arXiv:1307.1201v2 [math.AT] 19 Oct 2013

RESEARCH ARTICLE

Topology of Musical Data

William A. Sethares[a] and Ryan Budney[b]

[a]*Department of Electrical and Computer Engineering, University of Wisconsin, Madison, USA, sethares@wisc.edu*; [b]*Mathematics and Statistics, University of Victoria, PO Box 3060 STN CSC, Victoria, B.C., Canada V8W 3R4, rybu@uvic.ca*

The musical realm is a promising area in which to expect to find nontrivial topological structures. This paper describes several kinds of metrics on musical data, and explores the implications of these metrics in two ways: via techniques of classical topology where the metric space of all-possible musical data can be described explicitly, and via modern data-driven ideas of persistent homology which calculates the Betti-number bar-codes of individual musical works. Both analyses are able to recover three well known topological structures in music: the circularity of octave-reduced musical scales, the circle of fifths, and the rhythmic repetition of timelines. Applications to a variety of musical works (for example, folk music in the form of standard MIDI files) are presented, and the bar codes show many interesting features. Examples show that individual pieces may span the complete space (in which case the classical and the data-driven analyses agree), or they may span only part of the space.

Keywords: persistent homology, topological structure, metrics on musical data, melodic analysis, rhythm analysis, circle of fifths

1. Introduction

Music has rich internal structures. Embedded in any visualization of musical structure is a notion of the closeness and/or similarity (or equivalently a notion of the distance and/or dissimilarity) between various musical elements. This paper considers the implications of several different distance-measures applied at the level of notes, of chords/scales, and of rhythms. Two themes emerge. Ideas of musical similarity tend to be compound, to express relationships more complex than are easily visualized in (linear) Euclidean space; each notion of closeness corresponds to a space with different topological structure. Second,

even (apparently) small changes in the definition of a distance function can lead to large changes in the global (topological) shape of the resulting space. The implication is that there is not really a single "master" space in which all musical phenomena can be placed. Rather, different notions of distance between different kinds of musical events can be legitimately described in various ways, as having their own kind of local and global structure.

A large literature shows many ways to visualize and analyze musical structures. To name a few examples: musical scores present notes in a two-dimensional array, Euler's classical *Tonnetz* represents musical intervals in a lattice as do modern generalizations [1], Partch [2] draws the tonality diamond where the two axes represent significant musical intervals, Chew [3] visualizes musical progressions along a spiral array, and computer-based music visualizers [4] display moving patterns in real time as music progresses. Lewin [5] suggests the application of simplicial complexes and homology to music theory and Mazzola [6] considers actions by the symmetric group on n-tuples. More recently, Callender, Quinn, and Tymoczko [7]-[8] explore a topological approach to musical spaces using the voice-leading metric; the result is a space they call an orbifold. Topological considerations have also been used in [9] and [10] to cluster melodies and so play a role in the discovery of musical motives.

Several different notions of distance applicable to musical situations are discussed in Section 2 and are intended to motivate both the mathematical analyses in Section 3. Section 3.5 reviews the idea of persistent homology, which is then applied to musical data in Section 4.

Throughout the paper, several interesting spaces are encountered, not all of which can be described as open subsets of $\mathbb{R}^n$. Each of the distance functions corresponds to a particular topological space, and these spaces can often be described concretely as products of basic spaces like the circle S^1 and the closed ball D^n. Occasionally, somewhat more elaborate spaces arise, such as twisted products (fibre bundles) and higher-dimensional spheres.

A collection of simulations examine a series of increasingly complex sets of musical data using the Betti numbers of persistent homology. Applications begin with simple note and chord sequences to verify that the method is capable of displaying common-sense musical intuition. In some of the examples, the song data spans much of the underlying topological space while in other examples the musical data occupies only a small part of the space. Both situations may be interesting. Finally, an online supplement [11] is available which contains additional figures and examples that augment the presentation within the paper.

2.   Distance Measures

One of the oldest visualizations of musical structure represents temporal cycles as spatial circles: Safî al-Din al-Urmawî, the 13th century theoretician from Baghdad, represents both musical and natural rhythms in a circular notation in the *Book of Cycles* [12]. Time moves around the circle (usually in a clockwise direction) 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. Anku [13] argues that African music is perceived in a circular (rather than linear) fashion that makes the necklace notation, shown here in Figure 1, particularly appropriate.
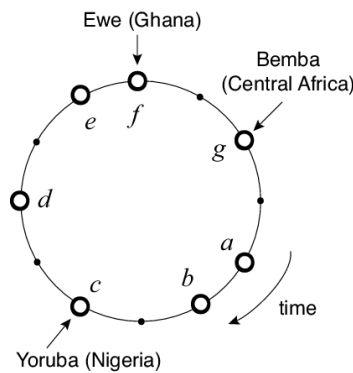
Figure 1. In the necklace notation, time progresses around the circle in a clockwise direction, marked off by a series of circles: small dots represent silent time points while open circles represent sounding events. Each passage around the circle represents one cycle through the rhythm. Traditional rhythms of the Ewe (from Ghana), the Yoruba (from Nigeria), and the Bemba (from Central Africa) are all variants of the "standard rhythm pattern" described by King [15], and can be represented in the necklace notation by the same pattern but with different starting points.

In this setting, musical events are sounds occurring at times specified by their position around the circumference of the circle. There are two different notions of the closeness of events in this representation. In one notion, two events are close if they occur near each other in time (for instance, events represented by the points labelled a and b are closer together than events represented by the points c and d. In the second notion, events are close if they occur at (or near) the same position in each cycle. Thus an event at a (which occurs the first time through the cycle) is identical to another event that occurs in a later cycle at a. With only the first notion of proximity in time, the natural representation would be a line segment $I^1 = [0, 1]$; with the repetitions of the cyclical pattern, the line is transformed into a circle S^1.

To be concrete, the distance between two events occurring at times f and g is

$$d_N(f, g) = \min(s, 1 - s) \text{ where } s = |f - g| \bmod 1, \tag{1}$$

and where one unit of time represents one period of the rhythm. Call this the *necklace distance*, indicated by the subscript N. The necklace metric on the space of repetitive rhythms is equivalent to the natural metric on the quotient space $S^1 = \mathbb{R}/\mathbb{Z}$. The slash notation indicates the conversion of the linear structure of the reals to a cyclic structure, by considering two numbers equivalent if they differ by an integer.

Music, of course, has aspects other than rhythm. The fundamentals of complex tones are generally perceived as a function of frequency in a logarithmic fashion. Thus the 15.6 Hz "distance" from C to $C\sharp$ is perceived to be the same size as the 24.7 Hz "distance" from $G\sharp$ to A (refer to Figure 2 for the origin of these numerical values). Accordingly, it is reasonable to consider a measure that operates on log frequency rather than on frequency itself. A metric like $|\log_2(f) - \log_2(g)|$ captures this. But there is also a second notion of closeness operating in the pitch domain, on the set of positive reals $\mathbb{R}_+$, that of octave equivalence. For example, the C note at 261.6 Hz and the high C at 523.2 Hz are closely identified. This is what happens when a man sings along with a woman (or when a woman sings along with a child): the "same" note is actually a factor of two apart in
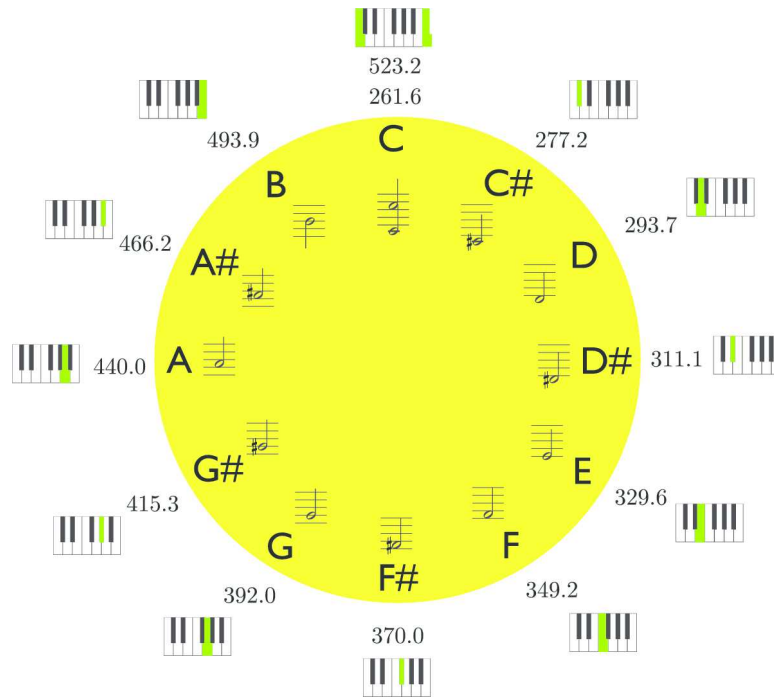
Figure 2. One way to picture two important aspects of musical perception is to place notes that are near each other in frequency (such as C and $C\sharp$) around the edge of a circle, and to overlay notes that are an octave apart (i.e, those which are distance zero apart in the pitch class metric (2)) such as low C and high C.

frequency. These two notions of nearness combine to suggest measuring distance between frequencies as

$$d_{PC}(f,g) = \min(s, 1-s) \text{ where } s = |\log_2(f) - \log_2(g)| \text{ mod } 1. \tag{2}$$

For example, the distance $d_{PC}(261.6Hz, 523.2Hz)$ between the two C notes an octave apart, is zero. This means that d_{PC} is not a metric (because metrics must satisfy the "identity of indiscernibles," that is, $d(x,y) = 0 \Leftrightarrow x = y$). But (2) does induce a metric on the quotient space determined by the equivalence relation $x \sim y \Leftrightarrow d_{PC}(x,y) = 0$. In musical terms, this may be interpreted as a measure of distance between "pitch classes" [16], since it identifies all Cs, all $C\sharp$s, etc. into equivalence classes and so (2) is called the *pitch-class* distance on $\mathbb{R}_+$, as indicated by the subscript PC. Again, two notions of closeness have combined to turn the space of pitch-class notes into the natural metric on the quotient space $S^1 = \mathbb{R}/\mathbb{Z}$.

Of course, music contains more than just isolated pitch classes and temporal cycles. One way to combine time and pitch information is with a time-delay embedding, which is a common practice in time series analysis. Suppose that a melody consists of a sequence of notes with fundamental frequencies at $f_1, f_2, f_3, f_4....$ These may be combined into pairs (a two-dimensional time-delay embedding) by creating the sequence of vectors (f_0, f_1), (f_1, f_2), (f_2, f_3), $(f_3, f_4)....$ The distances between such vectors can be calculated by adding the distances between the pitches element-wise using the pitch-class metric (2).

For example, the distance between pairs (f_n, f_{n+1}) and (f_m, f_{m+1}) is $d_{PC}(f_n, f_m) + d_{PC}(f_{n+1}, f_{m+1})$. The collection of all pairs of frequencies under this distance measure can be thought of as the torus $S^1 \times S^1$. Similarly, a d-dimensional time-delay embedding (where the f_i are gathered into vectors of size d) can be visualized as the d-fold product

$$\overbrace{S^1 \times S^1 \times ... \times S^1 \times S^1}^{d \ times}.$$

It may also be useful to generalize the distance measure to consider scalar harmony consisting of multiple pitches sounding simultaneously. The most straightforward generalization of the pitch-class metric is to add the (pitch-class) distances between all elements of the vectors, as was done for the time-delay embeddings. This distance measure distinguishes chord inversions: for instance, a C-major chord in root position (C-E-G) would be distant from a C-major chord in third position (G-C-E). While this may be desirable in some musical situations, it is undesirable when looking for structures that involve musical key, where (say) all C-major chords should be identified irrespective of inversion and all C major scales are identified irrespective of the order in which the pitches are listed. For example, the ascending C-major scale and the descending C-major scale are both the same entity, and the distance should reflect this realm of musical perception. Another example showing the shortcomings of (2) is when two voices are sounding simultaneously. The pairs (f_1, f_2) and (f_2, f_1) both mean that same two notes are heard. Using the pitch-class metric, these are unequal.

This can be addressed by permuting the elements and calculating the minimum over all the permutations. To be precise, let $f = (f_1, f_2, ... f_n)$ and $g = (g_1, g_2, ... g_n)$ be two n-tuples, and define the distance

$$d_{CC}(f, g) = \min_{\sigma} \sum_{i=1}^{n} d_{PC}(f_i, \sigma(g_i)) \tag{3}$$

where σ ranges over all possible permutations of size n, and $d_{PC}(\cdot, \cdot)$ is the pitch-class distance of (2). This *chord-class distance* is the smallest of the pitch-class distances between f and all permutations of g. Hence 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 (and have zero distance from each other). The quotient space induced by the chord-class distance is called the *symmetric product*. Specifically, given a space X, the n-fold symmetric product is denoted $Symm_n(X) = X^n / \Sigma_n$, and Σ_n is the symmetric group acting on X^n by permuting the coordinates of X^n. The chord-class distance induces a metric on $Symm_n(S^1)$, as described in Section 3. A further extension of these ideas combines a length j time-delay embedding of k-note chords. Describing the structure of these spaces concretely requires some work as in Section 3. Another complication is that there may not always be exactly n elements in each of the chords: for instance, one of the voices may become momentarily silent or two of the voices may converge to the same pitch. This can be handled by exploiting the n-th finite subset subspace, and leads to some of the more complex spaces encountered in the analysis.

As a further extension to rhythmic patterns such as those of Figure 1, consider extending the distance metric to one which measures the distance between complete rhythmic sequences (rather than the distance between individual timed events). In this case, an appropriate measure is one which considers any two sequences equal if they are circular

W. A. Sethares and R. Budney

shifts of each other. Let f and g be n-vectors that define the times of events in a cycle with period 1. Then the *rhythm distance* between f and g is

$$d_R(f,g) = \min_{\sigma_c} \sum_{i=1}^{n} d_N(f, \sigma_c(g)), \qquad (4)$$

where σ_c ranges over all cyclic shifts and $d_N(\cdot, \cdot)$ is the distance defined in (1). Under this measure, all three of the named rhythms in Figure 1 are the same since they are cyclical shifts of each other. In this case, the topology will be given by the subsymmetric product $Symm_n^G(X) = X^n/G$, which is the quotient of the n-times product of the circle $X = S^1$ and the subgroup $G \subset \Sigma_n$ of circular shifts.

Finally, a standard metric that can be applied to finite subset spaces is the *Hausdorff distance* which defines the distance between two finite subsets of a metric space $A, B \subset X$ as

$$d_H(A, B) = \max\{\max_{a \in A}\{\min_{b \in B} d(a, b)\}\}, \max_{b \in B}\{\min_{a \in A} d(a, b)\}\}.$$

3. Spaces and Homology

This section provides the tools needed to interpret the examples in Section 4, which analyze musical data using persistent homology. Sections 3.1, 3.2, and 3.3 describe auxiliary *natural* spaces on which it is possible to analyze the input (musical) data. These spaces admit their own notion of (metric) geometry compatible with Section 2, and it is through the lens of these geometries that persistent homology will be applied. The use of such auxiliary geometries is a commonly-occurring theme in topology. For example, [17] gives a detailed discussion and connections to subjects such as the *Shape Theory* of [18]. Section 3.4 summarizes the basics of homology in and Section 3.5 considers persistent homology. The examples of Section 4 show that the persistent homology of data is influenced by the topological features of the ambient metric spaces in which they live.

3.1. Products, Symmetric Products and Bundles

Given a pair of topological spaces X and Y, the *product* of the two is denoted $X \times Y = \{(x, y) : x \in X, \text{ and } y \in Y\}$. For example, if S^1 is the circle, $S^1 \times S^1$ is called a *torus*. Similarly, $S^1 \times \{1, 2\}$ is *two circles*. The n-fold product of X, denoted X^n, is the set of all n-tuples of elements of X. Thus $X^n = \{(x_1, \cdots, x_n) : x_i \in X \; \forall i \in \{1, 2, \cdots, n\}\}$. Observe that elements of X^n consist of n ordered elements of X. For example, $(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)$ if and only if $x_1 = y_1, x_2 = y_2, \cdots, x_{n-1} = y_{n-1}$ and $x_n = y_n$.

The n-fold *symmetric product* of X is the quotient of X^n, by the action of the *symmetric group* Σ_n, that is, $Symm_n(X) = X^n/\Sigma_n$ where $\Sigma_n \equiv \Sigma\{1, 2, \cdots, n\}$ is the set of all one-to-one onto functions $\{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\}$ acting on X^n on the right. Elements of $Symm_n(X)$ consist of n unordered elements of X and so $(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)$ in $Symm_n(X)$ if and only if there is a permutation $\sigma \in \Sigma_n$ such that $x_i = y_{\sigma(i)}$ for all $i \in \{1, 2, \cdots, n\}$. Given an arbitrary subgroup

$G \subset \Sigma_n$, the *sub-symmetric product* $Symm_n^G(X)$ is the quotient $Symm_n^G(X) = X^n/G$. In $Symm_n^G(X)$, $(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)$ if and only if there is a permutation $\sigma \in G$ such that $x_i = y_{\sigma(i)}$ for all $i \in \{1, 2, \cdots, n\}$.

A frequently-useful variant on the idea of the product of two spaces is the notion of a *fibre bundle*. A fibre-bundle is a continuous function $p : E \to B$ which is 'locally trivial' in the sense that for every $x \in B$ there is a neighbourhood $U \subset B$ and a homeomorphism $\phi : p^{-1}(U) \to U \times F$ where $F = p^{-1}(x)$. F is called 'the fibre over x' of the fibre bundle. The map ϕ must satisfy that $\pi \circ \phi = p$, where $\pi : U \times F \to U$ is the projection map onto the first coordinate $\pi(u, f) = u$.

A product $X \times Y$ is a fibre bundle with $E = X \times Y$, $B = X$ and $F = Y$, with $p : X \times Y \to X$ given by $p(x, y) = x$. One should think of a fibre bundle as a space that is locally a product in that the map ϕ gives a way of locally identifying nearby fibres. The Möbius band is an example of one such space, as we will see in Section 3.2. A fibre bundle with fibre a disc is called a *disc bundle*.

3.2.   Morton's Bundles

A classical result of Morton [19] states that $Symm_n(S^1)$ is a disc bundle over S^1 which is orientable if and only if n is odd. This means that when n is odd, the local homeomorphisms ϕ from the definition of a bundle assemble to give a homeomorphism $Symm_n(S^1) \cong S^1 \times D^{n-1}$. The space $Symm_2(S^1)$ is a Möbius band, and when n is even, $Symm_n(S^1)$ is similar to a Möbius band in that the local homeomorphisms of the fibres assemble into a mirror reflection of the fibre.

Morton gives a general prescription for studying sub-symmetric products $Symm_n^G(S^1)$. The basic idea is to consider the equivalence of S^1 and $\mathbb{R}/\mathbb{Z}$. This makes it possible to think of $(S^1)^n/G$ as $\mathbb{R}^n/(\mathbb{Z}^n \rtimes G)$ where $\mathbb{Z}^n \rtimes G$ denotes the semi-direct product[1] of G with $\mathbb{Z}^n$, equivalently it is denoted $\mathbb{Z} \wr G$ and denotes the wreath product[2] of $\mathbb{Z}$ with G. Complex multiplication gives a G-equivariant map $(S^1)^n \to S^1$ where the action of G on the target S^1 is trivial. Similarly, G acts on $\mathbb{R}^n$ via the inclusion $G \hookrightarrow \Sigma_n$ making $\mathbb{R}^n \to (\mathbb{R}/\mathbb{Z})^n$ G-equivariant. Denoting the kernel of the addition function $+ : \mathbb{R}^n \to \mathbb{R}$ by $\Delta_{\mathbb{R}} = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 0\}$ and $\Delta_{\mathbb{Z}} = \mathbb{Z}^n \cap \Delta_{\mathbb{R}}$, this means that there is a fibre bundle

$$\Delta_{\mathbb{R}}/(\Delta_{\mathbb{Z}} \rtimes G) \to \mathbb{R}^n/(\mathbb{Z}^n \rtimes G) \to \mathbb{R}/\mathbb{Z}.$$

Moreover, Morton identifies the monodromy of the bundle.

THEOREM 3.1   *The space* $Symm_n^G(S^1)$ *is a fibre bundle over* S^1, *with fibre* $\Delta_{\mathbb{R}}/(\Delta_{\mathbb{Z}} \rtimes G)$ *and monodromy induced by the map* $(t_1, t_2, \cdots, t_n) \longmapsto (t_1 + \frac{n-1}{n}, t_2 - \frac{1}{n}, \cdots, t_n - \frac{1}{n})$ *of* $\Delta_{\mathbb{R}}$.

The other main result in Morton's paper is the identification of the quotient $\Delta_{\mathbb{R}}/(\Delta_{\mathbb{Z}} \rtimes G)$ as a disc in the case $G = \Sigma_n$. A key argument is that the region $\{(t_1, \cdots, t_n) \in \Delta_{\mathbb{R}} :$

---

[1] Given a group G and a homomorphism $\phi : G \to Aut(H)$ the semi-direct product $H \rtimes_\phi G$ is defined as pairs (h, g) with multiplication defined as $(h_1, g_1) \cdot (h_2, g_2) = (h_1 \phi_{g_1}(h_2), g_1 g_2)$.

[2] The wreath product $\mathbb{Z} \wr G$ is a synonym for $\mathbb{Z}^n \rtimes G$ where G acts on $\mathbb{Z}^n$ via the regular representation $G \subset \Sigma_n$.

W. A. Sethares and R. Budney

$t_1 \leq t_2 \leq \cdots \leq t_n \leq t_1 + 1\}$ is a fundamental domain for the action of $\Delta_{\mathbb{Z}} \rtimes \Sigma_n$ on $\Delta_{\mathbb{R}}$. Since there are no identifications on the boundary, this *is* the fibre. It's an exercise to check the monodromy is orientation preserving if and only if n is odd. There are only two disc bundles over a circle: the trivial product bundle $D^{n-1} \times S^1$ and the Möbius bundle $D^{n-1} \rtimes S^1$. The 'center' (0-section) of the disc bundle consists of the equally-spaced points $(t_1, t_1 + \frac{1}{n}, t_1 + \frac{2}{n}, \cdots, t_1 + \frac{n-1}{n})$.

Depending on the group G, $Symm_n^G(S^1)$ fibers over a product of different numbers of circles. As will become clear, the fibres can be discs, spheres, and more varied spaces. The space $Symm_n^G(S^1)$ can be described explicitly, for all G with $n \leq 4$, and for a wide variety of 'useful' groups G with arbitrary n.

Example 3.2 For G the trivial group, $Symm_n^G(S^1) = (S^1)^n$. This is the space of ordered n-tuples of octave-reduced notes. For $n = 2$, this is the torus $S^1 \times S^1$.

Example 3.3 The space $Symm_n^{\Sigma_{n-1}}(S^1) \cong S^1 \times Symm_{n-1}^{\Sigma_{n-1}}(S^1)$ is a product of S^1 and the lower-dimensional symmetric product. This is the space of ordered pairs, where the first element is an octave-reduced note, and the 2nd element is an unrelated $(n-1)$-note chord. More generally, $Symm_n^{\Sigma_i \times \Sigma_j}(S^1) \cong Symm_i^{\Sigma_i}(S^1) \times Symm_j^{\Sigma_j}(S^1)$ provided $i + j = n$.

Example 3.4 The space of n-note chords (under chord-class distance measure (3)) is the symmetric product of the circle, i.e., $Symm_n(S^1) = (S^1)^n/\Sigma_n$. Morton's Theorem 3.1 shows that $Symm_n(S^1)$ fibers over S^1 with fiber a disc, and the bundle is orientable precisely when n is odd. There are only two disc bundles over a circle, the trivial (therefore orientable) one, and the non-orientable one. The first is a product $S^1 \times D^{n-1}$ while the latter can be thought of as a generalized Möbius band, $S^1 \ltimes D^{n-1}$, and in the case $n = 2$ it is the Möbius band.

Example 3.5 For $G = A_3 = \mathbb{Z}_3$, the space $Symm_3^G(S^1)$ represents the 3-note melody space, i.e. the space of triples of octave-reduced notes, where we only remember the cyclic ordering of the notes. We show in Section 3.5 this space is homeomorphic to $S^1 \times S^2$.

Example 3.6 The 4-part voice-leading measure of Callender et. al. [7]-[8] is equivalent to the chord-class distance function (3) and so the symmetric product $Symm_4(S^1)$ is the nonorientable orbifold identified in [7]-[8] and explored and popularized by Tymoczko in [20]-[21].

Similar arguments can be used to identify the sub-symmetric products $Symm_4^G(S^1)$. The subgroup lattice [1] for Σ_4 consists of: A_4, two different cyclic subgroups of order 2, $\Sigma_2 \times \Sigma_2$ in two different ways (disjoint cycles and not), Σ_3, $\mathbb{Z}_4$, and D_4. The corresponding spaces $Symm_4^G(S^1)$ can all be described as disc bundles over products of circles. Also, for example, it is possible to explicitly address the topological structure in the case of a j-term time-delay embedding of k-note chords.

PROPOSITION 3.7 *When* $G = (\Sigma_k)^j$, *and* $n = kj$,

$$Symm_n^G(S^1) \cong (Symm_k(S^1))^j = \prod_j Symm_k(S^1).$$

---

[1]We consider the subgroup lattice where the subgroups are taken up to conjugacy. Conjugate subgroups of the symmetric group result in homeomorphic sub-symmetric product spaces.

Proof: $(\Sigma_k)^j$ *can be considered a subgroup of* Σ_n *by partitioning* $n = kj$ *into* k *sets of size* j. *This is the subgroup of* Σ_n *that preserves the partition. There is a function*

$$Symm_n^G(S^1) \to \prod_j Symm_k(S^1)$$

given by grouping together the k-note chords. It is continuous by the properties of product and quotient space and one-to-one and onto by design. Since both spaces are compact Hausdorff spaces, this function is a homeomorphism.

Given two fibre bundles $\pi_i : E_i \to B_i$ *for* $i = 1, 2$ *the product is naturally a fibre bundle* $\pi_1 \times \pi_2 : E_1 \times E_2 \to B_1 \times B_1$, *where* $(\pi_1 \times \pi_2)(e_1, e_2) = (\pi_1(e_1), \pi_2(e_2))$. *The fibre of the product is* $(\pi_1 \times \pi_2)^{-1}(b_1, b_2) = \pi_1^{-1}(b_1) \times \pi_2^{-1}(b_2)$ *the product of the fibres of the individual maps.* ∎

Since $Symm_k(S^1)$ is a disc bundle over S^1 which is orientable only when k is odd, $Symm_n^G(S^1)$ is therefore a disc bundle over $(S^1)^j$ which is orientable if and only if k is odd. The case of $Symm_4^G(S^1)$, where $G = \mathbb{Z}_4$ is the subgroup of Σ_4 generated by a 4-cycle is analogous to Example 3.5, and is similarly involved. The space $Symm_4^{\mathbb{Z}_4}(S^1)$ represents 4-pitch melody space, and Example B.1 presents a brief derivation of its homotopy-type. The space $Symm_4^{\mathbb{Z}_4}(S^1)$ can be analyzed via its Morton bundle, as shown in Electronic supplement Appendix B.

3.3.   Finite Subset Spaces

The n-th *finite subset space* of a topological space X is the collection of all subsets of X with at least one, and no more than n elements, denoted $exp_n(X)$. It is given the quotient topology via the map $X^n \to exp_n(X)$ which sends $(x_1, x_2, \cdots, x_n) \longmapsto \{x_1, x_2, \cdots, x_n\}$. Thus $(0, 0, 1) = (0, 1, 1)$ in $exp_3(\{0, 1\})$, i.e. the subset space does not keep track of the number of times an element occurs but only keeps track of the binary information that 'an element is in the set, or not.' This differs from $Symm_3(\{0, 1\})$ in that $(0, 0, 1) \neq (0, 1, 1)$ in $Symm_3(\{0, 1\})$. Thus $Symm_n(X)$ can be thought of as being the 'multi-sets' with precisely n elements, counting repetition. Tuffley studied the finite subset spaces of S^1 in his dissertation.

THEOREM 3.8 (Tuffley [22]) *The space* $exp_1(S^1)$ *can be identified with* S^1. *The space* $exp_2(S^1)$ *is a Möbius strip, with* $\partial exp_2(S^1) = exp_1(S^1)$. *The space* $exp_3(S^1)$ *is homeomorphic to* S^3, *moreover if* $exp_3(S^1)$ *is visualized as the 3-sphere,* $exp_1(S^1)$ *sits in it as a trefoil knot. More generally,* $exp_n(S^1)$ *has the homotopy-type of an odd-dimensional sphere of dimension* n *or* $n - 1$ *if* n *is odd or even respectively. The inclusion* $exp_{2n-1}(S^1) \to exp_{2n}(S^1)$ *induces multiplication by 2 on* H_{2n-1}. *The complement of* $exp_{n-2}(S^1)$ *in* $exp_n(S^1)$ *has the homotopy-type of a* $(n-1, n)$-torus knot complement.

In particular, note that the quotient map $Symm_3(S^1) \to exp_3(S^1)$ is one-to-one except on the subspace mapping to $exp_2(S^1)$. Sitting inside $Symm_3(S^1)$ this is the subspace of the form $\{(z_1, z_1, z_2)\}$ where $z_1, z_2 \in S^1$. Note that (z_1, z_1, z_2) and (z_1, z_2, z_2) are mapped to the same object in $exp_3(S^1)$. So $Symm_3(S^1)$ is equivalent to $D^2 \times S^1$ and the map $Symm_3(S^1) \to exp_3(S^1)$ is one-to-one except on the boundary, where it collapses the torus boundary of $Symm_3(S^1)$ to the Möbius band $exp_2(S^1)$.

3.4. *Homology*

The idea of homology was first introduced by Poincaré; a modern treatment can be found in [23]. Singular homology is the result of a compromise between the geometric intuition of a 'hole' and the desire to have something that is readily computable. The i-th homology group of a space X, H_iX, is an abelian group whose non-zero elements are objects that detect combinatorial i-dimensional 'holes' in X. The precise definition is in the language of simplices and simplicial complexes. Let Δ^n be the standard n-simplex, then a *singular n-simplex* in a space X is a continuous function $\sigma : \Delta^n \to X$. The *boundary* of σ is denoted $\partial\sigma$ and is defined as the formal sum over the bounding $(n-1)$-singular simplices, precisely, $\partial\sigma = \sum_{k=0}^{n}(-1)^k \sigma_{|[0,1,\cdots,k-1,k+1,\cdots,n]}$ where $[0,1,\cdots,k-1,k+1,\cdots,n]$ indicates we are restricting to the k-th face of the n-simplex and parametrizing the face by lexicographically ordering the vertices. Thus the boundary of a 1-simplex is $\sigma_{|1} - \sigma_{|0}$. In general, this is the orientation assigned to the boundary as one learns in Stokes' Theorem, in multi-variable calculus. The *group of singular n-chains* in X is denoted C_nX and defined as the group of formal sums of singular n-simplices in X. The boundary map extends to a linear map $\partial : C_nX \to C_{n-1}X$, and the *n-th homology group $H_n(X)$* is defined as the kernel of $\partial : C_nX \to C_{n-1}X$ modulo the image of $\partial : C_{n+1}X \to C_nX$. In the case of a simplicial complex, for the purposes of computing homology it is known one can replace the arbitrary continuous functions σ with the characteristic maps of the simplicial complex. Thus if the simplicial complex is finite, computing the homology of the complex becomes a polynomial-time computation in the number of input simplices. Other work using homology in music theory can be found in [24] and [25].

Abelian groups have two standard invariants, the rank and torsion subgroup. The *torsion subgroup* of an abelian group is the subgroup of elements of finite order, so in the group $\mathbb{Z}^{100} \oplus \mathbb{Z}_{24}$, the torsion subgroup would be $\mathbb{Z}_{24}$. Taking the abelian group modulo its torsion subgroup (provided this group is finitely-generated), the minimal number of generators is called the *rank* of the original abelian group. If this group is not finitely generated, the group is said to have infinite rank. Hence the group $\mathbb{Z}^{100} \oplus \mathbb{Z}_{24}$ has rank 100.

The rank of the i-th homology group of X, H_iX, is called the *i-th Betti number* of X, and is typically denoted $\beta_i(X)$. For example, the only invariant of H_0X is $\beta_0(X)$, and this is the number of path-components of X. The number of linearly independent loops in X that do not bound discs is $\beta_1(X)$. When the space is a circle, $H_1 S^1 \cong \mathbb{Z}$ so there is one independent loop. Similarly, $H_1(S^1 \times S^1) \cong \mathbb{Z}^2$ and there are two linearly independent circles in the torus, correspondingly $\beta_1(S^1 \times S^1) = 2$. Homology can be somewhat subtle. If X is the Klein bottle, $H_1X \cong \mathbb{Z} \oplus \mathbb{Z}_2$, there is a torsion subgroup corresponding to the circle in the Klein bottle that does not reverse orientation; in this case $\beta_1(X) = 1$ since Betti numbers ignore torsion.

3.5. *Persistent Homology*

Introduced by Carlsson and his coworkers [26] as a way of parsing large data sets, persistent homology takes as its starting point the functoriality of homology. Precisely, if $m : X \to Y$ is a continuous function, there is an induced homomorphism of homology groups $m_* : H_iX \to H_iY$, given by composing the singular simplices in X with m. This allows a comparison between the homology groups of two spaces.

Let P be a collection of points in a space X where there is a notion of distance d :
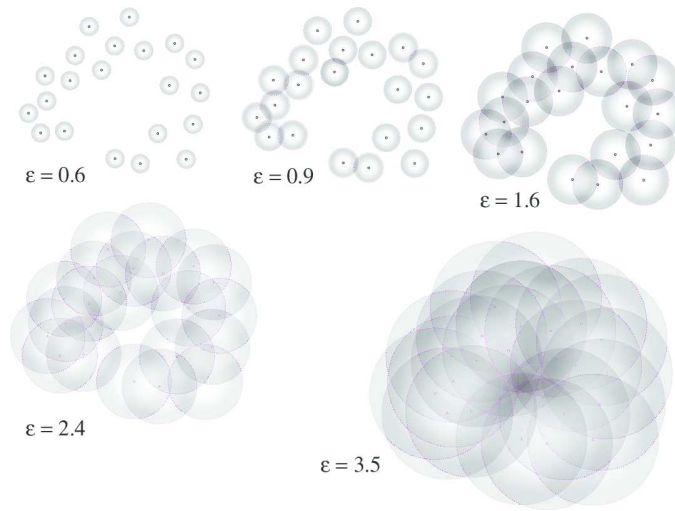
Figure 5. Representation of a point cloud of 21 points and P_ϵ for $\epsilon = 0.6, 0.9, 1.6, 2.4$, and 3.5

$X \times X \to \mathbb{R}$. The *singular homology groups* $H_* P_\epsilon$ are used to represent the general shape of the ϵ-neighborhood of P in X, where $P_\epsilon = \{x \in X : d(x,p) < \epsilon \text{ for some } p \in P\} \subset X$.

Figure 5 shows a set P consisting of 21 points and the ϵ-neighborhoods for five cases. For $\epsilon = 0.6$, the only non-trivial homology group is $H_0 P_{0.6} \cong \mathbb{Z}^{21}$. The top-middle figure shows $\epsilon = 0.9$, and $H_0 P_{0.9} \cong \mathbb{Z}^{12}$ is the only non-trivial homology group. In the third figure, all homology groups of $H_0 P_{1.6}$ are trivial (the homology groups of a point). As ϵ increases yet further to 2.4, a single large circle appears, generating a 1-dimensional homology class. For still larger ϵ (such as 3.5), P_ϵ is approximately a disc, and all homology groups disappear, with the exception of a single H_0 class.

The *barcode* of a point cloud P in X is a record of how the homology of the epsilon-neighborhood P_ϵ varies as a function of ϵ. If $\epsilon_1 < \epsilon_2$, $P_{\epsilon_1} \subset P_{\epsilon_2}$ so there is an induced map of homology groups $H_i P_{\epsilon_1} \to H_i P_{\epsilon_2}$. If a homology class in $H_i P_{\epsilon_1}$ is mapped to a non-trivial homology class in $H_i P_{\epsilon_2}$, it is said to *persist* over the interval $[\epsilon_1, \epsilon_2]$ and the *bar* associated to a homology class is the largest interval over which it persists. The *barcode* of the point cloud P is the record of all the bars over a range of ϵ. Reading a barcode from left-to-right over the ϵ-axis, bars may appear and disappear. A new bar means that a new homology class appears at that ϵ; the bar remains for as long as that homology class persists, and then it vanishes. For example, in Figure 5, consider the seven left-most points. For small ϵ, these are isolated and each appear as a separate line in the barcode. By $\epsilon = 0.9$, these seven lines merge together into a single line. All 21 lines merge into one by $\epsilon = 1.6$ and a single H_0 line persists for all larger ϵ. Figures 9, 10, 11, 13 etc. provide many examples of barcodes.

Unfortunately, as described above, the computation of persistent homology requires knowledge of more than just the point data P. The intermediate step of considering P_ϵ requires knowledge of the distance function on the entire space X, which can be difficult to compute. A simplicial complex P_ϵ^Δ (called the *Vietoris-Rips complex*), which depends only on the distance relation between individual points in P, can be used to approximate P_ϵ. The simplicial complex P_ϵ^Δ is constructed inductively. P is the vertex-set of P_ϵ^Δ. For every pair of points of P separated by less than ϵ one adds an edge to the simplicial

complex. Similarly, for every $(n+1)$-tuple of points of P which are pairwise distance less than ϵ apart, one adds an n-simplex to the simplicial complex. The simplicial complex P_ϵ^Δ is the result of this inductive procedure. Thus persistent homology should be viewed as a coarse measure of the shape of the data P *as an abstract space.* Theorem 2.3 of Adler et. al. [27] gives a sufficient criterion for H_*P_ϵ and $H_*P_\epsilon^\Delta$ to be isomorphic, although the preconditions of their theorem almost never hold with the data considered in this paper.

For ϵ small, the structure is always the same; $P_\epsilon^\Delta = P$, as all points are separated from all other points. For ϵ larger than the diameter of P, the structure is simple to describe; P_ϵ^Δ consists of a single simplex whose vertex set is P. In between small and large ϵ there may be a wide variety of homology groups. Homology classes that persist over a significant ϵ-interval are called *persistent* and frequently reflect interesting underlying structure in the data set P. A number of applications have begun to appear in areas such as image processing [26] and in the analysis of biological data [28].

4.   Simulations with Musical Data

This section analyzes a number of musical pieces (some contrived and some from various standard repertoires) using persistent homology by turning raw musical data (in this case, note-level MIDI data) into a collection of points. Using the appropriate distance measures (1)-(4), it is straightforward to calculate the distance between every pair of points in the data. The `javaPlex` software [29], designed to "calculate the persistent homology of finite simplicial complexes... generated from point cloud data," can be used to calculate the barcodes for a given set of data. See Section 3.5 for the basics of persistent homology, and how to read the barcodes. The input is a set of distances (between all points in the point-cloud) and the output is a set of plots that show the Betti numbers as a function of the persistence parameter ϵ.

4.1.   Rhythmic Barcodes

The first example translates the "Ewe" rhythm of Figure 1 into the set of time points

$$\left\{ 0, \frac{2}{12}, \frac{4}{12}, \frac{5}{12}, \frac{7}{12}, \frac{9}{12}, \frac{11}{12} \right\}. \tag{5}$$

The distance is calculated between every pair of these time points under the necklace metric (1), and this set of distances is the "point cloud" that is input into the javaPlex software. The resulting barcodes are shown in Figure 9.

The dimension 0 barcode shows the clustering of the points in time. In the sequence (5), the minimum distance is $\frac{1}{12}$, and this occurs in two places, between the third and fourth elements of the set, and again between the 11th and the first elements. Accordingly, the barcode shows two lines that vanish when ϵ reaches 0.08. Since the largest distance between any two adjacent time points is 0.16, all the points merge into one cluster at $\epsilon = 0.16$. To see why this happens, imagine that a balloon is placed around each of the small circles in Figure 1. When the radius of the balloon is small, each balloon is isolated and contains no other circles than the one at its center. But as the balloons are
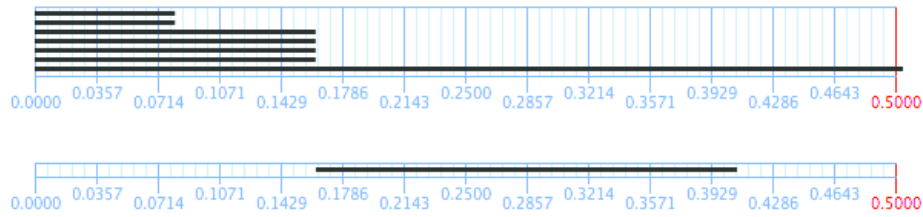
Figure 9. Barcodes calculated by the javaPlex software show the number of connected components in the dimension 0 plot (top) and the number of circles in the dimension 1 plot (bottom), as the size parameter ϵ varies. Barcodes for King's "standard rhythm" [15] of Figure 1 show the distribution of time intervals in the rhythm in the top (dimension 0) plot and show the circular structure with Betti numbers $\beta_0 = 1$ and $\beta_1 = 1$ in the bottom (dimension 1) plot. Using the distance function (1), the largest distance between elements is 0.5, and the diameter of the space, homeomorphic to S^1, is also 0.5.

inflated (which corresponds to increasing ϵ), the balloons around point a and f grow large enough to touch the points b and e. This is the radius at which the first two lines in the upper figure disappear, the points a and b are now encompassed by the same balloon (and similarly for the points e and f). As the balloons inflate still further, they grow to include other points that are two time units apart; at this size, all adjacent points are now connected. This situation is shown in the dimension 1 barcode which displays a persistent β_1 bar from $0.16 < \epsilon < 0.408$. This is the anticipated cycle around the necklace. With $\beta_0 = 1$ and $\beta_1 = 1$, the complete structure is indeed the circle S^1. While this is not surprising, it is useful to verify that the simulation method and the analysis agree.

All three of the rhythms indicated in Figure 1 have identical barcodes because the rhythms are circular shifts of each other and hence contain the same collection of rhythmic intervals.

4.2.   The C-Major Scale

To verify that the distance function (2) makes sense, consider the C-major scale consisting of the eight notes C, D, E, F, G, A, B, C with frequencies given starting at $C = 261.6$ Hz. The data is defined by a matrix of distances between all pairs of the eight notes using the pitch-class distance on $\mathbb{R}_+$. The resulting barcodes are shown in Figure 10.

These two plots are straightforward to interpret. When the size parameter 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 distance (2). Since the distance from the high C to the low C is zero, 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 $\beta_1 = 1$ feature is S^1, and corresponds to the circular structure illustrated in Figure 2.
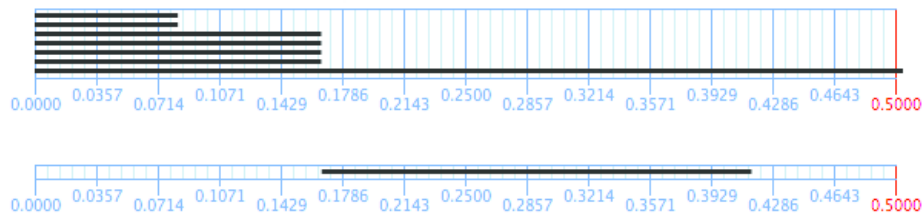
Figure 10. Barcodes for the C-major scale show seven distinct lines in the top (one corresponding to each note) for small ϵ. Two half steps merge at $\epsilon \approx 0.08$ and the whole steps merge at $\epsilon \approx 0.16$, at which point the circle S^1 with $\beta_0 = \beta_1 = 1$ appears. The distance function used is (2) and the diameter of the space S^1 is again 0.5.
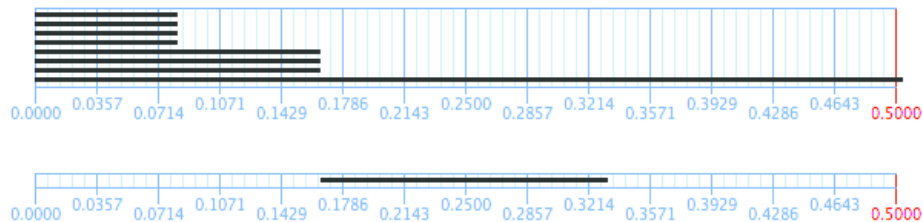

Figure 11. Barcodes for the traditional folk tune "Abbott's Bromley Horn Dance" (see Figure C1) under the distance function (2) show many of the same features as the major scale barcodes of Figure 10. The distribution of whole and half steps are clear from the β_0 code for small ϵ while S^1 appears again in β_1 for $0.16 < \epsilon < 0.33$.

4.3.   Pitch-Class Barcodes: Abbotts

The input data for example of Figure 10 was built specifically with the circular structure of Figure 2 in mind, so it is perhaps unsurprising that the circle appears. Will such shapes appear in real music? The website [30] contains a large selection of traditional melodies, with most tunes available in both sheet music and as standard MIDI files. The musical score for "Abbott's Bromley Horn Dance" is shown in Figure C1 and the corresponding barcodes are shown in Figure 11. The distance between note pairs is given by (2).

The top barcode in Figure 11 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 $\beta_1 = 1$, one circle. This is again the circular structure S^1. In fact, all the melodies from the website [30] show this same structure, though the number of half and whole steps changes to reflect the scale of the piece, and the exact extent of the $\beta_0 = \beta_1 = 1$ region is somewhat variable.
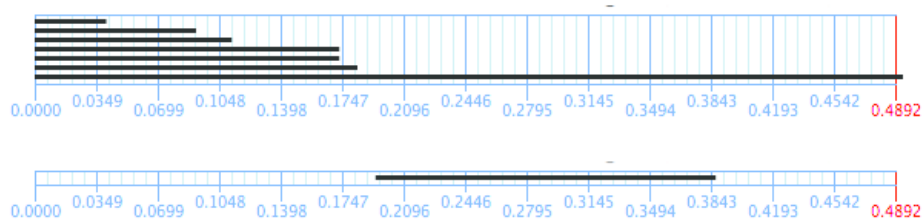
Figure 12. Barcodes for an Ussak makam under the distance function (2) show many of the same features as the barcodes of Figure 11, though some significant dissimilarities are also apparent. The occurrence of many different intervals is apparent in the staggered appearance of the dimension 0 barcode (top). Again, the circle appears in the bottom barcode for $0.18 < \epsilon < 0.38$.

4.4.   *Barcodes for an Ussak Makam*

While it might appear from the above examples that the analysis requires the pitches to be subsets of the semitones of the Western equal-tempered scale, they can assume any possible values. The Turkish makam tradition utilizes many tones/pitches that are distinct from the Western pitches, and it is instructive to look at the bar codes for one such piece. A data base of 1700 makams available online [31] contains MIDI and score data. About a hundred of the makams are from the mode ussak, and these were analyzed using the javaPlex software. The barcodes for a typical example are presented in Figure 12. Like the Western pieces, as ϵ increases, the pitches merge together. Unlike the equal tempered pieces, however, there are many different intervals between various pitches and so the bars of the code merge together at somewhat staggered values (rather than at one or two discrete ϵ values). Nonetheless, the structure of the underlying space is revealed by the merging of all the pitches in the dimension zero barcode by $\epsilon \approx 0.18$ and the emergence of the dimension 1 bar at about this same value. Again, this represents the circular S^1 structure.

4.5.   *Time-Delay Embedding: 2-D Abbotts*

The analyses of Figures 10 and 11 may be somewhat naive because they suppress temporal information in the melody. A time-delay embedding considers successive pairs (or more) of elements and so recovers some of the temporal relationships; it is a common procedure in time series analysis. For this approach the appropriate distance function is (2). Building a matrix of all such distances for "Abbott's Bromley Horn Dance" and calculating the barcodes gives Figure 13.

The dimension 0 barcodes (the top plot in Figure 13) can be interpreted as showing the distances between consecutive pairs of notes as the melody progresses over time. Thus there are 11 pairs of notes that are at a distance of one-half step, since 11 lines end at $\epsilon = 0.08$. There are 19 pairs that differ by a whole step since 19 lines end at $\epsilon = 0.16$. Above this value, all pairs have merged into a single connected component. This can be interpreted as saying that the melody progresses primarily by stepwise motion, and that no pairs of tones are isolated from any other pairs of tones (though of course there are
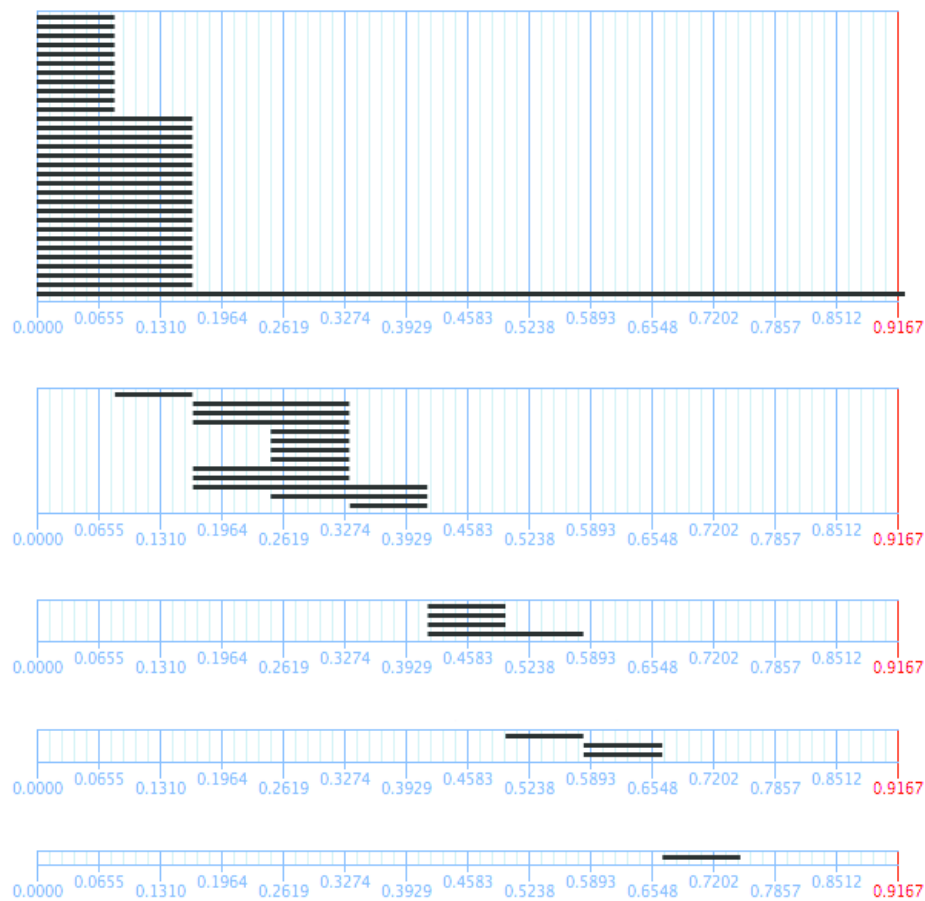
Figure 13. Barcodes for the two-dimensional time-delay embedding of "Abbott's Bromley Horn Dance" (see Figure C1) are considerably more interesting than those in Figure 11 without the time-delay. The top picture shows the dimension 0 barcodes, the second shows the dimension-one barcodes, continuing through the dimension-four barcodes. The distance function (2) is applied to each element of 2-vectors (which consist of consecutive pairs of note values), and then summed. The topological space in which this is embedded is $S^1 \times S^1$, the $n = 2$ case of Example 3.2. The maximum distance between elements of the performance is 0.9167, which is close to the diameter of the complete space.

many individual pairs with larger distances).

   The dimension 1 barcodes (the second plot in Figure 13) shows the number of independent circles present at each value of ϵ, the dimension 2 barcodes (the third plot) show the number of closed oriented 2-cycles (such as hollow spheres) as a function of ϵ, and the dimension 3 barcodes (the bottom plot) show the distribution of 4-cycles. For this example, the embedding space is the torus $S^1 \times S^1$, as in Example 3.2.

   This is our first encounter with barcodes beyond the dimension of the underlying metric space, which is two for the space $S^1 \times S^1$. In brief what we see is a decreasing number of bars (as the dimension increases) and the persistence intervals of these bars appears to shrink as the dimension increases. This is a typically-observed phenomenon with data that respects a simple distribution, like a Gaussian [14]. More precisely, what one typically

sees in the barcodes of a Gaussian field is a void supporting no homology in the centre of the distribution, while homology classes get sparser (and higher dimensional) as one approaches the void from the outside. In this case, there is one 1-dimensional homology class that persist over a fairly wide interval, indicating the data diverge somewhat from this expectation. In this case it is due to the relative diversity of two-note sequences starting with an E, and the relative sparsity of other combinations.

The higher dimensional barcodes are few and have relatively short persistence intervals so they represent relatively insignificant features of the data. Moreover, the fact that they represent cycles that can not exist in the parameter-space $S^1 \times S^1$ further indicates the idiosyncratic nature of these cycles. This is reinforced by analyses of other songs from the same set (not shown here) of 2D time-delay embeddings which have analogous (but different) groupings of higher order barcodes.

4.6.   *Time-Delay Embedding: 3-D Abbotts*

Longer temporal information can be incorporated by using longer time-delay embeddings. If a melody consists of a sequence of notes with pitches at $f_1, f_2, f_3, f_4...$, these can be combined into three-tuples (a three-dimensional time-delay embedding) by creating the sequence $(f_0, f_1, f_2), (f_1, f_2, f_3), (f_2, f_3, f_4), ...$ The distances between such three-tuples can be calculated by adding the distances between the notes element-wise using the pitch-class distance. Building a matrix of all such distances for "Abbott's Bromley Horn Dance" and calculating the barcodes gives Figure C2.

Again, it is straightforward to interpret the dimension 0 barcodes as distances between three-tuples of notes in the melody. Since the embedding here is $S^1 \times S^1 \times S^1$ by Example 3.2, the higher dimensional structures are again linked to the pieces being analyzed.

4.7.   *Barcodes for the Circle of Fifths*

A well known structure in music theory is the "circle of fifths," shown here in Figure 14, which is taken from the Wikipedia article of the same name [32]. The circle of fifths is a standard way musicians and composers talk about the close relationships between musical scales and keys, and represents another way of interpreting the distance between musical chords and scales. The appropriate distance function is (3), and it is reasonable to ask if the circle of fifths can be recovered from a data analysis.

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. For the present example, each key is represented by its major scale, a seven note set of pitches. Inputting these sets into javaPlex and calculating the barcodes gives Figure 15. Under the distance (3), scales that are a fifth apart (such as C major and G major) have a distance of 0.08. This explains the twelve lines that merge down to a single connected set at $\epsilon = 0.08$ in the dimension 0 (top) plot. For $0.08 < \epsilon < 0.33$, the dimension 1 barcode shows a single persistent bar; this is precisely the circle of fifths.

As shown in Example 3.4, the space in which scale (and/or chord) progressions lie is an n-fold symmetric product of the circle. As constructed in this section, the circle of fifths can then be interpreted as twelve points in $Symm_7(S^1)$. Observe that the circular
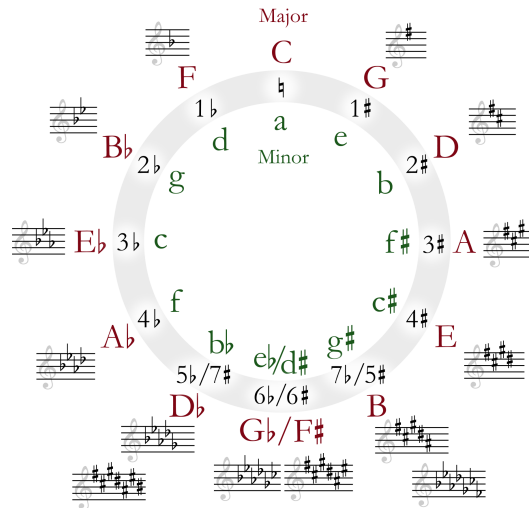
Figure 14. The circle of fifths shows the relationships among the tones of the equal tempered chromatic scale, standard key signatures, and the major and minor keys.

structure of Figure 14 is not used in any way to draw the barcodes: rather, a set of scales (along with a metric on those scales) are the only inputs to the software. The barcodes then exhibit a structure that can be interpreted as identical to the circle.

There are also some higher dimensional features for larger ϵ. The circle of fifths has dihedral symmetry – specifically, the symmetry group of a regular 12-sided planar polygon (a *dodecagon*). Interestingly the metric is not that of a regular dodecagon. In this case, the distance function is like the seam of a baseball – stretching off course in places, then turning back later. Precisely, the n-th and $(n + 4)$-th note are within 0.33 of each other *for all n*. So the triple consisting of the n-th, $(n + 4)$-th and $(n + 8)$-th notes form a triangle in the Vietoris-Rips complex for each n, provided $\epsilon > 0.33$, and they *do not* provided $\epsilon < 0.33$. Since there are 12 notes, that gives 4 triangles being attached to the Vietoris-Rips complex near $\epsilon = 0.33$. Attachment of any of these triangles removes the only H_1-class in the Vietoris-Rips complex for $\epsilon < 0.33$. This is because the generator of the H_1-class is the sum of the edges consisting of consecutive chords in the cyclic order of the circle of fifths. Moreover, the edge consisting of the n-th note and the $(n+4)$-th note is readily computed to be a boundary. Thus, all four of the triangles remove the same H_1-class from the chain complexes for $\epsilon < 0.33$, so the union of the four triangles not only removes the H_1-class but contributes three H_2-classes, via a direct Meyer-Vietoris argument.

4.8.   *Bar Codes for Chords in Bach's Chorale No. 19*

The musical score to Bach's Chorale No. 19 is shown in Figure D1. A MIDI file of this piece, from [33], is parsed to extract the four voices. The distances between all four-part chords are calculated according to the chord-class distance (3), and the results are input
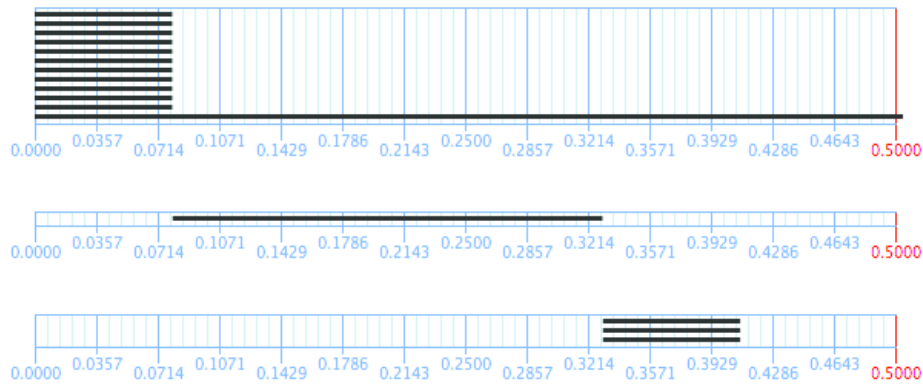
Figure 15. Barcodes for a sequence consisting of a major scale proceeding around the circle of fifths. The "circle" is the persistent line from $0.08 < \epsilon < 0.33$ in the dimension 1 plot. Under the chord-class distance function (3) the space is $S^1 \times D^2$ (this is the $n = 3$ case of Example 3.4) and the total diameter of the space is 3.5 while the maximum distance between any two elements is 0.5.

to the javaPlex software in order to draw the barcodes, which are shown in Figure D2. The dimension 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 dimension 1 barcode in Figure D2 shows $\beta_0 = 6$ connected components and one circle $\beta_1 = 3$ for $0.16 < \epsilon = 0.24$, and this structure then changes to $\beta_0 = 1$ and $\beta_1 = 3$ 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.

4.9.   *Time-Delay Embedding of Chords in Abbotts*

It is also easy to include temporal information with a time-delay embedding of the chords. Some care must now be taken with the distance function since it is a mixture of the chord-class (3) and the sum-of-elements distance useful in time-delay embeddings as in Section 4.5 and 4.6. To be explicit, suppose that a sequence of 3-note chords $F = (f_1, f_2, f_3)$, $G = (g_1, g_2, g_3)$, $H = (h_1, h_2, h_3)$, $I = (i_1, i_2, i_3)$, ... is gathered into time-delay pairs to form 6-tuples (F, G), (G, H), (H, I).... Then the distance between pairs of such 6-tuples is the time-delay embedding distance

$$d_{TDE}((F, G), (H, I)) = d_{CC}(F, H) + d_{CC}(G, I). \tag{6}$$

Observe that under this measure, the distance between notes within a single chord is measured differently from the distance between the same notes when they occur at different times. An example of the resulting barcodes is shown in Figure C1 for a 2D time-delay embedding of 3-note chords. The barcode shows between several circles for ϵ in the region

W. A. Sethares and R. Budney

Table 1. Some Afro-Cuban rhythms

bossa-nova	1	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
gahu	1	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0
rumba	1	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0
shiko	1	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0
son	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0
soukous	1	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0

of 0.5 indicating some structure in the chord pattern. The underlying space is described in Proposition 3.7 for the case of a $j = 2$-term time-delay embedding of $k = 3$-note chords.

4.10.   *Two-Dimensional Time-Delay Embedding of Chorale No. 19*

Figure D3 shows the two-dimensional time-delay embedding of the Bach's Chorale No. 19, where the zero dimensional plot is interpreted directly in terms of the distribution of chord pairs and how they cluster under the chord-class distance. Again, there is a collection of persistent β_1 circles and scattered entries in the dimension 2 and 3 barcodes.

4.11.   *Rhythms and Cyclic Permutations*

Repetitive rhythmic patterns with period n may be written in the timeline notation as elements of $\mathbb{Z}^N$ where each "1" represents the occurrence of a pulse and each "0" represents silence. For instance, the 16-beat *clave son* can be transcribed as

$$\textbf{Clave Son} = (1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0) \in \mathbb{Z}^{16}. \tag{7}$$

Table 1 shows a variety of Afro-Cuban rhythms each of which is the basis of a traditional or popular musical style. From a musical perspective, it is clear that some such rhythms can be thought of as "more alike" than others: they may be performed in similar musical styles, they may occur close together geographically, or there may be cultural ties between groups that prefer certain patterns over others. It is intuitively plausible that rhythms may be "close together" or "far apart" based on similarities and differences in the set of temporal intervals that define the rhythms. An appropriate notion of distance in this case is (4), which considers all patterns related by a cyclic shift to be identical. In this case, the underlying ambient space would be $Symm_5^{\mathbb{Z}_5}(S^1)$, a space consisting of 5-tuples that represent onset times in a temporal cycle. Technically, the rhythm space should be the subspace of distinct 5-tuples, but $Symm_5^{\mathbb{Z}_5}(S^1)$ has the correct metric on this subspace so we create no special notation for this situation.

Considering the six rhythmic patterns of Table 1 as 6 points in the space, the barcode is shown in Figure E1. The lack of Betti numbers for H_i with $i > 0$ is indicative of the underlying curiosity of this data as a metric space – the distance between the points is as if the points were the vertices of a graph with the distance being the *metric graph*

distance. Precisely, given a graph whose edges are labelled with positive numbers, the *metric distance* between two vertices is the minimum (over all paths) of the sum of all edge lengths along a path from one point to the other. In this case it is a metric graph where all edges have length 0.0625. See Figure E2.

5.   Conclusions and Discussion

There are two main contributions of this paper. First, it augments the current state of the art in understanding the topological structure of musical spaces by presenting several new and expanded examples of distance functions and the spaces they imply. The techniques are drawn from classical topology, and may be readily applied to a wide variety of distance measures. The development emphasizes the importance that even small details of the distance function can have on the global structure of the space, and shows that (generalized) circles and discs appear frequently among the spaces because of the dual nature of the distance functions: the pitch-based distances measure proximity in absolute pitch and proximity after octave reduction, the temporal distances measure proximity in time and proximity in phase around the rhythmic cycle.

The second contribution is the application of persistent homology to musical data. Time-delay embeddings allow the method to exploit the time-dependent nature of musical data. Using this tool, the Betti numbers (and hence homology classes) can be calculated directly from time-pitch data. To verify that the tool makes sense, we have presented a number of cases where features are known and then verified that the topological analysis is able to recover those features using only the note-level data. As far as we know, this is the first time such a validation has been attempted. When looking at more complicated musical data, the results of the low dimension homology classes are usually easy to interpret. In contrast, the meaning of the higher-dimensional Betti numbers is less clear. In part this is due to the developing nature of persistent homology as a field, and we suspect that once the persistent homology of Gaussian fields is better understood, it will help in the understanding of barcodes such as those presented here.

References

[1] Catanzaro M. "Generalized *Tonnetze.*" *J Math Music,* 2011;2:117–139.
[2] Partch H. *Genesis of a Music.* New York: Da Capo Press, 1974.
[3] Chew, E. "The spiral array: an algorithm for determining key boundaries." In Anagnostopoulou C, Ferrand M, Smaill A, editors. *Music and Artificial Intelligence - Second International Conference*, ICMAI 2002, Edinburgh, Scotland, UK, Sept, 2002.
[4] Music Visualization, `http://en.wikipedia.org/wiki/Music_visualization`
[5] Lewin D. *Generalized musical intervals and transformations.* New Haven: Yale University Press; 1987.
[6] Mazzola G. *Geometrie der Töne.* Birkhuäser, 1990.
[7] Callender C, Quinn I, Tymoczko D. "Generalized chord spaces." Presented at the John Clough Memorial Conference, University of Chicago, 9 July 2005.
[8] Callender C, Quinn I, Tymoczko D. "Generalized voice-leading spaces." *Science.* 2008;320(5874):3460348.

[9] Buteau C, Mazzola G. "Motivic analysis according to Rudolph Reti: formalization by a topological model." *J. Math. and Music*, 2008;2:117-134.

[10] Buteau C. "Motivic spaces of scores through RUBATOs MeloTopRUBETTE". In: Mazzola G, Noll T, Lluis Puebla E, editors. *Perspectives in mathematical and computational music theory*, Osnabrück Series on Music and Computation. Osnabrück: Electronic Publishing Osnabrück; 2004;330–342.

[11] Sethares W, Budney R, "Electronic supplement to *topology of musical data*," http://

[12] Safî al-Din al-Urmawî, *Kitâb al-Adwâr* 1252, trans. R. Erlanger in *La Musique arabe*, Paul Geuthner, Paris, 1938.

[13] Anku W. "Circles and time: a theory of structural organization of rhythm in African music." *Music Theory Online*, 2000;6(1).

[14] Balzano L, Ellenberg J. *Understanding persistent homology and plex using a networking dataset.* Oct. 2010. http://web.eecs.umich.edu/~girasole/wordpress/wp-content/uploads/2013/04/PersistentHomolo

[15] King A. "Employments of the "standard pattern" in Yoruba music." *African Music*, 1961;2(3).

[16] Rahn J. *Basic Atonal Theory*, Prentice-Hall International, 1980.

[17] Carlsson G, Zomorodian A, Collins A, Guibas L. *Persistence barcodes for shapes.* Eurographics Symposium on Geometry Processing. 2004.

[18] Kendal, D G. "A survey of the statistical theory of shape." *Statistical Science*, 1989;4(2):87–120.

[19] Morton H R., *Symmetric products of the circle.* Proc. Camb. Phil. Soc. 1967;63:349–352.

[20] Tymoczko D. "The geometry of musical chords." *Science* 2006;313(5783):72–74.

[21] Tymoczko D. *Geometry of Music*, 2010.

[22] Tuffley C. *Finite subset spaces of S^1*, Alg. Geom. Topology 2002;2:1119–1145.

[23] Hatcher A. "Algebraic Topology," Cambridge University Press, 2001.

[24] Nestke A. "Paradigmatic motivic analysis." In: Mazzola G, Noll T, Lluis Puebla E, editors. *Perspectives in mathematical and computational music theory*, Osnabrück Series on Music and Computation. Osnabrück: Electronic Publishing Osnabrück; 2004:343–365.

[25] Mazzola, G. "Singular Homology on Hypergestures." *J. Math. Music.* 2012;6:49–60.

[26] Carlsson G. "Topology and Data," *Bulletin of the American Mathematical Society*, 2009;46(2):255–308.

[27] Adler R J, Bobrowski O, Borman M S, Subag E, Weinberger S. *Persistent Homology for Random Fields and Complexes,* Institute of Mathematical Statistics, Beachwood, Ohio. 2010;124–143.

[28] Singh G, Memoli F, Ishkhanov T, Carlsson G, Sapiro G, Ringach D. "Topological structure of population activity in primary visual cortex." *Journal of Vision*, 2008;8(11):1–18.

[29] JavaPlex, http://code.google.com/p/javaplex/

[30] Peterson C. "Traditional Music." http://www.cpmusic.com/tradmus.html

[31] Karaosmanoğlu M K. "A Turkish Makam Music Symbolic Database for Music Information Retrieval: SymbTr." *Proceedings of ISMIR.* 2012. http://compmusic.upf.edu/node/140

[32] Circle of Fifths, http://en.wikipedia.org/wiki/Circle_of_fifths

[33] Classical MIDI Archives, http://www.classicalarchives.com/

Electronic Supplement to *Topology of Musical Data*

This document is the electronic supplement for the article 'Topology of Musical Data' by Sethares and Budney.

Appendix A. Example 3.5

For the subgroup $G = A_3 = \mathbb{Z}_3$, $Symm_3^G(S^1)$ represents the 3-note melody space. Analogous to Morton's study of symmetric products, this is homeomorphic to $S^1 \times S^2$. This can be shown by recognizing that Morton's bundle (Theorem 3.1) $Symm_3^{A_3}(S^1) \to S^1$ is trivial, and the fibre is a sphere. To see this, observe that a fundamental domain in Figure A1 for the action of $\Delta_{\mathbb{Z}} \rtimes A_3$ on $\Delta_{\mathbb{R}}$ is the quadrilateral with vertices $(0, 0, 0)$, $(-2/3, 1/3, 1/3)$, $(1/3, -2/3, 1/3)$ and $(-1/3, -1/3, 2/3)$ (the red quadrilateral). The argument proceeds in two steps. First, $\Delta_{\mathbb{Z}} \subset \Delta_{\mathbb{R}}$ is the dual of the hexagonal lattice; it is the triangulation of the plane by equilateral triangles. Take the hexagon centered at the origin, dual to the triangular lattice. This has vertices $(-2/3, 1/3, 1/3)$, $(1/3, -2/3, 1/3)$ together with their cyclic permutations (given by the action of A_3). Accordingly, the fundamental domain of the $\Delta_{\mathbb{Z}} \rtimes A_3$ is one third of the hexagon, the quadrilateral described above. Within the red polygon, the two top edges are glued together by a translation-rotation, while the two bottom edges are glued together by a counter-clockwise $4\pi/3$ rotation about the origin. From this perspective the fiber is naturally a triangulated S^2, using precisely two triangles. The triangulation can be viewed as the union of two identical triangles along their common boundary. The monodromy has two fixed points (the centres of the triangles) and is rotation by $\pm 2\pi/3$ at those fixed points.

Figure A1 shows a view 'down' the $\Delta_{\mathbb{R}}$ plane in $\mathbb{R}^3$, viewed from the perspective of the $x = y = z$ axis. The green dots indicate the points lying outside the $\Delta_{\mathbb{R}}$ plane, while the black dots are in the $\Delta_{\mathbb{R}}$ plane. The red quadrilateral is the fundamental domain of the $\Delta_{\mathbb{Z}} \rtimes A_3$ action on $\Delta_{\mathbb{R}}$, while the union of the yellow and red polyhedrons is the fundamental domain for the action of $\Delta_{\mathbb{Z}}$ on $\Delta_{\mathbb{R}}$ (the hexagonal lattice).

A section of the bundle is realized by all the 3-note melodies where all three notes equally spaced; these are the fixed-points of the monodromy $S^2 \to S^2$ above. There are two such natural sections: one where the three notes are cyclically increasing, and one where the notes are cyclically decreasing in pitch. Since the bundle is trivial, there are many other sections, but since the monodromy only has the two fixed points, no other sections are pitch-translates of a point in the fiber. This argument explains example 3.5.
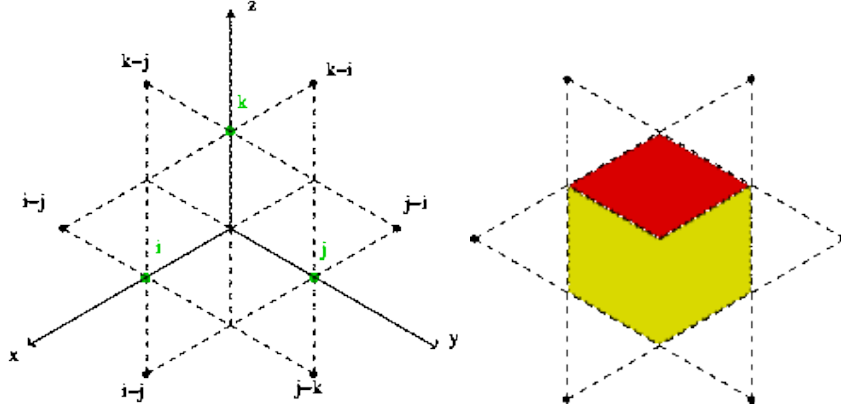
Figure A1. The fundamental domain for the 3-note melody space is described in detail in the text.

Appendix B. An alternative representation

A common way of representing the Betti numbers of a space used in topology is called the Poincaré Polynomial. Given a space X, the Poincaré Polynomial is the function $p_X(t) = \sum_{i=0}^{\infty} \beta_i(X)t^i$ where $\beta_i(X)$ is the i-th Betti number of X, i.e. $\beta_i(X) = rank(H_iX)$. The Poincaré Polynomial of a circle S^1 is $p_{S^1}(t) = 1 + t$ since the only non-trivial homology groups are H_0 and H_1, both are free abelian on one generator. Similarly, for the n-sphere $p_{S^n}(t) = 1 + t^n$. A consequence of the Künneth Theorem is that the Poincaré Polynomials of a product are the products of the Poincaré polynomials of the individual spaces. See [23] for details of the Künneth Theorem and Poincaré polynomials.

$$p_{X \times Y} = p_X \cdot p_Y$$

Thus, in the case of Theorem 3.7,

$$p_{Symm_n^G S^1}(t) = (1 + t)^j = \sum_{j=0}^{n} \binom{n}{j} t^j.$$

Similarly,

$$p_{Symm_3^{A_3} S^1}(t) = (1 + t)(1 + t^2) = 1 + t + t^2 + t^3.$$

Example B.1 $Symm_4^{\mathbb{Z}_4}(S^1)$ can be analyzed via its Morton bundle. The fibre of the map $Symm_4^{\mathbb{Z}_4}(S^1) \to S^1$ is $\Delta_{\mathbb{R}}/\Delta_{\mathbb{Z}} \rtimes \mathbb{Z}_4$. Let e_1, e_2, e_3, e_4 represent the standard basis vectors in $\mathbb{R}^4$. The lattice $\Delta_{\mathbb{Z}} \subset \Delta_{\mathbb{R}}$ is generated by the set $\{e_i - e_j : i \neq j\}$, which makes the vertices of a cuboctahedra. The tessellation of $\Delta_{\mathbb{R}}$ dual to this lattice is the rhombic dodecahedral tessellation. So a fundamental domain for the $\Delta_{\mathbb{Z}}$-action on $\Delta_{\mathbb{R}}$ is the rhombic dodecahedron centred at the origin, such that the centres of the faces correspond to the set $\{e_i - e_j : i \neq j\}$. The action of $\mathbb{Z}_4$ on the rhombic dodecahedron is given by rotation by $\pi/2$ about an axis through the origin and a vertex of valence 4,

followed by the antipodal map. The fundamental domain of the $\Delta_{\mathbb{Z}} \rtimes \mathbb{Z}_4$-action on $\Delta_{\mathbb{R}}$ is a parallepiped. This can be seen by expressing the rhombic dodecahedron as a union of 4 isometric parallepipeds which $\mathbb{Z}_4$ permutes cyclically. The parallelepipeds correspond to the vertices of the regular tetrahedron vertices, the orthogonal projects of $\{e_1, e_2, e_3, e_4\}$, in particular the intersection of the four parallelpipeds consists of only the origin. The identifications on the boundary of the parallelpiped folds two of the faces on themselves, and identifies two pairs of faces. This leaves the fiber with a CW-structure with 4 0-cells, 3 1-cells, 2 3-cells and a single 3-cell. The 2-cell attaching maps are null, so the fiber has the homotopy-type of the wedge $S^2 \vee S^2$ attach a 3-cell via the sum of twice the characteristic maps of the 2-spheres respectively. The result is the fibre $\Delta_{\mathbb{R}}/\Delta_{\mathbb{Z}} \rtimes \mathbb{Z}_4$ has the homology groups $H_0 \cong \mathbb{Z}$, $H_1 \cong 0$, $H_2 \cong \mathbb{Z} \oplus \mathbb{Z}_2$ and $H_3 \cong 0$. In particular, $Symm_4^{\mathbb{Z}_4}(S^1)$ is not a manifold. Since non-torsion H_2 elements are in the image of the quotient map $H_2 Symm_4(S^1) \to H_2 Symm_4^{\mathbb{Z}_4}(S^1)$, the monodromy acts trivially on the homology of the fiber. So the homology groups of $Symm_4^{\mathbb{Z}_4}(S^1)$ are isomorphic to that of a product bundle, $H_0 \cong \mathbb{Z}$, $H_1 \cong \mathbb{Z}$, $H_2 \cong \mathbb{Z} \oplus \mathbb{Z}_2$, $H_3 \cong \mathbb{Z} \oplus \mathbb{Z}_2$, and $H_4 \cong *$.

Appendix C. "Abbott's Bromley Horn Dance"

Figure C1. The traditional melody "Abbott's Bromley Horn Dance" is taken from Chris Peterson's collection [30]. The standard MIDI version of this melody is analyzed using the ideas of persistent homology in Figures 11, 13, C2, and C3.
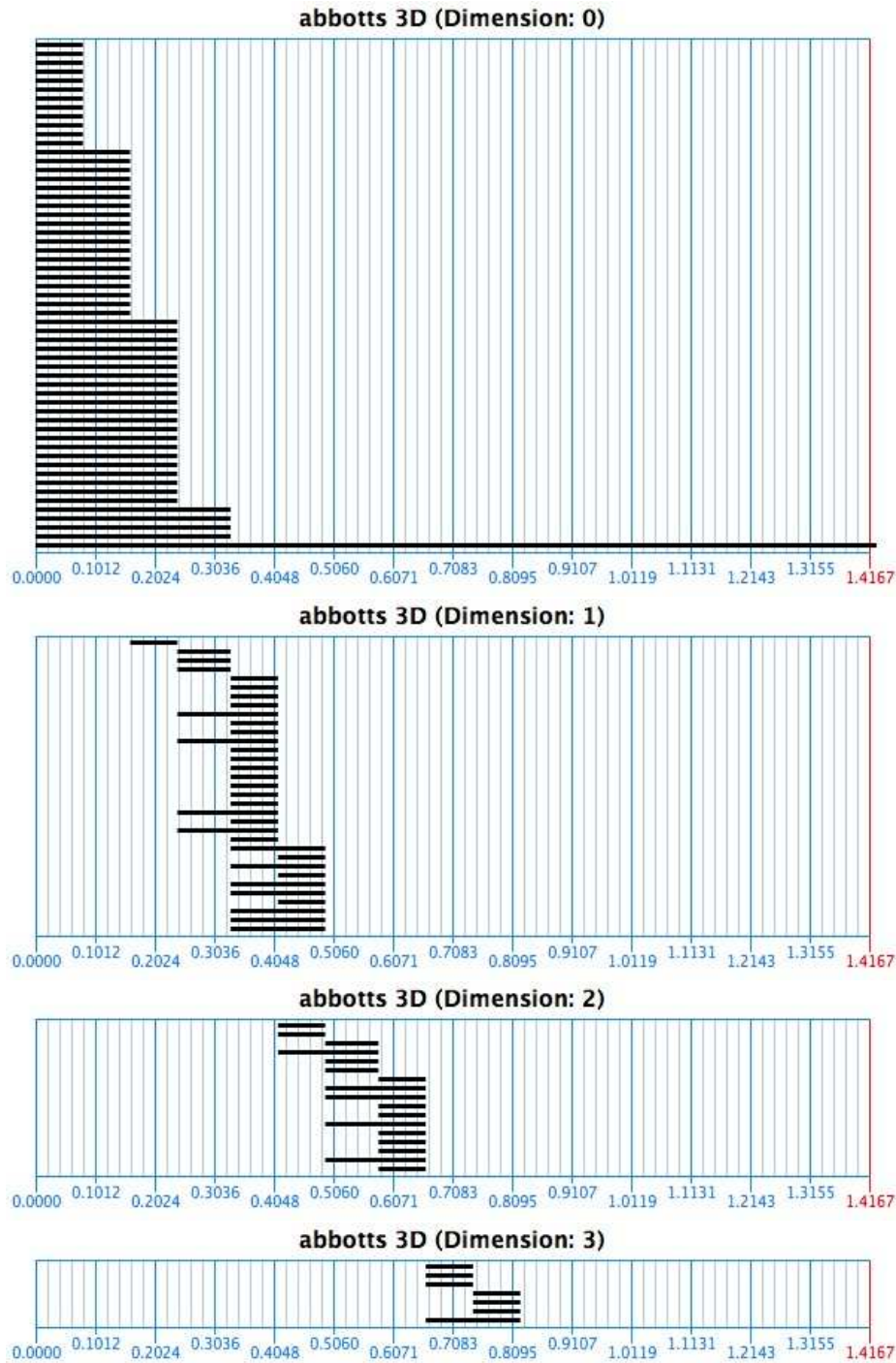
Figure C2. Barcodes for the two-dimensional time-delay embedding of "Abbott's Bromley Horn Dance" (see Figure C1) show a remarkable array of features that represent both the structure of the space and the structure of the individual piece. The dimension 0 plot is straightforward (showing the number of melodic 3-tuples at each distance), the Dimension 1 and 2 plots contain a fascinating collection of circles and higher dimensional analogs that persist over a nontrivial range of ϵ. The distance function is (2) applied to each element of the 3-vector, and then summed. The topological space in which this is embedded is $S^1 \times S^1 \times S^1$, the $n = 3$ case of Example 3.2. The maximum distance between elements of the performance is 1.4, which is close to the diamater 1.5 of the complete space.
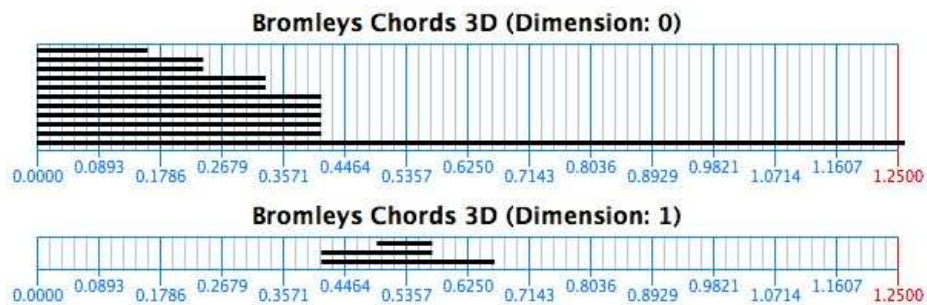
Figure C3. Barcodes for the dimension-two time-delay embedding of the chord progression from the score of "Abbott's Bromley Horn Dance." The distance function is the two dimensional time-delay embedding (6). Under this distance, the diamater of the space is 4.5 and the maximum distance between elements is 1.25, suggesting that the data sits in a small 'ball' and occupies only a small portion of the full space.

Appendix D. Bach's Chorale No. 19

Figure D1. The standard MIDI version of Bach's Chorale No. 19 is analyzed using the ideas of persistent homology in Figures D2 and D3.
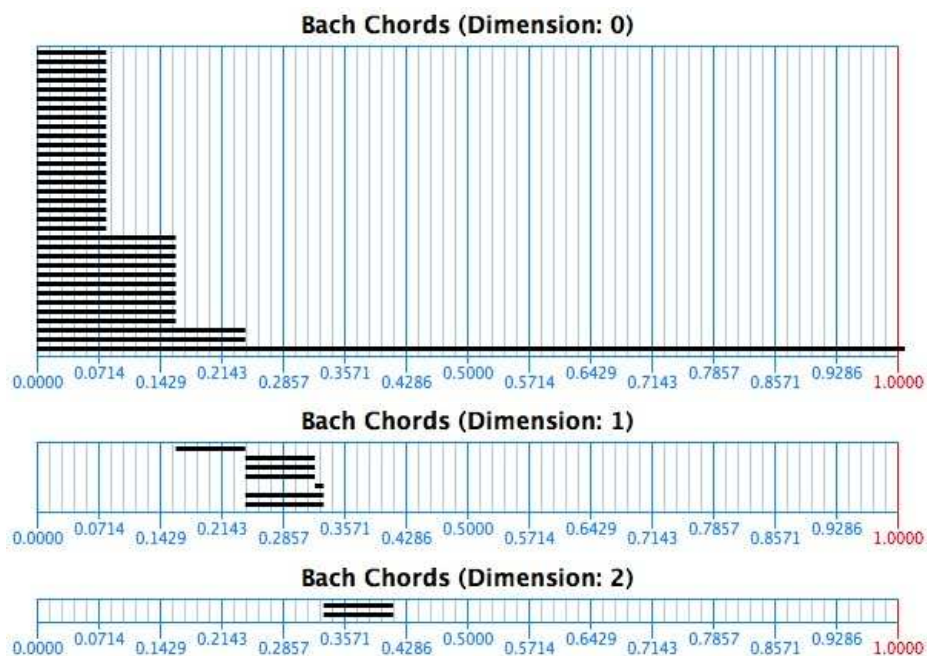
Figure D2. Barcodes for Bach's Chorale No. 19 under the chord class distance (**3**). The space is the nonorientable bundle $S^1 \times D^3$ (this is the $n = 4$ case of Example **3**.4) and the total diamater of the space is 2.0 while the maximum distance between any two elements is 1.0. This indicates that the data sits in a small 'ball' and fills only a small percent of the full space of chords.
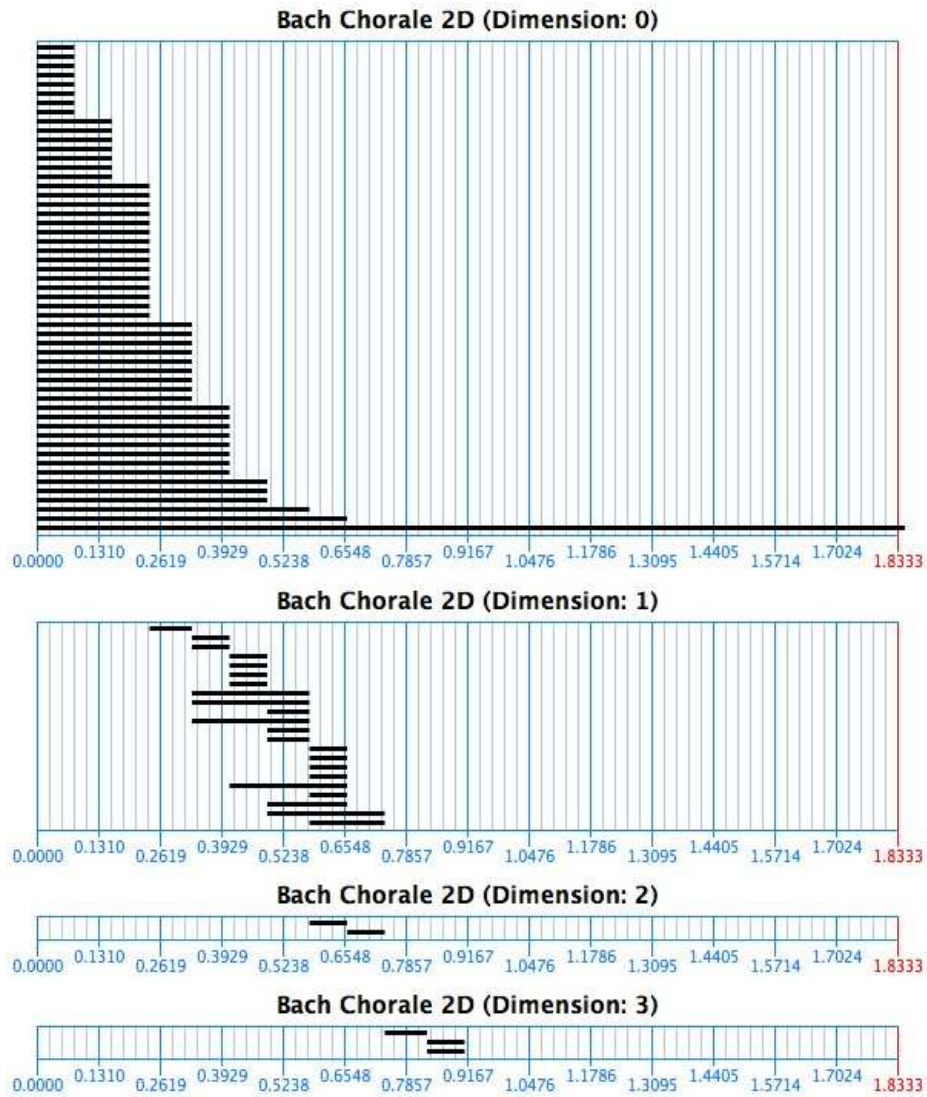
Figure D3. Barcodes for the two-dimensional time-delay embedding of Bach's Chorale No. 19. The distance function is (6). According to Proposition 3.7 with time delay $j = 2$ and chord-size $k = 4$. The diameter of the space is 4 and the maximum distance between elements is 1.833.

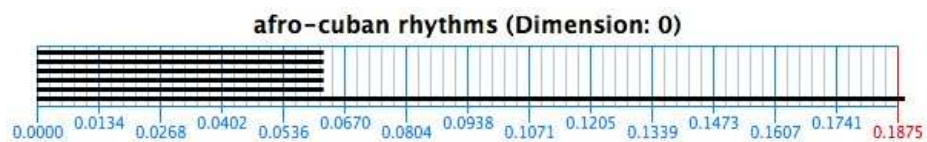## Appendix E. Rhythmic barcodes

Figure E1. Barcodes for the six rhythmic patterns in Table 1 under the rhythm distance (4). The max distance between elements is 0.1875.
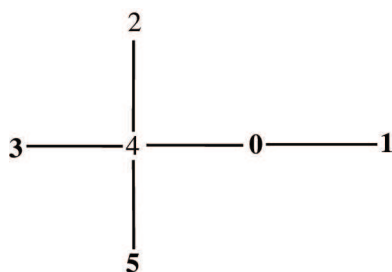

Figure E2. Representation of the underlying metric graph describing the data.

Appendix F. Topology of Musical Data