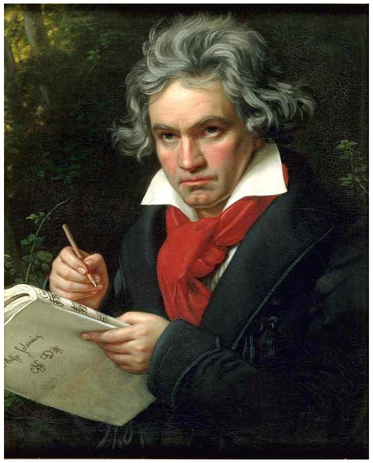
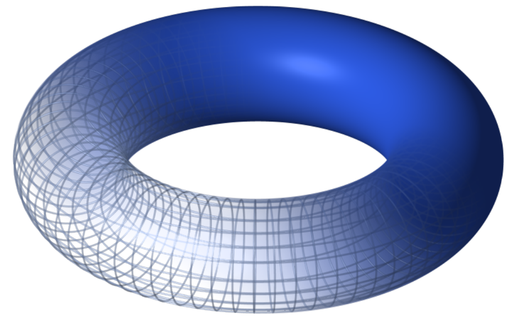


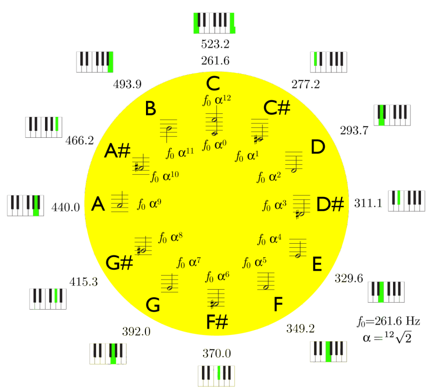
Topology of Musical Data



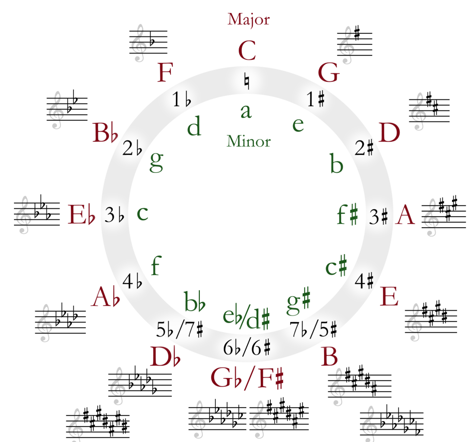
Techniques for discovering topological structures in large data sets are now becoming practical. This talk argues why the musical realm is a particularly promising arena in which to expect to find nontrivial topological features. The analysis is able to recover three important topological features in music: the circle of notes, the circle of fifths, and the rhythmic repetition of timelines, often pictured in the necklace notation.

Why Search for Topological Features in Music?

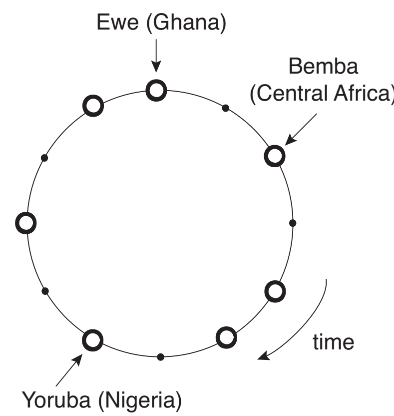
Because there is good reason to expect them to exist!



The Circle of Notes



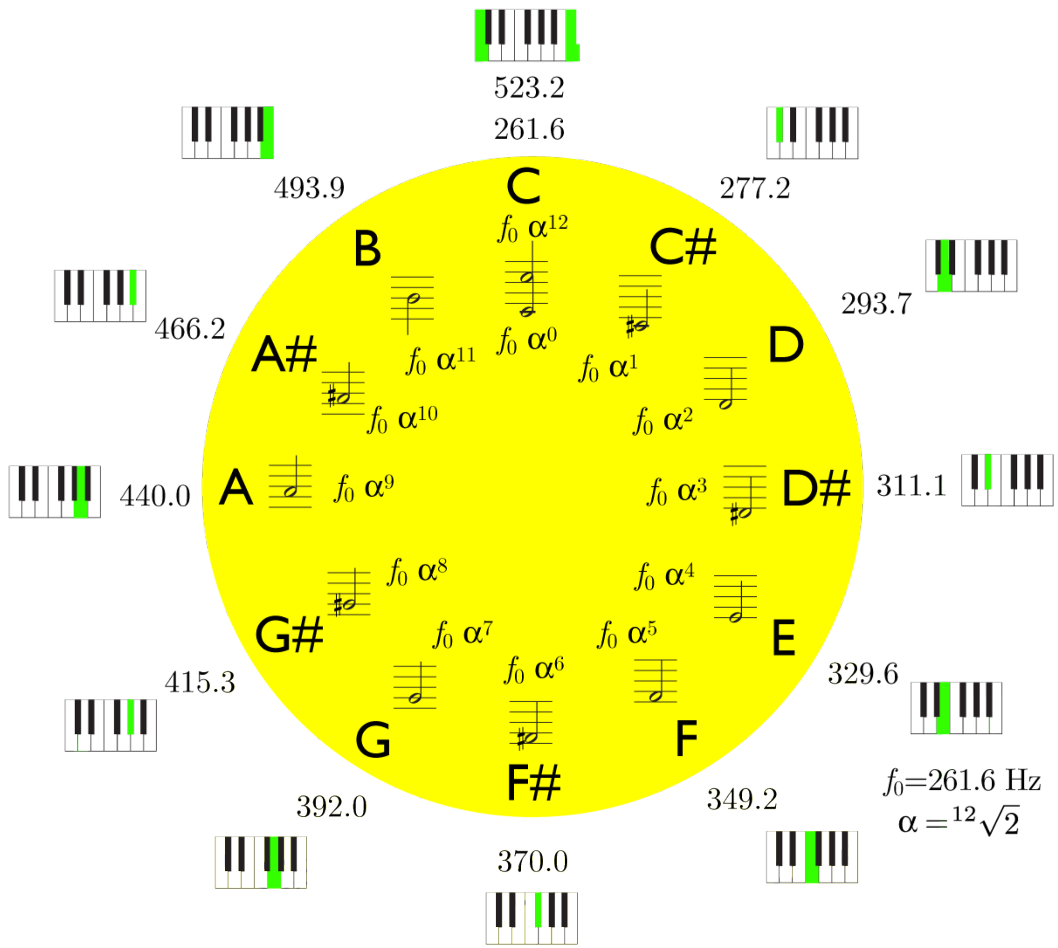
The Circle of Fifths



Rhythmic Cycles

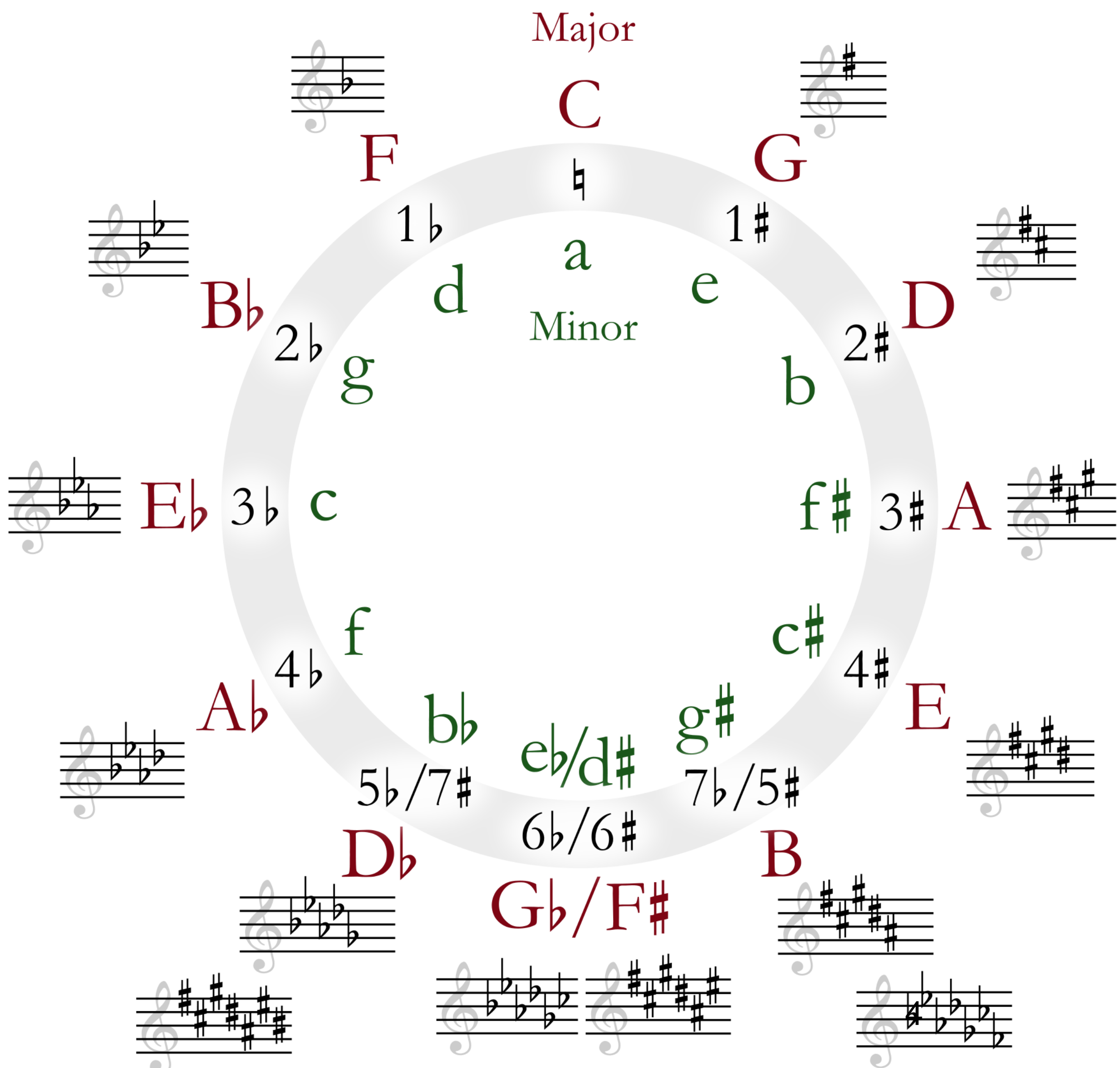
The Circle of Notes

There are two senses of "closeness" of tones (1) nearby in pitch or frequency, (2) having the same "chroma" or pitch class (equal but for a factor of two in frequency).



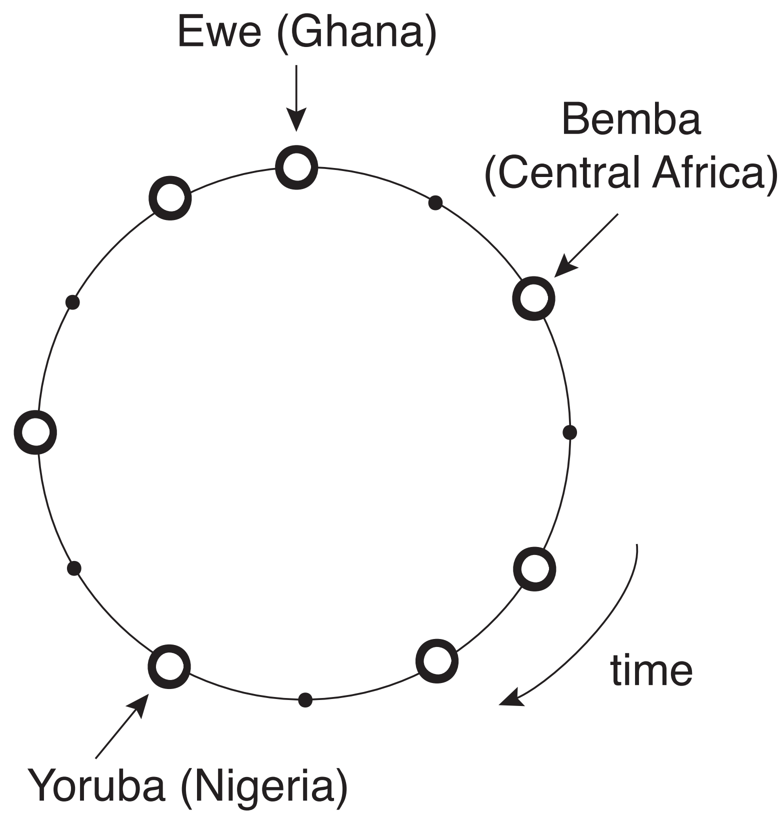
The Circle of Fifths

The circle of fifths shows relationships among the tones of the chromatic scale, standard key signatures (i.e., numbers of sharps and flats), and the major and minor keys.



Rhythmic Cycles

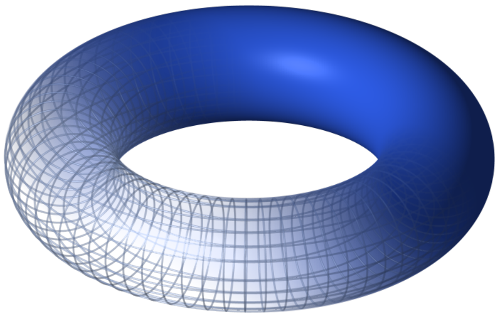
Time moves around the circle and events are depicted along the periphery. Since the "end" of the circle is also the "beginning," this emphasizes the repetition inherent in rhythmic patterns. Diagram depicts three traditional rhythmic variants of King's standard pattern in the "necklace notation."



Classic Algebraic Topology: Betti Numbers

A way of classifying surfaces: the Betti number β_k is the number of unconnected k-dimensional surfaces.

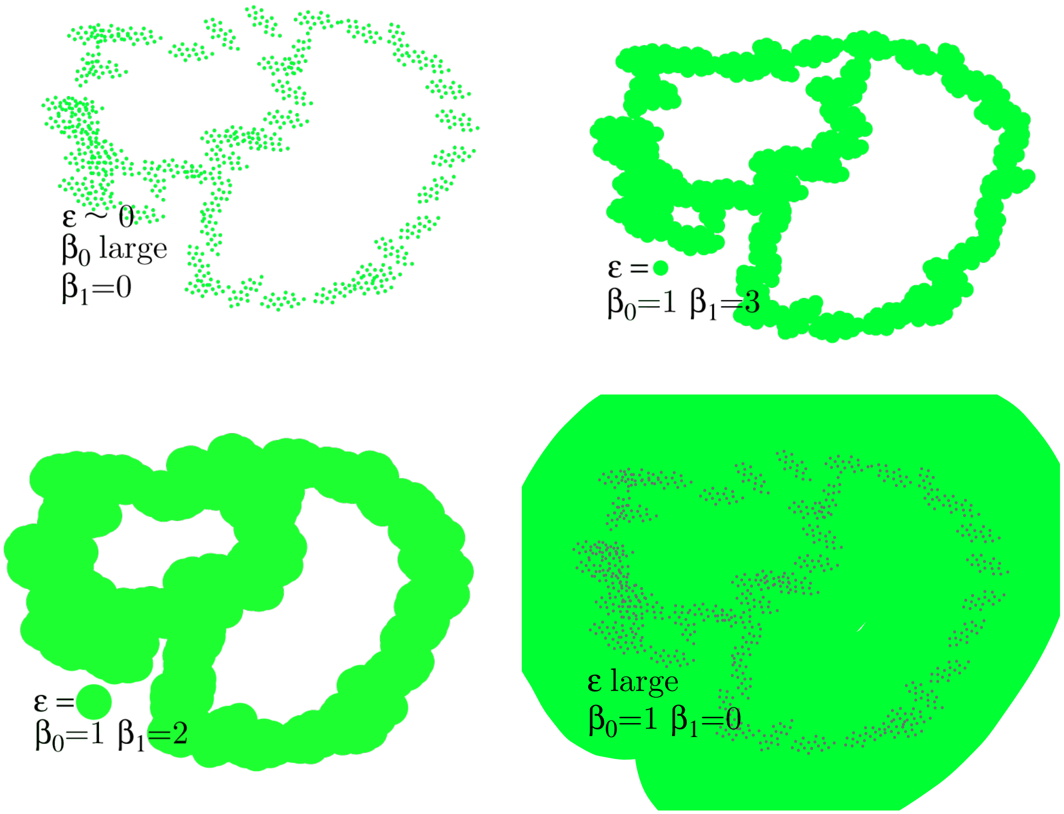
- β_0 is the number of connected components
- β_1 is the number of two-dimensional or "circular" holes
- β_2 is the number of three-dimensional holes or "voids"
- etc.



A torus has one connected component $\beta_0 = 1$, two circular holes $\beta_1 = 2$ (one in the center and one around the tube), and a three-dimensional void $\beta_2 = 1$ (the inside of the tube).

Topological Surfaces Defined by Data

The Betti numbers of a cloud of points depends on scale. For ϵ small, all points are separated from all other points, $\beta_0 =$ number of points. For ϵ large, all points are identified together and $\beta_0 = 1$. Interesting things happen in between small and large.



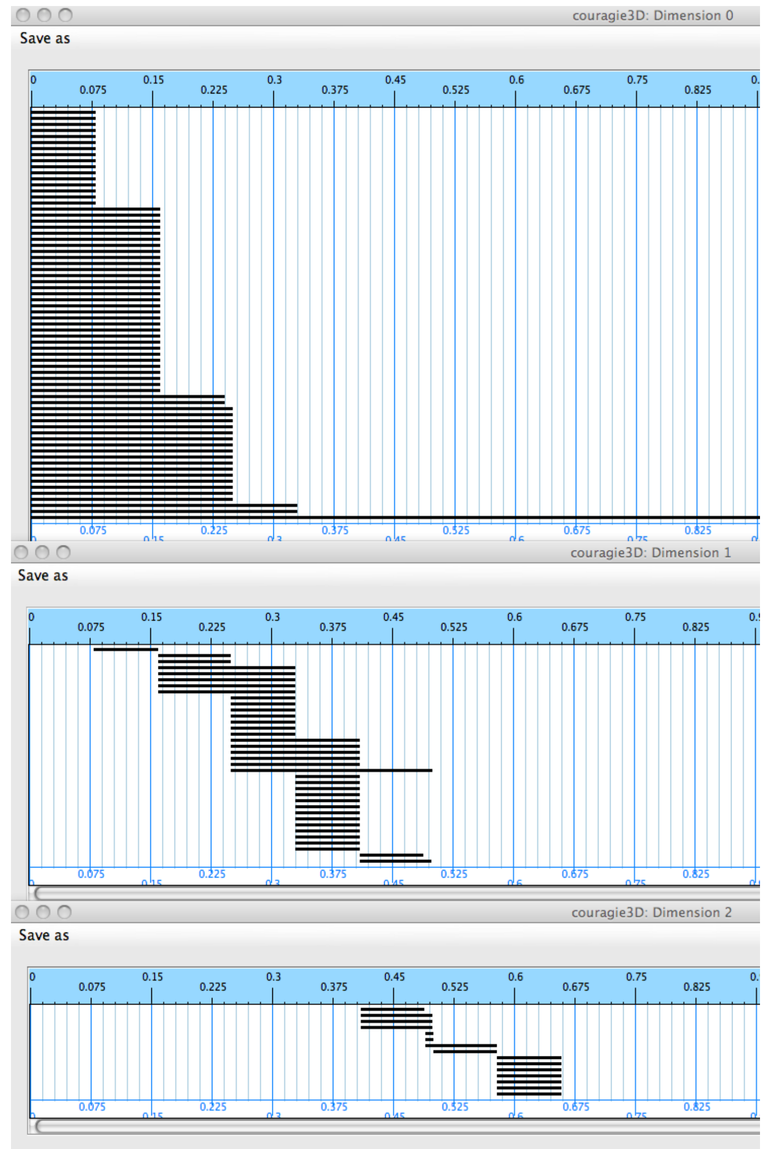
What Metric?

So far, the pictures have been drawn in $\mathbb{R}^2$, but the method does not require an embedding in $\mathbb{R}^n$, it just requires the ability to calculate distances.

The Plex software operates in two modes: in one mode the input is a collection of points in $\mathbb{R}^n$. In the second mode, the input is a matrix containing all distances between all points.

Bar Codes and Persistence

Carlsson's Plex software draws bar codes that show how the Betti numbers change as a function of scale ϵ. These show β_0, β_1, and β_2: features which persist over a range of ϵ are called persistent, and may reflect some underlying structure in the data.



Defining Distance Between Tones

The pitches of musical tones are generally perceived as a function of frequency in a logarithmic fashion. A metric like $|\log_2(f) - \log_2(g)|$ captures this along with the notion that nearby tones on the circle of notes should have a small numerical distance. To capture the idea that the "same" note recurs a factor of two apart in frequency, consider

$$s = \mathrm{mod}(|\log_2(f) - \log_2(g)|, 1).$$

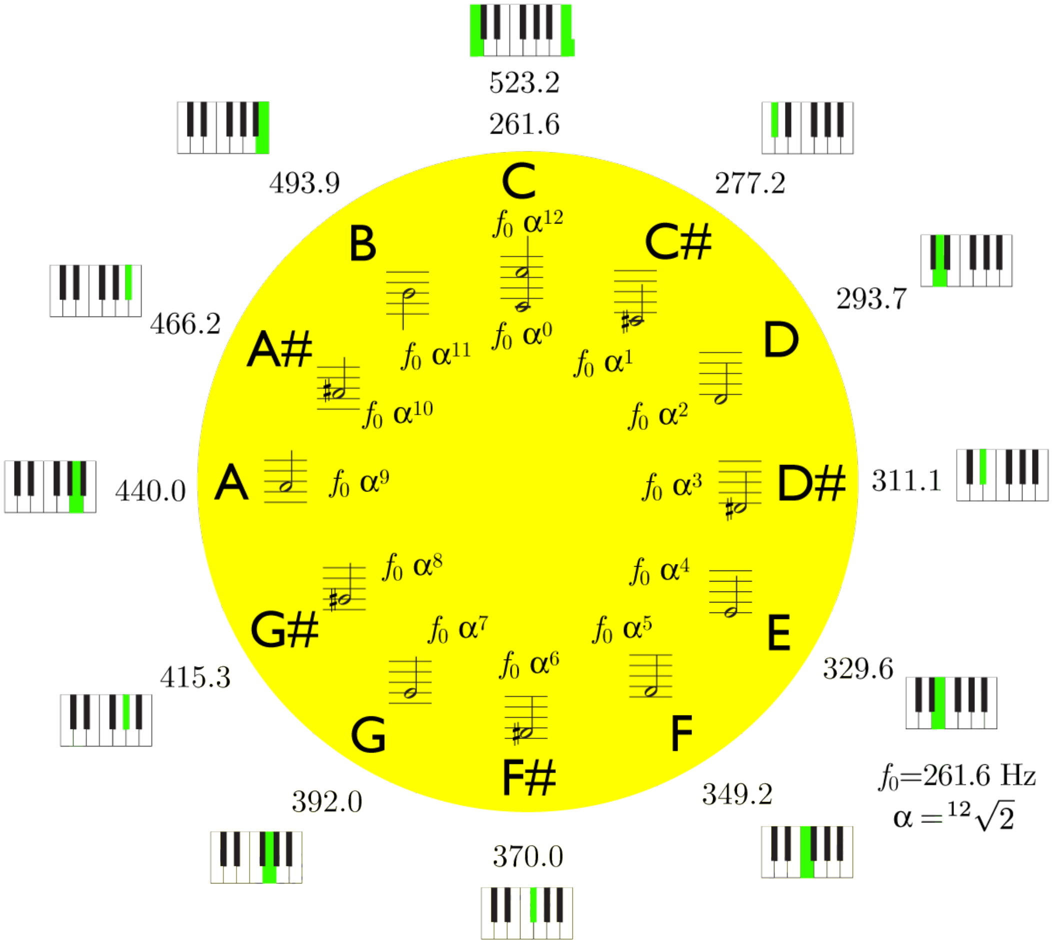
But this fails the triangle inequality. Instead, define the distance between two notes with fundamental frequencies f and g as

$$d(f, g) = \min(s, 1 - s).$$

This measures the distance between "pitch classes" and we call it the pitch class metric.

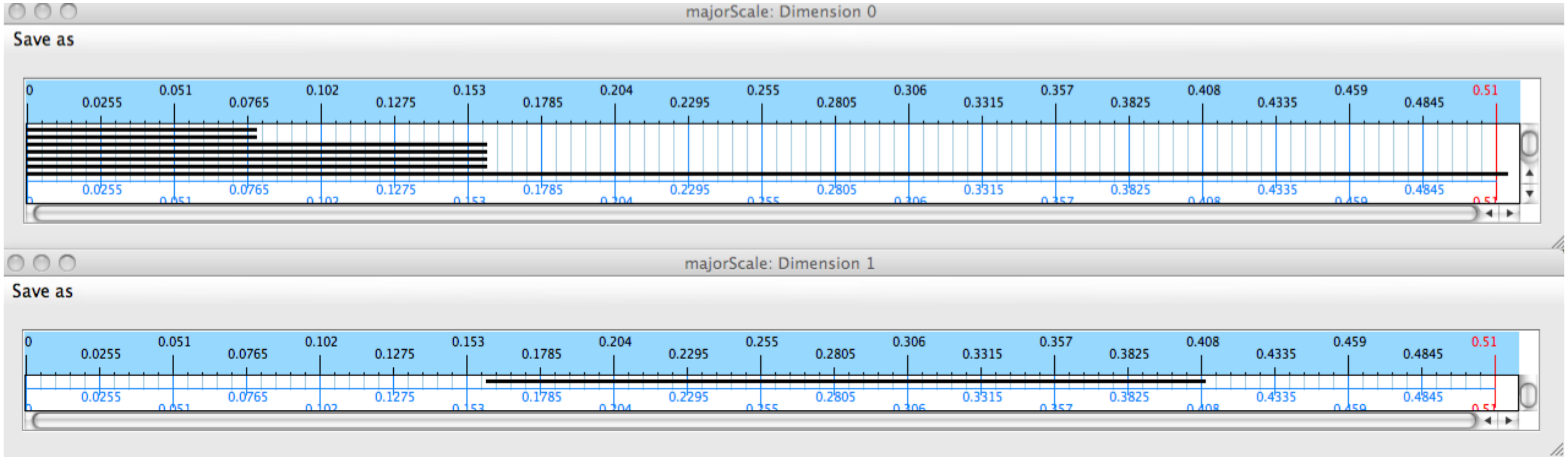
The Circle of Notes

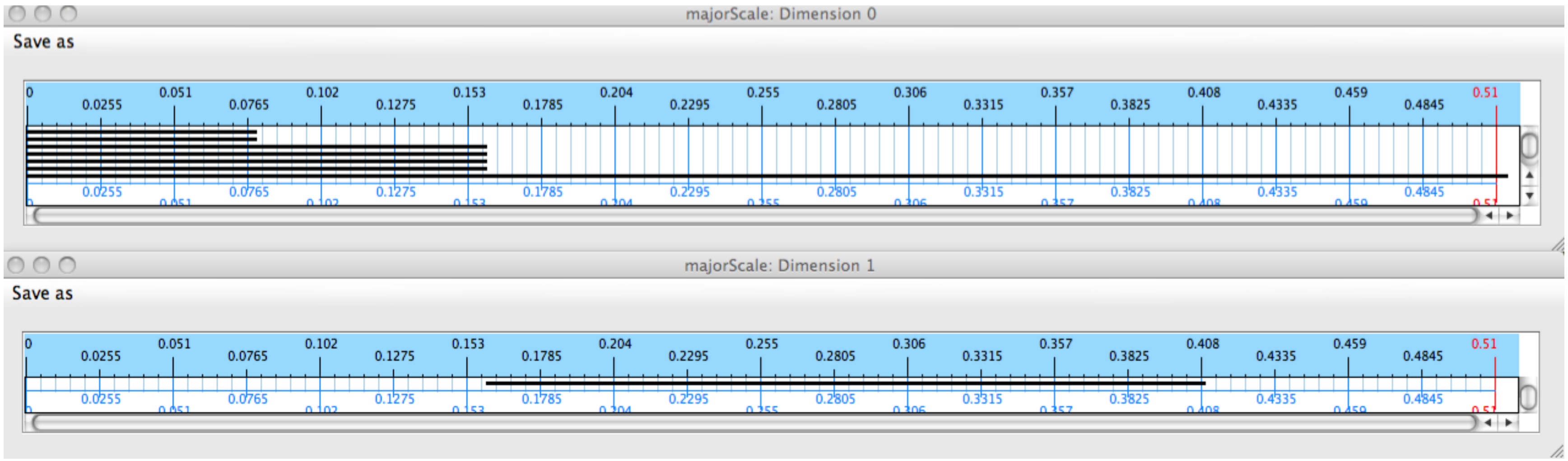
There are two senses of "closeness" of tones (1) nearby in pitch or frequency, (2) having the same "chroma" or pitch class (equal but for a factor of two in frequency).



Betti Numbers for the Major Scale

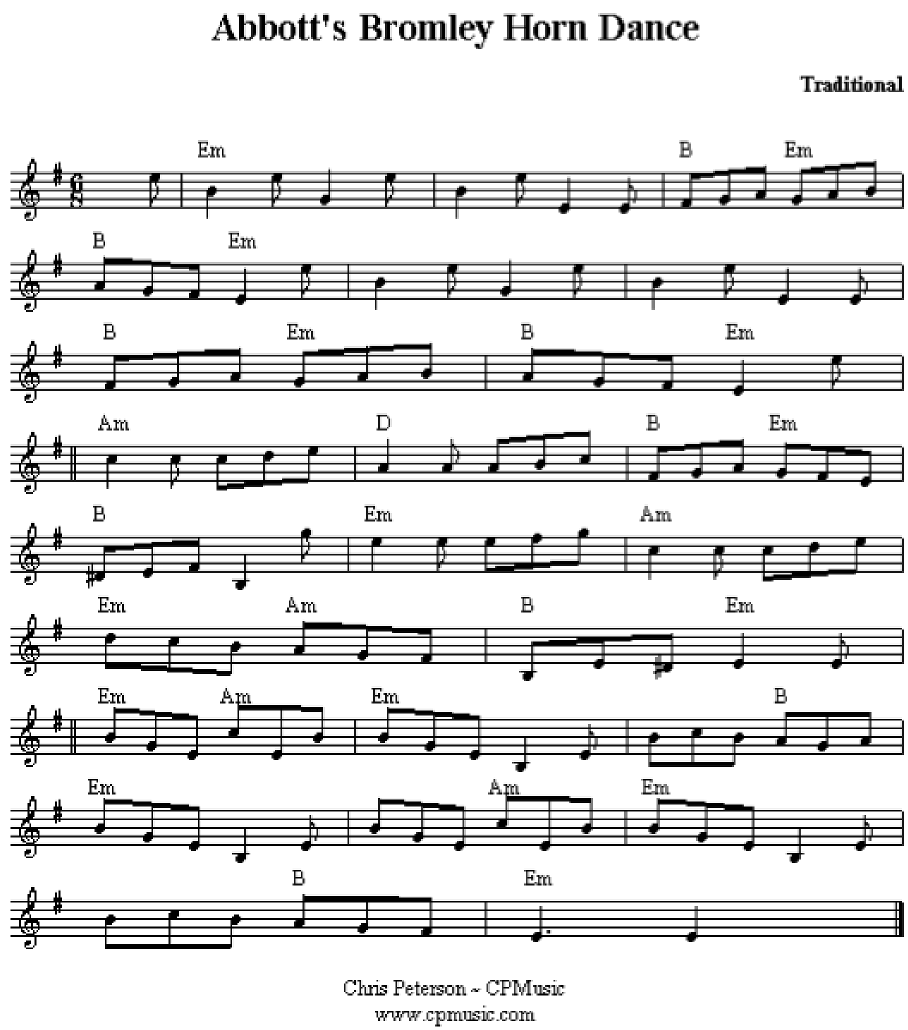
Consider the major scale made of the eight notes C, D, E, F, G, A, B, C with frequencies specified in the circle of notes. Calculating the distances between all pairs of notes (using the pitch-class metric) allows Plex to draw the barcodes. The persistent bar with $\beta_0 = \beta_1 = 1$ is the circle of notes!

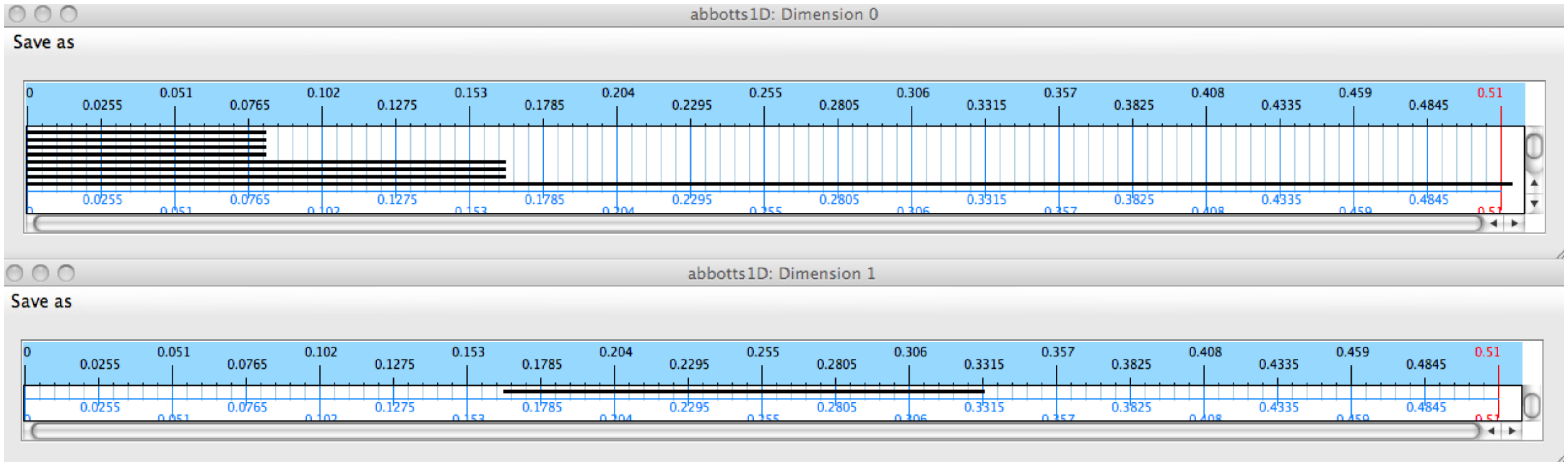




A more detailed interpretation: When ϵ is small, there are seven distinct notes. Though we input eight notes, the high C has exactly the same distances to all the other notes as the low C under the pitch-class metric, and thus the barcode merges these two tones even at $\epsilon = 0$. When ϵ reaches 0.08, the two half steps (the intervals between E-F and B-C) merge. When ϵ reaches 0.16, the five remaining connected components (all the major seconds) merge into one. Thus $\beta_0 = 1$ for all greater ϵ. At $\epsilon = 0.16$, the Dimension 1 code shows a single component, which persists until $\epsilon = 0.4$. This $\beta_0 = \beta_1 = 1$ feature is the circle of notes.

The major-scale example was built specifically with the circle of notes in mind, so it is perhaps unsurprising that the circle appears. Will such shapes appear in real music? Peterson's website contains a large selection of traditional melodies in both sheet music and standard MIDI files. Here's an example, and the corresponding barcodes are shown next.





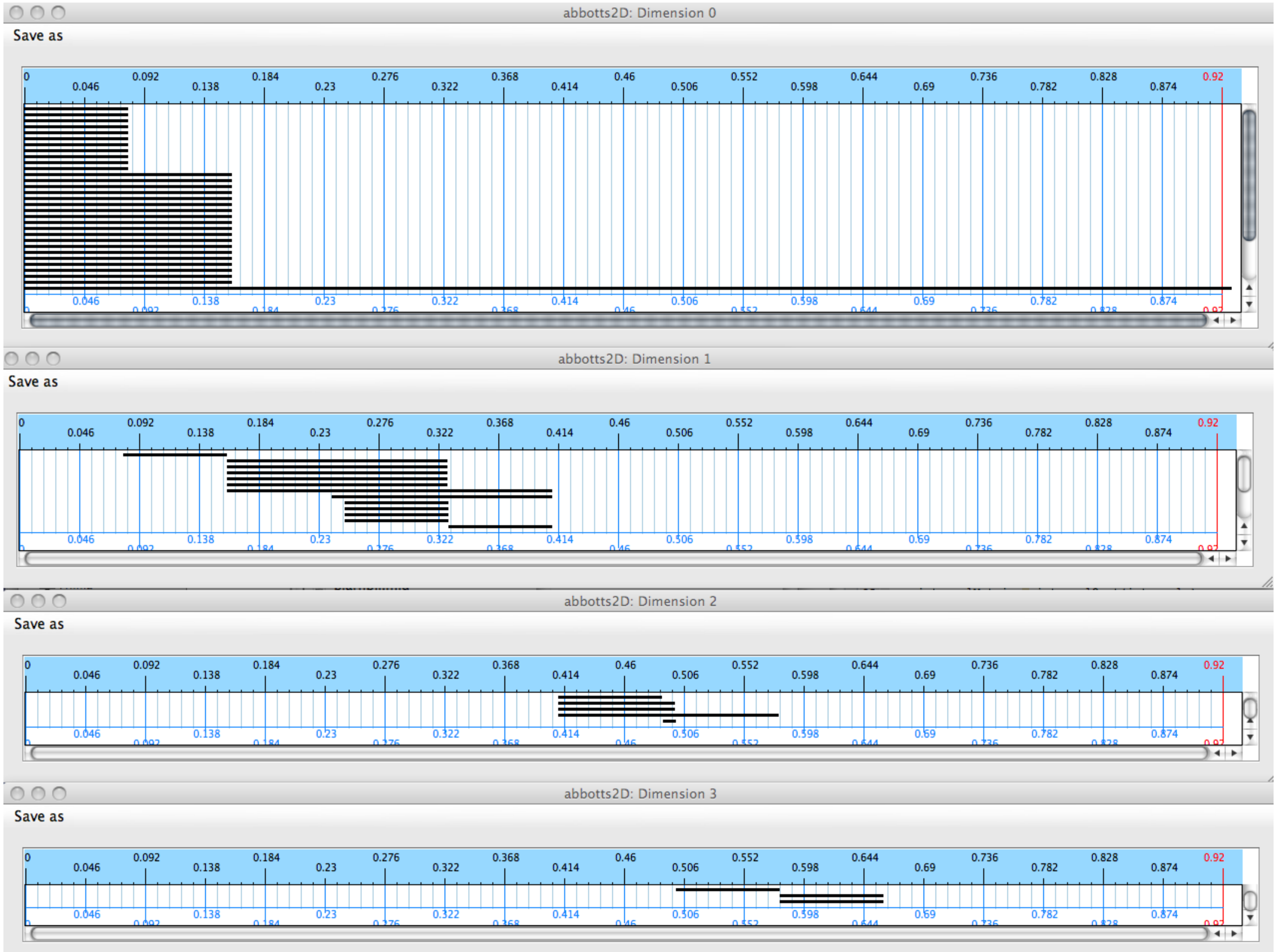
Interpreting Abbott's bar code: The top barcode shows eight lines, which correspond to the eight notes that appear in the score (observe again the insensitivity to octave). Four disappear at $\epsilon = 0.08$, which correspond to the four half steps ($F\sharp$-F, $D\sharp$-E, B-C, and D-$D\sharp$). Three more disappear at $\epsilon = 0.16$. Along with the constant bar, these correspond to the four whole steps (E-$F\sharp$, G-A, A-B, and C-D). All of these join into one bar for all larger ϵ. The region $0.16 < \epsilon < 0.33$ is characterized by $\beta_0 = 1$ (one connected component) and $\beta_1 = 1$. This is again the circle of notes.

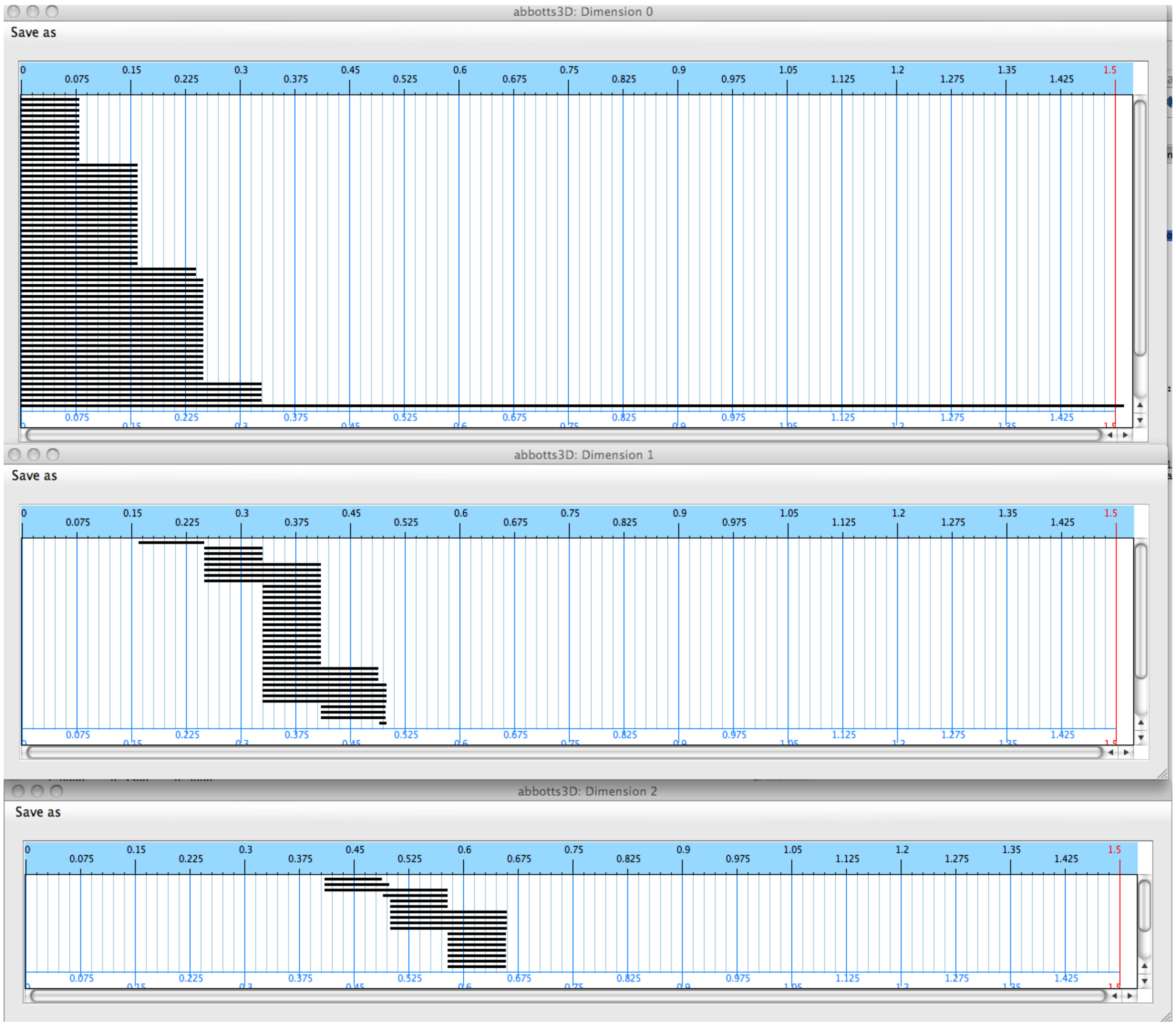
Incorporating Temporal Information

The above analyses may be somewhat naive because they suppress temporal information in the melody. This can be addressed using a time-delay embedding, which is common in time series analysis. Suppose that a melody consists of a sequence of notes with fundamentals at $f_1, f_2, f_3, f_4 \ldots$. These may be combined into pairs (a two-dimensional time-delay embedding) by forming

$$\begin{pmatrix} f_0 \\ f_1 \end{pmatrix}, \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, \begin{pmatrix} f_2 \\ f_3 \end{pmatrix}, \ldots$$

The distances between such pairs can be calculated by summing the distances between the notes elementwise using the pitch-class metric. These distances can now be used to form the barcodes.

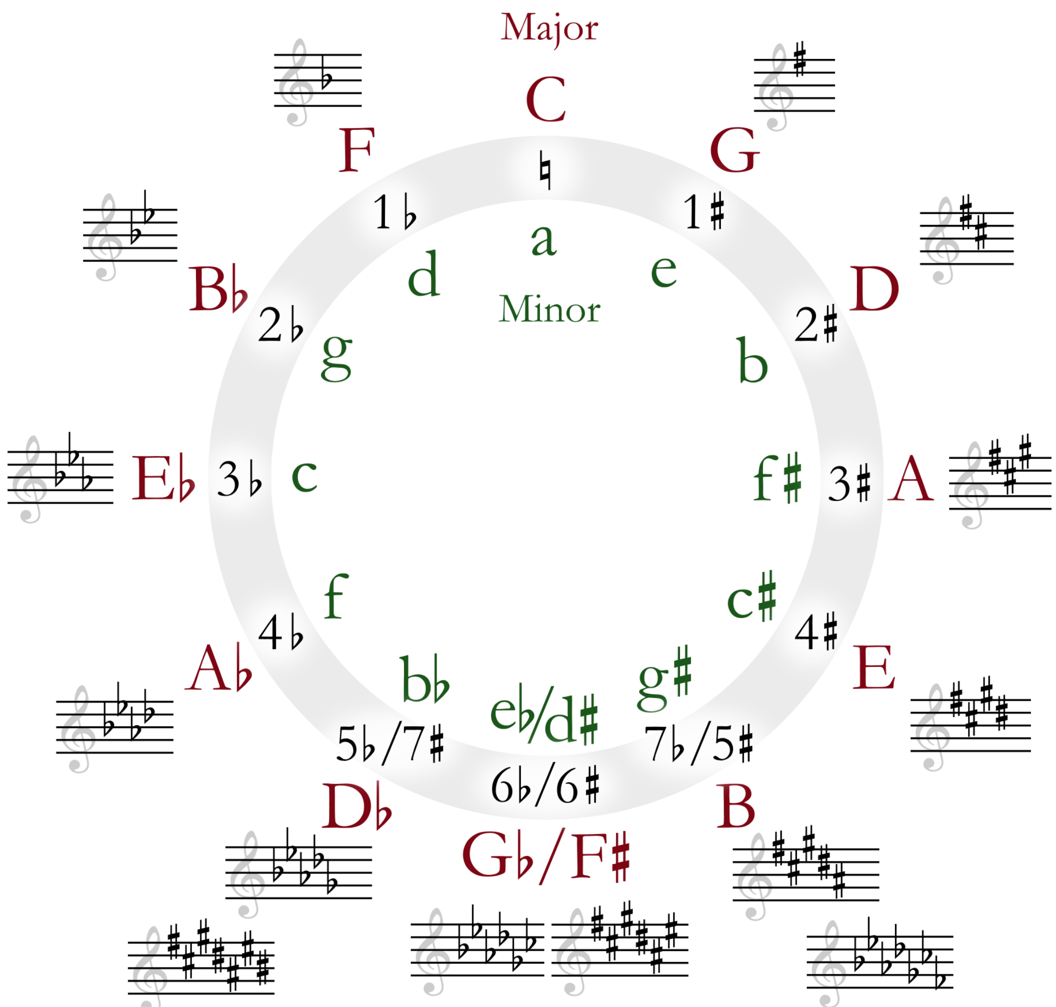

abbotts2D: Dimension 0
Save as
0 0.046 0.092 0.138 0.184 0.23 0.276 0.322 0.368 0.414 0.46 0.506 0.552 0.598 0.644 0.69 0.736 0.782 0.828 0.874 0.92
abbotts2D: Dimension 1
Save as
0 0.046 0.092 0.138 0.184 0.23 0.276 0.322 0.368 0.414 0.46 0.506 0.552 0.598 0.644 0.69 0.736 0.782 0.828 0.874 0.92
abbotts2D: Dimension 2
Save as
0 0.046 0.092 0.138 0.184 0.23 0.276 0.322 0.368 0.414 0.46 0.506 0.552 0.598 0.644 0.69 0.736 0.782 0.828 0.874 0.92
abbotts2D: Dimension 3
Save as
0 0.046 0.092 0.138 0.184 0.23 0.276 0.322 0.368 0.414 0.46 0.506 0.552 0.598 0.644 0.69 0.736 0.782 0.828 0.874 0.92



And using 3-D time delay embedding...

Harmonic Barcodes and the Circle of Fifths

The circle of fifths arises at the level of musical chords and scales, and so it is necessary to generalize the metric to consider multiple pitches simultaneously.



Generalizing the Pitch-Class Metric

Let $f = (f_1, f_2, ...f_n)$ and $g = (g_1, g_2, ...g_n)$ be two n-tuples, and define the distance

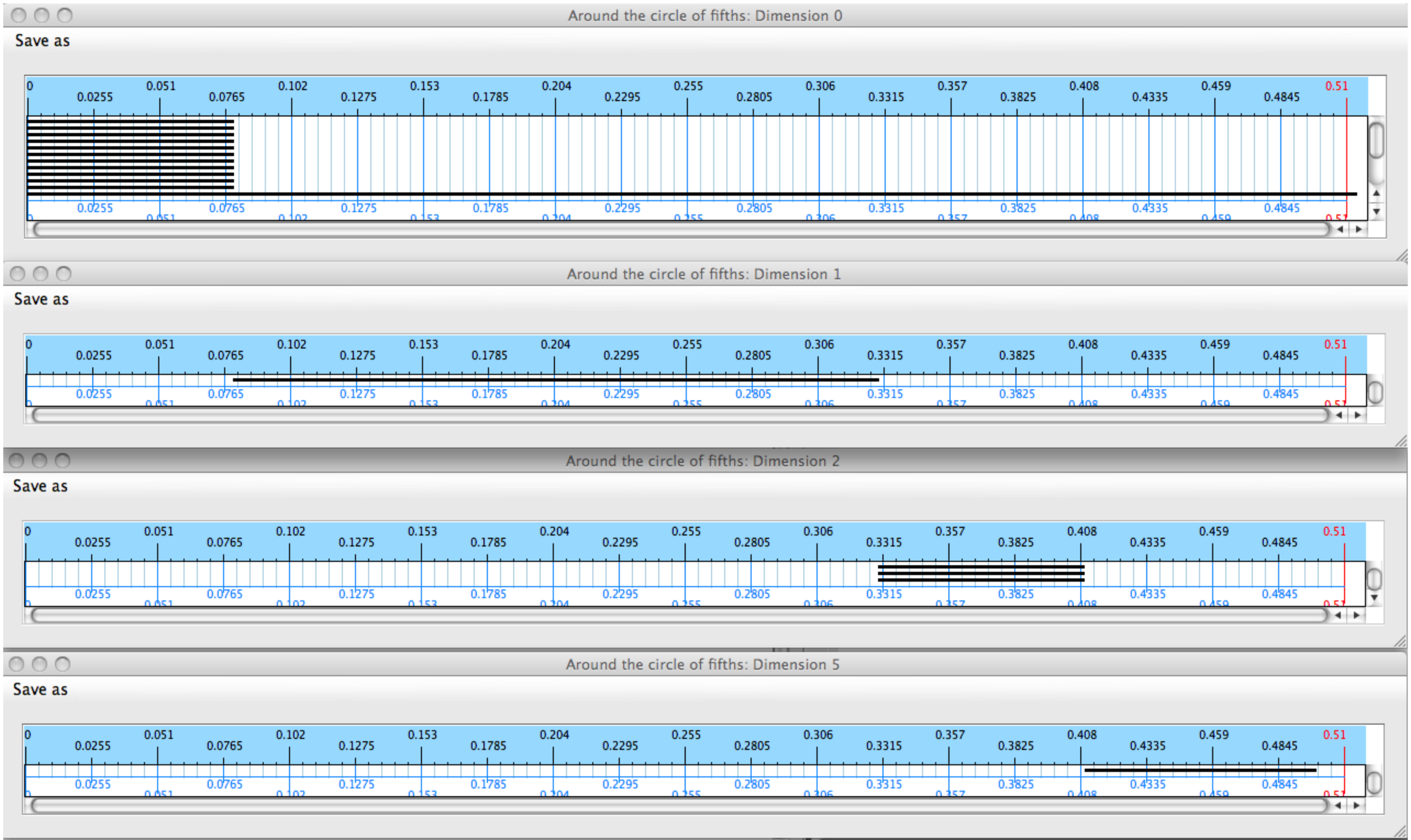
$$d_{cc}(f, g) = \min_P d(f, Pg)$$

where P ranges over all possible permutation matrices and where $d(\cdot, \cdot)$ is the pitch-class metric. This chord-class metric calculates the (elementwise) pitch-class distance between f and all the permutations of g. It is invariant with respect to chord and scale inversion; all reorderings of the elements of f and g are placed in the same equivalence class.

Finding the Circle of Fifths

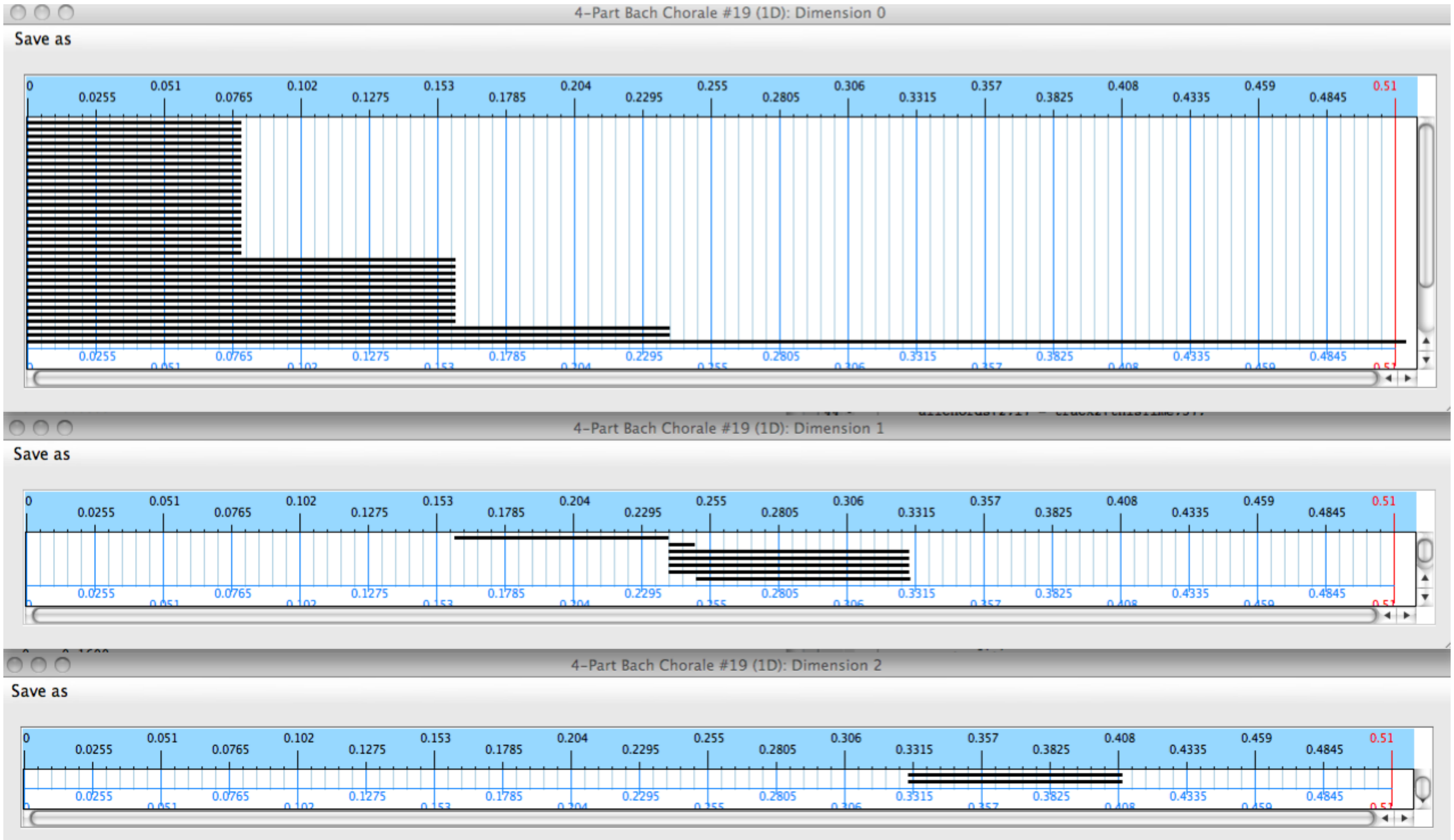
Consider a progression that moves around the circle of fifths: C major to G major to D major etc, all the way back to F and finally C. Inputting 12 such seven-note sets and calculating the barcodes gives the barcode on the next slide.

Scales that are a fifth apart (such as C major and G major) have a distance of 0.08 and this explains the twelve lines that merge down to a single connected set at $\epsilon = 0.08$ in the β_0 (top) plot. For $0.08 < \epsilon < 0.33$, the β_1 barcode shows a single persistent bar; this is the circle of fifths! (There are also some higher dimensional features for larger ϵ, but the exact meaning of these is not clear.)



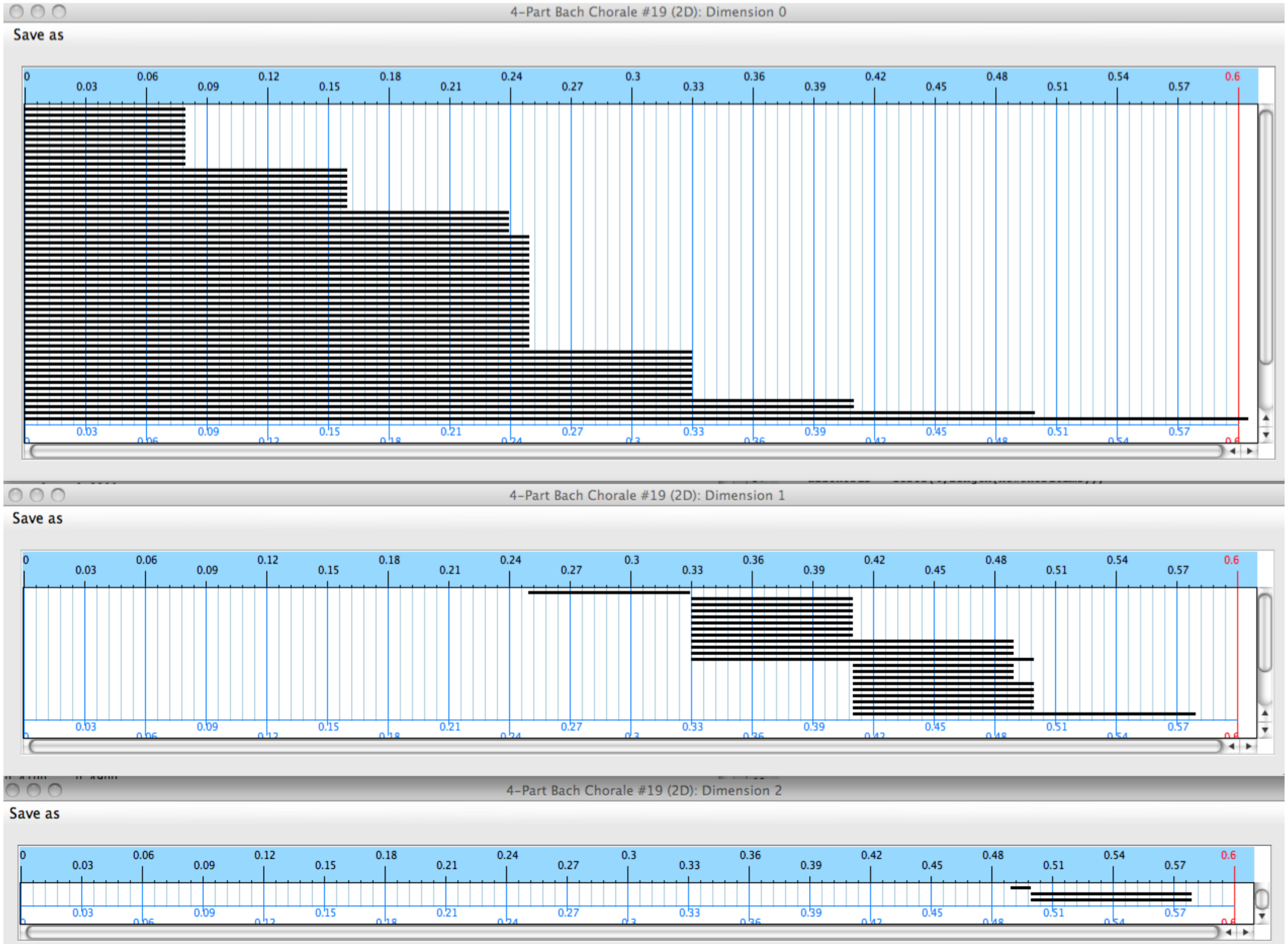
Around the circle of fifths: Dimension 0
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51
Around the circle of fifths: Dimension 1
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51
Around the circle of fifths: Dimension 2
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51
Around the circle of fifths: Dimension 5
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51

The progression around the circle of fifths example was built specifically with the circle of fifths in mind. Do such shapes appear in real music? The classical music archives website contains a large selection of Bach's chorales in standard MIDI file format. Chorale #19 is parsed to extract the four voices. The distances between all four-part chords are calculated according to the chord-class metric, and the results are used to draw the barcodes.


4-Part Bach Chorale #19 (1D): Dimension 0
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51
4-Part Bach Chorale #19 (1D): Dimension 1
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51
4-Part Bach Chorale #19 (1D): Dimension 2
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845 0.51

The β_0 barcode shows a large number of chords that are separated by $\epsilon = 0.08$, a somewhat smaller number of chords that are separated by a distance of $\epsilon = 0.16$, and two chords separated by $\epsilon = 0.23$. Above this value, all chords merge into one connected component.

The β_1 barcode shows one circle for $0.16 < \epsilon = 0.23$, and this structure then changes to $\beta_1 = 5$ for $0.24 < \epsilon = 0.33$. Features such as these appear to be unique identifiers of the particular pieces, meaning that other Bach Chorales from the same data set have different Betti numbers that occur over different ranges of ϵ. Finding the origin of such variations is an interesting challenge.

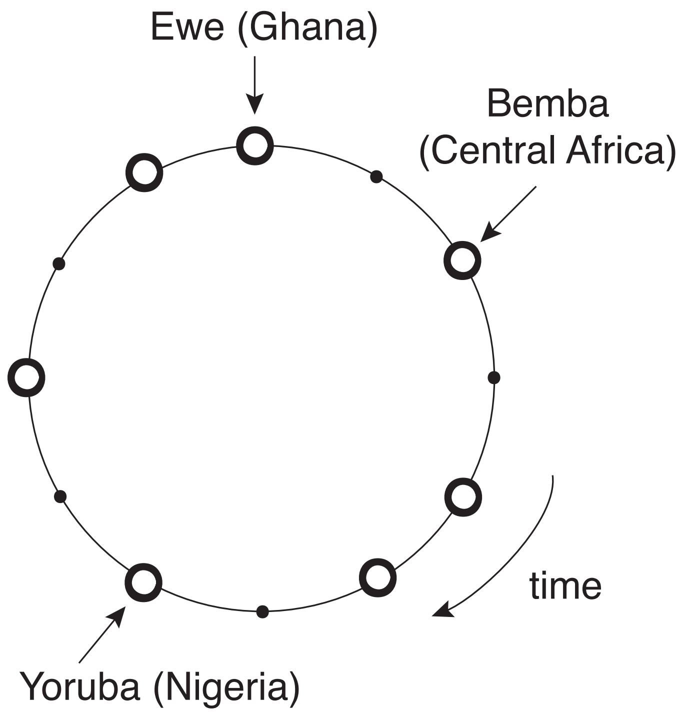

4-Part Bach Chorale #19 (2D): Dimension 0
Save as
4-Part Bach Chorale #19 (2D): Dimension 1
Save as
4-Part Bach Chorale #19 (2D): Dimension 2
Save as
0 0.03 0.06 0.09 0.12 0.15 0.18 0.21 0.24 0.27 0.3 0.33 0.36 0.39 0.42 0.45 0.48 0.51 0.54 0.57 0.6

Barcodes for Rhythmic Cycles

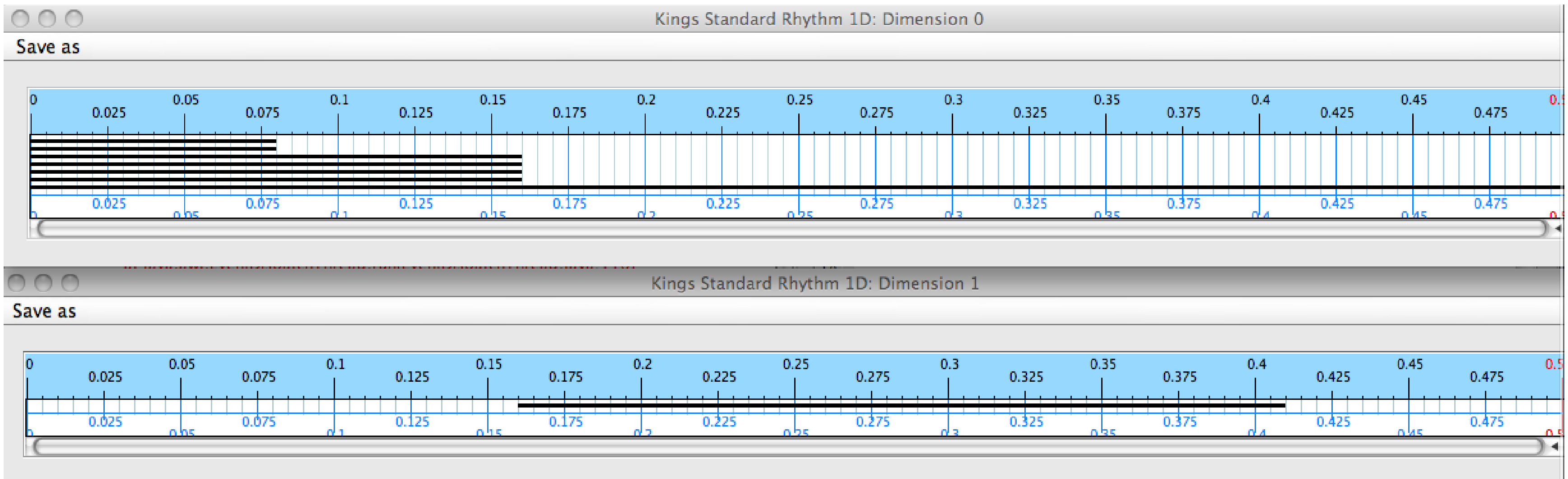
Rhythmic notations represent time via a spatial metaphor. The “Ewe” rhythmic variant is translated into the vector of time points

$$\{0, \frac{1}{6}, \frac{2}{6}, \frac{5}{12}, \frac{7}{12}, \frac{3}{4}, \frac{11}{12}\}$$

and the distance is calculated between all pairs



$d(f, g) = \min(s, 1 - s)$ where $s = \mathrm{mod}(|f - g|, 1)$.



Interpreting the rhythmic bar code: Barcodes for the Ewe variant of King's standard rhythm show the distribution of time intervals in the rhythm in the top plot and show the circular structure with $\beta_0 = \beta_1 = 1$ in the bottom.

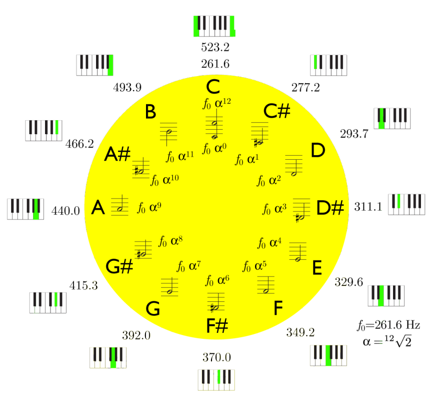
As might be expected, more complex rhythmic patterns and higher dimensional embeddings yield more complex barcodes. For instance, the rhythm of the first 4 measures (a 24 beat cycle) of “Abbott’s Bromley Horn Dance” is shown in the next slide. As usual, the distribution of short and long intervals is shown in the dimension 0 barcode while the circular structure of the rhythm appears in the dimension 1 barcode for $0.08 < \epsilon < 0.32$. There are also interesting features to this rhythm in the two and three dimensional barcodes.


Abbotts Rythhm 1D: Dimension 0
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845
Abbotts Rythhm 1D: Dimension 1
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845
Abbotts Rythhm 1D: Dimension 2
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845
Abbotts Rythhm 1D: Dimension 3
Save as
0 0.0255 0.051 0.0765 0.102 0.1275 0.153 0.1785 0.204 0.2295 0.255 0.2805 0.306 0.3315 0.357 0.3825 0.408 0.4335 0.459 0.4845

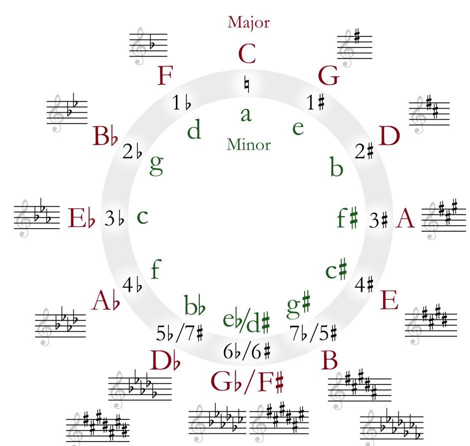
Summary

This talk argues that an investigation of the topological structures inherent in musical data is feasible using the ideas of persistent homology. Besides demonstrating that well known topological features can be derived from musical data sets, such analyses may be useful in information retrieval, in analysis of musical pieces, and in applications such as audio segmentation, melody recognition, and musical classification.

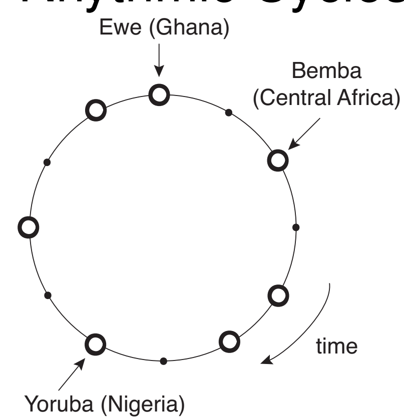
The Circle of Notes



The Circle of Fifths



Rhythmic Cycles



Some Questions...

- How to handle spectral data and audio `.wav` data?
- What kind of geometric shapes correspond to the higher level Betti numbers in the barcodes?
- How can these geometric shapes be interpreted in terms of the musical piece?
- There are other possible metrics... other ways of incorporating temporal information. How should the pitch and rhythmic analyses be combined?
- Rhythmic patterns often occur in hierarchies. Can the ideas of persistent homology locate such hierarchical structures from the musical data?

Some More Questions...

- Even if the high dimensional barcodes cannot be interpreted easily in terms of the musical pieces, they might be useful in segmentation.

- Even if the high dimensional barcodes cannot be interpreted easily, they might be useful in classification.

- It would be nice to have a metric that didn't require having the same number of elements in each term (so that 3 note chords could be measured against 4-note chords, for instance).

- Is it necessary that the distance be given between all pairs of points in the data set?