Characterizing and Interpreting Music Expressivity through Rhythm and Loudness Simplices

Paul Lascabettes[1], **Elaine Chew**[2], **Isabelle Bloch**[3]
[1] Sorbonne Université, Ircam, CNRS, Ministère de la Culture, STMS, Paris, France
paul.lascabettes@ircam.fr
[2] King's College London, London, United Kingdom
elaine.chew@kcl.ac.uk
[3] Sorbonne Université, CNRS, LIP6, Paris, France
isabelle.bloch@sorbonne-universite.fr

ABSTRACT

Characterizing and interpreting expressivity in performed music remains an open problem. In this paper, we explore the novel representation of recorded performances of triple time music using a 2-simplex, a graphical representation used to visualize three-interval rhythms. We analyze the MazurkaBL dataset, which contains beat-level tempo and loudness data of over 2000 recorded performances of 46 Chopin Mazurkas. Mazurkas' triple time lends themselves well to the 2-simplex; the expressive features of each three-beat bar map directly to unique points in the 2-simplex. We extend the rhythm simplex designed for beat durations to the representation of loudness. Each recorded performance is thus reduced to a set of points in 2-simplices based on beat-level duration or loudness. We provide the transformation to convert three-interval information to points in the 2-simplex; prove that inter-beat intervals and tempo representations in the 2-simplex are equivalent when timing variations are small; and, explain how smoothing the data impacts the coordinates of the points in the simplex. We demonstrate that the use of simplices can facilitate the analysis and interpretation of expressive music features; the method enables the identifying of bars with notable expressive variations such as temporal suspensions that form tipping points, and characterizing of performance regularity.

1. INTRODUCTION

Characterizing and interpreting expressivity in performed music is a fundamental problem in fields such as musicology, music perception, and music analysis. Moreover, the emergence of computational models has produced significant advances in expressive music research [1]. Computational modeling of expressive music performance requires large-scale databases; to create such databases, it is more expedient to focus on piano music. We focus on the piano pieces of the romantic period, which allow for greater expressive variations. A significant amount of this repertoire has been written for solo instruments, which makes comparative analysis of performances more straightforward. Furthermore, with the existence of computer-controlled pianos such as the Bösendorfer and Steinway Spirio, realistic piano performances can be readily captured with accuracy, which is not the case with other instruments. Therefore, the majority of performance research focuses on piano music.

Due to the resources amassed for and made available by the Mazurka Project [1], numerous studies have been based on Chopin's Mazurkas. A main purpose of the initial studies was to identify correlations between performed tempo and between performed loudness features [2, 3, 4, 5]. These analyses were based on selected Mazurka recordings. More recently, beat level tempo and loudness information of 2000 recorded Mazurkas have been extracted and made available for research [6]. Other studies have focused on generating Mazurka performances using machine learning models [7]. The two main expressive parameters considered are tempo and loudness variations [8]. Using these two parameters, trajectories in tempo-loudness space [9] were traced to represent the performance. Another representation of tempo and loudness curves for analysis is the arc model [10, 11]. Arc models have been successfully used to determine the segmentation of a piece based only on the loudness and tempo curves versus time [12, 13, 14].

In this paper, we propose to represent beat-level tempo and loudness information of a Mazurka performance in tempo and loudness simplices, a graphical representation used in music to analyze the perception of three-interval rhythms. This is the first time these simplices are applied to visualize Mazurka performances. This representation makes it straightforward to determine performance characteristics such as the accented beat in a bar or the regularity of a performance.

This paper is organized as follows: Section 2 introduces the rhythm simplex and reviews related work; Section 3 presents the dataset we used and explains why the 2-simplex representation is relevant to this dataset; Section 4 describes the proposed method, by first providing a transformation to convert a three-interval information into the 2-simplex, then proving that the choices of inter-beat intervals or tempo data are equivalent when timing variation are small, and explaining the impact of smoothing the data in the 2-simplex representation; and, Section 5 presents

[1] http://www.mazurka.org.uk/

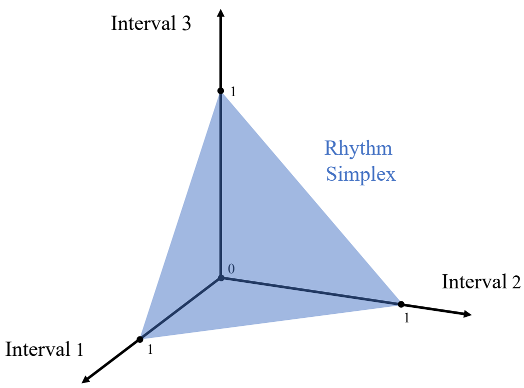
analyses and interpretations that can be made with the simplices, such as visualizing time suspensions, characterizing the notion of regularity of a performance, and identifying the bars with notable expressive variations.

2. THE RHYTHM SIMPLEX

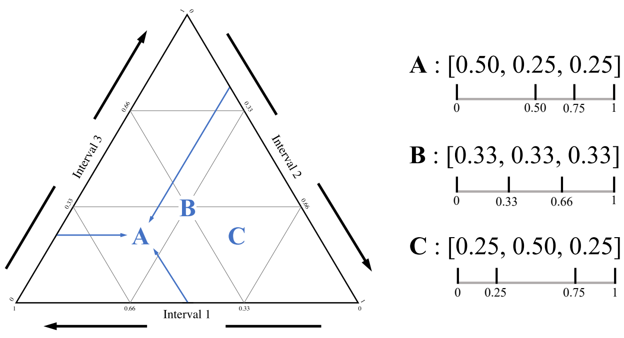
In this section, we introduce the rhythm simplex, a graphical representation used to visualize three-interval rhythms, and related work.

2.1 Presentation of the Rhythm Simplex

The *rhythm simplex*, also known as a *rhythm chart* or *chronotopological map*, is a graphical representation developed for visualizing three-interval rhythms. The idea of the rhythm simplex is that any three-interval rhythm can be represented in a 2-dimensional plot if the total duration of the rhythm is fixed (usually normalized to 1). For example, $A = [0.50, 0.25, 0.25]$ and $B = [0.33, 0.33, 0.33]$. By fixing the total duration, it is sufficient to know the first two intervals to deduce the third one. So a 2-dimensional plot can be used to visualize the rhythm. Thus, a three-interval rhythm is mapped to a unique point in the rhythm simplex. As illustrated in Figure 1(a), this 2-dimensional plot is a 2-simplex, i.e. a triangle. Figure 1(b) represents the rhythm simplex where each side of the triangle corresponds to one of the three intervals. In this figure, each three-interval rhythm is represented by a unique point. The rhythm A is located by the light blue arrows, while the rhythm B is in the middle of the rhythm simplex.



(a) The space of three-interval rhythms in the 2-simplex.



(b) Rhythm simplex where the three-interval rhythms A, B, and C are represented by unique points in the 2-simplex.

Figure 1. Three-interval rhythms in the rhythm simplex (adapted from Desain & Honing [15]).

2.2 Previous Work Related to the Rhythm Simplex

The use of the simplex representation in music was introduced by Desain and Honing [15, 16] to understand how listeners perceive rhythm categories. They asked the listeners to identify three-interval rhythms on a continuous scale to determine areas in the rhythm simplex representing equivalence classes. These rhythm equivalence classes can evolve according to parameters such as tempo [17], loudness or melodic structure [18]. Vaquero and Honing [19] created paths within these areas to generate performances. Based on a type of formal grammar called Lindenmayer systems [20], they defined paths in the rhythm simplex to generated expressivity in music. Bååth et al. [21] implemented a dynamical systems model to reproduce the categorical choices of listeners to retrieve these areas in the rhythm simplex. More recently, Jacoby and McDermott [22] ran a study where random rhythms in the rhythm simplex had to be reproduced by participants. Participants converged to rhythms having integer ratios after five iterations. They showed that the areas defined in [15] have little dependence on musical training but are highly dependent on cultural biases. For instance, listeners in the United States had different results than native Amazonian listeners. Nave et al. [23] conducted a similar experiment to the one in [22] with iterative reproductive tasks based on random rhythms in the rhythm simplex, but with children. They demonstrated the existence of rhythm priors in children, also related to cultural biases, which would develop in middle childhood.

Finally, these different studies used the rhythm simplex to understand the perception of three-interval rhythms, whether in populations of different cultures or of different ages. Only Vaquero and Honing have applied it to musical performances, but for generation rather than analysis. Here, we propose a new approach to use the rhythm simplex to represent musical performances in order to characterize and interpret music expressivity.

3. MAZURKA PERFORMANCES

In this section, we present the dataset that we use in this paper and explain how the rhythm simplex can be used to represent the essential expressive features of Mazurka performances.

3.1 The Dataset

To map a large number of Mazurka performances to the rhythm simplex for analysis, we used the MazurkaBL dataset [6]. This is currently by far the largest database of annotated performed classical music having multiple performances of each piece. The MazurkaBL contains beat level duration and loudness annotations for over 2000 recorded performances of 46 Chopin Mazurkas. There are, on average, more than 40 distinct performances per Mazurka. This dataset was made by manually annotating the beats of one recording of a Mazurka and automatically transferring these annotations to other recordings of the same piece through audio alignment. It is important to note that we used a smoothed version of the loudness data. Kosta et al. [6] filtered the loudness annotations by local regression using a weighted linear least squares and a 2nd degree polynomial model (the LOESS method of MATLAB's smooth function) with window sizes that are 1/30-th of the length of the recorded Mazurka while the tempo annotations are raw. It is therefore essential to take this into account, as it has an impact on the position of points in the 2-simplex and therefore on their interpretation, as explained in the following sections.

3.2 Representing Mazurka Performances in 2-Simplices

As described in Section 2.1, the rhythm simplex represents a three-interval rhythm as a unique point in a 2-dimensional plot. The use of the rhythm 2-simplex must therefore be applied to rhythm data that can be split into groups of three intervals, which is not the case for any general temporal data. However, Chopin's Mazurkas is particularly well suited to rhythm simplex representation because Mazurkas are folk dances mostly in triple meter, i.e., they are mostly pieces with three beats in each bar. On very rare occasions, some of Chopin's Mazurkas may contain bars that are not in triple meter, in which case these bars are removed from the analysis. Otherwise, the three beats can be seen as three intervals which have equal duration in the score but whose time is rendered differently in performance. Therefore, a bar of a performed Mazurka can be mapped into the rhythm simplex as a single point. By viewing the recording of a Mazurka as a sequence of performed three-beat bars, we can represent the recording by a set of points in the rhythm simplex. The rhythm simplex can show not only the timing variations in each bar, it can also display the beat accentuation because the loudness data of each beat of a three-beat bar can also be represented in a 2-simplex. For Mazurkas beginning with an anacrusis, we ignore the notes that come before the first full bar of music. Therefore, the essential expressive features of the musical performance of the Mazurkas can be visualized using 2-simplices because most of the bars have three beats.

4. THE METHOD TO REPRESENT MAZURKA PERFORMANCES IN 2-SIMPLICES

In this section, we give the transformation that we use to convert three-interval information into points in the 2-simplex. For the information related to rhythm, we prove that using inter-beat intervals or using tempo are equivalent in the 2-simplex representation. For the information related to loudness, we introduce the loudness simplex and explain how smoothing the data impacts the coordinates in the simplex.

4.1 Computing Points in the 2-Simplex

Previous articles on the rhythm simplex (see Section 2.2) have not explicitly provided the transformation to map a three-interval rhythm to a point in the rhythm simplex. Here, we provide the transformation we use to convert three-interval information to a point in 2-simplex.

First, we fix the vertices of the 2-simplex on the unit circle, i.e. the vertices are $(0, 1)$, $(\frac{\sqrt{3}}{2}, -\frac{1}{2})$, and $(-\frac{\sqrt{3}}{2}, -\frac{1}{2})$ (which are represented in Figure 1(b) by the top vertex, the bottom right vertex, and the bottom left vertex, respectively). In this case, the middle of the triangle is at the origin $(0, 0)$ (represented by point B in Figure 1(b)).

Let $B = (b_1, b_2, b_3)$ represent some properties of a three-beat bar such that $b_1 + b_2 + b_3 = 1$, where b_1, b_2, and b_3 are three positive numbers. The corresponding point (x, y) in the 2-simplex is defined by:

$$x = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2}b_3 - \sqrt{3}b_1, \quad (1)$$

$$y = \frac{3}{2}b_3 - \frac{1}{2}. \quad (2)$$

Reciprocally, given a point (x, y) in the 2-simplex, the corresponding bar $B = (b_1, b_2, b_3)$ is defined by:

$$b_1 = \frac{1}{3} - \frac{1}{\sqrt{3}}x - \frac{1}{3}y, \quad (3)$$

$$b_2 = \frac{1}{3} + \frac{1}{\sqrt{3}}x - \frac{1}{3}y, \quad (4)$$

$$\text{and} \quad b_3 = \frac{1}{3} + \frac{2}{3}y. \quad (5)$$

This defines a bijection between the 2-simplex and a normalized feature of the three-beat bars.

4.2 Inter-beat intervals and Tempo Data in the 2-Simplex

To represent a recorded Mazurka performance (or any performed piece in triple time) in the rhythm simplex, we compute inter-beat intervals from the beat onset information in the MazurkaBL. This gives us the duration of each beat, (d_1, d_2, d_3), in a bar. We scale this vector so that the elements sum to one to get (b_1, b_2, b_3), where $b_i = \frac{d_i}{d_1+d_2+d_3}$ for $i = 1, 2, 3$. The normalized durations in that bar then map to a point in the rhythm simplex according to Equations 1 and 2. In this way, we map the performed beat durations of all bars to the rhythm simplex. For example, in Figure 2, the different interpretations of the first four bars of Mazurka 6-1, obtained from the MazurkaBL dataset, are represented in the rhythm simplex. In this case, the first note is not considered, as Mazurka 6-1 begins with an anacrusis. Each point in a rhythm simplex corresponds to an interpretation of a bar, where the temporal deformation of the three beats (d_1, d_2, d_3) can be visualized. We can notice that all the performers follow a similar path. For instance, for the first bar, all points tend to be on the left in the simplex (meaning a long first beat and a short second beat). Whereas on the second and fourth bars, the points are on the right in the simplex (meaning a short first beat and a long second beat). Finally, for the third bar, interpretations are on average more regular as the points are concentrated in the middle of the simplex.

Previous studies on Mazurkas use tempo data rather than duration data. To generate comparable results, we can also compute the respective beat-to-beat tempo of a bar (t_1, t_2, t_3) with $t_i = 60/d_i$ for $i = 1, 2, 3$, and normalize the values for mapping to a 2-simplex. However, which data should we choose: inter-beat intervals or tempo? We show with the following proposition that the points in the 2-simplex based on inter-beat intervals or tempo data have approximately the same coordinates up to one symmetry

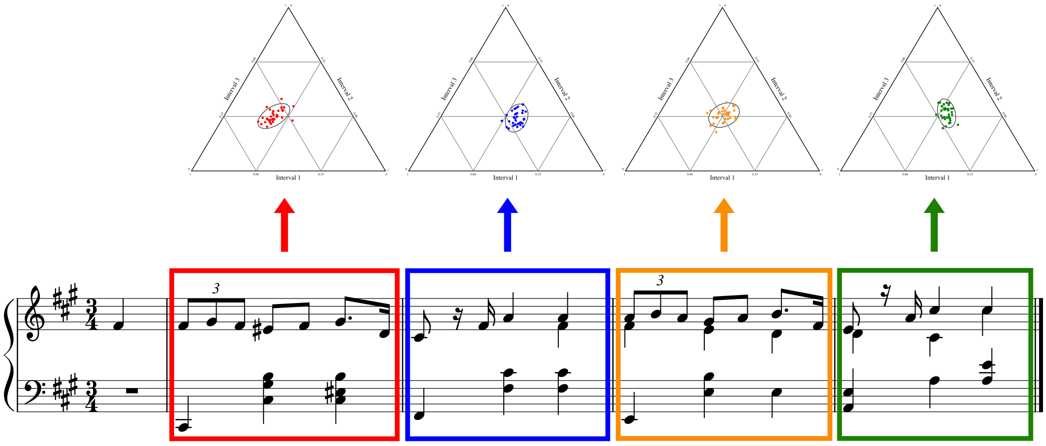


Figure 2. Visualization of different interpretations of the first four bars of Mazurka 6-1 in the rhythm simplex. Each point in a rhythm simplex corresponds to the interpretation of a bar by a performer, where the temporal deformation of the three beats of the bar is visualized.

with respect to the origin. This proposition shows that the choice of tempo or duration does not matter in the 2-simplex representation.

Proposition. Let (d_1, d_2, d_3) denote three inter-beat intervals of a bar and (x, y) the corresponding point in the 2-simplex. When (x, y) is close to the origin, the 2-simplex mapping of the beat-to-beat tempo of the same bar (t_1, t_2, t_3) is approximately $(-x, -y)$.

Proof. Let $d = d_1 + d_2 + d_3$ and $b_i = \frac{d_i}{d}$ for $i = 1, 2, 3$. Now, normalize the vector (t_1, t_2, t_3) and find its coordinates (x_T, y_T) in the 2-simplex.

$$\begin{aligned}
&\frac{1}{t_1 + t_2 + t_3}(t_1, t_2, t_3) \\
&= \frac{1}{\frac{60}{d_1} + \frac{60}{d_2} + \frac{60}{d_3}}\left(\frac{60}{d_1}, \frac{60}{d_2}, \frac{60}{d_3}\right) \\
&= \frac{1}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}}\left(\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3}\right) \\
&= \frac{d}{d}\frac{1}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}}\left(\frac{1}{d_1}, \frac{1}{d_2}, \frac{1}{d_3}\right) \\
&= \frac{1}{\frac{d}{d_1} + \frac{d}{d_2} + \frac{d}{d_3}}\left(\frac{d}{d_1}, \frac{d}{d_2}, \frac{d}{d_3}\right) \\
&= \frac{1}{\frac{1}{b_1} + \frac{1}{b_2} + \frac{1}{b_3}}\left(\frac{1}{b_1}, \frac{1}{b_2}, \frac{1}{b_3}\right) \\
&= \frac{1}{b_2 b_3 + b_1 b_3 + b_1 b_2}(b_2 b_3, b_1 b_3, b_1 b_2)
\end{aligned}$$

with (b_1, b_2, b_3) expressed with x and y as shown in Equations 3, 4 and 5. The vector (d_1, d_2, d_3) represents the duration of each beat of a bar. In a performance without large timing variations, the b_i's are close to $\frac{1}{3}$ and (x, y) is close to $(0, 0)$ in the 2-simplex (i.e. close to point B shown in Figure 1(b)). When x and y are close to zero, x^2, y^2, and xy are negligible. In this case, we get:

$$b_1 b_2 \approx \frac{1}{9}(-2y + 1), \qquad b_1 b_3 \approx \frac{1}{9}(y - \sqrt{3}x + 1),$$

$$b_2 b_3 \approx \frac{1}{9}(y + \sqrt{3}x + 1), \qquad b_2 b_3 + b_1 b_3 + b_1 b_2 \approx \frac{1}{3}.$$

Applying the transformation in Equations 1 and 2 to the vector (t_1, t_2, t_3) normalized with the approximations (i.e. x^2, y^2, and xy are negligible), we have:

$$x_T \approx \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2}\left(\frac{1}{3}(-2y+1)\right) - \sqrt{3}\left(\frac{1}{3}(y+\sqrt{3}x+1)\right) = -x,$$

and

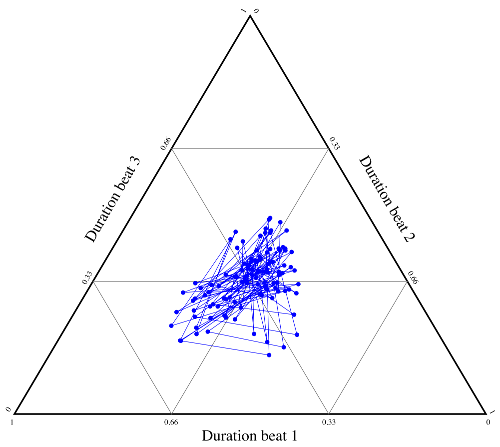
$$y_T \approx \frac{3}{2}\left(\frac{1}{3}(-2y + 1)\right) - \frac{1}{2} = -y.$$

To illustrate this result, Mazurka 6-1 performed by Luisada is represented in a 2-simplex in Figure 3(a) based on inter-beat intervals, and in Figure 3(b) based on beat-to-beat tempo. We can see that the point sets are similar in both figures up to one symmetry with respect to the origin.

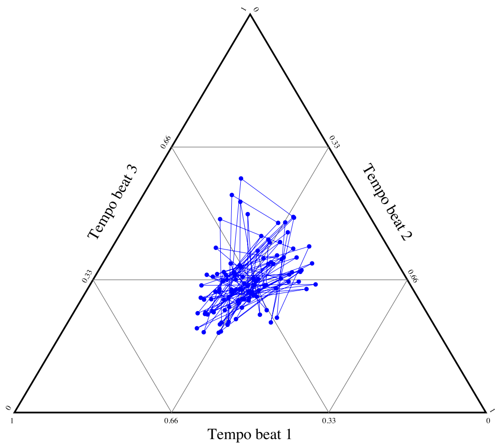
4.3 The Loudness Data in the 2-Simplex

Because loudness data are also available in the MazurkaBL dataset, we extend the rhythm simplex to the representation of loudness. We define the *loudness simplex* by considering the relative proportion of loudness for each beat in a three-beat bar. This is exactly the same idea as for rhythm; as we shall see later, it provides additional insights.

Because the loudness data have been smoothed with a LOESS filter by Kosta et al. [6], the loudness curves as a function of time tend to be smoother, involving fewer local maxima or minima. This has a direct consequence on the coordinates of the points in the loudness simplex. Figure 4 shows the triangle of the 2-simplex split into four regions: two areas, in blue, located at the top right and bottom left of the simplex; and, two areas, in red, located at the top left and bottom right. The two blue areas correspond to points in the simplex where the loudness increases or decreases with time within a bar. At the top right of the simplex, the elements are increasing, while at the bottom left they are decreasing. Musically, these areas translate to bars where the performer has an ascending (top right) or descending (bottom left) movement within that expressive property. Conversely, the points in the simplex are located in the red zone if the second beat is a local minimum or maximum, which is more seldom seen with loudness data because the data have been smoothed. Thus, as can be seen in Figures 6(a) and 6(b), when representing the loudness



(a) Representation of Mazurka 6-1 performed by Luisada in the 2-simplex based on inter-beat intervals.



(b) Representation of Mazurka 6-1 performed by Luisada in the 2-simplex based on beat-to-beat tempo.

Figure 3. Inter-beat intervals and tempo data represented in a 2-simplex to illustrate the symmetry with respect to the origin between the two sets of points.

data in the 2-simplex, the points tend to stay in the blue area because the data have been smoothed. However, this does not exclude performance information as we will see in the rest of this paper. Note that this reasoning applies to any type of smoothed data projected into a 2-simplex.

5. ANALYSIS AND INTERPRETATION OF MUSICAL EXPRESSIVITY USING THE SIMPLICES REPRESENTATION

In this section, we present analyses and interpretations that can be derived from the simplices representation. This includes visualizing tipping points (here realized as time suspensions) introduced by the performer, defining the regularity of a performance, and identifying bars that exhibit notable expressivity.

5.1 Visualize Time Suspensions in the Simplex

Some recorded Mazurkas display musical tipping points [24] which present as significant temporal suspensions. These time elongations may be indicated in the score as fermatas or may simply be inserted by the performer. Representing these Mazurka performances in the rhythm simplex allows these temporal deformations to be visible. For example, in Mazurka 24-3, bars 10, 22, 46, and 70 are identical and have a fermata marked and executed on the second beat. Thus, the duration of the second beat of these bars is significantly longer than others, making them more important. Therefore, the points corresponding to these bars are situated at the bottom right of the rhythm simplex. For example, by representing Uninsky's interpretation of Mazurka 24-3 in the rhythm simplex, we can easily extract these bars as shown in Figure 5.

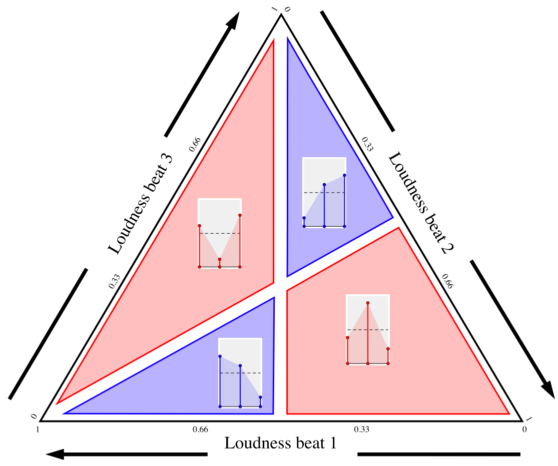


Figure 4. Area in the loudness simplex characterized by an increasing loudness curve (in blue) or local maxima or minima (in red) in the bar.

These temporal suspensions are present in most interpretations of Mazurka 24-3, so these four points in the bottom right of the simplex are seen in other interpretations as well. These tipping points [24] allow performers to pause to highlight certain notes or harmonies to create tension and anticipation [25]; in the case of Figure 5, it is the highest note of the bar on the second beat.

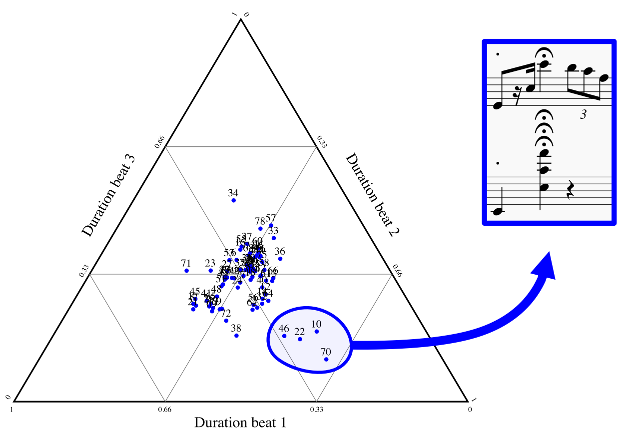
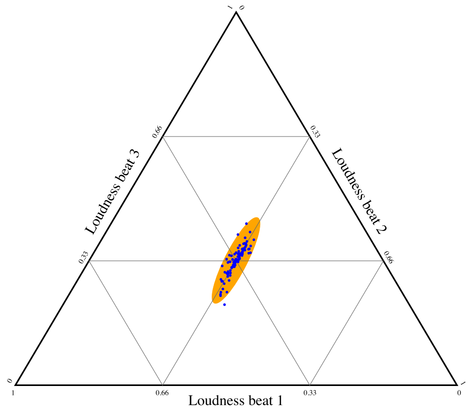


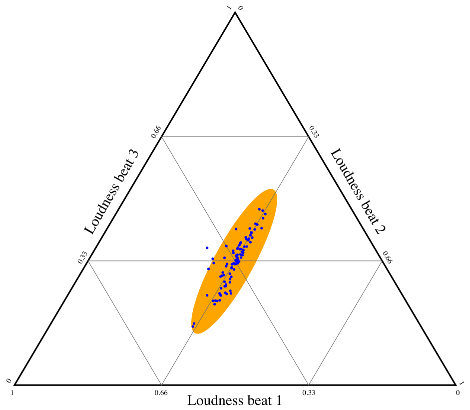
Figure 5. Temporal suspensions of the second beat of bar 10, 22, 46, and 70 of Mazurka 24-3 indicated as fermatas in the score (frame on the right) and visualized in Uninsky's interpretation by isolated points in the bottom right of the rhythm simplex.

5.2 Characterizing the Regularity of a Performance Using the Simplices Representation

During a music performance, some performers allow themselves greater latitude in their temporal and loudness variations than others. However, quantifying these variations to define the regularity of a performance is not easy [8]. If a performer chooses to always accentuate the first beat of each bar, the performance could be described as regular, even though the differences in tempo and volume per beat can be large during the piece. Therefore, it is not sufficient to characterize the regularity of a performance by the sum of the deviations per beat from the average of tempo and loudness. Hence, we propose to quantify the regularity of a performance using the 2-simplices. To get an intuition about this, see Figure 6(a) where the points are very dense in the loudness simplex meaning that the loudness variations are very regular. On the other hand, in Figure 6(b), the points are more spread out because the loudness varies less regularly.



(a) Small area of the ellipse in the loudness simplex from Kapell's interpretation of Mazurka 6-2 indicating an interpretation with regular loudness.



(b) Large area of the ellipse in the loudness simplex from Ohlsson's interpretation of Mazurka 6-2 indicating an interpretation with non-regular loudness.

Figure 6. Example of regular and non-regular performed loudness shown by a small and a large ellipse area in the loudness simplex. Kapell's ellipse area is three times smaller than Ohlsson's, pointing to greater loudness regularity.

We propose to define the regularity of a performance from the covariance matrix of the points in the 2-simplex as follows. Let $(x_i, y_i)_{1\leq i\leq n}$ denote the points in the simplex and Σ their covariance matrix:

$$\Sigma = \begin{pmatrix} \mathrm{Cov}(X,X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(Y,X) & \mathrm{Cov}(Y,Y) \end{pmatrix},$$

where $X = (x_i)_{1\leq i\leq n}$, $Y = (y_i)_{1\leq i\leq n}$,

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i, \\ \bar{y} &= \frac{1}{n}\sum_{i=1}^{n} y_i, \\ \text{and} \quad \mathrm{Cov}(X,Y) &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}). \end{aligned}$$

Let λ_1 and λ_2 denote the eigenvalues of Σ. They define an ellipse [2], whose semi-axes lengths are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$. The orientation is derived from Σ as well, but not involved in the regularity computation. We define the regularity value of a performance as the inverse of the area of that ellipse.

Since the area of an ellipse is equal to the multiplication of π by the length of the semi-major and the semi-minor axes, it is easy to compare the regularity of two performances in the 2-simplex. Using the example in Figure 5, Kappel (λ_1 and λ_2 as eigenvalues) is 2.995 times more regular than Ohlsson (λ_1' and λ_2' as eigenvalues) because $\sqrt{\lambda_1'\lambda_2'}/\sqrt{\lambda_1\lambda_2} = 2.995$. This reasoning can be applied across interpretations to other expressive features like beat-level tempo and duration.

5.3 The Most Distant Points in the Simplex Indicate Bars with Notable Musical Expressivity

The ellipses described in the previous section correspond to points having the same Mahalanobis distance to the center [26]. Distinct from Euclidean distance, the Mahalanobis distance takes into account the distribution of points in the simplex. Given two points (x_1, y_1) and (x_2, y_2) in the 2-simplex, the Mahalanobis distance is defined by:

$$d\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \sqrt{\begin{pmatrix} x_1-x_2 \\ y_1-y_2 \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} x_1-x_2 \\ y_1-y_2 \end{pmatrix}}.$$

This distance is useful in our case because there is often a correlation between the coordinates of the points in the simplex, in particular, for points in the loudness simplex, the x coordinate increases when the y coordinate increases (but the reasoning is also correct for the rhythm simplex). For example, in Figure 6, the points are aligned in the simplex because the loudness data have been smoothed, as detailed in Section 4.3. In this case, the Mahalanobis distance allows us to give more importance to the points that do not follow this alignment. Thus, the points with a high distance value compared to the average are the bars that present a strongly divergent musical expressivity according to the simplex feature. For example, we can identify the temporal elasticity in the interpretation of Mazurka 24-3 by Uninsky which is represented in the rhythm simplex in Figure 5.

[2] We used the matplotlib package `https://matplotlib.org/stable/gallery/statistics/confidence_ellipse.html`, with the default setting n_std=3.0, for the ellipses in Figure 6.

We computed the Mahalanobis distance between each simplex point and the mean and plotted the result in Figure 7. Bars that have an extreme time elongation on the second beat, i.e. bars 10, 22, 46, and 70 of Mazurka 24-3 represented by the four points at the bottom right of the rhythm simplex in Figure 5, are located almost at the end of each A section and have a high distance value due to their divergent interpretation as compared to other bars (Figure 7). Thus, by considering the maximum values of the Mahalanobis distance between the simplex points and their mean, we can detect the bars that have a notable expressive variation. For example, in Figure 7, bars 34 and 57 also have a high distance value because the first beat is considerably longer than the other two, which can be seen in the top corner of the rhythm simplex in Figure 5.

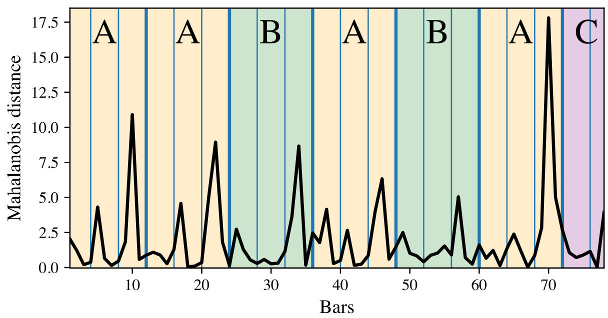


Figure 7. Mahalanobis distance between the points of Uninsky's interpretation of Mazurka 24-3 in the rhythm simplex shown in Figure 5. High distance values are bars with notable expressive variations. For example, bars 10, 22, 46, and 70 have high distance values because their second beats are elongated, while bars 34 and 57 have their first beats elongated.

6. CONCLUSION

In this paper, we proposed a novel method to represent performed music using beat duration, tempo and loudness simplices. We analyzed the MazurkaBL dataset using these simplices. MazurkaBL contains more than 2000 recorded performances of 46 of Chopin's Mazurkas with annotations of beat level tempo and loudness values. We provided the equations to map any bounded three-interval information into the 2-simplex. We proved that the choices of inter-beat intervals or tempo are nearly equivalent in the 2-simplex up to one symmetry with respect to the origin. We also explained the impact of smoothed data in the 2-simplex. Finally, we showed that the simplices facilitate the analysis and interpretation of music expressivity features. For example, by using the Mahalanobis distance, it is possible to identify bars with notable expressive variations such as temporal tipping points or to specify the regularity of a performance.

However, this method has some limitations. Firstly, the duration or loudness of each bar is normalized. As a result, information on the total duration or loudness of each bar is not available, i.e. only the proportional distribution of durations and loudness are shown. This means that two points with the same coordinates in the simplex might be two bars of completely different overall loudness or duration. For example, if the three beats of a bar are played equally fast or slow, the corresponding point is at the center of the rhythm simplex. Thus, this method is more appropriate for analyzing performance expressivity at the bar level. Nevertheless, we believe that variations on the scale of a musical phrase should also be considered. The bars corresponding to the beginning and end of the phrase are points distant from the center of the simplex, as they have increasing or decreasing tempo or loudness values. Whereas the bars in the middle of the phrase are bars with small variations in loudness or tempo, i.e. points in the center of the simplex. Another limitation of this method is that it can only be applied to three-interval data. While this approach is appropriate for the MazurkaBL dataset, which comprises of music pieces in triple time, it may not be suitable for more general types of data. However, since the MazurkaBL is the largest existing annotated dataset of performed music, we believe that developing tools like the 2-simplex, even if these methods are not applicable to all types of data, is crucial for analyzing and gaining a better understanding of music performances.

For future research, it would be valuable to study the movement of points within the simplex over time, in order to identify the trajectories of a section of music within the simplex and to capture the variation in interpretations amongst several performers. This would be useful to generate or modify performances by moving points in the 2-simplex and to identify the musical structure of a piece by knowing the trajectory patterns corresponding to the interpretation of sections. Finally, we expect that the development of methods based on formal mathematical representations, such as the 2-simplex in this paper, holds great promise for facilitating a more comprehensive analysis and understanding of musical performance.

Acknowledgments

P. Lascabettes is funded by a Contrat Doctoral Spécifique pour Normaliens (CDSN) scholarship. E. Chew is funded by a European Research Council (ERC) H2020 Advanced fellowship GA 788960 COSMOS. This work was partly supported by the chair of I. Bloch in Artificial Intelligence (Sorbonne Université and SCAI).

7. REFERENCES

[1] G. Widmer and W. Goebl, "Computational Models of Expressive Music Performance: The State of the Art," *Journal of New Music Research*, vol. 33, no. 3, pp. 203–216, 2004.

[2] N. Cook, "Performance Analysis and Chopin's Mazurkas," *Musicae Scientiae*, vol. 11, no. 2, pp. 183–207, 2007.

[3] C. S. Sapp, "Comparative Analysis of Multiple Musical Performances." in *International Society for Music Information Retrieval Conference (ISMIR)*, 2007, pp. 497–500.

[4] ——, "Computational Methods for the Analysis of Musical Structure," Ph.D. dissertation, Stanford University, 2011.

[5] N. Cook, "Between Art and Science: Music as Performance," *Journal of the British Academy*, vol. 2, pp. 1–25, 2014.

[6] K. Kosta, O. F. Bandtlow, and E. Chew, "MazurkaBL: Score-Aligned Loudness, Beat, Expressive Markings Data for 2000 Chopin Mazurka Recordings," in *Fourth International Conference on Technologies for Music Notation and Representation (TENOR)*, 2018, pp. 85–94.

[7] Z. Shi, "Computational Analysis and Modeling of Expressive Timing in Chopin's Mazurkas," in *22nd International Society for Music Information Retrieval Conference (ISMIR)*, 2021, pp. 650–656.

[8] C. E. Cancino-Chacón, M. Grachten, W. Goebl, and G. Widmer, "Computational Models of Expressive Music Performance: A Comprehensive and Critical Review," *Frontiers in Digital Humanities*, vol. 5, pp. 5–25, 2018.

[9] J. Langner and W. Goebl, "Visualizing Expressive Performance in Tempo-Loudness Space," *Computer Music Journal*, vol. 27, no. 4, pp. 69–83, 2003.

[10] N. P. M. Todd, "The Dynamics of Dynamics: A Model of Musical Expression," *Journal of the Acoustical Society of America*, vol. 91, no. 6, pp. 3540–3550, 1992.

[11] A. Gabrielsson, "Once Again: The Theme from Mozart's Piano Sonata in A Major," *Action and Perception in Rhythm and Music*, pp. 81–103, 1987.

[12] G. Widmer and A. Tobudic, "Playing Mozart by Analogy: Learning Multi-Level Timing and Dynamics Strategies," *Journal of New Music Research*, vol. 21, no. 1, pp. 259–268, 2003.

[13] D. Stowell and E. Chew, "Maximum a Posteriori Estimation of Piecewise Arcs in Tempo Time-Series," in *From Sounds to Music and Emotions*, M. Aramaki, M. Barthet, R. Kronland-Martinet, and S. Ystad, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 387–399.

[14] C. H. Chuan and E. Chew, "A Dynamic Programming Approach to the Extraction of Phrase Boundaries from Tempo Variations in Expressive Performances," in *8th International Society for Music Information Retrieval Conference (ISMIR)*, Vienna, Austria, 2007, pp. 305–308.

[15] P. Desain and H. Honing, "The Formation of Rhythmic Categories and Metric Priming," *Perception*, vol. 32, no. 3, pp. 341–365, 2003.

[16] H. Honing, "Structure and Interpretation of Rhythm and Timing," *Tijdschrift voor Muziektheorie*, vol. 7, no. 3, pp. 227–232, 2002.

[17] S. Handel, "The Effect of Tempo and Tone Duration on Rhythm Discrimination," *Perception & Psychophysics*, vol. 54, no. 3, 1993.

[18] H. G. Tekman, "Interactions of Perceived Intensity, Duration, and Pitch in Pure Tone Sequences," *Music Perception*, vol. 14, no. 3, pp. 281–294, 1997.

[19] C. V. Patricio and H. Honing, "Generating Expressive Timing by Combining Rhythmic Categories and Lindenmayer Systems," in *50th Anniversary Convention of the AISB*, London, UK, 2014, pp. 1–7.

[20] A. Lindenmayer, "Mathematical Models for Cellular Interactions in Development I. Filaments with One-Sided Inputs," *Journal of Theoretical Biology*, vol. 18, no. 3, pp. 280–299, 1968.

[21] R. Bååth, E. Lagerstedt, and P. Gärdenfors, "A Prototype-Based Resonance Model of Rhythm Categorization," *i-Perception*, vol. 5, no. 6, pp. 548–558, 2014.

[22] N. Jacoby and J. H. McDermott, "Integer Ratio Priors on Musical Rhythm Revealed Cross-Culturally by Iterated Reproduction," *Current Biology*, vol. 27, no. 3, pp. 359–370, 2017.

[23] K. Nave, C. Carrillo, N. Jacoby, L. Trainor, and E. Hannon, "The Development of Rhythmic Categories as Revealed Through an Iterative Production Task," *PsyArXiv*, 2021.

[24] E. Chew, "Playing with the Edge: Tipping Points and the Role of Tonality," *Music Perception*, vol. 33, no. 3, pp. 344–366, 2016.

[25] D. Herremans and E. Chew, "Tension Ribbons: Quantifying and Visualising Tonal Tension," in *Second International Conference on Technologies for Music Notation and Representation (TENOR)*, Cambridge, United Kingdom, 2016, p. 10.

[26] P. C. Mahalanobis, "On the Generalised Distance in Statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, p. 49–55, 1936.