MusMat

Brazilian Journal of Music and Mathematics

Vol. III - No. 2 - December. 2019

MusMat

Brazilian Journal of Music and Mathematics

Vol. III - No. 2

MusMat Research Group (Ed.)

Graduate Program in Music
Federal University of Rio de Janeiro (UFRJ)
Rio de Janeiro - December 2019

MusMat • Brazilian Journal of Music and Mathematics

{musmat.org}

MusMat - Brazilian Journal of Music and Mathematics. MusMat Research Group (Ed.). Rio de Janeiro: Federal University of Rio de Janeiro (UFRJ), School of Music, Graduate Program in Music, 2019. Vol. III, No. 2.

ISSN: 2526-3757

Contents

Foreword

The MusMat Group is very pleased to present the second number of 2019's volume of the *MusMat – Brazilian Journal of Music and Mathematics*. This issue opens with a paper by **David Clampitt** that presents an overview of Scale Theory via combinatorics on words, particularly the interaction between notes and words. **Hugo Tremonte de Carvalho** presents a music-oriented introduction to Markov chains and their application to Music Composition and Analysis. **Stephen Guerra'**s paper examines the relationships between solo and timeline rhythmic strucutres of Afro/diasporic musics, presenting four techniques to understand some Baden Powell's solos as cycles of a samba timeline. **Carlos Almada** reformulates the basic structure of a recursive algorithm proposed by Douglas Hofstadter in order to introduce transformational-musical tools. **Adolfo Maia and Igor Maia** present an analysis of Ligeti's *Musica Ricerta I and II* on a perspective of Information Theory and Complexity, including some of Ligeti's techniques of texture.

Liduino Pitombeira
December 2019

An Overview of Scale Theory via Word Theory: Notes and Words, Commutativity and Non-Commutativity

*David Clampitt
The Ohio State University
clampitt.4@osu.edu
Orcid: 0000-0001-5218-7186

Abstract: *About a decade ago, scale theory in music was invigorated by methods and results in the mathematical subfield of combinatorics on words. This paper reviews some of the insights stemming from this application. In particular, the interplay between notes and words is explored: the* words *of the theory capture musical intervals, but they conveniently "forget" the notes that constitute the intervals. For the purposes of mathematical music theory, this level of abstraction is usually convenient indeed, but reinstating the notes that give life to the musical interpretation in turn enriches the formal theory. Consideration of this dichotomy leads to the mathematical opposition of commutativity vs. non-commutativity, in terms of the relations between the non-abelian free group F_2 and the additive abelian group $\mathbb{Z}^2$, and between the free non-commutative and commutative monoids embedded, respectively, in those groups.*

Keywords: *Christoffel words, conjugation, Sturmian morphisms, well-formed scales*

I. Introduction

Word theory studies sequences of elements drawn from a set or *alphabet*; such sequences, finite or infinite, are what are meant by *words*. From around the 1980s, the subfield called algebraic combinatorics on words brought various mathematical disciplines to bear on these objects, with a great deal of growth in the past three decades. This happens to coincide with the flourishing of formal studies of musical scales, which proceeded largely in ignorance of the relevant work in mathematics and computer science. While the joining of word theory to scale theory within mathematical music theory is a fairly recent development, there were some applications to music from the beginning, principally in studies of rhythm and in certain compositional models. An historical survey may be found in [1], an editorial by the three guest editors of a recent special issue of *Journal of Mathematics and Music* devoted to music and combinatorics on words. Four articles ([2], [3], [4], [5]) in this special issue address the present state of scale theory in light of word theory (which in this article refers to algebraic combinatorics on words, given an exposition in the pseudonymous Lothaire series, [6], [7], [8]).

*Thank you to my colleagues in these studies, Norman Carey and Thomas Noll.

In the scale theory of the 1980s, the focus was typically on properties of certain distinguished pitch-class sets, drawn from the usual 12-tone equal-tempered universe (or, generalizing, from some equal-tempered universe of n pitch classes, modeled by $\mathbb{Z}_n$). As such, although the canonical examples qualified in some ill-defined sense as musical "scales," formally the objects studied were simply pitch-class sets, that is, subsets of $\mathbb{Z}_{12}$, or $\mathbb{Z}_n$. Although scale order was typically adduced, it was a cyclic order: no tonic or modal final was posited. By convention, the representative set is its prime form, which for the usual diatonic (Forte's set-class 7-35) is $\{0, 1, 3, 5, 6, 8, 10\}$.[1]

For the purposes of the present article the germinal 1985 work was [9], in which Clough and Myerson studied generalizations of the usual diatonic pitch-class set, which they called *diatonic systems*. To define diatonic systems they generalized the property that generic non-zero diatonic pitch-class intervals come in two specific sizes (generic diatonic seconds are specifically major or minor; similarly, generic thirds; generic fourths are specifically perfect or augmented, etc.), calling this Myhill's Property (MP). They proved that all diatonic systems, i.e., those with MP, of cardinality N, possess the apparently stronger but in fact equivalent property Cardinality equals Variety (CV), wherein ordered sets of cardinality $1 \leq k \leq N$ of a given generic description come in k specific varieties (e.g., diatonic triads, formed from 3 notes separated by two generic thirds, come in the 3 species major, minor, and diminished); for details, see [9]. Among other properties, they showed that diatonic systems admit a *generating interval* or generalized fifth, such that all of the pitch classes of the system may be ordered in a chain in which adjacent elements are separated by the generating interval. In the case of the usual diatonic, the generating interval is either the perfect fifth or perfect fourth. In the prime form of 7-35 the chain of perfect-fourth related pitch classes is 0, 5, 10, 3, 8, 1, 6; that is, the set is generated by $5j$ mod 12, $j = 0, 1, \ldots, 6$. Clough and Myerson derive an equivalent algorithm to generate the pitch classes in scale order (ascending in the usual order for integers), by taking floors of multiples of $\frac{12}{7}$: $\lfloor j\frac{12}{7} \rfloor$, $j = 0, 1, \ldots, 6$. In general, for a chromatic universe of cardinality n and an embedded diatonic system of coprime cardinality $N < n$, the algorithm is $\lfloor j\frac{n}{N} \rfloor$, $j = 0, 1, \ldots, N - 1$.

Observe that in this presentation, two orderings of the pitch classes, that is, the musical *notes*, are set in relief. The scale order comes into play because the generic interval notion uses the scale itself as a rough measurement, with the number of steps of the scale determining the generic interval, without regard for the quality or specific interval. The other ordering principle is that given by the generating interval. The relationship between these two orderings is essential to the construction, as we will see. For now, note that the generic measure for a cardinality N diatonic system is modeled by the cyclic group $\mathbb{Z}_N$ with addition modulo N. Then the generating-interval ordering is the image of an injective map from the scale ordering into the chromatic universe. In the case of the usual diatonic $f : \mathbb{Z}_7 \to \mathbb{Z}_{12} : z \to f(z) = 5z$ mod 12 is injective, preserving the order of ordinary arithmetic. Since f is injective, and since $5^{-1} = 5$ mod 12, multiplication by 5 modulo 12 applied to the image of f returns the generic scale order, 0, 1, ..., 6. This relationship is not explicit in [9], but it is implicit. Again, neither ordering confers a privileged status on a particular musical mode. We might consider that mathematically the orderings beginning with 0 should have priority, but musically we might resist that suggestion, since in the case of the usual diatonic the algorithms yield the (ill-favored) Locrian mode

In 1989, Carey and Clampitt in [10] took the relationship between the orderings as a point of departure, and generalized beyond rational generators, to define *well-formed scales*. Instead of taking

[1] The style of this article is rather discursive for a mathematical study, rather formal and algebraic for a music theory work. In order to avoid tedious and awkward formulations, certain musical concepts are taken as known, and similarly mathematical terminology and concepts that are common in mathematics are introduced without comment. No knowledge of word theory is assumed. Because this is in the nature of an overview, quite a bit of the development refers to recent literature.

MP as an axiom, they considered generated sets, and wished to determine which scales had desired properties, again taking the usual diatonic as a model. They allowed the generating interval to be either rational or irrational with respect to the octave, normalized as unity. That is, mathematically the situation is modeled by $\mathbb{R}/\mathbb{Z}$, with addition modulo 1. In this environment, we may generate the twelve-tone equal-tempered usual diatonic by taking fractional parts of multiples of $\frac{7}{12}$, representing the equal-tempered perfect fifth. If we notate the fractional part of a real number x as $\{x\} = x - \lfloor x \rfloor$, then the equal-tempered diatonic is modeled by $\{j\frac{7}{12}\}, j = 0, 1, \ldots, 6$; in scale order $0 < \{2(\frac{7}{12})\} = \frac{1}{6} < \{4(\frac{7}{12})\} = \frac{1}{3} < \{6(\frac{7}{12})\} = \frac{1}{2} < \{1(\frac{7}{12})\} = \frac{7}{12} < \{3(\frac{7}{12})\} = \frac{3}{4} < \{5(\frac{7}{12})\} = \frac{11}{12}$. But we might also generate the Pythagorean diatonic by fractional parts of multiples of the just perfect fifth, represented by the irrational $\log_2\frac{3}{2}$: $\{j(\log_2\frac{3}{2})\}$, $j = 0, 1, \ldots, 6$; in scale order $0 < \{2(\log_2\frac{3}{2})\} < \{4(\log_2\frac{3}{2})\} < \{6(\log_2\frac{3}{2})\} < \{1(\log_2\frac{3}{2})\} < \{3(\log_2\frac{3}{2})\} < \{5(\log_2\frac{3}{2})\}$. What characterizes both scales as well-formed diatonic scales is that the transformation on the index set $\mathbb{Z}_7$ representing the successive perfect fifth multiples to their position in scale order is the automorphism of the additive cyclic group $\mathbb{Z}_7 \to \mathbb{Z}_7$: $z \to 2z \bmod 7$. The definition of a well-formed scale is therefore: a generated set such that the mapping from the set in generation order to the set in scale order is a group automorphism of the cyclic group $\mathbb{Z}_n$ under addition modulo n. It follows that, equivalently, a generated set is well-formed if and only if the mapping of the set in scale order to the set in generation order is a group automorphism of the underlying cyclic group.

Since the n-note equal division of the octave (or whatever the interval of periodicity is) may be generated by any integer g coprime with n, it follows that the scale is well-formed. These cases are the *degenerate* well-formed scales; all other cases are non-degenerate, and unless otherwise specified, well-formed is by default non-degenerate.

The answer to the question of which scales, for a given generator, are well-formed was answered in [10]: all and only those generated scales with cardinalities equal to the denominator of a convergent or semi-convergent in the continued fraction representation of the size of the generator. For the just perfect fifth, $\log_2(\frac{3}{2})$, this sequence is 1, 2, 3, 5, 7, 12, 17, 29, 41, 53, 94, ... Among the musically significant scales in this hierarchy are the *tetractys* or tonic-subdominant-dominant roots, the usual pentatonic, the usual diatonic, and the chromatic, distinguishing between diatonic and chromatic semitones. Given any generating interval θ such that $4/7 < \theta < 3/5$, the continued fraction will begin with denominators 1, 2, 3, 5, 7, 12. In the case where $\theta = 7/12$, the final scale in the hierarchy is the degenerate well-formed 12-note equal division of the octave. The mathematical proof of the equivalence of well-formedness and the continued fraction property, and also the proof that a generalized MP (allowing for irrational specific interval sizes) is equivalent to non-degenerate well-formedness, are found in [11] and [12].

For the usual diatonic, we can establish the formal equivalence of scales in all tunings for generating intervals θ within the limit set above. As we have seen, the scales in this class are characterized by the mapping $\mathbb{Z}_7 \to \mathbb{Z}_7$: $z \to 2z \bmod 7$, so that the scale order in terms of the perfect fifths generation order is the sequence 0 2 4 6 1 3 5 (0). While generic step intervals are all represented as differences modulo 7 between adjacent elements of the sequence, i.e., 2 mod 7, the sequence of specific step intervals is represented by the sequence of differences in ordinary arithmetic between adjacent elements, 2 2 2 -5 2 2 -5. This is a way of encoding the sequence of tones and semitones of the Lydian mode, $TTTSTTS$, but so far in scale theory no particular rotation of this word was privileged over another, so in this first step toward word theory, the word $TTTSTTS$ was construed as a *circular word*, that is, as the equivalence class of all seven distinct rotations of the word. In word theory generally, though, the seven distinct rotations, or *conjugates* as they are referred to in this field, are considered distinct *words*, opening a possible avenue to a study of the diatonic modes, and to the modal varieties of well-formed scales in general. That there are N distinct modes of a (non-degenerate) well-formed scale of cardinality N

is a known consequence of CV, which follows from generalized MP, but it was not obvious that simply encoding the modes in N distinct words would lead to a more refined modal knowledge.

II. A Modal Refinement Through Word Theory

The subject matter of combinatorics on words arose, but only implicitly, in two 1996 articles. In [13], the concept of a *region* was introduced, equivalent to the *central word*, and in [14], an infinite word was constructed, which was equivalent to a *Sturmian word* (in this case of slope $\sqrt{2}$). In the latter article, equivalence classes of well-formed scales were defined as words characterized, up to rotation, by cardinality N and by the multiplicity of one of the step intervals, g. Thus, all possible tunings of the usual diatonic are in the equivalence class of all rotations of the word *aaabaab*, determined by the ordered pair $(N, g) = (7, 2)$. Even a dual class to the diatonic, all rotations of *abababb*, determined by $(N, g^{-1}_{\text{mod } N}) = (7, 4)$, was defined, but there was no acknowledgment yet that these objects had been defined abstractly within mathematics. The applications of word theory to musical scale theory began in 2007, at the First International Conference of the Society for Mathematics and Computation in Music in Berlin in May, and at the Sixth International Conference on Words (WORDS 2007) in Marseille-Luminy in September.

i. The monoid of words over an alphabet

The *monoid of words* over an alphabet A is the set A^* of all finite sequences or strings of letters drawn from a finite set of symbols or *alphabet* A: $A^* = \{w = w_1 \dots w_n | w_i \in A, i = 1, \dots, n, n \in \mathbb{N}\}$. The associative monoid operation is concatenation of words, and it is understood that the empty word ε is in A^* and is an identity. A^* is thus also a semigroup with identity. If $u, v \in A^*$ with $w = uv$, then we say that u and v are *factors* of w; ε is a factor of any word w. A *prefix* of a word is a factor that begins that word, and a *suffix* of a word is a factor that ends that word. The *length* $|w|$ of a word is the number of letters it contains: if $w = w_1 w_2 \dots w_k$, $|w| = k$. Another notation: given a letter ℓ from our alphabet A, $|w|_\ell$ refers to the number of occurrences of the letter ℓ in the word w.

Note that the monoid A^* is free. Two words $u, v \in A^*$ commute if and only if they are powers of the same word: $uv = vu \iff u = t^j, v = t^k$ for some $t \in A^*, j, k \in \mathbb{N}$. For example, if $A = \{a, b\}, u = aab, v = (aab)^2 = aabaab$, then $uv = vu = (aab)^3$.

For example, our alphabet might be $A = \{a, b, c\}$, and it might be convenient to consider A to be ordered, for example, $a < b < c$. If we wished to model the Japanese *hira joshi* mode, E-F-A-B-C-(E′), we might represent its sequence of scale step intervals by the word *acbac*, where $a = 1$ semitone interval, $b = 2$ semitones interval, and $c = 4$ semitones interval. Evidently, the information about the particular notes specifying this mode is lost in the representation as a word; it follows that this word with this interpretation may represent any or all musical transpositions of this mode. But different modes of this scale are represented by different words, different cyclic rotations of *acbac*, or as word theorists prefer to say, different *conjugates* of the word. Two words w and w' in A^* are *conjugate* (or conjugates of each other) if $w = uv$ and $w' = vu$ for some $u, v \in A^*$. The full set of conjugates of a word w is called its *conjugacy class*, and this set defines the *circular word* (w). The conjugacy class of $acbac = \{acbac, cbaca, bacac, acacb, cacba\}$. It should be evident that conjugacy is an equivalence relation on A^*. Note that for any of the conjugates r, we have $|r|_a = 2, |r|_b = 1, |r|_c = 2$.

ii. Sturmian endomorphisms

Our words for well-formed scales will be over two-letter ordered alphabets, $A = \{a < b\}$ or $A = \{x < y\}$. Any mapping $f : A^* \to A^*$ that replaces every occurrence of a with the word $f(a)$ and every occurrence of b with the word $f(b)$ is by definition an endomorphism of the monoid A^*: if $w = w_1w_2 \ldots w_n \in A^*, f(w) = f(w_1w_2 \ldots w_n) = f(w_1)f(w_2) \ldots f(w_n)$. It follows that, for any $u, v \in A^*, f(uv) = f(u)f(v)$, i.e., f is an endomorphism.

The distinguished endomorphisms (henceforth *morphisms*) of A^* are the Sturmian morphisms St in [7]. The generators of St are defined as follows:

$$\begin{array}{l|l|l|l} G(a) = a & \tilde{G}(a) = a & D(a) = ba & \tilde{D}(a) = ab \\ G(b) = ab & \tilde{G}(b) = ba & D(b) = b & \tilde{D}(b) = b \end{array}$$

$$E(a) = b, E(b) = a.$$

The set of all compositions of these morphisms (together with the identity mapping) forms the monoid St under composition of mappings. The submonoid St_0 of special Sturmian morphisms excludes the exchange morphism E, and is generated by the remaining four morphisms. Neither St nor St_0 are freely generated. For one thing, G and $\tilde{G}$ commute, and D and $\tilde{D}$ also commute. Moreover, the special Sturmian morphisms have the following presentation ([7], [15]):

$$St_0 \cong \left\langle G, \tilde{G}, D, \tilde{D} \quad | \quad GD^k\tilde{G} = \tilde{G}\tilde{D}^kG, DG^k\tilde{D} = \tilde{D}\tilde{G}^kD \text{ for all } k \in \mathbb{N} \right\rangle$$

.

Certain pairs of these morphisms generate distinguished free submonoids of St_0: the standard morphisms, $\langle G, D\rangle$; the anti-standard morphisms, $\langle \tilde{G}, \tilde{D}\rangle$; the Christoffel morphisms, $\langle G, \tilde{D}\rangle$; and the anti-Christoffel morphisms, $\langle \tilde{G}, D\rangle$.

Since the morphisms are completely defined by their actions on the single letters a and b, we must begin by applying them to either the word ab or ba. ab is considered to be a *Christoffel* word, because beginning with a single letter, $\tilde{D}(a) = ab$ and $G(b) = ab$; either way ab is the image under a Christoffel morphism. Similarly, ba is considered to be an *anti-Christoffel* word (and we classify other words as standard or anti-standard in the same way). As a matter of convention, we choose the Christoffel word ab as root word. We apply compositions of Sturmian morphisms to the root word ab, and in order to keep track separately of the images of a and b, we often introduce a divider symbol, writing "$a|b$" or, in this article, "a, b".

St and St_0 enter into mathematical music theory by offering another means of generating all and only well-formed scales, and moreover to address the modal identities. This approach also ratifies the distinction in the history of music theory between authentic and plagal modes. A few examples will illustrate these matters.

In the musical interpretation, we understand the root word a, b as expressing an octave, divided into perfect fifth and perfect fourth, a and b, respectively. This is known in modal theory as, going back to medieval chant, the authentic division of the octave (perfect fifth plus perfect fourth above the modal final). Just as ab is privileged over ba in the construction of the mathematical theory, in the history of modal theory, the authentic division is privileged over the plagal division (perfect fourth below the final, plus perfect fifth above the modal final). The final (*finalis*) is a *note*; as such it is has no role in word theory, but it is essential to the musical interpretation. Similarly, we may regard the divider symbol as standing in for a musical note: in the C-G-(C′) authentic division of the octave, we say that note G is the *divider*. If we apply morphism D to a, b, we have $D(a, b) = ba, b$. In the interpretation, the meaning of the letter b continues to be the interval perfect fourth, because D leaves b fixed, but a now represents the diatonic whole step, the difference between a perfect fifth and perfect fourth, as D replaces a by ba, perfect fourth followed by whole step. The word ba, b models, then, the ancient Greek *tetractys*, C-F-G-(C′). Here, C is again the final and G is the divider. Applying $\tilde{D}$, on the other hand, we have $\tilde{D}(a, b) = ab, b$, representing the other mode of the *tetractys*, divided authentically. Remark that ba, b may be considered a standard word and ab, b may be considered a Christoffel word. If we prepend G to compose it with D, we obtain the (now unambiguously) standard word $GD(a, b) = G(ba, b) = aba, ab$. This may be interpreted as the usual pentatonic scale, in the mode C-D-F-G-A-(C′). Under the action of G, a retains its meaning (fixed under G), while now the interpretation of b as perfect fourth is replaced by the (diatonic) minor third, which is a pentatonic step interval. Composing G with $\tilde{D}$, and composing $\tilde{G}$ with both D and $\tilde{D}$ we exhaust the possibilities, and the images of a, b form standard, Christoffel, anti-Christoffel, and anti-standard words, four of the five conjugates in the conjugacy class, representing four of the five modes of the usual pentatonic.

In his 1547 treatise *Dodecachordon*, Glarean expanded the system of diatonic modes, adding authentic and plagal Ionian and Aeolian modes, an increase from eight to twelve modes. Thus, six of the seven octave species are represented, each in authentic and plagal form. For example, authentic Dorian is divided $TSTT, TST$, or in our alphabet A, $abaa, aba$. The remaining modal possibilities (modern Locrian) do not support authentic and plagal divisions and were rejected by Glarean: (*hyperaeolius reiectus* or diminished fifth plus augmented fourth; *hyperphrygius reiectus* or augmented fourth plus diminished fifth).

The twelve Glarean modes are captured by elements of St as images of the root Christoffel word ab: the six authentic modes as images of special Sturmian morphisms, the six plagal modes as images of ba under the same morphisms; in other words, images of ab of morphisms in the complement of St_0 (the six morphisms associated with the authentic modes but preceded by E; see below). The rejected Locrian modes are also excluded in this environment: they are associated with the "bad conjugate" of the word theorists (the divided words in the conjugacy class of a Christoffel word that are not the images of a, b nor of b, a under an element in St_0; [16]). We refer to words such words as *amorphic*; words that may be realized as images of these morphisms are *morphic*. To reiterate, the authentic and plagal divisions are entailed by separating the images of a and b. For example, the authentic Ionian pattern of step intervals is captured transformationally by the standard morphism $GGD(a, b) = GG(ba, b) = G(aba, ab) = aaba, aab$. Word theorists refer to the division as a *standard factorization*, $f(a)f(b)$, for some special Sturmian morphism f (different meaning here of "standard")[18]. The transformational representations with their traditional modal finals and modern names are presented below (keeping the collection fixed as the "white notes" or C-major set; see [17] for details).

Transformation	Mode
$GGD(a, b) = aaba, aab$	C authentic Ionian
$\tilde{G}GD(a, b) = abaa, aba$	D authentic Dorian
$\tilde{G}\tilde{G}D(a, b) = baaa, baa$	E authentic Phrygian
$GG\tilde{D}(a, b) = aaab, aab$	F authentic Lydian
$\tilde{G}G\tilde{D}(a, b) = aaba, aba$	G authentic Mixolydian
$\tilde{G}\tilde{G}\tilde{D}(a, b) = abaa, baa$	A authentic Aeolian

Transformation	Mode
$GGDE(a,b) = aab, aaba$	C plagal Ionian
$\tilde{G}GDE(a,b) = aba, abaa$	D plagal Dorian
$\tilde{G}\tilde{G}DE(a,b) = baa, baaa$	E plagal Phrygian
$GG\tilde{D}E(a,b) = aab, aaab$	F plagal Lydian
$\tilde{G}G\tilde{D}E(a,b) = aba, aaba$	G plagal Mixolydian
$\tilde{G}\tilde{G}\tilde{D}E(a,b) = baa, abaa$	A plagal Aeolian

One may observe that the musical interpretations of ab, $D(ab)$, $GD(ab)$, and $GGD(ab)$ yield modes of well-formed scales in the hierarchy of such scales generated by the perfect fifth (to within some reasonable size) modulo the octave, those of cardinalities 2, 3, 5, 7. If one continues with $DGGD(ab)$, the well-formed 12-note chromatic follows. Indeed, we may identify modes of well-formed scales with the conjugacy classes of words generated in this fashion, as discussed in [17]. Where well-formed scales are defined by considering the distribution of integer multiples of a real number modulo 1, the same scales appear in the guise of words, which model the filling in of larger step intervals by smaller ones at each step. The procedures appear different, but they are equivalent.

So far the words representing modes of well-formed scales have been defined as the images of the minimal root word ab under Sturmian morphisms. One might well ask how to characterize such words more directly. One may also ask if there is a mathematically privileged member of the conjugacy class, to serve as an identifier for the class. A geometric definition, point of departure in [18], provides affirmative answers. In Figure 1 we see a construction which yields what Berstel et al. call lower and upper Christoffel words [18]; the terminology is adjusted here to be consistent within this article. Note that the word encoded by the lower path is $aaabaab$, the word representing Lydian mode, which was the mode suggested by the well-formed scale construction of the diatonic set generated by perfect fifths; see Section I. The word encoded by the upper path is $baabaaa$, the word representing Locrian mode, which was the mode suggested by the Clough-Myerson algorithm (or by the well-formed scale construction, taking perfect fourth as generator).

To integrate this geometric construction with the transformational perspective, consider a refinement of Figure 1 which considers the points on the paths, lower and upper, that approach closest to the line segment of slope 2/5. Make that point the divider of the respective words, and the lower word is $aaab, aab$, or the result of $GG\tilde{D}(a,b)$, a Christoffel word, representing authentic Lydian; the upper word is $baa, baaa$, or the result of $\tilde{G}\tilde{G}DE(a,b)$, an anti-Christoffel word (anti-Christoffel morphism applied to anti-Christoffel word b, a), representing plagal Phrygian.

Note that the upper and lower words are reversals of each other. The general Christoffel/anti-Christoffel word construction follows the example: given positive coprime integers p, q, one constructs the line segment of slope q/p from the origin to point (p, q), and the lower and upper polygonal paths through the nearest lattice points to the line segment, such that no lattice points are included within the region enclosed by the paths. The words are encoded over a two-letter alphabet; by convention, the lexicographically first letter is assigned to the horizontal line segments, the second letter to the vertical line segments. It follows that lower Christoffel words are lexicographically least of their conjugacy class ([18]). The words w are of length $|w| = N = p + q$, with $|w|_a = p$, $|w|_b = q$. The shortest non-trivial Christoffel/anti-Christoffel pair are those of slope 1, ab and ba, respectively. The trivial cases are the word of slope 0, a, and the word of slope ∞, b. Words defined by irrational slopes are infinite words of minimal complexity, called *Sturmian*. We won't discuss infinite words in this article (members of A^* are finite).

Foreshadowing the focus of this article, that is, commutativity vs. non-commutativity, we can propose that the preference in word theory for the lower, Christoffel word over the upper,

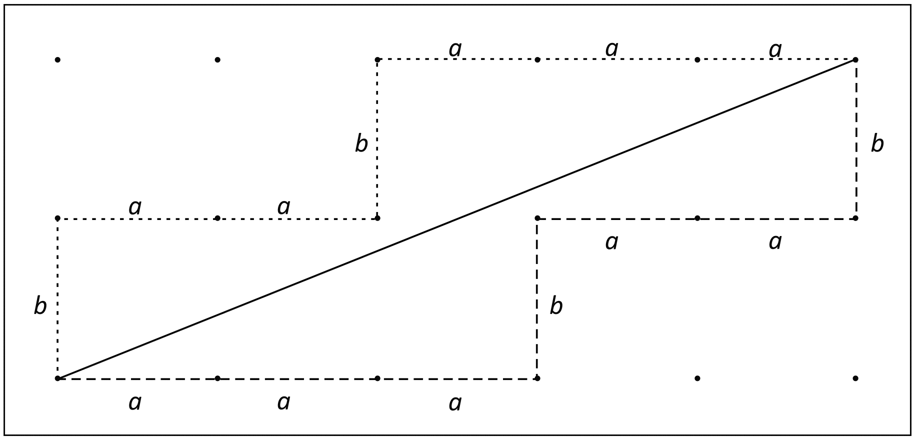

Figure 1: *Lower/(upper) Christoffel/(anti-Christoffel) words of slope $q/p = 2/5$, defined by paths (dashed and dotted, respectively) from the origin $(0,0)$ to $(5,2)$ that connect points of the lattice such that no points of the lattice lie in the region of the plane between the respective paths and the line segment of slope $2/5$. The words are encoded by labeling horizontal unit line segments a and vertical unit line segments b. These are also sometimes referred to as* digitized line segments.

anti-Christoffel word, has a basis in mathematics, part and parcel of a preference for special Sturmian morphisms over those in the complement of St_0 in St. This latter preference mirrors the music-theoretical bias towards the authentic modes over the plagal modes, as reflected in the very terminology. We will introduce a matrix associated with an element f of St, its *incidence matrix*, M_f, defined in terms of $f(a,b)$.

$$M_f = \begin{pmatrix} f(a)_a & f(b)_a \\ f(a)_b & f(b)_b \end{pmatrix}$$

Details of this construction will be given below, but for now consider $M_{GG\tilde{D}}$ and $M_{\tilde{G}\tilde{G}DE}$. Since $GG\tilde{D}(a,b) = aaab, aab$ and $\tilde{G}\tilde{G}DE(a,b) = baa, baaa$ we have:

$$M_{GG\tilde{D}} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$$

and

$$M_{\tilde{G}\tilde{G}DE} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}.$$

The mathematical distinction is that $det(M_{GG\tilde{D}}) = 1$, whereas $det(M_{\tilde{G}\tilde{G}DE}) = -1$. In other words, $M_{GG\tilde{D}}$ is in the special linear group $SL(Z)_2$. This is the motivation for the qualification *special* for St_0. The incidence matrices of elements in St_0 have determinant 1, while the incidence matrices of elements in St in the complement of St_0 have determinant -1. This distinction, in conjunction with the geometric construction of the lower Christoffel word, is the motivation for the word theory terminological choice of Christoffel to designate this class of words, and for the Christoffel word of a given slope to stand in for the conjugacy class of that word. In this article then, when the authentic/plagal distinction is at issue, the focus will be on the authentic side.

III. Christoffel Duality and Scale Foldings

In order to better understand distinctions among the modes, we triangulate between the words representing their patterns of step intervals and words to be introduced in this section called *scale foldings*. The latter follow what we earlier called the generation order for a well-formed scale, but in a way that is determined by the mode.

The usual diatonic may be generated by either the perfect fifth or the perfect fourth. Generalizing to well-formed scales generated modulo 1, the generator may be $\theta \in \mathbb{R}$, $0 < \theta < 1$, or $1 - \theta$. Translated into musical terms, we may generate the same (C-major) collection, either by rising perfect fifths (falling perfect fourths), departing from F, or by rising perfect fourths (falling perfect fifths), departing from B: F-C-G-D-A-E-B, or B-E-A-D-G-C-F. Whereas we by convention choose to conceive of and notate our scales as ascending, there is not necessarily a musical reason to prefer fifths to fourths, or, equivalently, rising perfect fifths to descending perfect fifths. Will the mathematics again tell us which to prefer?

In [17], scale foldings for the modes of usual diatonic were defined as follows (the generalization for modes of all well-formed scales follows this model). Given a diatonic mode, understood as an octave species, i.e., without the authentic/plagal distinction, within a fixed C-major collection, a *forward folding* is the sequence of upward perfect fifths and downward perfect fourths, departing from note F within the modal octave, such that all notes of the mode lie within the modal octave, including modal final and excluding the octave above the modal final. For example, for D-Dorian, from the F4 within the D4 to D5 octave one can extend a rising perfect fifth to C5, but then one is forced down a perfect fourth to G4, again down a perfect fourth to the modal final D4, up a perfect fifth to A4, down a perfect fourth to E4, up to B4. From B4 we complete the folding with a downward perfect fourth to the excluded note F♯4 (just as to complete the pattern *abaaaba* of scale step intervals for Dorian we need the excluded upper octave D5: D4-E4-F4-G4-A4-B4-C5-(D5). We choose a different alphabet to encode the forward folding word, with x for rising perfect fifth and y for falling perfect fourth. In the case of Dorian, the forward folding word is thus *xyyxyxy*. Proceeding through the modes in similar fashion, we have a one-to-one mapping between words in the conjugacy class of words representing scale step interval patterns and words representing the respective forward folding patterns (generally a different conjugacy class). Both conjugacy classes are the conjugacy classes of Christoffel words. Referring back to the opening paragraph of Section II, for the usual diatonic these are the conjugacy classes for the dual well-formed diatonic scale classes $(N, g) = (7, 2)$ and $(N, g^{-1}_{\text{mod } N}) = (7, 4)$, the conjugacy classes of lower Christoffel words *aaabaab* of slope $2/5$ and *xyxyxyy* of slope $4/3$, respectively. On the other hand, given a diatonic mode, the *backward folding* considers the generation by perfect fourth, departing from note B within the modal octave, such that all notes of the mode lie within the modal octave, again including the modal final and excluding the octave above the modal final. Again, for example in D-Dorian, from B4 within the B4 to B5 octave, one finds E4 a perfect fifth below B4, then a rising perfect fourth from E4 to A4, then a descending perfect fifth to the modal final, D4, then ascending perfect fourth to G4, again an ascending perfect fourth to C5, descending perfect fifth to F4, and finally ascending perfect fourth to (excluded) B♭. Encoding the backward folding word again over the alphabet $A = \{x, y\}$, now with x representing descending perfect fifths and with y representing ascending perfect fourths, the backward folding word for Dorian is *xyxyyxy*. This word is different from the forward folding word for Dorian, but it is a member of the same conjugacy class. If we go through the backward folding words for the diatonic modes, we again exhaust the conjugacy class of the lower Christoffel word of slope $4/3$.

Table 1 aligns the modes, in circle of fifths order, with their modal scale words, and with their associated forward and backward folding words. Observing the sequence of modal scale

Table 1: *The diatonic modes, their scale words, and associated forward and backward folding words*

final	mode	scale word	forward folding word	backward folding word
F	Lydian	*aaabaab*	*xyxyxyy*	*yxyxyxy*
C	Ionian	*aabaaab*	*yxyxyxy*	*xyxyxyy*
G	Mixolydian	*aabaaba*	*yyxyxyx*	*yxyxyyx*
D	Dorian	*abaaaba*	*xyyxyxy*	*xyxyyxy*
A	Aeolian	*abaabaa*	*yxyyxyx*	*yxyyxyx*
E	Phrygian	*baaabaa*	*xyxyyxy*	*xyyxyxy*
B	Locrian	*baabaaa*	*yxyxyyx*	*yyxyxyx*

words, we see that circle of fifths order coincides with lexicographic order (with $a < b$), beginning with the lower Christoffel word of the conjugacy class. The sequences of the associated folding words are rotations by one position, rotations by one to the left for forward foldings, one to the right for backward foldings. To be more precise, and to motivate the monoid terminology of conjugates/conjugacy, we extend the monoid $A^* = \{x, y\}^*$ to the free group on two letters, F_2. With the adjunction of the inverse letters, $F_2 = \{x, y, x^{-1}, y^{-1}\}^*$, becomes a group, again with concatenation of words as the closed, associative product, the empty word ε as group identity, and usually with the understanding that group elements are represented by reduced words, wherever a letter and its inverse are adjacent and may be canceled. We define *conjugation by* u for $u \in F_2$ by $conj_u : F_2 \to F_2 : w \to u^{-1}wu$. In any group such a mapping defines an inner automorphism of the group, so the group of inner automorphisms of F_2 is $Inn(F_2) = \{conj_u | u \in F_2\}$. The circle of fifths order defines a linear order, since the sequence of diatonic perfect fifths begins with F and ends with B–the lexicographic order on the modal scale words is similarly linear–but conjugation by initial letters defines a cyclic order through the conjugacy class. In the last column of Table 1, aligned with the circle of fifths order of the modes, the words for backward foldings follow the cyclic order of conjugation by initial letters, e.g., beginning with Lydian, $y^{-1}(yxyxyxy)y = xyxyxyy$, $x^{-1}(xyxyxyy)x = yxyxyyx$, $y^{-1}(yxyxyyx)y = xyxyyxy$, $x^{-1}(xyxyyxy)x = yxyyxyx$, $y^{-1}(yxyyxyx)y = xyyxyxy$, $x^{-1}(xyyxyxy)x = yyxyxyx$, and conjugation of $yyxyxyx$ by y returns to the Lydian backward folding word. Similarly, for the forward foldings, aligned with the circle of fifths order of the modes, the words follow a cyclic order, but now of conjugation by the inverses of initial letters (e.g., since $(y^{-1})^{-1} = y$, beginning with the forward folding word for Lydian, we have $conj_{y^{-1}}(xyxyxyy) = yxyxyxy$, forward folding word for Ionian). Put another way, conjugating by initial letters aligns backward foldings with ascending circle of fifths order for the modes, while conjugating by initial letters aligns forward folding with descending circle of fifths order for the modes. Forward and backward foldings coincide for Aeolian.

What does mathematics say as to forward foldings vs. backward foldings? Initially, we must recognize that mathematics knows nothing of foldings, because word theory does not recognize the notes that for the music theorist determine the words, the notes "between" the letters of the words, as it were. The foldings and their directions only arise in the interpretations assigned to the words.

But the mathematics does decide which interpretation makes for the better one-to-one mapping between the respective conjugacy classes. This clarification happens when we lift from the level of the words to the level of the morphisms in St_0. In the process, the cyclic ordering of the words is cut at a certain point to become a linear ordering of the morphisms. This is the level of *Sturmian involution*, first defined in [16]. Under Sturmian involution, a morphism f in St_0 is mapped to the morphism f^* such that f^* is the reversal of f, where every G and $\tilde{G}$ in the composition of

Table 2: *The diatonic modes, morphisms of modal scale words, and morphisms of backward folding words*

final	mode	morphism f	scale word	morphism f^*	backward folding word
C	Ionian	$GGD(a,b) =$	*aaba,aab*	$\tilde{D}GG(x,y) =$	*xy,xyxyy*
D	Dorian	$\tilde{G}GD(a,b) =$	*abaa,aba*	$\tilde{D}G\tilde{G}(x,y) =$	*xy,xyyxy*
E	Phrygian	$\tilde{G}\tilde{G}D(a,b) =$	*baaa,baa*	$\tilde{D}\tilde{G}\tilde{G}(x,y) =$	*xy,yxyxy*
F	Lydian	$GG\tilde{D}(a,b) =$	*aaab,aab*	$DGG(x,y) =$	*yx,yxyxy*
G	Mixolydian	$\tilde{G}G\tilde{D}(a,b) =$	*aaba,aba*	$DG\tilde{G}(x,y) =$	*yx,yxyyx*
A	Aeolian	$\tilde{G}\tilde{G}\tilde{D}(a,b) =$	*abaa,baa*	$D\tilde{G}\tilde{G}(x,y) =$	*yx,yyxyx*
B	Locrian	bad conjugate	*baabaaa*	bad conjugate	*yyxyxyx*

f is left fixed, and every D composing f is replaced in f^* by $\tilde{D}$, and every $\tilde{D}$ composing f is replaced in f^* by D. This is an anti-automorphism of St_0, and it is clear that $f^{**} = f$. For example, $(GGD)^* = \tilde{D}GG$. Moreover, from the definitions it is clear that Sturmian involution exchanges standard morphisms with Christoffel morphisms.

In Table 2 it becomes apparent that the backward folding assignment conforms with Sturmian involution, still using the diatonic modes as the canonical example. If we follow the forward folding assignment, there is a mismatch between the bad conjugate modal scale word, the representative of Locrian, which has a morphic folding word, and the bad conjugate folding word, which corresponds to the morphic scale word for authentic Mixolydian. However, for the major mode, whose historical ancestor is arguably authentic Ionian, there are music theoretical dividends paid by accepting the forward folding interpretation, as detailed in [17] and discussed in Section IV. In the general musical situation, the forward folding interpretation pairs standard scale words with standard folding words, and Christoffel scale words with Christoffel folding words, the associated morphisms being each other's reversal. For example, for the usual pentatonic, the standard word represents the scale C-D-F-G-A, $GD(a,b) = aba, ab$, whose forward folding F-C-G-D-A-(E) is the standard word $DG(x,y) = yx, yxy$. We have the freedom to choose alternative interpretations in the musical context because of the level of structure afforded by the sequences of notes that lie behind the words.

In Table 2 the modes are ordered in scale order as opposed to circle of fifths order. The *conjugation class* of a special Sturmian morphism may be defined (see Section IV), and it carries a natural linear ordering by virtue of the conjugations by single letters, $conj_a$ and $conj_b$ [19]. Starting from the standard morphism of the class as least element, for every morphism f in the class—except for the anti-standard morphism—either $conj_a \circ f$ or $conj_b \circ f$ is in the class and can be identified as the *successor* of f. The anti-standard morphism is the greatest element in the class, in *conjugation order*, $<_{conj}$. In Table 2 the morphisms are ordered from top to bottom in conjugation order, $GGD <_{conj} \tilde{G}GD <_{conj} \tilde{G}\tilde{G}D <_{conj} GG\tilde{D} <_{conj} \tilde{G}G\tilde{D} <_{conj} \tilde{G}\tilde{G}\tilde{D}$.[2] Note that applying Sturmian involution to these morphisms in conjugation order, the images of the resulting morphisms applied to $xy \in \{x < y\}^*$ are words in lexicographic order. This relationship was proven in [16]. If we define a co-lexicographic order (right-to-left, following the order of composition of mappings), where we order $G < \tilde{G}$ and $D < \tilde{D}$, we see that this co-lexicographic order matches the conjugation order of morphisms in the class. This was proven in general for the conjugation class of a special Sturmian morphism in [20]; see also [2].

Having introduced the inner automorphisms $conj_u$ of F_2, it should be admitted that the elements of St_0 may be extended to automorphisms of F_2, by defining the actions of the gen-

[2]In Table 1 the folding words are conjugated by initial letters, closely related to but distinct from conjugation order for morphisms.

erating morphisms on the negative letters. We set $G(a^{-1}) = a^{-1}, G(b^{-1}) = b^{-1}a^{-1}, \tilde{G}(a^{-1}) = a^{-1}, \tilde{G}(b^{-1}) = a^{-1}b^{-1}$, and $D(a^{-1}) = b^{-1}a^{-1}, D(b^{-1}) = b^{-1}, \tilde{D}(a^{-1}) = a^{-1}b^{-1}, \tilde{D}(b^{-1}) = b^{-1}$. Extended in this way to F_2, the elements of St_0 are automorphisms of F_2 (see [18]). We will not need to compute with these automorphisms, but that they sit within $Aut(F_2)$ will be relevant.

IV. Commutativity and Non-Commutativity

The monoid A^* and the free group F_2 are highly non-commutative structures. In Figure 2, drawn from Figure 9 in [17], the modal scales and foldings for Lydian and Ionian are displayed on the two-dimensional lattice, $\mathbb{Z}^2$. The basis $\{(1,0),(0,1)\}$ consists of units in the *width* and *height* dimensions, where the former represents a move along the line of ascending fifths, and the latter represents an ascending generic step on the diatonic scale. Red vectors $(2,1) = a$ and $(-5,1) = b$ represent tones and semitones, respectively, in the ascending modes, while blue vectors $(4,1) = x$ and $(1,-3) = y$ represent upward perfect fifths and downward perfect fourths, respectively, in the forward folding patterns associated with the modes. In the Lydian mode figure the modal final and the initial tone of the folding are identically the note F, and are set at the origin $(0,0)$. Ionian is displayed with a choice of coordinates. On the left, the same fundamental domain as for Lydian is used, with Ionian final C at the point $(0,1)$ and the initial folding tone F at the point $(3,0)$. On the right, the coordinates for the modal final C and initial folding tone F are now $(0,-3)$ and $(-1,0)$, respectively. In any choice of coordinates, though, the vector addition is commutative; any combination of 5 vectors $(2,1)$ and 2 vectors $(-5,1)$ yields the vector $(0,7)$, representing an octave: $5(2,1) + 2(-5,1) = (0,7)$. And yet the *path* through the lattice, *aaaaabb*, has a completely different musical meaning from the Lydian mode: F-G-A-B-C♯-D♯-E-F′. Similarly, we intuitively understand that the diatonic intervals D-F and E-G are both minor thirds, but that the short musical lines D-E-F and E-F-G are different from each other. While the lattice $\mathbb{Z}^2$ is embedded in the two-dimensional real plane, the lattice paths are elements of an infinite-dimensional vector space (see [5]).

The added level of interpretation afforded by the appeal to notes that lie behind the intervallic letters in words bears fruit when we consider the notes that stand in for the divider symbol, in relation to the notes assigned to modal final and initial tone of the folding. In Figure 2, with the change of coordinates for Ionian, where the width coordinate for the final (tonic) is set at zero, and the height coordinate of the initial tone of the forward folding is set at zero, we see that the coordinates for the shared divider note G are (1,1). This property, referred to in [17] as *divider incidence*, is a general property of authentic modes associated with standard words and with the words of the corresponding forward folding (see [21]). In divider incidence, not only does the divider coincide in scale and folding, but the final essential note (leading tone) also coincides for scale and folding, and initial scale tone and folding divider predecessor coincide, as do initial folding tone and scale divider predecessor tone. Divider incidence plays a role in the word-theoretical understanding of the privileged status of Ionian (and of modes of well-formed scales associated with standard words), with such properties as the Sensitive Interval Property and Double-Neighbor Polarity (see Section 4, [17]).

The explanatory power for tonal properties is the dividend, alluded to above, paid by the forward folding interpretation. What if we assume the alternative backward folding interpretation (recalling that by Sturmian involution this is mathematically preferred)? In this instance a form of divider incidence holds for authentic Dorian. As one can read off Figure 2, the shared divider for authentic D-Dorian is A, and the intervals from modal finals to dividers C-G and D-A are both expressed as vector sums $3(2,1) + (-5,1) = (1,4)$, but as words encoding paths, *aaba* and $abaa = rev(aaba)$, respectively, and the intervals from initial folding tones to dividers F-G and

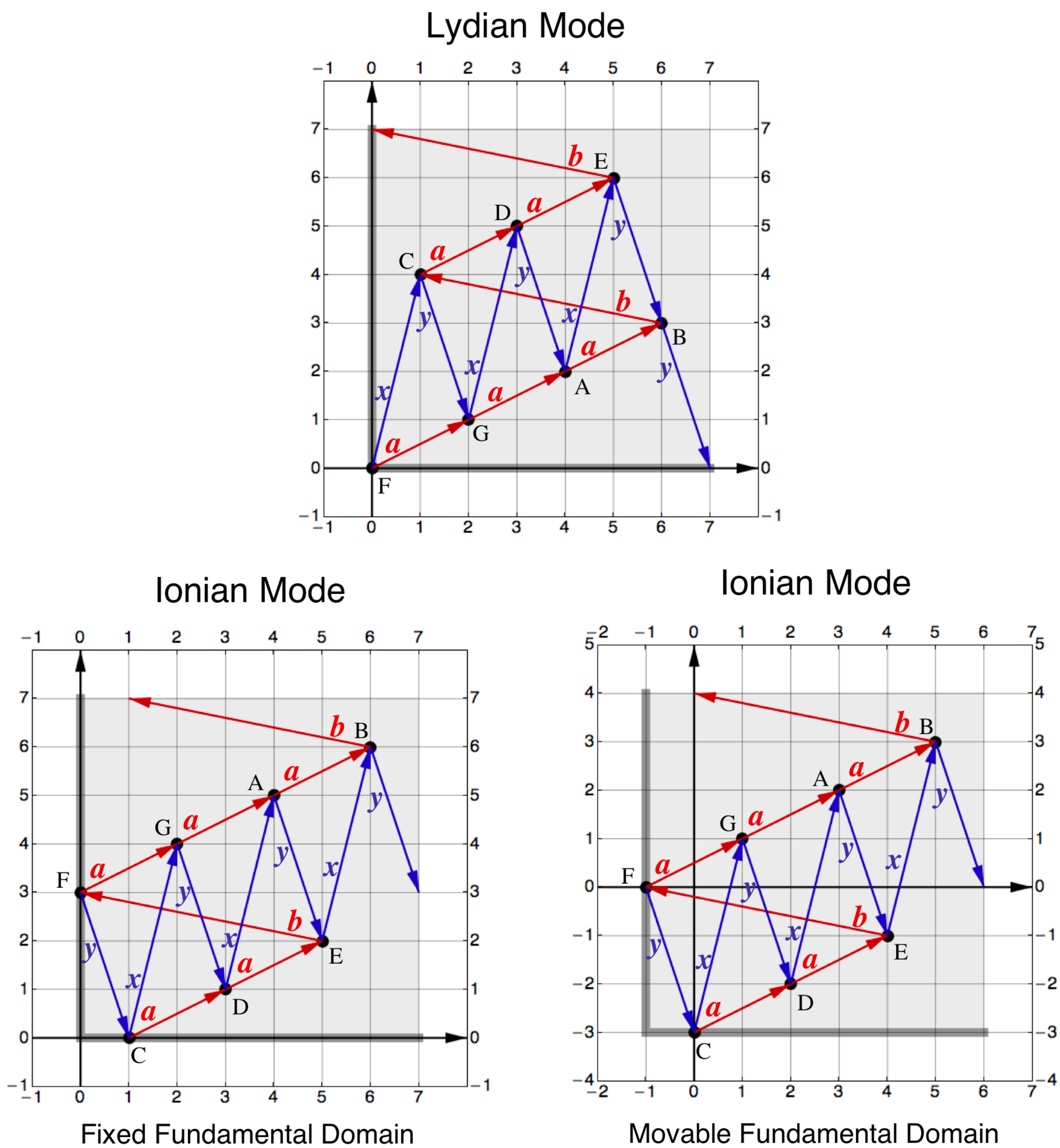

Figure 2: *Representations of Lydian and Ionian via Width-Height Vectors*

B-A are both expressed as vector sums $(1, 4) + (1, -3) = (2, 1)$, but as words encoding paths, yx and $xy = rev(yx)$, with the meanings of x and y in the latter word as descending perfect fifth and ascending perfect fourth, respectively. The music-historical meaning of the respective assignments must remain completely speculative, but it is suggestive that in the old eight church modes, authentic Dorian was mode 1, while after Glarean's 1547 expansion to the twelve *Dodecachordon* modes and Zarlino's 1571 reordering, authentic Ionian is labeled mode 1.

Returning to the free monoid $A^* = \{a, b\}^*$ of words over a two-letter alphabet, embedded in the free group $F_2 = \{a, b, a^{-1}, b^{-1}\}$, we consider relations with their commutative counterparts, the additive monoid of non-negative integer ordered pairs, $\mathbb{Z}^2_+$ and the additive group $\mathbb{Z}^2$. Consider the monoid homomorphism $V : A^* \to \mathbb{Z}^2_+ : w \to (|w|_a, |w|_b)$. $V(a) = (1, 0)$, $V(b) = (0, 1)$, which form a basis for $\mathbb{Z}^2_+$. The empty word is the kernel of V, but V is many-to-one: e.g., $V(aabb) = V(abab) = (2, 2)$. V is a projection based upon letter count, a monoid epimorphism: if $u, v \in A^*, w = uv$, then $V(w) = V(u, v) = V(u) + V(v)$, and if $z = (z_1, z_2) \in \mathbb{Z}^2_+$, any word w with $|w|_a = z_1$ and $|w|_b = z_2$ is in the inverse image of z under V.

We wish to show that V induces a linear map on $\mathbb{Z}^2_+$, such that we have a commutative square, where f is a morphism (not necessarily Sturmian) of A^*:

$$\begin{array}{ccc} A^* & \underset{f}{\rightarrow} & A^* \\ V\downarrow & & V\downarrow \\ \mathbb{Z}^2_+ & \underset{M_f}{\rightarrow} & \mathbb{Z}^2_+ \end{array}$$

where

$$M_f = \begin{pmatrix} f(a)_a & f(b)_a \\ f(a)_b & f(b)_b \end{pmatrix}.$$

That is, we wish to show that $M_f V(w) = V f(w)$, for all w in A^*. V sends ab to $(1, 1)$, while f sends ab to $f(ab) = f(a)f(b)$. Then $V(f(ab)) = V(f(a)f(b)) = V(f(a)) + V(f(b)) = (|f(a)|_a + |f(b)|_a, |f(a)|_b + |f(b)|_b) = M_f\binom{1}{1} = M_f V(ab)$.

Proposition 1 *Let $w = w_1 \ldots w_k$ where $w_j \in \{a, b\}, 1 \le j \le k$, and suppose that $|w|_a = m$ and $|w|_b = n$, and let f be a morphism of A^*, then $V(f(w)) = M_f V(w)$.*

Proof:

$f(w) = f(w_1) \ldots f(w_k)$, and $V(w) = (m, n)$. Then $V(f(w)) = V(f(w_1) \ldots f(w_k))$

$= (m|f(a)|_a + n|f(b)|_a, m|f(a)|_b + n|f(b)|_b) = M_f\binom{m}{n} = M_f V(w)$.

Exemplifying the commutative diagram, beginning with the root word ab and applying the standard morphism GGD we have $GGD(a, b) = aaba, aab$, and we observe that

$$\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

that is, $M_{GGD}(V(ab)) = V(GGD(ab))$.

In the following, f is assumed to be a member of St_0, generated by $\{G, \tilde{G}, D, \tilde{D}\}$. For G and $\tilde{G}$, we have $G(a,b) = a, ab$ and $\tilde{G}(a,b) = a, ba$, and for D and $\tilde{D}$, we have $D(a,b) = ba, b$ and $\tilde{D}(a,b) = ab, b$, so

$$M_G = M_{\tilde{G}} = R = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

and

$$M_D = M_{\tilde{D}} = L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

R and L freely generate the monoid $SL_2(\mathbb{N})$ (by which is meant the intersection of 2x2 matrices with non-negative entries with $SL_2(\mathbb{Z})$).

Recall that the inner automorphisms of F_2, $Inn(F_2) = \{conj_u | u \in F_2\}$ form a subgroup of the automorphisms of F_2, $Aut(F_2)$. From group theory, the inner automorphisms are a normal subgroup of the group of automorphisms, that is, they form the kernel of a group epimorphism. We appeal to the celebrated result of Nielsen (see chapter 5 of [18]; for a proof see [22]) that the quotient of $Aut(F_2)$ modulo $Inn(F_2)$ is isomorphic to the automorphism group $Aut(\mathbb{Z}^2) = GL_2(\mathbb{Z})$. Thus, $Aut(F_2)/Inn(F_2) = GL_2(\mathbb{Z})$, and the kernel of an epimorphism onto the general linear group is the group of conjugations. This in turn implies that the conjugation class of a special Sturmian morphism f is characterized by the result that all its members share the same incidence matrix M_f.

It follows that all the representatives of the conjugation class of $f \in St_0$ are products of the same sequence of letters G and D, to within the distribution of diacritic ˜ marks attached to these letters. Suppose that the conjugation class of f is characterized by the sequence of letters (morphisms) G and D, $X_1 X_2 \ldots X_n$, $X_i \in \{G, D\}, 1 \leq i \leq n$. Then the incidence matrix $M_f = M_{X_1} M_{X_2} \ldots M_{X_n}$. This matrix product is a product of generating matrices R and L, where $M_{X_i} = R$ or L as $X_i = G$ or D. For example, the conjugation class of the St_0 elements that yield the authentic diatonic modes consists of $\{GGD, \tilde{G}GD, \tilde{G}\tilde{G}D, GG\tilde{D}, \tilde{G}G\tilde{D}, \tilde{G}\tilde{G}\tilde{D}\}$. The associated incidence matrix is then $M_{GGD} = M_G M_G M_D = RRL$, i.e.,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}.$$

V. Concluding Remark

Musicians are used to adding intervals together commutatively to produce intervals, when only the size of the resulting interval is at issue. But more often we are interested in the process of arriving at the result. Even in the traditional one-dimensional pitch height conception of musical intervals this is the case. For a simple example, the perfect fifth followed by perfect fourth resulting in an octave is a very different process from perfect fourth followed by perfect fifth. In semitones, we are simply saying $7 + 5 = 12 = 5 + 7$; the endpoints of the musical line remain unchanged, but the path is different, C-G-C′ vs. C-F-C′. Consider a much more complicated example, the opening ascent in the subject of J. S. Bach's Fugue in F♯ minor from Book 1 of the WTC. The line is essentially 3 times a motive of tone-semitone, each time resetting one semitone lower: F♯-G♯-A, G♯-A♯-B, A♯ (…)-B♯-C♯, or, in semitones, $(2+1) - 1 + (2+1) - 1 + (2+1) = 7$. The goal of the line is the perfect fifth, C♯, 7 semitones above the tonic, F♯, but achieved in a very slow and highly chromatic ascent, and in a systematic, motivically oriented construction. And this is just with respect to the one-dimensional perspective. If the intersection of scale theory with word theory tells us anything it is the value of a two-dimensional framework. As Figure 2 suggested, both the commutative vector sums on the two-dimensional lattice and the paths traced in the course of these sums, are musically compelling. The points on the lattice are the musical notes that the mathematical words have "forgotten," but that breathe musical life to the study. The mathematics outlined above relating non-commutative algebraic objects with their commutative images suggests further interpretations in the study of scales, modes, and transformations among them.

References

[1] Brlek, S., Chemillier, M., Reutenauer, C. (2018). Music and combinatorics on words: A historical survey. *Journal of Mathematics and Music*, Vol. 12, No. 3, pp. 125–133.

[2] Clampitt, D., Noll, T. (2018). Naming and ordering the modes, in light of combinatorics on words. *Journal of Mathematics and Music*, vol. 12, No. 3, pp. 134–153.

[3] Carey, N., Clampitt, D. (2018). Variety and multiplicity for partitioned factors in Christoffel and Sturmian words. *Journal of Mathematics and Music*, Vol. 12, No. 3, pp. 154–170.

[4] Noll, T., Clampitt, D. (2018). Kaleidoscope substitutions and pairwise well-formed modes: major-minor duality transformationally revisited. *Journal of Mathematics and Music*, Vol. 12, No. 3, pp. 171–211.

[5] Noll, T. (2018). Dual lattice-path transformations and the dynamics of the major and minor exo-modes. *Journal of Mathematics and Music*, Vol. 12, No. 3, pp. 212–232.

[6] Lothaire, M. (1997). *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge: Cambridge University Press (First edition 1983).

[7] Lothaire, M. (2002). *Algebraic Combinatorics on Words*. Encyclopedia of Mathematics and Its Applications, Vol. 90. Cambridge: Cambridge University Press.

[8] Lothaire, M. (2005). *Applied Combinatorics on Words*. Encyclopedia of Mathematics and Its Applications, Vol. 105. Cambridge: Cambridge University Press.

[9] Clough, J., Myerson, G. (1985). Variety and multiplicity in diatonic systems. *Journal of Music Theory*, Vol. 29, No. 2, pp. 249–270.

[10] Carey, N., Clampitt, D. (1989). Aspects of well-formed scales. *Music Theory Spectrum*, Vol. 11, No. 2, pp. 189–207.

[11] Carey, N., Clampitt, D. (2012). Two theorems concerning rational approximations. *Journal of Mathematics and Music*, Vol. 6, No. 1, pp. 61–66.

[12] Carey, N., Clampitt, D. (2017). Addendum to 'Two theorems concerning rational approximations'. *Journal of Mathematics and Music*, Vol. 11, No. 1, pp. 61–63.

[13] Carey, N., Clampitt, D. (1996a). Regions: A theory of tonal spaces in early medieval treatises. *Journal of Music Theory*, Vol. 40, No. 1, pp. 113–147.

[14] Carey, N., Clampitt, D. (1996b). Self-similar pitch structures, their duals, and rhythmic analogues. *Perspectives of New Music*, Vol. 34, No. 2, pp. 62–87.

[15] Kassel, C., Reutenauer, C. (2007). Sturmian morphisms, the braid group B_4, Christoffel words and bases of F_2. *Annali di Matematica Pura ed Applicata*, Vol. 186, No. 2, pp. 317–339.

[16] Berthé, V., de Luca, A., Reutenauer. C. (2008). On an involution of Christoffel words and Sturmian morphisms. *European Journal of Combinatorics*, Vol. 29, No. 2, pp. 535–553.

[17] Clampitt, D., Noll, T. (2011). Modes, the height-width duality, and Handschin's tone character. *Music Theory Online*, Vol. 17, No. 1.

[18] Berstel, J., Lauve, A., Reutenauer, C., Saliola, F. (2009). *Combinatorics on Words: Christoffel Words and Repetitions on Words*. Providence, R.I.: American Mathematical Society.

[19] Séébold, P. (1998). On the conjugation of standard morphisms. *Theoretical Computer Science*, Vol. 195, No. 1, pp. 91–109.

[20] Clampitt, D., Noll, T. (2017). Matching lexicographic and conjugation orders on the conjugation class of a special Sturmian morphism. In *Proceedings of the Combinatorics on Words Eleventh International Conference, WORDS 2017, Montréal QC, Canada*, pp. 85–96. Brlek, S., Dolce, F., Reutenauer, C., Vandomme, E., eds. Springer.

[21] Noll, T. (2009). Ionian theorem. *Journal of Mathematics and Music*, Vol. 3, No. 3, pp. 137–151.

[22] Lyndon, R. C., Schupp, P. E. (2001). *Combinatorial Group Theory*. Berlin: Springer.

An Introduction to Markov Chains in Music Composition and Analysis

Hugo Tremonte de Carvalho
Federal University of Rio de Janeiro (UFRJ)
hugo@dme.ufrj.br
Orcid: 0000-0003-0776-0400

Abstract: *Since the second half of the 20th. century, the use of probabilistic structures to compose and analyze music became even more popular, and nowadays, with the widespread and easiness of use of high-level structures such as neural networks due to very efficient and intuitive computational packages, the area of algorithmic composition became even more popular. This text is an expansion of a lecture named "A brief overview of algorithmic composition", presented at the Music School of the Federal University of Rio de Janeiro in October 2019, and its main goal is to present a gentle and music-oriented introduction to Markov chains, a very intuitive and interpretable tool widely used in algorithmic composition.*

Keywords: *Markov chain, algorithmic composition, probability, statistics*

I. Introduction

Our lives are surrounded by randomness, and at each moment we face decisions that must be made without knowing the exact outcome. In a macroscopic scale, we usually resort to randomness to model extremely complex phenomena where it is infeasible to have all the necessary information required to predict the observed result; for example, the number observed when rolling a die is anything but random, since it is subjected to Newton's laws of motion, a set of deterministic equations. However, since this system is very sensitive to the initial conditions (i.e., initial position and velocity of the die) as well as a very accurate modeling of the die's geometry, the corresponding table where it will fall, atmospheric conditions, among others, it is much easier to simply model the observed value as purely random. On the other hand, the microscopical quantum world provides us real randomness, as it can be seen, for example, in Heisenberg's uncertainty principle, which says, intuitively, that one cannot measure accurately both the position and momentum of a particle at the same time. Despite this formulation it seems to be only a fancy way of translating one's ignorance with respect to some quantities of interest. It can be proven that it "actually states a fundamental property of quantum systems and is not a statement about the observational success of current technology" [4].

Focusing on the musical aspect, the role of the composer is essentially to organize musical material in order to make a pleasing sound, evoke some emotion on the listener or simply innovate artistically, giving form and substance to an idea that lives on his mind. Naively, one may think that this process is as easy to perform as it was easy to state it, but as it is common in Mathematics, some of the hardest problems to solve are the easiest to formulate[1]. However, as previously

[1] Recall Fermat's Last Theorem, that was an open problem during 358 years and can be stated as: no three positive integers a, b, and c satisfy the equation $a^n + b^n = c^n$ for any integer value of $n > 2$. Fermat himself claimed that he had a

exemplified, complex procedures could be modeled by stochastic[2] phenomena. Indeed, given a fixed style of composition, there is some "common sense" in the transition of one musical material to another in order to maintain some degree of musical coherence, and this coherence could be translated into probabilistic laws. The knowledge of these laws, which can be stated beforehand or inferred from a *corpus*, could be used to create new sounds or extract information about the respective *corpus* itself.

This text is an expansion of a lecture named "A brief overview of algorithmic composition", presented at the Music School of the Federal University of Rio de Janeiro in October 2019 for an audience composed mainly of musicians. The main goal in writing this note is to present some basic aspects of Probability theory in a mostly intuitive fashion, as well as showing its capabilities in producing new musical pieces and analyzing musical *corpora*, in order to stimulate future investigation in this direction and encouraging researchers to go more deeply into the Mathematics of modern methods not covered in this text (e.g., neural networks [5] and hidden Markov models [29]). This exposition is far from being exhaustive, and several references are given along the text in order to direct the reader to more specific topics of their interest. Since I am a mathematician and only an amateur musician, I apologize in advance if someone is offended by some naive musical analogy. I promise I will do my best!

The text is organized as follows: in Section II we introduce some aspects of Probability theory, its history, and Markov chains in an intuitive way followed by a more technical discussion on the same topics in Sections III and IV, respectively; Section V reviews some applications of Probability and Markov chains in music composition, including a discussion about the probabilistic structures underlying the piece *Analogique A* from Iannis Xenakis; in Section VI three other musical examples are discussed, namely an excerpt of *Brazilian Landscapes No. 20* for bassoon and string quartet by Liduino Pitombeira and inspired by the Binomial distribution, and two excerpts generated from a Markov chain inferred from some of the chorales from J. S. Bach; some words on the relationship between interpretability and flexibility of statistical methods are presented in Section VII, and conclusions are drawn in Section VIII.

II. Elements of Probability theory and Markov chains

In this section we will introduce some basic concepts of Probability theory and Markov chains, one of the first probabilistic model used in algorithmic composition (see [2] for an extensive review). To strictly follow the historical path is usually not the best way to learn Mathematics, but on the other hand, the history of Probability theory is very rich and interesting, and also deeply linked with our habit of using gambling scenarios to gain intuition (or learn that our prior intuition was wrong!) about its fundamental concepts. Therefore, our presentation will be an interchange of historical and more technical information.

Since our scientific history is largely influenced by the development of western civilization, I will mainly refer to developments in Europe and America before the 16th century, even though it is known that advances in Probability and Statistics also occurred in Chinese, Indian, Arabian, Egyptian civilizations, sometimes much earlier than in Europe. For more details, see [6, 25, 13, 26], being [25] also available in Portuguese.

proof of this result in 1637, but wrote in his copy of *Arithmetica*, an ancient text on Mathematics by Diophantus, that it was too large to fit the margin. Nowadays, it is believed that whatever was his proof it was wrong, since the result was only proven in 1994 by Andrew Wiles using very sophisticated mathematical techniques. It is widely believed that it is impossible to prove this result without these tools.

[2]This term comes from the greek $\sigma\tau o\chi o\zeta$, (*stókhos*), which means "aim, guess".

i. Early days of Probability

Our history begins long ago, in the 16th century with Gerolamo Cardano, an Italian polymath and gambler. Motivated by his addiction and lowering funds at the end of his life, he investigated probabilities associated with simple dice rolling games and wrote in 1564 the *Liber de ludo aleae* ("Book on Games of Chance"), which contains the first known systematic treatment of probability. The text was posthumously published in 1663.

Some years later, Blaise Pascal, a french mathematician, was asked to solve a problem, that was noted to be very difficult to solve at that time: what is the fairest way of dividing the stake if a game of chance is interrupted? Pascal noted that the tools necessary to answer this question were not available, and started a sequence of correspondences with Pierre de Fermat, a french lawyer and amateur mathematician.

It is important to note that these early developments were made under the hypothesis that some fundamental probabilities where known *a priori*. For example, Cardano was assuming that the dice used during the game were fair, and the question posed to Pascal assumed the hypothesis that at each round of the game each player was equally likely to win. More generally, as stated in [27], in this scenario we are aware of the mechanisms of the data generating process and we wish to forecast information about the observed data. However, in some problems the real probabilities or other quantities of interest are not known beforehand and must be inferred from observed data: that is, given some observed data we wish to infer information about the data generating process, and this is what we call *Statistics*. A deeper study of Statistics only became possible some years later, after the development of the laws of the large numbers by J. Bernoulli, which will be presented in subsection iii of this Section.

ii. Interlude: A little bit of technicalities

The formal treatment of Probability theory as we know today was developed much later in 1933 by the Russian mathematician Andrey Kolmogorov. However, it is convenient to introduce some terminology and concepts in order to follow more easily the forthcoming discussions [21] (also available in Portuguese).

A *random variable* X is a numerical outcome of some random experiment. In other words, it is a numerical variable whose value depends on a random phenomena. For example, the number observed when rolling a fair die is a random variable that can assume each value in the set $\mathcal{C} = \{1, 2, \ldots, 6\}$ with probability equal to $1/6$, that is, $\mathbb{P}(X = x) = 1/6$, for $x \in \mathcal{C}$.

Formally, the probability of any event is a number between 0 and 1, and this protocol will be followed in this text, except in Section V, where some probabilities will be measured in % to avoid inappropriate rounding to zero when dealing with small numbers. Each time a probability is measured in %, we will be explicit.

Example 1 *One could use a very simple probabilistic model to generate a music. Despite this example being not musically interesting, I will use it later to introduce other probabilistic concepts. Suppose one wishes to compose an infinite-length monophonic piece for piano only with the notes C, C♯, D, D♯ and E*[3]*, chosen at random with equal probability at each time instant*[4]*. If a listener is aware of the procedure employed by the composer, it is easy to see that the probability of a given note be played in a black key is* $2/5$.

[3]The specific pitch of the note is irrelevant here, since only the color of the keys being played are considered.

[4]This restriction is not necessary: the musician can choose the note n_i with some probability $0 \leq p_i \leq 1$, with the single restriction that $p_1 + \cdots + p_5 = 1$. This probability assignment is called the *distribution* of the random variable. However, we will keep the equiprobable scenario for simplicity. Note that in this particular case the random variables are also *identically distributed*, since all of them have the same probability assignment.

Mathematically, one can think that this music is a realization of a *stochastic process*, a sequence $(X_t)_{t\in\mathcal{T}}$ of random variables, where $\mathcal{T}$ contains the onset time of each note, arbitrarily chosen by the composer in principle. This structure is very general and important in Mathematics, but we will restrict ourselves to this particular case, in order to gain more intuition. Each of these random variables X_t assume values on the set $\mathcal{C} = \{\text{C}, \text{C}\sharp, \text{D}, \text{D}\sharp, \text{E}\}$[5], and each note is assumed with equal probability $1/5$. Moreover, *notes played at distinct times are independent*. This is defined as shown in Equation 1:

$$\mathbb{P}(X_{t_1} = n_{t_1}, \dots, X_{t_k} = n_{t_k}) = \mathbb{P}(X_{t_1} = n_{t_1}) \dots \mathbb{P}(X_{t_k} = n_{t_k}), \tag{1}$$

and this equality should hold for all $k \geq 2$ and $n_{t_1}, \dots, n_{t_k} \in \mathcal{C}$, if $t_1 < t_2 < \cdots < t_k$. That is, the probability of playing k specific notes $n_{t_1}, \dots, n_{t_k}$ at specific time instants $t_1 < \cdots < t_k$ is the product of the individual probabilities, for all choices of notes, time instants and quantities of notes. However, one does not need to be a musician to know that the vast majority of musics are not composed with total randomness like this! Let us go back to history in order to introduce more structure in our model.

iii. Calculus, Probability and the laws of large numbers

The formal development of calculus in the late 17th century by Issac Newton and Gottfried Leibniz[6] allowed immense advances in several areas of science, being Physics perhaps the most notable one, mainly because of its intimacy with the ideas developed by Newton and Leibniz themselves. Obviously, Probability theory also took advantage of this new tool, mainly in two works: "The doctrine of chances: or, a method for calculating the probabilities of events in play" published in 1718 by the french mathematician Abraham de Moivre and *Ars Conjectandi* ("The Art of Conjecturing") by Jacob Bernoulli, posthumously published in 1713.

The work of J. Bernoulli contains the first step towards one of the greatest achievements of Probability theory, the *laws of the large numbers*, a result for which he devoted about 20 years of his life. Recall the composition in Example 1, but suppose now that a listener is not aware of the procedure employed by the composer and wishes to *infer* the probability of a given note being played in a black key. Still assuming that different notes are independent, a very careful listener can listen the piece for an arbitrarily long time and at each note write down if it comes from a black or white key. Mathematically, the listener is observing a realization of *another* stochastic process $(Y_t)_{t\in\mathcal{T}}$, where each random variable Y_t assumes values in the set $\mathcal{D} = \{\text{white}, \text{black}\}$ and are also independent. However, he does not know the probability of observing black or white outcomes. Intuitively, he can simply count the occurrences of black keys and divide it by the total of notes being played and hope that this number will be close to the true proportion of $2/5$. This is a prototypical example of an *statistical inference* procedure, a scenario where one is interested in estimating quantities from observed phenomena instead of studying the possible outcomes of random variables knowing its distribution *a priori*. More generally, as stated in [27], now one is interested in in inferring information about the data generating process from observed data.

It is quite intuitive that if more notes are being played, more accurate is the listener's estimate. Indeed, as J. Bernoulli itself stated in his book in a quite presumptuous manner: "[even] the most stupid of men [...] is convinced that the more observations have been made, the less danger there

[5]Note that this does not contradict our definition of a random variable assuming only numerical values, since one can easily map this set to a numerical one, for example, its respective MIDI numbers. For a moment, we will make this abuse of notation for the sake of clarity.

[6]"Standing on the shoulders of giants", as said by Newton himself, since ideas of limits and integrals existed since the Greek mathematicians.

is of wandering from one's aim" [17] (also available in Portuguese). Again, this is an example of a mathematical result that is easy to formulate, at least intuitively, but quite hard to prove.

J. Bernoulli called this result the *law of the large numbers* and nowadays this name is given to an entire class of results about asymptotic behavior of sequences of random variables. We will not dive too much into this topic, but it is important to note two forms of the law, which will be stated in the scenario of our example:

- Weak law of large numbers: with a sufficient large sample of observed notes, the proportion computed by the listener is very likely to be close to the true value or $2/5$;
- Strong law of large numbers: the probability of observing an infinite song that leads to an incorrect estimation of the true value of $2/5$ is null.

On his work, J. Bernoulli only have proved the weak law of large numbers for a specific kind of random variable, and most important for us, *under the hypothesis of independence*, crucial to its proof. Nowadays, it is known that a very large class of random variables satisfy the laws, but the majority of them also assuming independence, being the Markov chains the first class which breaks this hypothesis [24].

The history of the laws of the large numbers is very rich and intimately linked to the development of the formal theory of Probability in the beginning of the 20th. century. For more details see [26, 17].

iv. Markov chains

In order to introduce the specific kind of dependence between random variables in a Markov chain, let us change somewhat the composition in Example 1.

Example 2 *Let us go back again to the composition we made in Example 1 and try to improve it imposing more structure. Instead of choosing randomly a note from the set $\mathcal{C}$, our notes now could be taken from the sets $\mathcal{C}_1 = \{G, G\sharp\}$ and $\mathcal{C}_2 = \{C, C\sharp, D, D\sharp, E\}$, using the following rules:*

1) *The first note n_1 is chosen at random from set $\mathcal{C}_2$*
2) *For $k = 2, \ldots, N$, where N is the number of notes the composer desires in his song:*
    - i) *If the previous note n_{k-1} comes from a black key, the note n_k is chosen at random from the set $\mathcal{C}_1$;*
    - ii) *Else, if the previous note n_{k-1} comes from a white key, the note n_k is chosen at random from the set $\mathcal{C}_2$.*

Assuming the same onset times as before, denote this song as a realization of another stochastic process $(X'_t)_{t\in\mathcal{T}}$. From the procedure above, it is clear that these random variables are not independent anymore, since a note depends on a feature of the previous one.

But assume that the listener also wishes to estimate the probability of a given note be played in a black key. He could repeat the same procedure as before, listening an arbitrarily long excerpt of the music, writing down if each note being played is black or white and computing the proportion of black keys being played. However, since the random variables being observed are not independent anymore, there is no guarantee that the listener's procedure will be close to the true proportion being estimated, for the result of J. Bernoulli needed the independence assumption.

In 1902, the Russian theologian and mathematician Pavel Nekrasov claimed that independence was a *necessary* condition for the laws of the large numbers to hold, that is, there is no possibility that our listener's estimate will be accurate in this new scenario. Being his work grounded not on

Mathematics itself but only in religious principles of predestination and free will, there was a lot of margin to more rigorous discussion on his claims.

Indeed, motivated by its mathematical accuracy and personal disputes with Nekrasov, the Russian mathematician Andrey Markov initiated a detailed study on dependent sequences of random variables, aiming the extension of J. Bernoulli's work to this new scenario. In 1906 he published his work which title can be loosely translated as "Extension of the law of large numbers to quantities that depend on each other", containing the beginning of an entire new theory in the field of Probability theory and proving that under some hypothesis, sequences of dependent random variables *can* satisfy the laws of the large numbers [24] . It is curious to note that even though Markov created its chains merely as a counterexample to Nekrasov's claim, its importance in applied fields is enormous [22].

Intuitively, Markov chains are stochastic processes such that at a particular time instant t, the random variable X_t is purely determined by its immediate preceding observed value plus a random effect, independent X_t. This is exactly the scenario our listener is faced with: in order to determine a given note, it is necessary only to know if the previous note came from set $\mathcal{C}_1$ or $\mathcal{C}_2$, plus a random choice which only depends on the set. We devote now some time formalizing some important concepts used from here on.

III. Some technical aspects of Probability theory

Recall that a random variable X was defined as a numerical outcome of some random experiment and probabilities associated with X are denoted using the letter $\mathbb{P}$. A particular assignment of probabilities to subsets of real numbers *via* X is called the *distribution* of the random variable. Some distributions of random variables are remarkably important and appears in many distinct and apparently uncorrelated scenarios that special names are given to them. We will now clarify this definition on random variables and present some examples.

Firstly, note that we can classify a random variable in two main classes: *discrete* or *continuous*. Discrete random variables assumes its values in finite or countable sets[7] of real numbers, and we can loosely say that continuous random variables assumes their values in uncountable sets. Despite this not being the formal definition of continuous random variables, the precise definition is quite involved and we will keep this intuition for a moment.

i. Discrete random variables

Discrete random variables are simplest than continuous one, since its manipulation requires only, in general, the basic arithmetic operations.

Example 3 *Some examples of important discrete distributions are:*

- <u>*Bernoulli distribution*</u>*: This is the simplest distribution of a random variable, and receives the name of J. Bernoulli since it was very important on his studies in the laws of the large numbers. We say that X is a Bernoulli random variable with parameter p if it assumes only the values 0 or 1, with respective probabilities $1-p$ and p, that is,*

$$\mathbb{P}(X=0) = 1-p \tag{2}$$

$$\mathbb{P}(X=1) = p. \tag{3}$$

[7] A set A of real numbers is said to be *countable* if there is an one-to-one correspondence between A and $\mathbb{N}$, the set natural numbers. Intuitively, A is countable if its elements can be counted. The sets of integer numbers $\mathbb{Z}$ and rational numbers $\mathbb{Q}$ are examples of countable sets, whereas the interval $[0,1]$, the set of irrational numbers $\mathbb{R}\setminus\mathbb{Q}$ and the set of real numbers itself $\mathbb{R}$ are examples of non-countable sets.

It is used to model the outcome of binary experiments, and the values 0 and 1 are usually denoted as "failure" and "success", respectively. The sentence "X has a Bernoulli distribution with parameter p" is abbreviated as $X \sim Bern(p)$.
Note that we already came across a Bernoulli random variable: recall Example 1 where each random variable Y_t indicates if the note X_t is played in a black key or not. Assuming that observing a black key is a success, we can state now, more formally, that $Y_t \sim Bern(2/5)$.

- *Binomial distribution: It is a generalization of the Bernoulli distribution, where a sequence of independent trials of a binary experiment with a probability of success equal to p is repeated n times and the quantity of successes are computed. Its probability function is given by Equation 4:*

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, \ldots, n, \tag{4}$$

where $\binom{n}{k}$ is given by $\frac{n!}{k!(n-k)!}$. Intuitively, this formula says that $p^x(1-p)^{n-x}$ is the probability of observing a particular sequence of x successes and $n-x$ failures, and the term $\binom{n}{k}$ accounts for the distinct ways we can organize these successes and failures. Finally, since n repetitions of the experiment are performed, one can obviously observe only 0 to n successes. We denote then $X \sim Bin(n, p)$.

- *Poisson distribution: Named after the French mathematician Siméon Poisson, it models the number of events of interest occurring in a fixed interval of time or space, assuming that these events occur with a known constant rate and are independent of the time since the last observation. Its probability function is given by Equation 5:*

$$\mathbb{P}(X = x) = \frac{\lambda^x e^{-x}}{x!}, \text{ for } x = 0, 1, 2, \ldots, \tag{5}$$

where e is the Euler number[8]. The parameter λ is the average numbers of occurrence of the event of interest in a fixed interval of time. We abbreviate $X \sim Poisson(\lambda)$.
This distribution appears quite naturally when one tries, for example, to model the telephone calls arriving in a system per time unit, the number of mutations on a strand of DNA per unit length, number of decays in a given time interval in a radioactive sample, among others.

More generally, a discrete random variable X assumes its values on a finite or countable set, here denoted by $\{x_1, x_2, \ldots\}$. The *distribution* of X is the particular assignment of probabilities of X to these numbers, that is, the values $\mathbb{P}(X = x_i) = p_i$, with the single restrictions that $p_i \geq 0$ for all $i = 1, 2, \ldots$ and $\sum_{i=1}^{\infty} p_i = 1$. The function $f_X(x) = \mathbb{P}(X = x)$ is called the *probability function* of X.

ii. Continuous random variables

A continuous random variable X, on the other hand, posses a *probability density function*, that is, a function $f_X : \mathbb{R} \to \mathbb{R}$ such that $f_X(x) \geq 0$ for all $x \in \mathbb{R}$ and satisfying the Equation 6:

$$\int_{-\infty}^{+\infty} f_X(x)\, dx = 1. \tag{6}$$

Therefore, the value $f_X(x)$ does not represent the probability of observing X assume the value x, but the *density* of the probability on the value x, and the probability of observing X between

[8] The appearance of the Euler number in this probability function is not obscure, since it can be proven that the Poisson distribution is approximately the same as the Binomial distribution with n large and small p, with $\lambda = np$, and it appears quite naturally in the derivation. For more details, see [21].

values a and b is given by the area below the graph of f_X, from a to b (Equation 7):

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x)\,dx. \tag{7}$$

It is important to note now that individual values have null probability of occurrence: that is, $\mathbb{P}(X = a) = \int_a^a f_X(x)\,dx = 0$, for all $a \in \mathbb{R}$. This is an apparent paradox, since we *can* observe $X = \pi$, for example, but this event has null probability of occurrence! It can be easily solved once we recall that a real number possess *infinite* decimal places, and any measurement device we have invented have a finite precision, that is, detects only a finite amount of decimal places: even if $X = \pi$ we may only observe its rounding on the first two decimal places $X \approx 3,14$, that is the same as saying $3,135 < X < 3,144$, an event with positive probability of occurrence.

Contrarily to discrete random variables, operating with continuous random variables it more involved, since it requires tools from Calculus, such as derivation and integration.

Example 4 *Some important examples of continuous random variables are:*

- *Normal distribution: Perhaps the most important probability distribution of all! Its first appearance was in the work of Abraham de Moivre, but it gained more visibility after Carl Gauss' work in 1809 about the movement of the planets around the Sun. He wished to estimate the orbit of Ceres, a dwarf planet between Mars and Jupiter, from a few observations and needed a way to model inaccuracies on its observations. He claimed that the "fairest" way of modeling errors in experiments is via the Normal distribution, whose probability density function is given by Equation 8:*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ for all } x \in \mathbb{R}. \tag{8}$$

  The parameters μ and σ^2 are called its mean and variance, respectively, and controls where the distribution is centered and its spreadness, respectively. We denote normality as $X \sim \mathcal{N}(\mu, \sigma^2)$. For more details about the history of the Normal distribution see [26].
- *Exponential distribution: The Exponential distribution is intimately related with the Poisson distribution we saw before. Consider the particular example when one uses the Poisson distribution with parameter λ to model the number of decays in a given time interval in a radioactive sample, and supposes that one is interested in the distribution of the time interval between consecutive radioactive emissions. Obviously this random variable must be continuous, and one can prove [21] that it has the Exponential distribution, whose probability density function is given by Equation 9:*

$$f_X(x) = \lambda e^{-\lambda x}, \text{ for all } x > 0, \tag{9}$$

  this fact being denoted as $X \sim Exp(\lambda)$.
- *Maxwell-Boltzmann distribution: This is a very important probability distribution in Statistical Mechanics, which describes the distribution of speeds of molecules from a gas at a certain temperature. Its probability density function is given by Equation 10:*

$$f_X(x) = \sqrt{\frac{2}{\pi}} \frac{x^2}{a^3} e^{-\frac{x^2}{2a^2}}, \text{ for all } x > 0. \tag{10}$$

iii. Expected values

The distribution of a random variable X provides us all the probabilistic information we may need. However, sometimes all we need is some *numerical summary* of it. For example, if X models the number of radioactive particles emitted by some source in an interval of one hour, it is reasonable to model it *via* a Poisson distribution. Recall that parameter λ refers to the rate of emission per unit of time (one hour, in this example). The whole information about the random variable allows us to compute, for instance, $\mathbb{P}(X > x_0)$, assuming that the value x_0 is important for heath security reasons. But recalling the interpretation of parameter λ, instead of informing the whole probability distribution of X, it may be enough to compare the value of λ with x_0.

Indeed, our former character Blaise Pascal noted the importance of single numerical summaries of random variables in November 23rd. of 1654, when he wrote a small note that he kept in his pocket for the last 8 years of life describing how "God came to him and set him free from the corrupted ways" [17]. When considering the pros and cons of his duties with God, he created a way of computing these quantities, which is now called the *expectation*, *expected value* or *mean* of a random variable. These names are used interchangeably in the literature, but here we will use only *expected value*. We define this quantity and many others only for discrete random variables, for simplicity. The interested reader may refer to [21] for more details.

Indeed, if X is a discrete random variable, its expected value is defined as shown in Equation 11:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i), \tag{11}$$

assuming that this summation is finite. The intuition behind this formula is quite simple: The value x_i is weighted by its probability of observance, and the final result is an weighted average of the values X can assume, interpreted as its "central value", again making an analogy with Physics, more specifically, the concept of center of mass.

Example 5 *Let us verify that some intuitive facts about expected values indeed hold:*

- *One can prove that if $X \sim Poi(\lambda)$, then $\mathbb{E}[X] = \lambda$, which formalize our claims on the example in the beginning of this Section.*
- *Recall Example 1. Let X denote the amount of black keys being played in n notes of our composition. Since the probability of a single note be played in a black key is $2/5$, we have that $X \sim Bin(n, 2/5)$, and one expects that approximately $2n/5$ notes played in black keys will be observed. Indeed, it can be show that if $X \sim Bin(n, p)$, then $\mathbb{E}[X] = np$.*

However, one does not always want to compute the expected value of X, but of some function $g(X)$ of X. For example, if X is a random variable denoting the radius of a circle, its area will also be a random variable, given by $g(X) = \pi X^2$. The *law of the unconscious statistician* [21] allows us to compute such expectations, without knowing the distribution of $g(X)$, which can be quite involved of obtaining, *via* the simple formula (Equation 12):

$$\mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(x_i) \mathbb{P}(X = x_i). \tag{12}$$

This Theorem allows us to define another important summaries of random variables, like the *variance*, denoted by $\mathbb{V}(X)$, which measures the spreadness of a random variable around its expected value (Equation 13):

$$\mathbb{V}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{i=1}^{\infty} (x_i - \mathbb{E}[X])^2 \mathbb{P}(X = x_i). \tag{13}$$

iv. Conditional probability

A necessary concept to properly define the Markov chains is the *conditional probability*, which essentially measures the probability of some event given that another event has occurred. More specifically, the probability of occurrence of the event A given that event B has occurred is called the *conditional probability of A given B*, denoted by $\mathbb{P}(A|B)$ and computed by Equation 14:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(B)}. \tag{14}$$

This definition can be understood essentially as the restriction of the possible outcomes of the experiment, given that event B has occurred.

In order to clarify this concept, let us see it in the formerly presented examples.

Example 6 *In Example 1, since the notes being played are independent, we have that*

$$\begin{aligned}
\mathbb{P}(X_t = C\sharp | X_{t-1} = E) &= \frac{\mathbb{P}(X_t = C\sharp, X_{t-1} = E)}{\mathbb{P}(X_{t-1} = E)} && (15)\\
&= \frac{\mathbb{P}(X_t = C\sharp)\mathbb{P}(X_{t-1} = E)}{\mathbb{P}(X_{t-1} = E)} && (16)\\
&= \mathbb{P}(X_t = C\sharp) && (17)\\
&= 1/5, && (18)
\end{aligned}$$

that is, the knowledge of the previous note being an E does not change the probability of the next note be a C♯. Note that the particular choice of C♯ and E is not relevant and the same result will be obtained for any two notes in set $\mathcal{C}$.

However, in Example 2, the quantity $\mathbb{P}(X_t = C\sharp | X_{t-1} = x)$ depends on the particular choice of note x: if $x = E$ this probability is $1/5$, since E is played in a white key and this implies that the note X_t is chosen from set $\mathcal{C}_2$; on the other hand, if $x = D\sharp$ this probability is zero, because in this scenario the note X_t will be chosen from set $\mathcal{C}_1$, which does not contain note C♯.

IV. Markov chains

As stated beforehand, the study of stochastic processes, i.e. sequences of random variables, is quite important in Probability theory, and we already came across the Markov chains in Example 2. We introduce it now more formally and then returns to analyze in more details this introductory example.

i. Basic definitions

Intuitively, a Markov chain is like a Monopoly match: a random walk on a finite set $\mathcal{C}$, where the next step depends only on where we are now. More formally, a sequence of random variables $(X_t)_{t\in\mathbb{N}}$ is a *Markov chain* if it assume values on a common set $\mathcal{C}$ and satisfies the Equation 19:

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_2 = x_2, X_1 = x_1) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n), \tag{19}$$

for all $n \in \mathbb{N}$ and $x_{n+1}, \ldots, x_1 \in \mathcal{C}$. The elements of set $\mathcal{C}$ are called *states* of the chain and we denote the *transition probabilities* $\mathbb{P}(X_{n+1} = x_j | X_n = x_i)$ as $p_{ij}(n)$. These probabilities must be interpreted as the probability of going to state x_j in time instant $n+1$ given that in time instant n the position is state x_i.

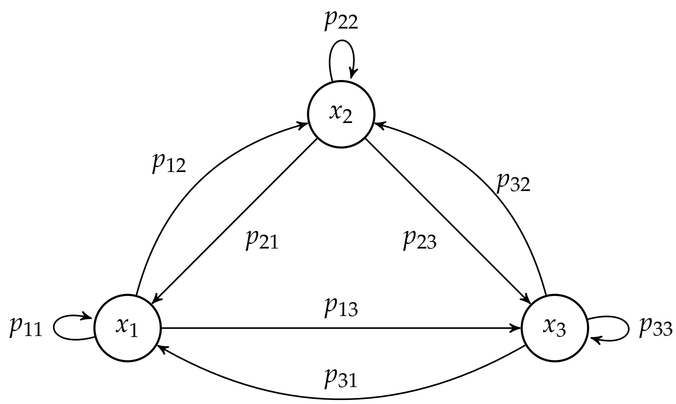

Figure 1: *Oriented graph associated with a Markov chain with states $\mathcal{C} = \{x_1, x_2, x_3\}$ and transition matrix in Equation 21.*

Note that in the previous definition we allowed the transition probabilities to depend on the time instant n, as if in the game of Monopoly we roll dice with different number of faces along the game. Chains with this characteristic will not be treated here, and we consider only the *homogeneous* ones, where the transition probabilities does not depend on time instant n, and we can simply write (Equation 20):

$$\mathbb{P}(X_{n+1} = x_j | X_n = x_i) = p_{ij}. \tag{20}$$

Since we will not deal with non-homogeneous chains, this adjective will be omitted from now on.

It is very instructive to represent Markov chains visually as an oriented graph, where the nodes are the states and the arrows indicates the allowed transitions, together with its respective probability of occurrence, as can be seen in Figure 1. The transition matrix of this chain, a concept to be shortly introduced in the beginning of next subsection, is given by

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}. \tag{21}$$

ii. Convergence of a Markov chain

A very important object associated to a Markov chain is its *transition matrix*: a square matrix $\mathbf{P}$ of size $M \times M$, where M is the cardinality of $\mathcal{C}$, whose element on the i-th row and j-th column is p_{ij}. The transition matrix uniquely determines the chain, and contains important information about its asymptotic behavior, which will lead us to a version of the laws of the large numbers for Markov chains. We will develop the reasoning to naturally arrive at the analogous result in this scenario.

Let $\boldsymbol{\pi}_1$ be an $M \times 1$ vector, containing the initial distribution probability of the chain, that is, the i-th entry of $\boldsymbol{\pi}_1$ is $\mathbb{P}(X_1 = x_i)$, the probability that the chain starts at state x_i, for all $i = 1, \ldots, M$. Note that in a Monopoly match the initial location of the players on the board is precisely determined by the game rules', but recalling Example 2, out first note could come from a random choice between sets $\mathcal{C}_1$ and $\mathcal{C}_2$. The distribution of the second step of the chain, X_2, is then given by Equation 22:

$$\mathbb{P}(X_2 = x_j) = \sum_{i=1}^{M} \mathbb{P}(X_2 = x_j | X_1 = x_i)\mathbb{P}(X_i = x_i), \tag{22}$$

for all $j = 1, \ldots, M$. Denoting the vector containing this information by $\boldsymbol{\pi}_2$, this is exactly the vector-matrix multiplication (Equation 23):

$$\boldsymbol{\pi}_2 = \boldsymbol{\pi}_1 \mathbf{P}. \tag{23}$$

It is easy to see, *via* recursion, that the distribution of the n-th step of the chain, X_n, is given by Equation 24:

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_{n-1}\mathbf{P} = \boldsymbol{\pi}_1 \mathbf{P}^{n-1}, \tag{24}$$

for all $n = 1, 2, \ldots$.

Recalling the intuition about Monopoly again, assume that some player decided to start not on the standard place, but somewhere on the opposite side of the board, for example. In the long-term it will be possible to determine where this player have started, or this information will be "forgotten" in the course of time? More formally, the distribution of X_n will eventually "forget" how the chain has started? Markov proved in 1906 that if the chain is *ergodic*, then the initial distribution will be little by little forgotten, that is,

$$\lim_{n \to \infty} \boldsymbol{\pi}_n = \boldsymbol{\pi}, \tag{25}$$

and this limit is independent of the initial distribution $\boldsymbol{\pi}_1$ of the chain.

Vector $\boldsymbol{\pi}$ is called the *stationary distribution* of the chain, and its i-th entry represents the proportion of time that the chain spends on state x_i in the long-term, and Markov also proved that it can be estimated as the average of time spent in this state in a realization of the chain, regardless of the initial state $\boldsymbol{\pi}_1$ being considered, disproving once and for all the non-rigorous argument of Nekrasov.

Ergodicity is another name borrowed from Physics, and it means, intuitively, that the chain is sufficiently connected, and its definition is beyond the scope of this text [22]. It is possible to prove that the chain is ergodic if some power of its transition matrix $\mathbf{P}$ has only positive entries. Following a reasoning analogous to the one in Equation 24, it can be proven that the entry in line i and column j of $\mathbf{P}^n$ is given by Equation 26:

$$[\mathbf{P}^n]_{ij} = \mathbb{P}(X_{n+1} = x_j | X_1 = x_i), \tag{26}$$

that is, the probability of going from state x_i to x_j in exactly n steps.

Example 7 *In order to clarify these new definitions, let us return to Example 2, building a Markov chain and studying its asymptotic properties.*

First note that we have several elements to choose as the states of the chain: the notes itself, the color of the key being played or the sets $\mathcal{C}_1$ and $\mathcal{C}_2$, abbreviated via its indices for simplicity. For this example we choose the latter. From the compositional rule therein stated, it is easy to verify that

$$\mathbb{P}(X_{n+1} = 1 | X_n = 1) = \mathbb{P}(\textit{playing a note from a black key in set } \mathcal{C}_1) = 1/2 = p_{11} \tag{27}$$
$$\mathbb{P}(X_{n+1} = 2 | X_n = 1) = \mathbb{P}(\textit{playing a note from a white key in set } \mathcal{C}_1) = 1/2 = p_{12} \tag{28}$$
$$\mathbb{P}(X_{n+1} = 1 | X_n = 2) = \mathbb{P}(\textit{playing a note from a black key in set } \mathcal{C}_2) = 2/5 = p_{21} \tag{29}$$
$$\mathbb{P}(X_{n+1} = 2 | X_n = 2) = \mathbb{P}(\textit{playing a note from a white key in set } \mathcal{C}_2) = 3/5 = p_{22}, \tag{30}$$

and therefore, the corresponding transition matrix is given by

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 \\ 2/5 & 3/5 \end{bmatrix}. \tag{31}$$

Since all of its entries are strictly positive, the chain is ergodic and Markov's theorem guarantee the existence of the stationary distribution. It can be proved that as n increases, matrix $\mathbf{P}^n$ becomes closer to

$$\begin{bmatrix} 4/9 & 5/9 \\ 4/9 & 5/9 \end{bmatrix}, \tag{32}$$

regardless of the initial distribution of the chain, that for completeness of the example, is given by $\boldsymbol{\pi}_1 = [0 \;\; 1]$, since our first note surely comes from set $\mathcal{C}_2$.

Therefore, independently on where our chain begins, in the long-term, our note has a probability of $4/9$ of being chosen from set $\mathcal{C}_1$ and of $5/9$ of being chosen from set $\mathcal{C}_2$. Since this chain is ergodic, the procedure employed by the listener indeed works, and his estimate will be, with high probability, quite close to $5/9$, provided that he listens the song for enough time.

iii. Learning the transition probabilities

The first widely known practical use Markov chains was in 1913, where Markov itself analyzed a sequence of 20,000 characters of the poem *Eugene Onegin* from Aleksandr Pushkin, inferring the transition probabilities between symbols by merely counting these transitions. Assuming that the chain is indeed homogeneous, this is known in Statistics as the *maximum likelihood estimator* [7]. For example, the probability that a vowel precedes another vowel is the ratio between the observed number of this transition and the total number of occurrences of vowels along the text.

Despite being quite intuitive this procedure possess some drawbacks, in particular when the observed sequence is not sufficient long and a big number of states is being considered, since in this scenario some transition probabilities could be quite underestimated. Techniques to overcome this and other difficulties exist, and are called *smoothing* procedures, that are beyond the scope of this text [16].

iv. Higher order Markov chains

Markov chains of higher order $N > 1$ can also be considered, that is, chains where the dependence of the observation in time instant n depends not only on X_{n-1} but also on $X_{n-2}, \ldots, X_{n-N}$.

Perhaps Markov preferred to use a higher-order chain to obtain more realistic information about Pushkin's poem, but in the absence of digital computers, estimating the transition probabilities became rapidly infeasible. For example, when considering a second-order chain, the transition matrix is not a matrix anymore, but a *tensor or order three*, a three-dimensional matrix $\mathbf{P}$ whose entry $\mathbf{P}_{ijk}$ represents the probability of the transition $x_i \to x_j \to x_k$. Whereas in usual Markov chains one needs to estimate M^2 transition probabilities, in second order chains this quantity increases to M^3 and it is easy to see that for a N-order chain, its respective tensor of order $N+1$ has M^{N+1} entries to be estimated!

V. Markov chains in music composition and analysis

Since its first use in algorithmic composition in the decade of 1950 [18][9], several applications and extensive well-written reviews can be found in the literature [2], and I specially recommend [16].

[9] Despite the paper being dated from 1961, the development of their work began in the early 1950, being published only a decade later.

In this section I will only briefly recall some remarkable applications and present in more detail the structure of *Analogique A*, by Iannis Xenakis [28].

i. A (very) small literature review

The first important aspect to note is that the notion of "state" is quite flexible to accommodate several musical aspects of interest. For instance, one can use a Markov chain to model transitions of order N between pitches, intensities, intervals, chords, vectors containing combinations of these features, among many others. Assuming that this modeling was performed intending the creation of a new musical piece[10], we came to the second point: will the transition probabilities be estimated from a *corpus* of interest or fixed beforehand, aiming some desired behavior? Combinations between these two aspects essentially classifies the works using Markov chains to create new musical material.

In the beginning of the decade of 1950, Harry Olson and Hebert Belar were the first researchers that used Markov models in algorithmic composition, developing the first known machine called a "synthesizer" [18]. Basically, by analyzing a *corpus* of 11 melodies by Stephen Foster, transposed to D major, they estimated transition probabilities of order 0[11], 1 and 2 between pitches of notes and of order 0 between rhythmic patterns in time signatures of $\frac{3}{4}$ and $\frac{4}{4}$. The notes of the new composition were generated accordingly to both models for pitch and rhythm simultaneously, reproduced by a speaker and recorded in a magnetic tape.

One of the main disadvantages of the Markov model in algorithmic composition is that it only captures short-term dependencies. For example, assume that the states of the chain are pitches, and that the probability transitions were estimated from a sufficiently large *corpus*, in order to overcome the difficulty of estimating transition probabilities in high-order chains. If the order being used is low, the learned model is quite simple and the new generated melodies usually does not resemble the *corpus* that one desires to mimic. However, increasing the order does not solve the problem, since only still short-term structures are being captured, only the definition of "short" being just a little enlarged. Therefore, it is possible that with a higher-order model (around 10, for example), the generated melodies will be only a plagiarism of some content already present in the *corpus*, a behavior one does not wish to observe, if the creation of new musical material is the researcher's goal.

However, the former paragraph does not mean that the Markov model needs to be abandoned at all! Instead of applying it to low-level musical aspects such as pitch and duration of notes, some researchers obtained good results in another structures, such as chord classes [1], harmony [20], interval between pair of notes [12] among others ([16, p. 24-43]). Indeed, forwarding somewhat in time, in 2002 François Pachet proposed the *Continuator* [19], a system that is not intended to create new music material, but to continue musical phrases played by a musician in a MIDI controller, responding to them in the same style of the excerpt being played, by learning transitions between pitches, velocity, beginning and length of the notes, among others. His work does not employed the usual transition matrices, but another structure called *prefix tree*, which can be proven to be equivalent of a varying-order Markov model. By only "filling the gaps" and composing short excerpts of music, his approach masks the drawbacks previously stated.

[10] One also could be interested in analyzing the style of a composer, for example, by means of its transition matrix.

[11] A Markov model of order 0 consider only the relative frequency of states within a *corpus*, as the black and white keys in Example 1.

ii. Iannis Xenakis' *Analogique A*

In 1958 and 1959, previously to the publication of the work of Olson and Belar [18], Iannis Xenakis, a Greek-French composer, music theorist, architect, performance director and engineer has pioneered the use of Markov chains in music composition. Moreover, instead of estimating the transition probabilities from a *corpus*, he fixed beforehand the transition probabilities and with calculations performed by hand[12] he created entirely new musical material in *Analogique A* and *Analogique B*, for string orchestra and sinusoidal sounds, respectively. This process is fully described in Chapters 2 and 3 of [28], with a quite cumbersome and intricate language, from both mathematical and musical viewpoints. In this subsection we briefly describe some remarkable aspects of this material, with a more accessible and updated terminology, mainly focused on *Analogique A* since the mathematical formulation of both compositions are similar. An expansion of this subsection to fully update some excerpts of [28] is addressed as a future work.

The first question posed by Xenakis is essentially about the adequate musical object over which impose the Markov structure. This leads him to a deep analysis of the nature of sounds, from which this quote is particularly interesting:

> All sound is an integration of grains, of elementary sonic particles, of sonic quanta. Each of these elementary grains has a threefold nature: duration, frequency, and intensity. All sound, even all continuous sonic variation, is conceived as an assemblage of a large number of elementary grains adequately disposed in time. [...] In the attack, body, and decline of a complex sound, thousands of pure sounds appear in a more or less short interval of time, Δt. Hecatombs of pure sounds are necessary for the creation of a complex sound. ([28, p. 43-44]).

It becomes clear that Xenakis was aware of the work of Dennis Gabor, a Hungarian-British electrical engineer and physicist who received the Nobel Prize in 1971, about time-frequency analysis and the nature of sound [11], since Xenakis' claim is essentially the same as saying that the sonic content in a given short interval of time can be decomposed as a superposition of more simple sounds. This is the principle behind the *spectrogram*, a visual representation of the frequency spectrum of a sound as it varies with time. Figure 2 illustrates a spectrogram of the subject of *Ricercar a 6*, the six-voice fugue from *The Musical Offering* from J. S. Bach.

Xenakis assumes that the length of the time frame being analyzed, Δt, is small but invariable, in order to ignore it and consider only the components of frequency and intensity, which he denotes as F and G, respectively. It is important to note that not every possible combination of F and G is audible to the human ear[13], a point that he is still aware, and without loss of generality, the audible region of the FG plane can be put in a one-to-one correspondence with a rectangle, as illustrated in Figure II-6 in [28, p. 49]. This leads us to his definition of *screen*:

> The screen is the audible area (FG) fixed by a sufficiently close and homogeneous grid [...], the cells of which may or may not be occupied by grains. In this way, any sound and its history may be described by means of a sufficiently large number of sheets of paper carrying a given screen S. These sheets are placed in a fixed lexicographical order. ([28, p. 51])

[12] The invention of the digital computer dates from the decade of 1950, but their widespread use was only possible several decades later.

[13] The way we process and perceive sounds is the main object of study of *psychoacoustics* [30]. The curves in Figure II-6 in [28, p. 49] are called the *equal loudness curves*. Psychoacoustics is quite important nowadays, since several lossy audio codecs such as MP3 depends on this theory in order to properly "throw away" information we do not perceive.

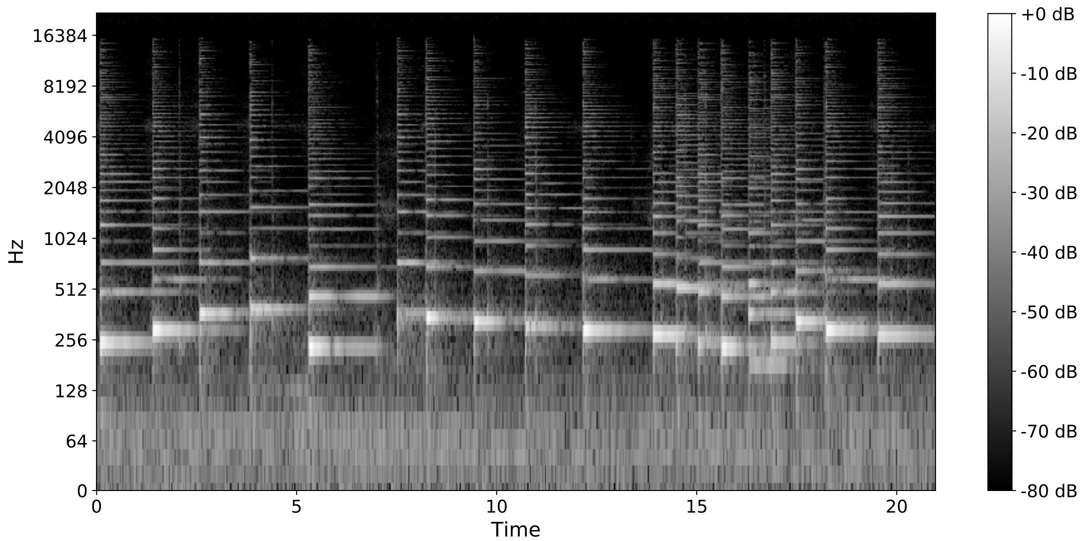

Figure 2: *Spectrogram of the subject of* Ricercar a 6*. Each point in the figure represents the intensity (in dB) of the respective frequency at time t (y and x coordinates of the point, respectively).*

Recalling again Figure 2, each vertical line can be understood as a screen, and the ordered set of screens forms the whole sound. In Figure 3 we can see one particular screen associated with this spectrogram.

One last variable that Xenakis considers important is the density of grains per unit of volume $\Delta F \Delta G \Delta t$, denoted by D. He suggests that it should also be measured in logarithmic scale, with base between 2 and 3. This particular choice is due to the fact that later in Chapter 2 Xenakis compares the density of grains, the complexity of a music and the *entropy* of a screen[14]. Along much of this Chapter he discuss several aspects of the screens and draws many parallels with Physics. In particular, he models the density of grains per screen *via* the Poisson distribution, and imagining that the continuous evolution of grains along time is similar to the movement of the molecules on a gas, its interaction can also be described by the Maxwell-Boltzmann distribution. He also operate screens with the usual operators of set theory like union, intersection and complements. Although very interesting, these discussion are outside the scope of this work, since they are not critical to have a good understanding of the stochastic structure of *Anaogique A*.

When analyzing the way screens are linked, he propose that transitions can be stochastic, in particular, following a Markov chain. Therefore, from Xenakis' viewpoint, screens are the most general object one can impose a Markov structure in order to create new music. However, as noted earlier in Chapter 2 *via* another parallel with Physics, the manipulation of individuals grains in order to achieve this goal is infeasible:

> Theoretically, a complex sound can only be exhaustively represented on a three-dimensional diagram F, G, t, giving the instantaneous frequency and intensity as a function of time. But in practice this boils down to saying that in order to represent a momentary sound, such as a simple noise made by a car, months of calculations and graphs are necessary. This impasse is strikingly reminiscent of classical mechanics, which claimed that, given sufficient time, it could account for all physical and even

[14]The entropy of a random variable will briefly appear in Section VII, and it is also measured in logarithmic scale, usually in base 2.

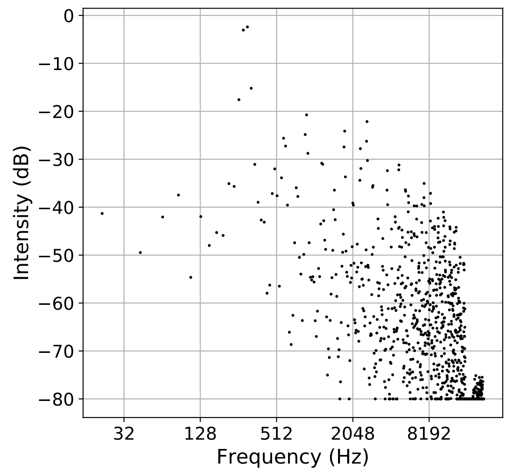

Figure 3: *Vertical line of the spectrogram in Figure 2, representing a screen.*

> biological phenomena using only a few formulae. But just to describe the state of a gaseous mass of greatly reduced volume at one instant t, even if simplifications are allowed at the beginning of the calculation, would require several centuries of human work! [...]
>
> The same thing holds true for complex as well as quite simple sounds. It would be a waste of effort to attempt to account analytically or graphically for the characteristics of complex sounds when they are to be used in an electromagnetic composition. For the manipulation of these sounds macroscopic methods are necessary. [...]
>
> Microsounds and elementary grains have no importance on the scale which we have chosen. Only groups of grains and the characteristics of these groups have any meaning. ([28, p. 49-50])

In order to compose *Analogique A*, Xenakis firstly restricts himself to impose a Markov structure over parameters F, G and D of the screens, each taking only two possible values. This will lead to a Markov structure on a set of screens as will become clear. He claims that more complex structures could lead to more interesting musical result, but the necessary volume of calculation is unfeasible to perform by hand, necessitating a computer to perform them.

Consider f_0 and f_1, g_0 and g_1, and d_0 and d_1 two distinct audible frequency, intensity and density regions[15], respectively. Each of these pair of variables is governed by one of the two transition matrices below:

$$\mathbf{P}_1 = \begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix} \qquad \mathbf{P}_2 = \begin{bmatrix} 0.85 & 0.15 \\ 0.4 & 0.6 \end{bmatrix}. \tag{33}$$

The Markov structure that Xenakis imposes on F, G, and D can be described in two steps: the current values of F, G and D determines if the next value of each one will be chosen accordingly to $\mathbf{P}_1$ or $\mathbf{P}_2$, and then, its respective values are chosen accordingly to the corresponding transition matrix. He calls this process as *coupling*, and it is completely described by the correspondence in Table 1, where the first line contains the current value of the variables, whereas the second contains the respective transition matrix for the variable in the third line.

[15] Here Xenakis measures density in *terts*, logarithm in base 3.

Table 1: *Coupling of transition matrices* $\mathbf{P}_1$ *and* $\mathbf{P}_2$.

f_0	f_1	d_0	d_1	g_0	g_1	g_0	g_1	f_0	f_1	d_0	d_1
$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_2$	$\mathbf{P}_1$	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_1$	$\mathbf{P}_2$
D	D	F	F	D	D	F	F	G	G	G	G

Table 2: *Transition matrix between the screens formed by the possible values of F, G and D. Letters from A to H are used to abbreviate the corresponding screens.*

	A $(f_0g_0d_0)$	B $(f_0g_0d_1)$	C $(f_0g_1d_0)$	D $(f_0g_1d_1)$	E $(f_1g_0d_0)$	F $(f_1g_0d_1)$	G $(f_1g_1d_0)$	H $(f_1g_1d_1)$
A	0.021	0.084	0.084	0.336	0.019	0.076	0.076	0.304
B	0.357	0.089	0.323	0.081	0.063	0.016	0.057	0.014
C	0.084	0.076	0.021	0.019	0.336	0.304	0.084	0.076
D	0.189	0.126	0.126	0.084	0.171	0.114	0.114	0.076
E	0.165	0.150	0.150	0.135	0.110	0.100	0.100	0.090
F	0.204	0.136	0.036	0.024	0.306	0.204	0.054	0.036
G	0.408	0.072	0.272	0.048	0.102	0.018	0.068	0.012
H	0.096	0.144	0.144	0.216	0.064	0.096	0.096	0.144

As an example, if the current state is $(f_0\ g_0\ d_1)$, the next value of F is conditioned to g_0 and d_1, and the table informs us that its next value should be chosen according to transition matrix $\mathbf{P}_2$. However, f_0 indicates that the next value of G should be chosen according to $\mathbf{P}_1$, but d_1 says that it should be chosen according to $\mathbf{P}_2$. Since in this case there is not agreement, one of the transition matrices are chosen at random with equal probability, and the next value of G will be chosen from this matrix.

Since each variable has two possible values, when combined we have 8 possibilities, that forms exactly 8 screens. From the methodology described above it is possible to compute the transition matrix between these screens, described in Table 2.

Therefore, this entire procedure is the *macroscopic method* that Xenakis mentioned before: there is a mechanism of transformation for frequency, intensity and density (transition matrices $\mathbf{P}_1$ and $\mathbf{P}_2$) and a interaction protocol between them (Table 1), that when combined implies in a mechanism of transformation between screens (transition matrix in Table 2).

Note that the transition matrix in Table 2 is ergodic, since each entry is strictly positive. Therefore, the respective Markov chain possess a stationary distribution, shown in Table 3. However, Xenakis notes that the chain converges quite quickly to the stationary distribution, and

Table 3: *Stationary distribution of transition matrix in Table 2.*

A $(f_0g_0d_0)$	B $(f_0g_0d_1)$	C $(f_0g_1d_0)$	D $(f_0g_1d_1)$	E $(f_1g_0d_0)$	F $(f_1g_0d_1)$	G $(f_1g_1d_0)$	H $(f_1g_1d_1)$
0.17	0.13	0.13	0.11	0.14	0.12	0.10	0.10

although not explicitly mentioned, he does not want that a initial screen freely evolve according to this transition matrix, maybe because of the "lack of innovation" of the stationary distribution. Therefore, an idea to overcome this difficulty can be loosely described as evolving a particular screen until it is close to the stationary distribution, applying a "perturbation", evolve the obtained screen again, and so on. Xenakis claims that this procedure does not diminish the importance of the Markov structure imposed by transition matrix on Table 2 but *confirms* its importance, since we are successively confirming and negating its structure, which he calls *mechanism* Z in the following quote:

> In effect the intrinsic value of the organism thus created lies in the fact that it must manifest itself, be. The perturbations which apparently change its structure represent so many negations of this existence. And if we create a succession of perturbations or negations, on the one hand, and stationary states or existences on the other, we are only *affirming* mechanism Z. In other words, at first we argue positively by proposing and offering as evidence the existence itself; and then we confirm it negatively by opposing it with perturbatory states. ([28, p. 94])

Therefore, the stochastic process underlying *Analogique A* can be described by the following *kinetic diagram*, as named by Xenakis:

$$E \to P_A^0 \to P_A' \to E \to P_C' \to P_C^0 \to P_B^0 \to P_B' \to E \to P_A', \tag{34}$$

where each of these symbols is a *protocol*, which means that screens are sampled according to a particular probability distribution:

- E is the stationary distribution on Table 3;
- P_A^0 is the distribution which assigns probability 1 to screen A and zero to the others (a *perturbation* towards screen A);
- P_A' is the result of applying the transition matrix in Table 2 to P_A^0;
- P_B^0, P_B', P_C^0, P_C' are similarly described.

Therefore, *Analogique A* is the result of some realization of this stochastic process, and the particular choices of f_0, f_1, g_0, g_1, d_0, and d_1 are shown in Figures III-8, III-9, and III-10 of [28, p. 98-99], respectively. The combinations between these features correspond to screens A from H in Figure III-13 of [28, p. 101], where the Roman numerals indicates the location of specific clouds of grains and Arabic numerals are the mean densities in grains per second. Finally, in order to properly transform this realization in a music, Xenakis sets the duration of each screen Δt as 1.11 s, the duration of one half note, and within this duration the densities of the occupied cells must be realized. Each of the protocol in the kinetic diagram in Equation 34 is explored with 30 screens, sampled from the corresponding probability distribution. This implies in 15 measures for each protocol, and since there are 10 protocols to be explored, *Analogique A* consists of 150 measures. Its underlying stochastic process can then be denoted as $S_1, \ldots, S_{300}$ where the individual probability distributions are described *via* the kinetic diagram in Equation 34, that is, $S_1, \ldots, S_{30}$ are independently sampled from distribution E, $S_{31} \ldots, S_{60}$ from P_A^0, and so on. Note that this process is not homogeneous, since the probability distribution being sampled changes at every 30 screens, despite its major inspiration, the Markov chain whose transition matrix is Table 2, being a homogeneous process.

It is important to note that since this composition is to be performed by string instruments, its execution does not correspond to the screens in Figure III-13 of [28, p. 101], because of the timbre of the particular instruments being played. Therefore, this stochastic structure corresponds only to the *fundamental frequencies*, and a particular execution possess much more complex screens.
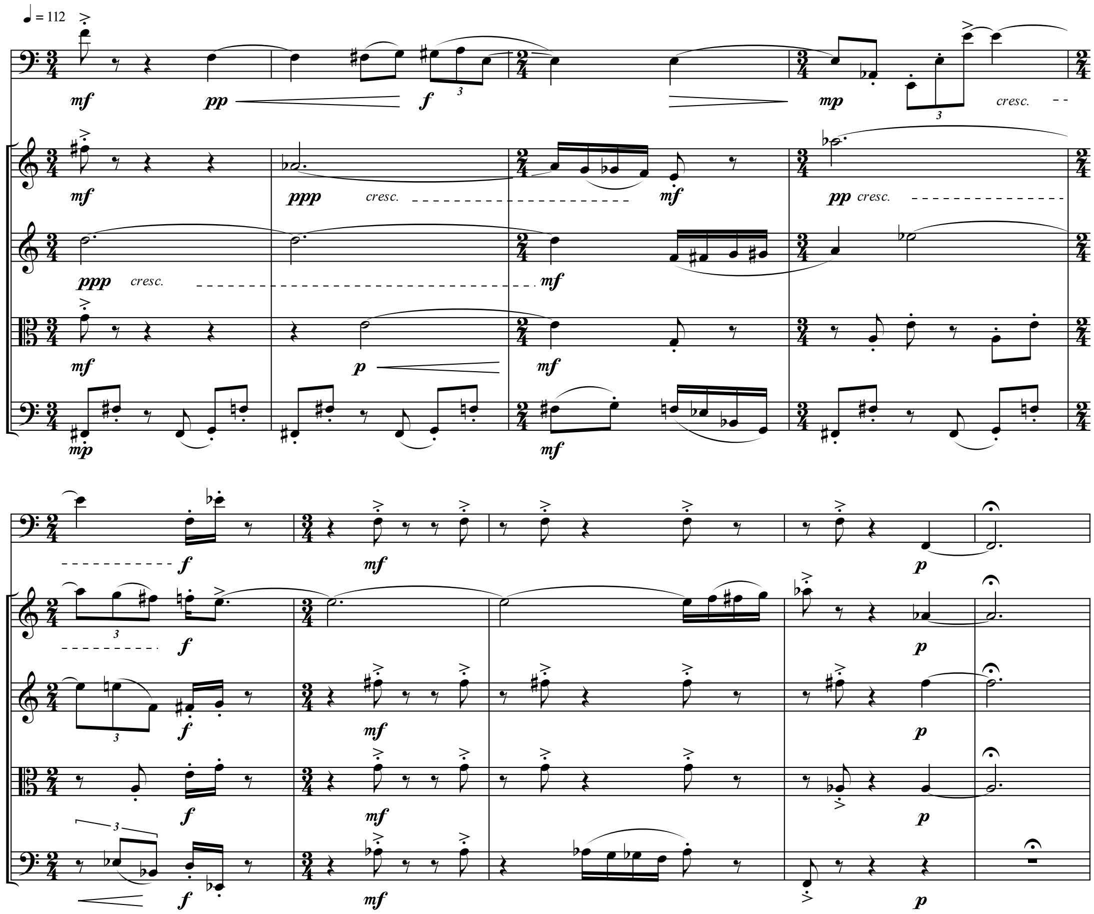
Figure 4: *Excerpt from Liduino Pitombeira's* Brazilian Landscapes No. 20 *for bassoon and string quartet.*

VI. Three more examples

The two musically-inspired examples previously developed, Examples 1 and 2, guided much of our intuition on mathematical concepts up to this point. However, one must agree that they are of low musical interest. In this section we present two applications of probabilistic tools in composition that, being less naive, leads to more interesting results.

i. Liduino Pitombeira's *Brazilian Landscapes No. 20* for bassoon and string quartet

The score on Figure 4 is an excerpt from a piece for bassoon and string quartet entitled *Brazilian Landscapes No.20* by Liduino Pitombeira. Its pitch classes comes from the Binomial distribution; the rhythm and specific pitch of the notes were not inspired by randomness, and we will explain only the first aforementioned aspect.

Firstly, consider $X \sim \text{Bin}(12, 1/2)$. With the intuition that X counts the number of heads obtained when tossing a fair coin 12 times, it becomes clear that its possible outcomes are $x = 0, \ldots, 12$. If the 12 pitch classes from C to B were put in a one-to-one correspondence with $\{0, \ldots, 11\}$ the possible outcomes of X except the last, we can loosely say that the notes in Figure

Table 4: *Probabilities associated with the notes in Liduino Pitombeira's* Brazilian Landscapes No. 20 *for bassoon and string quartet.*

x	$\mathbb{P}(X = x)$ (in %)	Pitch class	Quantity
0	0.02	C	0
1	0.29	C♯	0
2	1.61	D	2
3	5.37	D♯	5
4	12.08	E	12
5	19.34	F	19
6	22.56	F♯	23
7	19.34	G	19
8	12.08	G♯	12
9	5.37	A	5
10	1.61	A♯	2
11	0.29	B	0
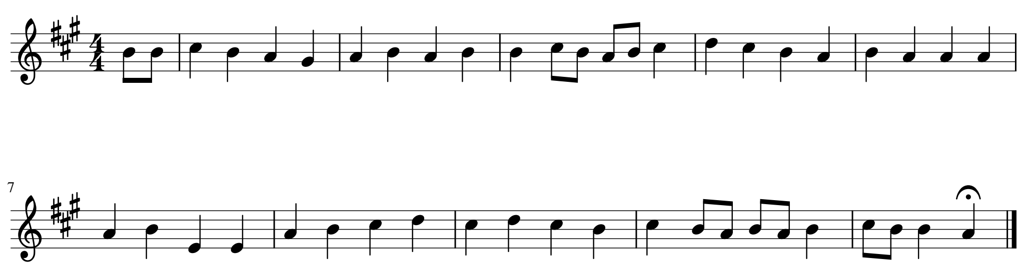
Figure 5: *A short composition based on the first order transitions from the chorales of J. S. Bach in the key of A major, with the rhythm of the chorale from BWV 104.*

4 were chosen accordingly to X', the random variable X truncated to only assume values in the set $\{0, \ldots, 11\}$. However, the procedure employed by Pitombeira was to list the probabilities of outcomes of X and then round its values to integer numbers, as in the second and fourth columns of Table 4, where the probability is measured in %. Therefore, the integer number in fourth column represents how many times the respective pitch class in third column will be present within the excerpt.

ii. Johann "Markov" Bach

On the other hand, the short pieces in Figures 5 and 6 were created with pure randomness, from a Markov structure of order 1 and 4, respectively, estimated from some of the chorales from Johann Sebastian Bach and using some tools from the Python package `music21` [8], a software tailored to perform computer-aided musicology. We now describe these excerpts in more details.

In order to estimate the transition probabilities, I considered only the soprano voice from Bach's chorales on the database of `music21` in the key of A major. This quite small database consisted of 1,637 notes, being only 13 of them distinct, whose MIDI numbers are 64, 66, 67, 68, 69,

Table 5: *Transition probabilities of order 1 estimated from the chorales of J. S. Bach in the key of A major. The states of the chain are the MIDI numbers of the notes, and probabilities are measured in %.*

	64	66	67	68	69	70	71	73	74	75	76	78	79
64	17.39	19.57	0	0	41.3	0	15.22	0	0	0	6.52	0	0
66	48.0	0	0	42.0	0	0	10	0	0	0	0	0	0
67	0	100	0	0	0	0	0	0	0	0	0	0	0
68	11.69	37.66	0	3.9	45.45	0	0	1.3	0	0	0	0	0
69	0.7	2.46	0	17.89	25.26	0	38.25	11.23	2.46	0	1.75	0	0
70	0	0	0	0	0	0	100	0	0	0	0	0	0
71	0.84	0.28	0.56	0.56	34.83	0.84	14.89	44.1	1.69	0	1.4	0	0
73	0	0.52	0	0	8.31	0	41.56	14.29	30.91	1.56	2.08	0.78	0
74	0	0	0	0	0	0	3.56	56.89	4.44	0	35.11	0	0
75	0	0	0	0	0	0	0	0	0	0	100	0	0
76	0	0	0	0	1.74	0	6.4	6.98	48.26	2.33	24.42	9.88	0
78	0	0	0	0	0	0	0	0	0	4.17	79.17	12.5	4.17
79	0	0	0	0	0	0	0	0	0	0	0	100	0

Table 6: *Stationary distribution of the transition matrix displayed in Table 5. Probabilities are measured in %.*

64 E4	66 F♯4	67 G4	68 G♯4	69 A4	70 A♯4	71 B4	73 C♯5	74 D5	75 D♯5	76 E5	78 F♯5	79 G5
2.81	3.05	0.12	4.7	17.41	0.18	21.75	23.52	13.74	0.67	10.51	1.47	0.06

70, 71, 73, 74, 75, 76, 78, and 79. The transition matrix of the chain of order 1 was estimated from this data by using the simple counting procedure described in Section IV and is displayed in Table 5, where the probabilities are measured in %. The first note on Figure 5 was chosen at random, uniformly from all the 13 possible notes, and 45 more notes were generated accordingly to the transition matrix in Table 5. The rhythm displayed in Figure 5 is from the chorale of BWV 104, one of the excerpts in the database analyzed. It is important to observe that the last note being an A was purely by chance, since I have not tried to generate the music several times until something "good" was obtained, but simply picked up the first sequence generated by the algorithm.

It can be shown that the transition matrix in Table 5 is ergodic, since its fourth power only has strictly positive entries. Therefore, it has a stationary distribution, displayed in Table 6 in %, together with the corresponding note names. Recall that this distribution can be interpreted as the proportion of appearance of each of these notes in a sufficiently large realization generated according to transition matrix in Table 5.

In order to illustrate the aforementioned drawback when considering higher order Markov chains, the short excerpt in Figure 6 was generated from a chain of order 4, whose transition probabilities were also inferred from the same *corpus*. Recall that now the estimation *via* the counting procedure is not quite reliable, since there are $13^5 = 371,293$ transition probabilities to estimate from the 1,637 observed notes, and several of these transitions does not even appear in the *corpus*. Indeed, from all of the $13^4 = 28,561$ distinct groups of four notes that can be formed with the 13 available notes, only 375 appears in the database, and for several of them there is

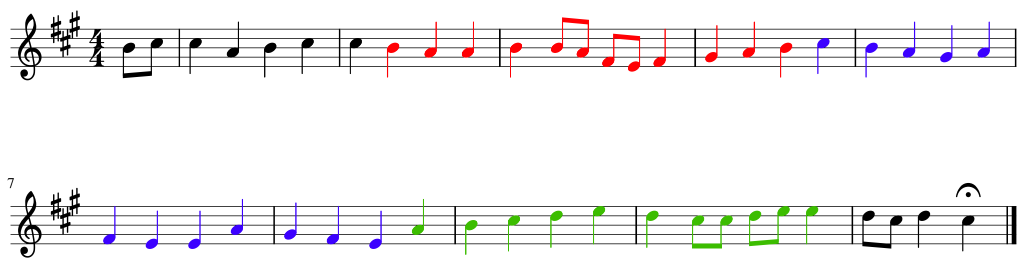

Figure 6: *A short composition based on the fourth order transitions from the chorales of J. S. Bach in the key of A major, with the rhythm of the chorale from BWV 104. The blocks of colored noted are already present in the database being analyzed.*

only one possibility of note to transition to, meaning that is very likely that some excerpt of some chorale is being exactly replicated. Indeed, the excerpts of 12 notes in red and blue in Figure 6 are already present within the *corpus* once, and the excerpt in green, consisting of 11 notes, appears eight times! This example illustrates that even with a low order chain, longer structures from the *corpus* can be replicated, without the creation of new musical material.

VII. Interpretability and flexibility

When using some statistical method to extract information from a dataset, one must have in mind the duality between intepretability and flexibility. There are some formal definitions of the flexibility of some model in the Statistics literature, but here let us understand the main idea, by comparing models in both extremes of this spectrum: Markov chains and neural networks.

Despite being outside the scope of this work, neural networks represent the state of the art in several Machine Learning techniques, in particular, algorithmic composition [10, 9]. However, it is extremely difficult to have an intuitive understanding of what the hidden layers are *exactly* doing, except in some special architectures. Usually, the only interpretable information in neural networks are the input and output layers, and if someone is interested in inferring information about the style of a composer, for example, is it very unlikely that this approach will lead him there, but it may have the capability of creating new musical pieces which passes in Turing tests with trained musicians [3].

On the other hand, we are convinced that Markov chains are highly interpretable, and some examples were presented illustrating that in order not to simply mimic the content already present in some *corpus* it must be wisely employed, in particular when dealing with its order and the definition of state, placing it then further to the flexible side of the spectrum and closer to the interpretable one.

A promising approach that claims to be in between the two extremes of this duality is [23]. Essentially, François Pachet and his collaborators propose that in order to properly capture some long-term information it is necessary only to model order two interactions but not only between adjacent notes, as in the Markov chain scenario. More specifically, they propose that a probability distribution that properly describes the information contained within a *corpus* is the *distribution of maximum entropy that honors the proportion of single notes and pair of notes up to some distance*.

This claim is one particular instance of a more general framework proposed in 1957 by the American physicist Edwin Jaynes. In this year he published two papers where some gaps between the notion of entropy in Statistical Mechanics and Probability were bridged, and also proposed an interpretation Thermodynamics in probabilistic terms [14, 15]. Essentially, he claimed that the probability distribution which best represents our current state of knowledge in some scenario is the one with maximum entropy, conditioned to the constraints obtained from the observed data.

The entropy of a random variable X can be understood as the *average surprise* contained therein, and is defined as shown in Equation 35:

$$\mathcal{H}(X) = \sum_{i=1}^{\infty} \mathbb{P}(X = x_i) \log\left(\frac{1}{\mathbb{P}(X = x_i)}\right) = -\sum_{i=1}^{\infty} \mathbb{P}(X = x_i) \log(\mathbb{P}(X = x_i)), \tag{35}$$

in the particular case of a discrete random variable X.

A deeper discussion of entropy of random variables and the applications of maximum entropy methods to algorithmic composition is outside the scope of this work, but it is important to remark the importance of this concept nowadays, not only in Science but directly in our lives. It was firstly introduced in 1948 by the American mathematician and electrical engineer Claude Shannon, in a landmark paper titled "A Mathematical Theory of Communication", and the concepts and theory introduced in this work are present everywhere, since it allows to efficiently compress and transmit information in a secure manner, and were crucial to the success of the space missions Voyager and Apollo, the invention of the compact disc and other medias, development of Internet, among several others. Moreover, the theory introduced by Shannon, nowadays known as *Information Theory*, is a field of knowledge with intersection with many others such as Statistics, Computer Science, Physics, Linguistics, Cryptography, and now, also Music!

VIII. Conclusion

In this work some basic aspects of Probability theory and Markov chains were introduced, mainly in a intuitive manner, in order to motivate researchers to employ these and more recent tools to perform music composition and analysis. Several examples were presented, with special attention to the brief analysis of Xenakis' *Analogique A* in Section V, and the excerpt of *Brazilian Landscapes No. 20* for bassoon and string quartet and both simulations from Markov chains in Section VI. My main goal with this work is to raise more questions and curiosity than answering them, in order to motivate the reader to go deeper in the literature on this subject. I hope that this goal was achieved.

Quoting again Isaac Newton, this text was only possible because I was standing on shoulders of giants. Without the references [16, 17, 28], some deep conversations with my doctorate advisor Luiz W. P. Biscainho about Music, Mathematics, life, the Universe and everything, and the meetings with Carlos Almada, Liduino Pitombeira, Pauxy Gentil-Nunes, Stefanella Boatto and Petrucio Viana to discuss Music and Mathematics, this work would not be possible. I am also deeply grateful to my student Nathalie Deziderio, for reviewing and making suggestions to improve this text.

References

[1] Almada, C. de L., *et. al.* (2019). Composição algorítmica de progressões harmônicas ao estilo de Antônio Carlos Jobim através de cadeias de Markov. In: *Proceedings of the 4th. International Congress of Music and Mathematics*, (in press). Rio de Janeiro, Brasil, 2019. Federal University of Rio de Janeiro.

[2] Ames, C. (1989). The Markov Process as a Compositional Model: A Survey and Tutorial. *Leonardo*, vol. 22, no. 2, pp. 175–187.

[3] BachBot. *The BachBot Challenge*. Available in https://bachbot.com/, accesed on 24/11/2019.

[4] Balakrishnan, V. (2009). *Lecture 1 - Introduction to Quantum Physics: Heisenberg's uncertainty principle*. Available in: https://www.youtube.com/watch?v=TcmGYe39XG0, accesed on 23/11/2019.

[5] Briot, J.-P., Hadjeres, G., and Pachet, F.-D. (2019) Deep Learning Techniques for Music Generation – A Survey. Available in: https://arxiv.org/abs/1709.01620.

[6] Broemeling, L. D. (2011). An Account of Early Statistical Inference in Arab Cryptology. *The American Statistician*, vol. 65, no. 4, pp. 255–257.

[7] Casella, G., and Berger, R. L. (2001). *Statistical Inference*. Boston: Cengage Learning.

[8] Cuthbert, M., and Ariza, C. (2010). music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. In: *Proceedings of the 11th. International Society for Music Information Retrieval Conference*, pp. 637–642. Utrecht, Netherlands, 2010. International Society for Music Information Retrieval.

[9] Eck, D., and Schmidhuber, J. (2002). A First Look at Music Composition using LSTM Recurrent Neural Networks. *Technical Report No. IDSIA-07-02 of the Instituto Dalle Molle di studi sull' intelligenza artificiale*.

[10] Hadjeres, G., Pachet, F., and Nielsen, F. (2017). DeepBach: a Steerable Model for Bach Chorales Generation. In: *Proceedings of the 34th. International Conference on Machine Learning*, pp. 1362–1371. Sydney, Australia, 2017. International Conference on Machine Learning.

[11] Gabor, D. (1947). Acoustical Quanta and the Theory of Hearing. *Nature*, vol. 159, pp. 591–594.

[12] Hiller, Jr., L. A., and Isaacson, L. M. (1957). Musical Composition with a High Speed Digital Computer. In: *Audio Engineering Society Convention 9*. New York, USA, 1957. Audio Engineering Society.

[13] Ibrahim, A. A.-K. (1992). The origins of cryptology: The Arab contributions. *Cryptologia*, vol. 16, no. 2, pp. 97–126.

[14] Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physical Review: Series II*, vol. 106, no. 4, pp. 620–630.

[15] Jaynes, E. T. (1957). Information Theory and Statistical Mechanics II. *Physical Review: Series II*, vol. 108, no. 2, pp. 171–190.

[16] Maia, L. S. (2016). *Formalismos da Composição Algorítmica – Um Experimento com Canções Folclóricas Brasileiras* Dissertation (M.Sc. in Electrical Engineering). Graduate Program in Electrical Engineering, COPPE, Federal University of Rio de Janeiro, Rio de Janeiro.

[17] Mlodinow, L. (2009). *The Drunkard's Walk: How Randomness Rules Our Lives*. New York: Vintage.

[18] Olson, H. F., and Belar, H. (1961). Aid to Music Composition Employing a Random Probability System. *The Journal of the Acoustical Society of America*, vol. 33, no. 9, pp. 1163–1170.

[19] Pachet, F. (2002) The Continuator: Musical Interaction With Style. In: *Proceedings of the 2002 International Computer Music Conference*, pp. 211–218. Göteborg, Sweden, 2002. International Computer Music Assotiation.

[20] Ponsford, D., Wiggins, G. and Mellish, C. (1999) Statistical learning of harmonic movement. *Journal of New Music Research*, vol. 28, no. 2, pp. 150-177.

[21] Ross, S. (1994). *A First Course in Probability*. Upper Saddle River: Prentice Hall.

[22] Ross, S. (2019). *Introduction to Probability Models*. Cambridge: Academic Press.

[23] Sakellariou, J., Tria, F., Loreto, V., and Pachet, F. (2017). Maximum entropy models capture melodic styles. *Nature Scientific Reports*, v. 7, no. 9172.

[24] Seneta, E. (1996). Markov and the Birth of Chain Dependence Theory. *International Statistical Review*, v. 64, no. 3, pp. 255-263.

[25] Singh, S. (2000). *The Code Book: The Science of Secrecy From Ancient Egypt to Quantum Cryptography*. New York: Anchor Books.

[26] Stigler, S. M. (1990). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge: Belknap Press.

[27] Wasserman, L. (2004). *All of Statistics: A Concise Course in Statistical Inference*. New York: Springer.

[28] Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Composition*. Hillsdale: Pendragon Press.

[29] Yanchenko, A. (2017). *Classical Music Composition Using Hidden Markov Models*. Dissertation (Ph.D in Statistics). Department of Statistical Science, Graduate School of Duke University, Duke University.

[30] Zwicker, E., and Fastl, H. (1999). *Psychoacoustics: Facts and Models*. New York: Springer.

Toward a Theory of Structuring Rhythm in Improvisation in Timeline-Based Musics

Stephen Paul Guerra
Nothern Arizona University
stephen.guerra@nau.edu
Orcid: 0000-0002-1907-112X

Abstract: *Among high-level performers of Afro/diasporic musics, it is generally assumed that solo and timeline rhythmic structures relate in some meaningful way(s). Despite the substantial ethnographic and music-theoretic research that has been done with regard to Afro/diasporic timelines in the past sixty years, little work has explicitly attempted to define and demonstrate the would-be nature of such relationships. This article intends to stimulate a new research agenda that will close this gap between theoretical and practical knowledge. It posits four techniques (two transformational and two non-transformational) that allow one of Afro-Brazilian guitarist Baden Powell's most rhythmically challenging recorded solos to be clearly and concisely understood as essentially six consecutive cycles of a samba timeline under various manipulations. Demonstrating the potential analytical and pedagogical economy that a theory of solo and timeline rhythmic structure relationships may underwrite, this article aims to inspire other authors to study the relationships of solo and timeline in the musics of their own specializations.*

Keywords*: Timeline. Samba. Improvisation. Baden Powell. Afro-disporic.*

I. Introduction

Figures 1 and 2 transcribe passages from Afro-Brazilian guitarist Baden Powell's (1937-2000) solos to *É de lei* (It's a given) [47] and *Samba triste* (Sad samba) [45] respectively. Based on my experience talking, studying, and performing with high-level musicians of Afro/diasporic timeline-based musics,[1] I assume that the rhythmic procedures of Baden's solos would be immediately recognizable to practitioners. Regarding Figure 1, they would say that Baden is "playing with" a samba timeline, or, if influenced by Afro-Cuban pedagogy, that he is playing "in [samba] clave." In either case, they would be drawing attention to the fact that the majority of onsets in each of Baden's two-bar phrases coincide with an instantiation of a prominent samba timeline, pictured floating above the staves in Figure 1. Baden's phrases in mm.1-2 and 5-6 are identical to the timeline except that they add a pickup to the latter's fourth onset. The third phrase of mm.3-4 has only a few more differences. In addition to the pickup to the timeline's fourth onset, it also fills the space between the timeline's fifth and sixth onsets

[1] I have studied and played Brazilian choro and samba (mainly as a guitarist but also as a percussionist) for more than ten years. During the summer months of 2016 and 2017, I conducted music-technical formal interviews with more than twenty high-level professional musicians based out of Rio de Janeiro and São Paulo. I also performed informally with many of the same. Over the last three years in New York City, I conversed and/or worked with advanced musicians (native to the US and not) of Brazilian and Afro-Cuban music. Then, I use "Afro/diasporic" to mean the whole of both African and Afro-diasporic traditions after Gerstin [16].

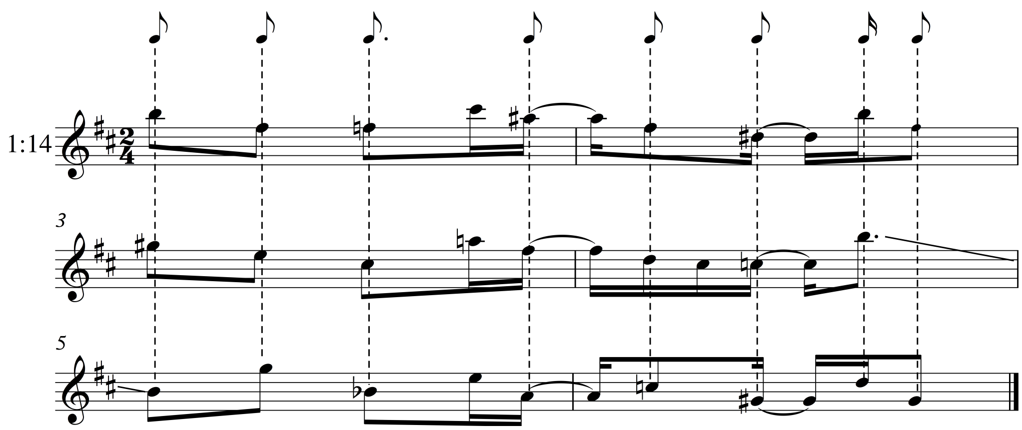

Figure 1: *Transcription of guitar part from É de lei [47].*

and deletes the timeline's final onset, resulting in a feel of slight syncopation relative to the timeline. Regarding Figure 2, I imagine that practitioners would say that Baden plays with the ♪ ♪ ♪. ♪ ♪ ♪ ♪. timeline twice across the first four bars and then takes the final ♪. duration of the same timeline and propagates it forward across the following eight bars. Because the 16-span cycle length of the timeline and the 3-span cycle length of the ♪. propagation are unequal, the cyclic onsets of the timeline and the propagation will conflict through time. In my conversations, performers tend to consider such conflicts as engendering forward-driving tensions that install an expectation for eventual resolution back to the timeline.

Although these two procedures would be (I expect) immediately recognizable to more advanced performers of the style, there is virtually no literature to which an academic study can point to support the same claims. That is, at present we have no theory to support analytical studies of rhythmic structure in improvisation in timeline-based musics. By "timeline-based" music I mean to include any tradition (whether from the Caribbean, South America, Africa, etc.) whose rhythmic language centers around one or more memorable and pervasive rhythmic patterns. I can think of two reasons that can explain this gap between theory and practice. First, transcribing solos by master drummers, guitarists, etc. can be prohibitively challenging. In this respect, (unwritten) theories of timeline-based rhythmic improvisational procedures may lie largely with high-level performers because they are among the few who can access the source materials (the solos) in all their detail. Second, timeline theory first had to address questions of the ensemble/metric function, traditional contexts, and structural properties of the timelines themselves before engaging more complex and potentially more speculative questions of how timelines can be manipulated to structure solos.

I understand the discourse surrounding timelines since the term's coinage by Nketia [39] to consist of, broadly speaking, three overlapping phases. Timeline theories first emerged in the ethnographic work of Africanists studying various types of ensemble musics of sub-Saharan West and Central African ethnic groups.[2] The primary charge of this work was documentary, sometimes comparative (as in theorizing general principles of temporal organization in "African" music), and ethnographic. Here, questions of improvisation are often limited to considerations of variation in and between ensemble parts. As Africanist work brought more source materials into printed English-language view, music theorists began to investigate the structural-mathematical properties of timelines: their being prime generated, maximally even, maximally individuated, and analogous

[2] See [20][39][40][21][42][27][28][29][30][32][33][8][4][5][23][24]

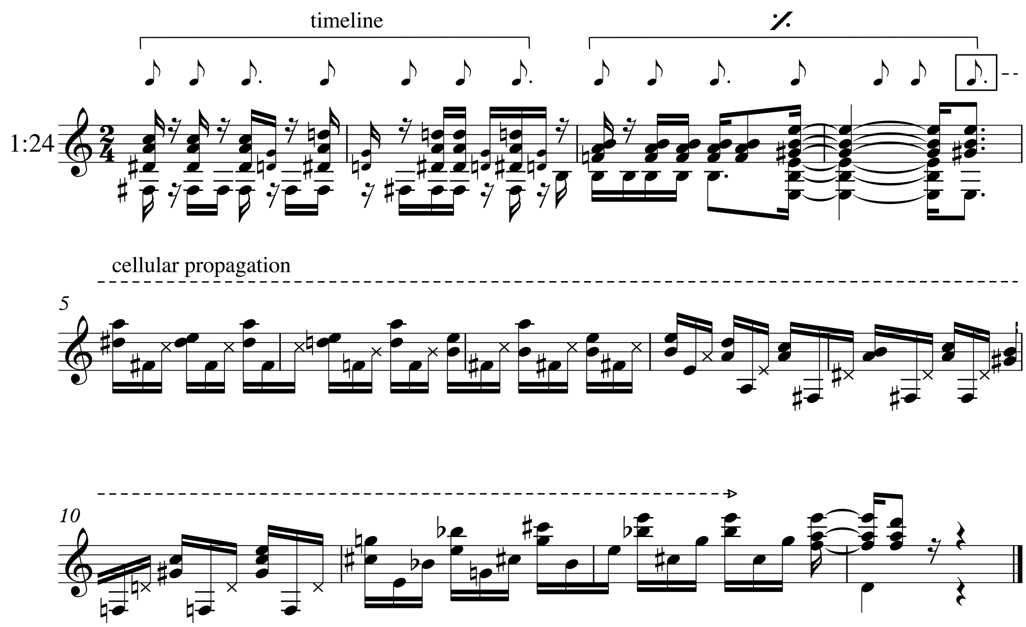

Figure 2: *Transcription of guitar part from Samba triste [45].*

to certain pitch collections to name a few [48][49][13][52] [53]. Such findings gradually stimulated other, more recent research lines investigating the possibility that timelines be heard as projecting non-isochronous meters, as in London [34] and Guerra [17]. The third and current phase has been one of critique and repertorial expansion. Agawu [1][2] investigates the tenability of claims made in music theoretic studies vis-à-vis the theory and language of native practitioners, and Gerstin [16] highlights the fact that most theories of Afro/diasporic rhythm and meter tend to extrapolate broadly from a data set that is essentially limited to only Ewe and Afro-Cuban traditions. In response, Gerstin and various Brazilianist scholars [22][50][43][55][14] have begun to expand the timeline studies corpus in their work on under-documented and contrasting Afro-diasporic traditions of the black Atlantic.

To my knowledge, only two prior studies–Anku [3] and Diaz [14]–have touched upon the topic of this paper. Anku argues that the complexity of a master drummer solo in "African music" can be analyzed in terms of two kinds of timeline set manipulation: "shifting set orientation" and "using successive sets."[3] The former would appear to be set rotation by another name, and the latter seems to be (or to at least include) a procedure similar to that which was discussed in the context of Figure 2 above–that is, improvising some rhythmic surface by weaving together various contiguous timeline duration subsets. Diaz studies how the timelines composed for the music of big band Orkestra Rumpilezz from Bahia, Brazil, can be related to traditional Afro-Bahian timelines by truncation/expansion and rotation transformations. While not explicitly their ambition, both articles suggest in their theoretical overlap the possibility of a coherent theory of rhythmic structure in improvisation in timeline-based music.

In this article, I define four techniques that together or separately can shape the rhythmic dimension of a solo out of a timeline. In an analysis of a passage from Baden Powell's guitar solo to the samba *Deixa*, which genre I consider to be a multiple-timeline-based music, I show

[3] Despite the breadth presupposed by Anku's stated repertory of "Africa music," it would seem by the examples he chooses that he has in mind specifically the dance-drumming traditions of Ghana.

how Baden transforms six consecutive cycles of an 8-span timeline across more than fifteen (2/4) bars. I end with a plot of Baden's trajectory through a transformational space, which informs a characterization of his improvisational strategy.

II. *Deixa* and samba timelines

The theme of *Deixa* is a 44-bar (in 2/4) parallel period in which each phrase is a sentence.[4] The 20-bar antecedent phrase comes to a half cadence in the tonic key of A-minor, and the parallel consequent 24-bar phrase comes to a full cadence. Baden's 1966 recording of *Deixa* [46] opens with a quasi-rubato setting of the 8-bar half cadence of the antecedent phrase of the period, presents the theme in tempo beginning at 0:15, and continues with a solo that extends across three cycles of the form. The first cycle begins at 1:11, the second at 2:04, and the third at 2:56. The performance ends in a brief fade out as Baden vamps on the tonic Am chord.

Deixa is a variant of the urban samba (*samba urbano*) native to Rio de Janeiro.[5] Like Guillot [18], I consider samba to be polymetric in the sense that in any given performance there are multiple timing frameworks available to both performer and listener.[6] The timing frameworks are the various timelines that have become canonic over the lifespan of samba through some process of spontaneous creation and imitation.[7] First among equals is the 8-span *surdo* timeline (see Figure 3), which enforces through its onsets and cycle length, respectively, the ♩ and 𝅗𝅥 pulses of samba's pure-duple metric background.[8] Surdos are large, cylindrical bass drums. The first hit is higher-pitched and often muted (represented with an "x" head in Figure 3); the second is lower-pitched and resonant. Even when they are not physically present, their timeline is assumed and internalized by knowledgeable performers and listeners. All other timelines in samba are learned and felt in terms of the surdo timeline. Arguably next in order of importance or prevalence is the 16-span timeline that Sandroni [50] calls the "Estácio paradigm," so named after an influential group of early samba composers and musicians from the Estácio neighborhood in Rio de Janeiro that defined the "new" and still current samba style (*estilo novo*). Note that only one of many possible instantiations is given in Figure 3.[9] Various instruments in a typical samba ensemble work with this timeline, including the *tamborim* (small, hand-held frame drum), *cuíca* (a friction drum), *violão* (guitar), and *cavaquinho* (ukulele-sized steel-string instrument). The asterisk

[4] Baden Powell composed the music for *Deixa* and poet/diplomat Vinícius de Moraes (1913-1980) wrote the lyrics.

[5] The adjective "urban" is commonly used to distinguish the samba that emerged and flourished in Rio de Janeiro from other "sambas," e.g., the Bahian samba-de-roda. Kubik and Pinto suggest that the number of types of Brazilian samba might be closer to a hundred [25, p. 153] but offer only a few examples by name: bossa nova, partido alto, samba-de-roda, samba chula, samba-de-viola [25, p. 156].

[6] Hudson [19] proposes a more subjective and dynamic way to metrically attend to music that may help to clarify what I mean by "polymetric" here. According to his theory, a samba timeline could serve as a "metering construction," a physically entrained or imagined background reference rhythm that shapes one's interpretation of a rhythmic surface. Performers and listeners can choose the same or different timelines-*cum*-metering constructions in any given moment and can switch freely among the same as they wish or as makes sense with respect to the exact character of the changing musical surface.

[7] For example, Didier [15] characterizes the change from old to new samba styles in the early 1920s as essentially a spontaneous and non-intellectual act of clever musical mischief. Over time and through a variety of reproductive means (e.g., radio broadcasting and federal government financial support), the exact rhythms of this mischief caught on and spread eventually coming to define the modern style. For more about this history, see Vianna [54], Moura [37], and McCann [36] among others.

[8] I measure spans of musical time in terms of the unit- or fastest-moving pulse–a 𝅘𝅥𝅯 pulse in notated samba. The term "pure-duple" comes from Cohn [9, p. 194]. It refers a span that can be represented as 2^n.

[9] Perhaps unlike other Afro/diasporic traditions, in samba there are no model timelines with accompanying sets of permissible variations. Rather, there seems to be an abstract possibility space determined by a small set of intuited rules that separates some rhythms from others as valid instantiations of a given timeline. See Guerra [17] for a theorization of this possibility space.

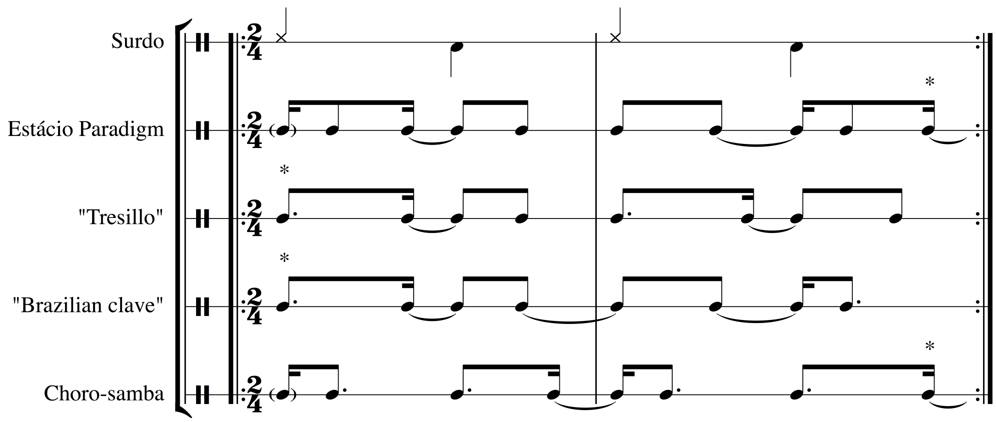

Figure 3: *Five samba timelines.*

in Figure 3 shows where the timeline phrase begins vis-à-vis a two-bar cycle of the surdo timeline. The third and fourth timelines are less ubiquitous. Third is what is known around the Western world as the *tresillo*, an 8-span timeline that is often carried by audience handclaps during a string of repeated choruses before the end of a samba song. Fourth is the 16-span, so-called "Brazilian clave." While this name is misleading ("Clave," whether by name or not, is not a well-formed and stable concept among practitioners of samba as it is among those of Afro-Cuban music.), it does appear to be well known and therefore expedient. The "Brazilian clave" is often heard in the rim shots of a *caixa* (snare) drum in a samba school or as part of a drum set in Brazilian jazz. Seldomly, it can be heard in a guitar part.[10]

To my knowledge, the fifth and last timeline printed in Figure 3 is nowhere described in the pedagogical and academic literature on either samba or choro–a sibling instrumental style-*cum*-genre that emerged in Rio de Janeiro around the same historical moment as samba. Despite this, one can easily hear this timeline emergent in many of the choro melodies written in a samba rhythm (sometimes designated "choro-samba" or "choro-sambado" or with such telling titles as "Receita de samba" or "Samba Recipe")[11] after the late-1950s, especially in those composed by Rio-native Jacob do Bandolim and his adherents, including Jonas da Silva (one of Jacob's accompanists) and Rossini Ferreira. Figure 4 suggests how one representative melody by each of these three composer-performers fits with what I will call the 8-span "choro-samba" timeline. The onset of the ♪ is in most cases suggested by its position at the end of a five-note run.[12] The onset of the first ♪. is suggested by its relatively isolated position, and the onset of the second ♪. is suggested by the implicit second and more resonant surdo beat mentioned above. The "downbeat" of the cycle, coincident with the onset of the ♪ duration, is induced by harmonic change (not shown). In both samba and choro, chords tend to change one sixteenth ahead of the notated bar.

Abstracted from their musical context, choro-samba and tresillo timelines are indistinguishable. In context, however, they differ by rotation and transposition. Different phrasings account for the difference in rotation. The onset of the ♪ duration in the tresillo generally functions anacrustically, while in the choro-samba timeline it acts as the cyclic "downbeat." Different alignments with the reference surdo timeline (indexed by the bar lines in the transcription) account for transposition differences. The downbeat of the tresillo coincides with the surdo's, while the downbeat of the

[10] The original 1961 recording of Jacob do Bandolim's choro-samba "Assanhado" begins in this way.

[11] For more on the history and conventions of this samba varietal, see Sève [51].

[12] See Povel and Essens [44, p. 415]. The authors report that in sequences of identical sounding tones those that are relatively isolated and those that begin and end a cluster of three or more tones become perceptually marked.

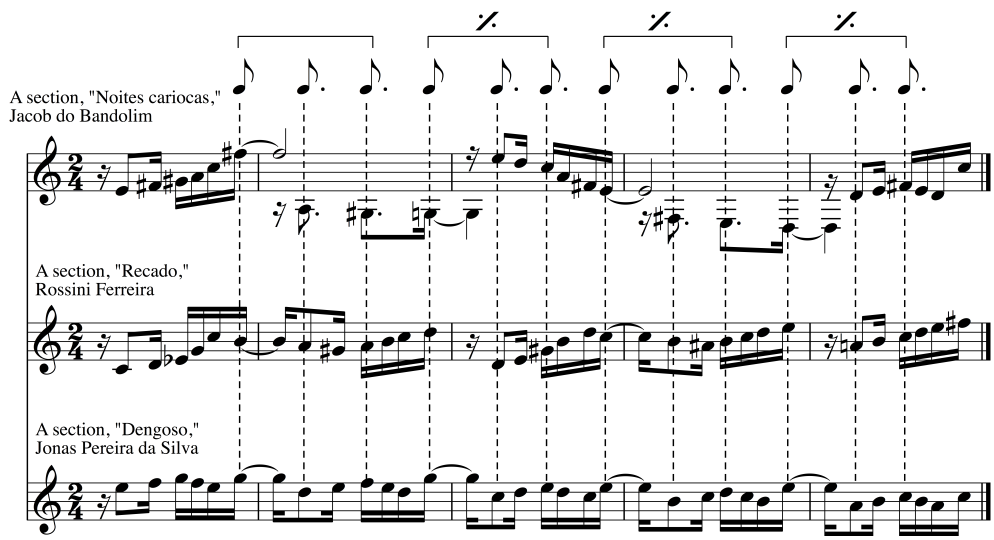

Figure 4: *Choro-samba melodies and the choro-samba timeline.*

choro-samba timeline anticipates the surdo's by one sixteenth.

III. Four techniques of timeline manipulation

In my analysis of the example passage from Baden's solo to *Deixa*, I recognize four techniques of improvised timeline manipulation: playing with the timeline, cellular propagation, augmentation/diminution, and rotation. While these four are all I need to explain my analysis completely, I in no way presume them to be exhaustive for all timeline-based improvisation. It is possible that other improvisations both within and without this style-genre may reveal other techniques.

The four techniques break down into two categories: transformational and non-transformational. The two belonging to the former category are augmentation (and its inverse, diminution) and rotation. Let an ordered set represent the interval structure of a timeline. For example, a tresillo could be represented with $(3, 3, 2)$. An augmentation transformation A_x applied to the set multiplies each interval in the set by the rational number subscript yielding the augmented product set. The double augmentation A_2 of the tresillo–what Biamonte [6] has called the "double tresillo" and Cohn [11] the "secondary rag"–is $(6, 6, 4)$. That is, $(3, 3, 2) * A_2 = (6, 6, 4)$. Rather than introduce another symbol, diminutions are represented with A fractions. For example, double diminution is written $A_{1/2}$. Applied to the double tresillo, $A_{1/2}$ yields the tresillo, or $(6, 6, 4) * A_{1/2} = (3, 3, 2)$.

Suppose an ordered timeline set were the dial of a combination lock. Rotations R_y turn the dial integral y notches clockwise or $-y$ notches counterclockwise. For example, $(3, 3, 2) * R_1 = (2, 3, 3)$. The modulus for y values depends on the cardinality of the timeline set. For example, a tresillo $(3, 3, 2)$ has a cardinality of three, so y is modulo-3. R_3 is the same as R_0 or not rotating at all. R_2 is the same as R_{-1}. Other timelines will have other cardinalities and therefore other rotation moduli.

Augmentation and rotation are commutative. Their order of application to the origin set does not affect the ultimate product set. In the analysis of Baden's solo, they effectively happen at the same time. Ordering their elapse would be arbitrary.

The two non-transformation techniques were discussed briefly at the head of this paper: playing with the timeline and cellular propagation. For the purposes of this paper, "to play with a timeline" means to play some figure that–via various cues of metric induction (after Lerdahl and

Jackendoff [26])–suggests a reduced rhythmic (or quasi-metric[13]) structure that is isomorphic to that timeline. Cellular propagation is where the performer takes one or more of the constituent durations from a timeline – in the same order, such that they would be recognizable as a contiguous subset to the original – and repeatedly propagates them forward through time.

IV. Six cycles of the choro-samba timeline

Figure 5 transcribes Baden's solo over the entire 24-bar consequent phrase in his second time through the form. This transcription is an edited version of the one by Magalhães [35] and follows a long-standing convention of notating samba in 2/4, which practice, if not owing to, at least reminds us of the primacy of the surdo timeline. My analysis will focus on mm.7-21, which passage I hear as six complete cycles (boxed and numbered) of the choro-samba timeline under various transformations.

In cycles 1 and 6 of Figure 5, Baden plays with the choro-samba timeline. The rhythmically identical phrases of 1 and 6 suggest the three onsets of the ♪ ♪. ♪. choro-samba timeline in various ways. The relatively isolated position of the chord change (one sixteenth before the bar line) suggests the first onset. The inception of the subsequent five-note run suggests the second. The final bass note (beat two in mm.7 and 21) suggests the third.

In cycles 2 and 3, Baden propagates the final ♪. cell of the choro-samba timeline from cycle 1 in a three-part treble/bass/midrange arpeggio figure twelve times over before this ♪. grouping of the sixteenth yields to that of the ♪ to be used in cycles 4 and 5. In cycle 2, the ♪. propagation provides the basis for a triple augmentation of the choro-samba timeline in rotation. That is, Baden transforms 8-span $(2,3,3)$ into 24-span $(9,6,9)$ by A_3R_1. The textural change from the block chords of m. 7 to the ♪. arpeggio articulates the first onset of transformed timeline $(9,6,9)$. The change from subdominant to dominant in the key of A-minor (m. 9) articulates the second. The descending-second G5-F5 melodic parallelism–downbeat of m. 9 and two sixteenths before m. 10–articulates the third onset. See Figure 6, which re-beams the first three choro-samba timeline cycles of Figure 6. Boxed numbers below the staff indicate which cycle is under way. The largest and boldened numbers above the staff identify the durations of the (transformed) choro-samba cycles. The lower row of smaller numbers above the staff shows how the ♪. propagation groups the unit 𝅘𝅥𝅯 pulse and how the propagated cell in turn participates in the timeline augmentations.

Cycle 3 transforms 24-span $(9,6,9)$ into 16-span $(6,6,4)$ by $A_{2/3}R_1$. The arrival of tonic Am harmony articulates the first onset of $(6,6,4)$. The F5 melodic parallelism–one sixteenth before m. 11 and three sixteenths before m. 12–articulates the second. The parallel C5s of m. 12, coincident with the change from ♪. to ♪ grouping of the unit pulse, articulate the third onset.

To build cycle 4, Baden propagates both the lower-level 2-unit and higher-level 6-unit durations from cycle 3. This is shown in the layered rows of smaller numbers above the staff in Figure 7. The largest and boldened numbers above the same staff show how these two simultaneous and hierarchic propagations allow for a triple augmentation of cycle 3 into cycle 4. That is, Baden transforms 16-span $(6,6,4)$ into 48-span $(18,18,12)$ by A_3 while he holds the rotation fixed. The 2-unit propagation can be heard in the parallel chord-melodic C5s, which continue the neighbor-note motive C5-B4 begun in the tail of the previous cycle. By m. 17, the parallel C5s of the motive are modified to parallels D5s. Bass notes further confirm the 2-propagation as three out of every four notes align with the ♪ propagation. The 6-unit propagation requires more explanation. Cycle 4 consists of back-to-back five-note runs in the treble blocks separated by a single sixteenth-note rest. These runs are located with brackets below the staff in Figure 7. One of the onsets from the

[13] Cohn [12] entertains the possibility of a cognitive/perceptual framework that sits somewhere between a rhythm and a meter. This follows various similar suggestions made by Butler [7], Osborn [41], and Cohn [11].
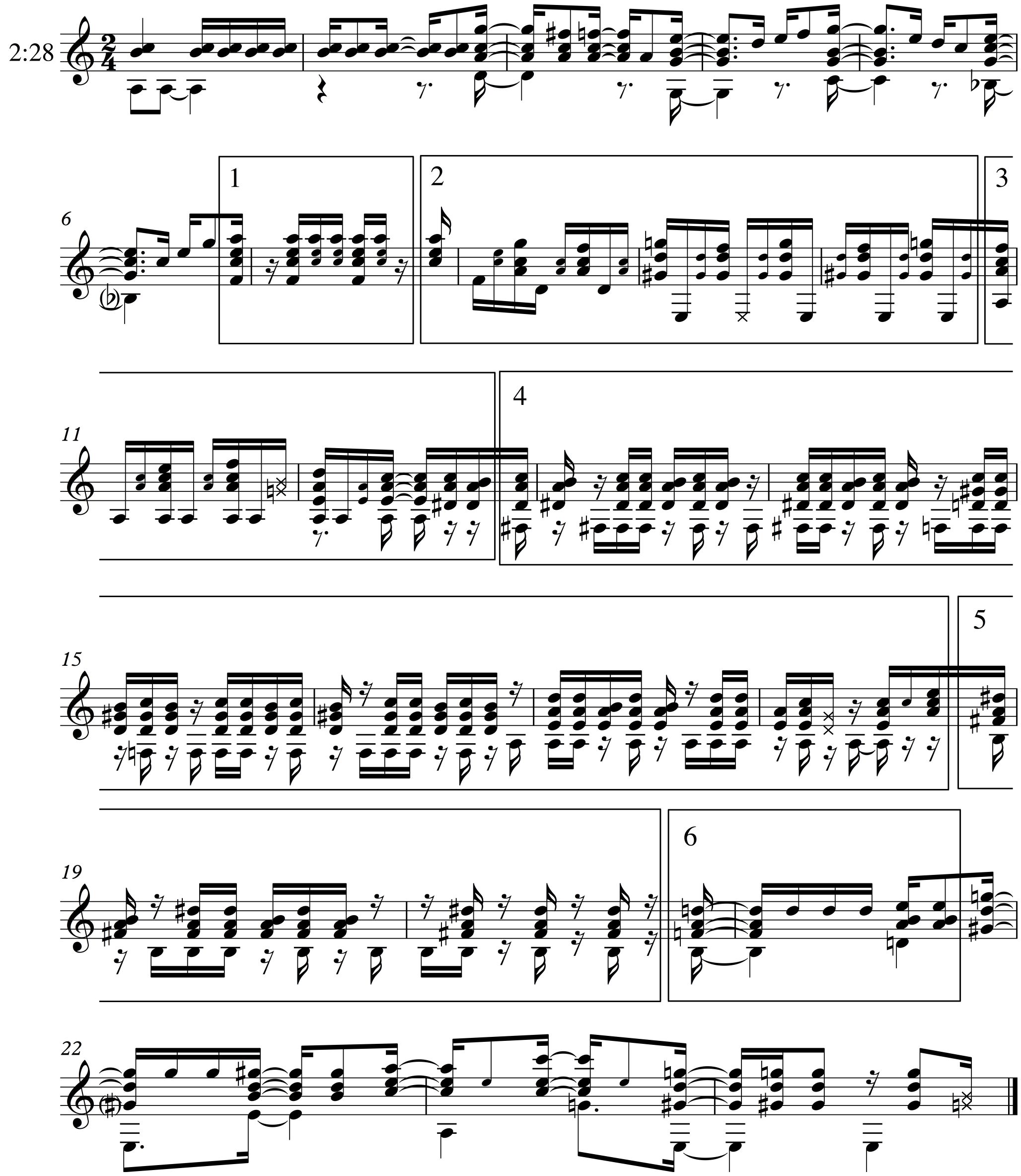
Figure 5: *Transcription of guitar part from Deixa [46].*

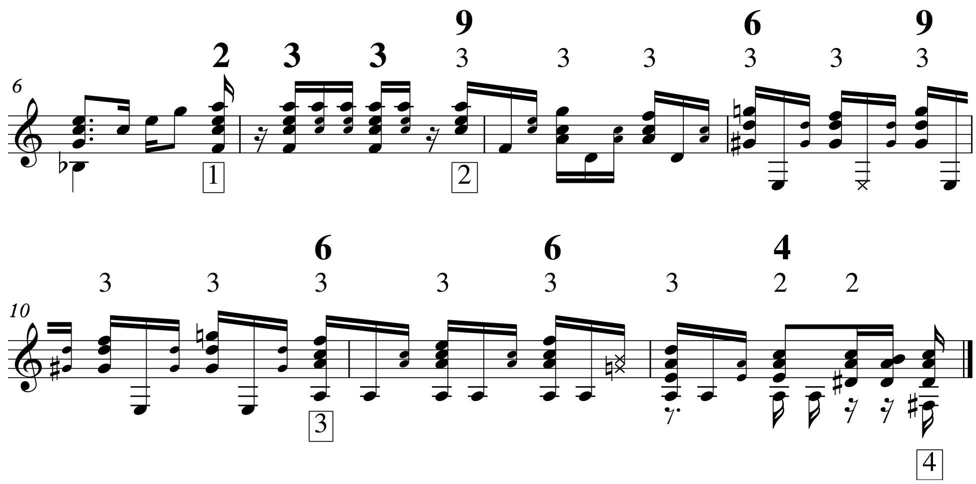

Figure 6: *(Transformed) choro-samba cycles 1-3.*

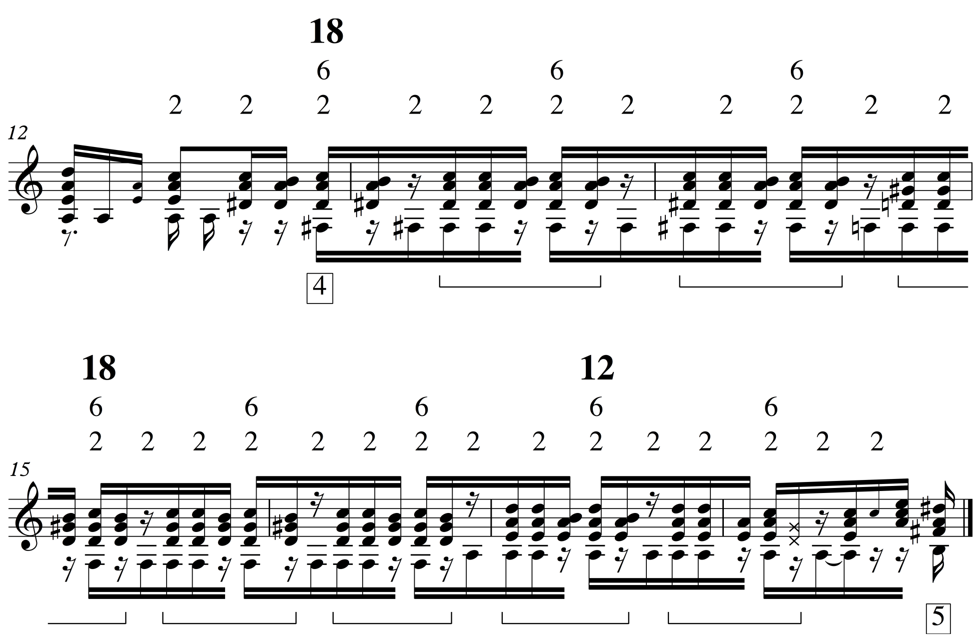

Figure 7: *Transformed choro-samba cycle 4.*

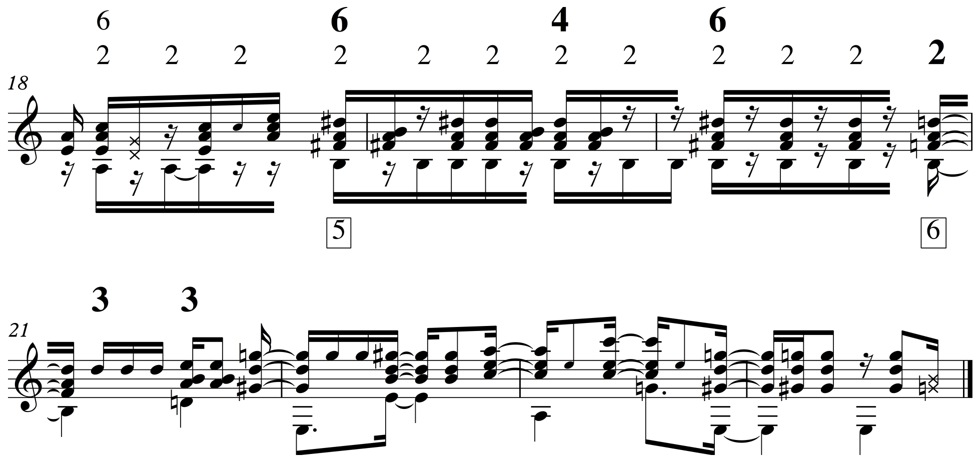

Figure 8: *(Transformed) choro-samba cycles 5-6.*

run will attract a "downbeat" status by virtue of the run's being cyclic, and it is preferable that such a downbeat should also coincide with an onset of the 2-unit ♪ propagation. Two onsets from each five-note run coincide with ♪ onsets, the second and fourth. There is a clear precedent in which to prefer. Cycle 1 (refer to m. 7 in Figure 6) also features an isolated five-note run.[14] There, the first and fourth onsets were relatively strong to the others, and the latter coincided with the surdo's low, resonant tone (whether imagined or real). Trained by cycle 1, I prefer to hear the fourth onset of each five-note cluster in cycle 4 as the cyclic downbeat, and this induces the sense of a 6-unit propagation. With clear perceptions of the 2- and 6-unit propagations, it is easy to hear the onsets of $(18, 18, 12)$. Each onset is articulated by a chord change, first from B7 to G♯dim7 (second sixteenth of m. 15) and then from G♯dim7 to tonic Am (fourth sixteenth of m. 17), where each change is preceded by a three-sixteenth-note anacrusis.[15]

Cycle 5 continues the 2-propagation in various ways: initially by the parallel D♯5s (a further development of the original C5-B4 neighbor-note motive begun in m. 12), and then by the isolated onsets of B7. Here, Baden organizes the 2-unit propagation into $(6, 4, 6)$, transforming the previous 48-span $(18, 18, 12)$ by $A_{1/3}R_{-1}$. The first two onsets of $(6, 4, 6)$ are projected by our existing hearing of the cyclic five-note run, and the third onset is suggested by the initiation (second sixteenth of m. 20) of the figural change to isolated block chords.

In cycle 6, Baden returns to his phrasing of the choro-samba timeline from cycle 1, transforming $(6, 4, 6)$ into $(2, 3, 3)$ by $A_{1/2}R_{-1}$. Once again playing with the timeline, Baden resolves the tension accumulated during his transformational journey and foreshadows the form's structural authentic cadence, which will arrive two bars later in m. 23.

V. Baden's solo in transformational space

My analysis identified six consecutive cycles of the choro-samba timeline connected by five transformations:

$$(2, 3, 3) * A_3 R_1 =$$
$$(9, 6, 9) * A_{2/3} R_1 =$$

[14] Indeed, mm. 13-18 could be further characterized as a propagation of the five-note run from cycle 1.
[15] Such anacruses are characteristic in choro and samba.

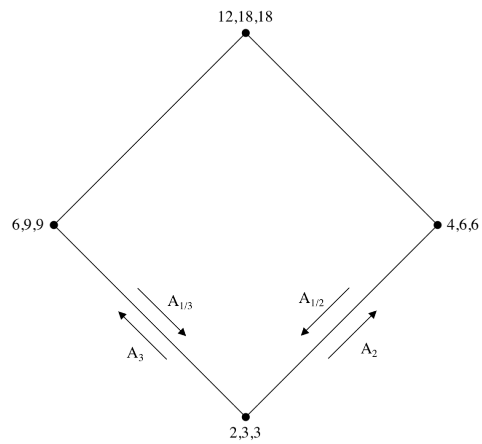

Figure 9: *Choro-samba augmentation lattice.*

$$\begin{gathered}(6,6,4) * A_3 = \\ (18,18,12) * A_{1/3}R_{-1} = \\ (6,4,6) * A_{1/2}R_{-1} = \\ (2,3,3)\end{gathered}$$

In this section, I will summarize Baden's timeline-based solo as movement in a transformational space.

While theoretically unlimited, Baden's augmentations (and diminutions) are based on factors of 2 and 3 only. For example, A_3 takes cycle 1 to cycle 2, $A_{2/3}$ takes 2 to 3, and so on. In this way, augmentation transformations of the $(2,3,3)$ timeline can be represented on a two-dimensional lattice analogous to Cohn's [10] ski-hill graph. See Figure 9. Nodes represent (transformed) choro-samba timelines. SW-NE edges connect nodes related by transformation A_2; NW-SE edges connect those related by A_3. The lattice is infinitely extensible to the NW and NE. However, because Baden's most augmented timeline is a rotation of $(12,18,18)$, I need only the diamond of Figure **??** for my analysis.

Integrating the R transformation into the space means that every timeline node must connect not only to its A_2 (or $A_{1/2}$) and A_3 (or $A_{1/3}$) transformations but also to its R_1 and R_{-1} transformations. Thus, every node in the expanded A and R transformation graph will have four incident edges. In other words, any timeline node is one transformation away from four other distinct timeline nodes. Figure 10 suggests a possible visualization for the space. The three-axis legend to the left shows which transformations apply to which directions in the space. Importantly, note that the four nodes of the highest diamond–$(3,2,3)$, $(9,6,9)$, $(18,12,18)$, and $(6,4,6)$–connect to those respective nodes of the lowest diamond by a single R_1 edge. For example, $(3,2,3)$ from the top diamond connects to $(3,3,2)$ from the bottom diamond by R_1. To make the graph easier to read, I left these edges out. Including them would take the cube-shaped graph and wrap it around a cylindrical mold to connect the top and bottom planes.

Two caveats should be made about the space. First, this space is not meant to be extrapolated outside of its original context, whether to Baden's other samba improvisations, samba more

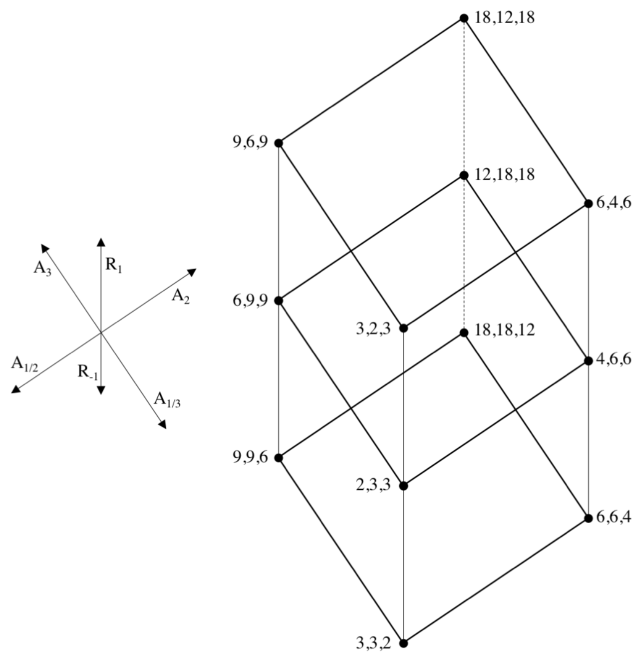

Figure 10: *Choro-samba augmentation-rotation space.*

broadly, or even Afro/diasporic music. The space is relevant to the analyzed passage only. Spaces for other improvisations inside or outside of samba could be similar or different; they could use the same two transformations or not. There is substantial work to be done ahead of any generalizing ambitions. Second, the construction of the graph suggests that we interpret the distance between any two timeline nodes as the minimal number of edges required to connect the two. This implicitly assumes that A_2, A_3, and R_1 transformational distances are all equal-weighted. Whether they are or not in perceptual terms I leave as an open question that is beyond the scope of this study.

Figure 11 plots Baden's trajectory through the transformational space. The gray-highlighted nodes are the five distinct timelines that Baden visits, and the dashed arrows – beginning and ending with $(2, 3, 3)$ – show the order in which he visits them. The plot helps visualize how the particular sequence of Baden's transformations achieves a sense of rhetorical balance in two related ways.

First, Baden's movement through the planes is palindromic.[16] He starts with $(2, 3, 3)$ on the middle plane and moves up to $(9, 6, 9)$ on the upper plane. From there, Baden moves up again, circling around in Figure 11 to $(6, 6, 4)$ on the bottom plane. He then stays on the bottom plane moving to $(18, 18, 12)$ before moving back down to $(6, 4, 6)$ on the top plane. Finally, Baden moves down again to $(2, 3, 3)$ resting on the middle plane. In short, Baden's palindromic movement follows the order middle-upper-lower-lower-upper-middle.[17]

[16]Recall that movement among the horizontal planes of Figure 11 is driven by the R transformation.

[17]Following the example set in Lewin [27] and carried forward by Cohn [10] and Murphy [38], it is tempting to model Baden's trajectory in Figure 11 as analogous to the three-termed dualism of subdominant-tonic-dominant, insofar as two nodes in the upper plane and two nodes in the lower plane flank the "tonic" $(2, 3, 3)$ of the middle plane. However

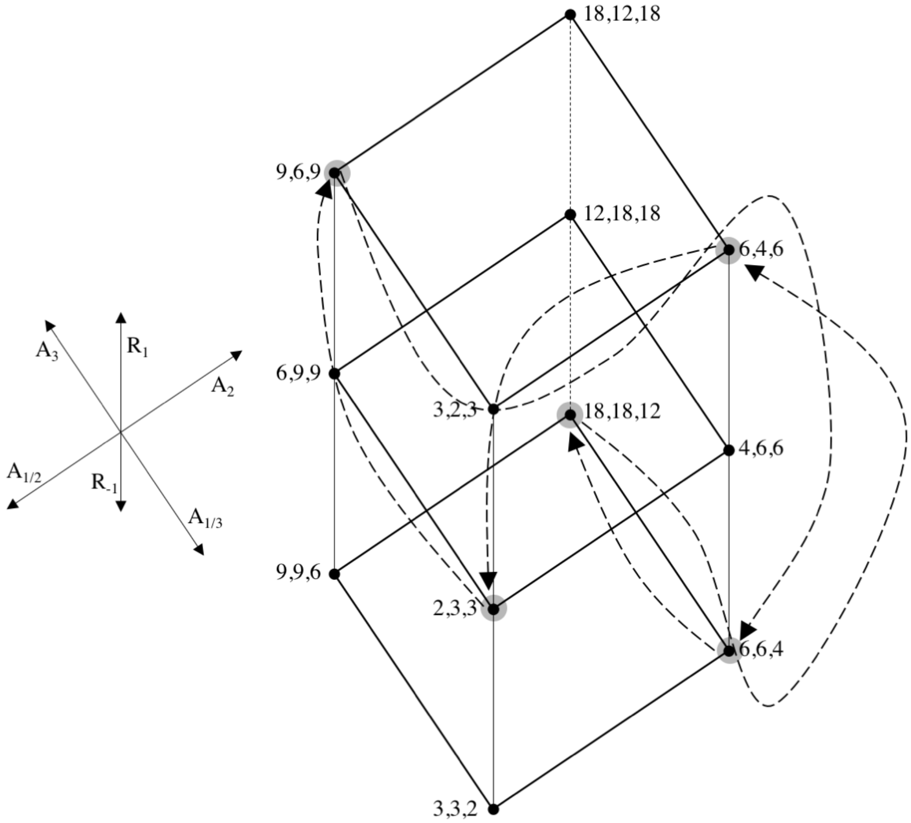

Figure 11: *Baden's trajectory through choro-samba augmentation-rotation space.*

A second way in which Baden's timeline-based improvisation achieves rhetorical balance is through two repetitions of a question-answer script. Baden's initial move away from $(2,3,3)$ to $(9,6,9)$ on the triple side of the upper plane seems to pose a question, which he then seems to answer by moving to $(6,6,4)$ on the duple side of the lower plane. Baden then poses another question, moving to $(18,18,12)$ in the duple-triple corner of the bottom plane, which he then answers moving to $(6,4,6)$ on the duple side of the upper plane.

VI. Closing remarks

In this study, I defined and analytically demonstrated four techniques that can meaningfully shape the rhythm of a solo out of a timeline. Two were transformational: augmentation/diminution and rotation; and two were non-transformational: playing with the timeline and cellular propagation. These techniques were all I needed to come to a more profound understanding of a single passage from a single recording. Zooming out to the larger world of Afro/diasporic music, however, many questions remain. Would analysis of other master instrumentalists, styles, and traditions, for example, suggest other techniques? Or are these the only four? And if they were the only four, why would there not be more? Are there some principles of cognition and perception which limit the range of techniques? Or perhaps we would find that these other instrumentalists, styles, and traditions emphasize some techniques while de-emphasizing or even excluding others. What would account for such differences? Culture? Language? Aesthetics? Instrumentation? Consider the fact that Anku [3] did not seem to need an augmentation transformation, whereas that was one of the primary drivers in my analysis of Baden's solo to *Deixa*. Could we assume that the rhythmic dimensions of solos developed on harmonic instruments have different capabilities than those developed on percussive instruments and therefore different tendencies? The list of questions easily goes on. And if this study has raised more questions than it has answers then it has done its job. As the field of music theory becomes more inclusive and as scholars continue to diversify the scope of their practice and study repertories, we are in an ever-better position to amass analytical data enabling more educated engagement with these and other questions concerning Afro/diasporic repertories. Moreover, as we come to a better understanding of how master soloists construct their solos to timeline-based musics, we will also be enrichening a larger, similarly underexplored discourse surrounding the question of what separates the intelligible from the unintelligible in a given style, whether it be Afro/diasporic or otherwise.

References

[1] Agawu, Kofi. (2003). *Representing African Music*: Post-Colonial Notes, Queries Positions. Routledge.

[2] Agawu, Kofi. (2006). Structural Analysis or Cultural Analysis? Competing Perspectives on the 'Standard Pattern' of West African Rhythm. *Journal of the American Musicological Society*, Vol. 59, pp. 1–46.

[3] Anku, Willie. (2000). Circles and Time: A Theory of Structural Organization of Rhythm in African Music. *Music Theory Online*, Vol. 6.

suggestive, the analogy is imperfect (if not misleading) for two reasons. First, the assignment of subdominant and dominant to upper and lower planes (or vice versa) would seem to be arbitrary. Second, neither assignment, whether T-S-D-D-S-T or T-D-S-S-D-T, would correctly mirror the standard asymmetric cadence model of T-S-D-T.

[4] Arom, Simha. (1989). Time Structure in the Music of Central Africa: Periodicity, Meter, Rhythm and Polyrhythmics. *Leonardo* Vol. 22, pp. 91–9.

[5] Arom, Simha. (1991). *African Polyphony and Polyrhythm*. Translated by Martin Thom, Barbera Tuckett, and Raymond Boyd. Cambridge University Press.

[6] Biamonte, Nicole. (2014). Formal Functions of Metric Dissonance in Rock Music. *Music Theory Online*, Vol. 20

[7] Butler, Mark. (2006). *Unlocking the Groove: Rhythm, Meter, and Musical Design in Electronic Dance Music*. Indian University Press.

[8] Chernoff, John Miller. (1981). *African Rhythm and African Sensibility*: Aesthetics and Social Action in African Social Idioms. University of Chicago Press.

[9] Cohn, Richard. (1992). The Dramatization of Hypermetric Conflicts in the Scherzo of Beethoven's Ninth Symphony. *19th-Century Music*, Vol. 15, pp. 188–206.

[10] Cohn, Richard. (2001). Complex Hemiolas, Ski-Hill Graphs and Metric Spaces. *Music Analysis*, Vol. 20, pp. 295–326.

[11] Cohn, Richard. (2016). A Platonic Model of Funky Rhythms. *Music Theory Online*, Vol. 22.

[12] Cohn, Richard. (2018). *Meter*. In Oxford Handbook of Concepts in Music Theory, ed. Alexander Rehding and Steven Rings. Oxford University Press.

[13] Demaine, Erik, et al. (2009). The Distance Geometry of Music. *Computational Geometry*, Vol. 42, pp. 429–54.

[14] Diaz, Juan Diego. (2017). Experimentations with Timelines: Strategies of Rhythmic Complication in Afro-Bahian Jazz. *Analytical Approaches to World Music*, Vol. 6, pp. 1–34.

[15] Didier, Caôla. (1984). O samba que veio do Estácio. *O Catacumba*, Vol. 1, p. 3.

[16] Gerstin, Julian. (2017). Comparisons of Africa and Diasporic Rhythms: The Ewe, Cuba, and Martinique. *Analytical Approaches to World Music*, Vol. 5, pp. 1–90. [

[17] Guerra, Stephen Paul. (2018). *Expanded Meter and Hemiola in Baden Powell's Samba-Jazz*. Dissertation (Ph.D in Music), Yale University.

[18] Guillot, Gérald. (2016). Hypothesis of Polymetric Structures in Afro-Brazilian Music. *Fourth International Conference on Analytical Approaches to World Music AAWM*. The New School, New York.

[19] Hudson, Stephen S. (2019). *Feeling Beats and Experiencing Motion*: A Construction-Based Theory of Meter. Dissertation (Ph.D in Music), Northwestern University.

[20] King, Anthony. (1960). Employments of the 'Standard Pattern' in Yoruba Music. *African Music*, Vol. 2, pp. 51–4.

[21] Koetting, James. (1970). Analysis and Notation of West African Drum Ensemble Music. *Selected Reports in Ethnomusicology*, Vol. 1, pp. 115–46.

[22] Kubik, Gerhard. (1979). Angolan Traits in Black Music, Games and Dances of Brazil: A Study of African Cultural Extension Overseas. *Junta de Investigações do Ultramar*.

[23] Kubik, Gerhard. (2010a). *Theory of African Music*—Vol. I. University of Chicago Press.

[24] Kubik, Gerhard. (2010b). *Theory of African Music*—Vol II. University of Chicago Press.

[25] Kubik, Gerhard, and Tiago de Oliveira Pinto. (2008). African Common Denominators across the South Atlantic: A Conversation. *AfricAmericas. Itineraries, Dialogues, and Sounds*, pp. 153–63.

[26] Lerdahl, Fred, and Ray Jackendoff. (1983). *A Generative Theory of Tonal Music*. MIT Press.

[27] Lewin, David. (1981). On Harmony and Meter in Brahms's Opus 76 #8. *19th-Century Music*, Vol. 4, pp. 261-–5.

[28] Locke, David. (1978). *The Music of Atsiagbeko*. Dissertation (Ph.D in Music), Wesleyan University.

[29] Locke, David. (1982). Principles of Offbeat Timing and Cross-Rhythm in Southern Ewe Dance Drumming. *Ethnomusicology*, Vol. 26, pp. 217–46.

[30] Locke, David. (1998). *Drum Gahu*. White Cliffs Media.

[31] Locke, David. (2009). Simultaneous Multidimensionality in African Music. *African Music*, Vol. 8, pp. 8-–37.

[32] Locke, David. (2010). Yewevu in the Metric Matrix. *Music Theory Online*, Vol. 16.

[33] Locke, David. (2011). The Metric Matrix: Simultaneous Multidimensionality in African Music. *Analytical Approaches to World Music*, Vol. 1, pp. 48–72.

[34] London, Justin. (2012). *Hearing in Time*. 2nd ed. Oxford University Press.

[35] Magalhães, Alain Pierre Ribeiro de. (2008). *Baden Powell*: Livro de partituras. Gryphus.

[36] McCann, Bryan. (2004). *Hello, Hello Brazil*: Popular Music in the Making of Modern Brazil. Duke University Press.

[37] Moura, Roberto M. (2004). *No princípio, era a roda*: um estudo sobre samba, partido-alto e outros pagodes. Rocco.

[38] Murphy, Scott. (2007). On Metre in the Rondo of Brahms's Op. 25. *Music Analysis*, Vol. 26, pp. 323–53.

[39] Nketia, J.H. Kwabena. (1963). *African Music in Ghana*. Northwestern University Press.

[40] Nketia, J.H. Kwabena. (1974). *The Music of Africa*. W.W. Norton.

[41] Osborn, Brad. (2014). Kid Algebra: Radiohead's Euclidian and Maximally Even Rhythms. *Perspectives of New Music*, Vol. 52, pp. 1-–25.

[42] Pantaleoni, Hewitt. (1972). Three Principles of Timing in Anlo Dance Drumming. *African Music*, Vol. 5, pp. 50–63.

[43] Pinto, Tiago de Oliveira. (1999/2000/2001). As cores do som: Estruturas sonoras e concepção estética na música afro-brasileira. *África*, Vol. 22-23, pp. 87–109.

[44] Povel, Dirk-Jan, and Peter Essens. (1985). Perception of Temporal Patterns. *Music Perception*, Vol. 2, pp. 411–40.

[45] Powell, Baden. (1964). *Le Monde Musical de Baden Powell*. Barclay.

[46] Powell, Baden. (1966). *Tempo feliz*. Forma.

[47] Powell, Baden. (1972). *É de lei*. Phonogram.

[48] Pressing, Jeffrey. (1983). Cognitive Isomorphisms between Pitch and Rhythm in World Music. *Studies in Music*, Vol. 17, pp. 38–61.

[49] Rahn, Jay. (1996). Turning the Analysis Around: Africa-Derived Rhythms and Europe- Derived Music Theory. *Black Music Research Journal*, Vol. 16, pp. 71–89.

[50] Sandroni, Carlos. (2001). *Feitiçco Decente, Transformações do samba no Rio de Janeiro*. Editora UFRJ.

[51] Sève, Mário. (2016). O choro no estilo sambado: Padrões rítmicas e fraseado musical. *Debates*, Vol. 17, pp. 219–249.

[52] Stover, Chris. (2009). *A Theory of Flexible Spaces for Diasporic African Music*. Dissertation (Ph.D in Music), University of Washington.

[53] Toussaint, Godfried. (2013). *The Geometry of Musical Rhythm*: What Makes a "Good" Rhythm Good?; CRC Press, Taylor & Francis Group.

[54] Vianna, Hermano. (1999). *The Mystery of Samba*: Popular Music and National Identity in Brazil. Edited and translated by Charles Chasteen. The University of North Carolina Press.

[55] Vicente, Alexandre. (2012). Moacir Santos, seus ritmos e modos: 'Coisas' do Ouro Negro. Thesis (Master in Music), UDESC.

P_{BACH} and Musical Transformations

Carlos de Lemos Almada *
Federal University of Rio de Janeiro (UFRJ)
carlosalmada@musica.ufrj.br
Orcid: 0000-0001-5888-2868

Abstract: *Departing from a specific passage of a text by Douglas Hofstadter which addresses recursive algorithmic application, this paper reformulates its basic structure, by using formal tools under a transformational-musical perspective. The last section of the study proposes some discussion about possible expansions and generalization from the obtained results.*

***Keywords**: Recursion. Ordered Sets. Operations and Functions. Transformational Theory.*

I. Introduction

In the dialogue that opens chapter VI (entitled "Canon by Intervallic Augmentation") of Douglas Hofstadter's well-acclaimed, Pulitzer-prized *Gödel, Escher, and Bach* ([1], pp. 153–157), the character Tortoise tries to convince his friend Achilles that two different songs played in his phonograph can be "coded inside the same record": the first "based on the famous old tune B-A-C-H", and a "totally different melody (...) C-A-G-E". This is the kern of Tortoise's argumentation:

> **Tortoise:** (...) What do you get if you list the successive intervals in the melody B-A-C-H?
> **Achilles:** Let me see. First it goes down one semitone, from B to A (where B is taken the German way); then it rises three semitones to C; and finally it falls one semitone, to H. That yields the pattern: $-1, +3, -1$.
> **Tortoise:** Precisely. What about C-A-G-E, now?
> **Achilles:** Well, in this case, it begins by falling three semitones, then rises ten semitones (nearly an octave), and finally falls three more semitones. That means the pattern is: $-3, +10, -3$.
> (...)
> **Tortoise:** They have exactly the same "skeleton", in a certain sense. You can make C-A-G-E out of B-A-C-H by multiplying all the intervals by 3⅓, and taking the nearest whole number.
> (...)
> **Tortoise:** The melody consisted of enormously wide intervals, and went B-C-A-H. (...) It can be gotten from the CAGE pattern by yet another multiplication by 3⅓, and rounding to whole numbers. (...)

*I am very grateful to Dr. Douglas Hofstadter for reading this paper, for his kind words about it, and for give me his personal permission for using an excerpt of his book.

Hofstadter's algorithm, described by Tortoise, could be succinctly expressed as follows:

1. Translate the sequence "BACH" into pitches (adopting the most compact disposition), using German musical notation. Let us name as w the pitch sequence. Thus,
   w = $<B\flat_4, A_4, C_5, B\natural_4>$[1] ;

2. Extract the melodic intervals from w, adopting as unity the semitone. Use the minus signal to indicate descending intervals, and the plus signal to ascending intervals. Let **INT** be the function used to determine intervals between sequential pitches. Let variable x represent the sequence of intervals. Thus,
   x = **INT**(w) = $< -1, +3, -1 >$;

3. Multiply any element of x by 3⅓, and approximate the result in case of fractioned number. Name y this product. Thus,
   y = x × 3⅓ = $< -1, +3, -1 >$ × 3⅓ = $< -3.33...$, $+9.99...$, $-3.33... > \approx < -3, +10, -3 >$;

4. Apply step 3 to y, and let z represent the resulted sequence. Thus,
   z = y × 3⅓ = $< -3, +10, -3 >$ × 3⅓ = $< -9.99...$, $+33.33...$, $-9.99... > \approx < -10, +33, -10 >$;

Figure 1 presents the pitch structure of referential "melody" BACH and of the two recursive transformations.

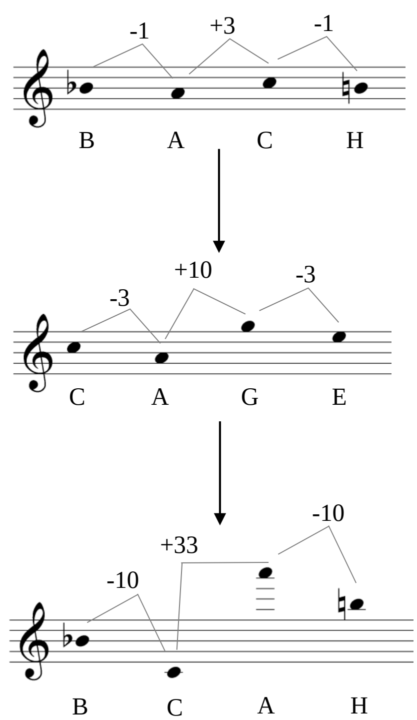

Figure 1: *Transformation of "melodies", according to Hofstadter's algorithm: BACH into CAGE, and CAGE into BCAH.*

[1] Angled brackets indicate that their content is ordered.

I was deeply impressed by this BACH/CAGE dialogue since the first time I read it (as well as the whole book, of course), but only recently, involved with a research based on Transformational and Group theories,[2] I started to conjecture if I could reformulate Hofstadter's clever idea using pitch classes instead of intervals, and if it would be possible to formalize more strictly the transformations.
This brief article was born as an attempt in this direction. The next section defines sets, functions, and operations needed to the transformations, which are implemented in section III, reaching the central aim of the article. An additional section explores the results obtained, attempting to propose some expansion and generalization.

II. Definitions

(1) Let X be a set formed by German symbolic representation for musical notes (upper-case letters or group of upper-case/lower-case letters).[3]

$$X = \{C, Cis, D, Es, E, F, Fis, G, As, A, B, H\}$$

(2) Let Y be a set formed by the twelve pitch classes, or else, Y is isomorphic to $\mathbb{Z}_{12}$.

$$Y = \{0,1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$$

(3) Let f be a bijective function that maps same-order members of X onto Y | f: X $\rightarrow$ Y.

(4) Let f^{-1} be the function inverse of f, that maps same-order members of Y onto X | f^{-1}: Y $\rightarrow$ X.

Figure 2 presents sets X and Y, exemplifying the actions of functions f and f^{-1}.

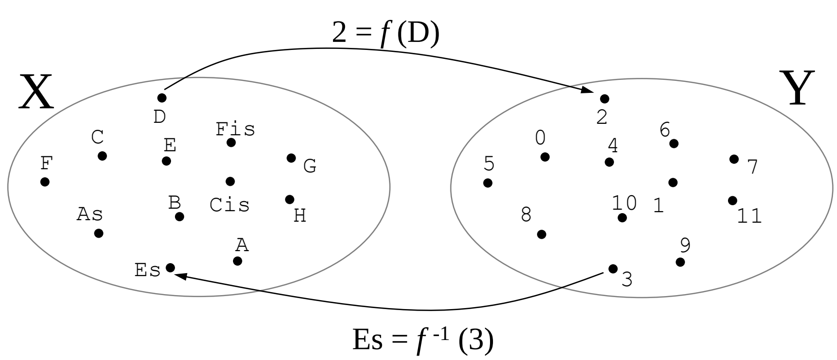

Figure 2: *Sets X and Y and two examples of application of functions f and f^{-1}.*

(5) Let g be the polynomial quadratic function $g(x) = x^2 + 3$ | g: $\mathbb{Z} \rightarrow \mathbb{Z}$.

(6) The operation *retrogradation*, labelled as **R**, flips leftwards the content of a given ordered set S. Ex.: Let set S = <i, j, k, ..., m, n>, then **R**(S) = <n, m, ..., k, j, i>.

[2]For some of my references, see, for example, [2], [3], [4], [5], [6], and [7].
[3]Suffixes "is" and "s" stand for, respectively, the accidents "sharp" and "flat".

(7) The operation *rotation*, labelled as $\mathbf{ROT}_t$, permutes t times (t is an integer greater than zero) the content of a given ordered set S, i.e., it sends at each application the first member of S to the last position of S, keeping unaltered the order of the remaining members.
Ex.: Let set S = <i, j, k, ..., m, n>, then $\mathbf{ROT}_2$(S) = **ROT**(**ROT**(S)) = **ROT**(<j, k, ..., m, n, i>) =
= <k,..., m, n, i, j>.

(8) The operation *extraction*, labelled as $\mathbf{EXT}_{t:u}$, extracts from a given ordered set S a subset formed by contiguous members, delimited by the $t^{th}$and the u^{th} members of S.
Ex.: Let set S = <i, j, k, ..., m, n>, then $\mathbf{EXT}_{2:3}$(S) = T = <j, k>.

(9) The operation *merging*, labelled as **MRG**(S, T), concatenates two ordered sets S and T, keeping unaltered their internal order, forming an ordered superset ST = <S, T>.
Ex.: Let sets S = <i, j, k, ..., m, n> and T = <o, p, q>, then **MRG**(S, T) = ST = <i, j, k, ..., m, n, o, p, q>.

(10) The operation *modulo12*, labelled as $\mathbf{MOD}_{12}$, maps members of set $\mathbb{Z}$ onto members set Y | $\mathbf{MOD}_{12}$: $\mathbb{Z} \rightarrow \mathbb{Z}_{12}$.
Ex.: Let set S = {2, 25, 13, 72, 0, 375}, then $\mathbf{MOD}_{12}$(S) = {2, 1, 1, 0, 0, 3}.

III. Transforming BACH into CAGE

Given the sets, functions, and operations previously defined, this section describes a sequence of nine transformations to be recursively applied. The initial input is a referential set (a_0), representing ”BACH”, whose output becomes the input of another transformation, with the process being then replicated until the target-set (a_9), namely the ”CAGE” string, is reached.

- Let a_0 be an subset of set X, in the following specific order:
  a_0 = <B, A, C, H>
- First transformation: a_1 = f(a_0) = {10, 9, 0, 11};
  Figure 3 provides a graphical representation of this transformation. It is possible to consider not only the individual mappings of the four members of a_0, but also the higher-level action of f on the whole set, in a kind of ”holistic” [4] transformation (indicated by the blue arrow).

[4]In reference to a frequent adjective in Hofstadter’s book.

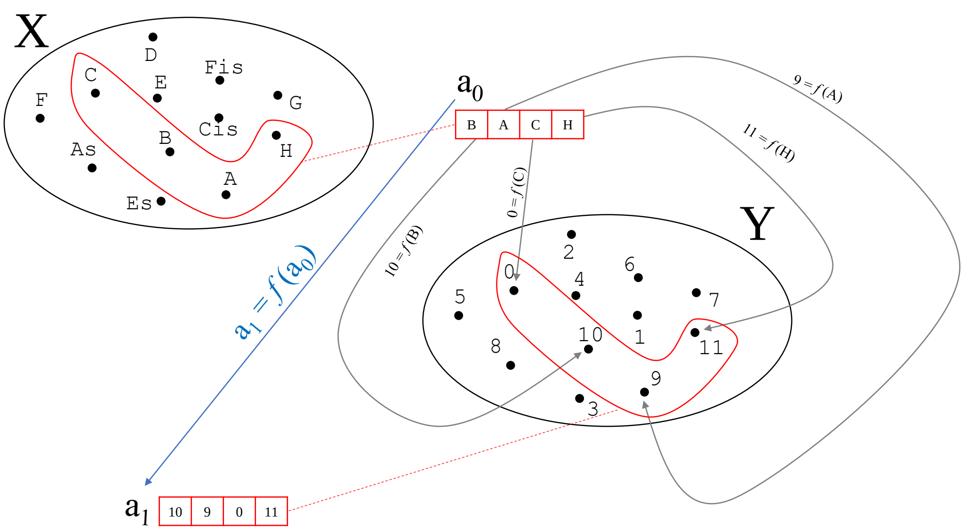

Figure 3: *Representation of the transformation of a_0 into a_1.*

- Second transformation (Figure 4): a_2 = **R**(a_1) = <11, 0, 9, 10>;

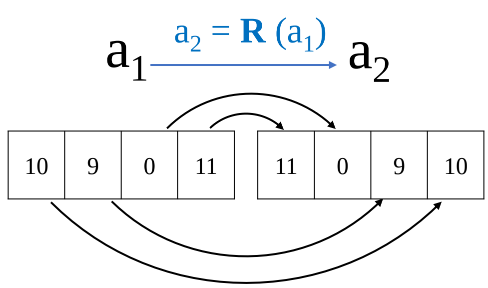

Figure 4: *Representation of the transformation of a_1 into a_2.*

- Third transformation (Figure 5): a_3 = **ROT**$_1$(a_2) = <0, 9, 10, 11>;

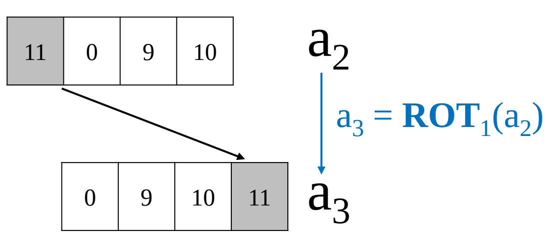

Figure 5: *Representation of the transformation of a_2 into a_3.*

- Fourth transformation (Figure 6): a_4 = **EXT**$_{3:4}$(a_3) = <10, 11>;

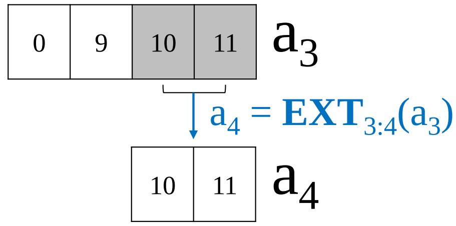

Figure 6: *Representation of the transformation of a_3 into a_4.*

- Fifth transformation (Figure 7): $a_5 = g(a_4) = <(10^2 + 3), (11^2 + 3)> = <103, 124>$;

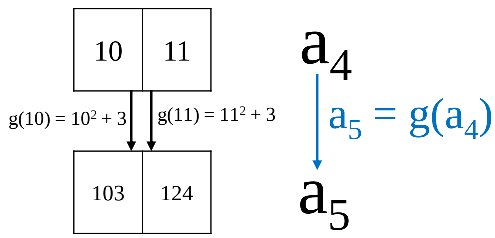

Figure 7: *Representation of the transformation of a_4 into a_5.*

- Sixth transformation (Figure 8): $a_6 = \mathbf{MOD}_{12}(a_5) = <\mathbf{MOD}_{12}(103), \mathbf{MOD}_{12}(124)> = <7, 4>$;

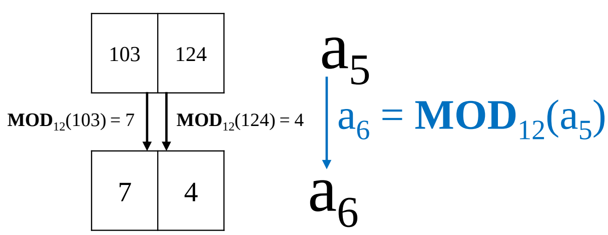

Figure 8: *Representation of the transformation of a_5 into a_6.*

- Seventh transformation (Figure 9): $a_7 = \mathbf{EXT}_{1:2}(a_3) = <0, 9>$;

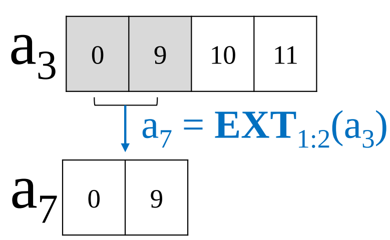

Figure 9: *Representation of the transformation of a_6 into a_7.*

- Eighth transformation (Figure 10): a_8 = **MRG**(a7, a6) = <0, 9, 7, 4>;

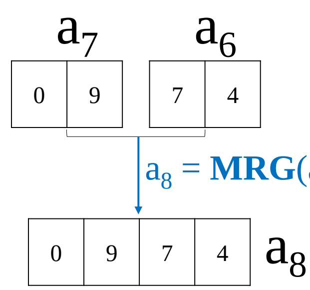

Figure 10: *Representation of the transformation of a_7 into a_8.*

- Ninth transformation (Figure 11): $a_9 = f^{-1}(a_8) = <f^{-1}(0), f^{-1}(9), f^{-1}(7), f^{-1}(4)> = <C, A, G, E>$;

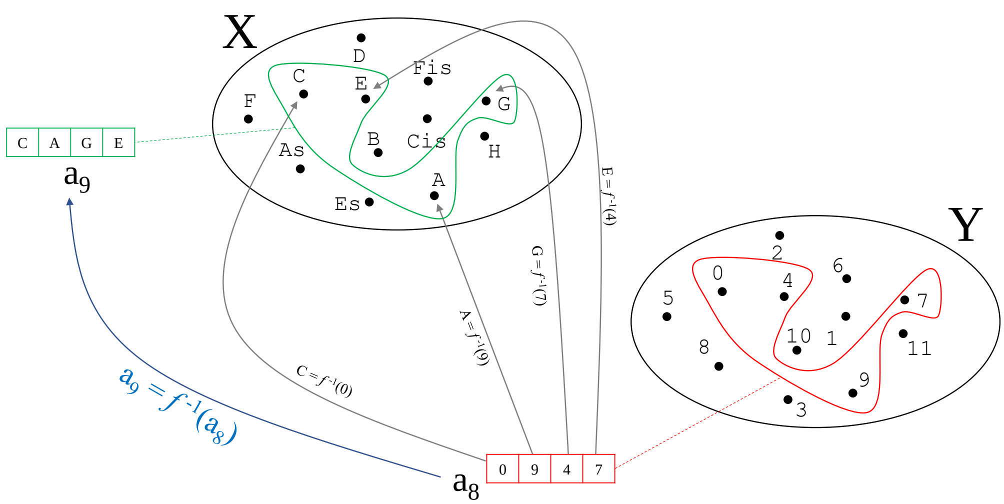

Figure 11: *Representation of the transformation of a_8 into a_9.*

Finally, Figure 12 summarizes the whole process with the aid of a *oriented transformational network*.[5] Under the same holistic perspective applied to the previous cases, I propose the creation of a high-level operation (called **B2C**)[6] that manages to map directly sequence BACH onto CAGE, bypassing the intermediary functions and operations.

[5]Concept coined by David Lewin ([2]). For a comprehensive typology of these networks, see [3], pp. 101-121)
[6]The label stands for *Bach-to-Cage*.

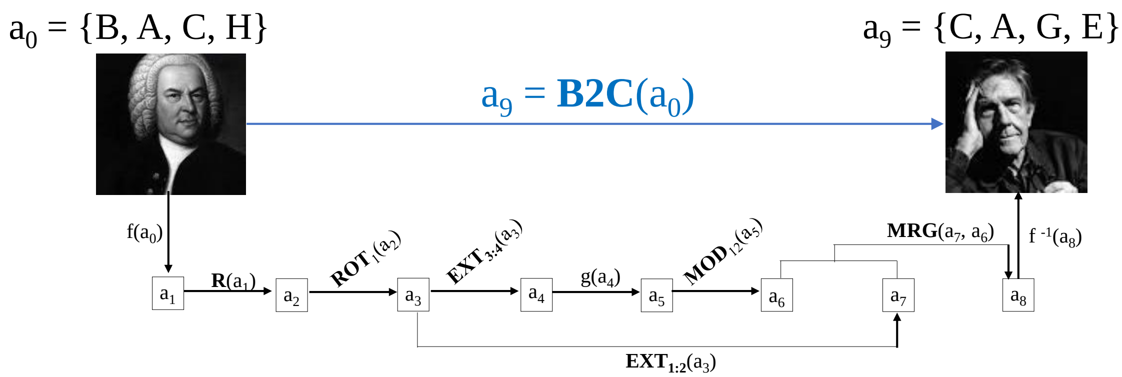

Figure 12: *Network of the nine transformations, including high-level operation **B2C**, that maps a_0 directly onto a_9.*

IV. Going a little further

After reaching the goal aimed by the article, namely, the formalization of the transformational process of the motive/sequence BACH into CAGE, some speculation and questions can arise. For example, what about the *individual* outputs of operation **B2C**? That is, in reverse to what has been done so far, could we implement a low-level function in such a manner that the elements of BACH could be sent to the corresponding members of CAGE? Or yet, it would be possible with this method, after reaching CAGE, turn back to BACH through recursive application of the same transformation (just as Hofstadter managed in his ”prove”)? If affirmative, could we generalize this function and use it to transform ”melodies” of any combination of notes in any possible extension?
Aiming to investigate these possibilities, I propose initially, for simplicity, to work with a subset of X (which ultimately represents the chromatic scale), and to adopt Guido d'Arezzo's hexachord (*Ut-Re-Mi-Fa-Sol-La*, or C-D-E-F-G-A), extended by the ”molle” and ”dur” versions of *Si* (B and H, in German notation).[7] Let us label this new set X' (Figure 13).

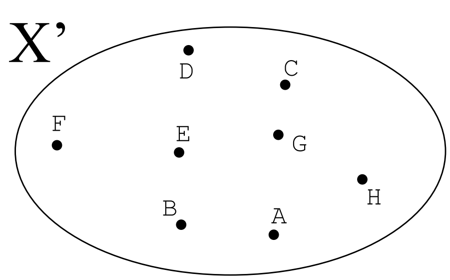

Figure 13: *Subset X'.*

Now, let *b2c* be the low-level function that maps members of X' onto members of itself | b2c: X' → X'

[7]Since this task deals with permutation, if maintained the original cardinality of 12 elements in the set, the number of possible arrangements of the letters would equal 479,001,600.

From the general action of operation **B2C** it is possible to deduce the individual ”behavior” of BACH’s notes (B→C, A→A, C→G, and H→E), but what about the remaining four letters that form subset X’ (D, E, F, and G)? Strictly speaking, there are 4^8 (65,536) possible solutions for this problem, but with the aid of logic, and by keeping in mind the idea of recursion (i.e., CAGE returning to BACH),[8] this number can be dramatically reduced to 2 alternatives.

Table 1 proposes an initial approach for the question. As it can be observed, is not possible to go back directly from CAGE to BACH, since ”C” is sent to ”G”, which will demand at least a second application of operation b2c.

Table 1: *Outputs obtained from two recursive applications of function b2c to members of subset X′.*

w	x = b2c(w)	y = b2c(x)
B	C	G
A	A	A
C	G	?
H	E	?
G	?	?
E	?	?
D	?	?
F	?	?

Assuming a minimal possible number of iterations for going from BACH to CAGE and back to BACH, it is easy to complete the output list: ”G” must sent to ”B” (which in turn goes to ”C”), and ”E” to ”D” (or ”F”), that goes to ”H” (and this returns to ”E”). Selecting ”D” as output of ”E”, ”F” must necessarily map to itself, like ”A”.[9] Table 2 depicts the definitive configuration of the cyclic transformations of the X′ members.

Table 2: *Outputs obtained from three recursive applications of function b2c to members of subset X′.*

w	x = b2c(w)	y = b2c(x)	z = b2c(y)
B	**C**	G	**B**
A	**A**	A	**A**
C	**G**	B	**C**
H	**E**	D	**H**
G	B	C	G
E	D	H	E
D	H	E	D
F	F	F	F

I will name P_{BACH} this special permutation in X′. It can be written in cyclic notation, as follows:[10]

$$(BCG)A(HED)F$$

[8]What must not necessarily be done in an immediate transformation.

[9]Or, alternatively, E→F and D→D. I chose arbitrarily the first option.

[10]It would also be possible to consider here the formation of a subgroup of the symmetrical group S_8, defined by set X′ and binary composition b2c, but do not intend to pursue this issue in the present article.

V. Concluding remarks

This paper aimed primarily, and unpretentiously to use Hofstadter's "prove" (certainly provocative, but humorous and extremely imaginative in his discussion about recursive algorithms) as a pretext to examine more deeply some of its effects in musical contexts. The system P_{BACH}, derived from the formalization of the transformational process, opens a promising connection between the notion of recursive transformation and musical variation, one of my current research interests.

Figure 14 presents a musical example, a pseudo-Haydnian theme, which is gradually "distorted" by recursive b2c-transformations of its notes (observe, however, that a third application of b2c leads the melody to its original state).[11] In spite of being a very simple (almost rudimentary, I would say) case of ("cyclic") variation,[12] the very essence of the process, namely, the application of recursive transformations, formalized as algebraic operations or functions, is potentially a powerful theoretical construct to be used both for analysis and composition, as it is being currently pursued in the course of my research.
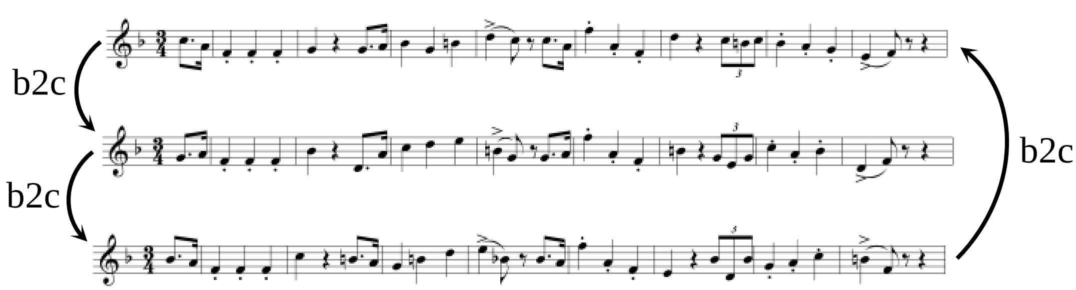
Figure 14: *Cyclic variations of a melody by recursive application of function b2c.*

References

[1] Hofstadter, D. (1979). *Gödel, Escher, and Bach*: An Eternal Golden Braid. New York: Basic Books.

[2] Lewin, D. (1987). *Generalized Musical Intervals and Transformations*. New Haven: Yale University Press.

[3] Rings, S. (2011). *Transformation and Tonality*. Oxford: Oxford University Press.

[4] Gollin, E. (2000). *Representations of Space and Conceptions of Distance in Transformational Music Theories.* Thesis (PhD in Music). Harvard University.

[5] Hoskinson, D. (2006). *The Grundgestalt and Network Transformations in the Late Choral Works of Anton Webern.* Thesis (PhD in Music). University of Oregon.

[6] Schiano, M. (1992). *Arnold Schoenberg's Grundgestalt and its Influence.* Thesis (PhD in Music). Brandeis University.

[11] Put another way, identity is achieved by applying $b2c^2$ to any member of X' (even "A" and "F"), matching one of the group properties.

[12] The complexity of the system could, of course, be increased if we relaxed the constraint of "minimal number of iterations" in the system, forming new, more extended permutations.

[7] Papadopoulos, H. (2014). Mathematics and Group Theory. *Mathematics*, [math.HO], arXiv: 149757v1, pp. 1-33.

An Information Theory Based Analysis of Ligeti's *Musica Ricercata*: Movements I and II

ADOLFO MAIA JR.
State University of Campinas (UNICAMP)
adolfo@unicamp.br
Orcid: 0000-0001-7224-8310

IGOR L. MAIA
Federal University of Minas Gerais (UFMG)
imaia@ufmg.br
Orcid: 0000-0001-9936-5023

Abstract: *In this work we present an analysis of* Musica Ricercata *I and II from the point of view of Information Theory and Complexity. We show that some patterns can be recognized and quantified using Information Theory methods such as Shannon Entropy and Statistical Correlation. In addition, we study some of Ligeti's techniques of texture formation and their time evolution. The analysis of Movement I is more concentrated on rhythmic aspect whereas for the case of Movement II, we study phrase variations along time with their timbre variations, as well some slight modifications. For both movements a suitable alphabet of symbols and coding is introduced in order to get quantitative results for comparison and analysis. We also present some comments on the power and limitations of the approach of Information Theory for the analysis of scores. Some suggestions for further work and potential applications in music composition are also briefly discussed.*

Keywords:*György Ligeti. Musica Ricercata. Information Theory. Computer Music Analysis.*

I. INTRODUCTION

> I am in a prison: one wall is the avant-garde, the other wall is the past, and I want to escape.
>
> *György Ligeti* [1]

Formal and mathematical methods have a fertile role in twentieth-century music. Since the pioneer works by Xenakis and the computer experiments of Hiller in the middle of last century with his *Illiac Suite*, this kind of formal approach has attracted many researches and composers and nowadays is a consolidated area of research in music composition and analysis. Mainly due to the increasing power of computer processing since that time, besides the development of new and efficient software, sophisticated mathematical, statistical and computer tools have been extensively used in areas such as Computational Music Analysis, Computer-assisted Composition and Sound Synthesis. In this work we are most interested in the application of the Mathematical Theory of Information to music analysis. The book edited by [2] brings contributions of several authors in the area. In general, computer methods aid the researcher to develop quantitative methods which allow a broader vision of the work under study. A recent example is the work by Jacoby *et al* on chord categorization in functional harmony using several kinds of representations [3].

This is in line with the increasing number of different styles in contemporary music. Since much of the traditional methods of analysis (Riemann, Schenker et al) are not always suitable to study the music of the 20th century and later, we think that mathematical and computer approaches can be an important tool to music analysis of the 20th and 21th centuries. Although abstract methods of mathematics can be applied to many different areas and models, the fact that music of those centuries have a huge number of different aesthetics, methods and styles, for sure, none mathematical/computer tool is comprehensive enough to encompass analyses of many different works. So we must be humble and to live with the reality that a method can be useful for a small set of works or perhaps for just a single work and that adaptations or extensions are needed most of times. Within this framework we analyse only the two first movements of György Ligeti's *Musica Ricercata* from the point of view of the Mathematical Theory of Information. In the spirit of the above comments we also discuss in some lenght the power and limitations of Information Theory in musical analysis and composition.

Musica Ricercata (MR, for short) was composed as a music problem-solving: how to compose a piece with "minimal resources" and "maximal results". It deserves attention not only due to its musicality and craftsmanship but also because it points to some strong characteristics in Ligeti's later music such as textural structures and processual procedures [4]. The title has a double meaning of ricercare as a musical form but most importantly as "researched music" [5] and, in fact, it was a very personal experiment on chromaticism [6]. Ligeti was always fond of science and, particularly, beautiful mathematical structures like Fractals and mechanical processes.

> Quelques-unes de mes oeuvres n'auraient pas été possibles sans la connaissance de la théorie du chaos. Je ne vois pas de limite de principe entre l'art et la science. Leurs méthodes sont quelque peu différentes: la science est orientée vers la réalité et développe à partir de là des hypothèses; l'art ets plus libre, il n'est pas soumis aux prémisses du monde réel. Et pourtant, l'art n'est pas arbitraire, car s'il a une certaine de liberté, il est cependant lié à l'histoire et à la société [7].

Within this collection of small pieces we chose to analyze the first two ones, since they have a small number of pitches and techniques, besides, many pattern repetitions, such characteristics which make them amenable to apply Information Theory methods as well as Complexity Theory using computer codes of musical information [8].

The freedom from the domain of tonality, which Ligeti imposed to himself, implied in a search for new methods of composition. About the gross structure of these two small first movements of MR, Ligeti wrote:

> Ainsi, le style de ces deux compositeurs [Bartók et Stravinski] est perceptible dans mes pièces pour piano, bien que j'espère que certaines d'entre elles - par exemple la dernière ou les deux premières - fassent apparaître déjà un style personnel. . . . Le statisme des trois premières pièces constitue une particularité stylistique qui devint ensuite

> caractéristique des "veritables" compositions de Ligeti apparues dans la seconde moitié des années cinquante [7].

The close interest Ligeti had with mathematics, and machinery repetitive patterns like clocks, were influential to a number of formal procedures he used in his compositions, mainly to those so called "Pattern-Mecannico" [9]. The quote below clearly express his thought about his method of composition

> Although I am an artist, my working method is that of a scientist active in basic research rather than in applied science. Or of a mathematician working on a new mathematical structure, or of a physicist looking for the tiniest particle of the atomic nucleus. I do not worry about the impact my music will make or what it will turn out to be like. What interests me is to find out the way things are. I am driven by curiosity to discover reality. Of course, there is no reality in art the way there is in science, but the working method is similar. Exactly as in basic research where the solution of a problem throws up innumerable new ones, the completion of a composition raises a host of new questions to be answered in the next piece [10].

Ligeti's approach in MR, as well as in other later works, is quite evident as an experiment of a minimal (simple) formal model and artistic craftsmanship. In our opinion, he seems fulfilling a famous quote attributed to Albert Einstein which appears, among many other sources, in a 1950 Roger Sessions article in the N.Y. Times *How a "Difficult" Composer Gets that Way*:

> A also remember a remark of Albert Einstein, which certainty applies to music. He said, in effect, that everything should be as simple as possible as it can be, but not simpler . . . I try only to put into each work as much as myself possible [11].

In this paper we make a partial analysis of Movements I and II of *Musica Ricercata* through some methods from Information Theory and Statistics also following the Einstein's premise Sessions prescribes. That means we search for patterns which can be, in some way, analyzed through Information Theory but also taking into account the expression "but not simpler" as related to the creativity of the composer, those other aspects which are not grasped by our code.

There exist many different measures of informational content of a string of symbols. We can mention, besides Shannon entropy, Kolmogorov Complexity, Self-correlation, Lempel-Ziv-Welch compression, among many others [12]. Many of them were applied to music encoding and extracting of quantitative information from a musical work. However, there is no common sense about what is more suitable for music or even for different styles of music. So, plenty of room for research.

However, as any area of science, the Information Theory has its limitations. Shannon Entropy returns a numerical value from a bunch of data (in our case, musical data) which in itself hasn't much utility. However, in the case of comparison of data sets, it indeed can have guide us in the analysis of related structure patterns of such data. For example, Information Theory can be used to compare variation of musical parameters in a corpus of music. For the case of *Musica Ricercata* an obvious approach is to compare patterns of right and left hands, or sections of the same hand, etc. Nevertheless, it is important to stress that Shannon Entropy does not take into account time order of the elements in a sequence. This is evident from the definition of entropy in Equation 1, which uses only probabilities which, in turn, are defined just by counting symbols with no order taken into account. So, Information Theory is not a well-tailored tool for a detailed analysis of a piece, since time order of symbols is very important for composition or analysis of a piece. In section III we make some additional account about these limitations.

We used MATLAB® as our mathematical tool for calculations and graphics.

II. Shannon Entropy

Shannon's Information, or Entropy, is based on the probability that a given symbol appears in a given sequence. Formally, if x is a sequence of symbols from an set commonly named alphabet $A = \{a_1, a_2, \ldots, a_N\}$ and p_i is the probability to find the symbol a_i in this sequence. The Shannon Entropy of the sequence x is given by the formula of Equation 1.

$$H(x) = -\sum_{i=1}^{N} p_i \log_2 p_i \tag{1}$$

Observe that the crucial problem here is to define probabilities for each symbol. In general, there isn't a unique way to do it. For example, to get the entropy of a written text in a language, it is necessary to count the relative frequencies of all symbols of the alphabet of the language, as appearing in the text. In the same way we can count frequency of words (higher structures) or even phrases. In general this is done by counting and calculating relative frequencies of all possible combinations of two letters, three letters, etc. These frequencies define the probabilities for symbols (first order) and higher structures, or words, (higher orders). So we have a distribution of probability for all symbols of the alphabet in that language. The entropy of the language is the entropy of the distribution of probability. Once we have the probabilities, we can use them to calculate, by plug them in the Shannon's formula, the entropy of any text in the language. This is pretty the same for a score where the musical symbols and higher structures as chords, motives, dynamics, etc., can also be counted and calculated their entropies. For the sake simplicity, we consider only the entropy taking into consideration only probabilities for symbols (first order).

In this work we are also interested with the time evolution of parameters such as rhythm and pitch patterns and their sequences of symbols in the score. So we prefer calculate entropy for short sequences of symbols representing rhythmic or pitch patterns in a bar and then to analyze the time evolution of the entropy along the bars. So, the probabilities we work are defined by the relative frequencies of symbols in short musical segments such as bars and we study the time evolution of these bar entropies along time, or, along bars. The definition below goes in this direction.

Definition: Let $A = \{a_1, a_2, a_3, \ldots, a_R\}$ be an alphabet and $s = [s_1\, s_2\, s_3\, \ldots s_N]$ a finite sequence of symbols of A, that is, for any $1 \leq i \leq N$, $s_i = a_k$, for some k, $1 \leq k \leq R$. Denote n_k the relative frequency of the element a_k in s, the probability of this element is defined as shown in Equation 2.

$$p_k = \frac{n_k}{N} \tag{2}$$

Formally, we're defining a *Distribution of Probability* X for the symbols of the alphabet A. So, in fact, we should write $H(X)$ for the entropy associated to the probability distribution X. We wrote $H(s)$ here because the distribution X is defined, as mentioned above, by relative frequency of symbols of sequences s, where s is the code sequence of a bar or a small set of bars.

As an example let's take $A = \{0, 1\}$ as an alphabet and consider the following sequence:

$$s = [0\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1]$$

Denoting $p_1(s)$ and $p_2(s)$ the probabilities of symbols 0 and 1 in the sequence s, it's easy to see, from Equation 2, they read: $p_1(s) = 5/11$ and $p_2(s) = 6/11$. Taking these values into Equation 1 we get the entropy of the sequences (Equation 3).

$$H(s) = -\frac{5}{11} \log_2 \frac{5}{11} - \frac{6}{11} \log_2 \frac{6}{11} \approx 0.99 \tag{3}$$

Observe that order is not taken into consideration to calculate entropy. So, for example, the sequence $s' = [1\,1\,0\,0\,1\,1\,0\,0\,0\,1\,1]$ has the same entropy of sequence s.

Clearly the above method of getting probability and calculate entropy can be applied to any musical parameter: pitches, durations, rhythm accents, rests, dynamics, and so on. Such general scheme seems an impressive tool for analysis or even composition but that is not necessarily true. Some limitations of the approach are shown in section III.

A note on units in Information Theory: in base 2, entropy measures the quantity of "bits per symbol" in a sequence. Take, for example, an alphabet with only two equally probable symbols $\mathcal{A} = \{x, y\}$. A bit can be viewed as a box in which we can write x or y. So, we have 2 possible "symbols per bit" which implies that the probability to find any of them is its inverse $\frac{1}{2}$ "bits per symbol". Thus, the information for symbol x, for example is $-\log_2 p(x) = -\log_2 \frac{1}{2} = -(-1) = 1$ bit. Other numerical basis can be used as well which define other information units. One can go from one representation to another through a constant of translation relating the logarithm of both basis.

III. Limits of Information Theory in Music

Edgar Varèse defined music as organized sound [13]. However, from a physical-mathematical point of view, sound is a multi-parameter physical phenomena and is detected in the human brain after a series of complex signal transformations (filters) through the ears. So any attempt to code music is very restrictive and the organization Varèse meant refers just to the surface of a thick structure [14]. Hopefully, Varèse definition is, in general, enough for music analysis. So, Information Theory is used here in the Varèse sense, that is, as time organization of symbols which represent sounds to be played. Taking this into account, we must describe the limits of our analysis. Firstly, it must be stressed that many aspects of performance and musical gestures are not grasped by our analysis, since we are most interested in pitch and duration only. So, we have, as any mathematical approach, a kind of reductionist analysis which, nevertheless, can shed some light on the overall construction. For example, the *accelerando* in the *Misurato* section, bars 6 to 13, is easily represented by our coding shown in section IV. However, the *accelerando* in bars 76 to 80 is only partially represented since there is not a well-defined way, with only our alphabet, to represent tuplets in $\frac{4}{4}$ signature. In order to do that we would need another symbol or perhaps a set of symbols. In fact, the increase in the number of symbols allows a better representation of the score, but the price for this is that the Information value is more difficult to interpret since it is a result, or sum, of many variables. For example, we did not represent harmonic notes (diamond shaped), since an additional symbol would be needed to represent the information that the note is a "harmonic". So, this kind of effect is not grasped by our coding of *Musica Ricercata* and thus we have only a partial information on the timbre actually indicated in the score. In general we can say that in Theory of Information the rule of thumb is "what isn't coded does not exist".

Another important point to mention here is about tempo. This is a very important aspect of musical material but our coding isn't able to give any information about it. In Information Theory, the calculations are about sequences of symbols, in general not directly related to the speed of time. In other words, if the same message, including music, is sent quickly or slowly, the entropy is exactly the same. Just to add a new symbol for tempo has not much utility since, in general, it would appear once or, at most, few times in a code of a musical work. On the same premise, articulation also is not represented by our code.

Likewise, Information Theory can be useful also to give some clues of the compositional process behind the musical surface in different contexts. Nevertheless we think that, in general, its use as a compositional tool is a poor one. In fact, since Xenakis and Hiller, many formal and mathematical approaches for composition as, for example, Stochastic Processes, Markov Chains associated to different statistical distributions, have been far more effective to composition than

Information Theory. In short, you need firstly an object to get information from it. Perhaps, it's better to compose music through other formal devices such as, for example, Combinatorics, and to use Information Theory as music analysis tool. Although Information Theory hasn't a high performance for music composition, since it doesn't take into account ordering in sequences but only symbols counting, there is plenty of room for creation of motives and phrases and textures, following prefixed entropy values, or just making permutations of the elements of a sequence of symbols keeping entropy fixed. Also, in composition this could be used for comparison of different sketches of a work in progress or even to compare them with material from other composers works.

Information Theory approach can be used, theoretically, for the analysis of any kind of music. Nevertheless, clearly it works better for music with repetitive structures or with small variations of them such as, for example minimalist music [16].

IV. Coding Musical Data as Symbolic Sequences in MR1

In this movement Ligeti uses only two pitch classes, that is A and D (this last pitch-class only at the end of the movement to create a cadence). All the pitches used in this movement can be seen in Figure 1.

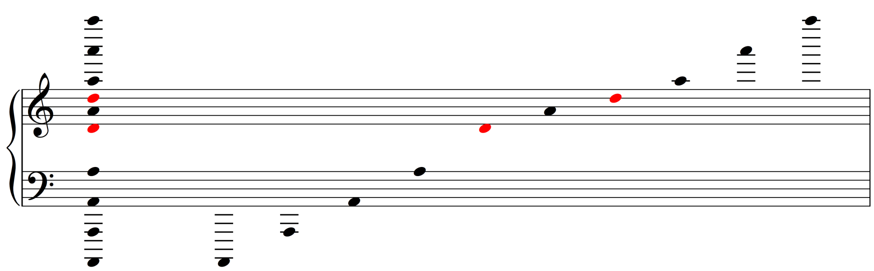

Figure 1: *The set of notes used in Movement I of Musica Ricercata.*

With such meagre resources, which Ligeti forced on himself, the other musical parameters like rhythm and timbre gain prominence. Nonetheless, Ligeti is able to find a fine balance between repetition and variation. Shannon Entropy gives an estimate of this balance. Nevertheless other statistical and complexity measures can be associated to symbolic sequences, such as Kolmogorov Complexity Correlation [15].

We show below how we code the musical material of Movement I of *Musica Ricercata*. Needless to say that a useful general code for analysis and composition is not yet available and our method is also plagued with limitations, mainly those ones of complementary performance indications as, for example, articulations, pedals, *crescendi*, etc. Therefore, although we intend to get the greatest possible generality, each code depends strongly on the piece or set of pieces under study. In order to describe our coding we firstly observe that, out of the four final bars which is a cadence in D from A, all material is obtained from the pitch class A. Each bar is coded as a set of two strings of integers, one string for each hand. The octaves range from A1 to A8. An isolated note is coded by its octave position, from 1 to 8. A chord is coded by the position of its two notes, that is, a pair of two integers in the interval from 1 to 8. Rest is denoted as the negative integer **-1**. In a string, the number of numerical symbol 0 after a positive number (isolated note), or pair of positive numbers (chord), denote duration of the note, or chord, in beats. We reserve the numerical symbol **9** to code any other needed parameter. So the set of symbols read as shown in Equation 4.

$$S = \{-1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9\} \tag{4}$$

With this set of symbols and the simple algorithm defined above, a typical bar as number 33 in Figure 2 is coded as two sequences, one for each hand

$$\begin{aligned} lh\{33\} &= [2 \quad 0 \quad 3 \quad 0 \quad 2 \quad 0 \quad 3 \quad 0 \quad 2 \quad 0 \quad 3 \quad 0 \quad 2 \quad 0 \quad 3 \quad 0] \\ rh\{33\} &= [5 \quad 0 \quad 0 \quad 4 \quad 0 \quad 5 \quad 0 \quad -1 \quad 0 \quad 6 \quad 0 \quad 5 \quad 0 \quad 6 \quad 0] \end{aligned}$$

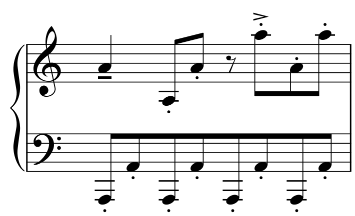

Figure 2: *Musica Ricercata, Movement I, Bar 33. Reprinted by permission of ©SCHOTT MUSIC, Mainz – Germany.*

For a simple example of calculation, using our symbols and coding algorithm, consider the right hand of bar 6 which is coded as:

$$rh\{6\} = [-1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

It uses only two symbols and the size of the string is 9. The relative frequency (probability) of each symbol is $p_1 = p(-1) = 1/9$ and $p_2 = p(0) = 8/9$. The entropy is then given by Equation 5.

$$H(rh\{6\}) = \sum_{i=1}^{2} p_i \log_2 p_i = -\frac{1}{9}\log_2(\frac{1}{9}) - \frac{8}{9}\log_2(\frac{8}{9}) \approx 0,503 \neq 0 \tag{5}$$

As a second example, observe that the left hand of bars 6 and 7 seem equivalent since both have 2 notes and 6 rest beats but in different positions. Our representation can differentiate them and, in fact, they have different entropies. Their codes read:

$$\begin{aligned} lh\{6\} &= [2 \quad 0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad -1 \quad 0] \\ lh\{7\} &= [-1 \quad 0 \quad 0 \quad 2 \quad 0 \quad -1 \quad 0 \quad 0 \quad 0 \quad 2 \quad 0 \quad -1 \quad 0] \end{aligned}$$

and their entropies are given, respectively, by $H(lh\{6\}) \approx 1.25$ and $H(lh\{7\}) \approx 1.33$). The difference comes from the additional symbol -1 in $lh\{7\}$.

V. Information Based Analysis of MR1

Below we make a comparative analysis from the point of view of Theory of Information through the calculation of Shannon Entropy for several sections along the work. We've calculate separately entropy coefficients of left and right hands.

The piece starts with an Introduction section indicated as *Sustenuto* from bars 1 to 4. Bar 5 is just a rest, functioning as a bridge for the next section.

1. **Bars 1 and 2**

   Clearly the two gestures in these bars are equal and, of course, their information measure as well. For example, $H_{left}\{1\} = H_{left}\{2\} = 1.55$, and the entropy ratio between right hand and left hand is equal to 1 (Equation 6).

$$C_{left-right} = \frac{H_{right}}{H_{left}} = 1 \tag{6}$$

   As mentioned in the Introduction, our alphabet is not rich enough to take into account the "tremolo" effect of this section.

2. **Bars 3 and 4**

   In this subsection there is only one chord on the right hand and a harmonic resonance on the chord A2-A3. Our code isn't able to get information from the harmonic chord. So we can only add it as a normal chord or simply ignore it in our analysis. Either way the information doesn't correspond to a more realistic representation of the real sound. A better solution which does approximate to the real sound is to assign new symbols for harmonic notes and chords (resonance). The downside for this is to increase the number of symbols which are not usable in our case since harmonics just appear in these bars and again in the final 4 bars of the cadence to D.

3. **Bar 5**

   This bar is just a pause to start of the next section, *Misurato*.

4. **Bars from 6 to 13**

   This subsection is the beginning of *Misurato* section which lasts from bars 6 to 59. It is played only with the left hand and is clearly an *accelerando* with asymmetrical rhythmic figures. The first 4 bars show an increasing number of notes, but their position do not seem to obey any specific order. From bar 10 to 13 a new pitch is added, namely, A3 and the number of notes continues to increase until bar 13, which contains no rests. The time evolution of the entropy along this subsection is given by the Figure 3.

5. **Bars from 14 to 58**

   This is the longest subsection which Ligeti keeps a more or less definite pattern. The left hand is an ostinato along the entire subsection, while the right hand develops an increasingly complex patterns including more notes in different octaves as well as chords. Figure 4 shows the initial bar of this subsection.

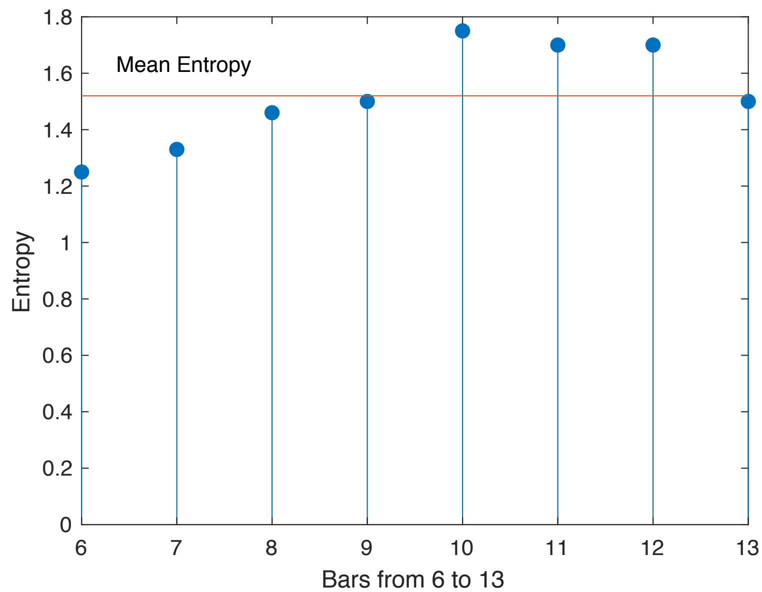

Figure 3: *Musica Ricercata, Movement I, Time evolution of Entropy of the left hand: Bars 6 to 13.*
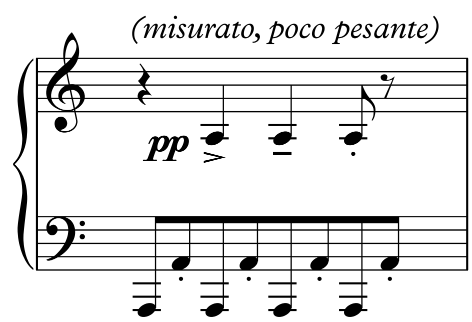
Figure 4: *Bar 14 of Movement I of Musica Ricercata. Reprinted by permission of ©SCHOTT MUSIC, Mainz – Germany.*

Clearly, the repetition of the same pattern in all bars of the left hand implies a constant entropy along it. So, it's more interesting to study the entropy variation of the right hand along this segment, although one can argue that is just the variation of the right hand against an *ostinato* (constant) background that strikes as new for the human ear. The behavior of the right hand against the *ostinato* can be seen in Figure 5 in terms of Shannon entropy.

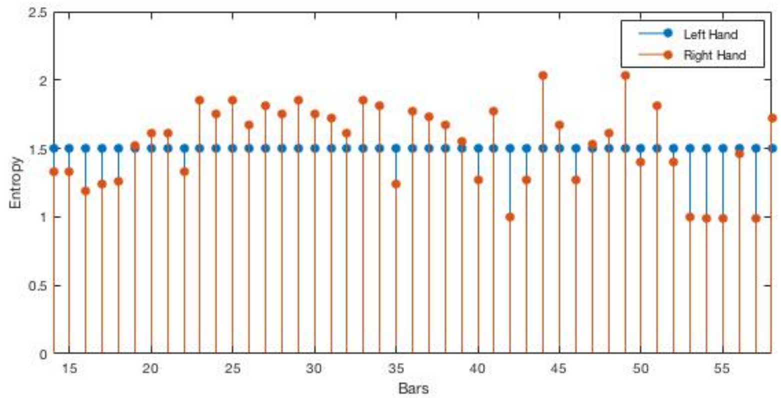

Figure 5: *Musica Ricercata, Movement I, Time evolution of Entropy for Bars 14 to 58.*

The ratio between the left and right hand entropies, for Bars 14 to 58, can also be seen in Figure 6.

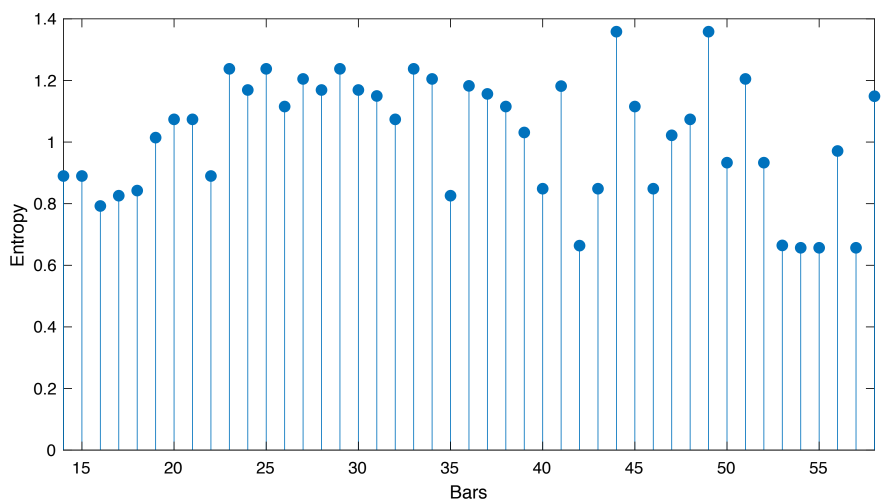

Figure 6: *Musica Ricercata, Movement I, Time evolution of the ratio between Left and Right Hand for Bars 14 to 58.*

6. **Bar 59**

   In this bar Ligeti keeps the same pattern of the previous bars but changes the time signature to $\frac{3}{4}$. The left and right entropies of this bar read $H_{left}\{59\} = 1.50$ and $H_{right}\{59\} = 1.56$.

7. **Bars from 60 to 65**

   These bars are the beginning of the section *Prestissimo* returning to signature $\frac{4}{4}$. The very similar rhythm patterns with the same notes lead to pretty the same entropy for left and right hands as shown in Figure 7.

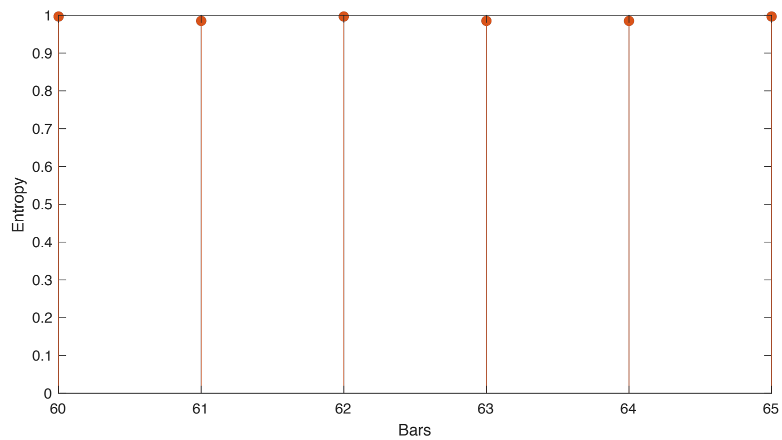

Figure 7: *Musica Ricercata, Mov. 1, Time evolution of Left and Right Hand Entropies for bars 60 to 65.*

8. **Bars from 66 to 80**

   This is an increasingly fast (*accelerando*) section preparing to the final cadence in D. It is not difficult to see that due the absolute parallelism between left and right hands, their entropy curves are the same. That is just case as shown in Figure 8.

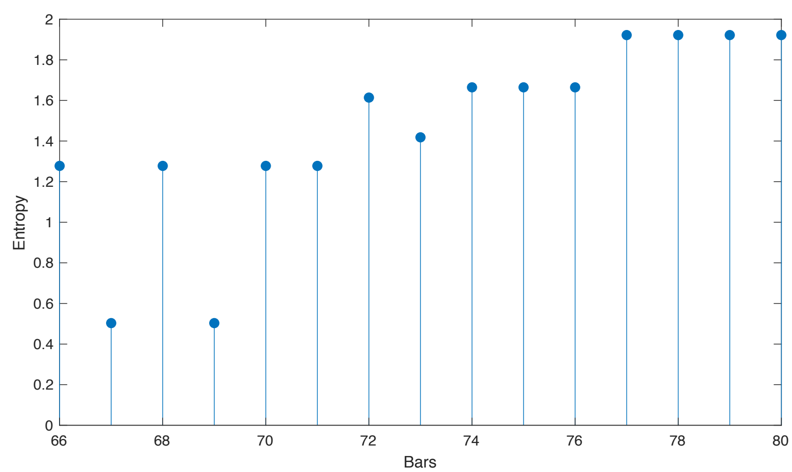

Figure 8: *Musica Ricercata, Mov. 1, Time evolution of Left and Right Hand Entropies for bars 66 to 80.*

Observe, as we have mentioned regarding to other parameters, our code is not able to see the clef changing in bar 68. So, this does not affect the calculation of the entropy. On the other hand it is possible, in this case, due to the parallelism, to measure the *accelerando* through our code simply counting the number of positive single or pair of numbers (notes or chords) per bar and taking into account that the time signature is $\frac{4}{4}$.

9. **Bars from 81 to 85**

   After a pause, bar 81, Ligeti makes a cadence in D. The basic notes are D3 and D5 late accompanied by harmonic chords D3-A4 and A6-D6. Our code isn't able to take into account both the addition of note D and the information "harmonic". So we must analyze this section in a different basis. Clearly this is a completely unexpected event, if we take into account all the previous bars where we have just A in different octaves, which is the only pitch information. So it is not a fault of our code, but simply the event implies an extension of our alphabet. If we keep the alphabet, this section cannot be coded and so there is no sense to calculate entropy for it. However, if we can consider bar section 81–85 as a new unique event (a unique sound event), extending our alphabet to include the pitch D, the Shannon entropy, restricted to it, is rigorously zero.

VI. Analysis of Movement II

MR2 is essentially melodic. The pitch-class set of this movement is $\{E\sharp - F\sharp - G\}$, the most compact of pitch-class sets Ligeti used in MR. Much as in MR1, the resources here are very limited. In a piece with 33 bars, the third pitch, in fact, the same note $G5$, is introduced only halfway, in bar 18.

In his Masters dissertation, D. Grantham makes a very good analysis of Musica Ricercata [6]. In particular, his analysis of MR2 is extensive in the description of the general structure of the movement. Here we present a different approach, based on Information Theory which can complete his one. Firstly we make some observations on the movement's structure.

Again, as in the MR1, Ligeti is able to find a fine balance between repetition, symmetry and variability in his two-note melodic construction along the movement. The melodic construction is based on a four-bar key phrase which can be thought as two question-answer bar motives. Figure 9 shows the key phrase.

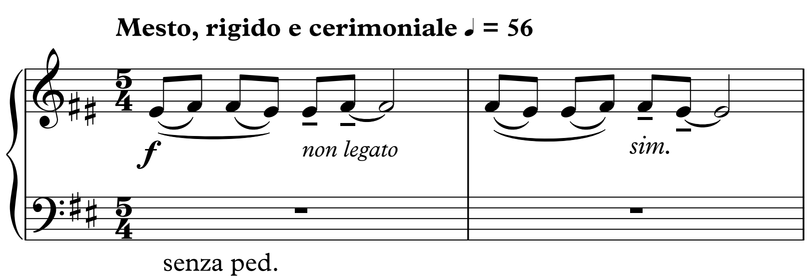

Figure 9: *Initial Phrase of Movement II of Musica Ricercata. Reprinted by permission of ©SCHOTT MUSIC, Mainz – Germany.*

Along the piece, slight variations of these two one bar motives appear in both hands, however with different vertical constructions, such as chords and against contrasting textures. Also it's worth to say that dynamics, most of times, changes wherever a new phrase starts which in turn is correlated also to changing in timbre. Overall taking into account, Ligeti gets a fine variability of material even using so restrictive pitch material.

In order to facilitate our quantitative analysis, we use here a simple code alphabet: for the pitch classes $\{E\sharp = 0, F\sharp = 1, G = 2\}$. The overall structure can be divided in the following sections:

Section 1: bars 1-16

The section can be thought as somewhat indolent and mysterious melody in tempo ♩= 56 composed of a sequence of variations of the above two one bar motives of Figure 10. These variations has a periodic behavior. The time signature has a periodic behavior each 4 bars as $\begin{bmatrix} 5 & 5 & 4 & 6 \\ 4 & 4 & 4 & 4 \end{bmatrix}$.

In fact, taking into account that each motive, comprising one or two bars, ends at an agogic (duration prolongation) accent of two or three beats, the sequence of motives, using the above code, is given by Table 1 in pairs.

Table 1: *Code of bars 1-16*

Bar Number	*Type*	*Bar Code*	*Signature*
1-2	Q-A	$\|011001\hat{\mathbf{1}}\|\|100110\hat{\mathbf{0}}\|$	$(5,4)-(5,4)$
3-4	Stat	$\|01101001\|\|100110\hat{\mathbf{0}}\|$	$(4,4)-(6,4)$
5-6	Q-A	$\|011001\hat{\mathbf{1}}\|\|100110\hat{\mathbf{0}}\|$	$(5,4)-(5,4)$
7-8	Stat	$\|01101001\|\|100110\hat{\mathbf{0}}\|$	$(4,4)-(6,4)$
9-10	Q-A	$\|011001\hat{\mathbf{1}}\|\|101001\hat{\mathbf{1}}\|$	$(5,4)-(5,4)$
11-12	Stat	$\|01101001\|\|101001\hat{\mathbf{1}}\|$	$(4,4)-(6,4)$
13-14	Q-A	$\|011001\hat{\mathbf{1}}\|\|100110\hat{\mathbf{0}}\|$	$(5,4)-(5,4)$
15-16	Stat	$\|01101001\|\|100110\hat{\mathbf{0}}\|$	$(4,4)-(6,4)$

In this table the duration accentuation is in bold face with a hat. We defined two types of the motive: *Q-A* means question-answer motive and *Stat* means statement motive as described below. As in MR1 Ligeti explores timbre with different uses of pitch classes of the pair $\{E\sharp, F\sharp\}$, such as, for example, right hand melody in bars 1-4, chords in both hands in bars 5-8, left hand melody in bars 9-12 and so on.

It is easy to check, from Table 1 that, except for bars 9-10, all pairs of bars with time signatures $\begin{bmatrix} 5 & 5 \\ 4 & 4 \end{bmatrix}$ are symmetrical under the exchange symbols $0 \leftrightarrow 1$. However bars 9-10 are symmetrical under the exchange of the two first digits. These pairs of bars sound as a kind of question-answer pattern, mostly due to the fluctuating sequence of notes ending with the "cadence" $F\sharp$ in the first bar of the pair and $E\sharp$ in the second one. Now the pairs with signature $\begin{bmatrix} 4 & 6 \\ 4 & 4 \end{bmatrix}$ seems to work as a two bar reinforcement of the previous two bars. As the previous case, except the bars 11-12, all other pairs are the same two-bar motive and don't have any apparent symmetry between the bars in each pair. Nevertheless, the first element of all pairs has symmetry $0 \leftrightarrow 1$ between the its half parts.

Once described the overall structure of the section we are interested to compare it with its information content. Much as we did in the case of MR, our approach here is just to calculate the Shannon Entropy defined in Equation 1 for each of the two-bar motives coded in Table 1 and plot its time evolution. We do not take into account rhythm, but only the pitch classes. So the alphabet $\{0,1\}$ only represents the change of pitch. The relative frequencies of symbols for *Q-A* and *Stat* motives are given by Table 2.

From Table 2 we get the Relative Probabilities (Table 3).

Since the alphabet has just two symbols $\{0,1\}$, Table 3 shows that they are fair balanced along the section and the Shannon Entropy for the eight motives gets almost constant:

$$H = [1.000 \quad 0.997 \quad 1.000 \quad 0.997 \quad 0.985 \quad 0.997 \quad 1.000 \quad 0.997]$$

So, from the point of view of pitch class, as expected, MR2 has a very low variation of information.

Table 2: *Relative Frequencies of Q-A and Stat Motives*

Relative Frequencies			
Q-A motives		*Stat* Motives	
7	7	8	7
7	7	8	7
6	8	7	8
7	7	8	7

Table 3: *Relative Probabilities of Q-A and Stat Motives*

Relative Probabilities			
Q-A motives		*Stat* Motives	
0.50	0.50	0.53	0.47
0.50	0.50	0.53	0.47
0.43	0.57	0.47	0.53
0.50	0.50	0.53	0.47

Section 2: bars 18-24

What do you do with just one pitch? In music, a lot, mainly rhythm patterns, but from the point of view of Theory of Information, given the alphabet, we use in MR2, with just 3 symbols in which $G = 2$, the entropy is exactly zero. In other words, in this section information on durations of notes is clearly detected by human ear, but our code cannot do it. This section, with time signature $\frac{4}{4}$, consists only of the accented note $G5$ played in *accelerando*. Ligeti's notation shows, in tempo of ♩ = 126, groupings of accented notes of shorter and shorter durations along the bars.

In the beginning of the section, bar 18, Ligeti wrote "tutta la forza" with dynamics ***ff***. This short section had an intentionally extramusical motivation

> ...Ligeti revealed that, as he composed the piece, the reiterated Gs had symbolised for him "a knife through Stalin's heart" [5].

In order to quantify the *accelerando* we use the concept of *Symbol Rate* from Information Theory which means the number of symbols per time unit which, in our case, we take the bar. In this section each symbol $G5$ corresponds to an event (note attack) from the right hand. So, the pace of attacks per bar along this section reads $1 - 1 - 2 - 3 - 4 - 5 - 96$, which shows a steady linear growth with a big discontinuity in the last bar where also Ligeti's indications "Senza Tempo" and "Rapido" leave some openness to the details of performance. Here we have "in nuce" the first example of the fuzzy zone of transition from discrete to continuum in time almost in the limit of human performance. This is to be compared with the last bar where we find the reverse transition, that is, from continuum to discrete. As it is well-known, Ligeti would explore these artifact much later in his 1968 piece *Continuum for Harpsichord*. Nevertheless, along all this section the Shannon Entropy is zero, since only one symbol $G5$ was used. This show that Shannon Entropy has some limitations when applied to some kinds of physical or artistic information, such as, for example, identical pulses from a source, or in our context, the same pitch from a musical instrument, since the human perception in fact gets information such as rhythm, although just one symbol is used. Another example is the obvious physical information (rhythm) coming from the use of odd n-tuples that get fractional durations for notes. That is not coded as a sequence of symbols. It is possible circumventing this problem extending the coding by choosing a very short time unit. However this will lead to a very artificial and long string representation of note durations which has not simple musical correspondence.

Section 3: bars 25-28

In this short section, spanned 3 bars, Ligeti superposes the two materials presented before: it starts with the initial melody by octave chords in the left hand and, halfway the bar 25, it is strongly contrasted, now on the right hand, with the octave $\{G4 - G5\}$, with $G4$ a half note, while $G5$ is articulated as the same sequence in the previous bar in tempo "rapido" on the right hand, but now with some notes of the sequence accented just playing the octave $\{G5 - G6\}$. These artifacts get a higher timbre complexity, as well rhythm variations. This pattern is kept until bar 28.

Section 4: bars 29-32

This four-bar section has the same initial melody line, a reaffirmation of the mysterious and somewhat nostalgic mood, but now with a stronger emphasis, since it is written with parallel octaves in both hands. Again it is a way Ligeti uses to get timbre variability. The code of this section is shown in Table 4.

Table 4: *Code of bars 29-32*

Bar Number	*Type*	*Bar Code*	*Signature*
29-30	Q-A	$\lvert 0110011\hat{1} \rvert \lvert 1001100\hat{0} \rvert$	$(5,4)-(5,4)$
31-32	Stat	$\lvert 01101001 \rvert \lvert 1001100\hat{0} \rvert$	$(4,4)-(6,4)$

Clearly the Entropy is the same of the bars 1-4, that is, $H = 1$.

Section 5: bar 33

On the left hand, a half tone descending "cadence" $\{F\sharp \rightarrow E\sharp\}$ in octave chords is played twice, while above it, on the right hand, the $G5$ "rapid" train of notes in *perdendosi* and continuously decelerating. The half-tone descending cadence works like a tension-release which is confirmed by the "perdendosi" sound attenuation (Figure 10).

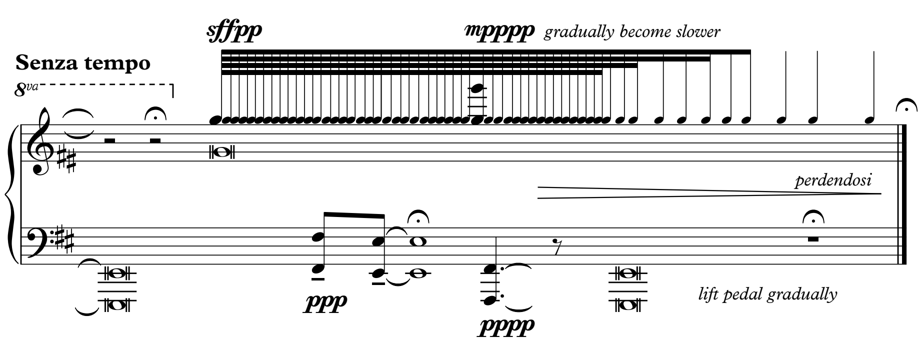

Figure 10: *Final Bar of Movement II of Musica Ricercata. Reprinted by permission of ©SCHOTT MUSIC, Mainz – Germany.*

VII. Conclusions

Musica Ricercata is nowadays, as several other of Ligeti's works, very famous. Although a work of Ligeti's early years as a composer, and tending to a prescient minimalism, some structures and indications in the score of *Musica Ricercata* make it a piece with some characteristics which are beyond the mathematical approach of Information Theory.

Of course, mainly for 20th-century music, any kind of musical analysis can be only partial and Information Theory based analysis is not different in this aspect. However, this work on the first two movements of *Musica Ricercata* is an example in which Information Theory can be useful for analysis of non-tonal music, since its quantitative approach explicitly shows the relative complexity of different musical structures and gestures and in which tonal aspects are not relevant.

As we mentioned in III, Information Theory based analysis depends strongly on the chosen variables for analysis. In this work, while our analysis of MR1 is concentrated on rhythm patterns of pitch octaves, which implies in some timbre variations, that one of MR2 explores more the pitch content.

It is worth to stress that we can make analysis of combined symbols from two or more alphabets with different distributions of probability. In this case, other Information Theory parameters are useful such as Joint Entropy, Conditional Entropy, Mutual Information and others [17].

References

[1] Griffiths, P. (2006). Gyorgy Ligeti, Central-European Composer of Bleakness and Humor, Dies at 83. https://www.nytimes.com/2006/06/13/arts/music/13ligeti.html. Last visited: Feb/16/2019.

[2] Meredith, D. (2016). *Computational Music Analysis*. Springer International Publishing.

[3] Jacoby, N., N. Tishby, and D. Tymoczko (2015). An Information Theoretic Approach to Chord Categorization and Functional Harmony. *Journal of New Music Research*, Vol. 44, No. 3, pp. 219–244.

[4] Kerékfy, M. (2008). A "New Music" from Nothing': György Ligeti's Musica Ricercata. *Studia Musicologica*, Vol. 49, No. 3/4, pp. 203–230.

[5] Steinitz, R. (2003). *György Ligeti: Music of Imagination*. Northeastern University Press.

[6] Grantham, D. (2014). Ligeti's Early Experiments in Compositional Process: Simple Structures in "Musica Ricercata". Master's thesis, University of North Texas.

[7] Ligeti, G. (2013). *György Ligeti: L'atelier du Compositeur*. Contrechamps Éditions.

[8] Temperley, D. (2007). *Music and Probability*. The MIT Press.

[9] Clendinning, J. P. (1993). The Pattern-Mecannico Compositions of György Ligeti. *Perspectives of New Music*, Vol. 31, No. 1, pp. 192–234.

[10] Varga, B. A. (2013). Gyögy Ligeti 1923-2006. In *From Boulanger to Stockhausen, Interviews and a Memoir*, Boydell and Brewer, University of Rochester Press, pp. 26–57.

[11] Prausnitz, F. (2002). *How a "Difficult" Composer Got That Way*. Oxford University Press.

[12] Hankerson, D., G. A. Harris, and P. D. Johnson Jr. (2003). *Introduction to Information Theory and Data Compression*. Discrete Mathematics and its Applications. CRC Press Company.

[13] Varèse, E. and C. Wen-chung (1966, Autumn - Winter). The Liberation of Sound. *Perspectives of New Music*, Vol. 5, No. 1, pp. 11–19.

[14] Huron, D. (2006). *Sweet Anticipation*. The MIT Press.

[15] Applebaum, D. (2008). *Probability and Information*. Cambridge University Press.

[16] Maia Jr., A. Clapping Music: Complexity and Information in Reich's Rhythm Space. *Perspectives of New Music* (accepted to publication)

[17] Cover, T. M., Thomas, J. A. (2008). *Elements of Information Theory*. John Wiley & Sons Inc.

FEDERAL UNIVERSITY OF RIO DE JANEIRO

Denise Pires de Carvalho
Rector

Carlos Frederico Leão Rocha
Vice Rector

Denise Maria Guimarães Freire
Pro-Rector of Graduate and Research

CENTER OF LITERATURE AND ARTS

Cristina Grafanassi Tranjan
Dean

SCHOOL OF MUSIC

Ronal Silveira
Director

Marcelo Jardim
Vice-director

Maria José Di Cavalcanti
Associate Director of Undergraduate Education

Marcelo Jardim
Director of the Cultural Artistic Sector

Andrea Adour
Director of Extension Course

João Vidal
Coordinator of the Graduate Program in Music

Aloysio Fagerlande
Coordinator of the Graduate Program Professional in Music

MusMat Research Group

Members:

- Adriel Viturino
- Alexandre Avellar
- Alexandre Ferreira
- Alexandre Schubert
- André Codeço
- André Pinto
- Bernardo Ramos
- Carlos Almada (leader / coordinator)
- Claudia Usai
- Daniel Moreira (coordinator)
- Daniel Mussatto
- Desirée Mayr
- Eduardo Cabral
- Erico Bonfim
- Fabio Adour
- Fabio Monteiro
- Filipe de Matos Rocha
- Gabriel Mesquita
- Helder Oliveira
- Igor Chagas
- João Penchel
- Jorge Santos
- Leandro Chrispim
- Liduino Pitombeira (coordinator)
- Max Kühn
- Marcel Castro-Lima
- Pauxy Gentil-Nunes (coordinator)
- Pedro Miguel de Moraes
- Pedro Proença
- Pedro Zizels Ramos
- Rafael Fortes
- Rafael Soares Bezerra
- Roberto Macedo
- Rodrigo Pascale
- Rodrigo Furman
- Sérgio Ribeiro
- Vinicius Braga