of a Listener." *Journal of the Acoustical Society of America* 107.3 (2000): 1627–1636.
- Smith, Peter H. "You Reap What You Sow: Some Instances of Rhythmic and Harmonic Ambiguity in Brahms." *Music Theory Spectrum* 28.1 (2006): 57–97.
- Snyder, Philip. "The Music of Alberto Ginastera: A Fertile Source for Guitar Transcriptions." Soundboard 31.4 (2006): 19–23.
 - ———. "The Music of Alberto Ginastera Transcribed for Guitar: A Performance Edition of Danzas Argentinas for Guitar Quartet." DMA diss., University of Georgia, 2003.
- Spence, Charles, and Jon Driver, eds. *Crossmodal Space and Crossmodal Attention*. Oxford: Oxford University Press, 2004.
- Spitzer, Michael. *Metaphor and Musical Thought*. Chicago and London: University of Chicago Press, 2004.
- Steffler, Alva William. *Symbols of the Christian Faith*. Grand Rapid, MI, and Cambridge, UK: William B. Eerdmans Publishing Company, 2002.
- Steinberg, Michael. "Notes on the Quartets." In *The Beethoven Quartet Companion*, edited by Robert Winter and Robert Martin, 143–282. Berkeley: University of California Press, 1994.
- Stockhausen, Karlheinz. "Five Revolutions Since 1950." In Contemporary Composers on Contemporary Music, expanded edition, edited by Elliot Schwartz and Barney Childs, 379–383. New York: Da Capo Press, 1998.
 - - -----. "Music in Space." Translated by Ruth Koenig. *Die Reihe* 5, English ed. (1961): 67–82. Originally published as "Musik im Raum" in *Die Reihe* 5 (1959).
 - ——. Stockhausen on Music: Lectures and Interviews Complied by Robin Maconie. New York: Marion Books, 1989.

- Stowell, Robin, ed. *The Cambridge Companion to the String Quartet*. Cambridge: Cambridge University Press, 2003.
- Straus, Joseph N. Introduction to Post-Tonal Music, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2000.
- "The Symphony: An Interactive Guide." Available at: http://library.thinkquest.org/22673/ orchestra.html (accessed January 2, 2007).
- Tarasati, Eero. *A Theory of Musical Semiotics*. Bloomington and Indianapolis: Indiana University Press, 1994.
 - ———. *Signs of Music: A Guide to Musical Semiotics*. Berlin and New York: Mouton de Gruyter, 2002.
- Ternus, Josef. "The Problem of Phenomenal Identity." In A Source Book of Gestalt Psychology, edited and translated by Willis D. Ellis, with an introduction by Kurt Koffka, 149–160. London: Routledge & Kegan Paul Ltd., 1967. Originally published as "Experimentelle Untersuchung über phänomenale Identität" in Psychol. Forsch. 7 (1926): 81–136.
- Truax, Barry. "Composition and Diffusion: Space in Sound in Space." Organised Sound 3.2 (1999): 141–146.
- Truscott, Harold. Beethoven's Late String Quartets. London: Dobson Books Ltd., 1968.
- Wagner, Michael J. Introductory Musical Acoustics. Raleigh: Contemporary Publishing Company, 1994.
- Wallace-Murphy, Tim. Cracking the Symbol Code: Revealing the Secret Heretical Messages within Church and Renaissance Art. London: Watkins Publishing, 2005.
- Warren, Richard M. *Auditory Perception: A New Analysis and Synthesis*. Cambridge and New York: Cambridge University Press, 1999.
- Watkins, Glenn. *Soundings: Music in the Twentieth Century*. New York: Schirmer Books, 1988.
- Webber, Frederick Roth. *Church Symbolism: An Explanation of the More Important Symbols of the Old and New Testament, the Primitive, the Mediaeval, and the Modern Church*, 2nd ed. Cleveland: J.H. Jansen, 1938. Reprint, Detroit: Omnigraphics, 1992.
- Weinberger, Norman R. "Music and the Brain." Scientific American Special Edition: Secrets of the Senses: How the Brain Deciphers the World Around Us 16.3 (2006): 37–43.

Wertheimer, Max. "Gestalt Theory." In *A Source Book of Gestalt Psychology*, edited and translated by Willis D. Ellis, with an introduction by Kurt Koffka, 1–11. London: Routledge & Kegan Paul Ltd., 1967.

"Laws of Organization in Perceptual Forms." In A Source Book of Gestalt Psychology, edited and translated by Willis D. Ellis, with an introduction by Kurt Koffka, 71–88.
London: Routledge & Kegan Paul Ltd., 1967. Originally published as "Untersuchungen zur Lehre von der Gestalt II" in *Psychol. Forsch.* 4 (1923): 301–350.

- ——. Productive Thinking, enlarged edition. New York: Harper & Brothers, 1959.
- "The General Theoretical Situation." In *A Source Book of Gestalt Psychology*, edited and translated by Willis D. Ellis, with an introduction by Kurt Koffka, 12–16. London: Routledge & Kegan Paul Ltd., 1967. Originally published as "Untersuchungen zur Lehre von der Gestalt I" in *Psychol. Forsch.* I (1922): 47–58.
- Winter, Robert. Compositional Origins of Beethoven's Opus 131. Studies in Musicology, No. 54. Ann Arbor: UMI Research Press, 1982.
- Winter, Robert, and Robert Martin, eds. *The Beethoven Quartet Companion*. Berkeley: University of California Press, 1994.
- Worrall, David. "Space in Sound: Sound of Space." Organised Sound 3.2 (1999): 93-99.
- Yost, William A., and George Gourevitch, eds. *Directional Hearing*. New York: Springer-Verlag, 1987.
- Zbikowski, Lawrence M. "Conceptual Models and Cross-Domain Mapping: New Perspectives on Theories of Music and Hierarchy." *Journal of Music Theory* 41.2 (1997): 193–225.
 - ——. Conceptualizing Music: Cognitive Structure, Theory, and Analysis. Oxford: Oxford University Press, 2002.
- Zvonar, Richard. "A History of Spatial Music." Available at: http://www.zvonar.com/writing/ spatial_music/History.html (accessed December 30, 2006).

SCORES

- Brant, Henry. *Desert Forests: Spatial Panoramas for Separated Orchestral Groups*, composed 1983. New York: Carl Fischer, 2000.
- Ives, Charles E. *The Unanswered Question for Chamber Orchestra or Chamber Ensemble*. New York: Southern Music Publishing Co. Inc., 1953.

Mahler, Gustav. Symphonie II. Revised version. Vienna and London: Universal Edition, 1971.

Polin, Claire. A Klockwrk Diurnal. Dorn Productions, 1977.

Stockhausen, Karlheinz. Der Kleine Harlekin für Klarinette. Kurten: Stockhausen-Verlag, 1975.

Xenakis, Iannis. Empreintes, composed 1975. Paris: Editions Salabert, 1987.

——. Persephassa pour six percussionists. Paris: Musique Contemporaine, 1970.

Zwilich, Ellen Taaffe. *Double Quartet for Strings*, composed 1984. Bryn Mawr: Merion Music, Inc., 1985.

RECORDINGS

The Arditti String Quartet. Karlheinz Stockhausen: Helikopter-Streichquartett. Montaigne MO 782097. Compact disc.

The Georgia Guitar Quartet. Mosaic. Sol Ponticello SP-015. Compact disc.

The London Symphony Orchestra & Chorus, Colin Davis, conductor. *Berloz: Te Deum*. Philips 839 790 LY. LP.

Romero, Pepe, and the Academy of St. Martin in the Fields Chamber Ensemble. *Boccherini: The Guitar Quintets*. Philips 438 769-2. Compact disc.

APPENDIX

INDEX OF SPATIAL GESTURES

This appendix explores the spatio-gestural affordances of three-, four-, and five-point ensemble spaces.¹ It shall only feature gestures in which the total event cardinality (e) equals the ensemble-point cardinality (n), and where each point is active once and only once.² This delimitation shall hold for all gestural kinds except two: oscillation gestures (which, by definition, feature the alternation of events at two distinct ensemble points) and strong closed gestures (which must begin and end at the same point in ensemble space). The following inventory of spatial gestures contains only one type of compound gesture: the sweeping gesture (SWEEP), which is classified as a strong closed gesture.³ Additionally, ensemble-segment space activity and abbreviated gestural forms (except those within the category of strong closed gesture) shall not be considered. Without such limitations, the appendix could extend indefinitely.

¹ As noted in Chapter 3, a two-point ensemble space (duo) affords little more than point shifts and adjacent-point oscillation gestures (strong or weak).

 $^{^{2}}$ Any time the event cardinality of a gesture exceeds the cardinality of the ensemble space in which the gesture occurs, one or more ensemble points must be active more than once within the gesture. Conversely, whenever the event cardinality is less than the total point cardinality, at least one ensemble point must be inactive.

³ A sweeping gesture comprises two strong open gestures with oppositional directionalities.

A.1 The Trio: Spatial Gestures in Three-Point Ensemble Space

0 1 2

Figure A-1: Integer diagram of three-point ensemble space

Oscillation Gestures:

Adjacent-point:

Strong:

$SG- 01010 ^4$
SG- 10101
SG- 12121
SG- 21212

Weak:

SG- 01	01
SG- 10	010
SG- 12	212
SG- 21	21

Distant-point:

Strong:

SG-	02020
SG-	20202

Weak:

SG-|020...2| SG-|202...0|

Open Gestures:

Strong:

⁴ The ellipses between integer points indicate that the oscillation gesture may continue indefinitely before terminating at the final integer point in the SG-label.

Weak:

*No three-event weak open gestures are possible in three-point ensemble space.

Closed Gestures:

Strong:

SWEEP Gestures: e = [n + (n - 1)]e = [3 + (3 - 1)]e = 5

> SG-|01210| (LEFT SWEEP, perfect) SG-|21012| (RIGHT SWEEP, perfect)

Abbreviated SWEEP Gestures: $e = [n + (n - 2)]^5$ e = [3 + (3 - 2)] $e = \underline{4}$

LEFT SWEEP (four-event abbr.)

SG-|0120| SG-|0210|

RIGHT SWEEP (four-event abbr.)

SG-|2102| SG-|2012|

Other (four-event, non-SWEEP)

SG-|1021| SG-|1201|

Weak:

SG-|021| (Clockwise inward spiral) SG-|102| (Counterclockwise outward spiral) SG-|120| (Clockwise outward spiral) SG-|201| (Counterclockwise inward spiral)

⁵ For present purposes, all abbreviated SWEEPs shall retain the boundary points of the total ensemble (which also serve as the spatial extremes of the gestures). All abbreviated forms shall be exclusively bi-directional and contain *at least* one instance of disjunct motion. The abbreviations result from the process of *deletion*, whereby an event point interior to the SG-label, and likewise occurring at an interior ensemble point, is omitted (as described in Chapter 4). Sweeping gestures unfolding within ensemble-segment space are not considered in this catalog of spatial gestures. This comment holds for the discussions of four- and five-point ensemble spaces below.

A.2 The Quartet: Spatial Gestures in Four-Point Ensemble Space

0 1 2 3

Figure A-2: Integer diagram of four-point ensemble space

Oscillation Gestures:

Adjacent-point:

Strong:

SG- 01010
SG- 10101
SG- 12121
SG- 21212
SG- 23232
SG- 32323

Weak:

SG- 0101
SG- 1010
SG- 1212
SG- 2121
SG- 2323
SG- 3232

Distant-point:

Strong:

SG- 02020
SG- 20202
SG- 13131
SG- 31313
SG- 03030
SG- 30303

Weak:

SG- 0202
SG- 2020
SG- 1313
SG- 3131

Open Gestures:

Strong:

Weak:

Closed Gestures:

Strong:

SWEEP Gestures: e = [n + (n - 1)]e = [4 + (4 - 1)]e = 7

> SG-|0123210| (LEFT SWEEP, perfect) SG-|3210123| (RIGHT SWEEP, perfect)

Abbreviated SWEEP Gestures: e = [n + (n - 2)]e = [4 + (4 - 2)] $e = \underline{6}$

LEFT SWEEP (six-event abbr.)⁶

SG-|023210| SG-|013210| SG-|012310| SG-|012320|

RIGHT SWEEP (six-event abbr.)

SG-|310123| SG-|320123| SG-|321023| SG-|321013|

⁶ In four-point ensemble space, a six-event abbreviated SWEEP gesture (in which total ensemble boundary points are retained) contains one (and only one) instance of disjunct motion.

Abbreviated SWEEP Gestures: e = [n + (n - 3)]e = [4 + (4 - 3)]e = 5

LEFT SWEEP (five-event abbr.)⁷

SG-|03210| SG-|02310| SG-|02320| SG-|03210| SG-|01310| SG-|01320| SG-|01310| SG-|01230| SG-|01230| SG-|01320| SG-|01230|

RIGHT SWEEP (five-event abbr.)

SG-|30123| SG-|31023| SG-|30123| SG-|32023| SG-|32023| SG-|31023| SG-|31023| SG-|32023| SG-|32103| SG-|32013| SG-|32013|

Other (five-event, non-SWEEP)

SG-|02130| SG-|03120| SG-|10231| SG-|10321| SG-|12031| SG-|12301| SG-|13021|

⁷ In four-point ensemble space, a five-event abbreviated SWEEP contains either one instance of disjunct motion in which two adjacent points are skipped-over (within a single, unidirectional point shift) or two instances of disjunct motion in which only one point is skipped-over (each within a different and opposite-directed point shift).

SG- 13201
SG- 20132
SG- 20312
SG- 21032
SG- 21302
SG- 23012
SG- 23102
SG- 30213
SG- 31203

Weak:

SG-	0132	
SG-	0231	
SG-	0312	(Clockwise inward spiral)
SG-	0321	*
SG-	1023	
SG-	1032	*
SG-	1203	(Clockwise outward spiral)
SG-	1230	*
SG-	1302	
SG-	1320	
SG-	2013	
SG-	2031	
SG-	2103	*
SG-	2130	(Counterclockwise outward spiral)
SG-	2301	*
SG-	2310	
SG-	3012	* ⁸
SG-	3021	(Counterclockwise inward spiral)
SG-	3102	
SG-	3201	

⁸ Although classified as weak closed gestures, SG-|1230|, SG-|2301|, SG-|3012|, SG-|2103|, SG-|1032|, and SG-|0321| (all indicated with an asterisk in the weak closed gesture column) are *displaced panning gestures*. SG-|0123| (L-R PAN) is a strong open gesture. When subjected to rotational slot displacement, wherein each event point is "displaced" one event slot to the right and point 3 (originally at event slot 4) is allowed to "wrap around" to event slot 1, SG-|0123| becomes SG-|3012|. Replicating this process twice will first yield SG-|2301| and then SG-|3012|. One final rotation reproduces the original SG-|0123|. The resultant displaced panning gestures are somewhat circular in nature, in the same way that SG-|1203| is spiral (clockwise, outward) in nature. Consecutive rotational slot displacement operations likewise transform SG-|3210| (R-L PAN) into SG-|0321|, SG-|1032|, and SG-|2103|.

A.3 The Quintet: Spatial Gestures in Five-Point Ensemble Space

0 1 2 3 4

Figure A-3: Integer diagram of five-point ensemble space

Oscillation Gestures:

Adjacent-point:

Strong:

SG- 01010
SG- 10101
SG- 12121
SG- 21212
SG- 23232
SG- 32323
SG- 34343
SG- 43434

Weak:

SG- 0101
SG- 1010
SG- 1212
SG- 2121
SG- 2323
SG- 3232
SG- 3434
SG- 4343

Distant-point:

Strong:

SG- 02020
SG- 20202
SG- 13131
SG- 31313
SG- 24242
SG- 42424
SG- 03030
SG- 30303
SG- 14141

SG-	4141.	4
SG-	0404.	0
SG-	4040.	4

Weak:

SG- 0202
SG- 2020
SG- 1313
SG- 3131
SG- 2424
SG- 4242
SG- 0303
SG- 3030
SG- 1414
SG- 4141
SG- 0404
SG- 4040

Open Gestures:

Strong:

Weak:

SG- 01324
SG- 02134
SG- 02314
SG- 03124
SG- 03214
SG- 42310
SG- 43120
SG- 41320
SG- 42130
SG- 41230

Closed Gestures:

Strong:

SWEEP Gestures:
$$e = [n + (n - 1)]$$

 $e = [5 + (5 - 1)]$
 $e = 9$

SG-|012343210| (LEFT SWEEP, perfect) SG-|432101234| (RIGHT SWEEP, perfect)

Abbreviated SWEEP Gestures: e = [n + (n - 2)]e = [5 + (5 - 2)]e = 8

LEFT SWEEP (eight-event abbr.)

SG-|02343210| SG-|01343210| SG-|01243210| SG-|01234210| SG-|01234310| SG-|01234320|

RIGHT SWEEP (eight-event abbr.)

SG-|42101234| SG-|43101234| SG-|43201234| SG-|43210234| SG-|43210134| SG-|43210124|

Abbreviated SWEEP Gestures: e = [n + (n - 3)]e = [5 + (5 - 3)] $e = \underline{7}$

LEFT SWEEP (seven-event abbr.)

SG-|0343210| SG-|0243210| SG-|0234210| SG-|0234310| SG-|0234320| SG-|0143210| SG-|0134210| SG-|0134310| SG-|0124310| SG-|0124310| SG-|0124320| SG-|0123410| SG-|0123420| SG-|0123430| RIGHT SWEEP (seven-event abbr.)

SG-|4101234| SG-|4201234| SG-|4210234| SG-|4210134| SG-|4210124| SG-|4301234| SG-|4310134| SG-|4310134| SG-|4320134| SG-|4320134| SG-|4320124| SG-|4321024| SG-|4321014|

Abbreviated SWEEP Gestures: e = [n + (n - 4)]e = [5 + (5 - 4)] $e = \underline{6}$

LEFT SWEEP (six-event abbr.)

SG-|043210| SG-|034210| SG-|034310| SG-|034320| SG-|024210| SG-|024310| SG-|024320| SG-|023410| SG-|023420| SG-|023430| SG-|014210| SG-|014310| SG-|014320| SG-|013410| SG-|013420| SG-|013430| SG-|012410| SG-|012420| SG-|012430| SG-|012340|

RIGHT SWEEP (six-event abbr.)

SG- 401234
SG- 410234
SG- 410134
SG- 410124
SG- 420234
SG- 420134
SG- 420124
SG- 421034
SG-421024
SG- 421014
SG- 430234
SG- 430134
SG- 430124
SG-431034
SG- 431024
SG- 432014
SG- 432104

Other (six-event, non-SWEEP)

SG-|013240| SG-|014230| SG-|021340| SG-|021430| SG-|023140| SG-|024130| SG-|031240| SG-|031420| SG-|032140| SG-|032410| SG-|034120| SG-|041230| SG-|041320| SG-|042130| SG-|042310| SG-|043120| SG-|102341| SG-|102431| SG-103241 SG-|103421| SG-|104231| SG-|104321| SG-|120341| SG-[120431]

SG-|123041| SG-|123401| SG-|124031| SG-|124301| SG-|130241| SG-|130421| SG-|132041| SG-|132401| SG-|134021| SG-|134201| SG-|140231| SG-|140321| SG-|142031| SG-|142301| SG-|143021| SG-|143201| SG-|201342| SG-|201432| SG-|203142| SG-|203412| SG-|204132| SG-|204312| SG-|210342| SG-|210432| SG-|213042| SG-|213402| SG-|214032| SG-|214302| SG-|230142| SG-|230412| SG-|231042| SG-|231402| SG-|234012| SG-|234102| SG-|240132| SG-|240312| SG-|241032| SG-|241302| SG-|243012| SG-|243102| SG-|301243| SG-|301423| SG-|302143| SG-|302413| SG-|304123| SG-|304213| SG-|310243| SG-|310423| SG-|312043| SG-|312403| SG-|314023| SG-|314203| SG-|320143| SG-|320413| SG-|321043| SG-|321403| SG-|324013| SG-|324103| SG-|340123| SG-|340213| SG-|341023| SG-|341203| SG-|342013| SG-|342103| SG-|401324| SG-|402134| SG-|402314| SG-|403124| SG-|403214| SG-|410324| SG-|412034| SG-|412304| SG-|413024| SG-|413204| SG-|420314| SG-|421304| SG-|423014| SG-|423104| SG-|430214| SG-|431204|

Weak:

SG- 01243
SG- 01324
SG- 01342
SG- 01423
SG- 01432
SG- 02134
SG- 02143
SG- 02314

SG- 02341	
SG- 02413	
SG- 02431	
SG- 03124	
SG- 03142	
SG- 03214	
SG- 03241	
SG- 03412	
SG- 03421	
SG- 04123	
SG- 04132	(Clockwise inward spiral)
SG- 04213	
SG- 04231	
SG- 04312	
SG- 04321	(Displaced panning gesture)
SG- 10234	
SG- 10243	
SG- 10324	
SG- 10342	
SG- 10423	
SG- 10432	(Displaced panning gesture)
SG- 12034	
SG- 12043	
SG- 12304	
SG- 12340	(Displaced panning gesture)
SG- 12403	
SG- 12430	
SG- 13024	
SG- 13042	
SG- 13204	
SG- 13240	
SG- 13402	
SG- 13420	
SG- 14023	
SG- 14032	
SG- 14203	
SG- 14230	
SG- 14302	
SG- 14320	
SG- 20134	
SG- 20143	
SG- 20314	
SG- 20341	
SG- 20413	
SG- 20431	
SG- 21034	

SG-21043 (Displaced panning gesture) SG-21304 (Counterclockwise outward spiral) SG-|21340| SG-|21403| SG-|21430| SG-|23014| SG-|23041| SG-|23104| SG-23140 (Clockwise outward spiral) SG-23401 (Displaced panning gesture) SG-|23410| SG-|24013| SG-|24031| SG-|24103| SG-|24130| SG-|24301 SG-|24310| SG-|30124| SG-|30142| SG-|30214| SG-|30241| SG-|30412| SG-|30421| SG-|31024| SG-|31042| SG-|31204| SG-|31240| SG-|31402| SG-|31420| SG-|32014| SG-|32041| SG-32104 (Displaced panning gesture) SG-|32140| SG-|32401| SG-|32410| SG-|34012| (Displaced panning gesture) SG-|34021| SG-|34102| SG-|34120| SG-|34201| SG-|34210| SG-40123 (Displaced panning gesture) SG-|40132| SG-|40213| SG-|40231| SG-|40312| (Counterclockwise inward spiral)

SG- 40321
SG- 41023
SG- 41032
SG- 41203
SG- 41230
SG- 41302
SG- 41320
SG- 42013
SG- 42031
SG- 42103
SG- 42130
SG- 42301
SG- 42310
SG- 43012
SG- 43021
SG- 43102
SG- 43120
SG- 43201

GLOSSARY

- **Abbreviation**: Any process by which the event cardinality of a gesture is reduced by at least one. Abbreviation often involves "shrinking" a gesture from the total ensemble space to a segment of ensemble space or introducing disjunct spatial motion.
- Acceleration: A temporal modification affecting oscillation gestures in which the rate of alternation between two points increases from one oscillation gesture to another. The two gestures may unfold over the same approximate amount of time, but the accelerated gestural form would contain more event points since it is oscillating at a faster rate. This type of acceleration is *intergestural*, which differs from the notion of *intragestural* acceleration.
- Accumulative Gesture: A gesture in which component sound events are continued or sustained following their entrances and allowed to amass. An accumulative gesture generally entails an expansion of the sonic space and often precedes an ALL PLAY section.
- Addition: A type of elaboration whereby an event at an interior ensemble point is added to the internal portion of a spatial gesture (between the temporal extremes of the gesture's SG-label).
- Adjacent-Point Oscillation Gesture: An oscillation gesture occurring between two contiguous points in ensemble space.
- Affordance: A term coined by ecological psychologist James J. Gibson to capture the complementarity (reciprocal relationship) between the perceiver and the environment (or objects in the environment). An affordance is based mutually upon the objective properties of an object and the needs and capabilities of the perceiver; affordances account for the prospective interactivity between an organism and an object. A single object may carry multiple affordances. The term is often applied more generally to describe the potential provisions or outcomes of a thing or situation, as in: "the gestural *affordances* of a five-point ensemble space."
- ALL PLAY: A portion of spatio-musical activity in which all ensemble points are actively producing sound.
- **Ambiguous Polystreaming**: When two or more feasible explanations of multigestural activity co-exist. If at least one of the gestures engaged in polystreaming is not likewise an auditory stream, the nature of the polystreaming might prove difficult to discern.

- **Apparent Motion**: The illusion of motion that occurs without the actual displacement of an object or event. The consecutive presentation of stationary stimuli often gives rise to apparent motion. Flashing lights may produce this effect in visual perception (stroboscopic movement), and a point shift (the temporal succession of sounds at two different points in space) may induce apparent motion in auditory perception.
- Apparent Polystreaming: When the temporal organization of a complex spatial gesture particularly its rate of unfolding—provides for the discernment of multiple, distinct streams. For example, under certain conditions (such as a rapid rate of presentation), SG-|031425| might be perceived (and analyzed) as the interwoven co-occurrence (echo streaming) of SG-|012| and SG-|345|.
- **Area**: Essentially an inflated ensemble point. An area exists when multiple performers play the same musical part on the same instrument and are locationally proximate in the ensemble space, such as the cello section of a symphony orchestra.
- Audience Space: The area(s) within the performance space potentially occupied by the audience.
- Auditory Scene: The analog in audition to the visual field in vision. The auditory scene comprises multiple sonic events, many of which are grouped together into distinct auditory streams.
- Auditory Sequence: The consecutive presentation of dissimilar sound events.
- **Auditory Stream**: The perceptual grouping of similar sonic events occurring over time. Multiple sounds are *integrated* to form a stream, and multiple streams are *segregated* from one another (within the auditory scene).
- Augmentation: A temporal modification whereby the time span of a spatial gesture is increased.
- **Binaural Fusion**: The process by which the nonidentical sound signals (produced by the same source) reaching the two ears at separate times are "fused" into a single auditory image.
- **Canonic Streaming**: A type of consonant polystreaming in which the same type of gesture occurs more than once in overlapping temporal succession.
- **Closed Gesture**: A spatial gesture that begins and ends at any point in ensemble space provided it does not begin at one extreme of ensemble space *and* conclude at the other extreme (this type of gestural activity would produce an open gesture). A closed gesture may *either* begin *or* end at one of the borders of the ensemble space. In addition, closed gestures may begin and end at the same ensemble point, and this point could be either of the spatial extremes of the ensemble space. A closed gesture will always feature motion that is not unidirectional, and unless one or more ensemble points are active more than once, a closed gesture will feature disjunct motion.

- **Complex Gesture**: A gesture in which nonadjacent event points within the SG-label (i.e., events that do not occur in direct temporal succession) are grouped together to form smaller, component gestures that share in a polystreaming relationship. For example, under certain conditions (described in section 4.2), complex gesture SG-|031425| might partition into SG-|012| and SG-|345|. The subgestures of a complex gesture may or may not be discrete in space.
- **Composite Gesture**: A gesture formed by events in ensemble space and those in another musical parameter, such as pitch space. For example, a left-to-right panning gesture conflated with a progressive rise in registral (pitch) space produces a "bottom-left-to-upper-right slanted (or, diagonal) gesture." This gesture is a "composite" of spatial and pitch activity.
- **Compound Gesture**: A gesture in which two or more small gestures (subgestures) occur in direct temporal succession to form a larger gesture. Unlike a complex gesture, the subgestures of a compound gesture are discrete in time (as well as within the SG-label).
- **Compression**: Occurs when the space that a gesture spans is decreased and the event-point cardinality of the original and the compressed gesture are the same.
- **Conceptual Metaphor**: A cognitive process by which something abstract and/or unfamiliar (the *target domain*) is structured and understood in terms of something concrete and/or familiar (the *source domain*). The mapping from source domain to target domain is termed *cross-domain mapping*.
- **Consonant Streaming**: A type of polystreaming that involves the coincident occurrence of two or more gestures of the same categorical kind and with the same directionality.
- **Contrary Streaming**: A type of polystreaming in which two gestures are categorically related but exhibit opposing directionalities.
- **Convergent Mirrored Streaming**: A type of contrary streaming in which two gestures (each in a separate segment of ensemble space) approach one another in ensemble space.
- **Cross-Domain Mapping**: The cognitive mapping from source domain to target domain that yields a conceptual metaphor.
- **Crossed Streaming**: A type of polystreaming in which related gestures with antithetical directionalities "cross" one another within the total ensemble space or within the same segment of ensemble space.
- **Crossmodal Binding**: The interaction between two or more modes of sensation (such as vision and audition) that forges a mental percept of an object or event.

- **Cyclic Panning Gesture**: A compound gesture in which two or more panning gestures are presented in direct temporal succession. Although the subgestures are all members of the strong open gesture category, the gesture as a whole is classified as a weak open gesture.
- **Deceleration**: A temporal modification affecting oscillation gestures in which the rate of alternation between two points decreases from one oscillation gesture to another. The two gestures may unfold over the same approximate amount of time, but the decelerated gestural form would contain less event points since it is oscillating at a slower rate. This type of deceleration is *intergestural*, which contrasts the notion of *intragestural* deceleration.
- Deflation: A type of abbreviation in which both of a gesture's spatial extremes are eliminated.
- **Deletion**: A type of abbreviation in which one or more events at interior event points are omitted from a gesture.
- **Diffraction**: The spreading out of a sound wave into spaces that are not in a direct path from the sound source. Low frequency sounds diffract around the head of the listener, causing an interaural time difference (ITD).
- **Diffusion**: The weakening of a sound's waves (the dissipation of sound energy) as the waves travel through a medium. The term likewise refers to the process by which a sound's energy spreads and disperses within a given space—often making the sound difficult to localize.

Diminution: A temporal modification in which the time span of a spatial gesture is decreased.

Directionality Change: When the directionality of motion within a spatial gesture changes.

Direct Signal: The sound wave approaching the listener directly from the sound source.

- **Displaced Panning Gesture**: A specific type of weak closed gesture. Displaced panning gestures are related to perfect panning gestures via the transformative procedure of rotational slot displacement.
- **Dissipative Gesture**: When an ALL PLAY scenario gradually disassembles into a smaller segment of the ensemble space or to a single point, which can occur as instruments successively cease their musical activity. A dissipative gesture is the opposite of an accumulative gesture and entails the dissolution of the sonic space.
- **Dissonant Streaming**: A type of polystreaming that occurs when two or more categorically different spatial gestures unfold either at the same time within different segments of the ensemble space or in overlapping succession within the total ensemble space. The "spatial dissonance" results from the categorical heterogeneity between the concomitant gestures.

- **Distant-Point Oscillation Gesture**: An oscillation gesture occurring between two nonadjacent points in ensemble space.
- **Divergent Mirrored Streaming**: A type of contrary streaming in which two gestures (each in a separate segment of ensemble space) recede from one another in ensemble space.
- **Doppler Effect**: The change in pitch brought about as a sound-emitting source approaches and subsequently passes a listener.
- **Double Inversion**: A type of inversion in which a gesture occurring within an ensemble-space segment is subjected to both type-1 and type-2 inversional operations.
- **Early Reflections**: Reflected sound waves that arrive within approximately 30 msec of the direct signal.
- Echo Stream: The latter-occurring gesture in an instance of echo streaming.
- **Echo Streaming**: A type of consonant polystreaming in which two identical gestures unfolding in separate regions of ensemble space are temporally staggered in their joint presentation.
- **Elaboration**: Any alteration that increases the event cardinality of a gesture by at least one. Elaboration often entails enlarging a gesture from a segment of ensemble space to either a larger ensemble-space segment or to the total ensemble space. Elaboration may likewise cause a previously disjunct gesture to recur in a more (or completely) conjunct form via the "filling-in" of interior gaps (spatial disjunctions). Elaboration is the procedural opposite of abbreviation.
- **Embodied Knowledge**: The understanding that results from repeated experience with our body's sensory apparatuses and motor programs within our physical and cultural environments.
- **Emergence**: The immediate and automatic perception of a group of objects or events as a unified whole.
- **Ensemble Point**: The discrete spatial position (within the total ensemble space) of a given ensemble member. A quartet comprises four distinct ensemble points and may thereby be referred to as a *four-point ensemble space*.
- **Ensemble Space**: The area within the performance space occupied by all of the performers of an ensemble. The ensemble space is the region from which all musical sounds ensue.
- **Ensemble-Space Segment**: A portion of the total ensemble space comprising contiguous ensemble points.
- **Event Point**: The term for each integer in an SG-label representing an event at a particular point in ensemble space.

- **Event Slot**: The (temporally) ordered slots in an SG-label, each of which contains a single event point.
- **Expansion**: When the space traversed by a gesture is increased and the event-point cardinality of the original gesture is preserved in the latter (expanded form) gesture.
- **Extension**: A type of elaboration in which a new spatial extreme is introduced to a gesture. A gesture may be left-extended or right-extended.
- **Externalization**: The recognition that a sound—as well as the source of that sound—is distally located relative to the perceiver.
- **Fused Oscillation Gesture**: An oscillation gesture in which a group of fused points alternates with another group of fused points. The fusion of points may or may not involve adjacent ensemble points (which forms a superpoint).
- **Gestalt**: The German term referring to the "whole" that emerges from the structural and/or dynamic interrelationships among various parts. The term may be translated as "figure," "pattern," "form," or "shape."
- **Gestaltqualität**: The prevailing "form-quality" of a whole that is absent when the whole is reduced to a collection of parts.
- **Hyperpoint**: When an ensemble space comprises two or more spatially separated ensembles, and each individual ensemble features multiple musical parts and/or instruments. The distinct ensembles may each function as a hyperpoint in the global context of the total ensemble space.
- **Image Schema**: Described by Mark Johnson as "a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience."¹
- Inflation: A type of elaboration in which a spatial gesture acquires two new spatial extremes.
- **Interaural Level Difference (ILD)**: The difference of intensity between the sound waves (produced by the same source) reaching each ear. ILDs aid in the localization of high frequency sounds.
- **Interaural Time Difference (ITD)**: The slight discrepancy of timing (phase) between the sound waves (produced by the same source) reaching each ear. ITDs aid in the localization of low frequency sounds.
- **Intercategorical Alteration**: A spatial alteration that changes a spatial gesture's category membership.

¹ Johnson, xiv.

- **Intracategorical Alteration**: A spatial alteration that preserves a spatial gesture's categorial affiliation. Includes intracategorical transformational operations.
- **Intragestural Acceleration**: When a spatial gesture's rate of unfolding gradually increases (as the onset-to-onset interval between consecutive events progressively decreases).
- **Intragestural Deceleration**: When a spatial gesture's rate of unfolding gradually decreases (as the onset-to-onset interval between consecutive events progressively increases).
- **Invariance**: The phenomenon in which a single object is recognized irrespective of its spatial (and temporal) orientation due to unchanging qualities. Invariance likewise enables similar items to be viewed as related by virtue of their commensurate properties.
- **Inversion**: An intracategorical transformational operation that maps ensemble points into one another across some axis of symmetry within the ensemble space. For example, SG-|04132| becomes SG-|40312| when reflected across an inversional axis abiding at (and equally dividing) the total ensemble space.
- Late reflections: The reflected signals that reach the listener quite late, after the sound wave has reflected multiple times and traveled greater distances. Late reflections are often perceived as echoes or reverberation, and they potentially degrade the intelligibility of a sound.

Lateral Reflections: Early reflections approaching the listener from his/her side.

- **Lateralization**: When listening to sounds over headphones, the illusory perception that the sounds are located somewhere along a lateral plane instead the head that extends from one ear to the other (along the binaural axis).
- Law of Closure: A Gestalt principle of Prägnanz stating that our perceptual preference is for closed, completed figures. A closed figure is judged to be "well-formed."
- Law of Common Fate: A Gestalt principle of Prägnanz that applies to grouped objects that are changing together in time. Elements moving together in the same direction, or otherwise changing in time concomitantly, group together in perception.
- Law of Continuity: A Gestalt principle of Prägnanz stating that figures and grouped objects that continue in one direction, without an abrupt shift in directionality, are strong and perceptually salient. The notion of continuity may pertain either to the sustainment of a sound particular's sonic attributes or to the continual, uninterrupted change of some property (or properties) of a sound.
- Law of Proximity: A Gestalt grouping principle stating that elements that are close together in time and/or space are readily grouped together in perception.

- Law of Similarity: A Gestalt grouping principle stating that similar or identical elements are readily grouped together in perception.
- Law of Symmetry: A Gestalt principle of Prägnanz stating that symmetrical, balanced objects exhibit good form.
- Lead Stream: In the context of canonic streaming, the primary spatial gesture that "leads" the spatio-canonic activity.
- **Localization**: The precise identification of a sound's location (azimuth, height, distance, and potential motion of sound source) in the environment.
- **Migratory Gesture**: A gesture in which constituent sonic events are not sustained (i.e., not allowed to accumulate).
- **Mirrored Streaming**: A type of contrary streaming in which two gestures unfolding in separate regions of the ensemble space are related by either retrograde, inversion, or retrograde-inversion. The mirrored streaming may be convergent or divergent.
- **Multistability**: Results from the fact that an object or image cannot be perceived as two entirely different things at the same time. A line, for example, may belong to only one figure at any given time (due to the principal of exclusive allocation), but its belongingness may shift from one figure to another—often giving rise to figure-ground reversals.
- **Musical Gesture**: Described by Robert Hatten as the "significant energetic shaping of sound through time."²
- Musical Space: The conceptual space constructed mentally during the experience of music.
- **Nudging**: When an entire gesture is shifted within the ensemble, always from one segment of ensemble space to another. A gesture may be left-nudged or right-nudged. For example, SG-|012| is right-nudged (+2) to become SG-|234| in a five-point ensemble space.
- **Open Gesture**: A spatial gesture that begins at one extreme of ensemble space and terminates at the other spatial extreme.
- Oscillation Gesture: The perpetual alternation of events at two different ensemble points.
- **Panning Gesture (PAN)**: Another term for a strong open gesture. Panning gestures are distinguished by their directionalities and may be qualified as either a left-to-right panning gesture (L-R PAN) or a right-to-left panning gesture (R-L PAN).
- **Parallel Streaming**: A type of consonant polystreaming in which the exact same gesture takes place in two or more regions of ensemble space at the same time (in point-for-point simultaneity).

² Hatten, Interpreting Musical Gestures, Topics, and Tropes, 95.

- **Partial Open Gesture**: A gesture that is classified as open in a segment of ensemble space but not in the total ensemble space. The gesture would begin and end at the two spatial extremes of the ensemble-segment space.
- **Partial Rotational Slot Displacement**: Occurs when only a portion of a gesture's SG-label undergoes rotational slot displacement.
- **Perfect Panning Gesture**: A perfectly conjunct panning gesture in which every ensemble point is active once and only once. (A panning gesture involving disjunct motion—in which one or more ensemble points is "skipped"—is imperfect.)
- Perfect Sweeping Gesture: A perfectly conjunct sweeping gesture, such as SG-|0123210|.
- **Performance Space**: The room in which a musical performance occurs and in which the sounds are, for the most part, entirely contained. The performance space contains the ensemble space and the audience space.
- **Pitch Space**: A component of a conceptual musical space wherein pitches are perceived to abide at varying heights.
- Plane Shift: The smallest unit of motion between two depth planes.
- **Point Exchange**: Occurs when any two event slots (within an SG-label) "exchange" event-point integers. The two event slots may be adjacent or nonadjacent to one another within the SG-label.
- **Point Fusion**: When two or more events occur simultaneously at different points in the ensemble space. The points are fused together in time. When two adjacent points are fused, a *superpoint* is formed. In the context of polystreaming, point fusion refers to the merging of two contemporaneous gestures into a single gesture (the two gestures fuse together at a specific point in the ensemble space).
- **Point Shift**: The smallest unit of motion within a spatial gesture, consisting of two temporally successive events—each occurring at a different point in ensemble space. A point shift may be *conjunct* (occurring between contiguous points) or *disjunct* (occurring between nonadjacent points).
- **Point-Shift Exchange**: Occurs when any two, non-overlapping point-shift slots (within an SG-label) "exchange" their component point shifts. The two point-shift slots may be adjacent or nonadjacent to one another within the SG-label, but they may not overlap one another within the label.
- **Point-Shift Reversal**: When a point exchange occurs between neighboring event slots (within an SG-label).

- **Point-Shift Slot**: The (temporally) ordered and overlapping slots in an SG-label, each of which contains a two-event point shift.
- **Point split**: A type of elaboration wherein two or more previously fused points are separated. In the context of polystreaming, point split refers to the bifurcation of a single gesture into two distinct and concomitant gestures (the two gestures "split-off" from one another at a specific point in ensemble space).
- Polystreaming: Two or more gesture unfolding simultaneously in a given ensemble space.
- **Prägnanz**: Refers to the "good form" of a gestalt. Includes the Law of Closure, the Law of Continuity, the Law of Common Fate, and the Law of Symmetry.
- Principal Gestural Form: The most dominant spatial gesture of a particular work.
- **Reflection**: When sound waves are reflected from a surface instead of being absorbed by (and transmitted through) that surface.
- **Refraction**: The bending of a wave front as it penetrates a medium of varying density.
- **Reification**: Reflects the constructive, generative aspect of perception and often involves the perceptual process of "connecting the dots" or "filling in the missing piece(s)." Reification occurs when the percept contains more spatial information than the distal, sensory stimulus (the object of perception) actually contains.
- **Retrograde-Inversion**: An intracategorical transformational operation in which a spatial gesture is subjected to both retrogression and inversion (either type-1 or type-2). For example, SG-|04132| becomes SG-|21304|.
- **Retrogression**: An intracategorical transformational operation in which the sequential ordering of event points composing a gesture is reversed. For example, SG-|04132| becomes SG-|23140|.
- **Rotational Slot Displacement**: When the event-point contents of every event slot within an SGlabel are displaced to the right by a certain number of slots, and the rightmost event point "wraps around" to become the leftmost event point. For example, displacing SG-|0123| (perfect panning gesture) one degree to the right produces SG-|3012| (a displaced panning gesture).
- Sequential Integration: Described by Albert Bregman as "the process that connects events that have arisen at different times from the same source."³
- Score Space-to-Ensemble Space Translation: The way in which the organization of parts (staves) in the score relates to the arrangement of points in ensemble space. In a "perfect" score space-to-ensemble space translation, there is a 1:1 correspondence

³ Bregman, 30.
between the vertical ordering of staves in the score and the horizontal series of points in the ensemble space. For example, it is common for the top-to-bottom ordering of parts in a string-quartet score to translate perfectly to the left-to-right sequence of ensemble points: Violin 1 (first staff/point 0), Violin 2 (second staff/point 1), Viola (third staff/ point 2), Cello (fourth staff/point 3). Under such conditions, there is often a correlation between an event in registral (pitch) space and an event in score space: an ascent in registral space that encompasses all four instruments of the string quartet would appear as a rising figure in the staff system(s) of the score. Such a rising figure (in both pitch and score space) would have a spatio-gestural counterpart exhibiting a right-to-left directionality (R-L PAN). A complex musical event need not traverse registral space to be dispersed through both the score space and, consequently, the ensemble space. Figure G-1 illustrates a perfect score space).



Figure G-1: Perfect score space-to-ensemble space translation (top-to-bottom = left-to-right)

- **SG-label**: Integer notation that depicts the temporal ordering of spatially discrete events constituting a spatial gesture.
- **Simultaneous integration**: The process that "takes acoustic inputs occur[ing] at the same time but at different places in the [frequency] spectrum or in space, and treats them as properties of a single sound."⁴
- **Sonic Space**: The portion of the ensemble space actively engaged in the production of sound at any given moment. The sonic space is nonexistent when none of the ensemble points are producing sound and is equivalent to the ensemble space when all ensemble members are producing sound simultaneously.

⁴ Ibid., 31.

- **Sound**: Described by Casey O'Callaghan as "the event of an object or interacting bodies disturbing a surrounding medium in a wave-like manner."⁵ The vibrating object is the cause of the sound, and the waves that follow from this event are the effects of the sound.
- Sound Event (or Sonic Event): Any sound produced at an ensemble point.
- **Source Domain**: In a conceptual metaphor, the source domain is some form of concrete, embodied knowledge that is typically grounded by one or more image schemata.
- **Spatial Alteration**: A procedure that alters the overall shape, spatial size, event cardinality, and/or directionality of a spatial gesture. Spatial alterations may be intracategorical or intercategorical.
- **Spatial Analysis**: The process of identifying spatial gestures, describing their interrelationships, and interpreting their various directionalities and interactions.
- **Spatial Extremes**: The outermost points of an ensemble space and/or the leftmost and rightmost point of a spatial gesture. The lowest and highest integer values in an SG-label depict the spatial extremes of the gesture: the lowest integer value represents the leftmost spatial extreme; the highest integer value corresponds to the rightmost spatial extreme. If a gesture unfolds within the entirety of the ensemble space (as opposed to within an ensemble-space segment), the spatial extremes of the gesture and of the ensemble space will be the same.
- **Spatial Gesture**: Directed motion through an ensemble space or loudspeaker array resulting from the consecutive sound-event activity of multiple, spatially discrete performers (or speakers). A spatial gesture may emerge from either an auditory stream or an auditory sequence that is spatially dispersed. Spatial gestures may be viewed as specific kinds of musical gesture (as formulated by Robert Hatten).
- **Spatialization**: The process by which musical parts are dispersed among the performers of an ensemble space such that spatial motion emerges. The term may describe the occurrence of spatial gestures in performance (irrespective of the composer's intentionality) or the composer's deliberate act of infusing the ensemble space with spatial shapes and motions.
- **Spatial Reconciliation**: A type of point fusion that occurs when the points contributing to an oscillation gesture or to a migratory open or closed gesture subsequently become fused in their spatio-musical activity.
- **Spatio-Temporal Extremes**: When the spatial extremes and temporal endpoints of a gesture are the same, i.e., when a spatial gesture begins at one extreme of ensemble (or ensemble-segment) space and terminates at the other spatial extreme.

⁵ O'Callaghan, "Sounds," iii.

- **Spiral Gestures**: A specific type of weak closed gesture exhibiting the shape of a spiral. Spiral gestures may be clockwise or counterclockwise and may spiral inward or outward. Thus, there are four specific spiral gestures for any one-dimensional ensemble space with an ensemble-point cardinality of three or higher: clockwise inward spiral, clockwise outward spiral, counterclockwise inward spiral, and counterclockwise outward spiral.
- **Strong Closed Gesture**: A closed gesture that begins and ends at the same point in ensemble space.
- **Strong Open Gesture**: An open gesture that moves continuously in one direction through the ensemble space. Also referred to as a panning gesture.
- **Strong Oscillation Gesture**: An oscillation gesture that concludes at the same point from which it initially commenced. Both an adjacent-point and a distant-point oscillation gesture may be classified as strong.
- **Subgesture**: A component of either a compound or a complex spatial gesture. A subgesture is a clearly discernible spatial gesture in its own right.
- **Summing Localization**: Occurs whenever two spatially separated sound sources emit nearly identical sound signals, allowing the listener to perceive a single sound located at a point intermediate between the two sources.
- **Superpoint**: When two adjacent ensemble points undergo point fusion. The two points are fused together in time as well as in space (by virtue of their spatial adjacency).
- **Sweeping Gesture (SWEEP)**: A compound gesture comprised of two linked panning gestures of opposing directionality, such as a L-R PAN followed by a R-L PAN (or vice versa).
- **Target Domain**: A thing or concept that we do not experience directly through sensorimotor interaction. Through the process of cross-domain mapping, the conceptualization of the target domain is structured by an appropriate source domain.
- **Temporal Endpoints**: The first and last integers in an SG-label depicting the first and last events of a spatial gesture.
- Temporal Modifications: Alterations affecting the temporal nature of a spatial gesture.
- **Tracking Stream**: In the context of canonic streaming, the gesture(s) that "follow" (in both time and space) the lead stream.
- **Tracking Stream Overtake**: In the context of canonic streaming, when a tracking stream moves faster than a lead stream (or a preceding tracking stream) and effectively "catches-up" to it—potentially terminating at the same time (and point) as the lead stream or perhaps surpassing it to conclude prior to the cessation of the initial lead stream.

- **Truncation**: A type of abbreviation in which one of a gesture's spatial extremes is eliminated. A gesture may be right-truncated or left-truncated.
- **Type-A Intercategorical Alteration**: Involves a change of subcategory membership while broad gestural categorization remains the same. For example, a weak *open* gesture becomes a strong *open* gesture.
- **Type-B Intercategorical Alteration**: Involves the maintenance of either the "strong" or "weak" qualifier across category boundaries. For example, a *strong* open gesture may be modified to become a *strong* closed gesture.
- **Type-C Intercategorical Alteration**: An alteration that preserves neither category membership nor the strong/weak descriptor between two gestural forms.
- **Type-1 Inversion**: A type of inversion that entails reflecting a gesture across an inversional axis abiding within the center of the total ensemble space.
- **Type-2 Inversion**: A type of inversion that involves reflecting a gesture across an axis of symmetry that does not equally divide the total ensemble space. Type-2 inversion can only affect a spatial gesture that occurs in an ensemble-space segment (i.e., one that does not contain both spatial extremes of the total ensemble space). Under type-2 inversion, a gesture will reflect within the same segment of ensemble space (whereas, under type-1 inversion, the same gesture would reflect into a different—yet potentially overlapping—ensemble-space segment).
- **Uneven Consonant Streaming**: When two like gestures unfold simultaneously but at completely different (i.e., "uneven") rates in separate regions of ensemble space.
- **Weak Closed Gesture**: A closed gesture that does not end at the same point from which it began (and which otherwise is not classifiable as an open gesture).
- Weak Open Gesture: An open gesture that does not feature strictly unidirectional motion.
- **Weak Oscillation Gesture**: An oscillation gesture that does not end at the same point from which it began. Both an adjacent-point and a distant-point oscillation gesture may be classified as weak.