Topological data analysis of Korean music in Jeongganbo: a cycle structure

Mai Lan Tran\textsuperscript{a,b}, Changbom Park\textsuperscript{c}, and Jae-Hun Jung\textsuperscript{a,b,s}

\textsuperscript{a}Department of Mathematics, Pohang University of Science & Technology, Pohang, Korea; \textsuperscript{b}POSTECH Mathematical Institute for Data Science (MINDS), Pohang University of Science & Technology, Pohang, Korea; \textsuperscript{c}School of Physics, Korea Institute for Advanced Study, Seoul, Korea

(Received 10 March 2021; accepted 24 September 2022)

Jeongganbo is a unique music representation invented by Sejong the Great. Contrary to the Western music notation, the pitch of each note is encrypted and the length is visualized directly in a matrix form. We use topological data analysis (TDA) to analyze the Korean music written in Jeongganbo for Suyeonjang, Songuyeo, and Taryong, those well-known pieces played among noble community. We define the nodes of each music with pitch and length and transform the music into a graph with the distance between the nodes defined as their adjacent occurrence rate. The graph homology is investigated by TDA. We identify cycles of each music and show how those cycles are interconnected. We found that the cycles of Suyeonjang and Songuyeo, categorized as a special type of cyclic music, frequently overlap each other in the music, while those of Taryong, which does not belong to the same class, appear only individually.

Keywords: Korean music; Jeongganbo music; topological data analysis; persistent homology; cycles; cyclic music

Classification codes: AMS 00A65; 55N31

1. Introduction

Topological data analysis (TDA) via persistent homology provides an efficient way of analyzing the cycle or loop structures embedded in multidimensional data. Particularly the one-dimensional homology structure is closely related to the repeating patterns in music flow when it is mapped to the proper topological space. We consider, in this research, three famous Korean old music pieces, namely Suyeonjangjigok (Suyeonjang hereafter), Songkuyeojigok (Songkuyeo hereafter), and Taryong, and study their cycle structures when represented in a music network; the cycle structures are the key element that characterizes these music pieces. They are categorized as “Jeong-Ak,” meaning the “Right Music,” mostly played at the palace and among the noble community and written in Jeongganbo, a unique music representation invented by Sejong the Great of Joseon dynasty. Contrary to the Western music notation, the pitch of each note is encrypted and the length is visualized directly in a matrix form. As the music is written in Jeongganbo, it sometimes provides a special musical pattern according to the matrix form. In this article,
we study the cycle structures and how those cycles are interconnected over the music flows of *Suyeonjang*¹ and *Songkuyeo* and compare them with those of *Taryong*.

According to the earliest literature in 1451 (Korea-sa-ak-ji (高麗史類志) volumes 70 & 71), the primitive form of *Suyeonjang* and *Songkuyeo* appeared before the Joseon dynasty (1392–1897) in Korea. *Suyeonjang* is originally a Chinese feast dance played at the palace wishing for the king’s longevity. It later evolved throughout the Joseon dynasty to an orchestral ensemble with variations, and became one of the most representative Korean Jeong-Ak musics. *Songkuyeo* is the music composed in a similar pattern as *Suyeonjang* but with one octave higher. *Suyeonjang* and *Songkuyeo* have their unique music structure known as “Dodeuri,” which means repeat-and-return. As *Songkuyeo* is composed with one octave higher, it is also called “Wut-Dodeuri” (Upper Dodeuri) and *Suyeonjang* is referred to as “Mit-Dodeuri” (Lower Dodeuri).

The “Dodeuri” is a unique music structure. Its simplest pattern has a form of A-B-C-B where the second pattern C-B is the variation of the first pattern of A-B while the second part B of each pattern is repeated in both patterns. While this “Dodeuri” music is old and has been studied for long, research is rare that studies how to define repeating cycles, how these repeating cycles are interconnected over the music flow, and what characteristics of the “Dodeuri” patterns make both *Suyeonjang* and *Songkuyeo* unique in Jeong-Ak. In this study, we use the notion of persistent homology to define cycles embedded in *Suyeonjang* and *Songkuyeo* after we represent them as a music network. In order to construct the music network of these music pieces, we first define and determine the node elements that constitute the music, characterized uniquely by its pitch and length. Then we transform the music into a graph and introduce a proper metric to define the distance between the notes using their adjacent occurrence rate. If two nodes appear side-by-side then we frequently consider that those nodes are close. The graph with the metric is then used as a point cloud. For the analysis, we compute persistent homology of the constructed point cloud, especially for the one-dimensional homology. Through this research we identified cycles from each music as the cycles identified through persistent homology. Among those cycles found, some match the music, that is, appear in the music in their found form, while some do not appear in the music directly. Then, by visualizing the distribution of those cycles in the music, we showed how those cycles are interconnected in the music. The main discovery of this research is that the cycles of *Suyeonjang* and *Songkuyeo*, defined using homology, frequently overlap each other when appearing over the music, while the cycles found in *Taryong*, which does not belong to “Dodeuri” class, appear rather individually when they appear in the music. This gives an interesting musical effect whereby the listener feels that they hear multiple cycles simultaneously when they appear in the music.

The outline of the article is as follows. In the next section we briefly explain TDA, particularly persistent homology, and how we associate our music data with TDA tools in order to study the cycle structure in the music. The following section reviews basic characteristics of Korean musical notation in Jeongganbo, in particular for *Suyeonjang*, *Songkuyeo*, and *Taryong*. In the fourth section we construct a music network from music data and define the distance matrix which will be used to generate barcode for each music by Javaplex (Adams et al. 2014) with Vietoris-Rips method. This is followed by a section that provides the frequency of the occurrence of each node in music. It is observed that the frequency decays more exponentially than algebraically. Detailed analysis of the barcodes and comparison between *Suyeonjang*, *Songkuyeo*, and *Taryong* is given in the final section. At the end of the article we provide the list of all nodes of *Suyeonjang*, *Songkuyeo*, and *Taryong*, according to our node definition, in Tables A1, A2, and A3 in the Appendix.

¹ https://www.youtube.com/watch?v=DKo8Fjl7Mg&t=461s. Readers can listen to *Suyeonjang* played by Haegaeum instrument with the link from 0:24–5:24.
2. Topological data analysis

In this section we briefly explain TDA, particularly persistent homology. For more detailed explanation we refer readers to (Zomorodian and Carlsson 2005). The main motivation of using TDA is to study the cycle structure in music data. In this article, we focus on the one-dimensional cycle, which is reflected as a repeated pattern in music. The two-dimensional cycle (or two-dimensional void) would also be of interest, but its interpretation is uncertain at the moment. Below, we explain the basic concepts of TDA used in this article. Our approach is to extract a point cloud with a metre from each music, then build the Vietoris-Rips complex to get a filtered complex and thus compute persistent homology and barcodes. The point cloud used for TDA is constructed in the fourth section. The cycle structure extracted from the barcodes is then visualized and studied in the final section.

2.1. Simplicial complexes and simplicial homology

A \( k \)-simplex, \( \sigma_k = [x_0, x_1, \ldots, x_k] \), is the convex hull of \( k + 1 \) geometrically independent points (or vertices), \( \{x_0, x_1, \ldots, x_k\} \subset \mathbb{R}^n \), \( k \leq n \). This is a generalization of the notion of a triangle or tetrahedron to arbitrary dimensions. Roughly speaking, \( \sigma_n \) is an \( n \) dimensional triangle. For example, \( \sigma_0 \) is a point or vertex, \( \sigma_1 \) an edge, \( \sigma_2 \) a triangle and \( \sigma_3 \) a tetrahedron. A higher dimensional simplex is the higher dimensional equivalent to a triangle. In this article, the independent point is replaced with the node in a music network in the following section. Figure 1 shows 0-simplex (vertex), 1-simplex (edge) and 2-simplex (triangle) from left to right, respectively. Here note that the inside of triangle (n-dimensional triangle) is filled.

A simplicial complex \( K \) is a set of simplices that satisfies the following conditions:

(i) If \( \sigma \in K \), then all faces of \( \sigma \) are also in \( K \).
(ii) If \( \sigma_1, \sigma_2 \in K \), then \( \sigma_1 \cap \sigma_2 \) is either the empty set or a face of both \( \sigma_1 \) and \( \sigma_2 \).

Consider a ring \( R \) with unity and a topological space \( X \). In our case, \( R \) will be \( \mathbb{Q} \) and the topological space \( X \) is the music network constructed from the Jeongganbo music. Given \( (X, R) \), let \( C_n(X, R) \) be the free \( R \)-module generated by all possible continuous images of \( \sigma_n \). Consider \( \delta_n : C_n \to C_{n-1} \), the boundary map defined as below

\[
\delta_n \sigma_n = \sum_{k=0}^{n} (-1)^k [p_0, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n]
\]

where \( \{p_i\} \) are the vertices, the music nodes in our case. Then it is easy to show that \( \delta_n \circ \delta_{n+1} = 0 \), for all \( n = 0, 1, \ldots \). For \( n = 0 \), we choose \( \delta_0 \) as a trivial map. Using the relation of vanishing consecutive boundary maps, we consider two subgroups, the kernel and image groups, \( \text{ker}(\delta_n) := Z_n \) and \( \text{im}(\delta_{n+1}) := B_n \), respectively. The kernel group is also called the cycle group and the image group is also called the boundary group. Obviously, \( Z_n \subset C_n \) and \( B_n \subset C_n \). Further, \( B_n \subset Z_n \), i.e. every boundary group is a subgroup of a cycle group.
Using $B_n \subseteq \mathbb{Z}_n$, the $n$th homology group of $X$ with the coefficient $R$ is defined by the following quotient group

$$H_n(X, R) = \mathbb{Z}_n / B_n = \ker(\delta_n) / \text{im}(\delta_{n+1}).$$

For a detailed, rigorous description of homology we refer the reader to (Hatcher 2002; Munkres 1984). Practically the $n$-dimensional homology group indicates how many $n$-dimensional holes are there in $X$. For example, for the sphere $H_0 = \mathbb{Z}, H_1 = \mathbb{Z}, H_2 = \mathbb{Z}$ while for a torus $H_0 = \mathbb{Z}, H_1 = \mathbb{Z}_2, H_2 = \mathbb{Z}$. The existence and the number of holes are useful information that we will use in this research.

It is, however, difficult to compute $H_n$ in general and we use the notion of simplicial complex out of the given data to understand the hole structures. Roughly speaking, with the simplicial complex, we approximate the original topological space $X$. Once the simplicial complex is constructed, we find its homology. Related articles on using topological tools in music analysis are found in (Bergomi and Baratè 2020; Bigo et al. 2013) and references therein.

2.2. Filtered complex and persistent homology

A filtration of a simplicial complex $\mathcal{K}$ is a nested subsequence of complexes $\emptyset = \mathcal{K}_0 \subset \mathcal{K}_1 \subset \ldots \subset \mathcal{K}_m = \mathcal{K}$. For generality, we let $\mathcal{K}_i = \mathcal{K}_m$ for all $i \geq m$. We call $\mathcal{K}$ a filtered complex. That is, a filtered complex is an increasing sequence of simplicial complexes. Figure 2 shows an example of a filtered complex.

Given a filtered complex $\emptyset = \mathcal{K}_0 \subset \mathcal{K}_1 \subset \ldots \subset \mathcal{K}_m = \mathcal{K}$, let $Z_n^i$ and $B_n^i$ be the associated cycle and image groups of $\mathcal{K}_i$, respectively. The $p$-persistent $n$th homology group of $\mathcal{K}_i$ is

$$H_n^{i+p} = Z_n^i / (B_n^{i+p} \cap Z_n^i).$$

The persistent homology of the filtered complex is a homological feature that persists within the filtration. See (Bergomi 2015; Cohen-Steiner et al. 2007; Edelsbrunner and Harer 2009) for more details.

2.3. Point clouds and Vietoris-Rips complex

The tuple composed of pitch and length in Jeongganbo is regarded as a node and all the nodes in the music constitute a point cloud, which we will use to construct the simplicial complex. The simplicial complex is a topological space constructed by glueing a finite number of simplices. Since the exact topological space $X$ is unknown, we use its approximate topological space – that is, the simplicial complex. In this research, for the simplicial complex, we consider the Vietoris-Rips complex (Ghrist 2008).

The basic procedure is as follows. From the given point cloud, we create one-simplices by connecting vertices resulting in edges, then create two-simplices by connecting edges and repeat this procedure for the construction of higher simplices. For the connecting condition, we rely on the notion of metric – that is, the topology is described as metric topology. For the metric, we need to introduce the proper metric. In many cases, Euclidean metric is used to measure
Figure 3. The Vietoris-Rips complex. Given a point cloud (top left) and the filtration parameter, \( \tau \), a union of closed radius \( \tau / 2 \) balls (top right) is used to check if the distance between two vertices is less than \( \tau \) to create 1-simplices (bottom left). The 2-simplices (bottom right) is created when the pairwise distance between three vertices is less than \( \tau \).

the pair-wise distance between simplices. For our research, we will use a different definition of metric defined in the following section based on the occurrence of consecutive nodes in the music flow. The procedure of connecting simplices is conducted if the connecting condition is satisfied. For the connecting condition, let us introduce an auxiliary parameter, so-called the filtration parameter, \( \tau \). If the distance between two simplices is less than the given value of the filtration parameter, \( \tau \), those two simplices will be connected. This is illustrated in Figure 3. Once such a simplicial complex is constructed, we compute the homology group for a particular value of \( \tau \). Then we collect all the homology group in the sequence of \( \tau \). This provides the notion of persistent homology.

2.4. Barcodes

We also use the \( n \)-dimensional Betti number, \( \beta_n \). The \( n \)-dimensional Betti number is the number of generators of the \( n \)th homology group – that is, \( \beta_n \) is the number of copies of \( R \). When \( n = 0 \), \( \beta_0 \) simply means the total number of connected components. For example, for a torus, \( \beta_0 = 1 \), \( \beta_1 = 2 \) and \( \beta_2 = 1 \). Once the simplicial complex is constructed from the Jeongganbo data, we will measure the Betti number in each dimension. One can compute the \( n \)th Betti number of the constructed simplicial complex at a particular value of \( \tau \). The barcode is the collection of the Betti number with respect to the given value of \( \tau \) in each dimension. For example, if \( \tau = 0 \), no two vertices are to be connected and the 0th Betti number is the number of all connected components, which is simply the total number of vertices, nodes in our case. And the \( n \)th Betti number will be zero for \( n \geq 1 \). For computational purpose and data analysis, one can plot \( \beta_n \) versus \( \tau \), known as barcode. There exists a certain interval of \( \tau \) where the homology is invariant in \( n \) dimension. Then in barcode, such an interval will be visualized as a persistent line, which we will call a persistence. The starting and ending values of \( \tau \) on such a line are called the birth and death. The plot of death versus birth is called the persistence diagram.
2.5. **An example**

We consider the following simple data composed of four music nodes on the plane, each a vertex of a unit square. Then the minimum distance between two random nodes will be one and the maximum distance will be $\sqrt{2}$, the diagonal distance between the orange and red nodes or the green and blue nodes.

Now using various values of $\tau \geq 0$, let us connect nodes. Here note that if a triangle is formed as a result, the triangle should be filled. The following shows how these four nodes are connected with different value of $\tau = 0, \ 0.5, \ 1, \ \sqrt{2}, \ 2$. The Betti numbers $\beta_0$ and $\beta_1$ are also provided for the give $\tau$. For example, when $\tau_1 = 1$, all the nodes are connected, so there is just one connected component, that is, $\beta_0 = 1$. As shown in the figure, the connected nodes also form one hole, a square. Thus $\beta_1 = 1$. But when $\tau = \sqrt{2}$, the square becomes two triangles as the blue and green nodes are connected with $\tau = \sqrt{2}$ and the hole disappear as those two triangles are filled inside. Note that it is also possible to connect the red and orange nodes instead of connecting the blue and green nodes. But once one of them is established, the other one cannot be established because the non-empty intersection of any two simplices is a simplex.

For the above example we used only five values of $\tau$. If we use multiple values of $\tau$ as many as we want and plot $\beta_0$ and $\beta_1$ versus $\tau$, the following barcodes are obtained. We plot the points at each $\tau$ as many as the given Betti number in the vertical axis.

For later use, we define the persistence $\Pi_i^n$ denoting the length of the $i$th barcode in the $n$th dimension. For example, for the above example,

$$\Pi_1^0 = \Pi_2^0 = \Pi_3^0 = 1,$$
As \( \tau \to \infty \), there remains only one connected component implying \( \Pi^0_4 \to \infty \).

### 3. Korean musical notation of Jeongganbo

Jeongganbo was created some time before June 1447 by Sejong the Great of the Joseon dynasty and many musicians in order to overcome the limitations of existing musical notations including Yukbo, Yuliabo, and Gongcheokbo (Hwang et al. 2010). It is the first notation method that shows the exact length and height of the sound to be played. In Jeongganbo, each column is usually composed of 32, 16, 12, or 6 squares called Jeonggan (井間). The pitches are displayed by the first letters of each of the 12 Yulmyeong (Hwang, Dae, Tae, Hyeop, Go, Jung, Yu, Im, I, Nam, Mu, and Eung) and the length of each note is displayed by the number of Jeonggans. Typically one Jeonggan represents the equivalent of one quarter note (Park and Geem 2017). So, as shown in Figure 4, for example, if only one Hwang is placed in one Jeonggan, it counts as one beat; if Hwang and Tae are both in one Jeonggan and occupy an equal amount of space, this divides the rhythm into eighth notes, and so on. Within each Jeonggan, the notes are read from left to right and from top to bottom. The rule of reading the length of a pitch in Jeongganbo is illustrated in Figure 5.
In this article we study in particular three music pieces: Suyeonjang, Songkuyeo, and Taryong, mentioned earlier, played by Haegum instrument. In Suyeonjang and Songkuyeo each column contains six Jeonggans, while in Taryong each column contains 12 Jeonggans. In these music pieces five notes including Jung (G♯3), Im (A♯3), Nam (C4), Hwang (D♯4) and Tae (F4) are used, where Jung (G♯3) can be played two octaves higher and the remaining four notes can be played one octave higher, making a total of eleven pitches, that are G♯3, A♯3, C4, D♯4, F4, G♯4, A♯4, C5, D♯5, F5 and G♯5, as shown in Figure 6. The lengths of notes range from 1/6-Jeonggan to 6-Jeonggan.

Beside the main letters that show the notes to be played in Jeongganbo, in particular Suyeonjang, Songkuyeo, and Taryong, various musical symbols are used to notate pitch and length of the note or provide information about how to play the note. For the sake of simplicity but still keeping the main characteristics of the music pieces, we have left out kkumim-eums (ornamenting tones) and instrument-specific notations – that is, Haegum in our case – which give the performer information about special playing techniques such as how to control the bow stick and horsehair. Development of diverse kkumim-eums is one of the main characteristics of Korean traditional music. Detailed analysis of the effects of kkumim-eums should be made in future studies. Figure 7 illustrates the musical symbols used in Suyeonjang, Songkuyeo, and Taryong that we interpret. The symbol L notates one note higher than the previous note, meaning that, for example, if it appears after note Hwang, then the symbol L stands for note Tae, which is one pitch higher than Hwang. The length of this note is read according to the rule of reading the length of a pitch in Figure 5. Symbol h represents two notes, one is one pitch lower than the main note and the other is two pitches lower than the main note. So, if h is placed after note Jung, then two notes Tae and Hwang with equal length should be played. Symbol ⬕ stands for the note that is two pitches higher than the previous note. So if the symbol ⬕ appears after note Jung, then it is supposed to be note Nam. Symbol ONGL simply stands for a repetition of the previous note. The last symbol, ↓ so-called Ingeojil, represents a short higher note. So if it appears after note Im, then Im is played quite longer followed by a short note Nam before proceeding to the next note. We directly interpreted Jeongganbo into data for the construction of a music network described in the next section instead of transcribing it for piano in an equal temperament for the interpretation.

4. Construction of music network

A conventional approach of studying music is to use the network representation (Bryan and Wang 2011; Gomez et al. 2014; Itzkovitz et al. 2006; Liu et al. 2010). In our case, by constructing a music network we aim to extract a point cloud from each music piece, and define a metric on it; this is exactly what allows us to build the associated Vietoris-Rips complex and also compute persistent homology and barcode.

---

2 https://www.youtube.com/watch?v=_DKo8FjL7Mg&t=461s. Readers can listen to Suyeonjang played by Haegum instrument with the link from 0:24–5:24. The rest of the pieces in the link are the variations of Suyeonjang composed with the technique of artificial intelligence.
Let $G = (N, E)$ be the music network that we want to construct, where $N$ is the set of nodes and $E$ is the set of edges in $G$. We define each node, $n \in N$ in $G$ to be a pair of a pitch and its length as a tuple below

$$n = \text{(pitch, length).}$$

As mentioned in the previous section, Suyeonjang, Songkuyeo, and Taryong consist of eleven pitches including, $A\#_3$, $C_4$, $D\#_4$, $F_4$, $G\#_4$, $A_4$, $C_5$, $D\#_5$, $F_5$ and $G\#_5$ with the lengths of notes ranging from 1/6-Jeonggan to 6-Jeonggan. For example, if the $i$th node, $n_i$, is a pair of a pitch, say $G\#_3$, and its length, say one Jeonggan, then the node is represented by $n_i = (G\#_3, 1)$. Using this construction, the network of Suyeonjang, Songkuyeo, and Taryong contains in total 33, 37 and 40 nodes, respectively. Two nodes are connected if they occur adjacent in time. Let $e_{ij} \in E$ be the edge whose end points are $n_i$ and $n_j$. The weight or the degree of the edge $e_{ij}$, $w_{ij}$, between $n_i$ and $n_j$ is the number of occurrences of those two nodes being adjacent in time. For two distinct nodes $n_i$ and $n_j$ with $i < j$, let $p_{ij}$ be the path with the minimum number of edges between $n_i$ and $n_j$ found by using Dijkstra algorithm. Then we define the distance $d_{ij}$ between nodes $n_i$ and $n_j$ and form the distance matrix $D = \{d_{ij}\}$ as follows:

$$d_{ij} = \begin{cases} 
\sum_{e_{kl} \in p_{ij}} w_{kl}^{-1}, & i < j \\
0, & i = j
\end{cases} \quad \text{and} \quad d_{ij} = d_{ji}, i > j,$$

where $w_{kl}$ represents the weight of the edge $e_{kl}$ and $p_{ij} = \cup e_{kl}$.

It is clear that if two nodes $n_k$ and $n_l$ are connected by the edge $e_{kl}$ then $w_{kl} \neq 0$ and further $w_{kl} \geq 1$. Thus $w_{kl}^{-1}$ is well defined in the above distance definition. Also note that since there is no isolated node which is not an end point of any of edges in $E$ (due to the fact that there is no empty Jeonggan where the music is not played), there always exists at least one path between any two nodes $n_i$ and $n_j$ even if they do not appear adjacently in the whole music, that is, $p_{ij}$ exists, $\forall i < j$. Thus for this definition it is not necessary that $n_i$ and $n_j$ are adjacent. Therefore the distance defined in Eq. (2) is well-defined. This observation is valid at least for those music pieces we consider in this article, played with Haegeum. The metric is chosen in this way because we want to construct a space where the more closely connected nodes in networks are the closer in the space. Therefore, we took the inverse of the weight since we look at networks where larger weights indicate more closely connected relationships between nodes.
Figure 8. Frequency versus node. The horizontal axis represents the rank and the vertical axis the frequency in logarithmic scale. Left: Suyeonjang. Middle: Songkuyeo. Right: Taryong. Note that the minimum frequency is unity as the minimum occurrence of a particular node is at least one by definition.

Remark 4.1: In general, the triangular inequality, $d_{ik} \leq d_{ij} + d_{jk}$, does not hold. For example, if nodes $n_i, n_j, n_k$ are pairwise connected with weights: $w_{ij} = 1$, $w_{jk} = 1$, $w_{ik} = 1$, i.e. $d_{ik} = d_{ij} = d_{jk} = 1$, then $d_{ik} < d_{ij} + d_{jk}$. However, if $w_{ij} = 5$, $w_{jk} = 5$, $w_{ik} = 1$, i.e. $d_{ij} = d_{jk} = \frac{1}{5}$, $d_{ik} = 1$, then $d_{ik} > d_{ij} + d_{jk}$. The motivation for defining distance this way is that, if nodes $n_i$ and $n_j$ are connected and $n_j$ and $n_k$ are connected, but $n_i$ and $n_k$ are not connected and the shortest path from $n_i$ to $n_k$ is $n_i - n_j - n_k$, then $d_{ik} = d_{ij} + d_{jk}$.

Remark 4.2: Because of the definition of the distance in Eq. (2), we observe that the point cloud obtained from the constructed network of Suyeonjang is in non-Euclidean pseudo-metric space. Thus, in generating the barcode we use the method ExplicitMetricSpace in Javaplex (Adams et al. 2014) which works for a point cloud in an arbitrary, in particular non-Euclidean, pseudo-metric space.

All nodes of Suyeonjang, Songkuyeo, and Taryong, according to the node definition above, are provided in Tables A1, A2, and A3 in the Appendix. All nodes in each table are listed in ascending order in terms of pitch first then length. That is, $n_j$ has a higher pitch than $n_i$ or both have the same pitch but $n_j$ has a longer length than $n_i$ if $j > i$. Suyeonjang, Songkuyeo, and Taryong have, in total, 33, 37 and 40 nodes, respectively.

5. Frequency

We first provide the frequency of occurrence of each node in music. Table 1 shows the node frequency for Suyeonjang, Songkuyeo, and Taryong, left to right, respectively. For each music, the left column shows the node number and the right column the corresponding occurrence frequency. The frequencies are listed from first to last in rank in the table. Figure 8 shows the frequency versus rank in semi-logarithmic scale. The horizontal axis represents the rank and the vertical axis the frequency in logarithmic scale. Note that the frequency is a positive integer by definition. As shown in the figure, the frequency decays more exponentially than algebraically. By looking at the frequency distribution we could know which nodes are important for the music flow. In Section 6 we will see that the frequently appearing nodes constitute more cycles. For example, nodes $n_{18}$ in Suyeonjang, $n_{20}$ in Songkuyeo, and $n_{16}$ in Taryong constitute five, six, and three cycles, respectively. However, analyzing music using only the frequency of the node occurrence would not be enough in our case since nodes which rarely appear in the music may still constitute a few cycles and play an important role in the analysis. For example, nodes $n_{24}$ in Suyeonjang, $n_{30}$ in Songkuyeo, and $n_{37}$ in Taryong also constitute a cycle in each music even though they occur only once.
Table 1. Frequency versus node.

| rank | Suyeonjang | Songkuyeo | Taryong |
|------|------------|-----------|---------|
| 1    | $n_{18}$  | $n_{20}$  | $n_{16}$ |
| 2    | $n_{6}$   | $n_{31}$  | $n_{11}$ |
| 3    | $n_{11}$  | $n_{13}$  | $n_{13}$ |
| 4    | $n_{22}$  | $n_{26}$  | $n_{26}$ |
| 5    | $n_{1}$   | $n_{8}$   | $n_{29}$ |
| 6    | $n_{20}$  | $n_{18}$  | $n_{31}$ |
| 7    | $n_{27}$  | $n_{4}$   | $n_{28}$ |
| 8    | $n_{3}$   | $n_{33}$  | $n_{3}$  |
| 9    | $n_{28}$  | $n_{6}$   | $n_{18}$ |
| 10   | $n_{12}$  | $n_{16}$  | $n_{15}$ |
| 11   | $n_{16}$  | $n_{25}$  | $n_{12}$ |
| 12   | $n_{26}$  | $n_{19}$  | $n_{22}$ |
| 13   | $n_{31}$  | $n_{24}$  | $n_{10}$ |
| 14   | $n_{2}$   | $n_{27}$  | $n_{32}$ |
| 15   | $n_{4}$   | $n_{28}$  | $n_{9}$  |
| 16   | $n_{23}$  | $n_{32}$  | $n_{20}$ |
| 17   | $n_{9}$   | $n_{2}$   | $n_{4}$  |
| 18   | $n_{10}$  | $n_{15}$  | $n_{9}$  |
| 19   | $n_{5}$   | $n_{7}$   | $n_{0}$  |
| 20   | $n_{8}$   | $n_{11}$  | $n_{14}$ |
| 21   | $n_{13}$  | $n_{12}$  | $n_{2}$  |
| 22   | $n_{9}$   | $n_{14}$  | $n_{5}$  |
| 23   | $n_{7}$   | $n_{17}$  | $n_{7}$  |
| 24   | $n_{17}$  | $n_{21}$  | $n_{8}$  |
| 25   | $n_{19}$  | $n_{23}$  | $n_{19}$ |
| 26   | $n_{23}$  | $n_{15}$  | $n_{21}$ |
| 27   | $n_{25}$  | $n_{0}$   | $n_{27}$ |
| 28   | $n_{29}$  | $n_{3}$   | $n_{33}$ |
| 29   | $n_{30}$  | $n_{9}$   | $n_{34}$ |
| 30   | $n_{32}$  | $n_{10}$  | $n_{35}$ |
| 31   | $n_{14}$  | $n_{36}$  | $n_{1}$  |
| 32   | $n_{15}$  | $n_{34}$  | $n_{10}$ |
| 33   | $n_{24}$  | $n_{1}$   | $n_{23}$ |
| 34   | $n_{5}$   | $n_{24}$  | $n_{1}$  |
| 35   | $n_{22}$  | $n_{25}$  | $n_{1}$  |
| 36   | $n_{29}$  | $n_{30}$  | $n_{1}$  |
| 37   | $n_{30}$  | $n_{38}$  | $1$     |
| 38   | $n_{36}$  | $1$       | $1$     |
| 39   | $n_{37}$  | $1$       | $1$     |
| 40   | $n_{39}$  | $1$       | $1$     |

6. Analysis of the barcodes

In this section we analyze the barcode and give an interpretation of the persistence intervals in the first dimension for Suyeonjang, Songkuyeo, and Taryong.

6.1. Suyeonjang

The barcode for Suyeonjang generated by Javaplex with Vietoris-Rips method is given in Figure 9.

In Figure 9 the horizontal axis is the filtration value $\tau$. Vertically we have multiple intervals that correspond to generators of the homology groups. In the zeroth dimension we have 33 generators that correspond to 33 components when $\tau$ is small, which eventually are connected into a single component when $\tau = 1$. The 33 components are actually those 33 nodes defined in Suyeonjang. All these components constitute a single component because of the fact that any node in the
network connects at least one time with another node, which means that at most when distance $d = 1$ all nodes in the network are connected. In the first dimension we see eight generators which topologically correspond to eight cycles. A natural question is whether these cycles are related to the repetition of music melodies. In the following section we will go into detailed analysis about the musical meaning of these cycles.

The persistence algorithm computing intervals is used to find a representative cycle for each interval. The annotated intervals computed by the method `computeAnnotatedIntervals` in Javaplex tell us what the nodes in the intervals of persistence are. In the one-dimension case, the annotated intervals consist of the components in the loops generated in the process of filtration.

Figure 10 shows eight cycles identified by TDA corresponding to eight persistence intervals in dimension one of the barcode. The order of cycles is the order of their appearance in the barcode. That is, the earlier the 1D barcode dies, the lower number is assigned to the corresponding cycle. For example, the death of Cycle $i$ is earlier than the death of Cycle $j$ if $i < j$. Note that this order is different from the order of their appearance in the actual music. In the figure, each cycle is shown with the persistence interval, the node information (its node number and note), edge weight (in normal size in blue), distance between node (in small size in blue in parentheses) and the average weight (in red in center) which is the simple mean of all edge weights. As shown in the figure, the minimum number of nodes that constitute a cycle is four and the maximum number is six. The average node number is 4.625. The average weight is 9.39375. Cycle 2 corresponds to the shortest persistence in 1D barcode. Cycle 4 corresponds to the longest persistence in 1D barcode. The information of corresponding music notes is given in Table 2. It should be kept in mind that Suyeonjang has in total 33 nodes, and the cycles show which node combinations are of interest.
Figure 10. The 8 cycles identified by TDA in Suyeonjang. Each cycle shows the persistence interval, the node number, edge weight (normal size in blue), distance between node (small size in blue in parentheses) and the average weight (red in center). Cycle 2 corresponds to the shortest persistence in 1D barcode. Cycle 4 corresponds to the longest persistence in 1D barcode. Cycles 6, 7 and 8 do not appear in actual music in their whole consecutive form. Their corresponding persistences start from $d = 1$ in 1D barcode. The average node number is 4.625. The average weight is 9.39375.

The first attempt is to find in the music sheet the sequence of music notes appearing in each cycle. We have found that except for Cycles 6, 7, and 8, five remaining cycles indeed represent consecutive sequences of notes in Suyeonjang. Among them two sequences of music notes actually make up two closed Cycles 4 and 5. They appear in Suyeonjang in the following orders:
Table 2. Information of all nodes that appear in cycles identified by TDA in Suyeonjang. All nodes are listed in ascending order in terms of pitch. There are all 20 nodes that appear in any of eight cycles. The table shows the music note, pitch and length.

| Node symbol in cycles (Figure 10) | Name    | Music note | Length (Jeonggan) |
|----------------------------------|---------|------------|-------------------|
| \( n_0 \) 傳 (Jung)             | G\#3    | 1/3        |
| \( n_1 \) 傳 (Jung)             | G\#3    | 1          |
| \( n_2 \) 傳 (Jung)             | G\#3    | 2          |
| \( n_3 \) 傳 (Im)               | A\#3    | 1/3        |
| \( n_6 \) 傳 (Im)               | A\#3    | 1          |
| \( n_7 \) 傳 (Im)               | A\#3    | 5/3        |
| \( n_{11} \) 傳 (Nam)          | C4      | 1          |
| \( n_{12} \) 傳 (Nam)          | C4      | 5/3        |
| \( n_{16} \) 黃 (Hwang)        | D\#4    | 1/3        |
| \( n_{18} \) 黃 (Hwang)        | D\#4    | 1          |
| \( n_{20} \) 太 (Tae)          | F4      | 1/3        |
| \( n_{21} \) 太 (Tae)          | F4      | 2/3        |
| \( n_{22} \) 太 (Tae)          | F4      | 1          |
| \( n_{23} \) 太 (Tae)          | F4      | 5/3        |
| \( n_{24} \) 太 (Tae)          | F4      | 2          |
| \( n_{25} \) 仲 (Jung)         | G\#4    | 1/3        |
| \( n_{26} \) 仲 (Jung)         | G\#4    | 2/3        |
| \( n_{27} \) 仲 (Jung)         | G\#4    | 1          |
| \( n_{29} \) 仲 (Jung)         | G\#4    | 2          |
| \( n_{30} \) 林 (Im)            | A\#4    | 2/3        |

All nodes are listed in ascending order in terms of pitch. There are all 20 nodes that appear in any of eight cycles. The table shows the music note, pitch and length.

for Cycle 4 (appearing three times) and

for Cycle 5 (appearing one time), respectively. Note sequences in Cycles 2 and 3 are also found in the music in the following orders:
Figure 11. Consecutive sequences of nodes corresponding to Cycles 1, 2, 3, 4, 5 found in Suyeonjang. The horizontal axis is the time sequence of music flow and the vertical axis is for the cycles. The top bars represents Cycle 1 and the bottom Cycle 5. The width of each bar is the number of nodes contained in each cycle. The height of each bar is only for the visual distinction purpose. Note that Cycles 6, 7 and 8 are not appearing. Also, the consecutive sequence of nodes corresponding to each cycle appears in the music in a unique order which is not necessarily the same as the order of nodes in the cycle.

for Cycle 2 (appearing two times) and

for Cycle 3 (appearing one time), respectively. Although these two sequences do not make up closed cycles, they do preserve the order of edge connectivity of the cycles. For Cycle 1 we found two sequences of the same nodes but in a slightly different order.

In summary, by using TDA tools we obtain a set of eight persistence intervals that can be interpreted as eight cycles in dimension one. Our approach is to find sequences of notes in the music that correspond to these cycles. We have found that out of the eight cycles, five of them appear as consecutive sequences of notes in Suyeonjang (single or multiple times) and the other three are not found. The order in which notes appear in the music is not necessarily the same as the order of edge connectivity of the cycles. Although it is not clear at the moment the underlying reason for this different ordering, we observe that the ordering may not directly relate to the closeness in pitch of the notes. For example, note Jung (G♯4) is closer to Tae (F4) than to Hwang (D♯4) but, as shown above, a part of the sequence of notes corresponding to Cycle 1 in Suyeonjang that appears in the music is in the order n_{27} − n_{18} − n_{22} which is Jung–Hwang–Tae.

Figure 11 shows the five cycles appearing in Suyeonjang with time. The horizontal axis represents the time sequence the music flows and the vertical axis represents the cycle number, from 1 to 5. Ignore the number displayed in the axis. The width of each is the total number of nodes involved in the cycle. The height of each bar in the figure has no meaning but is displayed only for visual distinction purpose. The monotonicity means that the sequence is the same as the cycle sequence. The variance in color means the slight variation of the cycle sequence when appearing. For example, the figure shows that Cycle 4 (the 4th bar from top) appears three times in the first half of the music while the rest of the cycles appear in the second half of the music. We observe that the occurrence of the cycles in the complete form is sparse. Also note that, the consecutive sequence of nodes corresponding to each cycle appears in the music in a unique order which is not necessarily the same as the order of nodes in the cycle.

Figure 12 shows the similar plot as Figure 11, the cycle occurrence with respect to music flow. The only difference is that each cycle is displayed in the figure whenever at least four
Figure 12. Consecutive sequences of at least four nodes involved in cycles of Suyeonjang. Similar plot as Figure 11. Each cycle is displayed in the figure whenever at least four consecutive columns are occupied by any node from the cycle. Notice that Cycles 6, 7 and 8 are appearing in the figure.

Figure 13. Number of cycles that each node in Suyeonjang belongs to. The horizontal axis shows all nodes in Suyeonjang and the vertical axis shows the total number of cycles that each node belongs to and the actual cycle number. Consecutive columns are occupied by any node from the cycle. The order may not be the same for each occurrence. The spectrum shown in the width shows the variation of the order of nodes in the cycle when appearing in the music. We observe that the occurrence is now rather dense compared to Figure 11. We also observe that the invisible Cycles 6, 7, and 8 in Figure 11 are now appearing in the figure. In fact, the number of the occurrences of Cycle 7 is not small. Also notice that multiple cycles can be found simultaneously.

We remark here that the consecutive sequence is named as an Overlap Matrix in the following work.\(^3\) This newly developed concept and method are advantageous compared to the existing TDA in that this method is efficient to apply to various problems such as machine composition and classification.

Figure 13 shows the number of cycles that each of 33 nodes in Suyeonjang in Table A1 belongs to. The horizontal axis shows all 33 nodes in ascending order in terms of pitch and the vertical axis shows the total number of cycles that the node belongs to and cycle number together. For examples, each of the first four nodes \(n_0, n_1, n_2, n_3\) belongs to only one cycle, i.e. Cycle 4 (c4), Cycle 6 (c6), Cycle 4 (c4), and Cycle 2 (c2), respectively while \(n_4\) and \(n_5\) do not belong to any cycles and \(n_6\) belongs to five cycles, Cycle 2 (c2), Cycle 4 (c4), Cycle 5 (c5), Cycle 6 (c6), and Cycle 7 (c7).

\(^3\) M. Tran, D. Lee, J.-H. Jung, Machine composition of Korean music via topological data analysis and artificial neural network, arXiv:2203.15468, 2022.
We provide the similar data for Songkuyeo. Figure 14 shows the barcodes in 0D, 1D and 2D. There are 37 components when $\tau = 0$, which corresponds to the total number of nodes in Songkuyeo. In 1D barcode, there are eight non-zero persistences implying that there are eight cycles in music network as Suyeonjang. It is interesting to observe that there is one persistence in 2D. This implies that there exists a void-like structure in Songkuyeo network, although the size of the void is small as the size of the persistence is small.

Figure 15 shows the eight cycles in Songkuyeo. Cycle 1 corresponds to the shortest persistence in 1D barcode. Cycle 6 corresponds to the longest persistence in 1D barcode. Cycles 2, 3, 5, 6, 7 and 8 do not appear in actual music in their whole consecutive form. The persistences of Cycles 7 and 8 start from $\tau = 1$ in 1D barcode. The number of nodes in the cycles in Songkuyeo ranges from four to eight. Recall that the maximum number of nodes in a cycle for Suyeonjang is six. The average node number is 5.375 which is larger than the average node number of Suyeonjang. The average weight is 7.84125, which is smaller than the average weight of Suyeonjang. The information of corresponding music notes is given in Table 3.

Unlike Suyeonjang, there are only two cycles that appear in the music as a whole consecutive form, Cycle 1 and Cycle 4. Both of these sequences appear in the music in the exact same order as in the corresponding cycles.

The following Cycle 1 appears in the music but is not a closed cycle.
Figure 15. The 8 cycles identified by TDA in Songkuyeo. Each cycle shows the persistence interval, the node number, edge weight (normal size in blue), distance between node (small size in blue in parentheses) and the average weight (red in center). Cycle 1 corresponds to the shortest persistence in 1D barcode. Cycle 6 corresponds to the longest persistence in 1D barcode. Cycles 2, 3, 5, 6, 7 and 8 do not appear in actual music in their whole consecutive form. The persistences of Cycles 7 and 8 start from \( d = 1 \) in 1D barcode. The average node number is 5.375. The average weight is 7.84125.

The following Cycle 4 also appears in the music but is not a closed cycle.
Table 3. Information of all nodes that appear in cycles identified by TDA in Songkuyeo.

| Node symbol in cycles (Figure 15) | Music note |
|-----------------------------------|------------|
|                                   | Name       | Pitch | Length (Jeonggan) |
| 太 n6                             | 太 (Tae)   | F4    | 1/3               |
| 太 n8                             | 太 (Tae)   | F4    | 1                 |
| 仲 n11                            | 仲 (Jung)  | G♯4   | 1/6               |
| 仲 n12                            | 仲 (Jung)  | G♯4   | 2/3               |
| 仲 n13                            | 仲 (Jung)  | G♯4   | 1                 |
| 仲 n14                            | 仲 (Jung)  | G♯4   | 1.5               |
| 仲 n15                            | 仲 (Jung)  | G♯4   | 5/3               |
| 林 n17                            | 林 (Im)    | A♯4   | 1/6               |
| 林 n18                            | 林 (Im)    | A♯4   | 1/3               |
| 林 n20                            | 林 (Im)    | A♯4   | 1                 |
| 林 n21                            | 林 (Im)    | A♯4   | 5/3               |
| 南 n23                            | 南 (Nam)   | C5    | 1/6               |
| 南 n24                            | 南 (Nam)   | C5    | 1/3               |
| 南 n26                            | 南 (Nam)   | C5    | 1                 |
| 南 n27                            | 南 (Nam)   | C5    | 5/3               |
| 南 n28                            | 南 (Nam)   | C5    | 2                 |
| 南 n30                            | 南 (Nam)   | C5    | 6                 |
| 漢 n31                            | 漢 (Hwang) | D♯5   | 1                 |
| 漢 n32                            | 漢 (Hwang) | D♯5   | 2                 |
| 汰 n33                            | 汰 (Tae)   | F5    | 1                 |
| 汰 n34                            | 汰 (Tae)   | F5    | 11/6              |
| 中 n36                            | 中 (Jung)  | G♯5   | 1/6               |

All nodes are listed in ascending order in terms of pitch. There are all 20 nodes that appear in any of eight cycles. The table shows the music note, pitch and length.

Figure 16 shows the actual cycles appearing in Songkuyeo in their whole consecutive form. As shown in the figure, there are only two cycles, Cycle 1 and Cycle 4 appearing in the music.
Cycle 4 appears in the first half of the music and Cycle 1 in the second half of the music. As in Suyeonjang, the distribution of these cycles in the music is sparse.

Figure 17 shows the cycle occurrence with respect to music flow whenever at least four consecutive columns are occupied by any nodes in the cycle. The order may not be the same for each occurrence. The spectrum shown in the width shows the variation of the order of nodes in the cycle when appearing in the music. We observe that the occurrence becomes now dense. We also observe that all the cycles except Cycle 5 are appearing. It is interesting to notice that Cycle 4 still appears only three times. Also notice again that multiple cycles can be found simultaneously.

Figure 18 shows how many cycles are there that each of 37 nodes in Songkuyeo in Table A2 belongs to. The horizontal axis shows all 37 nodes in ascending order in terms of pitch and the vertical axis shows the total number of cycles that the node belongs to and cycle number together. The visualization follows the same format as Figure 13.

6.3. Taryong

Figure 19 shows the barcode of Taryong. As expected, there are 40 components in 0D barcode as there are 40 nodes in Taryong. We observe that there are 10 cycles in 1D barcode and one cycle (void) in 2D barcode. The number of 1D cycles of Taryong is larger than both those numbers of Suyeonjang and Songkuyeo (Figure 20).

The following figure shows the 10 cycles identified by TDA in Taryong. Each cycle shows the persistence interval, the node number, edge weight (normal size in blue), distance between nodes.
Figure 18. Number of cycles that each node in Songkuye helps to. The horizontal axis shows all nodes in Songkuye and the vertical axis shows the total number of cycles that each node belongs to and the actual cycle number.

Figure 19. Barcode of Taryong using Vietoris-Rips method. Notice that there is a 2D persistence in 2D barcode corresponding to the 2D void in music network. There are 10 cycles in 1D barcode. (small size in blue in parentheses), and the average weight (red in center). Cycle 1 corresponds to the shortest persistence in 1D barcode. Cycle 10 corresponds to the longest persistence in 1D barcode. Cycles 2, 3, 7, and 8 do not appear in actual music in their whole consecutive form. The corresponding persistences of Cycles 7 and 8 start from $\tau = 1$ while the persistences of Cycles 2 and 3 start earlier in 1D barcode. The average node number is 4.8. The average weight is 4.385. Notice that the number of cycles of Taryong is larger than those numbers of Suyeonjang and
Figure 20. The 10 cycles identified by TDA in Taryong. Each cycle shows the persistence interval, the node number, edge weight (normal size in blue), distance between node (small size in blue in parentheses) and the average weight (red in center). Cycle 1 corresponds to the shortest persistence in 1D barcode. Cycle 10 corresponds to the longest persistence in 1D barcode. Cycles 2, 3, 7 and 8 do not appear in actual music in their whole consecutive form. The corresponding persistences of Cycles 7 and 8 start from $\tau = 1$ while the persistences of Cycles 2 and 3 start earlier in 1D barcode. The average node number is 4.8. The average weight is 4.385.
Songkuyeo, but the average node number is smaller than that of Songkuyeo. In fact, despite the larger number of cycles, the occurrence plot of cycles appearing in the music is much sparser than those plots of Suyeonjang and Songkuyeo, shown in Figures 22 and 23.

Figures 21 and 22 show the consecutive sequences of nodes corresponding to Cycles 1, 4, 5, 6, 9, and 10 found in Taryong. Among them Cycles 1 and 9 are closed cycles. The top bars represent Cycle 1 and the bottom bars represent Cycle 10. Note that six cycles among all cycles are appearing in the music. As in Suyeonjang and Songkuyeo, the occurrence distribution is sparse.

Figure 23 shows the consecutive sequences of at least four nodes involved in cycles of Taryong. Cycles are displayed when at least four consecutive columns are occupied by any node in the cycle. Notice that all cycles except Cycle 2 are appearing in the figure. It is interesting to observe that unlike Suyeonjang and Songkuyeo the plot is still spare although Taryong has more cycles than Suyeonjang and Songkuyeo.
Figure 22. Consecutive sequences of nodes corresponding to Cycles 1, 4, 5, 6, 9 and 10 found in Taryong. The horizontal axis is the time sequence of music flow and the vertical axis is for the cycles. The top bars represent Cycle 1 and the bottom Cycle 10. The width of each bar is the number of nodes contained in each cycle. The height of each bar is only for visual distinction purposes. Note that six cycles among all cycles are appearing in the music.

Figure 23. Consecutive sequences of at least four nodes involved in cycles of Taryong. Cycles are displayed when at least four consecutive columns are occupied by any node in the cycle. Notice that all cycles except Cycle 2 are appearing in the figure.

Figure 24. Number of cycles that each node in Taryong belongs to. The horizontal axis shows all nodes in Taryong and the vertical axis shows the total number of cycles that each node belongs to and the actual cycle number. The visualization follows the same format of Figure 13.

Figure 24 shows how many cycles are there that each of 40 nodes in Taryong in Table A3 belongs to. The horizontal axis shows all 40 nodes in ascending order in terms of pitch and the vertical axis shows the total number of cycles that the node belongs to and the corresponding cycle number together. The visualization follows the same format of Figure 13.
Table 4. Information of all nodes that appear in cycles identified by TDA in Taryong.

| Node symbol in cycles (Figure 20) | Music note | Pitch | Length (Jeonggan) |
|----------------------------------|------------|-------|-------------------|
| n₀                               | атегор (Jung) | G⁴3   | 1/2               |
| n₂                               |  Paginator (Im) | A⁴3   | 1/2               |
| n₃                               |  Paginator (Im) | A⁴3   | 1                 |
| n₄                               |  Paginator (Im) | A⁴3   | 3/2               |
| n₆                               |  Paginator (Nam) | C⁴    | 1/2               |
| n₁₀                              | Paginator (Hwang) | D⁴4   | 1/2               |
| n₁₁                              | Paginator (Hwang) | D⁴4   | 1                 |
| n₁₂                              | Paginator (Hwang) | D⁴4   | 3/2               |
| n₁₃                              | Paginator (Hwang) | D⁴4   | 2                 |
| n₁₅                              | Paginator (Tae) | F⁴    | 1/2               |
| n₁₆                              | Paginator (Tae) | F⁴    | 1                 |
| n₁₈                              | Paginator (Tae) | F⁴    | 2                 |
| n₁₉                              | Paginator (Tae) | F⁴    | 3                 |
| n₂₀                              |  Paginator (Jung) | G⁴4   | 1/2               |
| n₂₁                              |  Paginator (Jung) | G⁴4   | 1                 |
| n₂₂                              |  Paginator (Jung) | G⁴4   | 3/2               |
| n₂₃                              |  Paginator (Jung) | G⁴4   | 2                 |
| n₂₄                              |  Paginator (Jung) | G⁴4   | 5/2               |
| n₂₅                              |  Paginator (Jung) | G⁴4   | 3                 |
| n₂₆                              |  Paginator (Im) | A⁴4   | 1                 |
| n₂₇                              |  Paginator (Im) | A⁴4   | 3/2               |
| n₂₉                              |  Paginator (Im) | A⁴4   | 2                 |
| n₃₂                              |  Paginator (Nam) | C⁵    | 1                 |
| n₃₃                              |  Paginator (Nam) | C⁵    | 2                 |
| n₃₄                              |  Paginator (Nam) | C⁵    | 6                 |
| n₃₇                              |  Paginator (Hwang) | D⁴5   | 17/6              |
| n₃₈                              |  Paginator (Tae) | F⁵    | 1/6               |

All nodes are listed in ascending order in terms of pitch. There are 27 nodes that appear in any of 10 cycles. The table shows the music note, pitch and length.
### Table 5. Comparison of cycle distributions.

|          | # of cycles | Average node # | Average weight | Occurrence/Cycle | Denseness   | Overlap (%) |
|----------|-------------|----------------|----------------|-----------------|-------------|-------------|
| Suyeonjang | 8           | 4.625          | 9.39375        | 1.125           | 0.09489     | 36.23188    |
| Songkuyeo | 8           | 5.375          | 7.84125        | 0.625           | 0.10153     | 32.05128    |
| Taryong   | 10          | 4.8            | 4.385          | 0.7             | 0.03194     | 0           |

### 6.4. Discussion

We compare Figures 12, 17, and 23 in Table 5. The denseness is defined as follows

\[ \text{Denseness} = \frac{A_c}{A_f} \]

where \( A_c \) is the total area that consecutive sequences of at least four nodes belonging to the occupying cycles, and \( A_f \) is the total area of the corresponding Overlap Matrix. We see that the denseness of cycles of Suyeonjang and Songkuyeo are about three times larger than that of Taryong. It is also observed that, unlike Suyeonjang and Songkuyeo, the consecutive sequences of at least four nodes involved in cycles of Taryong are non-overlapping, meaning that only one cycle occurs at a time as music flows. So, if we define the overlapping percentage by

\[ \text{Overlap} = \frac{N_s}{N_c} \times 100\% \]

where \( N_s \) is the number of times at least two cycles occurred simultaneously, and \( N_c \) is the number of times a cycle occurred in the time sequence of music flow, then the consecutive sequences of at least four nodes involved in the cycles of Suyeonjang and Songkuyeo overlap approximately 36.2% and 32.1%, respectively, and those of Taryong overlap 0%. The results are shown in the last column of Table 5. It is reasonable to hypothesize that the cycle overlapping gives Suyeonjang and Songkuyeo a feeling of melodic repetition which is well matched with the fact that they are classified as “Dodeuri,” literally translated as “repetition,” in Korean traditional music.

### 7. Conclusion

In this work we used persistent homology to identify the cycles in three famous old Korean music pieces, namely Suyeonjang, Songkuyeo, and Taryong, by constructing the simplicial complex of music network of those three pieces. By calculating homology at different scales, we found several one-dimensional cycles in each music. By visualizing the distribution of the cycles in the music flow, we showed how those identified cycles appear through the music flow. We found that the first two music pieces, Suyeonjang and Songkuyeo, show similar overlapping patterns of those cycles in a way that cycles appear simultaneously, while Taryong shows different patterns that cycles only appear individually. Our finding provides a systematic explanation of the unique structure embedded in Suyeonjang and Songkuyeo known as “Dodeuri” (repeat-and-return) pattern. That is, “Dodeuri” pattern does not only repeat or variate themes but also juxtaposes those in parallel to maximize its musical inspiration of cyclic patterns. In this article, we showed that persistent homology provides a useful tool to analyze the Korean music written in Jeongganbo. Our current work did not consider the ornamenting tones and the analysis was based on the instrument-specific notes. In our future work, we plan to consider the ornamenting tones and expand the analysis to other instruments using persistent homology.
Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Research Foundation under Grant No. 2018R1D1A1A02086097 and 2021R1A2C3009648.

References

Adams, H., Tausz, A., and Vejdemo-Johansson, M. 2014. “javaPlex: A Research Software Package for Persistent (Co)Homology.” Lecture Notes in Computer Science 8592: 129–136. doi:10.1007/978-3-662-44199-2_23

Bergomi, Mattia. 2015. “Dynamical and Topological Tools for (Modern) Music Analysis.” Ph.D. thesis. Université Pierre et Marie Curie, Paris.

Bergomi, Mattia G., and Adriano Baratè. 2020. “Homological Persistence in Time Series: An Application to Music Classification.” Journal of Mathematics and Music 14 (2): 204–221. doi:10.1080/17459737.2020.1786745.

Bigo, Louis, Moreno Andreatta, Jean-Louis Giavitto, Olivier Michel, and Antoine Spicher. 2013. “Computation and Visualization of Musical Structures in Chord-Based Simplicial Complexes.” In MCM 2013 – 4th international conference mathematics and computation in music, Montreal, Canada, June, 38–51. Springer.

Bryan, Nicholas J., and Ge Wang. 2011. “Musical Influence Network Analysis and Rank of Sample-Based Music.” In Proceedings of the 12th international society for music information retrieval conference, Miami, United States, Oct., 329–334. ISMIR. doi:10.5281/zenodo.1415768.

Cohen-Steiner, D., J. Harer, and H. Edelsbrunner. 2007. “Stability of Persistence Diagrams.” Discrete and Computational Geometry 37:103–37120.

Edelsbrunner, Herbert, and John L. Harer. 2009. Computational Topology: An Introduction. American Mathematical Society.

Ghrist, Robert. 2008. “Barcodes: The Persistent Topology of Data.” Bulletin (New Series) of The American Mathematical Society 45.

Gomez, Florian, Tom Lorimer, and Ruedi Stoop. 2014. “Complex Networks of Harmonic Structure in Classical Music.” In Nonlinear Dynamics of Electronic Systems, edited by Valeri M. Mladenov and Plamen Ch. Ivanov, 262–269. Cham: Springer International Publishing.

Hatcher, A. 2002. Algebraic Topology. Cambridge: Cambridge University Press.

Hwang, Jun-yon, Jin-Ah Kim, and Yong-sik Yi. 2010. Musical Notations of Korea. Seoul, Korea: National Gugak Center.

Itzkovitz, Shalev, Ron Milo, Nadav Kashtan, Reuven Levitt, Amir Lahav, and Uri Alon. 2006. “Recurring Harmonic Walks and Network Motifs in Western Music.” Advances in Complex Systems (ACS) 9: 121–132. doi:10.1142/S021952590600063X

Liu, Xiao Fan, Chi K. Tse, and Michael Small. 2010. “Complex Network Structure of Musical Compositions: Algorithmic Generation of Appealing Music.” Physica A: Statistical Mechanics and its Applications 389 (1): 126–132. https://www.sciencedirect.com/science/article/pii/S0378437109006827. doi:10.1016/j.physa.2009.08.035

Munkres, James R. 1984. Elements of Algebraic Topology. Vol. 2. Cambridge, Massachusetts: Addison-Wesley Reading.

Park, Y. J., and Zong Woo Geem. 2017. “Korean Traditional Music Instrument and Educational Mobile Tool.” Information (Japan) 20: 3889–3899.

Zomorodian, Afra, and Gunnar Carlsson. 2005. “Computing Persistent Homology.” Discrete & Computational Geometry 33: 249–274. doi:10.1007/s00454-004-1146-y
Appendix A. All nodes in Suyeonjang, Taryong, Songkuyeo

Table A1. All nodes in Suyeonjang.

| Node symbol | Music note | Name | Pitch | Length (Jeonggan) |
|-------------|------------|------|-------|-------------------|
| \( n_0 \)  | 両 (Jung)  | G\#3 | 1/3   |                   |
| \( n_1 \)  | 両 (Jung)  | G\#3 | 1     |                   |
| \( n_2 \)  | 両 (Jung)  | G\#3 | 2     |                   |
| \( n_3 \)  | 両 (Im)    | A\#3 | 1/3   |                   |
| \( n_4 \)  | 両 (Im)    | A\#3 | 2/3   |                   |
| \( n_5 \)  | 両 (Im)    | A\#3 | 5/6   |                   |
| \( n_6 \)  | 両 (Im)    | A\#3 | 1     |                   |
| \( n_7 \)  | 両 (Im)    | A\#3 | 5/3   |                   |
| \( n_8 \)  | 両 (Nam)   | C4   | 1/6   |                   |
| \( n_9 \)  | 両 (Nam)   | C4   | 1/3   |                   |
| \( n_{10} \) | 両 (Nam) | C4   | 2/3   |                   |
| \( n_{11} \) | 両 (Nam) | C4   | 1     |                   |
| \( n_{12} \) | 両 (Nam) | C4   | 5/3   |                   |
| \( n_{13} \) | 両 (Nam) | C4   | 2     |                   |
| \( n_{14} \) | 両 (Nam) | C4   | 3     |                   |
| \( n_{15} \) | 両 (Nam) | C4   | 5     |                   |
| \( n_{16} \) | 黄 (Hwang) | D\#4 | 1/3   |                   |
| \( n_{17} \) | 黄 (Hwang) | D\#4 | 2/3   |                   |
| \( n_{18} \) | 黄 (Hwang) | D\#4 | 1     |                   |
| \( n_{19} \) | 黄 (Hwang) | D\#4 | 2     |                   |
| \( n_{20} \) | 太 (Tae)  | F4   | 1/3   |                   |
| \( n_{21} \) | 太 (Tae)  | F4   | 2/3   |                   |
| \( n_{22} \) | 太 (Tae)  | F4   | 1     |                   |
| \( n_{23} \) | 太 (Tae)  | F4   | 5/3   |                   |
| \( n_{24} \) | 太 (Tae)  | F4   | 2     |                   |
| \( n_{25} \) | 両 (Jung) | G\#4 | 1/3   |                   |
| \( n_{26} \) | 両 (Jung) | G\#4 | 2/3   |                   |
| \( n_{27} \) | 両 (Jung) | G\#4 | 1     |                   |
| \( n_{28} \) | 両 (Jung) | G\#4 | 5/3   |                   |
| \( n_{29} \) | 両 (Jung) | G\#4 | 2     |                   |
| \( n_{30} \) | 林 (Im)   | A\#4 | 2/3   |                   |
| \( n_{31} \) | 林 (Im)   | A\#4 | 1     |                   |
| \( n_{32} \) | 林 (Im)   | A\#4 | 2     |                   |
Table A2. All nodes in Songkuyeo.

| Node symbol | Name | Pitch | Length (Jeonggan) |
|-------------|------|-------|------------------|
| n₀          | 侻 (Im) | A₄3  | 1                |
| n₁          | 健 (Nam) | C₄  | 1                |
| n₂          | 黃 (Hwang) | D₅  | 1/3              |
| n₃          | 黃 (Hwang) | D₅  | 2/3              |
| n₄          | 黃 (Hwang) | D₅  | 1                |
| n₅          | 黃 (Hwang) | D₅  | 2                |
| n₆          | 大 (Tae) | F₄  | 1/3              |
| n₇          | 大 (Tae) | F₄  | 2/3              |
| n₈          | 大 (Tae) | F₄  | 1                |
| n₉          | 大 (Tae) | F₄  | 5/3              |
| n₁₀         | 大 (Tae) | F₄  | 2                |
| n₁₁         | 佔 (Jung) | G₅  | 1/6              |
| n₁₂         | 佔 (Jung) | G₅  | 2/3              |
| n₁₃         | 佔 (Jung) | G₅  | 1                |
| n₁₄         | 佔 (Jung) | G₅  | 1.5              |
| n₁₅         | 佔 (Jung) | G₅  | 5/3              |
| n₁₆         | 佔 (Jung) | G₅  | 2                |
| n₁₇         | 林 (Im) | A₄4  | 1/6              |
| n₁₈         | 林 (Im) | A₄4  | 1/3              |
| n₁₉         | 林 (Im) | A₄4  | 2/3              |
| n₂₀         | 林 (Im) | A₄4  | 1                |
| n₂₁         | 林 (Im) | A₄4  | 5/3              |
| n₂₂         | 林 (Im) | A₄4  | 2                |
| n₂₃         | 南 (Nam) | C₅  | 1/6              |
| n₂₄         | 南 (Nam) | C₅  | 1/3              |
| n₂₅         | 南 (Nam) | C₅  | 2/3              |
| n₂₆         | 南 (Nam) | C₅  | 1                |
| n₂₇         | 南 (Nam) | C₅  | 5/3              |
| n₂₈         | 南 (Nam) | C₅  | 2                |
| n₂₉         | 南 (Nam) | C₅  | 3                |
| n₃₀         | 南 (Nam) | C₅  | 6                |
| n₃₁         | 漢 (Hwang) | D₅  | 1                |
| n₃₂         | 漢 (Hwang) | D₅  | 2                |
| n₃₃         | 汀 (Tae) | F₅  | 1                |
| n₃₄         | 汀 (Tae) | F₅  | 11/6             |
| n₃₅         | 汀 (Tae) | F₅  | 2                |
| n₃₆         | 汀 (Jung) | G₅  | 1/6              |
Table A3. All nodes in Taryong.

| Node symbol | Music note | Pitch | Length (Jeonggan) |
|-------------|------------|-------|--------------------|
|             | Name       |       |                    |
| n0          | 旬 (Jung)  | G♯5  | 1/2                |
| n1          | 旬 (Jung)  | G♯5  | 1                  |
| n2          | 旬 (Im)    | A♯5  | 1/2                |
| n3          | 旬 (Im)    | A♯5  | 1                  |
| n4          | 旬 (Im)    | A♯5  | 3/2                |
| n5          | 旬 (Im)    | A♯5  | 2                  |
| n6          | 旬 (Nam)   | C4   | 1/2                |
| n7          | 旬 (Nam)   | C4   | 1                  |
| n8          | 旬 (Nam)   | C4   | 2                  |
| n9          | 旬 (Nam)   | C4   | 3                  |
| n10         | 黄 (Hwang) | D♯5  | 1/2                |
| n11         | 黄 (Hwang) | D♯5  | 1                  |
| n12         | 黄 (Hwang) | D♯5  | 3/2                |
| n13         | 黄 (Hwang) | D♯5  | 2                  |
| n14         | 黄 (Hwang) | D♯5  | 3                  |
| n15         | 太 (Tae)   | F4   | 1/2                |
| n16         | 太 (Tae)   | F4   | 1                  |
| n17         | 太 (Tae)   | F4   | 3/2                |
| n18         | 太 (Tae)   | F4   | 2                  |
| n19         | 太 (Tae)   | F4   | 3                  |
| n20         | 仲 (Jung)  | G♯4  | 1/2                |
| n21         | 仲 (Jung)  | G♯4  | 1                  |
| n22         | 仲 (Jung)  | G♯4  | 3/2                |
| n23         | 仲 (Jung)  | G♯4  | 2                  |
| n24         | 仲 (Jung)  | G♯4  | 5/2                |
| n25         | 仲 (Jung)  | G♯4  | 3                  |
| n26         | 林 (Im)    | A♯4  | 1                  |
| n27         | 林 (Im)    | A♯4  | 3/2                |
| n28         | 林 (Im)    | A♯4  | 11/6               |
| n29         | 林 (Im)    | A♯4  | 2                  |
| n30         | 林 (Im)    | A♯4  | 3                  |
| n31         | 南 (Nam)   | C5   | 1/6                |
| n32         | 南 (Nam)   | C5   | 1                  |
| n33         | 南 (Nam)   | C5   | 2                  |
| n34         | 南 (Nam)   | C5   | 6                  |
| n35         | 黃 (Hwang) | D♯5  | 1/2                |
| n36         | 黃 (Hwang) | D♯5  | 2                  |
| n37         | 黃 (Hwang) | D♯5  | 17/6               |
| n38         | 沈 (Tae)   | F5   | 1/6                |
| n39         | 沈 (Tae)   | F5   | 2                  |