

VISUALIZATION OF LOW DIMENSIONAL STRUCTURE IN TONAL PITCH SPACE

J. Ashley Burgoyne and Lawrence K. Saul
Department of Computer and Information Science
University of Pennsylvania
{burgoyne, lsaul}@cis.upenn.edu

ABSTRACT

In his 2001 monograph *Tonal Pitch Space*, Fred Lerdahl defined a distance function over tonal and post-tonal harmonies distilled from years of research on music cognition. Although this work references the toroidal structure commonly associated with harmonic space, it stops short of presenting an explicit embedding of this torus. It is possible to use statistical techniques to recreate such an embedding from the distance function, yielding a more complex structure than the standard toroidal model has heretofore assumed. Nonlinear techniques can reduce the dimensionality of this structure and be tuned to emphasize global or local anatomy. The resulting manifolds highlight the relationships inherent in the tonal system and offer a basis for future work in machine-assisted analysis and music theory.

1. INTRODUCTION

Since Gottfried Weber introduced the chart in Figure 1 early in the nineteenth century [12], music theorists have acknowledged two pivotal axes controlling the relationships among the major and minor keys of the diatonic tonal system in Western art music: the cycle of fifths, represented on the vertical axis of the figure, and the cycle of thirds, represented on the horizontal.¹ Capital letters designate major keys and lowercase minor, as is traditional. These axes are most often considered to be periodic, defining a topological space isomorphic to $\mathbb{S}^1 \times \mathbb{S}^1$, and by the end of the twentieth century, Carol Krumhansl's pioneering psychological experiments had demonstrated a cognitive basis for this toroidal structure [4]. Krumhansl's work also explored topological relationships among harmonies and pitch classes within each key [5], which Fred Lerdahl integrated into the framework of *A Generative Theory of Tonal Music* [7], his 1983 monograph coauthored with Ray Jackendoff, in a 2001 monograph entitled *Tonal Pitch Space* [6].

Although Lerdahl makes much of Krumhansl's data and the toroidal topology of harmonic space, he defines that torus only implicitly, by way of a distance function over harmonies. No other research to date has attempted

¹ Vial had organized the keys in a similar fashion three decades earlier, but Weber's treatise proved to be more influential.

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

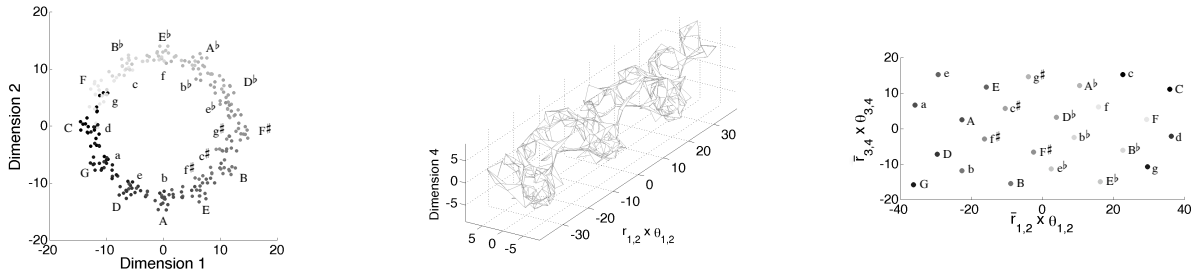
Figure 1. Gottfried Weber's diagram of regions (keys) in tonal space. The cycles of fifths run vertically and the cycles of minor thirds horizontally in all directions.

to embed it explicitly. David Temperley used a MIDI-based approach to implement many components of the theory [11] but has not yet treated its topology. The Mathematical Music Theory Group at the Technical University of Berlin uncovered some inconsistencies in Lerdahl's theory while developing their *HarmoRubette* software tool [9] but did so strictly in terms of distance functions; Guerino Mazzola's monograph has treated the topic in more detail [8]. Elaine Chew's spiral model [1, 2] is an explicit representation of tonal space that has aided the development of intelligent musical systems, most notably for key finding and pitch spelling, but it is founded on music theoretical principles (the Riemannian *Tonnetz*) that, despite the apparent similarities, are incompatible with Lerdahl's and Krumhansl's.

In this paper, we use statistical techniques to produce explicit embeddings based on Lerdahl's harmonic topology. The visualizations of tonal pitch space presented complement Chew's model and should be especially useful for machine-assisted harmonic and hierarchical analysis.

2. LERDAHL'S DISTANCES

One of the distinguishing features of Lerdahl's model is that it treats pitch classes, chords, and regions (keys) as unified and inseparable. There is no well defined notion of distance between pitch classes *qua* pitch classes or chords *qua* chords. Pitch classes have meaning only as elements



(a) Dimensions 1 and 2: A circle of fifths. Only harmonies from major keys are plotted, each key given its own shade of gray and labelled outside the circle.

(b) Dimensions 3 and 4: An inner spiral. The circle from the previous figure has been broken at C major and straightened to form the long axis here. Lines connect neighbors.

(c) A regional *Tonnetz*. Both pairs of dimensions have been broken and unfolded here and tonic chords plotted only. The shades of gray match 2(a).

Figure 2. Tonal pitch space as viewed with multidimensional scaling

of the sets that define chords and regions, and chords are always understood as functioning within some region. An important corollary is that there is always a nonzero distance, albeit usually small, between two instances of the same nominal chord when these instances are heard in distinct regions: C/C is not the same as C/F and certainly not the same as C/D^b .² This corollary gives Lerdahl’s model more nuance than most alternatives.

For two harmonies $x = C_1/R_1$ and $y = C_2/R_2$, the simple distance is given by the equation

$$\delta(x \rightarrow y) = i + j + k \quad (1)$$

where i is the smallest number of steps along the circle of fifths between R_1 and R_2 (or their relative majors in case one or both is minor), j is the smallest number of steps along the circle of fifths between the roots of C_1 and C_2 within each region, and k is a specially weighted Hamming distance between the sets of pitch classes that define each chord and region. Lerdahl’s formulation of the k parameter is asymmetric, and so to create a symmetric distance function, we have taken the average of the two directions. Lerdahl restricts δ to prevent implausible modulations, allowing it to be defined only when either x and y are in the same region or at least one of C_1 and C_2 is a tonic chord and R_1 and R_2 are in each other’s set of “pivot regions,” $\{\mathbf{i}, \mathbf{ii}, \mathbf{iii}, \mathbf{IV}, \mathbf{V}, \mathbf{vi}\}$ for major keys and for minor keys $\{\mathbf{I}, \mathbf{bIII}, \mathbf{iv}, \mathbf{v}, \mathbf{bVI}, \mathbf{bVII}\}$. The general distance function is

$$\begin{aligned} \Delta(C_1/R_1 \rightarrow C_2/R_2) = & \delta_1(C_1/R_1 \rightarrow I/P_1) \\ & + \delta_2(P_1 \rightarrow P_2) + \delta_3(P_2 \rightarrow P_3) + \dots \\ & + \delta_n(I/P_n \rightarrow C_2/R_2) \quad (2) \end{aligned}$$

where $\delta(P_i \rightarrow P_j)$ is shorthand for $\delta(I/P_i \rightarrow I/P_j)$, and the chain of regions P_1, P_2, \dots, P_n is chosen to minimize Δ within the constraint that $\delta_1, \delta_2, \dots, \delta_n$ are defined.

² Lerdahl uses Roman type for chords and boldface type for regions, and we follow the convention here, e.g., \mathbf{ii}/G for the minor supertonic chord in G major or E^b/F for an E^b major chord understood in the key of F major.

3. MULTIDIMENSIONAL SCALING

Metric multidimensional scaling (MDS) is a linear statistical technique that produces an explicit geometric map of a set of objects given a complete matrix of distances between any pair of them. This map replicates the distances as closely as possible in a space of arbitrary dimensionality. For most data, there is a trade-off between keeping the number of dimensions low and preserving the original distances; normal practice is to apply the technique over a range of possible dimensionalities and then select an optimum for the task at hand. If only some of the pairwise distances are available, the remainder can be estimated by computing the shortest paths through a graph with a vertex for each of the original objects and weighted edges connecting every pair of points for which the distance is known; this technique is analogous to the Isomap algorithm for nonlinear dimensionality reduction, which deliberately discards and recomputes distance measurements for all but the closest neighbors [10].

We used MDS to analyze the matrix of Lerdahl distances among a set of common harmonies. These harmonies were restricted to the major and minor triads of each of the 24 major and minor keys and the chromatically altered major and minor triads that can be reached via simple, secondary, or double mixture. The set includes 22 triads per key (every triad except the major and minor triads rooted a tritone away from the tonic) for a total of 528 harmonies. The constraints on Equation 1 define a partial distance matrix on this set, and the generalized distance in Equation 2 is precisely the shortest path algorithm described above.

While two, three, or even four dimensions of the result are insufficient to represent the distances well – fifteen would be required to capture even 75 percent of the information in Lerdahl’s model – we can still use them for visualization. Figure 2(a) is a plot of major-key harmonies in the leading two dimensions. Each tonal region is colored in its own shade of gray and labeled with capital letters. The pattern is a tightly organized regional circle of fifths, which confirms that Lerdahl’s model conforms in at least one regard to the Weber-Krumhansl model it seeks

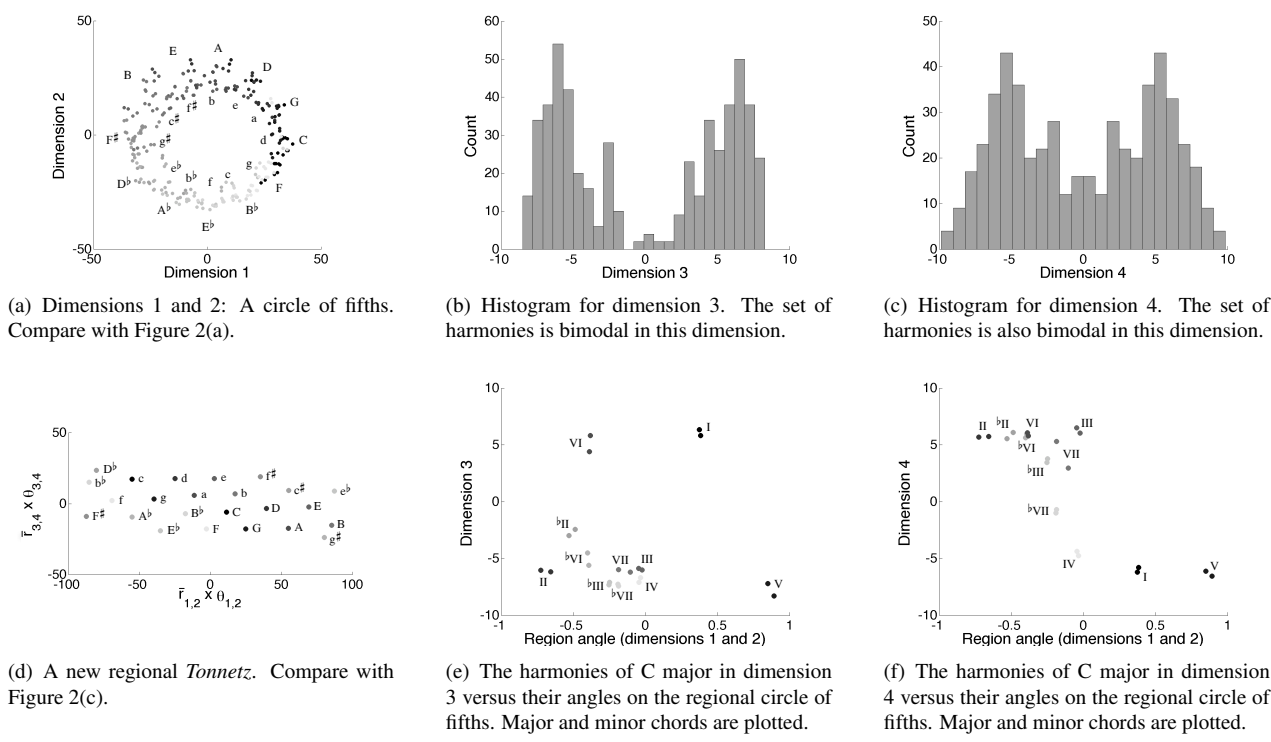


Figure 3. Tonal pitch space as viewed with maximum variance unfolding

to emulate.³ One thus would expect the second and third dimensions to trace a cycle of thirds, but the structure is more complicated than that. Figure 2(b) unwraps the circle from Figure 2(a) by converting the first two dimensions to polar coordinates and plots the third and fourth dimensions of the embedding with respect to the polar angle. These dimensions form a spiral with three periods to the first dimension's one. Figure 2(c) attempts to clarify the relationship between these structures by converting the third and fourth dimensions to polar coordinates as well and plotting their angular component against the angular component of the first two dimensions, scaled in each case to the average radius. Here, the cycles of thirds emerge. Crisscrossing the spiral in a form that looks much like a Riemannian *Tonnetz*, the traditional cycles of minor thirds in regional space travel from the top left to the bottom right of the figure and cycles of major thirds move horizontally.

The presence of the circle of fifths and the cycles of thirds is sufficient for isomorphism to the toroidal regional model of the psychological literature. The remaining two dimensions, the radial components of the two sets of polar coordinates, distinguish our model. Together, they or-

³ If space had permitted the minor keys to be plotted in the form of 2(a), they would be interspersed with the major keys but form their own circle of fifths corresponding to the lowercase labels. In contrast to both the Krumhansl model and most theory textbooks, major keys are paired with neither their relative or parallel minors but the minor key a whole tone higher. This relationship arose early in harmonic theory with David Heinichen's *General-Bass in der Composition* [3]. It arises as a neutral statistical compromise between the parallel and relative key relationships so as to allow the third and fourth dimensions to account for them properly.

ganize the chords within each region around the toroidal structure of the regions themselves. Weber first presented the regional structure, Krumhansl and Kessler assigned an embedding to it, Lerdahl developed a theory to incorporate inter-regional relationships, and our work derives a new embedding for them. It should allow machine analysis systems to synthesize key finding and harmonic analysis more smoothly.

4. MAXIMUM VARIANCE UNFOLDING

The inter-regional structures in this embedding are less consistent than the intra-regional ones. This shortcoming is tied to MDS, which must optimize over the global structure of its input data. There are nonlinear algorithms, however, that can shift the emphasis to local structures. One very effective such technique is maximum variance unfolding (MVU) [10]. Like MDS, this algorithm produces an embedding from a matrix of pairwise distances, but while maximizing the variance of the output embedding, it seeks to preserve only the distances between nearest neighbors. This subset of distances is locked, and a nonlinear optimization technique is used to expand the data as much as possible given these locks, analogous to stretching a ball-and-stick model in which the balls correspond to harmonies and the sticks correspond to the locked distances. By tuning the size of the neighborhoods, one can control the level of structure in the output embedding. Large neighborhoods yield more global structures and behave comparably to algorithms like MDS, while smaller neighborhoods preserve local structure and can provide

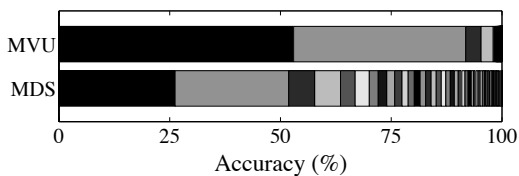


Figure 4. Dimensionality-accuracy trade-offs. Each segment of the bars represents the additional accuracy afforded by an additional dimension.

much better dimensionality reduction.

MVU is not designed to analyze non-Euclidean distances, and as Noll and Garbers note, Equation 2 is not a true distance function because its special handling of pivot regions causes it to violate the triangle inequality. For our experiments, we Euclideanized the Δ -derived distance matrix before computing neighborhoods by converting it to a Gram matrix of inner products, replacing all negative eigenvalues with zeros, and converting back.

Figure 3(a) presents the leading two dimensions of the MVU embedding. As in the linear case, they form a circle of fifths. The third and fourth dimensions, however, serve different purposes. The histograms in Figures 3(b) and 3(c) show that the data is bimodal in each of these dimensions. The third dimension separates regions into two isomorphic planes a semitone apart; the fourth dimension separates the major keys from the minor keys. These patterns are evident in Figure 3(d) and form a very different regional network than the one from the MDS embedding in Figure 2(c). These dimensions also preserve consistent chordal structures across the regional structure. As seen in Figures 3(e) and 3(f), dimension 3 keeps tonics with the tonics of their relative keys while dimension 4 puts them closer to the dominant and subdominant. Notably absent from the nonlinear embedding are the cycles of thirds. Instead, the MVU embedding prioritizes cycles of major seconds and binary oppositions between parallel and relative region pairs. This unexpected result calls for experimentation on corpuses of real music to see which model is more successful.

MVU has a strong advantage over MDS when it comes to dimensionality reduction, however; Figure 4 illustrates the difference. The top bar represents the distribution of information after using MVU on the Euclideanized distance matrix with neighborhoods including the four nearest neighbors to each harmony; the leading four dimensions account for 98 percent of the distance information. After MDS, represented in the lower bar, the leading four dimensions account for only 57 percent. Thus, the structure shown in Figure 3 should be a much better representation of the structure of Lerdahl’s space than Figure 2, further challenging the notion of a cycle of thirds.

5. SUMMARY AND FUTURE WORK

Linear and nonlinear statistical methods can produce embeddings that emphasize the global or local structure of

data defined by pairwise distances and can help visualize models of tonal pitch space, including structures more complex than the three-dimensional toroidal model commonly cited in psychological literature. These higher dimensional embeddings incorporate more subtle details of the harmonic system that are helpful to visualize and can serve as foundational models for machine-assisted analysis. They free analysis from the finite set of harmonic labels, replacing them instead with four continuous parameters that can encode more nuanced musical concepts. The labeling or classification problem becomes one of spatial localization; ambiguous tonalities can be assigned explicit locations between more stable regions. Moreover, Lerdahl’s theory includes extensions for hexatonic, octatonic, and other non-diatonic tonal models that we hope to incorporate into our existing framework.

6. ACKNOWLEDGEMENTS

This work was supported by NSF Award 0238323. We also would like to thank Fei Sha, Kilian Weinberger, Chris Hasty, and Fred Lerdahl for sharing their ideas.

7. REFERENCES

- [1] E. Chew, “Towards a mathematical model of tonality,” Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2000.
- [2] —, “Thinking out of the grid and inside the spiral: Geometric interpretations of and comparisons with the spiral array model,” University of Southern California, Tech. Rep. 03-002, 2003.
- [3] J. D. Heinichen, *General-Bass in der Composition*, Dresden, 1728.
- [4] C. L. Krumhansl, *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press, 1990.
- [5] C. L. Krumhansl and E. J. Kessler, “Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys,” *Psychological Review*, vol. 89, pp. 334–368, 1982.
- [6] F. Lerdahl, *Tonal Pitch Space*. New York: Oxford University Press, 2001.
- [7] F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press, 1983.
- [8] G. Mazzola, *The Topos of Music*. Basel: Birkhauser, 2003.
- [9] T. Noll and J. Garbers, “Harmonic path analysis,” in *Perspectives of Mathematical and Computational Music Theory*, G. Mazzola, T. Noll, and E. Luis-Puebla, Eds. Os-nabrück: epOs-Music, 2004.
- [10] L. K. Saul, K. Q. Weinberger, F. Sha, J. Ham, and D. D. Lee, “Spectral methods for dimensionality reduction,” in *Semi-Supervised Learning*, O. Chappelle, B. Schölkopf, and A. Zien, Eds. Cambridge, MA: MIT Press, in press.
- [11] D. Temperley, *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press, 2001.
- [12] G. Weber, *Versuch einer geordneten Theorie der Tonsetzkunst*. Mainz: B. Schotts Söhne, 1821–24.