Graph theory tools applied to the harmonic analysis of J. S. Bach: towards an integration of network models and conventional analytical approaches

Juan Felipe Vásquez R.[1]**, Hanner Cajar C.**[2]

[1]Universidad Nacional de Colombia - Universidade Estadual de Campinas
Bogotá, Colombia - Campinas, SP, Brasil

[2]Universidad Distrital Francisco José de Caldas - Universidad Nacional de Colombia
Bogotá, Colombia

jufvasquezri@unal.edu.co, hacajarc@udistrital.edu.co

Abstract. *This paper applies network analysis methodologies—specifically graph theory and associated centrality measures—to the study of two compositions by Johann Sebastian Bach from the Baroque period (17th–18th centuries). Each composition is segmented into discrete harmonic events, which are then modeled as nodes within a graph, with edges representing their functional or temporal relationships. The structural roles of individual events are assessed through centrality metrics (such as degree, closeness, and betweenness), revealing the presence of hierarchical or functionally pivotal elements within the harmonic network. Results indicate that network-derived centralities can align with, and potentially enrich, traditional music-theoretical interpretations, offering an alternative yet complementary analytical framework.*

1. Introduction

Throughout history, various musical systems have been created and characterized, where musical expression, performance, and interpretation are represented through graphic notation. However, traditional staff notation can present significant difficulties for analytical understanding among individuals without formal music education—a type of training that, in the Latin American context, is often absent from the core curricula of general education [1].

Over time, analytical tools have also been developed to classify works from specific historical periods into different ”musical styles.” This gave rise to terms such as ”Baroque,” ”Classicism,” ”Romanticism,” ”Impressionism,” among others, where composers share certain aesthetic and compositional features. However, these analytical tools are also largely inaccessible to the general public, as they require prior knowledge of music theory to understand the relationships among musical, harmonic, and temporal events.

In this study, we aim to shift the discourse slightly away from traditional academic music analysis and toward network analysis, as we believe it can provide a complementary perspective on the development of a particular musical work. However, we are also aware that graph and network analysis requires a certain degree of mathematical and geometrical knowledge, which may distance it from a general audience. For this reason, we will attempt, as much as possible, to provide clear explanations of all the concepts involved.

Our analysis is grounded in interdisciplinary tools. While it relies substantially on traditional academic music analysis—particularly on the study of the relationships among harmonic degrees (chords) within a musical work—it introduces a complementary approach through the application of network theory. In this framework, each harmonic degree is represented as a node, and the transitions between them are represented as edges in a graph.

We analyzed the chorales “Du Friedensfürst, Herr Jesu Christ” (Chorale #42) and “Werde munter, mein Gemüte” (Chorale #233), both taken from *371 vierstimmige Choralgesänge für ein Tasteninstrument (Orgel, Klavier, Cembalo)*. We employed some analytical set theory tools developed by [2, 3] for the numerical classification of pitch classes. For the network analysis associated with each piece, we were guided by the methods proposed by Liu [4, 5]. The `music21` Python library [6] was also essential, as it enabled both pitch classification and its integration with network theory.

For the network analysis, we also relied on definitions and concepts from Wasserman [7] and Newman [8]. Finally, the software Gephi [9] was used to visualize the networks extracted from the chorales, making it easier to understand their harmonic and compositional processes.

2. Methodology

The data used for this study were extracted directly from the musical scores of the chorales (see Fig. 1 and Fig.2). These scores were subjected to a standard harmonic analysis as traditionally taught in academic music institutions. Based on this information, the data were encoded to perform the analysis presented in this work.

We employed several tools from network analysis as well as elements from set theory to examine two chorales from the Baroque period. The analysis was carried out vertically (i.e., it is a harmonic rather than melodic analysis), in order to establish the relationships between harmonic degrees and their role in the structural development of the musical pieces.

For data collection and characterization, we used the scores of the two chorales and, through a series of steps, encoded the musical information into a numerical system. This representation proved useful for modeling harmonic (vertical) processes within the musical compositions.

2.1. Characterization of pitches (or tonal degrees)

Since both chorales are in the key of A major, it was possible to establish a reference system in which tonal and harmonic degrees can be treated as numbers. First, we express the musical notes using the American notation, in which tonal degrees are represented by letters from A to G.

La:A Si:B Do:C Re:D Mi:E Fa:F Sol:G

This representation is quite useful graphically, as it helps us understand the order of the notes through alphabetical order. After having represented the notes using American notation, we proceeded to assign a number from 0 to 11 to each note (or pitch) found within the musical octave from A_n to A_{n+1} :

$$\begin{gathered} A_n = 0 \\ A\#_n = 1 \\ B_n = 2 \\ C_n = 3 \\ C\#_n = 4 \\ D_n = 5 \\ D\#_n = 6 \\ E_n = 7 \\ F_n = 8 \\ F\#_n = 9 \\ G_n = 10 \\ G\#_n = 11 \\ A_{n+1} = A_n = 0 \end{gathered} \tag{1}$$

The last term means that the register (or octave) is not considered when characterizing the pitch, always applying the condition that for any note x_n with octave number n, then $n = n + 1$, and therefore $x_n = x_{n+1} = x$. Since octaves are not taken into account, we call x the **pitch class** [2, 3] .

2.2. Harmonic characterization

Using the numerical coding previously established for each note, it is possible to understand musical chords as sets containing the numbers that represent each pitch class. It is common to find chords of three or four notes in music: this is very important for our analysis since the sets representing chords will have three or four elements.

Our sets, when characterizing chords, require certain rules to create a solid foundation for musical harmonic representation:

2.2.1. *The first element of the set is the root note of the chord*

This means the chord will be represented in the form

$$X\,(x, c_1, c_2, \ldots, c_n)$$

where X is the chord name in American notation and x is the number representing the root pitch class of the chord.

Example: If we want to write the A minor chord (Am), the first element of the set must be the pitch class A; considering that in this work we established note A as our reference zero.

$$Am(0, 3, 7)$$

2.2.2. *If the score contains a chord with notes repeated in different octaves, the set only includes the number representing the pitch class*

If a chord Y contains notes c_n and c_{n+1}, the set only includes one of them; that is, its **pitch class**

$$Y(c_n, a, b, c_{n+1}) = Y(c, a, b) \quad ; \quad c \neq a \neq b$$

Example:

If the chord in the score is C major, in which the note C appears twice in different octaves, the set only lists the pitch class C once.

Figure 3: C major chord on the staff

Based on Figure 3, and taking into account the characterization of pitch classes established in (1), the C major chord can be constructed as:

$$Cmajor : C(3, 7, 10)$$

2.3. Chord Construction

In music, chords are characterized by the intervals between the different notes that compose them.

For this work, we have considered **major, Maj7, minor, minor 7, 7, diminished**, and **suspended** chords.

The intervallic characteristics of each chord type are specified in Table 1.

Harmonic correspondences in Chorale #1

The sets extracted from Chorale #1 are presented in Table 2.

To relate the chords, their temporal sequence in the score was taken into account; that is, there is only a relationship between chords if they are directly connected in time. Now, the relationship between these sets is represented in Table 3.

Figure 1: Score of Chorale #1

Figure 2: Score of Chorale #2

Graphical representation of chord Y (American notation)	Harmonic meaning	Intervallic characteristics
Y	Major	Y(y,y+4,y+7)
YMaj7	Major, minor 7th	Y(y,y+4,y+7,y+11)
Ym	Minor	Y(y,y+3,y+7)
Ym7	Minor, minor 7th	Y(y,y+3,y+7,y+10)
Y7	Major, minor 7th	Y(y,y+4,y+7,y+10)
Ydim	Diminished	Y(y,y+3,y+6,(y+9))
Ysus	Suspended	Y(y,y+5,y+7)

Table 1: Intervallic characteristics for each chord type

ID	Nodes Chorale 1
1	$\mathrm{A}(0,4,7)$
2	$\mathrm{Bm7}(2,5,9,0)$
3	$\mathrm{C\#}(4,8,11)$
4	$\mathrm{D}(5,9,0)$
5	$\mathrm{DMaj7}(5,9,0,4)$
6	$\mathrm{E}(7,11,2)$
7	$\mathrm{E7}(7,11,2,5)$
8	$\mathrm{F\#m}(9,0,4)$
9	$\mathrm{G\#\,dim}(11,2,5,8)$

Table 2: Correspondence between the different harmonic degrees of Chorale #1 and the ID assigned to each chord for recognition by the graphing software.

Harmonic correspondences in Chorale #2

For Chorale #2, the same procedure was followed as for Chorale #1. First, the chords were represented using the sets of pitch classes that compose them (see Table 4). Then, the direct temporal directed relationships were identified, as shown in Table 5.

3. Results

Using the graphing software *Gephi*, we obtained the graphs for each chorale (see Figures 4 and 5), which clearly show **centrality and significant movements** between the different sets.

3.1. Parameters and analysis

For the creation of tables, the relationships between chords—represented as nodes—will be taken into account. These relationships are represented by edges, which go from sources (initial chord) to targets (arrival chord). The weight indicates how many times the relationship occurs between the source and the target. Type D (Directed) refers to the relationship between a starting node (source) and an ending node (target) (see Table 3).

3.1.1. Degree Centrality (G)

Degree centrality is one of the most important metrics, as it indicates the number of chords related to a given chord X. Since chord relationships in chorales are temporal, it is pertinent to introduce the **In-degree** (G_{in}) and the **Out-degree** (G_{out}) of each chord, which establish the number of different chords that lead into X and those that X transitions to, respectively.

For **chorale #1**, the chord with the highest G_{in} and G_{out} is $G_{1inA(0,4,7)} = G_{1outA(0,4,7)} = 4$, with a total

Source (Node ID)	Target (Node ID)	Weight	Type
1	5	2	D
5	9	2	D
9	1	2	D
1	6	4	D
6	4	3	D
4	7	2	D
7	1	3	D
1	4	2	D
4	1	3	D
1	2	3	D
2	6	3	D
6	1	3	D
6	7	2	D
7	8	1	D
8	9	1	D
9	3	1	D
3	8	1	D
8	6	1	D

Table 3: Relationships between the chords presented in *Table 2*.

G of $G_{1A(0,4,7)} = 8$. This means that 4 chords lead into $A(0,4,7)$, and $A(0,4,7)$ transitions to 4 different chords. Therefore, there is a total of 8 chordal connections for the chord $A(0,4,7)$. Following $A(0,4,7)$, the chord with the next highest G_{in} and G_{out} is $E(7,11,2)$.

We can also introduce the **Weighted In-degree** ($G_{in.P}$) and the **Weighted Out-degree** ($G_{out.P}$), which indicate the total number of chords (including temporally repeated chords) that lead into X and to which X transitions, respectively. It is important to clarify that this measure may include repeated chords, and thus:

$$G_{in} \leq G_{in.P}$$
$$G_{out} \leq G_{out.P}$$

For chorale #1, the chord with the highest $G_{1in.P}$ and $G_{1out.P}$ is $A(0,4,7)$, with a total of $G_{1in.P,A(0,4,7)} = G_{1out.P,A(0,4,7)} = 11$; therefore, a total weighted degree can be defined as $G_{Total.P} = G_{in.P} + G_{out.P}$.

The Degree Centrality for each chord in chorale #1 is illustrated in Table 6, which clearly displays G_{1in}, G_{1out}, G_1, $G_{1in.P}$, $G_{1out.P}$, and $G_{1Total.P}$.

For **chorale #2**, the chord with the highest G_{in}, G_{out}, and therefore G, is $A(0,4,7)$, with total values:

$$G_{2in,A} = G_{2out,A} = 6$$

$$G_{2A} = 12$$

Similarly, the weighted in-degree and out-degree for A in chorale #2 are:

$$G_{2in.P,A} = G_{2out.P,A} = 15$$

ID	Nodes Chorale 2
1	A (0,4,7)
2	A7 (0,4,7,10)
3	Asus (0,5,7)
4	A#dim (1,4,7)
5	Bm7 (2,5,9,0)
6	Bm (2,5,9)
7	C#m (4,7,11)
8	D (5,9,0)
9	E (7,11,2)
10	Esus (7,0,2)
11	E7 (7,11,2,5)
12	F#m7 (9,0,4,7)
13	F# (9,1,4)
14	F#sus (9,2,4)
15	F#7 (9,1,4,7)
16	G#dim (11,2,5,8)

Table 4: Correspondence between the various harmonic degrees of Chorale #2 and the ID assigned to each chord for recognition by the graphing software.

Source	Target	Weight	Type
1	8	4	D
8	7	2	D
7	12	2	D
12	5	2	D
5	16	2	D
16	1	2	D
1	9	4	D
9	1	4	D
8	9	3	D
1	10	3	D
10	9	3	D
9	11	2	D
11	1	4	D
9	4	1	D
4	6	1	D
6	14	1	D
14	13	1	D
13	15	1	D
15	6	1	D
6	1	1	D
1	2	2	D
2	8	2	D
8	3	1	D
3	1	1	D
8	1	3	D
9	7	1	D
7	8	1	D
1	12	1	D
12	8	1	D
1	11	1	D

Table 5: Table of Sources, Targets, and Weights for Chorale #2

And therefore:

$$G_{2Total.P,A} = 30$$

The Degree Centrality for each chord in chorale #2 is illustrated in Table 7, which clearly displays G_{2in}, G_{2out}, G_2, $G_{2in.P}$, $G_{2out.P}$, and $G_{2Total.P}$.

Average Degrees $< G_{in} >, < G_{out} >$

The average degrees $< G_{in} >$ and $< G_{out} >$ for each network are defined as:

$$< G_{in} > = \frac{1}{N}\sum_{i}^{N} G_{in}(X_i)$$
$$< G_{out} > = \frac{1}{N}\sum_{i}^{N} G_{out}(X_i)$$

where N is the total number of chords and X_i is the i-th chord in the chorale.

Following the definitions above, we find the average degrees for each chorale:

$$< G_{1in} > = \frac{1}{9}(4+3+2+2+2+2+1+1+1) = 2$$
$$< G_{1out} > = \frac{1}{9}(4+3+2+2+2+2+1+1+1) = 2$$

Which reveals:

$$< G_{1in} > = < G_{1out} >$$

This indicates that in chorale #1, the average number of chords leading into a given chord X is the same as the average number of chords to which chord X moves. The fact that $< G_{1in} > = < G_{1out} > = 2$ shows that, on average, a chord X in chorale #1 transitions to two other chords in the development of the chorale, and approximately two chords lead into X.

For chorale #2, we have:

$$< G_{2in} > = 1.875 \approx 2$$
$$< G_{2out} > = 1.875 \approx 2$$

Thus,

$$< G_{2in} > = < G_{2out} >$$

We observe that, even though chorale #2 has more nodes than chorale #1, the average in-degree and out-degree are the same for both chorales. The values of $< G_{in} >$ and $< G_{out} >$ for both chorales indicate that, in these compositions, a given chord is expected to transition to approximately two other chords.

3.1.2. Eigenvector Centrality (E.C.)

This centrality measure indicates the global prominence of each node within the network. In other words, it reveals how important a chord is within the musical piece based on the relationships it maintains with other chords.

Mathematically, eigenvector centrality is defined as:

$$x_i = \frac{1}{\lambda}\sum_{j=1}^{n} A_{ij}x_j, \qquad (2)$$

where:

- x_i is the centrality score of node i,
- A_{ij} is the element of the adjacency matrix representing the connection between nodes i and j,
- x_j is the centrality score of node j,
- λ is the largest eigenvalue of the adjacency matrix,
- n is the total number of nodes in the network.

This equation expresses that the centrality of a node is proportional to the sum of the centralities of its neighboring nodes.

For chorale 1, the highest E.C. value is found in chord A, with a normalized value of 1.0, followed by chords E(7,11,2) with 0.732 and D(5,9,0) with 0.726. For the remaining chords, the values are below 0.7. It is expected that chord A would have the highest value, since it is the tonal center of the piece, which is also reflected in the definition of the model, where this chord plays a leading role, as all other degrees are defined in relation to it. Musically speaking, it is also consistent that these three chords would show greater centrality, as they correspond to the three main degrees of the key and help establish the tonal regions.

Similarly, for chorale 2, E.C. is concentrated in the same three chords. However, unlike chorale 1, the values for E(7,11,2) and D(5,9,0) are slightly higher. Additionally, many more nodes appear in this chorale compared to the first, and the distribution of relationships is slightly more complex. These two fundamental differences can be interpreted musically by considering that chorale 2 features brief modulations (or tonicizations) towards B minor and E; in the model, this is represented by the presence of two additional ”sinks” besides the main one (A). For these new tonal centers, there are several shared chords (nodes), meaning that—even if they belong to a different system—they can still relate to the original one without needing to venture far or seek connections outside of the accessible domain. Musically, chorale 2 offers more variety than chorale 1. As we will see in the graphs later, the relationships between nodes are more complex in chorale 2 due to the greater number of nodes and the roles played by each chord in different tonal contexts.

Since this is a normalized quantity, it is possible to directly compare the values obtained for each chord across the two chorales. Right away, this centrality reveals that there is significantly more movement in chorale 2, as

Label	In-deg.	Out-deg.	Deg.	Weighted in-deg.	Weighted out-deg.	Weighted deg.
A (0,4,7)	4	4	8	11.0	11.0	22.0
E (7,11,2)	3	3	6	8.0	8.0	16.0
D (5,9,0)	2	2	4	5.0	5.0	10.0
E7 (7,11,2,5)	2	2	4	4.0	4.0	8.0
F#m (9,0,4)	2	2	4	2.0	2.0	4.0
G#dim (11,2,5,8)	2	2	4	3.0	3.0	6.0
Bm7 (2,5,9,0)	1	1	2	3.0	3.0	6.0
C# (4,8,11)	1	1	2	1.0	1.0	2.0
DMaj7 (5,9,0,4)	1	1	2	2.0	2.0	4.0

Table 6: Degree and weighted degree measures for each chord note in Chorale #1.

Label	In-deg.	Out-deg.	Deg.	Weighted In-deg.	Weighted Out-deg.	Weighted deg.
A (0,4,7)	6	6	12	15.0	15.0	30.0
D (5,9,0)	4	4	8	8.0	9.0	17.0
E (7,11,2)	3	4	7	10.0	8.0	18.0
Bm (2,5,9)	2	2	4	2.0	2.0	4.0
C#m (4,7,11)	2	2	4	3.0	3.0	6.0
F#m7 (9,0,4,7)	2	2	4	3.0	3.0	6.0
E7 (7,11,2,5)	2	1	3	3.0	4.0	7.0
A7 (0,4,7,10)	1	1	2	2.0	2.0	4.0
Asus (0,5,7)	1	1	2	1.0	1.0	2.0
A#dim (1,4,7)	1	1	2	1.0	1.0	2.0
Bm7 (2,5,9,0)	1	1	2	2.0	2.0	4.0
Esus (7,0,2)	1	1	2	3.0	3.0	6.0
F# (9,1,4)	1	1	2	1.0	1.0	2.0
F#sus (9,2,4)	1	1	2	1.0	1.0	2.0
F#7 (9,1,4,7)	1	1	2	1.0	1.0	2.0
G#dim (11,2,5,8)	1	1	2	2.0	2.0	4.0

Table 7: Degree and weighted degree measures for each chord node in Chorale #2.

Label	Eigenvector Centrality
$\mathrm{A}(0,4,7)$	1.000000
$\mathrm{E}(7,11,2)$	0.731889
$\mathrm{D}(5,9,0)$	0.725831
$\mathrm{E7}(7,11,2,5)$	0.608676
$\mathrm{Bm7}(2,5,9,0)$	0.417307
$\mathrm{DMaj7}(5,9,0,4)$	0.417307
$\mathrm{F\#m}(9,0,4)$	0.319668
$\mathrm{G\#dim}(11,2,5,8)$	0.314582
$\mathrm{C\#}(4,8,11)$	0.139792

Table 8: Eigenvector centrality values for each chord node in Chorale #1.

there is a greater number of nodes and we observe relationships such as $E.C.(E)_2 > E.C.(E)_1$ and $E.C.(D)_2 > E.C.(D)_1$ for the main nodes within each chorale.

3.1.3. Betweenness Centrality (B.C.)

This centrality highlights the nodes through which the most shortest paths pass in our network. That is, among all possible paths between any two nodes, it quantifies how many of them pass through a particular node X. In the current model, this coincides with the chord most frequently used by the composer as a bridge within the work, across different harmonic progressions.

Betweenness centrality $C_B(v)$ of a node v is defined as:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{3}$$

where σ_{st} is the total number of shortest paths from node s to node t, and $\sigma_{st}(v)$ is the number of those paths that pass through node v.

For chorale 1, we observe that this centrality is concentrated on node A(0,4,7), which is expected, since every chord progression eventually passes through the chord that serves as the tonal center of the piece (even if the two chords being connected are far apart). This is not the case for other chords in the network. For example, node Bm7(2,5,9,0) has a betweenness centrality value of zero, which will be explained later.

On the other hand, in chorale 2, we observe a similar behavior. Considering tonicizations, we find that the nodes with the highest B.C. are precisely those that served as tonal centers at some point. Additionally, these nodes can act as secondary chords in particular harmonic progressions depending on the tonal context.

In this case, we also see a greater number of nodes

Label	Eigenvector Centrality
$\mathrm{A}(0,4,7)$	1.000000
$\mathrm{A7}(0,4,7,10)$	0.357263
$\mathrm{Asus}(0,5,7)$	0.309015
$\mathrm{A\#dim}(1,4,7)$	0.278185
$\mathrm{Bm7}(2,5,9,0)$	0.196117
$\mathrm{Bm}(2,5,9)$	0.114919
$\mathrm{C\#m}(4,7,11)$	0.587200
$\mathrm{D}(5,9,0)$	0.883702
$\mathrm{E}(7,11,2)$	0.794193
$\mathrm{Esus}(7,0,2)$	0.357263
$\mathrm{E7}(7,11,2,5)$	0.635448
$\mathrm{F\#m7}(9,0,4,7)$	0.559670
$\mathrm{F\#}(9,1,4)$	0.028647
$\mathrm{F\#sus}(9,2,4)$	0.051526
$\mathrm{F\#7}(9,1,4,7)$	0.020164
$\mathrm{G\#dim}(11,2,5,8)$	0.070932

Table 9: Eigenvector centrality values for each chord node in Chorale #2.

Label	Betweenness Centrality
$\mathrm{A}(0,4,7)$	0.455357
$\mathrm{G\#dim}(11,2,5,8)$	0.285714
$\mathrm{E}(7,11,2)$	0.285714
$\mathrm{F\#m}(9,0,4)$	0.223214
$\mathrm{E7}(7,11,2,5)$	0.125000
$\mathrm{DMaj7}(5,9,0,4)$	0.089286
$\mathrm{D}(5,9,0)$	0.035714
$\mathrm{C\#}(4,8,11)$	0.035714
$\mathrm{Bm7}(2,5,9,0)$	0.000000

Table 10: Betweenness centrality values for each chord node in Chorale #1.

(compared to chorale 1) with a B.C. value equal to zero. To understand this, we can refer to Graphs 1 and 2. When examining which nodes have B.C. = 0 (Bm7 in chorale 1, and A7, Esus, E7, and Asus in chorale 2), we see they all share the characteristic of connecting only a single pair of chords—and in all cases, those pairs are already directly connected. In terms of network efficiency, this makes perfect sense: if node 1 already has a direct connection to node 3, there is no need to go through an intermediary node 2 to reach it.

However, this is not the case in music. While theoretically one could move (almost) freely from one chord to another, depending on the musical style, it is sometimes necessary to pass through certain chords to reach a target chord. Moreover, the use of these chords—despite being structurally irrelevant according to this indicator—provides invaluable variety and artistic richness, making a piece or musical fragment much more pleasing than it would be without them.

3.1.4. Closeness Centrality (C.C.)

This centrality is associated with the node (or nodes) that have the shortest relative path length to all other nodes in

Label	Betweenness Centrality
$\mathrm{A}(0,4,7)$	0.614286
$\mathrm{Bm}(2,5,9)$	0.409524
$\mathrm{E}(7,11,2)$	0.333333
$\mathrm{D}(5,9,0)$	0.252381
$\mathrm{A\#dim}(1,4,7)$	0.209524
$\mathrm{F\#sus}(9,2,4)$	0.128571
$\mathrm{F\#}(9,1,4)$	0.128571
$\mathrm{F\#7}(9,1,4,7)$	0.128571
$\mathrm{F\#m7}(9,0,4,7)$	0.128571
$\mathrm{Bm7}(2,5,9,0)$	0.066667
$\mathrm{G\#dim}(11,2,5,8)$	0.066667
$\mathrm{C\#m}(4,7,11)$	0.038095
$\mathrm{A7}(0,4,7,10)$	0.000000
$\mathrm{Esus}(7,0,2)$	0.000000
$\mathrm{E7}(7,11,2,5)$	0.000000
$\mathrm{Asus}(0,5,7)$	0.000000

Table 11: Betweenness centrality values for each chord node in Chorale #2.

the network. In our model, it indicates the chord that is closest to all others in terms of distance and connectivity.

The closeness centrality $C_C(v)$ of a node v is defined as:

$$C_C(v) = \frac{1}{\sum_{t \neq v} d(v,t)} \tag{4}$$

where $d(v,t)$ is the length of the shortest path between node v and node t.

For both chorale 1 and chorale 2, the results align with expectations. The chords that served as tonal centers at some point have the highest C.C. In chorale 1, the C.C. values cluster around 0.21, a relatively low value that suggests all chords are roughly equidistant from one another and from the tonal center. This implies that there are various options to connect one chord to another, and from a practical standpoint, it's not very difficult to travel from any chord X to any chord X' (following the style's conventions, which are inherently encoded in the graph and the measures). If we were to analyze other musical styles with different paradigms, we would certainly observe very different relationships.

In chorale 2, we find very similar values. The range for C.C. is also approximately 0.21, which again indicates that all chords are well-connected to one another. The average distance between any node and the rest of the chords is relatively small.

3.1.5. Significant Movement

The most significant movement evidenced in the adjacency matrix (as well as in the weight matrix) and visually represented in the graphs is the one formed by the triangle [A(0,4,7) → E(7,11,2) → E7(7,11,2,5)]. This can serve as a compositional tool by suggesting recurrent movements within a musical piece.

3.2. Graphs

Graphs were constructed to support the analysis of the chorales. In each graph, chords are represented as nodes and directed edges represent the temporal connections (or relationships) that exist between different chords. The size of the nodes was set proportional to their degree, while the size of the edges was set proportional to their weight.

It was gratifying and quite exciting to see how both graphs show centrality focused on the same chord, as well as similar chord movements.

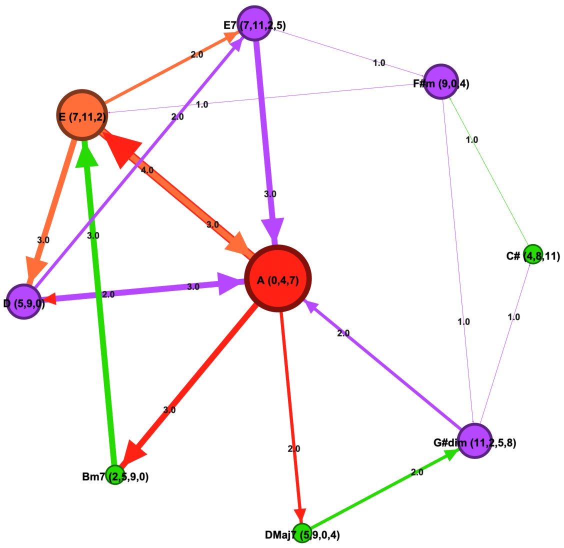

Figure 4: Graph generated for chorale #1

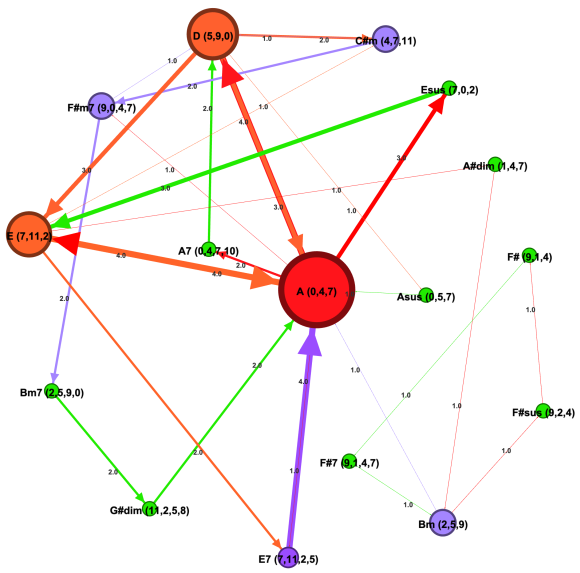

Figure 5: Graph generated for chorale #2

4. Conclusions

- Using centrality measures from network analysis, a clear centrality towards the chord $A(0, 4, 7)$ is evidenced, which is coherent given that the tonality of both chorales is A major.
- Centrality measures, in addition to being effective analytical tools, can also serve as vital compositional tools when used as parameters for the creation and connection of musical chords.
- The centrality measures for tonal musical analysis that yield the most relevant information are **Degree, Eigenvector Centrality, Betweenness Centrality, and Closeness Centrality**, since a proper interpretation of each can allow them to be proposed as parameters for musical creation.
- Musical encoding into numerical sets allows for the analysis of any musical piece in any key, as long as the model formulation and the coding proposal are based on solid and unambiguous principles.
- The generation of graphs emerges as a valuable resource for analyzing chordal movements, as it makes temporal and harmonic relationships more evident and offers deeper insight into the organizational principles of the music, complementing and enriching traditional analytical methods.

References

[1] Rolando Angel-Alvarado. The crisis in music education resulting from the demise of educational institutions. *Educación*, 44(1):598–611, June 2020. Accessed: 2025-08-26.

[2] Joseph N. Straus. *Introduction to Post-Tonal Theory*. W. W. Norton & Company, 4th edition, 2016.

[3] John Rahn. *Basic Atonal Theory*. Schirmer Books, New York, 1987.

[4] Xiao Fan Liu, Chi K. Tse, and Michael Small. Composing music with complex networks. In Jie Zhou, editor, *Complex Sciences*, pages 2196–2205, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[5] Xiao Fan Liu, Chi K. Tse, and Michael Small. Complex network structure of musical compositions: Algorithmic generation of appealing music. *Physica A: Statistical Mechanics and its Applications*, 389(1):126–132, 2010.

[6] Michael Scott Cuthbert and Christopher Ariza. music21: A toolkit for computer-aided musicology and symbolic music data. In J. Stephen Downie and Remco C. Veltkamp, editors, *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pages 637–642, Utrecht, Netherlands, August 9-13 2010. Author's final manuscript.

[7] Peter J. Carrington, John Scott, and Stanley Wasserman, editors. *Models and Methods in Social Network Analysis*, volume 28 of *Structural Analysis in the Social Sciences*. Cambridge University Press, Cambridge, 2005.

[8] Mark Newman. *Networks*. Oxford University Press, 07 2018.

[9] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1):361–362, Mar. 2009.