Thinking Out of the Grid and Inside the Spiral - Geometric Interpretations of and Comparisons with the Spiral Array Model

Elaine Chew

Assistant Professor, University of Southern California

Senior Investigator, Integrated Media Systems Center

Daniel J Epstein Department of Industrial and Systems Engineering

3715 McClintock Avenue GER240 MC:0193

Los Angeles, CA 90089-0193

http://www-rcf.usc.edu/~echew

echew@usc.edu

15 October 2003

Thinking Out of the Grid and Inside the Spiral -

Geometric Interpretations of and Comparisons with the Spiral Array Model

Abstract

This paper presents a geometric interpretation of the Spiral Array model and its comparison to Lerdahl's Tonal Pitch Space and Krumhansl's spatial representation of pitch relations. The Spiral Array model is based on the Harmonic Network. The fundamental idea underlying the model is the representing of higher level objects in the spiral's interior as convex combinations of the representations of the lower level components. By using the interior of the spiral, the original discrete space is relaxed to one that is continous. Geometric mappings are demonstrated among Lerdahl's Tonal Pitch Space, Krumhansl's spatial representation of pitch relations and the Spiral Array model. The interior point approach of the Spiral Array model is shown to generate higher level structures that are consistent with the results of these other approaches. The advantage of the interior point approach is that it facilitates comparisons across different hierarchical levels and problems that were previously combinatorial in nature can be modeled more efficiently using the continuous space in the interior.

1 Introduction

Tonality is the system of relationships that generates a hierarchy among pitches, resulting in one pitch being the most stable. Numerous geometric models for these pitch relations have been proposed. According to Shepard (1982), any "cognitive representation of musical pitch must have properties of great regularity, symmetry, and transformational invariance." Not surprisingly, many of these models are based on lattices that can be wrapped around cylinders to form spiral structures. One such lattice is the Harmonic Network, also known as the *tonnetz*. The Harmonic Network clusters pitch classes that form higher level structures in the tonal system, such as triads and keys.

This paper argues for the representing of higher level objects in a systematic fashion as spatial points inside the pitch class cylinder generated by the Harmonic Network. The edges between the vertices mark the distance between any two pitch classes represented on the network. The Harmonic Network cannot easily provide metrics for measuring the distance between objects other than pitch classes and triads, a problem addressed by using the interior space to represent objects. The continuous three-dimensional space inside the spiral provides a metric for quantifying the distance between any two objects represented in the same space.

In the spirit of the adage "a picture's worth a thousand words," the first part of the paper presents an image-driven guide to the geometry of the Spiral Array model (Chew, 2000). Representing objects out of the grid and inside the spiral is the fundamental idea behind the Spiral Array model. The second part of the paper compares the Spiral Array's geometric structures with other spatial representations of tonal pitch space proposed by Krumhansl (1978) and Lerdahl (2001). Mappings among the three models are demonstrated geometrically, validating the use of the interior point approach to representing tonal objects in space.

A distinct advantage of the interior approach is that objects from different hierarchical levels are represented in the same space, thus facilitating inter-level object comparisons. By utilizing the interior space, problems of pattern recognition can be reduced from a combinatorial one to a simple nearest neighbor search, as in the case of key-finding and in pitch spelling as discussed in previous papers (see Chew, 2001; Chew & Chen, 2003a,b) .

The inspiration for this interior point approach came from the field of Operations Research. In the domain of linear optimization, for many decades, the method of choice to solve linear programming problems was the Simplex Method, invented in 1947 by George Dantzig (see Dantzig, 1963). The Simplex Method finds the optimal solution by pivoting through adjacent corner point solutions, vertices on the convex hull of the solution space. This proved to be a reasonable approach in practice although it's computational complexity was exponential. In 1984, Karmarkar proposed the interior point approach. By traveling though the interior of the solution space, the new algorithm was shown to be polynomial in complexity.

2 From Grid to Spiral Representations

The Harmonic Network is a network representation of pitch relations where each node represents a pitch class, that is to say, a set of pitches related by some multiple of an octave. In network terminology, each node is a vertex of degree six (having six edges incident on the node). Each opposing pair of edges connects the pitch class to other nodes related by one of three intervals – Perfect fifth (P5), major third (M3) and minor third (m3) – as shown in Figure 1. The Harmonic Network, also known as the *tonnetz*, forms the foundation of Neo-Riemannian Theory and has been attributed to the mathematician Euler (see Cohn, 1998; Lewin, 1987, 1982).

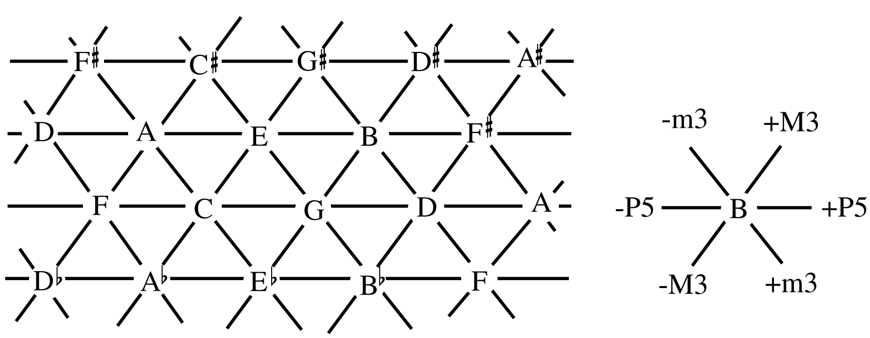


Figure 1: The Harmonic Network, also known as the *Tonnetz*.

Each triangle in the Harmonic Network forms a triad (major or minor depending on its orientation). The network of triads forms the dual graph of the Harmonic Network (see Figure 2). Each new edge that cuts across an arc in the original lattice represents a distance-minimizing transformation between two triads. Transformations on the dual graph have been used to analyze triadic movement in tonal and post-tonal music (see Cohn, 1996, 1997).

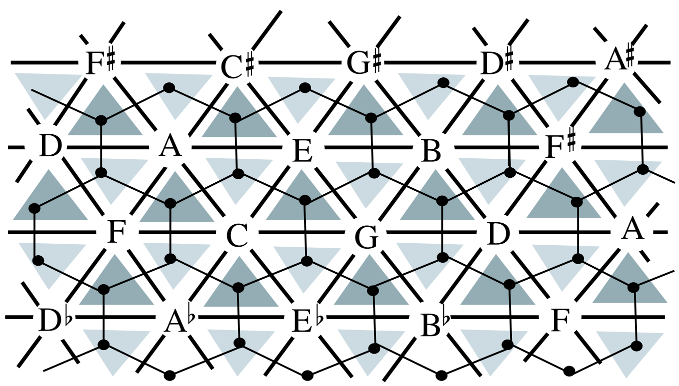
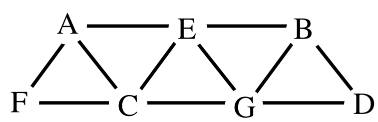


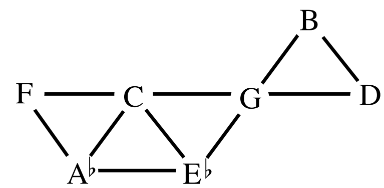
Figure 2: Triads form the dual graph of the Harmonic Network.

Pitches that belong to a given key also form compact sets of connected components with unique shapes (see Figures 3(a) and 3(b)). This property was exploited in Longuet-Higgins and Steedman's shape matching algorithm for key-finding (Longuet-Higgins & Steedman, 1971; Longuet-Higgins, 1976). Transformations on key shapes can be used as a metric for comparing keys, but would be less suitable for comparing objects from different hierarchical levels, for example, keys and triads.

Most literature on the Harmonic Network alude to the spiral structure (and toroid under the assumption of enharmonic equivalence) inherent in the grid (See Figure 4). However, the three-dimensional realization of the model is hardly used and not necessary for performing transformations (see Lewin, 1987) and deriving group-theoretic properties of the network (see Balzano, 1980). Thus, the three dimensional spiral configuration of the Harmonic Network is rarely used for more



(a) C major key shape.



(b) C (harmonic) minor key shape.

Figure 3: Uniquely shaped connected components representing major and minor keys.

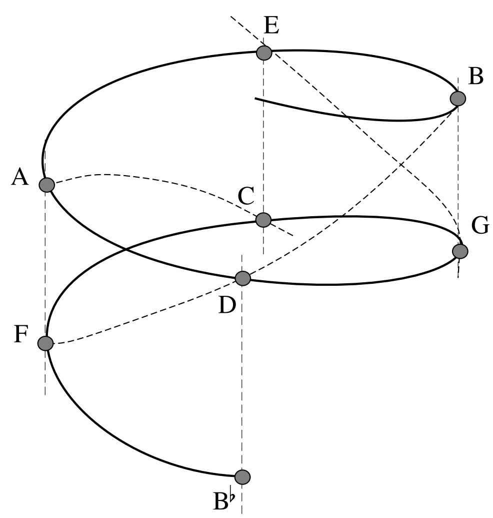


Figure 4: Spiral representation of the Harmonic Network.

than illustrative purposes.

Other grid models that map to cylindrical spiral structures include Lerdahl's Tonal Pitch Space (2001) based on a distance metric defined on a network of pitch classes on a cone similar to the one discovered by Krumhansl by experimental means (1978). The resulting lattice of chord and key relationships also wrap nicely onto cylindrical spirals (that fold over into tori). Another example is Shephard's double helical model for pitch relations (Shepard, 1982). These models utilize only the discrete space. Strict adherence to the lattice structure often leads only to integral values for inter-object distance produced by counting edges or transformations on the lattice.

3 Getting Inside the Spiral: Geometry of the Spiral Array

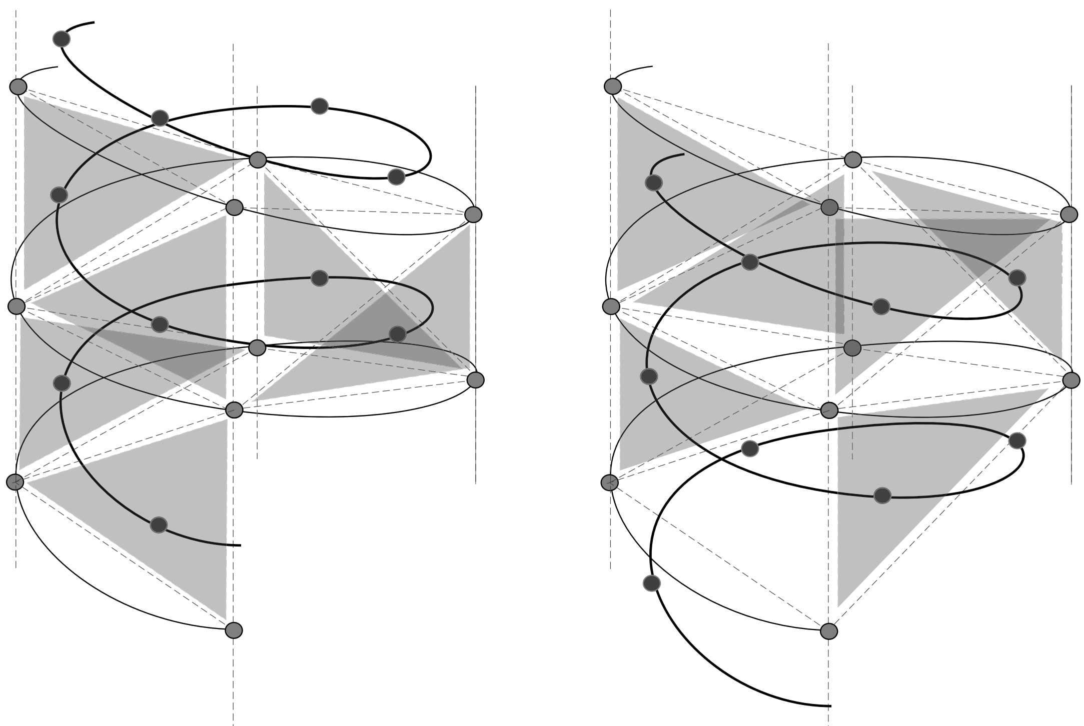
The Spiral Array is a geometric model that spatially represents pitches, chords and keys as points on the spiral configuration, as well as inside the spiral, of the Harmonic Network. The fundamental insight behind the model is that any collection of pitches can generate a *center of effect* (*c.e.*) that is an interior point in the convex hull of its component pitch representations and whose distance from any other element can then be measured. By using the interior space, the Spiral Array is able to represent pitches, intervals, chords (major and minor triads) and keys (major and minor) in the same spatial framework. It also represents the inter-relations among these objects as distance measured through the interior of the spiral.

The Spiral Array model begins with the spiral configuration of pitch classes as shown in Figure 4.

The equation for the pitch classes is as follows:

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

where k marks the pitch's distance from C on the line of fifths and C is arbitrarily set at position [0,1,0].



(a) Major triad spiral. (b) Minor triad spiral.

Figure 5: Triad Representations in the Spiral Array.

Each triad is represented as a point on the face of the triangle outlined by its component pitches.

Each triad is a convex combination of its root, fifth and third. Note that the triad representation generated in this fashion is a point in the interior of the spiral. The set of major triads form a spiral inside the pitch spiral (shown in Figure 5(a)), as do the set of minor triads (shown in Figure 5(b)).

The major triad equation is:

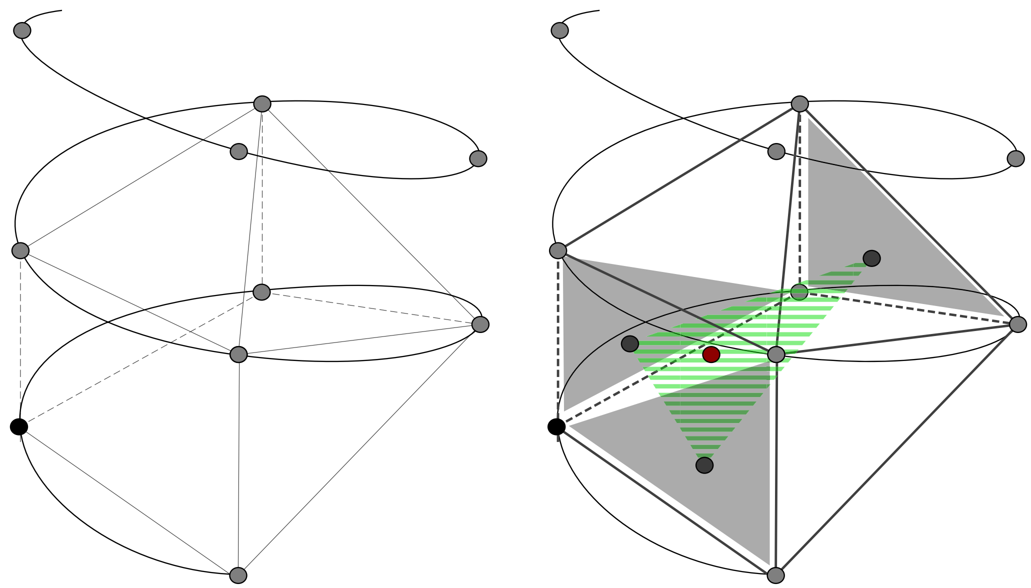
$$\mathbf{C}_{\mathrm{M}}(\mathrm{k}) \stackrel{\mathrm{def}}{=} w_1 \cdot \mathbf{P}(\mathrm{k}) + w_2 \cdot \mathbf{P}(\mathrm{k}+1) + w_3 \cdot \mathbf{P}(\mathrm{k}+4),$$

where $w_1 \geq w_2 \geq w_3 > 0$ and $\sum_{i=1}^{3} w_i = 1$. The minor triad is generated by a similar equation:

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

where $u_1 \geq u_2 \geq u_3 > 0$ and $\sum_{i=1}^{3} u_i = 1$. The weights, w_i and u_i determine where on the triangle the point representing the triad resides. By choosing these weights carefully, the distance of the triad representation to its component pitch classes can reflect the desired relations between these objects. By design, the range of possible distance relations is constrained by the structure of the original Harmonic Network as well as the way in which the triad representations are defined.

As in the Harmonic Network, pitch classes belonging to a given key form compact clusters in the Spiral Array model. Figure 6 and 7 show the convex hull of the pitch classes and the way in which the key representations are generated for the major and minor key respectively. The major key is represented by a spatial point in the interior of the three-dimensional spiral structure. Figure 6 shows the convex hull of the spatial representations of the pitch classes in a given major key. Since each major key is uniquely defined by its I, V and IV triads, the major key representation is defined to be a point on the face of the triangle outlined by the spatial representations of its tonic, dominant and subdominant triads as shown in Figure 6(b). Figure 7(a) and (b) show the corresponding geometric objects for the minor (harmonic) key.



(a) Convex hull of pitches in major key. (b) Representing a major key.

(The tonic is marked by a black sphere on the pitch spiral.)

Figure 6: Major key span and representation in the Spiral Array.

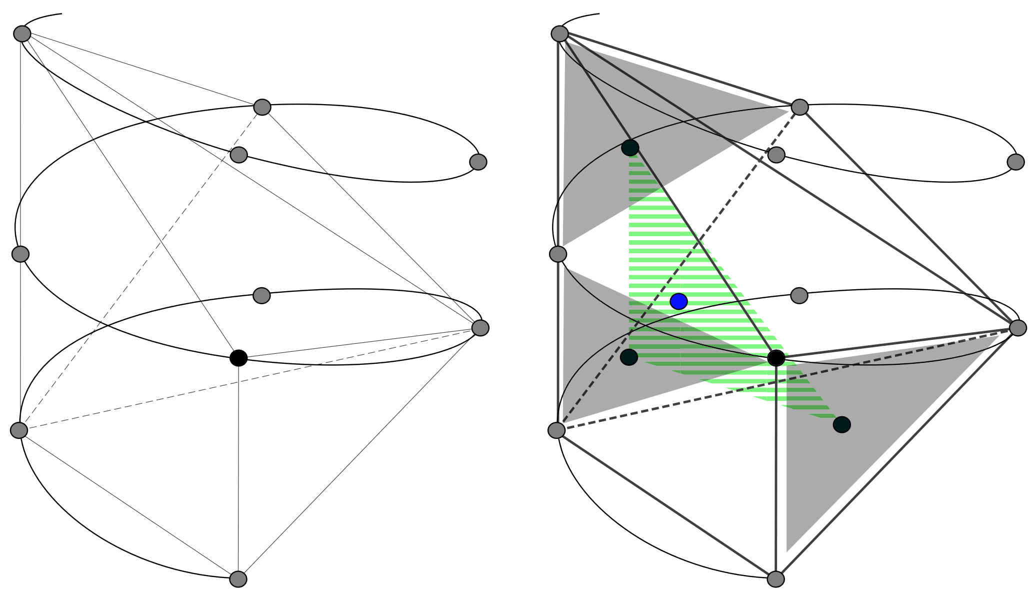
The equation for the major key representation is as follows:

$$\mathbf{T}_{\mathrm{M}}(\mathrm{k}) \stackrel{\mathrm{def}}{=} \omega_1 \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}) + \omega_2 \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}+1) + \omega_3 \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}-1),$$

where $\omega_1 \geq \omega_2 \geq \omega_3 > 0$ and $\sum_{i=1}^{3} \omega_i = 1$. The minor key definition is as follows:

$$\begin{aligned}\mathbf{T}_{\mathrm{m}}(\mathrm{k}) \stackrel{\mathrm{def}}{=} \upsilon_1 \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}) \quad &+ \quad \upsilon_2 \cdot [\alpha \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}+1) + (1-\alpha) \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}+1)] \\ &+ \quad \upsilon_3 \cdot [\beta \cdot \mathbf{C}_{\mathrm{m}}(\mathrm{k}-1) + (1-\beta) \cdot \mathbf{C}_{\mathrm{M}}(\mathrm{k}-1)],\end{aligned}$$

where $\upsilon_1 \geq \upsilon_2 \geq \upsilon_3 > 0$ and $\upsilon_1 + \upsilon_2 + \upsilon_3 = 1$, and $0 \leq \alpha \leq 1, 0 \leq \beta \leq 1$. The constraints on the weights are chosen to reflect each chord's significance in the key. When $\alpha = \beta = 1$, $\mathbf{T}_{\text{M}}(\text{k})$ represents the harmonic minor key.

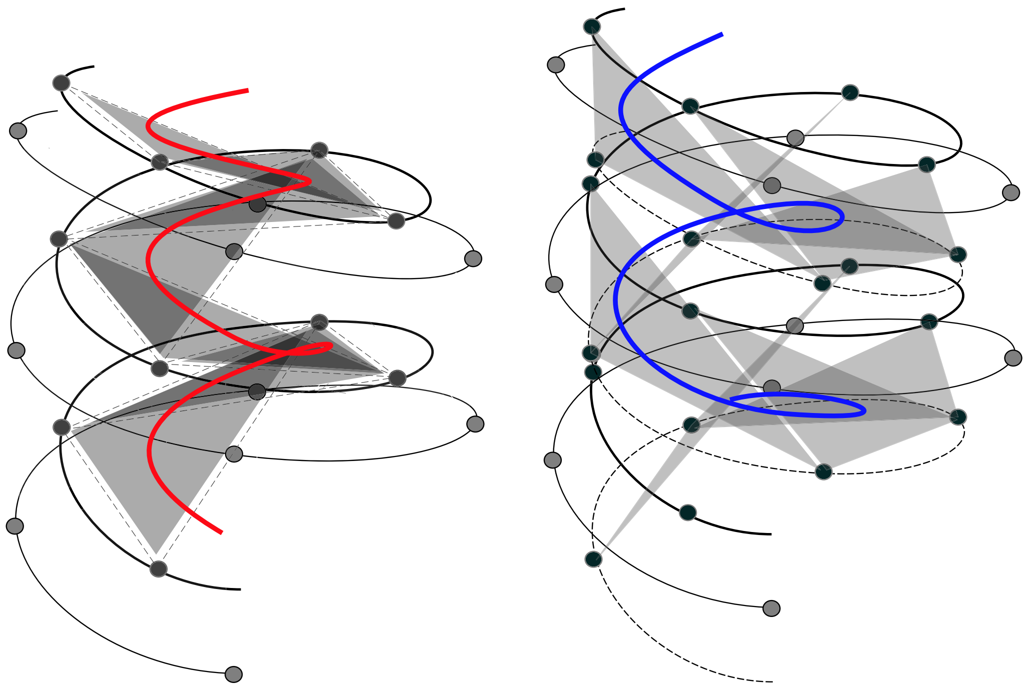


(a) Convex hull of pitches in minor (harmonic) key. (b) Representing a (harmonic) minor key.

(The tonic is marked by a black sphere on the pitch spiral.)

Figure 7: Minor key span and representation in the Spiral Array.

Like the major and minor triads, the major and minor keys also form spiral structures as shown in Figure 8. Figure 8(a) shows a spiraling sequence of major key triangles and Figure 8(b) shows the same for the minor key triangles. The key representation is chosen to be a point on the triangle with its defining triads as vertices.



(a) Major keys. (b) Minor (harmonic) keys.

Figure 8: Key Representations in the Spiral Array.

4 A Metric to Compare Distance Between Objects

The Spiral Array model can be visualized as a set of nested spirals as shown in Figure 9. Because objects from different hierarchical levels are represented in the same space, the model can be calibrated so that the perceived distance between any two objects can be quantified.

One of the advantages of using the interior of the spiral is that Euclidean space can now provide a metric for quantifying the distance between any two objects from any hierarchical level. For example, Figures 10 shows the relative positions of major and minor spirals for triad representations

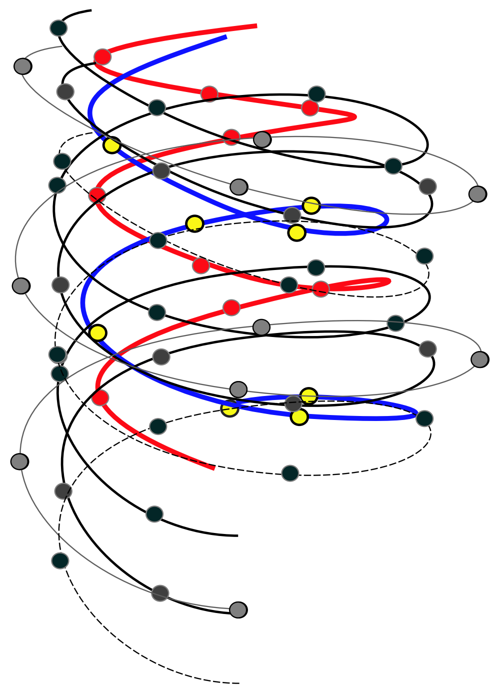


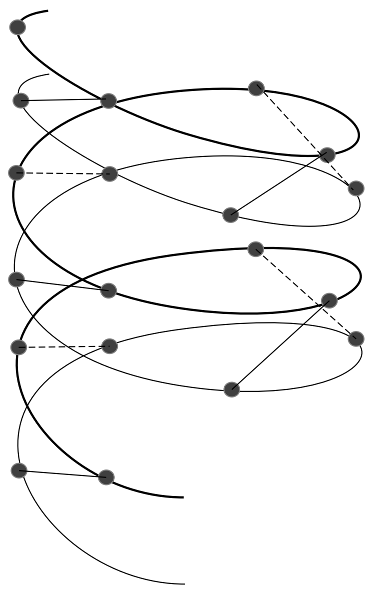
Figure 9: The Spiral Array model visualized as an array of spirals.

and key representations respectively.

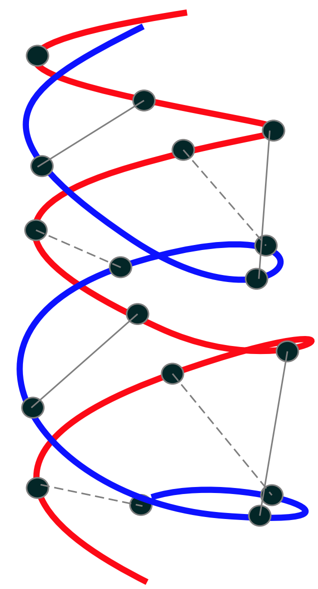
5 Comparing the Results of the Interior Point Approach

The previous section described the interior point approach to representing higher level tonal objects inside the pitch class spiral. This method of representing higher level objects in the same space results in an array of spirals, each representing a different type of tonal object. Since the Spiral Array begins with the Harmonic Network, it inherits all the spatial properties of the network. In this section, the geometry of the Spiral Array will be compared to that of two other models, namely, Krumhansl's spatial representation of pitch relations and Lerdahl's Tonal Pitch Space.

Krumhansl's results were formed by applying multi-dimensional scaling techniques to experimental data. An inverted version of Krumhansl's pitch cone using pitch class notation is shown in



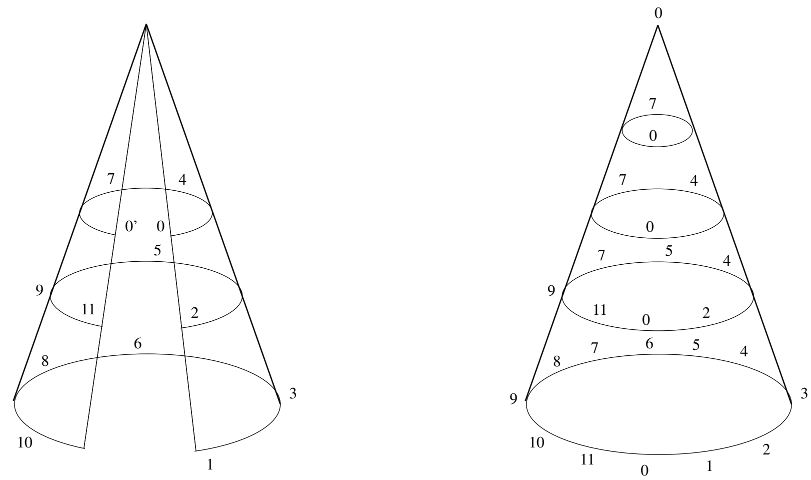
(a) Major and minor triads superimposed.



(b) Major and minor keys superimposed.

(Transverse lines across spirals connect objects with the same name: the solid lines connect objects in the foreground and the dotted lines connect objects in the background.)

Figure 10: Major and minor representations in the Spiral Array.



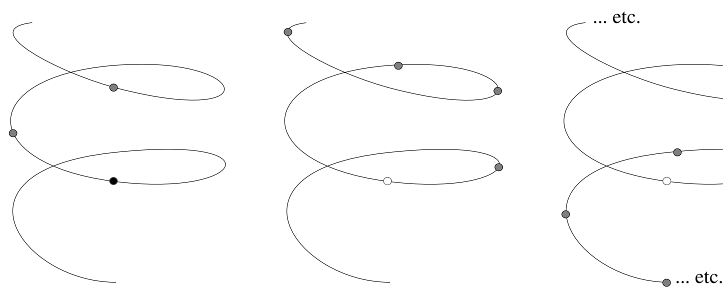
(a) Krumhansl's pitch cone (inverted).

(b) Lerdahl's pitch class cone.

Figure 11: Pitch Class Cones.

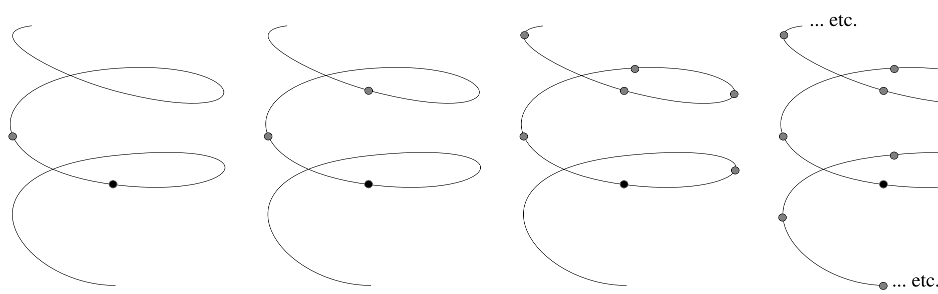
Figure 11(a). The top-most layer contains pitches in the tonic triad. The second layer contains the other pitches in the diatonic scale, and the base layer contains the remaining five pitches outside of the diatonic scale. Lerdahl's Tonal Pitch Space is based on Deutsch and Feroe's idea of hierarchically organized pitch classes. Lerdahl's pitch class cone is diagrammed in Figure 11(b). The main difference between the two cones is that Lerdahl's pitch cone contains an additional layer highlighting the Perfect Fifth interval relation. To map the layers from Krumhansl's to Lerdahl's model, one simply allows pitch classes to carry over to the next layer down on the cone.

As shown in Figure 11, the pitch arrangements in the Spiral Array mirror that of Krumhansl's and Lerdah's pitch class cones. Figure 12 shows the Spiral Array pitch classes that correspond to each layer in Krumhansl's cone – the tonic is shown as a white sphere in the figures other than the first as a reference although it is not one of the pitches in that layer. Figure 13 shows the Spiral Array pitch classes that correspond to each layer in Lerdahl's pitch class cone.



(a) Layer closest to apex. (a) Second layer. (a) Third layer.

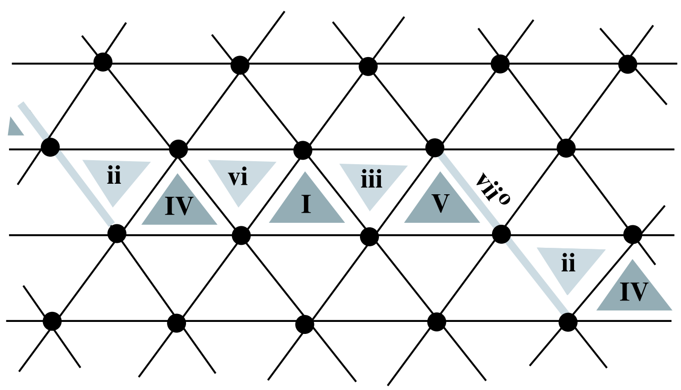
Figure 12: From the Spiral Array to Krumhansl's Pitch Class Cone.



(a) Layer closest to apex. (a) Second layer. (a) Third layer. (b) Fourth layer.

Figure 13: From the Spiral Array to Lerdahl's Pitch Class Cone.

Observe that the hierarchical ordering of the distance from the tonic to each level of pitches in the cone representation is part of the Spiral Array structure. Compare Figures 12 and 11(a). Each layer that is closer to the base of the cone contains pitches that map to positions on the Spiral Array that are progressively farther away from the tonic. Compare Figures 13 and Figure 11(b). Each layer closer to the base of the cone maps to positions on the Spiral Array that define an expanding compact set.

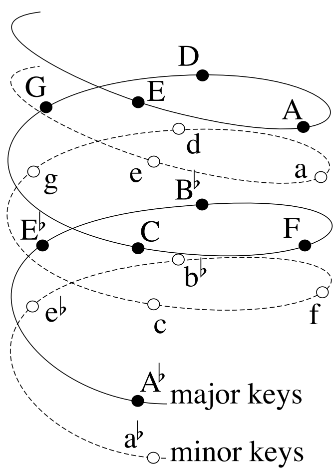


vii^o	ii	IV	vi	I	iii	V
iii	V	vii^o	ii	IV	vi	I
vi	I	iii	V	vii^o	ii	IV
ii	IV	vi	I	iii	V	vii^o
V	vii^o	ii	IV	vi	I	iii
I	iii	V	vii^o	ii	IV	vi
IV	VI	I	iii	V	vii^o	ii

Figure 14: Relations among chord functions in the Harmonic Network and Lerdahl's Chordal Space.

Lerdahl derives chord relations and key relations based on steps between objects in the pitch cone space. It is not surprising that the chord and key relations that are derived in Lerdahl's Tonal Pitch Space correspond to structures in the Spiral Array model. The table in Figure 14 shows the chord relations that are derived and Figure 15 tabulates the key relations that are abstracted from the pitch class cone. This table of key regions parallels Krumhansl and Kessler's (1982) chart of the multidimensional scaling solution for key relations.

Each row in the chordal space traces the path highlighted in the Harmonic Network in Figure 14. Because the Spiral Array inherits the pitch relations shown in the Harmonic Network, this implies that the same chord relations can be shown in the Spiral Array.



d$\sharp$	g$\sharp$	c$\sharp$	f$\sharp$	b	e	a
F$\sharp$	B	E	A	D	G	C
f$\sharp$	b	e	a	d	g	c
A	D	G	C	F	B$\flat$	E$\flat$
a	d	g	c	f	b$\flat$	e$\flat$
C	F	B$\flat$	E$\flat$	A$\flat$	D$\flat$	G$\flat$
c	f	b$\flat$	e$\flat$	a$\flat$	d$\flat$	g$\flat$

Figure 15: Key Representations in the Spiral Array and Lerdahl's Regional Space.

The connection between Lerdahl's Regional Space (and Krumhansl and Kessler's key chart) and the Spiral Array is more apparent because the regional space is not limited to the keys of the diatonic pitch set. As can be seen in Figure 15, the key relations represented in the different models are equivalent. Assuming enharmonic equivalence, the spiral structures would wrap around to form a torus. Findings on mental models of key relations have been remarkably consistent. For example, recent experiments using Magnetic Resonance Imaging of brain activity with subjects listening to melodies in different keys have further confirmed the toroid structure of key relations (Zatorre & Krumhansl, 2002).

6 Conclusions

This paper posits and validates the use of the interior point approach to modeling higher level structures using the spiral configuration of the Harmonic Network. To demonstrate this concept, a geometric interpretation of the Spiral Array model and its fundamental idea of representing higher level objects in the interior as convex combinations of the representations of the lower level components was presented. This interior point approach not only preserves the pitch relations of the original lattice, it also generates higher level structures that are consistent with other researchers' results. In particular, mappings were shown among Lerdahl's Tonal Pitch Space, Krumhansl's spatial representations of pitch relations and the Spiral Array model.

Note that the interior point approach is not limited to the Harmonic Network. The idea extends to any geometric model that clusters objects that form higher level structures. A distinct advantage of the interior approach is that objects from different hierarchical levels are represented in the same space, thus facilitating inter-level object comparisons. By creating and utilizing a continuous interior space through linear relaxation of a discrete space, problems of tonal recognition can be reduced from combinatorial ones in discrete space to computationally simpler ones, such as nearest-neighbor searches, in continuous space.

References

Balzano, G.J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal*, **4**, 66–84.

Chew, E. (2000). *Towards a Mathematical Model of Tonality*. Ph.D. thesis, Massachussetts Institute of Technology, Cambridge, MA.

Chew, E. (2001). Modeling Tonality: Applications to Music Cognition. In J.D. Moore & K. Stenning, eds., *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society*, 206–211, Lawrence Erlbaum Assoc. Pub, Mahwah, NJ/London, Edinburgh, Scotland, UK.

Chew, E. & Chen, Y.C. (2003a). Determining context-defining windows: Pitch spelling using the spiral array. In *Proceedings of the 4th International Conference on Music Information Retrieval*, Baltimore, MD.

Chew, E. & Chen, Y.C. (2003b). Mapping MIDI to the Spiral Array: Disambiguating Pitch Spellings. In H.K. Bhargava & N. Ye, eds., *Computational Modeling and Problem Solving in the Networked World - Proceedings of the 8th INFORMS Computer Society Conference*, 259–275, Kluwer, Chandler, AZ.

Cohn, R. (1996). Maximally smooth cycles, hexatonic systems, and the analysis of late-romantic triadic progressions. *Music Analysis*, **15**, 9–40.

Cohn, R. (1997). Neo-riemannian operations, parsimonious trichords, and their tonnetz representations. *Journal of Music Theory*, **41**, 1–66.

Cohn, R. (1998). Introduction to Neo-Riemannian Theory: A Survey and a Historical Perspective. *Journal of Music Theory*, **42**, 167–180.

Dantzig, G.B. (1963). *Linear Programming and Extensions*. Princeton University Press, Princeton.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. *Combinatorica*, **4**, 373–395.

Krumhansl, C.L. (1978). *The Psychological Representation of Musical Pitch in a Tonal Context*. Ph.D. thesis, Stanford University, Stanford, CA.

Krumhansl, C.L. & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, **89**, 334–368.

Lerdahl, F. (2001). *Tonal Pitch Space*. MIT Press, Cambridge, MA.

Lewin, D. (1982). A formal theory of generalized tonal functions. *Journal of Music Theory*, **26**, 32–60.

Lewin, D. (1987). *Generalized Musical Intervals and Transformations*. Yale University Press, New Haven, CT.

Longuet-Higgins, H.C. (1976). Perception of melodies. *Nature*, **263**, 646.

Longuet-Higgins, H.C. & Steedman, M.J. (1971). On interpreting bach. In B. Meltzer & D. Michie, eds., *Machine Intelligence (Vol. 6)*, 221, Edinburgh University Press, Edinburgh.

Shepard, R.N. (1982). Structural representations of musical pitch. In D. Deutsch, ed., *The Psychology of Music*, 335–353, Academic Press, New York, NY.

Zatorre, R.J. & Krumhansl, C.L. (2002). Mental models and musical minds. *Science*, **298**, 2138–2139.