Tension ribbons: Quantifying and visualising tonal tension.

HERREMANS, D; Chew, E; Second International Conference on Technologies for Music Notation and Representation

To be presented at Second International Conference on Technologies for Music Notation and Representation

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/xmlui/handle/123456789/11798

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk
Tension ribbons: Quantifying and visualising tonal tension

Dorien Herremans
Centre for Digital Music
School of Electronic Engineering
and Computer Science
Queen Mary University of London
d.herremans@qmul.ac.uk

Elaine Chew
Centre for Digital Music
School of Electronic Engineering
and Computer Science
Queen Mary University of London
elaine.chew@qmul.ac.uk

ABSTRACT

Tension is a complex multidimensional concept that is not easily quantified. This research proposes three methods for quantifying aspects of tonal tension based on the spiral array, a model for tonality. The cloud diameter measures the dispersion of clusters of notes in tonal space; the cloud momentum measures the movement of pitch sets in the spiral array; finally, tensile strain measures the distance between the local and global tonal context. The three methods are implemented in a system that displays the results as tension ribbons over the music score to allow for ease of interpretation. All three methods are extensively tested on data ranging from small snippets to phrases with the Tristan chord and larger sections from Beethoven and Schubert piano sonatas. They are further compared to results from an existing empirical experiment.

1. INTRODUCTION

Musical tension forms an essential part of the experience of listening to music. According to [1], increasing tension can be qualitatively described as “a feeling of rising intensity or impending climax, while decreasing tension can be described as a feeling of relaxation or resolution”. However, defining tension in a more quantitative, formalized way is a difficult problem. In previous studies, different characteristics have been used to try to model musical tension. These aspects are usually rooted in either the domain of psychology or that of music. From the psychological point of view, models look at influential factors such as expectation and emotion [2, 3, 4]; and semantic meaning of lyrics [5]. From a more low-level musical point of view, examined features include rhythm and timing [6, 7, 1]; harmonic tonal perception through Lerdahl’s tonal tension model [8, 9, 10]; pitch height/melodic contour [11, 12]; dynamics [13, 12]; timbral elements (roughness, brightness, and density) [14, 6]; and pitch register [12, 7]. It must be noted that most of the above mentioned low-level musical features can also be linked to expectation.

No one particular feature, however, seems to be decisive in predicting the experience of tension [14]. Listening to music is an aggregate experience that requires the integration of many different features. A listener’s attention can focus on one feature at a particular time and then shift to a different feature or combination of features [1].

In this research we explore tonality as one of the dimensions of musical tension. Three methods for quantifying aspects of tension are developed based on the spiral array [15], a geometric model for tonality. The system developed outputs the results of the methods as ribbons over a musical score. In the next section, these different methods are discussed, followed by an analysis of selected musical fragments, which include snippets previously analysed in an empirical study by [1].

2. THEORY

In this paper, three methods that capture aspects of perceived tonal tension are developed and discussed based on the spiral array, a model for tonality. We first give a brief review of the spiral array, then introduce the three methods: cloud diameter, cloud momentum, and tensile strain.

2.1 Spiral array

The spiral array is a three dimensional representation of pitch classes, chords and keys. Each pitch class is represented as spatial coordinates along a helix [16]. The three dimensional representations allows higher level musical entities such as chords and keys to be embedded in the helix. The exact formula of the pitch class helix implemented in this paper is as follows:

\[ x = r \times \sin(t), y = r \times \cos(t), z = a \times t, \]

where \( r = 1, a = \sqrt{\frac{2}{15}} \times \frac{z}{2}, t \in [\infty, \infty], \) and \( t \in \mathbb{R} \).

Close tonal relationships (such as the perfect fifth) are mirrored by their spatial proximity in the spiral array. Figure 1 shows that notes which sound tonally close are in fact positioned close to each other inside the array. This is illustrated by the C major chord, which only consists of spatially close pitches. Notes are positioned one perfect fifth away from each other (a quarter turn in the spiral), which results in notes positioned “above” each other representing a major third [15].

Pitches can be spelled in multiple ways, for instance G♯ and A♭. The spiral array takes pitch spelling into account...
by assigning a different geometric position to enharmonically equivalent (but differently-spelled) pitches. Because the pitch class representations are a helical arrangement of the line of fifths, pitches with sharps are located above the D and those with flats are located below D.

2.2 Cloud diameter

Tension in musical pieces is a property that varies over time. Therefore, a sliding window approach was used, whereby a musical piece is divided into equal length windows. Within each of those windows, all of the notes can be represented as a cloud in the spiral array.

The idea of cloud diameter is to capture the largest distance between any two notes in a cloud. When a chord or a cloud of notes contains intervals that are tonally far apart (i.e. dissonant), the distance between these pitches in the spiral array will be large. A first method tries to capture this type of harmonic tension by looking at the largest Euclidean distance within the cloud, or the cloud diameter. To illustrate this, Figure 2 shows the cloud diameter of the C major triad and its diminished counterpart. The larger diameter of the diminished triad can be explained by the large tonal distance between C and G♭, which is a diminished fifth. This is illustrated in the spiral array in Figure 3.

\[
\text{ce} = \frac{\sum_{i=1}^{N} d_i}{D} \times p_i \text{ whereby } D = \sum_{i=1}^{N} d_i. \tag{2}
\]

The idea of cloud momentum is to capture how large the distance between the centres of effect of two clouds of points is, thus capturing the movement in tonality. The ce’s of tonally similar chords or groups of notes are positioned close to each other in the spiral. When there is a change in tonality, this will cause the centres of effect to shift to a new area from one cloud to the next, thus resulting in a larger cloud momentum. The cloud momentum measures this type of tonal tension by calculating the Euclidean distance between the centres of effect of each window or cloud of notes. In the example in Figure 4, a large movement in tonality between the C chord and the C♯ chord can be seen. This is followed by no movement to the inverted C♯ chord.

Cloud momentum is a characteristic of movement. It can therefore be seen that its value for the first note of a fragment...
ment is non-existent (represented as zero). In the case that the window size is smaller than the duration of a note, it might occur that the cloud momentum drops during the span of that note, as it represents the “movement” in the spiral array and there is no movement within a note or cluster of notes.

2.4 Tensile strain

The previous methods capture the span of the cloud and the distance between adjacent centres of effect. The tensile strain captures the tonal distance between the $ce$ of a cloud of notes and the key. By implementing the key detection algorithm developed in [15], the Euclidean distance between $ce$ of the cloud of notes and the $ce$ of the global key (as described in [15]) is calculated. This distance represents the tensile strength. The short example in Figure 6 has a (given) global key of C major. The tensile strain is largest on the C$\#_7$ major chord, which is to be expected since it is tonally more distant than both C major or A minor from the given key. Figure 7 illustrates how the tensile strain was calculated, by marking the distance from the $ce$ of all three chords (in green) to the $ce$ of the key of C major (in orange) in the spiral array.

![Figure 6](image)

**Figure 6**: Tensile strain of C major – C$\#_7$ major – A minor chord given that the key is C major (min 1.0, max 1.5).

The distance to the $ce$ of the key in the spiral array is illustrated in Figure 6.

![Figure 7](image)

**Figure 7**: C major – C$\#_7$ major – A minor chord together with their distance from the $ce$ of the key (in orange).

In the next section, all three methods are studied in greater detail by means of examples.

3. EXAMPLES

In this section a number of example pieces are discussed. Short snippets from an empirical study by Farbood [1] are first analysed, followed by a few famous phrases and more extensive sections of music.

The methods above were implemented in java and are available online [1]. The system developed takes a musicXML file as input and outputs an INScore file [18] which represents the results of each of the methods as a coloured ribbon overlaid on the score. To represent these ribbons, the output of each method was normalised within a piece, or, in the case of the examples based on Farbood (Section 3.1.1), over all samples. In order to emphasise the changes in tension visibly, zero values (which occur when there are rests) are not taken into account when normalising.

3.1 Snippets

This section considers the application of the tension measures—cloud diameter, cloud momentum, and tensile strain—to the harmonically-motivated tension examples from Farbood, and compares the results to known user annotations of perceived tension for these samples.

3.1.1 Tension examples from Farbood

The above methods each capture some aspect of harmonic tension. In order to validate this claim, the methods were applied to examples from an empirical study. While each of the proposed methods captures part of the harmonic tension, one could argue that it is not possible to capture tension as an aggregate feature, as the attention of the listener constantly shifts between different features when evaluating tension [1].

Mary Farbood performed an extensive online questionnaire, whereby a total of 2,661 participants (17% of which self-categorised as musicians) annotated the perceived tension after listening to a snippet of music [1]. The participants were asked to select one out of six possible shapes for the perceived tension (see Figure 8).

![Figure 8](image)

**Figure 8**: Possible responses in Farbood’s study. Figure adapted from [1].

In this study, relevant stimuli pertaining to harmonic tension were selected and the results of the three methods were represented as tension ribbons overlaid on the score.
Figure 9: Computed tension parameters for selected stimuli from [1]. In sequence: cloud diameter (orange), cloud momentum (yellow), and tensile strain (red).
snippets in Figure 9. The colour-coded ribbons represent (in sequence): the cloud diameter (orange), the cloud momentum (yellow), and the tensile strain (red). The algorithm was run with one window per measure, except for stimulus A14, whereby a window size of two quarter notes was selected. In order to make the changes in the results of the methods clearly visible, the data for each type of ribbon was normalised over the results of all the stimuli.

<table>
<thead>
<tr>
<th>Stimulus ID</th>
<th>cloud diameter</th>
<th>cloud momentum</th>
<th>tensile strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>A03</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A04</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A05</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A06</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A07</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A08</td>
<td>(d)</td>
<td>(a)</td>
<td>(d)</td>
</tr>
<tr>
<td>A11</td>
<td>(a)(d)</td>
<td>(b)</td>
<td>(d)</td>
</tr>
<tr>
<td>A12</td>
<td>(d)</td>
<td>(b)</td>
<td>(a)(d)</td>
</tr>
<tr>
<td>A13</td>
<td>(a)(d)</td>
<td>(a)</td>
<td>(a)(b)</td>
</tr>
<tr>
<td>A14</td>
<td>(f)</td>
<td>(e)</td>
<td>(f)(g)</td>
</tr>
</tbody>
</table>

Table 1: Correspondence between computed tension patterns and templates in Figure 8.

<table>
<thead>
<tr>
<th>ID</th>
<th>% 1st</th>
<th>% 2nd</th>
<th>% 3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>A03</td>
<td>54% (b)</td>
<td>15% (a)</td>
<td>11% (d)</td>
</tr>
<tr>
<td>A04</td>
<td>41% (b)</td>
<td>16% (a)</td>
<td>13% (d)</td>
</tr>
<tr>
<td>A05</td>
<td>50% (d)</td>
<td>13% (g)</td>
<td>11% (f)</td>
</tr>
<tr>
<td>A06</td>
<td>46% (d)</td>
<td>19% (g)</td>
<td>14% (a)</td>
</tr>
<tr>
<td>A07</td>
<td>26% (g)</td>
<td>18% (e)</td>
<td>18% (f)</td>
</tr>
<tr>
<td>A08</td>
<td>25% (e)</td>
<td>20% (a)</td>
<td>16% (d)</td>
</tr>
<tr>
<td>A11</td>
<td>24% (a)</td>
<td>21% (g)</td>
<td>17% (d)</td>
</tr>
<tr>
<td>A12</td>
<td>27% (g)</td>
<td>25% (a)</td>
<td>18% (d)</td>
</tr>
<tr>
<td>A13</td>
<td>29% (a)</td>
<td>15% (d)</td>
<td>15% (i)</td>
</tr>
<tr>
<td>A14</td>
<td>78% (d)</td>
<td>7% (g)</td>
<td>5% (a)</td>
</tr>
</tbody>
</table>

Table 2: Top three responses for perceived tension for each stimulus (shown in Figure 8) as found by [1].

The results in Figure 9 are mapped to the patterns in Figure 8, as presented in Table 1, for direct comparison with the top three responses of Farbood’s empirical study, shown in Table 2. In the analysis below, a few results from less popular responses (not in the top three) are used. For a full overview of the empirical results, the reader is referred to [1]. (When examining the cloud momentum, we have to take into account the fact that the first window will not have a momentum, and hence will always be zero. The first value can therefore be ignored when analysing the results.) The comparison confirms that the above defined methods for tonal tension do capture different nuances of tension.

For the first two stimuli (A03 & A04), the cloud diameter and tensile strain have the same movement as the third most popular pattern, (d), identified by 11% and 13%, respectively, of the respondents. Another 15% and 16%, respectively, chose the same response as cloud momentum, (a). It is worth noting here that some of the features that may play a role in listeners’ evaluation of perceived tension may not be captured in the spiral array. Therefore it is extremely difficult to simplify the cause of tension perception to just one feature. For example, since the last chord is in essence the same as the first chord, with a doubling of the third at the octave, they are represented by the same cloud of points in the spiral array. An increasing line of tension in the first two fragments might have been captured by applying a method that takes into account the highest pitch or melodic contour [19], both features which are not covered in this paper. Since the spiral array uses pitch classes, our model will not capture tension which arises from chord inversions.

The three methods defined in this paper capture 58% (50% (d) + 8% (a)) and 60% (46% (d) + 14% (a)), respectively, of the responses for stimuli A05 and A06 [1]. For the next two stimuli (A07 & A08), 30% (17% (d) + 13% (a)) and 36% (20% (a) + 16% (d)), respectively, is captured [1]. This decrease can be due to the fact that many participants selected response (e) and (f), both of which require at least four points. More experiments with the window size might be useful in the future to test this influence, however, for now a window size per onset produces reasonable results.

The fact that tension is a characteristic which is perceived through different, alternative features is yet again confirmed in stimuli A11 and A12, where the top response (being (a)) is only selected by 24% and 27%, respectively. In the case of A11, 17% of the participants selected (d) and chose 14% (e) [1]. These two answers could even be seen as a reversed or opposite movement of tension. The calculated tensile strain varies over the fragments, as indicated by 21% and 27%, respectively, of the participants. The cloud diameter captures the perceived increase followed by a decrease very well. This tension profile corresponds to 17% and 18%, respectively, of the responses. These results yet again confirm the multidimensional aspect of perceived tension.

The cloud momentum is (perhaps) surprisingly small in stimuli A13. This can be attributed to the fact that the ce is calculated from many notes at the same time, thus diminishing the effect of changing only one note from one cloud to the next. When many notes are sounding at any given time in a slow changing sequence, cloud diameter and tensile strain might be more sensitive to tonal tension than cloud momentum.

For stimulus A14, a window size of two notes was selected. Due to the chromatic nature of this stimulus, the distance to the key is, as to be expected, very large. Most participants indicate an increase in tension followed by a decrease. This is also apparent when looking at the cloud momentum and cloud diameter. Although these ribbons also register tension in the beginning and ending.

From this analysis, it is apparent that the three measures—cloud diameter, momentum, and tensile strain—can help us understand different aspects of perceived tension in music.
3.2 Phrases

Next, we consider slightly longer examples in the form of phrases excerpted from Wagner’s Tristan and Isolde and Beethoven’s Sonata Op. 31 No. 3.

3.2.1 Tristan chord

A famous tension-inducing chord, which typically has an unusual relationship to the implied key of its surroundings, is the Tristan chord. Its unusual composition has been the topic of many musicological works [20]. [21] described it as “that Sphinx-chord, which has already occupied so many minds”. The Tristan chord consists of an augmented fourth, augmented sixth, and augmented ninth above a bass note, for example, \{F, B, D\#, G\#\}. The chord was given its name because the leitmotif associated with Tristan in Richard Wagner’s opera “Tristan und Isolde” contains this chord [22]. In the opera, Tristan and Isolde fall in immortal love after drinking a magic potion when they try to commit suicide together. Wagner uses the Tristan chord every time the potion or its effects are mentioned, thus connecting it to the build-up of suspense in the story. When the Tristan chord is represented in the spiral array (see Figure 10) it becomes clear that it is a very dispersed chord in tonal space.

An example of the Tristan chord is displayed in the first beat of bar 3 in the excerpt shown in Figure 11. This figure also displays the results when applying the three methods (with 6 windows per bar) to the fragment. A large increase in cloud momentum and cloud diameter are seen when the chord appears, indicating its tonally disperse nature and large tonal distance to the previous chord/notes. The tensile strain is more difficult to evaluate, as the key of this short phrase is not entirely clear. The example shows the tensile strain with A minor.

3.2.2 Beethoven Sonata Op. 31 No. 3 in E\# major

The Tristan chord has appeared in other compositions before Wagner wrote the opera. In Beethoven’s Sonata No. 18, Op. 31 No. 3, in E\# major (see Figure 12) the Tristan chord appears with the exact intervals albeit with a different spelling in the fourth bar. When enharmonically rewriting the Tristan chord in this way, it becomes less disperse in tonal space, as can be seen in Figure 13. This is reflected by the spiral array, which takes pitch spelling into account. For example, The G\# in the Tristan chord is much farther away from F than the A used by Beethoven (see Figures 10 and 12). We assume that the composers choices reflect the pitch relations they intend for the listener to perceive. The tension values calculated reflect these choices.

Figure 12 shows the cloud diameter and momentum, and tensile strain behaviour for the Beethoven Op. 31 No. 3 example. Recall that there was a global peak in the tensile strain and a local peak in the cloud momentum values for Wagner’s Tristan phrase. The peak tension as evaluated by the tensile strain has now shifted to the chord preceding the Tristan chord. For cloud diameter, although the Tristan chord produces a high value, the value remains high for the following chord. Thus this different spelling may lead to different interpretations of the tension, even though the notes are enharmonically equivalent.

3.3 Sections

Here, we turn to tension in larger sections of music: namely, Adagio (in A-flat major), the second movement of Schubert’s Piano Sonata in C minor D958 (beginning and end), and the first sixteen bars of Beethoven’s Les Adieux (Sonata No. 26, opus 81a, in E\# major).

To obtain the results, the methods used eight windows per bar and the results were normalized within each piece. During normalisation, rests were ignored, so as to make the difference between higher tensions more readily visible.

3.3.1 Schubert

Figure 14 and 15 display the harmonic tension ribbons for the first 17.25 bars and the last 14 bars of Schubert’s Pi-
ano Sonata in C minor D958, second movement: Adagio in $A\flat$ major. The last bars contain basically the same material, but are harmonised differently to sound even more tense. It is noticeable that sections with increased tension are mostly captured by one of the three suggested methods, yet not always by all three methods at the same time.

For example, the cloud diameter does a good job of capturing the heightened tension in the middle of the first bar of the ending segment, but less so the cloud momentum or the tensile strain because there is very little movement. Alternatively, in the second and third systems of the ending segment, the tensile strain has elevated values due to the tension of the pull away from the $A\flat$ major key to the more distant regions of $D$ major/minor and $A$ major. For the $D$ major/minor zone, the cloud diameter is low because the pitches are fairly self-consistent even though the keys vacillate between $D$ major and minor, the tonic remains the same and the cloud momentum is only higher at the beginning. The cloud diameter is also high in the $A$ major region, but not the cloud momentum.

3.3.2 Beethoven

Figures 16 and 17 show the harmonic tension ribbons for the first 16 bars of Beethoven’s Les Adieux (Sonata 26, opus 81a, in $E\flat$ major). This is an example that is fraught with tension, as reflected in the high tensile strain throughout. The tensile strain is only not high at regions with the $E\flat$ major chord (as in the second eighth note of bar 6) or $E\flat$ major seventh chord (towards the end of the excerpt). The cloud diameter varies, with the highest values at the right hand’s melodic turn in bars 3, 9, and 10, where the accompanying harmonies are also moving chromatically. The cloud momentum turns out to be less informative for this example as the highest values are associated with melodic leaps when there are very few notes, as in the end of bar 4.

4. CONCLUSIONS

Tension is a complex concept which is not easy to define or quantify. In this paper we have developed and implemented three methods for measuring aspects of tonal tension. These methods are based on the spiral array, a model for tonality. The implemented system is able to display tension ribbons over the input musical scores, a model for tonality. The implemented system is able to display tension ribbons over the input musical scores, thus allowing for easy interpretation. An analysis of existing pieces and a comparison with an empirical study [1] revealed that cloud diameter, cloud movement and tensile strain all contribute to capturing the composite feature humans refer to as tension.

This work only attempted to model aspects of tonal tension. The proposed measures relate to perceived distance between notes in a cluster, between consecutive clusters of notes, and between the global and local tonal contexts. They fail, however, to consider tension that is caused by other kinds of expectation, such as that due to delay of cadential closure—modeling this kind of tension is trickier because not all dominant-tonic pairs form cadences. Much work remains to understand the many different parameters that contribute to the perception of musical tension.

Future research includes conduct a more thorough empirical study of how the quantitative measures produced by the methods discussed in this paper correlate with what listeners describe as tension. The current model could be expanded to more completely capture the composite characteristics of tension. Further extensions could take into account features related to melodic contour, rhythm and timbre. Beyond score features, another interesting expansion would be to capture the influence of performance (e.g. timing and dynamics) variations on tension.

Acknowledgments

This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 658914.
Figure 14: Comparison of the beginning and ending of Schubert’s Piano Sonata in C minor D958, second movement: Adagio (in A-flat major).
Figure 15: Comparison of the beginning and ending of Schubert’s Piano Sonata in C minor D958, second movement: Adagio (in A♭ major).

Figure 16: Analysis of Sonata 7826, op 81a (first 16 bars).
Figure 17: Analysis of Sonata 26, op 81a (first 16 bars).

5. REFERENCES


