1

Algorithmic, Geometric, and Combinatorial Problems in Computational Music Theory

Godfried T. Toussaint
godfried@cs.mcgill.ca
McGill University
School of Computer Science
3480 University St., Montreal, Canada

Published in: *Proceedings of X Encuentros de Geometría Computacional*, University of Sevilla, Sevilla, Spain, June 16-17, 2003, pp. 101-107.

Abstract

Computational music theory offers a wide variety of interesting geometric, combinatoric, and algorithmic problems. Some of these problems are illustrated for the special cases of rhythm and melody. In particular, several techniques useful for the teaching, analysis, generation and automated recognition of the rhythmic components of music are reviewed. A new measure of rhythm-evenness is described and shown to be better than previous measures for discriminating between rhythm timelines. It may also be more efficiently computed. Several open problems are discussed.

Key Words: rhythm analysis, clave and bell patterns, measuring rhythmic and melodic similarity, swap-distance, phylogenetic analysis, rhythmic oddity, rhythmic evenness, visualization, computational rhythm, music informatics.

1. Introduction

Mathematics and music theory have a long history of collaboration dating back to at least Pythagoras [9]. More recently the emphasis has been mainly on analysing mathematical problems that arise in music theory [18], [6], [19], [2],

[1]Research partially supported by NSERC

[4], and applying new tools from discrete mathematics to music analysis [3], [12], [7], [5]. Since the advent of computers, the field of artificial intelligence has also had much to say about computer music [13].

The artificial intelligence approach to the study of music has concentrated almost exclusively on the automated generation (composition) of music or the recognition of written scores or acoustic signals. The mathematics treatment of music has been limited almost entirely to its *vertical* aspects (pitch or scales), thus virtually ignoring the *horizontal* dimension of time and rhythm. One notable exception is the work of Simha Arom [2].

Toussaint [22], [23] initiates the theoretical investigation of rhythm with mathematical tools along several fronts by introducing a variety of geometric, graph-theoretical and combinatorial techniques useful for the visualization, teaching, analysis, generation, and automated recognition of rhythms. Combinatorial techniques based on permutations of multisets are used to generate new interesting rhythms from old. A new measure of rhythm complexity is compared to older measures. Several methods for measuring the similarity between two rhythms are compared. Tools from computational biology are applied to rhythm analysis. In particular, the study of the *evolution* of rhythms is initiated by applying phylogenetic analyses to distance matrices of groups of rhythms.

In this paper several fruitful directions for the algorithmic study of music in general, and rhythm in particular, are reviewed. A new measure of rhythmic evenness is described, and it is shown that this measure is more discriminating than previous measures when applied to rhythm timelines, and can be computed more efficiently. Finally, several open problems are discussed.

2. Melodic and Rhythmic Similarity

2.1. Music as Strings of Symbols.— One of the most popular rhythms is the *Clave Son* heard a lot in Son and Salsa music as well as much other music around the world. It is traditionally played with two wooden sticks. The *Clave Son* rhythm is usually notated for musicians using standard music notation which affords many ways of expressing a rhythm. Four examples are given in the top four lines of Figure 1. The fourth line shows the rhythm with music notation using the smallest convenient notes and rests. The bottom line shows a popular way of representing rhythms for percussionists that do not read music. It is called the *Box Notation Method* developed by Philip Harland at the University of California in Los Angeles in 1962 and is also known as TUBS (Time Unit Box System).

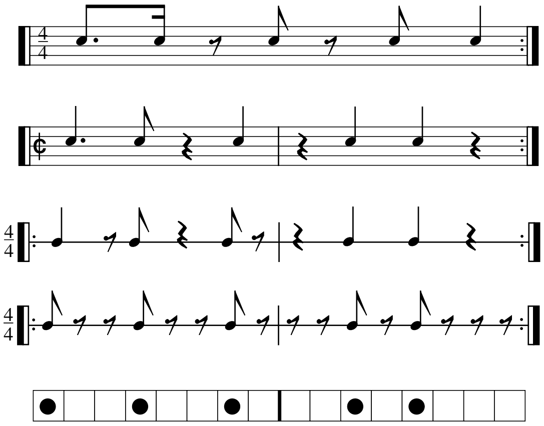

Figure 1: Five ways of representing the *clave Son* rhythm [22].

The box notation method is convenient for simple-to-notate rhythms like bell and clave patterns as well as for experiments in the psychology of rhythm perception, where a common variant of this method is simply to use one symbol for the note and another for the rest. Thus for the clave Son a common way to write it is simply as [x . . x . . x . . . x . x . . .]. In computer science the clave Son would be written as the 16-bit binary sequence: [1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0]. Here this rhythm becomes a point in a 16-dimensional space (the hypercube). A natural measure of the difference between two rhythms represented as binary sequences is the well known Hamming distance, which counts the number of positions in which the two rhythms disagree. Another useful representation of a rhythm is as an *interval-vector*. This representation consists of a list of the time intervals between successive notes in the cyclic rhythm. For example, for the clave Son the interval vector is: (3 3 4 2 4). Here this rhythm becomes a point in a 5-dimensional lattice space because the rhythm contains five intervals. Measures of the difference between two rhythms represented as interval-vectors include the Manhattan metric, the Euclidean distance or more generally any Minkowski metric.

2.2. The Edit Distance.— One of the most popular measures for comparing two arbitrary musical sequences represented as notes (strings of symbols) is the *edit-distance* [14] well known in text and string processing. The edit distance between two strings is the minimum number of insertions, deletions and substitutions (mutations) needed to convert one string into the other. This distance measure is well known and computed in practice using dynamic programming [17].

2.3. The Swap Distance.— For the special case of two rhythm timelines that contain the same number of of notes, Toussaint [23] proposed a distance measure termed the *swap* distance. A swap is an interchange of a one and a zero (note duration and rest interval) that are adjacent in the sequence. The swap distance between two rhythms is the minimum number of swaps required to convert one rhythm to the other. For example the rhythm [x . x . x x . x . x . x] can be converted to the rhythm [x . x x . x x . x . x .] with a minimum of four swaps, namely interchanging the third, fifth, sixth, and seventh notes with the corresponding adjacent rests preceeding them. It was shown in [23] that such a measure of dissimilarity appears to be more appropriate than the Hamming distance between the binary vectors or the Euclidean distance between the interval vectors.

Given two cyclic binary sequences, each of length n with k notes (one's), and each with it's starting note identified, it is straightforward to verify that the swap-distance may be computed in $O(n)$ time. Also of interest is to compute the minimum swap-distance over all possible rotations of the rhythms. The naive method of trying all possible rotations leads obviously to an $O(n^2)$ time algorithm. Tom Shermer [20] has shown that this can be improved to $O(n + k^2)$ time. It is an open problem whether his algorithm is optimal.

2.4. The Area-Difference Distance.— O'Maidín [15] proposed a geometric measure of the distance between two melodies modelled as monotonic pitch-duration rectilinear functions of time as depicted in Fig. 2. O'Maidín measures the distance between the two melodies by the area between the two polygonal chains (shown shaded in Fig. 2). Note that if the area under each melody contour is equal to *one*, the functions can be viewed as probability distributions, and in this case O'Maidín's measure is identical to the classical Kolmogorov *variational distance* used to measure the difference between two probability distributions [21]. If the number of vertices (vertical and horizontal segments) of the two polygonal chains is n then it is trivial to compute O'Maidín's distance in $O(n)$ time using a line-sweep algorithm [16].

In a more general setting such as music retrieval systems we are given a short query segment of music denoted by the polygonal chain $Q = (q_1, q_2, ..., q_m)$, and a longer stored segment $S = (s_1, s_2, ..., s_n)$, where $m < n$. Furthermore, the query segment may be presented in a different *key* (transposed in the vertical direction) and in a different *tempo* (scaled linearly in the horizontal direction). Note that the number of keys (horizontal levels) is a small finite constant. Time is also quantized into fixed intervals (such as eighth or sixteenth notes). In this context it is desired to compute the minimum area between the two contours under vertical translations and horizontal scaling of the query. Francu and Nevill-Manning [8] claim that this distance measure can be computed in $O(mn)$ time but they do not

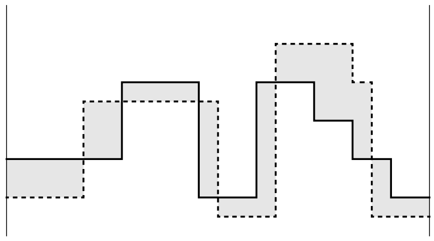

Figure 2: Two melodies as rectilinear pitch-duration functions of time.

describe their algorithm in detail.

The work of O'Maidín [15] and Francu and Nevill-Manning [8] immediately suggests several interesting open problems. In the acoustic signal domain the key of the melody loses significance and hence the vertical transposition is continuous rather than discrete. The same can be said for the time axis. What is the complexity of computing the minimum area between a query $Q = (q_1, q_2, ..., q_m)$ and a longer stored segment $S = (s_1, s_2, ..., s_n)$ under these more general conditions?

Another simpler variant of this problem concerns acoustic rhythmic melodies, i.e., cyclic rhythms with notes that have pitch as a continuous variable. Here we assume two rhythmic melodies of the same length are to be compared. Since the melodies are cyclic rhythms they can be represented on the surface of a cylinder. What is the complexity of computing the minimum area between the two rectilinear polygonal chains under rotations around the cylinder and translations along the length of the cylinder? Aloupis et al. [1] present an $O(n)$ time algorithm to compute this measure if rotations are not allowed, and an $O(n^2 \log n)$ time algorithm for unrestricted motions (rotations around the cylinder and translations along the length of the cylinder). Can this complexity be improved?

3. Vizualization of Rhythms

Another geometric representation of rhythms useful for mathematical analysis is as convex polygons [22]. If we map the sequence: [x . . x . . x . . . x . x . . .] (the clave Son) onto a circle, and connect all the notes adjacent to each other in clockwise order with line segments, we obtain the convex polygon depicted in Figure 3. A variety of geometric features of such polygons are immediately evident. For example, the triangle determined by the notes labelled 0, 3 and 6

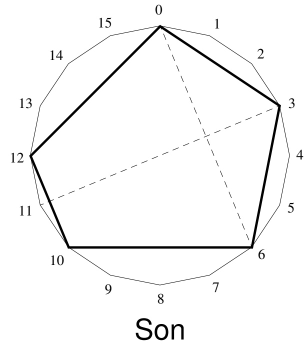

Figure 3: The *clave Son* rhythm as a convex polygon.

is an *isoceles* triangle. The line through locations 3 and 11 determines an axis of mirror symmetry for this rhythm. If the rhythm is started at location 3 then it sounds the same whether it is played forwards or backwards. No internal angle of the polygon is a right angle.

In [22] it was shown that the six fundamental 4/4 time *clave* and bell timelines (of which the Son is a member) can be classified with these simple geometric features without knowing the starting note of each rhythm. The six fundamental timelines are: the *Son* [x . . x . . x . . . x . x . . .], the *Shiko* [x . . . x . x . . . x . x . . .], the *Rumba* [x . . x . . . x . . x . x . . .], the *Bossa-Nova* [x . . x . . x . . . x . . x . .], the *Soukous* [x . . x . . x . . . x x . . . .], and the *Gahu* [x . . x . . x . . . x . . . x .].

It was also shown in [22] with the phylogenetic analysis techniques of Gaston Gonnet [10], using both the Hamming distance between binary vectors and the Euclidean distance between the interval-vectors, that the *clave Son* rhythm is most like all the others, thus offering an explanation for its worldwide popularity. A similar analysis of the ten principal 12/8 time ternary African bell patterns, using the *SplitsTree* phylogenetic algorithm of Daniel Huson [11] with the swap-distance was reported in [23].

4. Rhythmic Oddity

Simha Arom [2] defines a rhythm as having the *rhythmic oddity* property if no two onsets (notes) partition the entire cyclic interval into two subintervals (bipartition) of equal length. Toussaint [23] went further by defining a measure of

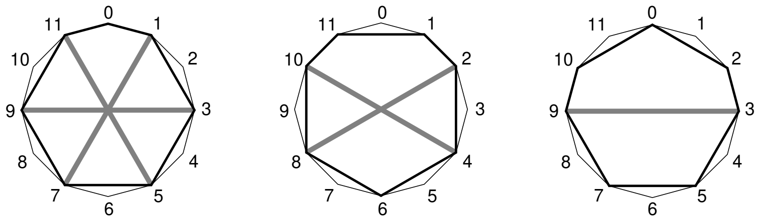

Figure 4: Three rhythms showing the number of equal bi-partitions in each.

the amount of rhythmic oddity of a rhythm in terms of the *number* of bi-partitions of equal length that the rhythm admits. The fewer bi-partitions a rhythm admits, the more rhythmic oddity it posesses. Figure 4 shows three seven-beat rhythms with the number of equal bi-partitions contained in each. From left to right the three rhythms admit three, two, and one equal bi-partitions, respectively.

For a rhythm with k notes represented as a binary sequence of length n it is trivial to either decide if a rhythm has the *rhythmic oddity* property, or to compute its rhythmic oddity, in $O(n)$ time by rotating a diameter and counting the diametral antipodal pairs of vertices.

A more difficult problem is that of counting and enumerating rhythms that have the *rhythmic oddity* property, as a function of n and k. Chemillier and Truchet [5] characterize such rhythms and describe an algorithm for enumerating them. On the other hand, it is an open problem to count and enumerate rhythms that admit p equal bi-partitions for any *fixed* integer p greater than zero.

5. Rhythmic Evenness

Consider the following three 12/8 time ternary rhythms: [x . x . x . x . x . x .], [x . x . x x . x . x . x] and [x . . . x x . . x x x .]. It is intuitively clear that the first rhythm is more *even* than the second, and the second is more even than the third. In passing we note that the second rhythm is internationally the most well known of all the African ternary timelines. It is traditionally played on an iron bell and is known mainly by its Cuban name *Bembé* [23].

In music theory much attention has been devoted to developing the notion of the evenness of pitch scales, but no work has been done on measuring the even-

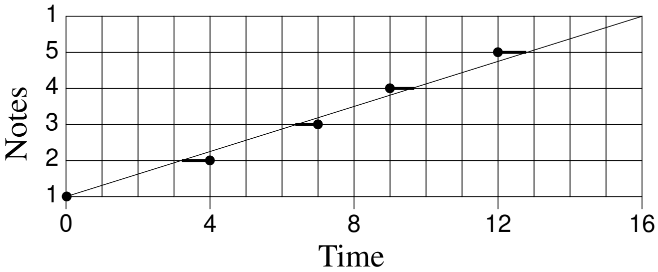

Figure 5: The linear-regression evenness measure of a rhythm.

ness of rhythm. Clough and Duthett [6] introduced the notion of *maximally even sets* with respect to scales represented on a circle. According to Block and Douthett [4], Douthet and Entringer went further by constructing several mathematical measures of the amount of *evenness* contained in a scale (see the discussion on p. 41 of [4]). One of their measures simply adds all the interval arc-lengths determined by all pairs of pitches in the scale. Their definitions may be readily transferred to rhythms. The reader may verify that according to this measure the *Bembé* is a maximally even set among all seven-note 12/8 time rhythms [23]. For a rhythm represented as a binary sequence of length n with k notes, the measure of Douthet and Entringer can be computed in $O(n + k^2)$ time.

The arc-length evenness measure of Douthet and Entringer however, is too coarse to be useful for comparing rhythm timelines such as those studied in [22] and [23]. For rhythms that differ widely from each other there is no problem. For example, the rhythms [x . . . x . . . x . . . x . . .] and [x . x . x . . x . . . .] yield evenness values of 32 and 23, respectively, reflecting clearly that the first rhythm is more even than the second. However, *all* six fundamental 4/4 time clave and bell patterns discussed in the preceeding have an equal evenness value of 48, and yet the *Rumba* clave is clearly more uneven than the *Bossa-Nova* clave. Using interval chord-lengths instead of arc-lengths, as proposed by Block and Douthet [4], yields a more discriminating measure but does not change its computational complexity. However, it is possible to define a measure of rhythmic evenness which is not only more efficient to compute than these measures, but which is sensitive enough to discriminate well between all six clave rhythms.

Michael Keith [12] proposed a measure of the *idealness* of a scale which in fact measures the evenness of the pitch intervals present in the scale. Keith's idea may be applied to measure the evenness of rhythms as follows. Consider the following 5-note example rhythm on a 16-unit time scale: [x . . . x . . x . x . . x . . .]. This sequence is mapped onto a two-dimensional grid of size 16 by 5 as pictured in Figure 5. The x-axis represents the 16 units of time (16th notes) at which the five notes are played and the y-axis indexes the five notes. The example rhythm is shown in solid black circles on the 0, 4, 7, 9, and 12 time positions. The intersections of the horizontal note-lines with the diagonal line indicate the times at which the five notes should be played to obtain a perfectly even (and boring) pattern. The deviations between these intersections and the actual positions of the notes are shown in bold line segments. The sum of these deviations serves as a measure of the un-evenness of the rhythm. Because of its similarity to linear regression fitting of data points in statistics this measure is termed the *regression-evenness* of the rhythm. The reader may readily verify that the six clave rhythms discussed in the preceeding have the following values of regression-evenness: *Bossa Nova* = 1.2, *Son* = 1.8, *Rumba* = 2.0, *Gahu* = 2.2, *Shiko* = 2.4 and *Soukous* = 2.8. Furthermore, the regression-evenness measure may be computed trivially in $O(n)$ time, much faster than the Block-Douthet measure.

References

[1] Greg Aloupis, Thomas Fevens, Stefan Langerman, Tomomi Matsui, Antonio Mesa, Yurai Nuñez, David Rappaport and Godfried Toussaint, "Computing a geometric measure of the similarity between two melodies," *Proceedings of the 15th Canadian Conference on Computational Geometry*, Dalhousie University, Halifax, Nova Scotia, Canada, August 11-13, 2003, pp. 81-84.

[2] Simha Arom, *African Polyphony and Polyrhythm*, Cambridge University Press, Cambridge, England, 1991.

[3] G. Balzano, "The group theoretic description of 12-fold and microtonal systems", *Computer Music Journal*, Vol. 4, pp. 66-84, 1980.

[4] Steven Block and Jack Douthett, "Vector products and intervallic weighting", *Journal of Music Theory*, Vol. 38, pp. 21-41, 1994.

[5] M. Chemillier and C. Truchet, "Computation of words satisfying the "rhythmic oddity property" (after Simha Arom's works). Technical Report, GREYC, University of Caen, France, 2002.

[6] J. Clough and J. Douthett, "Maximally even sets", *Journal of Music Theory*, Vol. 35, pp. 93-173, 1991.

[7] Andrew Duncan, "Combinatorial music theory", *Journal of the Audio Engineering Society*, Vol. 39, pp. 427-448, 1991.

[8] Cristian Francu and Craig G. Nevill-Manning, "Distance metrics and indexing strategies for a digital library of popular music", *Proc. IEEE International Conference on Multimedia and EXPO (II)*, 2000.

[9] Joscelyn Godwin, *The Harmony of the Spheres: A Sourcebook of the Pythagorean Tradition in Music*, Inner Traditions Intl. Ltd, 1993.

[10] Gaston H. Gonnet, "New algorithms for the computation of evolutionary phylogenetic trees", *Computational Methods in Genome Research*, S. Suhai, Ed., Plenum Press, 1994.

[11] Daniel H. Huson, "SplitsTree: analysing and visualizing evolutionary data", *Bioinformatics*, Vol. 14, pp. 68-73, 1998.

[12] Michael Keith, *From Polychords to Pólya: Adventures in Musical Combinatorics*, Vinculum Press, Princeton, 1991.

[13] E. R. Miranda, *Readings in Music and Artificial Intelligence*, Harwood Academic Publisher, 2000.

[14] M. Mongeau and D. Sankoff, "Comparison of musical sequence", *Computers and the Humanities*, Vol. 24, pp. 161-175, 1990.

[15] D. O'Maidín, "A geometrical algorithm for melodic difference", *Computing in Musicology*, Vol. 11, pp. 65-72, 1998.

[16] Joseph O'Rourke, *Computational Geometry in C*, Cambridge University Press, Cambridge, 1998.

[17] Keith S. Orpen and David Huron, "Measurement of similarity in music: a quantitative approach for non-parametric representations", *Computers in Music Research*, Vol. 4, pp. 1-44, 1992.

[18] David L. Reiner, "Enumeration in music theory", *American Mathematical Monthly*, Vol. 92, pp. 51-54, January 1985.

[19] Ronald C. Read, "Combinatorial problems in the theory of music", *Discrete Mathematics*, Vol. 167-168, pp. 543-551, April 1997.

[20] Thomas Shermer, *Personal communication*, November 2002.

[21] Godfried T. Toussaint, "Sharper lower bounds for discrimination information in terms of variation", *IEEE Transactions on Information Theory*, pp. 99-100, January 1975.

[22] Godfried T. Toussaint "A mathematical analysis of African, Brazilian and Cuban *Clave* rhythms" *Proc. BRIDGES: Mathematical Connections in Art, Music and Science*, Towson University, Towson, Maryland, pp. 157-168, July 27-29, 2002.

[23] Godfried T. Toussaint "Classification and phylogenetic analysis of African ternary rhythm timelines" *Proc. BRIDGES: Mathematical Connections in Art, Music and Science*, Granada, Spain, July 23-27, 2003.