MusMat

Brazilian Journal of Music and Mathematics

[m]

MusMat

Brazilian Journal of Music and Mathematics

Vol. V - No. 2

MusMat Research Group (Ed.)

[m]

MusMat • Brazilian Journal of Music and Mathematics

Contents

Foreword

We are glad to announce the release of the second number of the fifth volume of *MusMat – Brazilian Journal of Music and Mathematics*. This issue presents five original articles discussing different aspects of the intersection between Music and Mathematics. **Juan Sebastián Arias-Valero**, **Octavio Alberto Agustín-Aquino**, and **Emilio Lluis-Puebla** present a theoretical description of a model based on the generalization of first-species counterpoint considering arbitrary rings, which results in a broader mathematical theory for contrapuntal intervals. **Carlos Mathias** and **Carlos Almada** introduce an original proposal to encode timelines as univocal integers by using arithmetic mapping so that drum-set timelines are encoded by using Gödel's Numbering algorithm. **Paul Lombardi** presents an interesting discussion on feathered beams, examining their concept and notation to propose a graphing system to deconstruct them using examples from George Crumb's *Night Music I*. **Hugo Carvalho** proposes a new tool for performing time-frequency analysis on audio signals, the probabilistic spectrogram, that may allow for probabilistic interpretations related to the Discrete Fourier Transform and also the creation of new features for audio signal processing and music information retrieval. Finally, **Juan Sebastián Arias-Valero and Emilio Lluis-Puebla** present a specific and didactic application for gestural presheaves in the language of abstract gestures, dealing specifically with the relation thereof to the Yoneda embedding and Mazzola's idea and gestural sheaves, demonstrating the application in Mozart and Beethoven. In this issue, we are also glad to inaugurate a new section with an interview with **Severine Neff**, discussing her work on Schoenberg's music and theory.

Daniel Moreira
December 2021

On First-Species Counterpoint Theory

*Juan Sebastián Arias-Valero

Universidad Nacional Autónoma de México

jsariasv1@gmail.com

Orcid: 0000-0001-6812-7639

Octavio Alberto Agustín-Aquino

Universidad Tecnológica de la Mixteca

octavioalberto@mixteco.utm.mx

Orcid: 0000-0002-0556-6236

Emilio Lluis-Puebla

Universidad Nacional Autónoma de México

lluisp@unam.mx

Orcid: 0000-0002-2633-5955

Abstract: *We generalize first-species counterpoint theory to arbitrary* rings *and obtain some new* counting and maximization results *that enrich the theory of admitted successors, pointing to a* structural approach, *beyond computations. The generalization encompasses an* alternative theory *of contrapuntal intervals. We also propose several* variations *of the model that intend to deepen into its principles. The original* motivations *of the theory, as well as all* technical passages, *are carefully reviewed so as to provide a complete exposition.*

Keywords: *Counterpoint. Rings. Modules. Combinatorics.*

1. Introduction

This article is a theoretic exercise on a model of *first-species counterpoint* and the mathematical support of the musicological and computational analysis in [6]. This kind of counterpoint is the simplest one and the didactic basis of *Renaissance counterpoint*, as taught by Johann

1

J. Fux in [8]. The model was introduced by Mazzola in [12], where the discussion confined to the case when the ground ring is $\mathbb{Z}_{12}$; a ring that can be used to model the algebraic behavior of the twelve intervals between tones in the *chromatic scale* of Western musical tradition. Then, the model was re-exposed with some additional computational results by Hichert in [11, Part VII]. Further generalizations to the case when the rings are of the form $\mathbb{Z}_n$ were considered in [1, 9], and then included in a collaborative compendium of mathematical counterpoint theory and its computational aspects [3]. The motivation for such a generalization to $\mathbb{Z}_n$ was the existence of *microtonal scales* with more than twelve tones, which have been used for making real music [2].

The contributions of this article to the previous theory are the following.

- Complete re-exposition of the model with a careful review of all its principles, clarifying important points, like the contrapuntal symmetry definition (Section 4).
- Generalization of the model to *noncommutative rings*. This offers the possibility of defining counterpoint notions on noncommutative rings of generalized intervals, which can be worth exploring in musical terms; see Example 3.10 and Section 9.
- Some theoretic advances on previous computational approaches. They include the determination of self-complementary dichotomies with the *groupoid of intervals* (Section 3.1), a criterion for the strength of dichotomies (Lemma 3.4), the *counting formulas* for successors sets cardinalities (Section 6), a *maximization criterion* for these cardinalities (Section 6.2), and the derivative admitted successors computation, *by hand*, for Renaissance counterpoint (Section 9).
- Some *variations* of the model. First, our generalization includes an alternative model of contrapuntal intervals, namely the *product ring $R \times R$*, which takes the place of the dual numbers ring $R[\epsilon]$. These structures only differ in their products; the first one regards intervallic variations as elements of R, whereas the second one treats these variations as infinitesimals; see Section 6.2. The other two variations intend to deepen into the second principle of the model, and coincide in the case of Renaissance counterpoint (Section 11).

We organize this paper as follows. We start with the basic definitions: symmetries (natural operations in music) in Section 2, hierarchy of dichotomies and polarities (consonance/dissonance partitions and their symmetries) in Section 3, and contrapuntal intervals rings in Section 3.5. In particular, in Section 3.4, we prove that several dichotomies are strong. Then, in Sections 4.1, 4.2, and 4.3 we discuss the principles that lead to a formal definition of first-species counterpoint in Section 4.4. The remaining Section 4 develops the mathematical details of the principles. In Section 5, we simplify this definition, by reducing it to the case when the cantus firmus is 0 and use a smaller group of symmetries. Based on this simplification, we prove that, for the dual numbers case, it is always possible to find successors of a given consonant interval. This result is the *Little theorem of counterpoint* (Section 7). A crucial lemma (Lemma 6.1) leads us to the counting formulas and the maximization criterion in Section 6. With these tools, in Section 8, we obtain the admitted successors of consonant intervals in Renaissance counterpoint. In Section 9, we explore a noncommutative notion of counterpoint. In Section 10, we prove that our contrapuntal symmetry condition 2 is equivalent to Mazzola's original one in the commutative case, although it is weaker in general. Finally, in Section 11 we discuss two conceptual variations of the model and close this article by proving that they coincide in the case of the Renaissance dichotomy. Section 12 contains some conclusions and suggestions for further research. The appendix includes the proof of a variation of the *rearrangement inequality*, which is crucial for the maximization criterion.

Throughout this paper, R denotes an arbitrary ring with unity, not necessarily commutative. When needed, commutativity hypothesis on R is explicitly stated. Similarly, modules considered in this paper are right R-modules, unless we indicate otherwise. So as to easily locate a numbered

equation in this paper, we urge the reader to use hyperlinks.

2. Ring symmetries

The main goal of this section is to recall basic facts about rings and their symmetries. An introduction to noncommutative rings is [14, Chapter B-1].

Affine homomorphisms are the natural correspondences that occur in music. They are the formalization that encompasses musical transformations like transposition and inversion in the ring $\mathbb{Z}_{12}$ of tones.[1]

Let R be a ring. Given two R-modules M and N, an *affine homomorphism* from M to N is the composite $e^a \circ f$, where $f : M \longrightarrow N$ is an R-homomorphism, $a \in N$, and $e^a : N \longrightarrow N : x \mapsto x + a$ is the *translation* associated with a. We write a typical affine homomorphism $e^a \circ f$ as $e^a f$, for short. Observe that the composition rule for affine homomorphisms is $(e^a f) \circ (e^b g) = e^{f(b)+a} fg$.

The monoid (with respect to the composition) $\text{End}_R(M)$ of R-endomorphisms of an R-module M is a ring with the usual sum of homomorphisms. However, the monoid of affine endomorphisms of M is not a ring since the distributivity fails, except when M is the trivial module. We denote by $\text{Aut}_R(M)$ the group of R-automorphisms of an R-module M.

An affine homomorphism $e^a f : M \longrightarrow N$ is an isomorphism if and only if f is an R-isomorphism.[2] Thus, affine automorphisms of R are of the form $e^a f$, where $f \in \text{Aut}_R(R)$ and $a \in R$. On the other hand, recall that there is a ring isomorphism

$$\text{End}_R(R) \cong R,$$

which identifies an element $r \in R$ with the R-endomorphism obtained by left multiplication with r. This isomorphism restricts to a group isomorphism

$$\text{Aut}_R(R) \cong R^*,$$

which establishes a bijective correspondence between R-automorphisms of R and invertible elements of R. From this discussion we conclude that an affine automorphism of R is the function associated with the linear polynomial $bx + a$, where $b \in R^*$ and $a \in R$.

In what follows, we will call affine automorphisms of R **symmetries** of R and denote the group of symmetries of a ring R by **Sym**(R). We use the notation $e^a b$ for the symmetry associated with $bx + a$. Two symmetries $e^a b$ and $e^{a'} b'$ are equal if and only if $a = a'$ and $b = b'$. Observe that $e^0 1$ is the identity symmetry and that the composition of symmetries is given by the formula

$$e^a b \circ e^{a'} b' = e^{ba'+a} bb'. \tag{1}$$

The inverse of a symmetry $e^a b$ is $e^{-b^{-1}a} b^{-1}$.
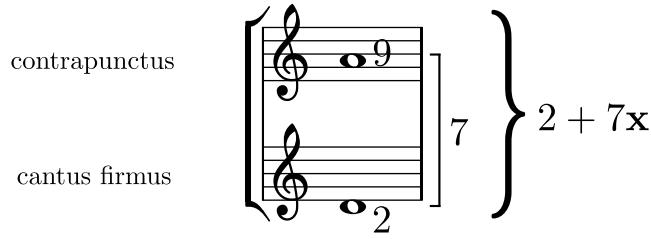
3. Basic motivations and definitions

In Renaissance counterpoint we divide the ring $\mathbb{Z}_{12}$ of **intervals** into two disjoint subsets, namely the set K of *consonances* and the set D of *dissonances*. The consonances are unison, minor third,

---

[1] The ring $\mathbb{Z}_{12}$ models the twelve tones of the chromatic scale and, at the same time, the twelve intervals between these tones. The distinction is usually clear from the context.

[2] First, suppose that $e^b g : N \longrightarrow M$ is the inverse of $e^a f$. Then $(e^a f)(e^b g) = e^{f(b)+a} fg = id_N$ and $(e^b g)(e^a f) = e^{g(a)+b} gf = id_M$. Therefore, $fg = id_N$, $gf = id_M$, and $b = -g(a)$. Conversely, if $fg = id_N$ and $gf = id_M$, then $(e^a f)(e^{-g(a)} g) = id_N$ and $(e^{-g(a)} g)(e^a f) = id_M$.

Figure 1: *A contrapuntal interval.*

major third, perfect fifth, minor sixth, and major sixth. The dissonances are minor second, major second, perfect fourth, tritone, minor seventh, and major seventh. The corresponding mathematical definitions are

$$K = \{0,3,4,7,8,9\} \text{ and } D = \{1,2,5,6,10,11\}.$$

A composition of *first-species counterpoint* consists of two voices, *cantus firmus* and *discantus*, whose notes have the same length and where each interval between the voices is a consonance. See [8, pp. 19-29] and [6, Section 2] for details.

The germ of Mazzola's model is the observation that there is a *unique* symmetry p of the ring $\mathbb{Z}_{12}$, namely $e^2 5$, that sends consonances to dissonances. In fact,

$$e^2 5(0) = 2, \ e^2 5(3) = 5, \ e^2 5(4) = 10, \ e^2 5(7) = 1, \ e^2 5(8) = 6, \ e^2 5(9) = 11.$$

We postpone the proof of the uniqueness to Section 3.4. Moreover, if a partition $\{K', D'\}$ of $\mathbb{Z}_{12}$ satisfies the conditions that $e^2 5$ is the unique symmetry in $\mathbf{Sym}(\mathbb{Z}_{12})$ sending K' to D' and that $7K'$ is a multiplicative monoid, then $K' = K$ and $D' = D$ according to [10, Section 13.1]. Thus, we say that these two conditions *characterize* $\{K, D\}$.

Up to now, we have the intervals, consonances, and dissonances of counterpoint, but we need to model the voices in a composition. We achieve this by means of the **countrapuntal intervals** ring $\mathbb{Z}_{12}[\mathbf{x}]$, which consists of all linear polynomials $c + d\,\mathbf{x}$, with $c, d \in \mathbb{Z}_{12}$, where $\mathbf{x}$ is a dummy variable that will be further discussed in Section 3.5. A contrapuntal interval $c + d\,\mathbf{x}$, represents a pitch class c, of the cantus firmus, together with the interval d between c and the pitch class $c + d$ from the superior discantus. As already said, in a piece of Renaissance counterpoint, it is mandatory that d be a consonance. For instance, the contrapuntal interval $2 + 7\,\mathbf{x}$ in $\mathbb{Z}_{12}[\mathbf{x}]$ comes from the musical example in Figure 1.

We have a partition of the ring $\mathbb{Z}_{12}[\mathbf{x}]$ into **contrapuntal consonances and dissonances**, namely $\{K[\mathbf{x}], D[\mathbf{x}]\}$, where $X[\mathbf{x}]$ consists of all $c + d\,\mathbf{x} \in \mathbb{Z}_{12}[\mathbf{x}]$ with $d \in X$ for $X = K, D$. We would want properties for this partition analogous to those that characterize $\{K, D\}$. Certainly, the symmetry $e^{2\mathbf{x}} 5$ of $\mathbb{Z}_{12}[\mathbf{x}]$ is a quite natural[3] extension of $e^2 5$ sending $K[\mathbf{x}]$ to $D[\mathbf{x}]$ since $e^{2\mathbf{x}} 5$ looks simple and acts on the interval part d of a contrapuntal interval $c + d\,\mathbf{x}$ just as $e^2 5$. But in this case *it is not the unique* sending $K[\mathbf{x}]$ to $D[\mathbf{x}]$. For example, $e^{2\mathbf{x}+1} 5$ also does. However, there is an induced *local uniqueness property* that characterizes the partition $\{K[\mathbf{x}], D[\mathbf{x}]\}$, as discussed in Section 4.2.

The first-species compositions deal with finite sequences of contrapuntal consonances with cantus firmus and discantus in a given *scale*; see Figure 3. However, not all sequences are valid since counterpoint has some *rules*. In the model to be exposed in the following sections, we aim to predict when a given contrapuntal consonance η can follow another one ξ, by means of the *admitted successor* concept. We base the latter on the *alternation principle* (Section 4.1), the local uniqueness property (Section 4.2), and the *variety principle* of counterpoint (Section 4.3).

---

[3]Precise statements of this naturality claim are Propositions 4.7 and 10.2.

3.1. Dichotomies

We start to develop our theory by formalizing the common properties of the rings of intervals $\mathbb{Z}_{12}$ and $\mathbb{Z}_{12}[\mathbf{x}]$ that are relevant for counterpoint.

Let R be a ring.

- A partition $\{K, D\}$ of R is a **dichotomy** of R if $|K| = |D|$. Note that a finite ring R has dichotomies if and only if its cardinality is even.
- A **self-complementary dichotomy**[4] of R is a triple (K, D, p), where $\{K, D\}$ is a partition of R, p is a symmetry of R, and $p(K) = D$. Note that necessarily $\{K, D\}$ is a dichotomy of R.

The triples

$$(K, D, e^2 5) \text{ and } (K[\mathbf{x}], D[\mathbf{x}], e^{2\mathbf{x}} 5)$$

from Renaissance counterpoint are self-complementary dichotomies. The dichotomy

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

of $\mathbb{Z}_{12}$ is not part of a self-complementary dichotomy of $\mathbb{Z}_{12}$ as proved at the end of this section, so it is not a good candidate for providing notions of contrapuntal dissonance/consonance with the characterizing properties of the Renaissance dichotomy.

We would like to construct self-complementary dichotomies other than the basic one of Renaissance counterpoint, so as to create new notions of counterpoint. Given an arbitrary finite ring R of even cardinality, we could start with a symmetry p of R and aim to construct a self-complementary dichotomy (K, D, p). We pick some $k_1 \in R$ to be in K, and define $d_1 := p(k_1)$, with d_1 to be in D. Then, if $|R| > 2$, we take some $k_2 \in R \setminus \{k_1, d_1\}$ to be in K and define $d_2 := p(k_2)$ to be in $D \setminus \{d_1\}$, and so on. Note that this process produces a dichotomy $\{K, D\}$ provided *p has no fixed points*. We say that a symmetry p without fixed points is a **derangement**.

The following proposition contains some concluding properties of self-complementary dichotomies.

Proposition 3.1. *Each self-complementary dichotomy (K, D, p) of R has the following properties:*

1. *The identity $p(D) = K$ holds, so (D, K, p) is a self-complementary dichotomy.*

2. *The symmetry p is a derangement.*

Proof. 1. Since p is a bijection, the properties of the inverse image imply that $p(D) = p(R \setminus K) = R \setminus p(K) = R \setminus D = K$.

2. Let $x \in R$. If $x \in K$, then $x \neq p(x) \in D$. If $x \in D$, then, by 1, $x \neq p(x) \in K$. □

Thus, *self-complementary dichotomies are strongly related to derangements.*

The following is a short reflection about the problem of deciding when a dichotomy $\{K, D\}$ of a ring R is not self-complementary.

Let $X \subseteq R$. We associate with X a categorical **groupoid**[5] **of intervals** $\mathcal{G}(X)$ as follows. Its set of objects is X. For each $x, y \in X$, there is a unique morphism from x to y, namely the triple $(x, y, y - x)$. We define the composition by

$$(y, z, z - y) \circ (x, y, y - x) = (x, z, z - y + y - x) = (x, z, z - x).$$

---

[4]This definition is essentially the same of an **autocomplementary** dichotomy in [11, Definition 92].

[5]A groupoid is a category [11, Appendix G.1] whose morphisms are invertible. This example is an invitation to the structural advantages of category theory. The groupoid of intervals is a very simple instance of a category in music. However, it is the unique appearance of a category in this paper.

The identities are of the form $(x, x, 0)$. The inverse of $(x, y, y - x)$ is $(y, x, x - y)$.

Given a self-complementary dichotomy $(K, D, e^a b)$ of R, we have an induced groupoid isomorphism [11, p. 1118] $F : \mathcal{G}(K) \longrightarrow \mathcal{G}(D)$ sending[6] $(x, y, y - x)$ to $(e^a b(x), e^a b(y), b(y - x))$, which acts linearly on intervallic variations. Hence, *a self-complementary dichotomy induces a linear correspondence between the intervals of K and D*.

This fact can be used to prove that the dichotomy $\{K, D\}$ with $K = \{0, 1, 3, 6, 8, 11\}$ and $D = \{2, 4, 5, 7, 9, 10\}$ is not self-complementary. In fact, the sequence $(11, 0, 1)$ of elements of K has the maximum length among sequences $(x_1, \ldots, x_n)$ such that $x_{i+1} - x_i = 1$ for $i = 0, \ldots n - 1$. If a symmetry $e^a b$ sends K to D, then the groupoid isomorphism F sends $(11, 0, 1)$ to a sequence (d_1, d_2, d_3) in D with maximum length among sequences $(y_1, \ldots, y_n)$ such that $y_{i+1} - y_i = b$. But the sequences of maximum length with $y_{i+1} - y_i = b$ are $(4, 5)$ and $(9, 10)$(for $b = 1$), $(5, 4)$ and $(10, 9)$(for $b = 11$), $(4, 9, 2, 7)$ (for $b = 5$), and $(7, 2, 9, 4)$ (for $b = 7$), whose length is not 3; a contradiction.

3.2. Strong dichotomies and polarities

We formalize the additional uniqueness property of the Renaissance dichotomy $(K, D, e^2 5)$ with the following definitions. Let R be a ring. A self-complementary dichotomy (K, D, p) of R is a **strong dichotomy** of R if p is the unique symmetry of R such that $p(K) = D$. In such a case, we also say that p is a **polarity**.

As we will show in the next sections, once we have a strong dichotomy of a ring (made up of generalized intervals), we have an induced self-complementary dichotomy of the contrapuntal intervals ring, and an associated theory of admitted successors. Thus, strong dichotomies lead to new *counterpoint worlds* for composing non-traditional counterpoint.

We need to determine whether a given self-complementary dichotomy (K, D, p) is strong. We approach this problem by determining the number of symmetries sending K to D. Given another symmetry q with $q(K) = D$, note that

$$q^{-1} \circ p(K) = q^{-1}(D) = K$$

and hence $q^{-1} \circ p$ is in the stabilizer $\theta(K)$, where

$$\theta(K) := \{g \in \mathbf{Sym}(R) \mid g(K) = K\}.$$

This suggests the right action of $\theta(K)$, by composition, on the set $\mathbf{Sym}(K, D)$ of symmetries of R sending K to D, as shown in the following diagram.

$$\circ : \quad \mathbf{Sym}(K, D) \times \theta(K) \quad \longrightarrow \quad \mathbf{Sym}(K, D)$$
$$(q, g) \quad \longmapsto \quad q \circ g$$

Every $q \in \mathbf{Sym}(K, D)$ is in the same orbit of p because $q^{-1} \circ p \in \theta(K)$, so the action is transitive. Also, the stabilizer of p under this action is the identity since $p \circ g = p$ implies $g = id_R$. Hence, the following proposition.

Proposition 3.2. *Let (K, D, p) be a self-complementary dichotomy of a ring R. There is a bijective correspondence between $\mathbf{Sym}(K, D)$ and $\theta(K)$. The bijection sends an element $g \in \theta(K)$ to $p \circ g$.*

Proof. The function $\theta(K) \longrightarrow \mathbf{Sym}(K, D) : g \mapsto p \circ g$ is surjective since the action is transitive and is injective since the stabilizer of p is trivial. $\square$

---

[6]It is the unique possible definition if the correspondence on objects is $e^a b$.

Thus, we have reduced the study of **Sym**(K, D) to that of the stabilizer group $\theta(K)$ of K, which coincides with $\theta(D)$ whenever $\{K, D\}$ is a partition of R. Besides, we have determined when a self-complementary dichotomy is strong, as established in the following corollary. We say that a partition $\{K, D\}$ of R is **rigid** if $\theta(K)$ is the trivial group. Similarly, we say that a subset K of R is **rigid** if $\theta(K)$ is trivial. Note that in a rigid partition $\{K, D\}$ of R both K and D are rigid.

Corollary 3.3. *A self-complementary dichotomy (K, D, p) is a strong dichotomy if and only if $\{K, D\}$ is rigid.*

Now, our objective is to prove a pair of lemmas that help us to decide whether a given dichotomy is rigid, by discarding non-identity symmetries that could be in $\theta(K)$. These lemmas could be a first approximation to a more conceptual argument.[7] Consider the left action $\cdot$, by multiplication, of the group R^* on R.

Lemma 3.4. *Let $\{K, D\}$ be a dichotomy of a ring R. Suppose that $a \in R$ and that C is either K or D. If there is $r \in R$ such that*

   1. *the orbit R^*r of r under the action $\cdot$ is contained in C and*

   2. *$r + a \notin C$,*

then $e^a b \notin \theta(K)$ for each $b \in R^*$.

Proof. If $r + a \notin C$ for r as above, then for each $b \in R^*$, $b(b^{-1} \cdot r) + a \notin C$ with $b^{-1} \cdot r \in R^*r \subseteq C$, and hence $e^a b \notin \theta(C) = \theta(K)$. □

In particular, if $r = 0$, we obtain the following result.

Lemma 3.5. *Suppose that C is either K or D and that $0 \in C$. Each symmetry of the form $e^a b$ with $a \in R \setminus C$ is not in $\theta(K)$.*

Before using these lemmas to check that some dichotomies are strong, we first study a possible way to construct self-complementary dichotomies that are good candidates for strong ones.

3.3. Constructing strong dichotomies: quasipolarities

A symmetry p of a ring R is **involutive** if $p \circ p = id$. If $p = e^a b$, by Equation (1), this condition is equivalent to $e^{ba+a} b^2 = e^0 1$, that is, to $ba + a = 0$ and $b^2 = 1$. In the next proposition we prove that all polarities are involutive, so if we want a polarity, we could start with an involutive derangement, then define a suitable self-complementary dichotomy following the strategy in Section 3.1, and finally check whether the uniqueness property holds.

Proposition 3.6. *If a self-complementary dichotomy (K, D, p) is strong, then p is involutive.*

Proof. If (K, D, p) is a self-complementary dichotomy and p has the uniqueness property, then $p \circ p(K) = p(D) = K$ by 1 in Proposition 3.1, and hence $p \circ p = id$ by Corollary 3.3. □

For example, the Renaissance polarity $e^2 5$ is involutive. The converse of this proposition does not hold since, in the Renaissance dichotomy $(K[\mathbf{x}], D[\mathbf{x}], e^{2\mathbf{x}}5)$, $e^{2\mathbf{x}}5$ is involutive, but the dichotomy is not strong as already observed.

---

[7]This conceptual argument can be related to the groupoid of intervals of K, briefly introduced in Section 3.1.

We define a **quasipolarity** as an involutive derangement. A **quasipolarization** of R is a self-complementary dichotomy (K, D, p) of R where p is involutive. Therefore, in view of Proposition 3.1, a self-complementary dichotomy (K, D, p) is a quasipolarization if and only if p is a quasipolarity.

Note that not all self-complementary dichotomies are quasipolarizations. For example, the self-complementary dichotomy $(\{0, 2, 4, 6, 8, 10\}, \{1, 3, 5, 7, 9, 11\}, e^1)$ of $\mathbb{Z}_{12}$ is not a quasipolarization since $e^1 \circ e^1 = e^2 \neq id$.

Thus, the construction of quasipolarizations is a good beginning if we want a strong dichotomy, though quasipolarizations need not be strong dichotomies.

3.4. Examples of strong dichotomies

In this section, we use Lemmas 3.4 and 3.5 to open up a series of counterpoint worlds. First, consider the case when R is the commutative ring $\mathbb{Z}_{12}$ and different examples of quasipolarizations that we will check to be strong dichotomies. The set of orbits of the action of the multiplicative group $\{1, 5, 7, 11\}$ of invertible elements of $\mathbb{Z}_{12}$ on $\mathbb{Z}_{12}$ is

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

In the following example we prove the claim that the Renaissance quasipolarization is a strong dichotomy.

Example 3.7 (Renaissance counterpoint). Consider the dichotomy $(K, D, e^2 5)$, where

$$K = \{0, 3, 4, 7, 8, 9\} \text{ and } D = \{1, 2, 5, 6, 10, 11\}.$$

Let us prove that the dichotomy is strong. According to Corollary 3.3, it is enough to show that our dichotomy is rigid, or equivalently, that $\theta(K) = \{e^0 1\}$. In fact, by Lemma 3.5, $e^a b \notin \theta(K)$ for each symmetry $e^a b$ with $a \in D$. The cases when $a \in K$ remain. If $a = 0$, then since $5 \times 7 = 11 \notin K$, $7 \times 7 = 1 \notin K$, and $11 \times 7 = 5 \notin K$, we conclude that the symmetries of the form $e^0 b$ with $b = 5, 7, 9$ are not in $\theta(K)$. Finally, since

$$3 + 8 = 8 + 3 = 11 \notin K, \ 4 + 9 = 9 + 4 = 1 \notin K, \text{ and } 3 + 7 \notin K,$$

Lemma 3.4 implies that all symmetries $e^a b$ with $a \in \{3, 4, 7, 8, 9\}$ are not in $\theta(K)$. We have exhausted all possibilities, except the identity, and hence $\theta(K)$ is trivial. ◇

Example 3.8 (Scriabin's mystic chord, see [5]**).** Consider the dichotomy $(K, D, e^5 11)$, where $K = \{0, 2, 4, 6, 9, 10\}$ and $D = \{1, 3, 5, 7, 8, 11\}$. Let us show that $e^5 11$ is a polarity by using the same strategy of the preceding example.

As before, $e^a b \notin \theta(K)$ for each symmetry $e^a b$ with $a \in D$, and we need to discard the cases when $a \in K$. If $a = 0$, then since $5 \times 4 = 8 \notin K$, $7 \times 9 = 3 \notin K$, and $11 \times 4 = 8 \notin K$, the symmetries of the form $e^0 b$ with $b = 5, 7, 11$ are not in $\theta(K)$. Also, since

$$6 + 2 = 2 + 6 = 8 \notin K, \ 5 + 4 = 9 \notin D, \ 2 + 9 = 11 \notin K, \text{ and } 10 + 10 = 8 \notin K,$$

Lemma 3.4 implies that all symmetries $e^a b$ with $a \in \{2, 4, 6, 9, 10\}$ are not in $\theta(K)$. Hence, $\theta(K)$ is trivial. ◇

The next computation establishes the existence of at least a strong dichotomy in $\mathbb{Z}_{2k}$ for each $k \geq 3$. These dichotomies are the beginning of a lot of microtonal counterpoint worlds, which were first studied in [1]. Recall that all invertible elements of $\mathbb{Z}_{2k}$ are odd since they are coprime with $2k$.

Example 3.9 (Cf. [3, Proposition 2.1] and [1, Proposition 2.6]). Let[8] $k \geq 3$. In $\mathbb{Z}_{2k}$, consider the dichotomy $(K, D, e^{-1}(-1))$, where

$$K = \{0, 1, 3, \ldots, 2k - 5, 2k - 3\} \text{ and } D = \{-1, 2k - 2, 2k - 4, \ldots, 4, 2\}.$$

We already know (Lemma 3.5) that $e^a b \notin \theta(K)$ for each symmetry $e^a b$ with $a \in D$. If $a = 0$, then $b(-1) = -b \neq -1$ with $-b$ odd whenever $b \in \mathbb{Z}_{2k}^* \setminus \{1\}$ and hence $-b \notin D$, so the symmetries of the form $e^0 b$ with $b \in \mathbb{Z}_{2k}^* \setminus \{1\}$ are not in $\theta(D) = \theta(K)$. Also, since all orbits of even numbers, except the orbit $\{0\}$ of 0, under the action of $\mathbb{Z}_{2k}^*$ on $\mathbb{Z}_{2k}$ are contained in D, the equations

$$2 + 1 = 3 \notin D, \ldots, 2 + (2k - 5) = 2k - 3 \notin D, \text{ and } 4 + (2k - 3) = 1 \notin D,$$

imply (Lemma 3.4) that all symmetries $e^a b$ with $a \in \{1, 3, \ldots, 2k - 5, 2k - 3\}$ are not in $\theta(K)$. Hence, $\theta(K)$ is trivial. ◇

Next, we provide a noncommutative example.

Example 3.10 (**A noncommutative strong dichotomy**). Let R be the noncommutative ring of all upper triangular matrices 2×2 with entries in $\mathbb{Z}_2$. The symmetry e^I, where I is the identity matrix, is a quasipolarity, so we can construct self-complementary dichotomies with the procedure of Section 3.1. A possible choice is the dichotomy with $K = \{\mathbf{0}, A_1, A_2, A_3\}$, where

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \text{ and } A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

In this case, $R^* = \{I, A_3\}$ and A_1 and A_2 are invariant under the left action of R^*. Let us show that K is rigid. As usual, $e^A B \notin \theta(K)$ for each symmetry $e^A B$ with $A \in D$. The cases when $A \in K$ remain. If $A = 0$, then since $A_3 A_3 = I \notin K$, the symmetry $e^0 A_3$ is not in $\theta(K)$. Also,

$$A_2 + A_1 = A_1 + A_2 \notin K, \text{ and } A_1 + A_3 \notin K,$$

so, by Lemma 3.4, all symmetries $e^A B$ with $A \in \{A_1, A_2, A_3\}$ are not in $\theta(K)$. Thus, $\theta(K)$ is trivial, and $(K, R \setminus K, e^I)$ is a strong dichotomy. The reader can obtain other examples of strong dichotomies from quasipolarities of R, whenever the chosen K is not invariant under the left action of A_3, by using a similar argument. ◇

Finally, note that there is no strong dichotomy of $\mathbb{Z}_4$ to define a counterpoint world with a four-tone scale (which can be thought of as a diminished seventh chord). In fact, the possible dichotomies $\{\{0, 1\}, \{2, 3\}\}$, $\{\{0, 2\}, \{1, 3\}\}$, and $\{\{0, 3\}, \{1, 2\}\}$ are invariant under $e^1(-1)$, e^2, and $e^{-1}(-1)$, respectively, so they are not rigid and hence not strong (Corollary 3.3).

3.5. The contrapuntal intervals ring

Given an arbitrary ring R, we can construct the polynomial ring $R[x]$ and the *two-sided* ideal $\langle p(x) \rangle$ consisting of all (left or right) multiples of $p(x)$, where $p(x)$ is the polynomial $x^2 - \alpha x$ for some fixed $\alpha = 0, 1$. Thus, we have the quotient ring[9] $R[x] / \langle p(x) \rangle$, which we call *the contrapuntal intervals ring* associated with R. If we denote by $\mathbf{x}$ the class of x, then each element of this ring can uniquely be written[10] as $c + d \mathbf{x}$, and hence it makes sense to denote the ring by $R[\mathbf{x}]$. Thus,

---

[8]Thanks to the hypothesis $k \geq 3$, the six distinct elements 0, 1, 3, −1, 2, and 4 are in $\mathbb{Z}_{2k}$.

[9]See [14, p. 278] for details.

[10]In this representation we use the fact that R can be regarded as a subring of $R[x] / \langle p(x) \rangle$. Also, the representation follows from the fact that each polynomial $f(x)$ in $R[x]$ can uniquely be written (division algorithm) as $q(x)p(x) + r(x)$, where either $r(x) = 0$ or $deg(r(x)) < deg(p(x)) = 2$.

$R[\mathbf{x}]$ consists of all linear polynomials $c + d\,\mathbf{x}$ subject to the relation $\mathbf{x}^2 = \alpha\,\mathbf{x}$. The ring $R[\mathbf{x}]$ is noncommutative if R is.

Explicitly, the operations in $R[\mathbf{x}]$ are defined by

$$(c + d\,\mathbf{x}) + (c' + d'\,\mathbf{x}) = c + c' + (d + d')\,\mathbf{x}$$

and

$$(c + d\,\mathbf{x})(c' + d'\,\mathbf{x}) = cc' + (cd' + dc' + dd'\alpha)\,\mathbf{x}.$$

In the case when $\alpha = 0$, $R[\mathbf{x}]$ is the *dual numbers* ring $R[\epsilon]$, which was the ring used in the original model [12, p. 556]. In the case when $\alpha = 1$, $R[\mathbf{x}]$ is isomorphic to the *product* ring $R \times R$, whose elements can be interpreted in musical terms as pairs consisting of cantus firmus and discantus notes, by means of the assignment

$$R \times R \longrightarrow R[\mathbf{x}] : (r, s) \mapsto r + (s - r)\,\mathbf{x}.$$

Both variations share the same structure as Abelian groups, but they differ in their products. In the dual numbers case, intervallic variations are *infinitesimals*[11] so $d\epsilon d'\epsilon = dd'\epsilon^2 = 0$, whereas in the other case, intervallic variations are *just elements of R*, so $d\,\mathbf{x}\,d'\,\mathbf{x} = dd'\,\mathbf{x}^2 = dd'\,\mathbf{x}$. The exploration of the latter alternative is one of the contributions of this paper.

An element $c + d\,\mathbf{x} \in R[\mathbf{x}]$ is invertible if and only if c and $c + d\alpha$ are invertible in R. In fact, if $(c' + d'\,\mathbf{x})(c + d\,\mathbf{x}) = 1$ and $(c + d\,\mathbf{x})(c' + d'\,\mathbf{x}) = 1$, then $c'c + (c'd + d'c + d'd\alpha)\,\mathbf{x} = 1$ and $cc' + (cd' + dc' + dd'\alpha)\,\mathbf{x} = 1$, so c is invertible and $cd' + dc' + dd'\alpha = 0$. From the last equation we deduce the following ones.

$$(c + d\alpha)d' + dc' = 0$$
$$(c + d\alpha)d'\alpha + d\alpha c' = 0$$
$$(c + d\alpha)d'\alpha + d\alpha c' + cc' = cc'$$
$$(c + d\alpha)(c' + d'\alpha) = 1$$

Similarly, $c + d\alpha$ is left-invertible. Reciprocally, if c and $c + d\alpha$ are invertible, define $c' = c^{-1}$ and $d' = -(c + d\alpha)^{-1}dc'$. Thus, $c' + d'\,\mathbf{x}$ is a two-sided inverse for $c + d\,\mathbf{x}$.

Alternatively, in the case when $\alpha = 1$, by the isomorphism between $R[\mathbf{x}]$ and $R \times R$, the inverse of $c + d\,\mathbf{x}$ also has the description $c^{-1} + ((c + d)^{-1} - c^{-1})\,\mathbf{x}$.

We have the following proposition regarding dichotomies of $R[\mathbf{x}]$ induced by dichotomies of a ring R. Given subsets $X, Y \subseteq R$, we define

$$X + Y\,\mathbf{x} = \{c + d\,\mathbf{x} \in R[\mathbf{x}] \mid c \in X \text{ and } d \in Y\}.$$

Also, we denote by $X[\mathbf{x}]$ the set $R + X\,\mathbf{x}$.

Proposition 3.11. *Let $\{K, D\}$ be a dichotomy of a ring R.*

1. *The pair $\{K[\mathbf{x}], D[\mathbf{x}]\}$ is a dichotomy of $R[\mathbf{x}]$.*

2. *If $(K, D, e^a b)$ is a self-complementary dichotomy of R, then $(K[\mathbf{x}], D[\mathbf{x}], e^{a\,\mathbf{x}} b)$ is a self-complementary dichotomy of $R[\mathbf{x}]$.*

3. *If $(K, D, e^a b)$ is a quasipolarization of R, then $(K[\mathbf{x}], D[\mathbf{x}], e^{a\,\mathbf{x}} b)$ is a quasipolarization of $R[\mathbf{x}]$.*

As previously commented, strong dichotomies need not induce strong dichotomies on contrapuntal intervals.

---

[11] The reader can found a musical justification of the relation $\epsilon^2 = 0$ in [12, pp. 557-558].

Proof. 1. First, $\{K[\mathbf{x}], D[\mathbf{x}]\}$ is a partition of $R[\mathbf{x}]$ since

$$K[\mathbf{x}] \cup D[\mathbf{x}] = (R + K\mathbf{x}) \cup (R + D\mathbf{x}) = R + (K \cup D)\mathbf{x} = R + R\mathbf{x} = R[\mathbf{x}]$$

and

$$K[\mathbf{x}] \cap D[\mathbf{x}] = (R + K\mathbf{x}) \cap (R + D\mathbf{x}) = R + (K \cap D)\mathbf{x} = R + \varnothing\mathbf{x} = \varnothing.$$

On the other hand, if $f : K \longrightarrow D$ is a bijection, then $K[\mathbf{x}] \longrightarrow D[\mathbf{x}] : r + k\mathbf{x} \mapsto r + f(k)\mathbf{x}$ is a bijection, so $|K[\mathbf{x}]| = |D[\mathbf{x}]|$.

2. This follows from 1 and the equalities

$$
\begin{aligned}
e^{a\mathbf{x}}b(K[\mathbf{x}]) &= \{e^{a\mathbf{x}}b(r + k\mathbf{x}) \mid r \in R \text{ and } k \in K\} \\
&= \{br + (bk + a)\mathbf{x} \mid r \in R \text{ and } k \in K\} \\
&= bR + (e^a b(K))\mathbf{x} \\
&= R + D\mathbf{x} \\
&= D[\mathbf{x}].
\end{aligned}
$$

As to the fourth equality above, note that $bR = R$ since b is invertible.

3. This follows from 2 and the equation

$$e^{a\mathbf{x}}b \circ e^{a\mathbf{x}}b = e^{(ba+a)\mathbf{x}}b^2 = e^0 1,$$

which is a consequence of the identity $e^a b \circ e^a b = e^0 1$, the latter being equivalent to $ba + a = 0$ and $b^2 = 1$. □

4. A MATHEMATICAL MODEL OF FIRST-SPECIES COUNTERPOINT
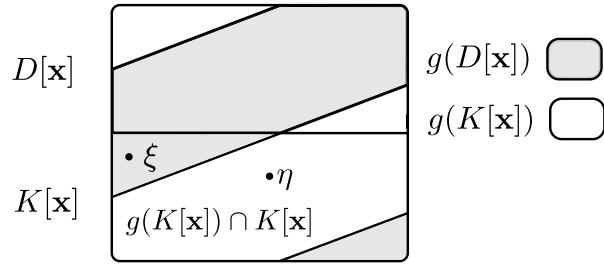
We start with a *strong dichotomy* (K, D, p) of a ring R, where $p = e^a b$, and construct the contrapuntal intervals ring $R[\mathbf{x}]$, which models the two voices of first-species counterpoint. Then, we consider **progressions** (ξ, η) of contrapuntal consonances in $K[\mathbf{x}]$ and aim to determine when they are valid for counterpoint. The three principles in Sections 4.1, 4.2, and 4.3 serve this purpose.

4.1. Alternation

This principle intends to formalize the idea of *tension/resolution* in counterpoint. It requires the existence of a symmetry $g \in \mathbf{Sym}(R[\mathbf{x}])$ such that $\xi \in g(D[\epsilon])$ and $\eta \in g(K[\epsilon])$. In such a case, we say that (ξ, η) is **polarized**. Thus, (ξ, η) is the deformation, by means of g, of a progression from a dissonance to a consonance; see Figure 2. Under a musical intuition, ξ moves to η since dissonances resolve to consonances. However, how do we ensure that deformed consonances and dissonances are consonances and dissonances on their own right? In the following section we solve this problem. In Section 4.5 we give a simple characterization of polarized progressions.

4.2. Local characterization of consonances and dissonances

As observed in Section 3, the strength property of the Renaissance dichotomy characterizes it. In general, given a strong dichotomy (K, D, p) as above, the induced dichotomy of the contrapuntal intervals is not strong, but it has the following local strength property in each **fiber** I_z of $R[\mathbf{x}]$, where $I_z = z + R\mathbf{x}$ and $z \in R$, as shown in Theorem 4.6 below:

Figure 2: *Here, g is a deformation symmetry and η an admitted successor of ξ.*

- There is a unique symmetry $p^z[\mathbf{x}] \in \mathbf{Sym}(R[\mathbf{x}])$ of the form $e^{u+v\mathbf{x}}c$, defined by $p^z[\mathbf{x}] = e^{(1-b)z+a\mathbf{x}}b$, such that $p^z[\mathbf{x}](z + K\mathbf{x}) = z + D\mathbf{x}$.

In the case of the Renaissance strong dichotomy $(K, D, e^2 5)$ of $\mathbb{Z}_{12}$, this condition locally determines the induced dichotomy $\{K[\mathbf{x}], D[\mathbf{x}]\}$. This means that the following two conditions *characterize* $\{z + K\mathbf{x}, z + D\mathbf{x}\}$ among partitions $\{z + K'\mathbf{x}, z + D'\mathbf{x}\}$ of I_z:

1. The symmetry $p^z[\mathbf{x}]$ is the unique $p' \in \mathbf{Sym}(R[\mathbf{x}])$ of the form $e^{u+v\mathbf{x}}c$ such that $p'(z + K'\mathbf{x}) = z + D'\mathbf{x}$.

2. The set $7K'$ is a multiplicative monoid.

In fact, 1 is equivalent to the condition that p is the unique symmetry of $\mathbb{Z}_{12}$ sending K' to D', so according to the characterization of the Renaissance dichotomy in Section 3, the conditions 1 and 2 imply $K' = K$ and $D' = D$. Consequently, if we want to ensure that a certain partition $\{z + K'\mathbf{x}, z + D'\mathbf{x}\}$ of I_z offers a relaxed consonance/dissonance notion, we only require 1; otherwise we back to $\{z + K\mathbf{x}, z + D\mathbf{x}\}$. Thus, for a general strong dichotomy, we locally characterize contrapuntal dissonances and consonances with 1.

Now, by alternation, ξ (respectively η) is in a half of the partition

$$\{g(K[\mathbf{x}]) \cap I_z, g(D[\mathbf{x}]) \cap I_z\}, \tag{2}$$

where z is the cantus firmus of ξ (respectively η). Therefore, we require the condition 1 on the partition in Equation (2). In Theorem 4.9, we prove that this particular condition is equivalent to the equation

$$p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = g(D[\mathbf{x}]) \cap I_z. \tag{3}$$

Mazzola [12] requires it only for the cantus firmus[12] z of ξ.

4.3. Variety

In this model, the variety principle of counterpoint [8, p. 21] corresponds to the condition that there is a maximum of alternations from ξ, that is, the cardinality of $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is maximum among all $g \in \mathbf{Sym}(R[\mathbf{x}])$ such that

1. $\xi \in g(D[\mathbf{x}]) \cap K[\mathbf{x}]$ (alternation) and

2. Equation (3) holds for the cantus firmus z of ξ (local dissonance).

---

[12]In Section 11, we discuss other possibilities, like also requiring 1 for the cantus firmus of η, which need not coincide with z. On the other hand, this condition is *weaker* than Mazzola's second requirement [11, Definition 95] $p^z[\mathbf{x}](g(K[\mathbf{x}])) = g(D[\mathbf{x}])$, and they are equivalent if R is commutative, as shown in Section 10.

4.4. Defining first-species counterpoint

Now we can give the definition of admitted successor.

Definition 4.1. Let (K, D, p) be a *strong dichotomy* of a ring R.

- A **contrapuntal symmetry** for a consonance $\xi \in K[\mathbf{x}]$, with $\xi = z + k\,\mathbf{x}$, is a symmetry g of $R[\mathbf{x}]$ such that

  1. $\xi \in g(D[\mathbf{x}])$,
  2. $p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = g(D[\mathbf{x}]) \cap I_z$, and
  3. the cardinality of $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is maximum among all g satisfying 1 and 2.

  Note that the contrapuntal symmetry for a given consonance is not required to be unique.
- An **admitted successor** of a consonance $\xi \in K[\mathbf{x}]$ is an element η of $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$ for some contrapuntal symmetry g. See Figure 2.
- Progressions (ξ, η) usually occur in a subset X of R, which we call **scale**. If η is an admitted successor of ξ we say that (ξ, η) is **allowed**. If it does not happen and (ξ, η) is polarized, we say that it is **forbidden**.

♣

Since the model chooses suitable pairs (ξ, η) as allowed among all polarized pairs, we say that *the model does not decide on non-polarized pairs*. Next we characterize polarized pairs and observe that, in the case of dual numbers $(\alpha = 0)$, they are all non-constant progressions. In the case $\alpha = 1$, they could be more particular.

4.5. Polarized progressions characterization

Let $\xi = z + k\,\mathbf{x}$ and $\eta = z' + k'\,\mathbf{x}$, where $\xi, \eta \in K[\mathbf{x}]$ and (K, D, p) is a strong dichotomy of a *finite* ring R. We want to determine when there is $g \in R[\mathbf{x}]$ such that $z + k\,\mathbf{x} \in g(D[\mathbf{x}])$ and $z' + k'\,\mathbf{x} \in g(K[\mathbf{x}])$. This property is equivalent to $k\,\mathbf{x} \in e^{-z}g(D[\mathbf{x}])$ and $z' - z + k'\,\mathbf{x} \in e^{-z}g(K[\mathbf{x}])$, so we reduce the problem to the case when $\xi = k\,\mathbf{x}$ and $\eta = y + k'\,\mathbf{x}$. Now, the following properties are equivalent.

1. There is $g \in \mathbf{Sym}(R[\mathbf{x}])$ such that $k\,\mathbf{x} \in g(D[\mathbf{x}])$ and $y + k'\,\mathbf{x} \in g(K[\mathbf{x}])$.

2. There is $g \in \mathbf{Sym}(R[\mathbf{x}])$ such that $g(k\,\mathbf{x}) \in D[\mathbf{x}]$ and $g(y + k'\,\mathbf{x}) \in K[\mathbf{x}]$.

3. There is $e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x}) \in \mathbf{Sym}(R[\mathbf{x}])$ such that $(c + d\alpha)k + v \in D$ and $(c + d\alpha)k' + v + dy \in K$.

4. There are $e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x}) \in \mathbf{Sym}(R[\mathbf{x}])$ and $\kappa \in K$ such that $v = \kappa - (c + d\alpha)k' - dy$ and $\kappa + (c + d\alpha)(k - k') - dy \in D$.

5. There is $c + d\,\mathbf{x}$ invertible such that $e^{(c+d\alpha)(k-k')-dy}(K) \nsubseteq K$.

6. There is $c + d\,\mathbf{x}$ invertible such that $e^{(c+d\alpha)(k-k')-dy}(K) \neq K$.

7. There is $c + d\,\mathbf{x}$ invertible such that $(c + d\alpha)(k - k') - dy \neq 0$.

The equivalence between 5 and 6 is given by the finiteness of R. Certainly, given two finite[13] sets A and B with the same number of elements, $A \subseteq B$ if and only if $A = B$ (that is, $A \nsubseteq B$ if and only if $A \neq B$). The equivalence between 6 and 7 holds since K is rigid by Corollary 3.3.

Note that in turn 7 implies $k\,\mathbf{x} \neq y + k'\,\mathbf{x}$. In the dual numbers case $(\alpha = 0)$, we can prove that it is also a sufficient condition.

---

[13] This fails for infinite sets: $\mathbb{N}$ and $\mathbb{Z}$ have the same cardinality but $\mathbb{N}$ is strictly contained in $\mathbb{Z}$.

Proposition 4.2. *Let $\xi, \eta \in K[\epsilon]$, where (K, D, p) is a strong dichotomy of a finite ring R. The pair (ξ, η) is polarized if and only if $\xi \neq \eta$.*

Proof. By 7 above, it only remains to prove that if $k\epsilon \neq y + k'\epsilon$, then there are $c \in R^*$ and $d \in R$ such that $c(k - k') - dy \neq 0$. We have two cases. *First case:* $k - k' \neq 0$. Take $d = 0$ and $c = 1$. *Second case:* $y \neq 0$ and $k - k' = 0$. Take $d = 1$ and any c. □

In the case when $\alpha = 1$, this result is not true. Let us suppose that $R = \mathbb{Z}_{12}$. If $k - k' \neq 0$, we polarize (ξ, η) as in the previous proof. However, if $y \neq 0$ and $k - k' = 0$, we have two possibilities. First, if $y \neq 6$, then $-2y \neq 0$ with $5 + 2\mathbf{x}$ invertible. Second, if $y = 6$, then $dy = 0$ for all $c + d\mathbf{x}$ invertible since d is even always. Thus, the polarized pairs (ξ, η) are all those such that $k \neq k'$, or $k = k'$ and $y \notin \{0, 6\}$. In other words, the non-polarized progressions are all repetitions and parallelisms with $y = 6$ (tritone skip).

4.6. The structural role of polarities

Next, we prove that there is a unique symmetry $p^z[\mathbf{x}] \in \mathbf{Sym}(R[\mathbf{x}])$ of the form $e^{u+v\,\mathbf{x}}c$ that interchanges contrapuntal consonances and dissonances in the fiber I_z. In particular, it leaves I_z unchanged. In the case when $z = 0$, $p^0[\mathbf{x}]$ is the extension $e^{a\,\mathbf{x}}b$ studied in Proposition 3.11.

Definition 4.3. Let g be a symmetry of $R[\mathbf{x}]$ and $z \in R$. We say that g is z-**invariant** if it leaves invariant the fiber I_z, that is, if $g(I_z) = I_z$. ♣

We first compute some candidates to $p^z[\mathbf{x}]$.

Proposition 4.4. *A symmetry g of $R[\mathbf{x}]$, where $g = e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x})$, is z-invariant if and only if $u = (1 - c)z$.*

Proof. Note that

$$g(z + R\,\mathbf{x}) = e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x})(z + R\,\mathbf{x})$$
$$= cz + u + ((c + d\alpha)R + dz + v)\,\mathbf{x} = cz + u + R\,\mathbf{x}.$$

Thus, g is z-invariant if and only $cz + u = z$, the latter condition being equivalent to $u = (1 - c)z$. □

We denote by H_z the *group*[14] of all z-invariant symmetries of $R[\mathbf{x}]$. In the case when $z = 0$, we simply write H instead of H_0. The conjugation automorphism $e^z \circ (-) \circ e^{-z}$ of $\mathbf{Sym}(R[\mathbf{x}])$ restricts to an isomorphism $H \longrightarrow H_z$ for each z, as established in the following proposition.

Proposition 4.5. *For each $z \in R$, the conjugation homomorphism $e^z \circ (-) \circ e^{-z} : H \longrightarrow H_z$ is an isomorphism of groups.*

Proof. First, let us prove that the conjugation homomorphism has its images in H_z. If $h \in H$, then

$$e^z \circ h \circ e^{-z}(z + R\,\mathbf{x}) = e^z \circ h(R\,\mathbf{x}) = e^z(R\,\mathbf{x}) = z + R\,\mathbf{x}.$$

Moreover, the inverse of $e^z \circ (-) \circ e^{-z}$ is $e^{-z} \circ (-) \circ e^z$. □

The *uniqueness* in the following theorem offers an explanation of the structural role of the strong dichotomy (K, D, p) with $p = e^a b$.

---

[14]In fact, this set is a group since it is an stabilizer one.

Theorem 4.6. *Let $(K, D, e^a b)$ be a strong dichotomy of R. For each $z \in R$, there is a unique z-invariant symmetry $p^z[\mathbf{x}]$ of the form $e^{u+v\,\mathbf{x}}c$, defined by $p^z[\mathbf{x}] = e^z \circ e^{a\,\mathbf{x}}b \circ e^{-z} = e^{(1-b)z+a\,\mathbf{x}}b$, such that*

$$p^z[\mathbf{x}](z + K\,\mathbf{x}) = z + D\,\mathbf{x}.$$

Proof. Since symmetries $p' \in H_z$ of the form $e^{u+v\,\mathbf{x}}c$ such that $p'(z + K\,\mathbf{x}) = z + D\,\mathbf{x}$ are in correspondence (conjugation) with symmetries in H, of the same form, sending $K\,\mathbf{x}$ to $D\,\mathbf{x}$, the proof of the theorem reduces to the case when $z = 0$ (Proposition 4.7). In that case, $p^0[\mathbf{x}]$ is $e^{a\,\mathbf{x}}b$, so $p^z[\mathbf{x}] = e^z \circ e^{a\,\mathbf{x}}b \circ e^{-z} = e^{(1-b)z+a\,\mathbf{x}}b$. □

Proposition 4.7. *Let $(K, D, e^a b)$ be a strong dichotomy of R. The symmetry $e^{a\,\mathbf{x}}b$ is the unique 0-invariant symmetry of the form $e^{u+v\,\mathbf{x}}c$ sending $K\,\mathbf{x}$ to $D\,\mathbf{x}$.*

Proof. First, by Proposition 4.4, necessarily $u = 0$. Now, $e^{v\,\mathbf{x}}c$ sends $K\,\mathbf{x}$ to $D\,\mathbf{x}$ if and only $e^v c(K) = D$. Hence, $e^v c = e^a b$ and $e^{u+v\,\mathbf{x}}c = e^{a\,\mathbf{x}}b$. □

4.7. The local condition on the deformed dichotomy

Lemma 4.8. *Let g be a symmetry of R[$\mathbf{x}$]. For each $z \in R$, there is $\gamma_z \in \mathbf{Sym}(R)$ such that*

$$g(X[\mathbf{x}]) \cap I_z = z + \gamma_z(X)\,\mathbf{x}$$

for any subset X of R. Concretely, if $g = e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x})$, we define

$$\gamma_z = e^{v+dc^{-1}(z-u)}(c + d\alpha).$$

Proof. If $g = e^{u+v\,\mathbf{x}}(c + d\,\mathbf{x})$, then

$$g(X[\mathbf{x}]) \cap I_z = z + ((c + d\alpha)X + v + dc^{-1}(z - u))\,\mathbf{x} = z + \gamma_z(X)\,\mathbf{x}.$$

□

Theorem 4.9. *Let (K, D, p) be a strong dichotomy of R and $g \in \mathbf{Sym}(R[\mathbf{x}])$. The following conditions are equivalent.*

1. *The symmetry $p^z[\mathbf{x}]$ is the unique $p' \in H_z$ of the form $e^{u+v\,\mathbf{x}}c$ such that*

$$p'(g(K[\mathbf{x}]) \cap I_z) = g(D[\mathbf{x}]) \cap I_z. \tag{4}$$

2. *The equation $p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = g(D[\mathbf{x}]) \cap I_z$ holds.*

3. *The equation $p\gamma_z = \gamma_z p$ holds, where γ_z is as in Lemma 4.8.*

Proof. Equation (4), for $p' = e^{u+v\,\mathbf{x}}c \in H_z$, is equivalent to the equations $(e^v c)\gamma_z(K) = \gamma_z(D)$, $\gamma_z^{-1}(e^v c)\gamma_z(K) = D$, and $(e^v c)\gamma_z = \gamma_z p$, by Lemma 4.8 and the fact that p is a polarity. The last equation (and hence Equation (4)) has a unique solution for $e^v c$ (respectively p') and Equation (4), for $p' = p^z[\mathbf{x}]$, is equivalent to $p\gamma_z = \gamma_z p$, so 1 is equivalent to 3 and 2. □

5. SIMPLIFIED COMPUTATION OF ADMITTED SUCCESSORS

In this section we characterize the admitted successors of a contrapuntal consonance $z + k\,\mathbf{x}$ as the translations, by z, of the admitted successors of $k\,\mathbf{x}$ associated with contrapuntal symmetries in H. This offers an important simplification of the original computation.

First, we transfer the symmetries satisfying conditions 1, 2, and 3 in Definition 4.1, and admitted successors, between the consonances $z + k\,\mathbf{x}$ and $k\,\mathbf{x}$. The basic tool is the **translation permutation** $e^z \circ (-) : \mathbf{Sym}(R[\mathbf{x}]) \longrightarrow \mathbf{Sym}(R[\mathbf{x}])$, whose inverse is $e^{-z} \circ (-)$.

5.1. Transfer of the first condition

Let $z + k\,\mathbf{x}$ be a contrapuntal consonance. We define $S^1_{z+k\mathbf{x}}$ as the set of all $g \in \mathbf{Sym}(R[\mathbf{x}])$ satisfying condition 1 in Definition 4.1, namely $z + k\,\mathbf{x} \in g(D[\mathbf{x}])$. Note that the translation $e^z \circ (-)$ restricts to a bijection $S^1_{k\mathbf{x}} \longrightarrow S^1_{z+k\mathbf{x}}$. This means that $k\,\mathbf{x} \in g(D[\mathbf{x}])$ if and only if $z + k\,\mathbf{x} \in e^z g(D[\mathbf{x}])$.

5.2. Transfer of the second condition

We define S^2_z as the set of all $g \in \mathbf{Sym}(R[\mathbf{x}])$ satisfying condition 2 in Definition 4.1, namely $p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = g(K[\mathbf{x}]) \cap I_z$. The translation $e^z \circ (-)$ restricts to a bijection $S^2_0 \longrightarrow S^2_z$, as proved by the following equivalences. Recall that $p^0[\mathbf{x}] = e^{-z} \circ p^z[\mathbf{x}] \circ e^z$ by Theorem 4.6.

$$p^0[\mathbf{x}](g(K[\mathbf{x}]) \cap I_0) = g(D[\mathbf{x}]) \cap I_0 \Leftrightarrow$$
$$e^{-z}p^z[\mathbf{x}]e^z(g(K[\mathbf{x}]) \cap I_0) = g(D[\mathbf{x}]) \cap I_0 \Leftrightarrow$$
$$p^z[\mathbf{x}]e^z(g(K[\mathbf{x}]) \cap I_0) = e^z(g(D[\mathbf{x}]) \cap I_0) \Leftrightarrow$$
$$p^z[\mathbf{x}](e^z g(K[\mathbf{x}]) \cap I_z) = e^z g(D[\mathbf{x}]) \cap I_z$$

In fact, e^z, as a bijection, commutes with intersections and $e^z(I_0) = I_z$.

5.3. Transfer of the third condition and admitted successors

Let $z + k\,\mathbf{x}$ be a contrapuntal consonance and $S_{z+k\mathbf{x}}$ the set of all $g \in \mathbf{Sym}(R[\mathbf{x}])$ satisfying 1 and 2 in Definition 4.1. Since the translation $e^z \circ (-)$ restricts to bijections between sets of symmetries satisfying 1 and 2, respectively (Sections 5.1 and 5.2), then it restricts to a bijection $S_{k\mathbf{x}} \longrightarrow S_{z+k\mathbf{x}}$.

On the other hand, note that

$$e^z(g(K[\mathbf{x}]) \cap K[\mathbf{x}]) = e^z g(K[\mathbf{x}]) \cap e^z(K[\mathbf{x}])$$
$$= e^z g(K[\mathbf{x}]) \cap K[\mathbf{x}].$$

The second equation holds because $X[\mathbf{x}]$ (for any X) is invariant under transformations of the form e^y. In particular, this means that $|g(K[\mathbf{x}]) \cap K[\mathbf{x}]| = |e^z g(K[\mathbf{x}]) \cap K[\mathbf{x}]|$ and hence, in that sense, the bijection $S_{k\mathbf{x}} \longrightarrow S_{z+k\mathbf{x}}$ preserves the cardinality of the sets of the form $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$.

Thus, $|g(K[\mathbf{x}]) \cap K[\mathbf{x}]|$ is maximum, among all g satisfying 1 and 2 in Definition 4.1, for the consonance $k\,\mathbf{x}$, if and only $|e^z g(K[\mathbf{x}]) \cap K[\mathbf{x}]|$ is for $z + k\,\mathbf{x}$. This means that $S_{k\mathbf{x}} \longrightarrow S_{z+k\mathbf{x}}$ restricts to a bijective correspondence between contrapuntal symmetries.

In particular, $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is a set of admitted successors of $k\,\mathbf{x}$ if and only if $e^z g(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is for $z + k\,\mathbf{x}$. This implies the following lemma.

Lemma 5.1. *The sets of admitted successors of $z + k\,\mathbf{x}$ can be computed as those of the form*

$$e^z(g(K[\mathbf{x}]) \cap K[\mathbf{x}]),$$

where g is a contrapuntal symmetry for $k\,\mathbf{x}$.

5.4. Restricting symmetries

Given a symmetry g, we next prove that $g(K[\mathbf{x}]) = h(K[\mathbf{x}])$, where $h \in H$. Note that this immediately implies that h is contrapuntal for $z + k\,\mathbf{x}$ whenever g is. This suggests drawing our attention to the contrapuntal symmetries in H.

Recall that H_z consists of all z-invariant symmetries of $R[\mathbf{x}]$, that is, the symmetries of the form $e^{(1-c)z+v\mathbf{x}}(c + d\,\mathbf{x})$ with $c, c + d\alpha \in R^*$ (Proposition 4.4).

Proposition 5.2. *Let g be a symmetry of $R[\mathbf{x}]$. For each $z \in R$, there is $g_z \in H_z$ such that*

$$g(X[\mathbf{x}]) = g_z(X[\mathbf{x}])$$

for any subset X of R. Concretely, if $g = e^{u+v\mathbf{x}}(c + d\mathbf{x})$, then g_z is ge^w, where $w = c^{-1}(z - u) - z$.

Proof. Note that the symmetry g_z defined above is in H_z, by Proposition 4.4, since $cw + u = (1 - c)z$. Thus,

$$g_z(X[\mathbf{x}]) = ge^w(X[\mathbf{x}]) = g(X[\mathbf{x}]).$$

$\square$

5.5. Main theorem

Our previous observations lead to a considerably simpler characterization of contrapuntal symmetries and admitted successors, only involving symmetries in H.

Proposition 5.3. *The admitted successors of a consonance $z + k\mathbf{x} \in K[\mathbf{x}]$ can be computed as the elements of the sets of the form*

$$e^z(h(K[\mathbf{x}]) \cap K[\mathbf{x}]), \tag{5}$$

where $h \in H$ and

    (a) $k\mathbf{x} \in h(D[\mathbf{x}])$,
    (b) $p^0[\mathbf{x}](h(K[\mathbf{x}]) \cap I_0) = h(D[\mathbf{x}]) \cap I_0$, and
    (c) the cardinality of $h(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is maximum among all $h \in H$ satisfying (a) and (b).

Proof. By Lemma 5.1, it is enough to show that the collection of all sets of the form $g(K[\mathbf{x}]) \cap K[\mathbf{x}]$, with g contrapuntal symmetry for $k\mathbf{x}$, is equal to the set of all intersections $h(K[\mathbf{x}]) \cap K[\mathbf{x}]$ with h satisfying (a), (b), and (c) above.

If g is a contrapuntal symmetry for $k\mathbf{x}$, then, since $g(K[\mathbf{x}]) = g_0(K[\mathbf{x}])$ (Proposition 5.2), g_0 is a contrapuntal symmetry and, in particular, satisfies (a) and (b). Also, $g(K[\mathbf{x}]) \cap K[\mathbf{x}] = g_0(K[\mathbf{x}]) \cap K[\mathbf{x}]$ and g_0 satisfies (c) because $H \subseteq \mathbf{Sym}(R[\mathbf{x}])$.

Conversely, if h satisfies (a), (b), and (c), we claim that h is a contrapuntal symmetry and hence $h(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is an usual admitted successors set. To prove the claim, take g satisfying 1 and 2 in Definition 4.1 for the consonance $k\mathbf{x}$. The symmetry g satisfies (a), (b), and $g(K[\mathbf{x}]) = g_0(K[\mathbf{x}])$, so

$$|g(K[\mathbf{x}]) \cap K[\mathbf{x}]| = |g_0(K[\mathbf{x}]) \cap K[\mathbf{x}]| \leq |h(K[\mathbf{x}]) \cap K[\mathbf{x}]|,$$

and h is contrapuntal. $\square$

If we replace condition (b) in Proposition 5.3 by its equivalent 3 in Theorem 4.9, we immediately obtain the main theorem.

Theorem 5.4. *The admitted successors of a consonance $z + k\mathbf{x} \in K[\mathbf{x}]$ can be computed as the elements of the sets of the form*

$$e^z(h(K[\mathbf{x}]) \cap K[\mathbf{x}]), \tag{6}$$

where $h = e^{v\mathbf{x}}(c + d\mathbf{x}) \in H$ and

    1. $k\mathbf{x} \in h(D[\mathbf{x}])$,

    2. $p \circ e^v(c + d\alpha) = e^v(c + d\alpha) \circ p$, and

    3. the cardinality of $h(K[\mathbf{x}]) \cap K[\mathbf{x}]$ is maximum among all $h \in H$ satisfying 1 and 2.

This theorem says that, to compute the admitted successors of a consonance $z + k\,\mathbf{x}$, it is enough to do so for $k\,\mathbf{x}$, and then apply the transposition e^z (Equation (6)). In this simplification we only use symmetries in H and reduce the condition 2 to a suitable commutativity. This situation agrees with the fact that in Renaissance and the common-practice period contrapuntal consonances are invariant with respect to musical transposition.

6. Counting formulas and maximization

Let R be a *finite* ring, (K, D, p) a strong dichotomy of R with $p = e^a b$, and $\{K[\mathbf{x}], D[\mathbf{x}]\}$ the induced dichotomy of $R[\mathbf{x}]$; see Proposition 3.11.

According to Theorem 5.4, it is desirable to count the number of elements of the sets of the form $h(K[\mathbf{x}]) \cap K[\mathbf{x}]$, for $h \in H$, so as to choose suitable maximum cardinals.

If $e^v\mathbf{x}(c + d\,\mathbf{x})$, with $c, c + d\alpha \in R^*$, belongs to H and $r + k\,\mathbf{x} \in K[\mathbf{x}]$, then

$$e^v\mathbf{x}(c + d\,\mathbf{x})(r + k\,\mathbf{x}) = cr + ((c + d\alpha)k + v + dr)\,\mathbf{x}.$$

Let $\sqcup$ denote the disjoint union of sets. Thus,

$$e^v\mathbf{x}(c + d\,\mathbf{x})(K[\mathbf{x}]) = \bigsqcup_{r \in R} cr + ((c + d\alpha)K + v + dr)\,\mathbf{x}$$

and

$$e^v\mathbf{x}(c + d\,\mathbf{x})(K[\mathbf{x}]) \cap K[\mathbf{x}] = \bigsqcup_{r \in R} cr + (((c + d\alpha)K + v + dr) \cap K)\,\mathbf{x} \tag{7}$$

because $e^0 c$ is a permutation of R, and hence

$$|e^v\mathbf{x}(c + d\,\mathbf{x})(K[\mathbf{x}]) \cap K[\mathbf{x}]| = \sum_{r \in R} |((c + d\alpha)K + v + dr) \cap K|. \tag{8}$$

Moreover, we can assume that $\alpha = 0$ in the last equation in view of the following ones, where $\chi_K : R \longrightarrow \{0, 1\}$ is the characteristic function[15] of K.

$$\begin{aligned}
\sum_{r \in R} |((c + d\alpha)K + v + dr) \cap K| &= \sum_{r \in R} \sum_{k \in K} \chi_K((c + d\alpha)k + v + dr) \\
&= \sum_{r \in R} \sum_{k \in K} \chi_K(ck + v + d(\alpha k + r)) \\
&= \sum_{k \in K} \sum_{r \in R} \chi_K(ck + v + d(\alpha k + r)) \\
&= \sum_{k \in K} \sum_{r \in R} \chi_K(ck + v + dr) \\
&= \sum_{r \in R} |(cK + v + dr) \cap K|
\end{aligned}$$

Given an element d in a ring R, we have the R-endomorphism $d \cdot - : R \longrightarrow R : r \mapsto dr$, which is useful in the following discussion.

Lemma 6.1. *Let K and K' be subsets of a finite ring R and $d \in R$. The equation*

$$\sum_{r \in R} |(K' + dr) \cap K| = \rho \sum_{\gamma \in \mathrm{Coker}(d \cdot -)} |K'_\gamma||K_\gamma| \tag{9}$$

---

[15]It sends r to 1 if $r \in K$ and sends r to 0 otherwise.

holds, where $\mathrm{Coker}(d \cdot -) = R/dR,$

$$X_\gamma = \{x \in X \mid [x] = \gamma\}$$

for each $X \subseteq R$, *and* $\rho = |\mathrm{Ker}(d \cdot -)|$. *Here,* $[x]$ *denotes the class of* x *in* $\mathrm{Coker}(d \cdot -)$.

Proof.

$$
\begin{aligned}
\sum_{r \in R} |(K' + dr) \cap K| &= \rho \sum_{r \in dR} |(K' + r) \cap K| \\
&= \rho \sum_{r \in dR} \sum_{s \in K'} \chi_K(s + r) \\
&= \rho \sum_{s \in K'} \sum_{r \in dR} \chi_K(s + r) \\
&= \rho \sum_{s \in K'} |K_{[s]}| \\
&= \rho \sum_{\gamma \in K'/dR} |K'_\gamma||K_\gamma| = \rho \sum_{\gamma \in \mathrm{Coker}(d \cdot -)} |K'_\gamma||K_\gamma|.
\end{aligned}
$$

As to the last equality, note that if γ is not in the image K'/dR of K' under the canonical projection onto the cokernel, then $|K'_\gamma| = 0$.   □

Let us remark a curiosity. In Lemma 6.1, the cases $d = 1$ and $d = 0$ correspond to the formulas

$$\sum_{r \in R} |(K' + r) \cap K| = |K'||K|$$

and

$$\sum_{r \in R} |K' \cap K| = |R||K' \cap K|$$

respectively; see [11, Lemma 48]. We can exemplify the first formula by saying that the total number of common notes between the C major triad $\{C, E, G\}$ and all its transpositions in $\mathbb{Z}_{12}$ is 3×3: three common notes with itself and a common note with E, G, $A\flat$, $E\flat$, F, and A major chords.

The following corollary illustrates the lemma in the important case when the ring is $\mathbb{Z}_n$.

Corollary 6.2. *Let K and K' be subsets of $\mathbb{Z}_n$ and $d \in \mathbb{Z}_n$. The equation*

$$\sum_{r=0}^{n-1} |(K' + dr) \cap K| = \rho \sum_{i=0}^{\rho-1} |K'_i||K_i| \tag{10}$$

holds, where

$$X_i = \{x \in X \mid x \equiv i \pmod{\rho}\}$$

for each $X \subseteq \mathbb{Z}_{12}$ and $\rho = \gcd(d, n)$.

Proof. Note that for each finite ring R (isomorphism and index theorems for groups)

$$|\mathrm{Ker}(d \cdot -)| = |R|/|dR| = |R/dR| = |\mathrm{Coker}(d \cdot -)|.$$

If $R = \mathbb{Z}_n$, then by [7, Theorem 6.14], $R/dR = \mathbb{Z}_n/d\mathbb{Z}_n = \mathbb{Z}_n/\rho\mathbb{Z}_n \cong \mathbb{Z}_\rho$, where $\rho = \gcd(d, n)$, so $|\mathrm{Ker}(d \cdot -)| = |\mathrm{Coker}(d \cdot -)| = \rho$. Moreover, $[x] = \gamma$ for $\gamma \in \mathrm{Coker}(d \cdot -)$ if and only if $x \equiv i \pmod{\rho}$ by identifying γ and i through the isomorphism $\mathrm{Coker}(d \cdot -) \cong \mathbb{Z}_\rho$.   □

d	ρ	$\sum\limits_{r=0}^{11} |(K'+dr)\cap K|$
0	12	$12(|K'_0|+|K'_3|+|K'_4|+|K'_7|+|K'_8|+|K'_9|)$
$1,5,7,11$	1	36
$2,10$	2	36
$3,9$	3	$3(|K'_0|3+|K'_1|2+|K'_2|)$
$4,8$	4	$4(|K'_0|3+|K'_1|+|K'_3|2)$
6	6	$6(|K'_0|+|K'_1|+|K'_2|+|K'_3|2+|K'_4|)$

Table 1: *Values of the right-hand side of Equation* (8) *for the Renaissance dichotomy.*

Example 6.3. Let (K,D,e^25) be the Renaissance strong dichotomy of $\mathbb{Z}_{12}$, with $K=\{0,3,4,7,8,9\}$ and $D=\{1,2,5,6,10,11\}$. Table 1 contains the particular values of the right-hand side of Equation (8), by using Equation (10) with $K'=cK+v$, for d ranging over $\mathbb{Z}_{12}$. Next, we briefly justify each row.

Case $d=0$. The number of elements in K congruent to i modulo 12 is just given by the characteristic function of K as a subset of $\mathbb{Z}_{12}$.

Case $d=1,5,7,11$. Since all integers are congruent modulo 1, we obtain the expression $|K'||K|$ for the sum, whose exact value is 6×6.

The *case $d=2,10$* is quite interesting. The integers congruent to 0 (respectively 1) modulo 2 are the even (respectively odd) ones. Now, there are three even numbers (0, 4, and 8) and three odd numbers (3, 7, and 9) in K. Thus, the counting formula becomes $2(|K'_0|3+|K'_1|3)$. But multiplying K by c (which is always odd because it is invertible) does not alter the parity of its elements, and adding v to cK does not alter the number of odd or even elements. For this reason, $|K'_0|=|K'_1|=3$ and the exact value of the counting formula is $2((3\times 3)+(3\times 3))$.

Case $d=3,9$. In this case, $|K_0|=|\{0,3,9\}|=3$, $|K_1|=|\{4,7\}|=2$, $|K_2|=|\{8\}|=1$, and the remaining terms of the form K_i are empty.

Case $d=4,8$. Here $|K_0|=|\{0,4,8\}|=3$, $|K_1|=|\{9\}|=1$, $|K_3|=|\{3,7\}|=2$, and the remaining terms of the form K_i are empty.

Case $d=6$. Here $|K_0|=|\{0\}|=1$, $|K_1|=|\{7\}|=1$, $|K_2|=|\{8\}|=1$, $|K_3|=|\{3,9\}|=2$, $|K_4|=|\{4\}|=1$, and the remaining terms of the form K_i are empty. ◇

Another important consequence of the equation

$$\sum_{r\in R}|(K'+dr)\cap K| = \rho_d \sum_{r\in dR}|(K'+r)\cap K|$$

in the proof of Lemma 6.1 is that if $dR=d'R$, then

$$\rho_d = |\operatorname{Ker}(d\cdot -)| = |R|/|dR| = |R|/|d'R| = \rho_{d'}$$

and hence

$$\sum_{r\in R}|(K'+dr)\cap K| = \sum_{r\in R}|(K'+d'r)\cap K|.$$

Moreover, in the case when $R=\mathbb{Z}_n$, we observe from Equation (10) that we can reduce the computation of the sums of the form $\sum\limits_{r\in R}|(K'+dr)\cap K|$ for all d with the same ρ (recall that $\rho=\gcd(d,n)$) to that of the sum $\sum\limits_{r\in R}|(K'+\rho r)\cap K|$, since these sums coincide.

6.1. Main counting formulas

Now, according to Equation (8) (with $\alpha = 0$), we will focus on the case when $K' = cK + v$ with c invertible in R. Let us assume that d is in the *center*[16] $Z(R)$ of R. We claim that the sequence $(|K'_\gamma|)_\gamma$ is a *rearrangement*[17] of $(|K_\gamma|)_\gamma$, where γ ranges over $\mathrm{Coker}(d \cdot -)$. In fact,

$$|K'_{[r]}| = |\{e^v c(k) \mid k \in K \text{ and } [e^v c(k)] = [r] \text{ in } \mathrm{Coker}(d \cdot -)\}| \tag{11}$$

$$= |\{e^v c(k) \mid k \in K, ck + v - r \in dR\}| \tag{12}$$

$$= |\{e^v c(k) \mid k \in K, k - (c^{-1}r - c^{-1}v) \in dR\}| \tag{13}$$

$$= |\{e^v c(k) \mid k \in K, k - (e^v c)^{-1}(r) \in dR\}| \tag{14}$$

$$= |\{e^v c(k) \mid k \in K, [k] = [(e^v c)^{-1}(r)] \text{ in } \mathrm{Coker}(d \cdot -)\}| \tag{15}$$

$$= |\{k \in K \mid [k] = [(e^v c)^{-1}(r)]\}| \tag{16}$$

$$= |K_{[(e^v c)^{-1}(r)]}|.$$

Equation (11) holds by the definitions in Lemma 6.1. Equation (12) corresponds to the definition of equality in the quotient $\mathrm{Coker}(d \cdot -)$. Equation 13 follows from the fact that $z \in dR$ implies $c^{-1}z \in c^{-1}dR = dc^{-1}R = dR$ whenever $d \in Z(R)$. Equation (14) holds since the inverse of $e^v c$ is $e^{-c^{-1}v}c^{-1}$ according to Section 2. Equation (15) is the definition of equality in the cokernel again. The equality of cardinalities in Equation (16) holds because $e^v c$ is a bijection.

Second, since $d \in Z(R)$, the function (actually a ring symmetry according to [14, p. 278] because $\mathrm{Coker}(d \cdot -)$ is a two-sided ideal of R)

$$[e^v c] : \quad \mathrm{Coker}(d \cdot -) \quad \longrightarrow \quad \mathrm{Coker}(d \cdot -)$$
$$[r] \quad \longmapsto \quad [e^v c(r)]$$

is well defined (check) with inverse $[(e^v c)^{-1}]$, so $[e^v c]$ and $[(e^v c)^{-1}]$ are permutations of the cokernel. Thus, regarding $(|K'_\gamma|)_\gamma$ and $(|K_\gamma|)_\gamma$ as functions from the cokernel to $\mathbb{N}$, the former is the composite of the latter with $[(e^v c)^{-1}]$ by the equality $|K'_{[r]}| = |K_{[(e^v c)^{-1}(r)]}|$ above. Also, note that $[(e^v c)^{-1}] = [e^{-c^{-1}v}c^{-1}] = e^{-[c]^{-1}[v]}[c]^{-1}$. This proves the main counting formulas.

Theorem 6.4 (Main counting formulas). *Let R be a finite ring, $d \in Z(R)$, $K \subseteq R$, and $e^{v\mathbf{x}}(c + d\mathbf{x})$ a symmetry in H. The equation*

$$|e^{v\mathbf{x}}(c + d\mathbf{x})(K[\mathbf{x}]) \cap K[\mathbf{x}]| = \rho \sum_{\gamma \in \mathrm{Coker}(d\cdot-)} |K_{[(e^v c)^{-1}](\gamma)}||K_\gamma| = \rho \sum_{\delta \in \mathrm{Coker}(d\cdot-)} |K_\delta||K_{[e^v c](\delta)}| \tag{17}$$

holds, where $K_\gamma = \{x \in K \mid [x] = \gamma\}$ and $\rho = |\mathrm{Ker}(d \cdot -)|$. In particular, if $R = \mathbb{Z}_n$, the right-hand term of Equation (17) coincides with

$$\rho \sum_{i=0}^{\rho-1} |K_i||K_{e^v c(i)}|, \tag{18}$$

where $K_i = \{x \in K \mid x \equiv i \pmod{\rho}\}$, $\rho = \gcd(d, n)$, and $e^v c$ is reduced modulo ρ.

Proof. Combine the previous discussion with Equation (8), Lemma 6.1, and Corollary 6.2. The second equality in Equation (17) follows from the first one and the change of variable $\delta = [(e^v c)^{-1}](\gamma)$. $\square$

---

[16]It consists of those r in R such that $rs = sr$ for all $s \in R$.

[17]A rearrangement of a function (sequence) $f : S \longrightarrow X$ is a function of the form $f \circ \sigma$, where σ is a permutation of S.

6.2. Maximization criterion

The following maximization criterion helps to find symmetries h such that $|h(K[\mathbf{x}]) \cap K[\mathbf{x}]|$ is maximum among all symmetries with $h = e^{v\mathbf{x}}(c + d\,\mathbf{x})$ and $d \in Z(R)$ fixed. It is important to emphasize that it need not find symmetries with maximum values among all symmetries satisfying conditions 1 and 2 of contrapuntal symmetry, but is very useful to discard a number of symmetries whose values are not maximum.

Theorem 6.5 (Maximization criterion). *Assume the hypotheses of Theorem 6.4. The right-hand side of Equation* (17)

$$\rho \sum_{\gamma \in \mathrm{Coker}(d \cdot -)} |K_\gamma||K_{[e^v c](\gamma)}|$$

is maximum, for $e^v c$ ranging over all symmetries of R, if and only if $|K_{[e^v c](\gamma)}| = |K_\gamma|$ for each $\gamma \in$ $\mathrm{Coker}(d \cdot -)$. *In particular, the sum is maximum if $e^v c \equiv e^0 1 \pmod{d}$, that is,[18] if $v \equiv 0 \pmod{d}$ and* $c \equiv 1 \pmod{d}$. *Further, the maximum value is*

$$\rho \sum_{\gamma \in \mathrm{Coker}(d \cdot -)} |K_\gamma|^2.$$

Proof. Apply the rearrangement inequality (Theorem 13.1) to the sequences

$$(|K_\gamma|)_\gamma \text{ and } (|K_{[e^v c](\gamma)}|)_\gamma,$$

the latter being a rearrangement of the former by Section 6.1. □

Certainly, the criterion is useful because if a symmetry $e^{v\mathbf{x}}(c + d\,\mathbf{x})$ satisfies the conditions 1 and 2 for a contrapuntal symmetry and $e^v c \equiv e^0 1 \pmod{d}$, then we get rid of all symmetries with the same d that induce a rearrangement of $(|K_\gamma|)_\gamma$, which are usually all but those satisfying $e^v c \equiv e^0 1 \pmod{d}$.

7. The little theorem of first-species counterpoint

The task of finding maximum cardinals subject to the conditions 1 and 2 seems to be difficult in general. In the *Little theorem of counterpoint*, which we prove for the dual numbers case, we establish that each consonance has at least $|K|^2$ admitted successors. For example, in Renaissance counterpoint, each consonance has at least 36 admitted successors. However, this approximation theorem does not provide the admitted successors explicitly, which requires a greater effort.

The proof idea of the little theorem of counterpoint is the following. If we prove that there is at least an $h \in H$ satisfying the conditions 1 and 2 in Theorem 5.4 for all consonances $k\,\mathbf{x} \in K[\mathbf{x}]$, then we deduce that the number of admissible successors is at least $|h(K[\mathbf{x}]) \cap K[\mathbf{x}]|$ by the maximum property of admitted successors sets associated with contrapuntal symmetries. Finally, Theorem 6.4 gives the exact value of $|h(K[\mathbf{x}]) \cap K[\mathbf{x}]|$.

So as to prove the existence of such an h, we first need to give concrete criteria for deciding when an $h \in H$ satisfies 1 or 2. We do this in Sections 7.1 and 7.2. Finally, we translate our result to any consonance $z + k\,\mathbf{x}$, by using the bijective translation e^z.

---

[18]**Definition.** We say that $x \equiv y \pmod{d}$ if $[x] = [y]$ in R/dR.

7.1. The first condition criterion

Regarding the condition 1 in Theorem 5.4, given $h = e^{vx}(c + dx)$, we have the following equivalences, where p is the polarity $e^a b$ of $(K, D, e^a b)$.

$$
\begin{aligned}
kx \in h(D[x]) &\Leftrightarrow k \in (c + d\alpha)D + v \\
&\Leftrightarrow k \in (c + d\alpha) \cdot p(K) + v \\
&\Leftrightarrow v = k - (c + d\alpha) \cdot p(s) \text{ for some } s \in K
\end{aligned}
$$

7.2. The second condition criterion

As to 2 in Theorem 5.4, given $h = e^{vx}(c + dx)$, we have the following equivalences.

$$
\begin{aligned}
p \circ e^v(c + d\alpha) = e^v(c + d\alpha) \circ p &\Leftrightarrow e^a b \circ e^v(c + d\alpha) = e^v(c + d\alpha) \circ e^a b \\
&\Leftrightarrow e^{bv+a}b(c + d\alpha) = e^{(c+d\alpha)a+v}(c + d\alpha)b \\
&\Leftrightarrow bv + a = (c + d\alpha)a + v \text{ and } b(c + d\alpha) = (c + d\alpha)b
\end{aligned}
$$

7.3. The little theorem

In the dual numbers case ($\alpha = 0$), we can prove the existence of an $h \in H$ with $d = 1$ that satisfies 1 and 2 in Theorem 5.4, for the consonance $k\epsilon$. In fact, according to Sections 7.1 and 7.2, it is enough to show that the following system of equations, in the unknowns $v \in R$, $c \in R^*$, $d \in R$ and $s \in K$, has at least a solution.

$$
\begin{cases}
v = k - cbs - ca \\
bv + a = ca + v \\
bc = cb
\end{cases}
$$

Certainly, it has the solution $v = -ba = a$, $c = b$, $d = 1$, $s = k$ (note that $b^2 = 1$ and $ba + a = 0$ since $e^a b$ is involutive by Proposition 3.6). To conclude, $e^{-ba\epsilon}(b + \epsilon)$ is the desired h.

Lemma 7.1. *Let R be a finite ring and (K, D, p) a strong dichotomy. Each consonance $k\epsilon$ has at least $|K|^2$ admitted successors.*

Proof. If $h \in H$ satisfies $d = 1$, then, according to Lemma 6.1,

$$
|h(K[\epsilon]) \cap K[\epsilon]| = \sum_{r \in R} |(cK + v + r) \cap K| = |cK + v||K| = |K|^2.
$$

Hence, h, with $h = e^{-ba\epsilon}(b + \epsilon)$, satisfies 1, 2, and the previous equation. Now, let N be the number of admitted successors of $k\epsilon$. If h' satisfies the conditions of Theorem 5.4, for $k\epsilon$, then

$$
N \geq |h'(K[\epsilon]) \cap K[\epsilon]| \geq |h(K[\epsilon]) \cap K[\epsilon]| = |K|^2.
$$

$\square$

An upper bound for a single contrapuntal symmetry can be obtained as well. First, suppose that d is not 0. Note that

$$
\sum_{r \in R} |(cK + v + dr) \cap K| = \rho \sum_{r \in dR} |(cK + v + r) \cap K|,
$$

where $d \cdot - : R \longrightarrow R$ is the R-endomorphism that sends an element $r \in R$ to dr, and $\rho = |\operatorname{Ker}(d \cdot -)|$.

In the case when $e^{v+r}c$ is not the identity $e^0 1$, note that $|(cK + v + r) \cap K|$ is at most $|K| - 1$ since K is rigid. On the other hand, there is at most a value of r in dR that makes $e^{v+r}c$ the identity, and hence there is at most an r such that $|(cK + v + r) \cap K| = |K|$. Thus,

$$\rho \sum_{r \in dR} |(cK + v + r) \cap K| \leq \rho[(|dR| - 1)(|K| - 1) + |K|]$$

$$= \rho \left[\left(\frac{|R|}{\rho} - 1 \right) (|K| - 1) + |K| \right]$$

$$= \rho \left[\left(\frac{2|K|}{\rho} - 1 \right) (|K| - 1) + |K| \right]$$

$$= 2|K|^2 - 2|K| + \rho$$

$$\leq 2|K|^2 - 2|K| + |K| = 2|K|^2 - |K|.$$

As to the last inequality note that ρ, as a divisor of $|R|$ that is not[19] $|R|$, must be less than or equal to $|K|$ (which coincides with $|R|/2$), the greatest divisor of $|R|$ different from $|R|$.

Now suppose that $d = 0$. So as to find an upper bound for

$$\sum_{r \in R} |(cK + v) \cap K|$$

we consider two cases. If $c \neq 1$, then $e^v c$ is not the identity $e^0 1$. If $c = 1$, then, since we require h (for $h = e^{v\epsilon}c$) to satisfy condition 1 in Theorem 5.4, by Section 7.1, $v = k - p(s)$ for some $s \in K$. This means that $v \neq 0$ since $p(s) \in D$, and hence $e^v c$ is not the identity. In both cases, $|(cK + v) \cap K| \leq |K| - 1$ since K is rigid. Thus,

$$\sum_{r \in R} |(cK + v) \cap K| \leq |R|(|K| - 1) = 2|K|(|K| - 1) < 2|K|^2 - |K|.$$

To sum up, by collecting the results for $d \neq 0$ and $d = 0$ and using Theorem 5.4, we obtain the upper bound in the following theorem.

Theorem 7.2 (Little theorem of counterpoint). *Let R be a finite ring and (K, D, p) a strong dichotomy. Each consonance $x + k\epsilon$ has at least $|K|^2$ admitted successors, and at most $2|K|^2 - |K|$ for a single contrapuntal symmetry.*

Proof. By Theorem 5.4, e^z is a bijection between the sets of admitted successors of $k\epsilon$ and $z + k\epsilon$. Thus, the number of admitted successors of $z + k\epsilon$ is equal to the number of admitted successors of $k\epsilon$. The lower bound now follows from Lemma 7.1. The upper bound corresponds to the preceding discussion. □

7.4. Steps for the computation of contrapuntal symmetries

The steps for calculating the contrapuntal symmetries $e^{v\mathbf{x}}(c + d\,\mathbf{x})$ in H, for a consonance of the form $k\,\mathbf{x}$, are the following. We start with a strong dichotomy $(K, D, e^a b)$.

1. Solve the following system of equations in the unknowns $c \in R^*$, v, and d.

$$\begin{cases} bv + a = (c + d\alpha)a + v \\ b(c + d\alpha) = (c + d\alpha)b \end{cases}$$

This corresponds to condition 2 of contrapuntal symmetry (Section 7.2).

---

[19] In fact ρ, which is by definition $|\mathrm{Ker}(d \cdot -)|$, is equal to $|R|$ if and only if $d = 0$; but we are assuming $d \neq 0$.

2. For each consonance $k\mathbf{x}$, among the symmetries of the form $e^{v\mathbf{x}}(c + d\mathbf{x})$ obtained in 1, choose those with $v \in k - cD$. This corresponds to condition 1 of contrapuntal symmetry (Section 7.1). The reason for first computing the symmetries satisfying condition 2 is that those symmetries are usually less than those satisfying condition 1, so we perform less operations.

3. For each consonance $k\mathbf{x}$, among the symmetries $e^{v\mathbf{x}}(c + d\mathbf{x})$ obtained in 2, choose those such that the right-hand term of Equation (9)

$$\rho \sum_{\gamma \in \mathrm{Coker}(d \cdot -)} |(cK + v)_\gamma| ||K_\gamma|$$

with $K' = cK + v$ is maximum. This ensures that the condition 3 holds by Equation (8). In the case when R is commutative, maximize the right-hand side of Equation (17). In the case when $R = \mathbb{Z}_n$, maximize Equation (18). So as to discard a number of symmetries, if R is commutative, Theorem 6.5 can be used.

In the next section we apply the previous steps to computing the contrapuntal symmetries in the case of first-species Renaissance counterpoint.

8. FIRST-SPECIES RENAISSANCE COUNTERPOINT

Let $(K, D, e^2 5)$ be the Renaissance strong dichotomy of $\mathbb{Z}_{12}$, with $K = \{0, 3, 4, 7, 8, 9\}$ and $D = \{1, 2, 5, 6, 10, 11\}$. Next, we compute the respective contrapuntal symmetries and admitted successors, in the dual numbers case ($\alpha = 0$). In Section 11.1, we explain a simple procedure for obtaining the contrapuntal symmetries in the case $\alpha = 1$.

8.1. Condition 2

By the commutativity of $\mathbb{Z}_{12}$, the condition 2 for symmetries $e^{v\epsilon}(c + d\epsilon) \in H$, $c \in R^*$, reduces to solve the equation

$$5v + 2 = c2 + v,$$

which is equivalent to

$$4v = 2(c - 1).$$

If $c = 1, 7$, then the equation becomes $4v = 0$, so $v = 0, 3, 6, 9$. If $c = 5, 11$, then the equation becomes $4(v - 2) = 0$ and hence $v = 2, 5, 8, 11$. The solutions are summarized in the following table.

c	v
1, 7	0, 3, 6, 9
5, 11	2, 5, 8, 11

8.2. Condition 1

Now, among these solutions we choose, for each consonance $k\epsilon$, those satisfying

$$v \in k - cD.$$

In the following table, we organize the results of the operations involved. Specifically, we compute $3\mathbb{Z}_{12} \cap (k - cD)$ for $c = 1, 7$ and $(2 + 3\mathbb{Z}_{12}) \cap (k - cD)$ for $c = 5, 11$.

$\diagdown\quad k$ c	0	3	4	7	8	9
1	$\{6\}$	$\{9\}$	$\{3,6\}$	$\{6,9\}$	$\{3,6,9\}$	$\{3\}$
5	$\{2,5,11\}$	$\{2,5,8\}$	$\{2,11\}$	$\{2,5\}$	$\{2\}$	$\{2,8,11\}$
7	$\{6\}$	$\{9\}$	$\{6,9\}$	$\{0,9\}$	$\{3,6,9\}$	$\{3\}$
11	$\{2,5,11\}$	$\{2,5,8\}$	$\{2,5\}$	$\{5,8\}$	$\{2\}$	$\{2,8,11\}$

We can easily fill in the table as follows. There are six entries, namely those labelled by (c,k) with $c = 1, 5$ and $k = 0, 4, 8$, which we can start from, the other being obtained by using certain symmetries. In fact, note that

$$7(3\mathbb{Z}_{12} \cap (k - D)) = 3\mathbb{Z}_{12} \cap (7k - 7D)),$$

$$7((2 + 3\mathbb{Z}_{12}) \cap (k - 5D)) = (2 + 3\mathbb{Z}_{12}) \cap (7k - 11D),$$

$$\pm 3 + (3\mathbb{Z}_{12} \cap (k - cD)) = 3\mathbb{Z}_{12} \cap (\pm 3 + k - cD),$$

and

$$\pm 3 + ((2 + 3\mathbb{Z}_{12}) \cap (k - cD)) = (2 + 3\mathbb{Z}_{12}) \cap (\pm 3 + k - cD).$$

For example, the entry $(1,3)$ is obtained by adding 3 to the entry $(1,0)$, and the entry $(7,0)$ is obtained by multiplying the entry $(1,0)$ by 7.

8.3. Maximization

In the case of the Renaissance dichotomy, Theorem 6.5 allows us to obtain the following table of maximum values, without the restrictions of conditions 1 and 2 of contrapuntal symmetry, and their corresponding symmetries. We do not include the value $d = 0$, whose maximum is not useful because the identity does not satisfy the condition 1. After the table we justify the results.

d	ρ	maximum $\sum\limits_{r=0}^{11} |(cK + v + dr) \cap K|$	$e^v c$
1, 5, 7, 11	1	36	any
2, 10	2	36	any
3, 9	3	42	$\equiv e^0 1 \pmod 3$
4, 8	4	56	$\equiv e^0 1 \pmod 4$
6	6	48	$\equiv e^0 1 \pmod 6$

If $\rho = 3$, since all entries of the vector $(|K_0|, |K_1|, |K_2|)$ are different (Example 6.3), then the unique rearrangement that coincides with it is the composition with $e^0 1$, and hence the maximum is only taken for the identity modulo 3. The same is true for $\rho = 4$. If $\rho = 6$, then $(|K_0|, |K_1|, |K_2|, |K_3|, |K_4|, |K_5|) = (1, 1, 1, 2, 1, 0)$. Now, if a symmetry $e^v c$ modulo 6 induces a rearrangement that leaves the vector invariant, then the rearrangement leaves K_3 invariant, but $e^v c(3) = 3 + v = 3$ and hence $v = 0$. Moreover, $e^0 5(1) = 5$ and hence it does not induce a rearrangement that leaves K_1 invariant, so $e^v c = e^0 1$.

Contrapuntal symmetries for 0ϵ.

According to the table in Section 8.2, there are exactly two symmetries for $\rho = 6$, namely $e^{6\epsilon}(1 + 6\epsilon)$ and $e^{6\epsilon}(7 + 6\epsilon)$, with $e^v c$ congruent to $e^0 1$ modulo 6. This allows us to discard the cases $\rho = 3, 2, 1$. Up to now the maximum is 48, and it remains to examine the cases $\rho = 4, 12$.

In the case when $\rho = 4$, the residues modulo 4 of the candidates in the table produce the following sums, calculated with Equation (18).

c \ v	1	2	3
1	36	16	36
3	24	20	48

Thus, we have found two new symmetries with sum 48 (the maximum up to now), namely $e^{11\epsilon}(11 + 8\epsilon)$ and $e^{11\epsilon}(11 + 4\epsilon)$. Here, $e^{11}11 \equiv e^3 3 \pmod 4$.

It remains to examine the case $\rho = 12$. We have the following sums.

c \ v	6	2	5	11
1	24			
7	36			
5		0	36	36
11		12	24	48

This means that the maximum sum is 48 and that there is yet another symmetry $e^{11\epsilon}11$.

Contrapuntal symmetries for 3ϵ

According to the table in Section 8.2, there are exactly two symmetries for $\rho = 4$, namely $e^{8\epsilon}(5 + 4\epsilon)$ and $e^{8\epsilon}(5 + 8\epsilon)$, with $e^v c$ congruent to $e^0 1$ modulo 4. We thus discard the cases $\rho = 6, 3, 2, 1$. Up to now the maximum is 56.

The remaining case is $\rho = 12$. We have the following sums.

c \ v	9	2	5	8
1	36			
7	24			
5		0	36	48
11		12	24	36

The maximum is 56 and there are no more contrapuntal symmetries.

Contrapuntal symmetries for 4ϵ

There are exactly two symmetries for $\rho = 6$, namely $e^{6\epsilon}(1 + 6\epsilon)$ and $e^{6\epsilon}(7 + 6\epsilon)$, with $e^v c$ congruent to $e^0 1$ modulo 6. This allows us to discard the cases $\rho = 3, 2, 1$. Up to now the maximum is 48. The sums are less than 48 for $\rho = 4$ as shown in the following table.

c \ v	1	2	3
1		16	36
3	24	20	

The case $\rho = 12$ yields the following sums.

c \ v	3	6	2	11	9	5
1	36	24				
5			0	36		
7		36			24	
11			12			24

Hence, the maximum is 48 and there are no more contrapuntal symmetries.

Contrapuntal symmetries for 7ϵ

The case $\rho = 12$ yields the following sums.

c \ v	0	6	2	8	9	5
1		24			36	
5			0			36
7	60				24	
11				36		24

Thus, we discard all remaining cases for ρ and the maximum is 60, corresponding to a unique contrapuntal symmetry $e^0 7$.

Contrapuntal symmetries for 8ϵ

There are exactly two symmetries for $\rho = 6$, namely $e^{6\epsilon}(1 + 6\epsilon)$ and $e^{6\epsilon}(7 + 6\epsilon)$, with $e^v c$ congruent to $e^0 1$ modulo 6. This allows us to discard the cases $\rho = 3, 2, 1$. Up to now the maximum is 48.

The sums for $\rho = 4$ are the same of the consonance 0ϵ. In the table from Section 8.2, we observe that there are two additional symmetries, namely $e^{3\epsilon}(7 + 4\epsilon)$ and $e^{3\epsilon}(7 + 8\epsilon)$, with $e^v c$ congruent to $e^3 3$ modulo 4.

The case $\rho = 12$ yields the following sums.

c \ v	2	3	6	9
1		36	24	36
5	0			
7		48	36	24
11	12			

Hence, the maximum is 48 and there is an additional contrapuntal symmetry $e^{3\epsilon} 7$.
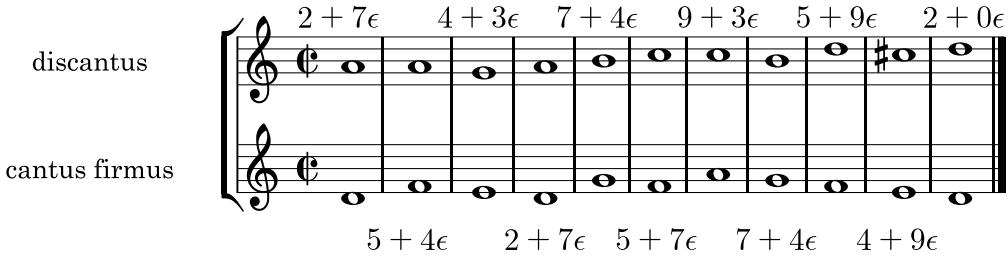
Contrapuntal symmetries for 9ϵ

There are exactly two symmetries for $\rho = 4$, namely $e^{8\epsilon}(5 + 4\epsilon)$ and $e^{8\epsilon}(5 + 8\epsilon)$, with $e^v c$ congruent to $e^0 1$ modulo 4. We thus discard the cases $\rho = 6, 3, 2, 1$. Up to now the maximum is 56.

For $\rho = 12$ we have the following sums.

c \ v	3	2	11	8
1	36			
7	48			
5		0	36	48
11		12	48	36

The maximum is 56 and there are no more contrapuntal symmetries.

Figure 3: *A first-species counterpoint example given by Fux, which is in the Dorian mode, and the contrapuntal intervals involved.*

8.4. Admitted successors

We can finally obtain the list of admitted successors, by directly using Equation (7) for each contrapuntal symmetry h obtained in Section 8.3; see Table 2.

The **prohibition of parallel fifths** is an important conclusion that we can draw from the table. In fact, the consonance 7ϵ has as set of admitted successors $\mathbb{Z}_{12} + (K \setminus \{7\})\epsilon$, that is, any interval can follow a fifth, except a fifth.

Figure 3 is an example of first-species composition in Dorian mode by Fux [8, p. 29]. After suitably transposing the transition from each bar to the next, we have the following counterpoint symmetries. The symmetry $e^0 7$ occurs from bars 1 to 2, 4 to 5, and 6 to 7. The symmetry $e^{8\epsilon}(5 + 4\epsilon)$ occurs from bars 3 to 4 and 9 to 10. The symmetry $e^{6\epsilon}(7 + 6\epsilon)$ occurs from bar 5 to bar 6. The symmetries $e^{6\epsilon}(7 + 6\epsilon)$ and $e^{6\epsilon}(1 + 6\epsilon)$ occur from bars 2 to 3 and 8 to 9. The symmetries $e^{8\epsilon}(5 + 4\epsilon)$ and $e^{8\epsilon}(5 + 8\epsilon)$ occur from bars 7 to 8 and 10 to 11.

9. A non-commutative counterpoint

Consider the data of Example 3.10. There, we have the strong dichotomy (K, D, e^I) of the ring R of upper triangular matrices on $\mathbb{Z}_2$, with $K = \{0, A_1, A_2, A_3\}$ and $D = \{I, B_1, B_2, B_3\}$, where $B_i = A_i + I$ for $i = 1, 2, 3$. In this case, $b = I$ and $a = I$. Let us compute the contrapuntal symmetries and admitted successors in the dual numbers case.

9.1. Condition 2

We solve the equation

$$v + I = c + v,$$

so $c = I$, and v and d range over R.

9.2. Condition 1

Now, among the previous solutions, we take, for each consonance $k\epsilon$, those satisfying

$$v \in k - D = k + D.$$

The following table contains the results.

k	$\mathbf{0}$	A_1	A_2	A_3
$k + D$	D	$\{B_1, I, A_3, A_2\}$	$\{B_2, A_3, I, A_1\}$	$\{B_3, A_2, A_1, I\}$

k	$\|h(K[\epsilon]) \cap K[\epsilon]\|$	h	admitted successors of $k\epsilon$
0	48	$e^{6\epsilon}(1+6\epsilon)$	$r + \{3,9\}\epsilon, r$ even $r + K\epsilon, r$ odd
		$e^{6\epsilon}(7+6\epsilon)$	$r + \{3,7,9\}\epsilon, r$ even $r + (K \setminus \{7\})\epsilon, r$ odd
		$e^{11\epsilon}(11+8\epsilon)$	$\{0,3,6,9\} + \{3,4,7,8\}\epsilon$ $\{1,4,7,10\} + \{0,3,7,8\}\epsilon$ $\{2,5,8,11\} + \{0,3,4,7\}\epsilon$
		$e^{11\epsilon}(11+4\epsilon)$	$\{0,3,6,9\} + \{3,4,7,8\}\epsilon$ $\{1,4,7,10\} + \{0,3,4,7\}\epsilon$ $\{2,5,8,11\} + \{0,3,7,8\}\epsilon$
		$e^{11\epsilon}11$	$\mathbb{Z}_{12} + \{3,4,7,8\}\epsilon$
3	56	$e^{8\epsilon}(5+8\epsilon)$	$\{0,3,6,9\} + \{0,4,7,8\}\epsilon$ $\{1,4,7,10\} + (K \setminus \{7\})\epsilon$ $\{2,5,8,11\} + (K \setminus \{9\})\epsilon$
		$e^{8\epsilon}(5+4\epsilon)$	$\{0,3,6,9\} + \{0,4,7,8\}\epsilon$ $\{1,4,7,10\} + (K \setminus \{9\})\epsilon$ $\{2,5,8,11\} + (K \setminus \{7\})\epsilon$
4	48	$e^{6\epsilon}(1+6\epsilon)$ $e^{6\epsilon}(7+6\epsilon)$	see $k=0$
7	60	$e^0 7$	$\mathbb{Z}_{12} + (K \setminus \{7\})\epsilon$
8	48	$e^{3\epsilon}7$	$\mathbb{Z}_{12} + \{0,3,4,7\}\epsilon$
		$e^{6\epsilon}(1+6\epsilon)$ $e^{6\epsilon}(7+6\epsilon)$	see $k=0$
		$e^{3\epsilon}(7+4\epsilon)$	$\{0,3,6,9\} + \{0,3,4,7\}\epsilon$ $\{1,4,7,10\} + \{3,4,7,8\}\epsilon$ $\{2,5,8,11\} + \{0,3,7,8\}\epsilon$
		$e^{3\epsilon}(7+8\epsilon)$	$\{0,3,6,9\} + \{0,3,4,7\}\epsilon$ $\{1,4,7,10\} + \{0,3,7,8\}\epsilon$ $\{2,5,8,11\} + \{3,4,7,8\}\epsilon$
9	56	$e^{8\epsilon}(5+8\epsilon)$ $e^{8\epsilon}(5+4\epsilon)$	see $k=3$

Table 2: *Contrapuntal symmetries and admitted successors for the cantus firmus 0. We obtain the admitted successors of $z + k\epsilon$ by adding z to the cantus firmus of the results.*

9.3. Maximization

Since $c = I$, and $I \in Z(R)$, the maximization criterion (Theorem 6.5) remains valid in this case and we use it.

We start by computing rR, ρ, K/rR, and $(|K_\gamma|_\gamma)$ in the following table. The ordering on γ is that written for the classes in K/rR. As before, the maximization criterion is not useful for the case $d = \mathbf{0}$ since the identity does not satisfy the condition 1.

r	rR	ρ	K/rR	$(|K_\gamma|_\gamma)$
A_1, A_2	$\{\mathbf{0}, A_1, A_2, B_3\}$	2	$\{A_1 R, A_1 R + A_3\}$	$(3,1)$
A_3, I	R	1	$\{R\}$	(4)
B_i for $i = 1,2,3$	$\{\mathbf{0}, B_i\}$	4	$\{\{\mathbf{0}, B_i\}, \{A_i, I\}, \{A_1, A_2, A_3\} \setminus \{A_i\}\}$	$(1,1,2,0)$

We have the following maximum values with the restriction $c = I$. In the case $\rho = 2$, the unique $[e^v]$ inducing the identity rearrangement of $(3,1)$ is the identity modulo A_1. In the case $d = B_1$, if $[e^v]$ induces an identity rearrangement of $(1,1,2,0)$, then necessarily $[e^v]$ leaves invariant $K_{\{A_2, A_3\}}$, which is equal to $\{A_2, A_3\}$. The possibilities are either $v = \mathbf{0}$ or $v = B_1$, that is, e^v is the identity modulo B_1. The cases $d = B_2$ and $d = B_3$ are similar.

d	ρ	maximum $\sum\limits_{r \in R} |(K + v + dr) \cap K|$	v
A_1, A_2	2	20	$\mathbf{0}, A_1, A_2, B_3$
A_3, I	1	16	any
B_1	4	24	$\mathbf{0}, B_1$
B_2	4	24	$\mathbf{0}, B_2$
B_3	4	24	$\mathbf{0}, B_3$

Now we must compute the cardinalities for $d = \mathbf{0}$, that is, those of the form $8|(K + v) \cap K|$ with $v \neq \mathbf{0}$ by the table in Section 9.2. We already have the computation of $K + I + k$, which is just $D + k$, for $k \in K$ by the same table. Moreover, $K + A_i = \{\mathbf{0}, A_i\} \cup (\{B_1, B_2, B_3\} \setminus \{B_i\})$. Thus, $8|(K + v) \cap K| = 0$ if $v = I$ and $8|(K + v) \cap K| = 16$ if $v \notin \{I, \mathbf{0}\}$, so the maximum cardinals (subject to conditions 1 and 2) occur for the symmetries with maximum value 24 in the preceding table. The following table shows the contrapuntal symmetries directly obtained.

k	h
$\mathbf{0}$	$e^{B_i \epsilon}(I + B_i \epsilon)$ for $i = 1,2,3$
A_i (with $i = 1,2,3$)	$e^{B_i \epsilon}(I + B_i \epsilon)$

9.4. Admitted successors

The direct use of Equation (7) for each contrapuntal symmetry h obtained in Section 9.3 produces the following table. In this example, the maximization criterion is useful to determine all contrapuntal symmetries.

k	h	admitted successors of $k\epsilon$
$\mathbf{0}$	$e^{B_i \epsilon}(I + B_i \epsilon), 1 \leq i \leq 3$	see $k = A_i$
A_i (with $i = 1,2,3$)	$e^{B_i \epsilon}(I + B_i \epsilon)$	$\{A_3, I, B_1, B_2\} + K\epsilon$ $\{\mathbf{0}, A_1, A_2, B_3\} + (\{A_1, A_2, A_3\} \setminus \{A_i\})\epsilon$

Note that some parallelisms of the consonance $A_i \epsilon$ are forbidden in this notion of counterpoint. The counterpoint notion that results from multiplication on the right is very similar.[20]

10. Local characterization equivalences

In this section we compare our condition 2 in Definition 4.1 with Mazzola's original condition [11, Definition 95].

The following result inspires the latter. It is a local version of the uniqueness property of a strong dichotomy. Observe that the conjugation in Proposition 4.5 restricts to bijections $H(K[\mathbf{x}], D[\mathbf{x}]) \longrightarrow H_z(K[\mathbf{x}], D[\mathbf{x}])$ between respective sets of symmetries sending $K[\mathbf{x}]$ to $D[\mathbf{x}]$ (check).

Theorem 10.1. *Let $(K, D, e^a b)$ be a strong dichotomy of R. For each $z \in R$, $p^z[\mathbf{x}]$ is the unique z-invariant symmetry such that $(K[\mathbf{x}], D[\mathbf{x}], p^z[\mathbf{x}])$ is a self-complementary dichotomy of $R[\mathbf{x}]$. In particular, by Lemma 10.3,*

$$p^z[\mathbf{x}](z + K\mathbf{x}) = z + D\mathbf{x}.$$

Proof. Since self-complementary dichotomies $(K[\mathbf{x}], D[\mathbf{x}], p')$ with p' z-invariant are in correspondence (conjugation) with self-complementary dichotomies $(K[\mathbf{x}], D[\mathbf{x}], p)$ where p is 0-invariant, the proof of the theorem reduces to the case when $z = 0$ (Proposition 10.2). There, $p^0[\mathbf{x}] = e^{a\mathbf{x}} b$, so $p^z[\mathbf{x}] = e^z \circ e^{a\mathbf{x}} b \circ e^{-z} = e^{(1-b)z + a\mathbf{x}} b$. $\square$

Proposition 10.2. *Let $(K, D, e^a b)$ be a strong dichotomy of R. The symmetry $e^{a\mathbf{x}} b$ is the unique 0-invariant symmetry such that $(K[\mathbf{x}], D[\mathbf{x}], e^{a\mathbf{x}} b)$ is a self-complementary dichotomy of $R[\mathbf{x}]$.*

Proof. Let h be a 0-invariant symmetry, where $h = e^{u + v\mathbf{x}}(c + d\mathbf{x})$. By Proposition 4.4, $u = 0$. Moreover, if $h(K[\mathbf{x}]) = D[\mathbf{x}]$, from Lemma 10.3, we obtain that $(c + d\alpha)K + v = D$, so the condition that $e^a b$ is a polarity implies that $c + d\alpha = b$ and $v = a$. Further, we claim that $d = 0$. In fact, by the condition that $h(K[\mathbf{x}]) = D[\mathbf{x}]$,

$$e^{v\mathbf{x}}(c + d\mathbf{x})(1 + K\mathbf{x}) = c + ((c + d\alpha)K + v + d)\mathbf{x} = c + (D + d)\mathbf{x} \subseteq D[\mathbf{x}]$$

and

$$e^{v\mathbf{x}}(c + d\mathbf{x})(-1 + K\mathbf{x}) = -c + ((c + d\alpha)K + v - d)\mathbf{x} = -c + (D - d)\mathbf{x} \subseteq D[\mathbf{x}],$$

so $D + d \subseteq D$ and $D - d \subseteq D$, and hence $D + d \subseteq D$ and $D \subseteq D + d$. Thus, $D = D + d$ and, since both K and D are rigid (Corollary 3.3), $d = 0$.

Up to now, we know that each 0-invariant symmetry h such that $h(K[\mathbf{x}]) = D[\mathbf{x}]$ is necessarily $e^{a\mathbf{x}} b$. It remains to show that $(K[\mathbf{x}], D[\mathbf{x}], e^{a\mathbf{x}} b)$ is a self-complementary dichotomy. This was done in Proposition 3.11. $\square$

Lemma 10.3. *If h is a z-invariant symmetry of $R[\mathbf{x}]$ such that $h(K[\mathbf{x}]) = D[\mathbf{x}]$, then $h(z + K\mathbf{x}) = z + D\mathbf{x}$.*

Proof. Under our assumptions,

$$h(z + K\mathbf{x}) = h(K[\mathbf{x}] \cap (z + R\mathbf{x})) = h(K[\mathbf{x}]) \cap (z + R\mathbf{x}) = D[\mathbf{x}] \cap (z + R\mathbf{x}) = z + D\mathbf{x}.$$

$\square$

We have the following corollary of Theorem 10.1.

---

[20]To be more exact, this notion is obtained by applying the previous theory to the opposite ring R^{op} instead of R. In that case, all right R-modules considered become left R-modules.

Corollary 10.4. *Let $(K, D, e^a b)$ be a strong dichotomy of R. For each $z \in R$, $K[\mathbf{x}]$ is H_z-rigid, and $f(K[\mathbf{x}]) = f'(K[\mathbf{x}])$ implies $f = f'$ whenever $f, f' \in H_z$.*

Proof. As in the proof of Proposition 3.2, there is a bijection between $H_z(K[\mathbf{x}], D[\mathbf{x}])$ and the stabilizer $\theta(K[\mathbf{x}])$ with respect to H_z. Thus, Theorem 10.1 implies that $\theta(K[\mathbf{x}])$ is trivial.

Now, suppose that $f(K[\mathbf{x}]) = f'(K[\mathbf{x}])$ with f, f' in H_z. This implies that $f'^{-1} \circ f(K[\mathbf{x}]) = K[\mathbf{x}]$, so $f'^{-1} \circ f = id$, and $f = f'$. $\square$

If we apply the condition of Theorem 10.1 to a dichotomy $\{\tilde{K}, \tilde{D}\}$ of $R[\mathbf{x}]$, we obtain the following one.

- The symmetry $p^z[\mathbf{x}]$ is the unique $p' \in H_z$ such that $(\tilde{K}, \tilde{D}, p')$ is a self-complementary dichotomy.

As explained in [10, Section 14.2], it is a sort of characteristic property of the partition $\{z + K\,\mathbf{x}, z + D\,\mathbf{x}\}$ in the case of the Renaissance dichotomy (K, D, p).

If we apply it to a deformed dichotomy $\{g(K[\mathbf{x}]), g(D[\mathbf{x}])\}$, for $g \in \mathbf{Sym}(R[\mathbf{x}])$, we get the following condition.

- The symmetry $p^z[\mathbf{x}]$ is the unique $p' \in H_z$ such that

$$(g(K[\mathbf{x}]), g(D[\mathbf{x}]), p')$$

 is a self-complementary dichotomy.

First, let us prove that the latter is just Mazzola's condition 2.

Proposition 10.5. *Let (K, D, p) be a strong dichotomy, $z \in R$, and $g \in \mathbf{Sym}(R[\mathbf{x}])$. The following conditions are equivalent.*

1. *The symmetry $p^z[\mathbf{x}]$ is the unique $p' \in H_z$ such that $(g(K[\mathbf{x}]), g(D[\mathbf{x}]), p')$ is a self-complementary dichotomy.*

2. *The triple $(g(K[\mathbf{x}]), g(D[\mathbf{x}]), p^z[\mathbf{x}])$ is a self-complementary dichotomy.*

3. *The equation $p^z[\mathbf{x}]g_z = g_z p^z[\mathbf{x}]$ holds, where g_z is as in Propostion 5.2.*

Proof. For each $p' \in H_z$, by Proposition 5.2 and Corollary 10.4, we have the following equivalences.

$$p'(g(K[\mathbf{x}])) = g(D[\mathbf{x}]) \Leftrightarrow$$
$$p'(g_z(K[\mathbf{x}])) = g_z(D[\mathbf{x}]) \Leftrightarrow$$
$$p'g_z(K[\mathbf{x}]) = g_z p^z[\mathbf{x}](K[\mathbf{x}]) \Leftrightarrow$$
$$p'g_z = g_z p^z[\mathbf{x}]$$

The last (and hence the first) equation has a unique solution for the unknown p', so 1 is equivalent to 3 and 2. $\square$

The condition of Proposition 10.5 is stronger than our characterizing condition in Definition 4.1. In fact, if $p^z[\mathbf{x}](g(K[\mathbf{x}])) = g(D[\mathbf{x}])$, then

$$p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = p^z[\mathbf{x}](g(K[\mathbf{x}])) \cap p^z[\mathbf{x}](I_z) = g(D[\mathbf{x}]) \cap I_z.$$

Finally, we establish that under the commutativity of R, both conditions are equivalent.

Theorem 10.6. *Let (K, D, p), with $p = e^a b$, be a strong dichotomy of a commutative ring, $z \in R$, and $g \in \mathbf{Sym}(R[\mathbf{x}])$. The following equations are equivalent.*

1. $p^z[\mathbf{x}](g(K[\mathbf{x}])) = g(D[\mathbf{x}])$

2. $p^z[\mathbf{x}](g(K[\mathbf{x}]) \cap I_z) = p^z[\mathbf{x}](g(D[\mathbf{x}])) \cap I_z$

Proof. Suppose that $g = e^{u+v\mathbf{x}}(c + d\,\mathbf{x})$ and write $w = v + dc^{-1}(z - u)$. By unraveling the condition 3 in Proposition 10.5, which is just 1, it is equivalent to the following set of equations.

$$\begin{cases} bw + a = (c + d\alpha)a + w \\ \quad\quad bc = cb \\ \quad\quad bd = db \end{cases}$$

On the other hand, 2 is equivalent to the condition 3 in Theorem 4.9, that is, to the following system.

$$\begin{cases} bw + a = (c + d\alpha)a + w \\ b(c + d\alpha) = (c + d\alpha)b \end{cases}$$

Hence, 1 and 2 are equivalent when R is commutative. $\square$

11. Variations of the model

The validity of Definition 4.1 for the case $\alpha = 1$ of the contrapuntal intervals ring, and the motivations in Sections 4.1, 4.2, and 4.3, give rise to the following questions on the classical model.

- Are dual numbers ($\alpha = 0$) the most appropriate structure to model contrapuntal intervals? According to Mazzola, the discantus is a tangential alteration of the cantus firmus, so this principle leads to dual numbers, which are the natural tangents in algebraic geometry. However, we could ask why this kind of alteration is involved in counterpoint and not another one.
- Why do we not require the local characterization condition on $\{g(K[\mathbf{x}]), g(D[\mathbf{x}])\}$ for the cantus firmus of a possible successor η?
- More radically, why do we not require the local characterization on $\{g(K[\mathbf{x}]), g(D[\mathbf{x}])\}$ for all fibers?

These questions lead to the following three variations of the model. We indicate how to compute the contrapuntal symmetries for the Renaissance dichotomy in each case. The musicological analysis of these results is in [6].

11.1. First variation

This variation corresponds to Definition 4.1 with $\alpha = 1$.

First, let us note that we can relate the models for $\alpha = 1$ and $\alpha = 0$. Certainly, there is an injection $\phi : e^{v\mathbf{x}}(c + d\,\mathbf{x}) \mapsto e^{v\epsilon}((c + d) + d\epsilon)$ from the set of symmetries in H satisfying[21] 1 and 2 in Theorem 5.4 for $\alpha = 1$ to the set of symmetries in H satisfying 1 and 2 for $\alpha = 0$. Moreover, ϕ preserves the cardinality of successors sets. In fact, if $h = e^{v\mathbf{x}}(c + d\,\mathbf{x})$, then

$$|h(K[\mathbf{x}]) \cap K[\mathbf{x}]| = \sum_{r \in R} |((c + d)K + v + dr) \cap K| = |\phi(h)(K[\epsilon]) \cap K[\epsilon]|.$$

---

[21]See Sections 7.1 and 7.2 for explicit forms of these conditions.

Thus, if M_0 and M_1 are the maximum cardinalities according to 3 in Theorem 5.4 for $\alpha = 0$ and $\alpha = 1$, respectively, then $M_0 \geq M_1$.

Moreover, if there is a contrapuntal symmetry in the image of ϕ, like in Table 2, then $M_0 = M_1$. In this case, we can compute the contrapuntal symmetries for $\alpha = 1$ as all inverse images of contrapuntal symmetries for $\alpha = 0$.

Also, note that $e^{v\epsilon}(c + d\epsilon)$ is in the image of ϕ if and only if $c - d \in R^*$, in which case its inverse image is $e^{v\mathbf{x}}((c - d) + d\mathbf{x})$.

11.2. Second variation

If we also require the local characterization on the fibers of the possible successors of a consonance, then we change condition 3 in Definition 4.1 for the following one, where $P(z)$ is the property 2 in Definition 4.1.

- The cardinality of $\{z' + k'\mathbf{x} \in g(K[\mathbf{x}]) \cap K[\mathbf{x}] \mid P(z')\}$ is maximum among all g satisfying 1 and 2.

By following the same steps to prove Theorem 5.4, we obtain (check) the following simplification of this definition.

Theorem 11.1. *The admitted successors of $k\mathbf{x} \in K[\mathbf{x}]$ are the elements of the sets T of the form*

$$\{z + k\mathbf{x} \in h(K[\mathbf{x}]) \cap K[\mathbf{x}]) \mid P(z)\},$$

where $h = e^{v\mathbf{x}}(c + d\mathbf{x}) \in H$ and

(a2) $k\mathbf{x} \in h(D[\mathbf{x}])$,
(b2) $p \circ e^v(c + d\alpha) = e^v(c + d\alpha) \circ p$, and
(c2) the cardinality of T is maximum among all $h \in H$ satisfying (a2) and (b2).

Suppose that $h = e^{v\mathbf{x}}(c + d\mathbf{x}) \in H$. Note that once (b2), which is just $P(0)$, holds, the proof of Theorem 10.6 implies that $P(z)$ is equivalent to $bdc^{-1}z = dc^{-1}z$. In turn, $P(cr)$ is equivalent to $bdr = dr$, and the latter is equivalent to $P((c + d)r)$ whenever $(c + d) \in R^*$ and R is commutative.

Assume that R is commutative. We relate the cases $\alpha = 1$ and $\alpha = 0$ by means of the injection ϕ (Section 11.1) again. In fact, it preserves (a2) and (b2) as it did for the first variation. Also, the cardinalities of the successors sets T and T' associated with h (subject to (a2) and (b2)) and $\phi(h)$, respectively, are equal by the following equations, where $V = \{r \in R \mid bdr = dr\}$.

$$|T| = \left| \bigcup_{r \in V} cr + (((c + d)K + v + dr) \cap K)\epsilon \right|$$

$$= \left| \bigcup_{r \in V} (c + d)r + (((c + d)K + v + dr) \cap K)\mathbf{x} \right| = |T'|$$

For the Renaissance dichotomy, in the following section we observe that the dual numbers model has all contrapuntal symmetries in the image of ϕ, so, as in Section 11.1, the contrapuntal symmetries for $\alpha = 1$ are the inverse images of those for $\alpha = 0$. But, by collecting these inverse images, we obtain the same original set of symmetries up to indeterminate names $\mathbf{x}, \epsilon$ (see Table 3), so *the contrapuntal symmetries coincide for $\alpha = 0$ and $\alpha = 1$ in the case of the Renaissance dichotomy.*

11.3. Third variation

If we require the local characterization property on the deformed dichotomy for all fibers, which amounts to a *global property*, then we change condition 2 of Definition 4.1 for the following one.

- For all $z \in R$, $P(z)$ holds.

The respective simplification is the following theorem.

Theorem 11.2. *The admitted successors of $k\,\mathbf{x} \in K[\mathbf{x}]$ are the elements of the sets S of the form*

$$h(K[\mathbf{x}]) \cap K[\mathbf{x}],$$

where $h \in H$ and

(a3) $k\,\mathbf{x} \in h(D[\mathbf{x}])$,
(b3) for all $z \in R$, $P(z)$ holds, and
(c3) the cardinality of S is maximum among all $h \in H$ satisfying (a3) and (b3).

According to the proof of Theorem 10.6, if $h = e^{v\,\mathbf{x}}(c + d\,\mathbf{x})$, the condition (b3) just says that (b2) is true and that $bdc^{-1}z = dc^{-1}z$ holds for all $z \in R$. Hence, (b3) is equivalent to (b2) and $bd = d$.

For the dual numbers case, we next prove that *the second and third variations coincide if $R = \mathbb{Z}_n$ and each consonance has at least an admitted successor in the third variation*, like in the case of the Renaissance dichotomy; see Table 3. On the one hand, if h satisfies (a3) and (b3), then it satisfies (a2) and (b2) and the successors sets (cardinalities) are the same for both definitions. In particular, each consonance has at least an admitted successor in the second variation too. On the other hand, if $h \in H$ with $h = e^{v\epsilon}(c + d\epsilon)$ is a contrapuntal symmetry in the second variation, then the set of all $z \in R$ such that $P(z)$ holds, that is, such that

$$bdz = dz,$$

is a principal ideal[22] $\langle o \rangle$ in R. We have two possibilities. If $\langle o \rangle = R$, then h satisfies (a3) and (b3). Otherwise, $\langle o \rangle \subsetneq R$ and the set T (which is nonempty) from Theorem 11.1 satisfies

$$
\begin{aligned}
|T| &= \left| \bigcup_{z \in \langle o \rangle} z + ((cK + v + dc^{-1}z) \cap K)\epsilon \right| \\
&= \sum_{z \in \langle o \rangle} \left| (cK + v + dc^{-1}z) \cap K \right| \\
&< \rho \sum_{z \in \langle o \rangle} \left| (cK + v + dc^{-1}z) \cap K \right| \\
&= \sum_{r \in R} \left| (cK + v + dc^{-1}or) \cap K \right| \\
&= \left| \bigcup_{r \in R} r + ((cK + v + dc^{-1}or) \cap K)\epsilon \right| \\
&= |e^{v\epsilon}(c + do\epsilon)(K[\epsilon]) \cap K[\epsilon]|,
\end{aligned}
$$

where $\rho = |\operatorname{Ker}(o \cdot -)| = |R|/|\langle o \rangle| > 1$ and $bdo = do$. We deduce that h satisfies (a3) and (b3), or there is h' satisfying (a3) and (b3) such that its successor set has a strictly greater cardinality than

---

[22]In fact, the solution set of $ax = 0$ in $\mathbb{Z}_n$ is $\langle o \rangle$, where $o = n/\gcd(a, n)$.

k	$\|h(K[\epsilon]) \cap K[\epsilon]\|$	h
0	48	$e^{6\epsilon}(1 + 6\epsilon)$ $e^{6\epsilon}(7 + 6\epsilon)$ $e^{11\epsilon}11$
3	48	$e^{8\epsilon}5$
4	48	$e^{6\epsilon}(1 + 6\epsilon)$ $e^{6\epsilon}(7 + 6\epsilon)$
7	60	$e^{0}7$
8	48	$e^{3\epsilon}7$ $e^{6\epsilon}(1 + 6\epsilon)$ $e^{6\epsilon}(7 + 6\epsilon)$
9	48	$e^{11\epsilon}11$ $e^{3\epsilon}7$ $e^{8\epsilon}5$

Table 3: *Contrapuntal symmetries for the Renaissance dichotomy in the third variation and the dual numbers case. The computation, by hand, is akin to that of Table 2—now we just add the condition $5d = d$ for contrapuntal symmetries, that is, $d = 0, 3, 6, 9$ or $\rho = 3, 6, 12$.*

the successors set of h. Our observations imply that the contrapuntal symmetries coincide for the second and third variations.

In the Renaissance dichotomy case, by the computation of those symmetries (Table 3), they are in the image of ϕ. The latter preserves the conditions (a3) and (b3) and successors sets cardinalities from $\alpha = 1$ to $\alpha = 0$, so the contrapuntal symmetries coincide for $\alpha = 1$ and $\alpha = 0$, as in Section 11.2.

To summarize, *the second and third variations coincide for the Renaissance dichotomy.*

12. Conclusions and further research

We generalized Mazzola's first-species counterpoint to arbitrary rings and refined some of its theoretic aspects. The first author recommends a simplification of the model, preserving its strengths. Some of these strong features are the following:

- **Economy**. Essentially, we deduce all mathematical results of the theory from just a fact: the uniqueness property of the consonance/dissonance partition. In particular, we deduce the parallel fifths prohibition in Renaissance counterpoint.
- **Generality and universality**. This generalization is just an expression of the original intention of establishing a universal counterpoint by detecting the essential features of the Renaissance incarnation. Thus, the model paves the way to many other forms of counterpoint, which opens up a lot of fields of musical experimentation.

Further developments may include the following points.

Generalizations and structural approaches

Further generalizations to categories could be developed. The use of the product ring $R \times R$ (case $\alpha = 1$ in Section 3.5) to model contrapuntal intervals, as an alternative to the dual numbers, helps

express the model in terms of universal constructions in the category of rings, so it could be translated to other categories.

Also, there are several open questions like whether there is an entirely structural proof of the *Counterpoint theorem*, which establishes all counterpoint symmetries and admitted successors of a given consonance, beyond algorithmic computations. This is suggested by the counting formulas for successors sets cardinalities (Theorem 6.4) and the maximization criterion (Theorem 6.5).

Study of new counterpoint worlds

The musical study of our generalizations is an open field of research. Composition processes should generate new music in these counterpoint worlds. The counterpoint world induced by Scriabin's mystic chord has also been subject of study [5].

Musicology

An initial comparison of the model's results with the *Missa Papae Marcelli* can be found in [13]. The analysis of other works of Renaissance polyphony is also a pending task.

Remaining species

Although there are some advances regarding the second-species [4], a mathematical theory of the remaining species, and the cases of three or more voices, are not at hand.

13. APPENDIX

Theorem 13.1 (Rearrangement inequality, variation). *Suppose given an ordered list of real numbers*

$$a_1 \geq a_2 \geq \cdots \geq a_n.$$

The inequality

$$\sum_{i=1}^{n} a_i^2 \geq \sum_{i=1}^{n} a_i a_{\sigma(i)}$$

holds for each permutation σ of $\{1, \ldots, n\}$. Moreover, the equality holds if and only if $(a_1, \ldots, a_n) = (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$.

Proof. Induction on n. The case $n = 1$ is just the equality $a_1^2 = a_1^2$. Now suppose the result for $n - 1$. We consider two cases.

First case: $a_{\sigma(1)} = a_1$. We can assume $\sigma(1) = 1$ (switch $\sigma(1)$ and 1) since this do not alter the sum $\sum_{i=1}^{n} a_i a_{\sigma(i)}$. In this case,

$$\sum_{i=1}^{n} a_i a_{\sigma(i)} = a_1^2 + \sum_{i=2}^{n} a_i a_{\sigma(i)} = a_1^2 + \sum_{i=1}^{n-1} a_i' a_{\sigma'(i)}',$$

where $a_i' = a_{i+1}$ and $\sigma'(i) = \sigma(i+1) - 1$ for each $i = 1, \ldots, n - 1$. Thus, by applying the induction hypothesis,

$$\sum_{i=2}^{n} a_i a_{\sigma(i)} \leq \sum_{i=2}^{n} a_i^2$$

and the equality holds if and only if $(a_2, \ldots, a_n) = (a_{\sigma(2)}, \ldots, a_{\sigma(n)})$. Hence, by adding a_1^2 on both sides,

$$\sum_{i=1}^{n} a_i a_{\sigma(i)} \leq \sum_{i=1}^{n} a_i^2$$

and the equality holds if and only if $(a_1, \ldots, a_n) = (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$.

Second case: $a_{\sigma(1)} \neq a_1$. Note that there is k with $a_{\sigma(k)} = a_1$ such that $a_k < a_1$; otherwise $\{k \mid a_{\sigma(k)} = a_1\} \subseteq \{k \mid a_k = a_1\}$ and, since these sets have the same finite cardinality, they are equal, which is a contradiction since 1 is in the latter but not in the former. Thus,

$$\sum_{i=1}^{n} a_i a_{\sigma(i)} = a_1 a_{\sigma(1)} + a_k a_{\sigma(k)} + \sum_{i \in \{1, \ldots, n\} \setminus \{1, k\}} a_i a_{\sigma(i)}$$

and, if σ' is obtained from σ by switching $\sigma(1)$ and $\sigma(k)$, then

$$\sum_{i=1}^{n} a_i a_{\sigma'(i)} - \sum_{i=1}^{n} a_i a_{\sigma(i)} = (a_1 - a_k)(a_{\sigma(k)} - a_{\sigma(1)}) = (a_1 - a_k)(a_1 - a_{\sigma(1)}) > 0.$$

This means that

$$\sum_{i=1}^{n} a_i a_{\sigma(i)} < \sum_{i=1}^{n} a_i a_{\sigma'(i)} \leq \sum_{i=1}^{n} a_i^2,$$

where the right-hand inequality corresponds to the first case. Finally, note that here the inequality is always strict and $(a_1, \ldots, a_n) \neq (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$. □

REFERENCES

[1] Agustín-Aquino, O. 2011. *Extensiones Microtonales de Contrapunto (English Version)*. Dissertation (PhD in Mathematics), Universidad Nacional Autónoma de México. Available at: `https://drive.google.com/file/d/0Bw88kkzzRdcUY2FyRHFnZHcxTWM/view`.

[2] Agustín-Aquino, O. 2015. *Tod der Homophonie! Es lebe der Kontrapunktsatz!*, Available at: `https://www.youtube.com/watch?v=2nh8xoLwC0o`.

[3] Agustín-Aquino, O.; Junod, J.; Mazzola, G. 2015. *Computational Counterpoint Worlds: Mathematical Theory, Software, and Experiments*. Cham, Switzerland: Springer International Publishing.

[4] Agustín-Aquino, O.; Mazzola, G. 2018. A Projection-Oriented Mathematical Model for Second-Species Counterpoint. *Preprint*, Available at: `https://arxiv.org/abs/1810.00306`.

[5] Agustín-Aquino, O.; Mazzola, G. 2019. Contrapuntal aspects of the mystic chord and Scriabin's Piano Sonata No. 5. In: Mathematics and Computation in Music. Cham, Switzerland: Springer International Publishing, pp. 3–20.

[6] Arias-Valero, J.; Agustín-Aquino, O.; Lluis-Puebla, E. 2021. Musicological, Computational, and Conceptual Aspects of First-Species Counterpoint Theory. *Preprint*, Available at: `https://arxiv.org/pdf/2102.11767.pdf`.

[7] Fraleigh, J. 2003. *A First Course in Abstract Algebra, 7th edition*. New York: Pearson Education.

[8] Fux, J.; Mann, A.; Edmunds, J. 1965. *The Study of Counterpoint from Johann Joseph Fux's Gradus ad Parnassum*. New York: W. W. Norton.

[9] Junod, J. 2010. *Counterpoint Worlds and Morphisms : A Graph-Theoretical Approach and Its Implementation on the Rubato Composer Software*. Dissertation (PhD), University of Zurich. Available at: `https://www.zora.uzh.ch/id/eprint/163946/`.

[10] Mazzola, G. 2007. *La Vérité du Beau dans la Musique*. Paris: Editions Delatour France.

[11] Mazzola, G. et al. 2002. *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance*. Basel: Birkhäuser.

[12] Mazzola, G.; Wieser, H-G.; Brunner, V.; Muzzulini, D. 1989. A Symmetry-Oriented Mathematical Model of Classical Counterpoint and Related Neurophysiological Investigations by Depth EEG. *Computers & Mathematics with Applications*, 17/4–6, pp. 539–594.

[13] Nieto, A. 2010. *Una Aplicación del Teorema del Contrapunto*. Bachelor's thesis, ITAM.

[14] Rotman, J. 2015. *Advanced Modern Algebra: Third Edition, Part 1*. Providence, RI: American Mathematical Society.

Prime Decomposition Encoding: An Analytical Tool by the Use of Arithmetic Mapping of Drum-Set Timelines

*Carlos Mathias

Fluminense Federal University

profcarlosmathias@gmail.com

Orcid: 0000-0001-7949-2324

Carlos Almada

Federal University of Rio de Janeiro

carlosalmada@musica.ufrj.br

Orcid: 0000-0001-5888-2868

Abstract: *This paper introduces an original proposal intended to encode timelines as univocal integers, by the use of arithmetic mapping, exploring inherent properties of prime numbers. A group of algorithms were developed in order to encode drum-set timelines (as well as to retrieve the original rhythms from the codes), either individually or taken together, forming grooves. Additional parameters (dynamic levels and timbre) are also included in the encoding process. Geometrical representation of the grooves, adopting terminology and methodology proposed by Gottfried Toussaint[19], is also provided. Some practical application, addressing analysis and composition, are suggested at the last section of the study.*

Keywords: *Fundamental Theorem of Arithmetic. Prime encoding of rhythms. Drum-set timelines. Geometrical representation of rhythms.*

I. Introduction

The historical blood links that connect music and mathematics are of general knowledge, dating back to about 500 BC, with Pythagoras' formulations on acoustic proportions, basically associating numbers and sounds. This represents the very foundation of what would be known as the *Musica Speculativa* tradition, to be pursued along the subsequent 1000 years by great

---

mathematicians/physicists like Euclid, Ptolemy, Galileo, Kepler, Descartes, Newton, and Euler, involving a vast range of aspects, strongly concentrated on pitch relations.

This noble tradition was reborn in a way during the twentieth century, with Schoenberg's twelve-tone method, which can be considered as a starting point and source of inspiration for a number of mathematical-musical formalizations that have changed and influenced profoundly the thought about musical creation. Among these, we can mention three decisive landmarks: (1) the Pitch-Set Class Theory, consolidated by the efforts of Milton Babbitt, Allen Forte, John Rahn, Jack Boss, and Joseph Straus; (2) David Lewin's Transformational Theory, which gave rise to the (3) Neo-Riemannian Theory, consolidated through collaborative and complementary approaches by Bryan Hyer, Richard Cohn, Jack Douthett and Peter Steinbach, Clifton Callender, Adrian Child, Edward Gollin, David Clampitt, Henry Klumpenhouwer, Robert Peck, David Kopp, among others. Since the advent of this true "cambrian explosion" in the realm of music theory, the interest on the use of mathematical models for formalizing compositional and analytical studies, branching or not from these three basic axes, has increased a lot among scholars around the world, evidenced by the great number of published dissertations, articles, and books on correlated topics in the last years.

In spite of the fact pitch subjects are certainly at the very center of gravity of such formulations (ultimately reflecting the focus of the three landmarks), an increase interest in respect to rhythm corresponds to a very recent and promising research trend. Such an avenue was opened especially by the implications that arise from isomorphic relations between pitch and rhythm realms. This is epitomized in a well-known article written by Justin London [15], who discusses the isomorphic relations that connect pitch and temporal domains based on a graph-theory perspective. The central element of London's proposal is the concept of *Zeitnetz* (associated with Cohn's idea of ski-hill graphs), as an analogous metric counterpart of the neo-Riemannian *Tonnetz*.

Certainly, meter has been being the most explored aspect of time in music, aiming at mathematical-like formalization. Richard Cohn [3] addresses Harald Kreb's concept of metric dissonance (extending it also to hypermetric levels) in an analysis of the *Menuetto* of Mozart's G-minor symphony. He uses some of his own reflections arisen in a previous study [2] as a base for a formal approach, searching for a new theory constructed from properties shared by pitch and time. Cohn [4] expands his metric theory by addressing the notion of *complex hemiolas* (i.e., involving the superimposition of two or more rhythmic layers conflicting by double and triple metric organizations). Aiming at a formal analytical treatment for the issue, Cohn idealizes an algebraic description of rhythms as products of powers of the two first prime numbers, 2 and 3, capturing the very essence of hemiolas, a strategy which in a sense keeps some relation to that proposed in this article. A graphic representation of conflicting layers in action in complex hemiolas (called *sky-hill graph*) is also provided by Cohn, as a helpful analytical tool. In a PhD dissertation, Stephen Guerra [10] adopts a theoretical framework mostly based on Cohn's formulations on metrical hierarchical levels to explore the music of Brazilian guitarist Baden Powell. In this work, Guerra proposes a model destined to address the concept of *extended hemiolas*, which extrapolates the conventional duple/triple rhythmic opposition. In a recent article, Guerra [11] uses his previous study as basis for another theoretical investigation motivated by the analytical observation of Powell's music, considering the manners with which six distinct samba timelines can be related by transformational classes of operations (labeled as augmentation and rotation). Meter is also one of the parameters (aside phrase structure, contrapuntal structure, pitch spelling, harmony, and key) selected by David Temperley [18] for the elaboration of a computational model focused on the cognition of musical structures, based on a system of preference rules. Mark Gotham [7] proposes to identify the principles that govern the notion of *attractor tempos*, responsible for the optimization of the salience of certain metric structures. For this purpose, Gotham elaborates a

mathematical model capable of quantifying pulse salience, extended then to establish metrical salience values, according to hierarchical levels considered. The author provides an exhaustive list of metric structures based on consecutive levels of 2 and 3 groupings (p. 32), including the respective attractor indices.

Evidently, meter is not the only subject of scholarship interest with respect to musical time studied in mathematical terms. There is considerable literature that addresses specifically rhythmic issues. Catherine Guastavino and colleagues [9] propose an empirical-based investigation on flamenco rhythms directed to measurement of rhythmic similarity. The authors compare two mathematical methods for measuring similarity, namely *chronotonic distance* and the *directed swap distance*, the latter considered as the most efficient in the conclusion of the article. A distinctive trait of this approach is the use of *phylogenetic analysis* for depicting similarity relations between the rhythmic varieties. Francisco Gómez-Martín, Perouz Taslakian, and Godfried Toussaint [8] address two rhythmic categories, namely interlocking and Euclidean rhythms. While the latter category corresponds to "rhythms where the onsets are distributed among the pulses *as evenly as possible*" (p.18, italics in the original), the notion of interlocking rhythms refers to interaction of two or more rhythmic lines resulting into complex textures (as in African-Cuban music or in a heavy-metal groove), an aspect to be explored in the present study. In this sense, *complementary interlocking rhythms*, that form one of the possible types listed, are here of special interest. A recent empirical study by Florian Hoesl and Olivier Senn [14] is focused on the measurement of syncopation in drum-set grooves. It consists of a listening experiment, in which pairs of drums patterns, combined from a set of six popular grooves, were presented to 25 participants, who ranked them according to the degree of syncopation. The findings allow the authors to elaborate a formal model able to quantify syncopation in popular-music genres. Beyond the focus on the same instrument, there are other interesting shared characteristics by Hoesl's and Senn's article and the present study, especially associated with mathematical solutions for depicting rhythmic information. This will be commented in due time. Jay Hardesty [12] investigates the issue of similarity between rhythmic variants, generated from a collection of basic units. An original aspect of his work is the mapping of rhythmic derivations on Pascal's triangle, intended to provide a precise measurement of similarity between variants. In a recent PhD dissertation, Julio Herrlein [13] proposes the idea of a rhythmic set theory, in analogy to pitch-class sets, which includes a modulo-12 classification. Cohn [5] introduces the concept of *funky rhythms*, referred to the superimposition of triple-generated rhythmic patterns to duple-meter cycles formed by 8, 16, 32, and 64 units. Cohn's model is deeply associated with Plato's ideas about acoustic relations. The author elects for initiating his exposition the so-called *tresillo*,[1] present in several musical contexts around the world, and which represents perfectly the 3-over-2 duality in question. This model is adopted as basis in a study by Scott Murphy [16], whose main focus lies on the presence of funky rhythms in film music. As a theoretical framework for his analyses, Murphy proposes a detailed typology and catalogue 104 successions of 2s an/or 3s (durations, taking the sixteenth note as unit) with cardinalities from two to six elements, organizing them inside an Euler diagram, according to five categories by him established: the binary oppositions "maximally/minimally even" and "trochaic"/"iambic", and the unary "platonic".

The present article proposes a new and original manner to address rhythmic patterns, by associating their respective configurations of onsets with particular codes. This is achieved

---

[1] *Tresillo* is how is mostly known a rhythmic cell formed by the asymmetric division <3, 3, 2> of an eight-pulse temporal span. Normally associated with Cuban music (commonly as the "*tresillo* clave"), this pattern is also present in a wide geographical spectrum, especially in Sub-Sahara musical cultures and Latin-America popular-music genres (*baião, milonga, habanera, rumba, cha-cha, tango, maxixe, guajira*, etc.).

through some principles and formulations of Number Theory, notably the Fundamental Theorem of Arithmetic and a new function inspired by the function known as Gödel numbering. Aiming at obtaining concrete practical results and immediate applications of what will be discussed in theoretical terms along this study we direct our focus initially on the particular universe of drum-set grooves. As it will be demonstrated, the strategy of encoding rhythms (either individually or superimposed) as integers have multiple advantages, considering both analytical as compositional perspectives, opening some possible new avenues for further exploration. Before properly describing the mathematical framework that grounds this proposal, it is necessary to present the main conceptual and terminological references with respect to musical constructs adopted in our approach. These are derived from the studies of mathematical properties of rhythm as discussed by Toussaint in his 2013 book, *The Geometry of Musical Rhythm* [19].

II. Toussaint's geometrical approach of rhythm

In his book, Toussaint proposes a very original manner to address rhythmic and metrical organizations. Adopting essentially an approach which involves several mathematical domains, but focusing especially on geometrical representation, Toussaint aims to systematically investigate the reasons why some rhythms are more preferable (or "good", on his terms) than others – regardless of musical style, genre, epoch, or cultural context – simply by exploring some inherent mathematical properties that these rhythms have.

Four concepts of paramount importance for the present study, namely the notions of *pulse*, *onset*, *inter-onset intervals*, and *timeline*, are defined in separated passages of Toussaint's introductory notes:

> The term "pulse" is used in this book to denote the location at which a sound or attack may be realized [19, p. 5].
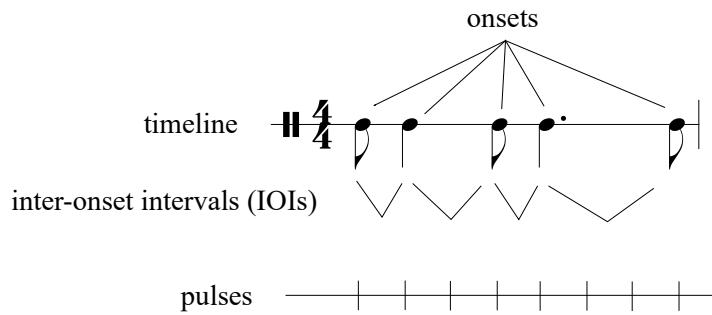
> The starting and ending times of the notes are the *onsets* and *offsets*, respectively. In the case of pure rhythms consisting of beats (attacks), it is assumed that there are no sustained notes, and thus we dispense with the offsets altogether. In this setting, the inter-onset-duration intervals are simply the durations between two consecutive attacks [19, p. 10].

> In much traditional, classical, and contemporary music around the world, one hears a distinctive and characteristic rhythm that appears to be an essential feature of the music, that stands out above the other rhythms, and that repeats throughout most if not the entire piece. Sometimes this essential feature will be merely an isochronous pulsation without any recognizable periodicity. At other times, the music will be characterized by unique periodic patterns. These special rhythms are generally called *timelines*. Timelines should be distinguished from the more general term *rhythmic ostinatos* [19, p. 13].

Figure 1 tries to combine these concepts in a unique simple illustration, using musical notation. Observe that the five onsets of the rhythm (or timeline) are distributed inside a time-span of eight pulses, considering the eighth note – the shorter value in this context – as the unit of measurement.

Among 4,368 possible five-onset-in-sixteen-pulse timelines, Toussaint distinguishes six special "good rhythms", identified by him as *Shiko*, *Son*, *Rumba*, *Soukous*, *Gahu*, and *Bossa-nova*.[2] A

---

[2]Toussaint informs that the names adopted "reflect the terminology perhaps most well known in the Western popular

Figure 1: *Example of timeline, pulses, onsets, and inter-onset intervals.*

number of mathematical properties of these rhythms are discussed along the chapters, allowing for Toussaint to provide systematical comparisons. In order to properly studying these (and many other) rhythms, Toussaint presents several notation types, adopting some of them as preferential. Figure 2 illustrates the use of four of these types (binary, box, interval-vector, and polygonal) for describing the rhythm of Figure 1.

The two first types are positional descriptions of rhythms, that is, they indicate presence (1s in the former, dots in the latter) or absence (0s, empty spaces) of onsets in a given universe of pulses. On the other hand, interval-vector notation informs durational distances (measured as multiples of the considered unit) between onsets. Polygonal notation combines the two basic categories–positional and intervallic–in a geometrical representation. It will be preferentially employed in the present article, due to its special advantages and clarity, eventually complemented by the information conveyed by IOI vectors.[3] Pulses are represented as little gray-line circles that divide the large circle. They are numbered from zero to n-1, where n is the number of pulses the timeline is associated with. Onsets form vertices of a polygon (that ultimately represents the rhythm in question). The sizes of the polygon's edges are proportional to the IOIs, with their magnitude corresponding to the number of segments in the circle that they cover (for example, the edge connecting pulses 4 and 7 covers three segments, namely 4-5, 5-6, and 5-7, corresponding therefore to a IOI of three unities). From now on, let us call this notation a necklace representation, adopting Toussaint's terminology. Given the adopted set of basic concepts, terms, and rhythmic representations, we can go on with the mathematical framework which grounds the proposal.

III. Rhythmic encoding

In his article of 1931 entitled "On Formally Undecidable Propositions of Principia Mathematica and Related Systems", Austro-Hungarian mathematician Kurt Gödel (1906-1978) proved the First Incompleteness Theorem. As presented in the Britannica Encyclopaedia homepage :[4] "[The theorem] states that any integer greater than 1 can be expressed as the product of prime numbers in only one way." It is of central importance to add that Gödel numberings are used for expressing any sequence of symbols, as mathematical formulas and even formal proofs of theorems.

The proof of this theorem has three well-defined parts. In the first one, Gödel proposes an innovative strategy, by associating each symbol/formula considered in the formal system F with

---
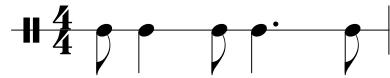
media, and I use them here purely for convenience rather than the establishment of any historical priority or cultural entitlement" [19, p. 28].

  [3]Polygonal representation is also preferential for Toussaint, which is especially associated with the possibility of evidencing symmetrical relations of several types, an omnipresent and very important topic in his book.

  [4]https://www.britannica.com/topic/Godels-first-incompleteness-theorem
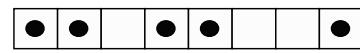
REPRESENTATIONS

musical

binary

<11011001>
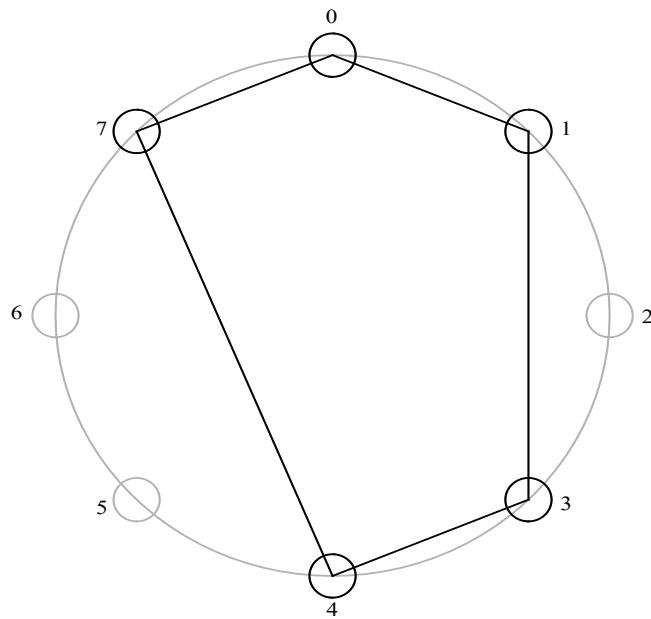
box notation

interval-vector

<12131>

as a polygon (necklace)

Figure 2: *Four equivalent representations for the example of Figure 1.*

Figure 3: *A simple rhythmic motive with three onsets.*

an integer (called Gödel number), so that such symbol/formula is only retrievable from that number.[5] The strategy adopted by Gödel inspires similar intentions in this article, more specifically with respect to the creation of a unique association between formal musical representations (more specifically, rhythmic) and codes expressed as integers, in order to allow the recovery of the former from the latter.[6] In fact, more than just encoding numerically the rhythms, we are also interested in relating mapping and overlapping of musical streams through arithmetic operations involving their generating codes.

Keeping in the focus this central goal, it is paramount at this moment to recall the main result of the integer arithmetic, the *Fundamental Theorem of Arithmetic* (FTA, for short). This result was first proved by German mathematician Carl Friedrich Gauss (1777-1855) at the very end of the eighteenth century and published in 1801 on his treatise *Disquisitiones Arithmeticae*. The proposition 30, 31, and 32 of the book VII of Euclid´s Elements presents results that show how close the Greeks were to the FTA, 2000 years before Gauss.

The Fundamental Theorem of Arithmetic: Let N be an integer other than 0, 1, and -1. Then there exist positive prime numbers $p_1 < p_2 < \ldots < p_k$ and positive integers $n_1, n_2, \ldots, n_k$ such that

$$N = \pm p_1^{n_1} \times p_2^{n_2} \times ... \times p_k^{n_k} \tag{1}$$

Disregarding permutations of the factors, this decomposition is unique. If we extend the perception of such a result beyond the specificities involved, we will realize that, more than a single decomposition of an integer as a product of prime-number powers, FTA explains the possibility of encoding a vector containing k numeric information $< n_1, n_2, \ldots, n_k >$ as a single number, in a bijective way. For example, the decomposition presented by equality $360 = 2^3 \times 3^2 \times 5^1$ can be understood as the encoding of the information conveyed by the vector $<3, 2, 1>$, formed by the exponents of the prime product, in the number 360. In this sense, the involved prime numbers address the number of coordinates/information conveyed by the vector, while its exponents constitute the information itself. In the present article, musical representations will adopt this same encoding process.

As a very simple example, consider the following rhythmic motive, corresponding to a possible manifestation of the *tresillo* clave (Figure 3).

Considering the duration of an eighth note as a unit of informational space, this representation configures Toussaint's binary vector $<10010010>$. If we use FTA for encoding this vector, we have this unique possible factoring:
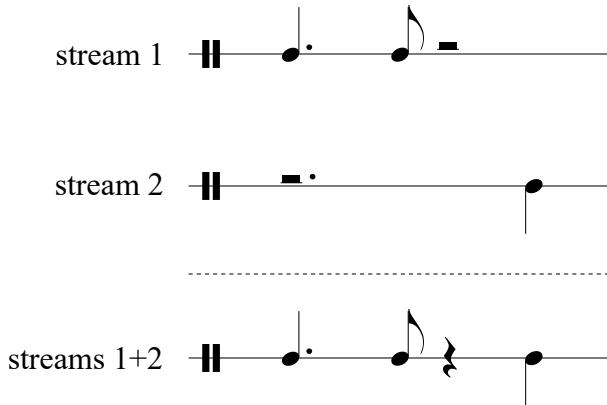
$$2^1 \times 3^0 \times 5^0 \times 7^1 \times 11^0 \times 13^0 \times 17^1 \times 19^0 = 238$$

---

[5]For a detailed presentation of this algorithm, which came to be known as Gödel numbering, as well as a comprehensive explanation of Gödel's achievements, see [17].

[6]As far goes our knowledge, the idea of expressing rhythmic formulas through the use of the Gödel-numbering algorithm was not yet explored in academic studies. Considering other aspects of musical construction, see the strategy proposed by [1] for codifying as univocal integers strings of genealogical descriptions of variations algorithmically produced. In this case, the order number of a variant located in the context of a given sequence of previous generation is represented as a series of exponents raising a sequence of prime bases, whose product corresponds to an integer that stands for the precise and formal identification of the variation in question.

Therefore, the number 238 uniquely identifies the rhythmic motive of Figure 3. In other terms, 238 is a numeric code univocally related to this specific musical representation.

A fundamental point to be observed is that the *product* of the codes referring to two distinct musical representations (since both share the same durational unit) will be the code related to the representation of the superimposition of those two representations, if they start at the same instant and have the same time-span pattern. Consider, for example, the distribution of the onsets of the *tresillo* of Figure 3 in two streams, as if played by two distinct instruments (Figure 4).

Figure 4: *Distribution of the onsets of the rhythm of Figure 3 in two separated streams.*

Thus, we have:
$$2^1 \times 3^0 \times 5^0 \times 7^1 \times 11^0 \times 13^0 \times 17^0 \times 19^0 = 14 \text{ (code of the stream 1)}$$
$$2^0 \times 3^0 \times 5^0 \times 7^0 \times 11^0 \times 13^0 \times 17^1 \times 19^0 = 17 \text{ (code of the stream 2)}$$

In this case, the product of the two codes, $17 \times 14 = 238$, equals the representation code referring to superimposed rhythms, as shown in Figure 3. Thus, we can affirm that the greatest common divisor (gcd) of the codes related to N representations in a same informational context will be the representation code referring to the commonly shared information. In the absence of any shared information, gcd equals 1, which is the code for the representation associated with the null vector.

For specific purposes, let us define an *informational unit* (IU, for short) as the smallest durational value considered in a given rhythmic context, which is equivalent to Toussaint's notion of pulse, as presented in the previous section. The strategy of setting IUs as contextual-dependent instead of adopting a fixed unit (say, the 16th note) is motivated by an intention of minimizing computational cost. Figure 5 illustrates the contextual nature of IUs with different situations.[7] In two distinct depictions of the *tresillo* (labeled as a and b), the informational unit corresponds to the 8th note, disregarding if it is explicitly or not present in the rhythm. In (c) the smallest value is clearly the 16th note. The last rhythm, however, combines groups of 8th notes (figure that divides the beat by 2) and of triplets (a ternary division). Since the least common multiple referred to these two values (2 and 3) is 6, this means that the implicit sixteenth triplet must be assigned as IU in the case.

In addition, it is also possible to incorporate other types of information in a code as, dynamics, for example. This is reflected in the algorithm by defining other possible values (beyond 0 and 1) to be assumed by the exponents. We can, for example, establish three possible dynamic levels for

---

[7]Observe how in each rhythm the pulses are counted from 0 to n-1, where n is the number of IUs present in the rhythm.
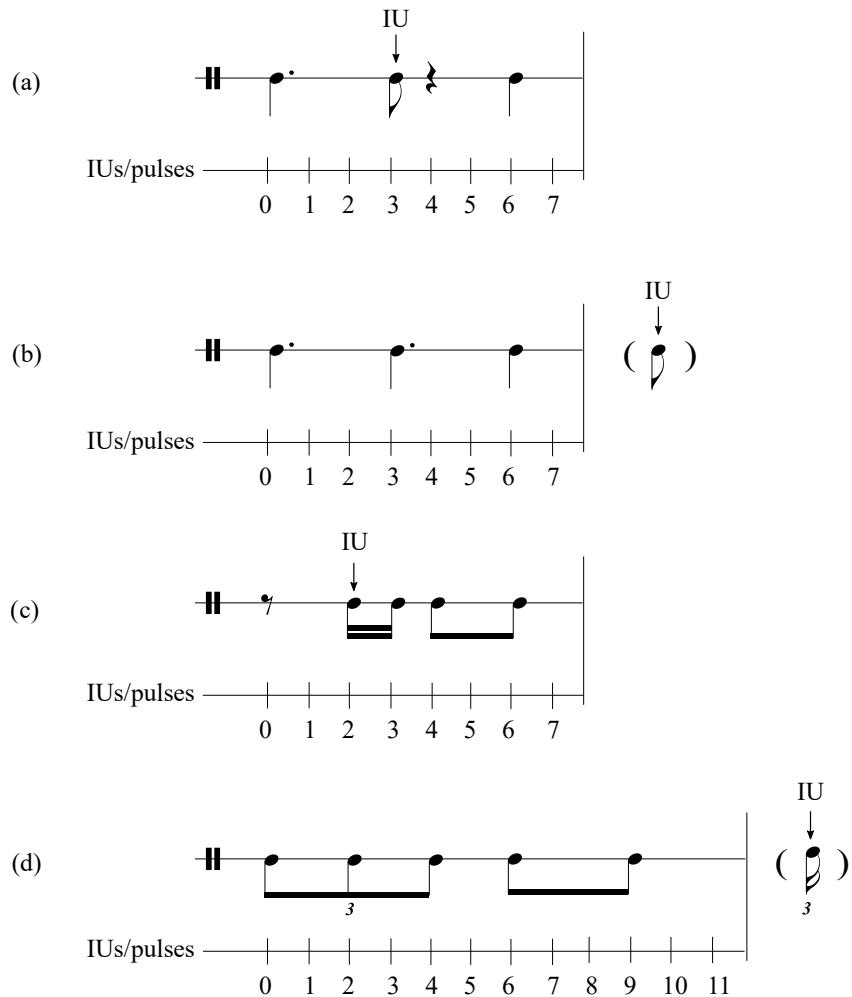
Figure 5: *Four rhythms with different IUs.*

any onset of a given rhythm: 1 (default or "flat"), 2 ("ghost" note) or 3 (accented).[8] In this sense, the intention of adding a phenomenal accent to the third onset of the *tresillo* of Figure 3 would result in a new code: $2^1 \times 3^0 \times 5^0 \times 7^1 \times 11^0 \times 13^0 \times 17^3 \times 19^0 = 68782$. Therefore, with each unit mapped by a prime number, it is possible to assign different sets of informational values, depending on the level of detail that one intends. Evidently, the more information is included, the bigger the integers that represent the codes. Anyway, and this is of paramount importance, there will be always a one-to-one correspondence between a code and a rhythmic context which it is intended to represent.

We propose the following formalization of the code, associated with a measurable rhythmic representation by a determined unit (IU), retrievable with k information.

Definition: Let R be a rhythmic representation measurable by an informational unit U, with

---

[8]These numbers were arbitrarily chosen. As it will be discussed in due time, the same strategy can be applied to other parameters, like timbral identification, for example.
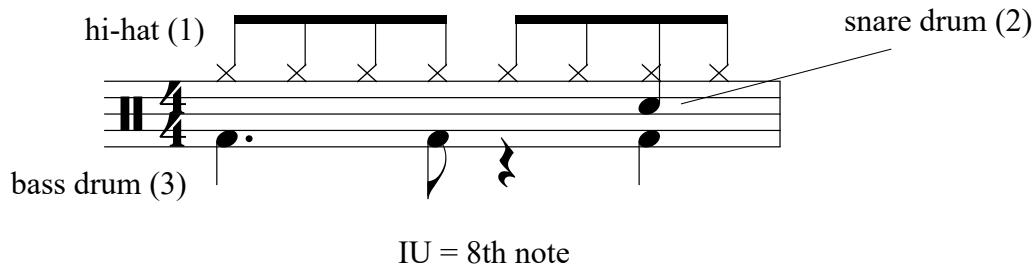
measure k.[9] Let $I_n \subset \mathbb{Z}^*_+$, $1 \leq n \leq k$ be the sets of informational values of the coordinates of the vector $(c_1, c_2, \ldots, c_k)$, associated with the representation R.

Let B = $\{p_1, p_2, \ldots, p_k\}$ a set made up of ordered, distinct, and positive prime numbers (p_1 = 2). The code M of the representation R, in relation to the set B, is given by:

$$M = p_1^{c_1} \times p_2^{c_2} \ldots p_{k-1}^{c_{k-1}} \times p_k^{c_k} = \prod_{n=1}^{k} p_n^{c_n} \tag{2}$$

The M code is unique and retrieves all information contained in the information sets associated with a given rhythmic representation.

As an illustration of the type of application that will be considered in this study, consider a simple eight-pulse drums pattern composed by the superimposition of three rhythmic lines, played by the hi-hat (stream 1), snare drum (stream 2), and bass drum (stream 3), as depicted in Figure 6.

IU = 8th note

Figure 6: *Example of a simple three-stream drums pattern.*

The respective codes are calculated as following:
$$M_{\text{hi-hat}} = 2^1 \times 3^1 \times 5^1 \times 7^1 \times 11^1 \times 13^1 \times 17^1 \times 19^1 = 9699690$$
$$M_{\text{snare}} = 2^0 \times 3^0 \times 5^0 \times 7^0 \times 11^0 \times 13^0 \times 17^1 \times 19^0 = 17$$
$$M_{\text{bass}} = 2^1 \times 3^0 \times 5^0 \times 7^1 \times 11^0 \times 13^0 \times 17^1 \times 19^0 = 238$$

It should be noted that gcd(17, 238, 9699690) = 17, which is precisely the prime number associated with the seventh pulse, the unique point where onsets of the three drum pieces coincide. In this sense, the concept of greater common divisor and the disposition of the rhythmic attacks in the pattern are closely related.[10]

IV.   Algorithmic implementation

The combination of the ideas of the Gödel-numbering algorithm with vector/necklace representation provides the necessary means for the elaboration of a robust theoretical-methodological apparatus addressing rhythm, with both analytical and compositional applications. As suggested in the previous section, this strategy allows for encoding any possible rhythmic configuration (or even superimposed rhythmic lines, as happens in drum-set grooves) as a unique integer (or a tuple of integers, as it will be discussed), which ultimately will constitute a precise identifier of such rhythm. Put another way, given a rhythm R, expressed in one of the notational alternatives proposed by Toussaint (we will adopt in this study vector and necklace types), there is a particular

---

[9]Depending on the musical context in question, any durational value can be considered as rhythmic unit. By default, we propose the sixteenth note for this role.

[10]It is important to add that this model for drum-groove encoding is a simple, initial version that to be be replaced by a more complex one, whose description is presented along the next section.

integer M (standing for "M... code")[11] that will represent univocally R. Conversely, considering a given integer as M, playing backwards the same process, we can obtain algebraic and geometric representations of the related rhythm R. The present section describes the algorithms and processes employed for the accomplishment of both tasks. In addition, practical illustration of their potential application in analysis and composition will be provided.

i.   Isolated rhythms

Aiming at a more clear and gradual explanation of the structure of the algorithms, let us start with a simple rhythm played by a single drum-kit piece, say, the snare. As shown in Figure 7, the timeline in question consists of four onsets in a groove of sixteenth pulses. In this context, the eighth note is the lesser durational value and, therefore, it is the IU to be considered along the process.

| 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
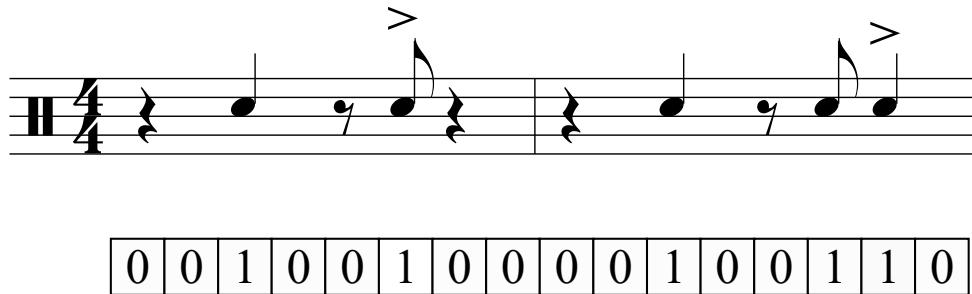
Figure 7: *Example of an isolated rhythm written for snare drums, in musical and binary notation.*

Firstly, the rhythm must be transcribed from musical to binary (positional) notation, with 1s corresponding to onset points, as shown beneath the score. The conventions adopted with respect to dynamic levels (introduced in the previous section) are incorporated into the positional vector, according to the following instructions: (a) replace 1 by 2 in the case of a ghost-note onset; (b) replace 1 by 3 in the case of an accented onset (and do nothing in normal onsets). Thus, there are two substitutions in our example (Figure 8).

	0	0	1	0	0	1	0	0	0	0	1	0	0	1	1	0
W	0	0	1	0	0	3	0	0	0	0	1	0	0	1	3	0
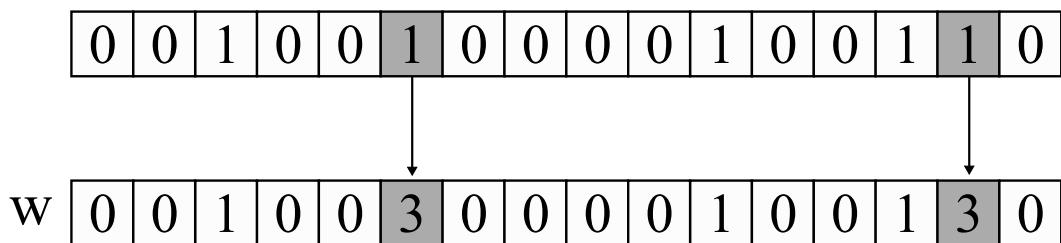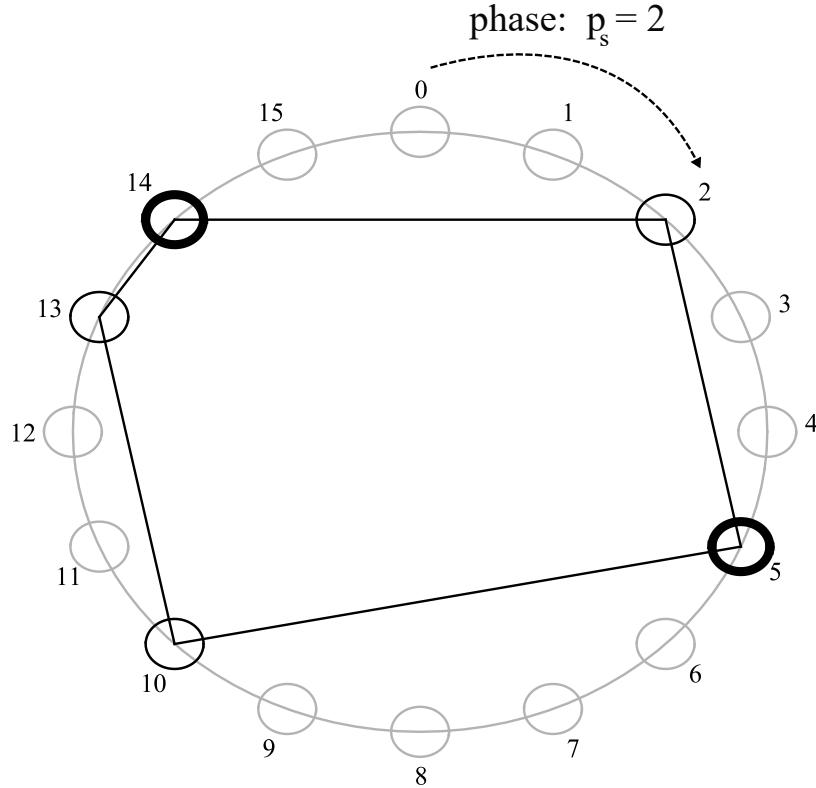
Figure 8: *Formation of vector w, using as example the rhythm of Figure 7.*

Let us refer to this latter structure as the *weighted vector* w, which ultimately represents a given rhythm R with respect to its onset points and micro-dynamic fluctuations. Vector w is then inputted in the program *PlotRhythm*,[12] which produces the necklace representation of R as

---

[11]This designation refers to Dr. Carlos Mathias, the researcher who idealized the code, and one of the co-authors of this article.

[12]This and the all programs which will be mentioned in this article were implemented in Matlab language.

exemplified in Figure 9 (thick-line nodes correspond to the accented onsets). Nodes of normal width refer to normal onsets. Ghost-note onsets would be represented by dashed-line nodes.

$$w_s = <0010030000100130>$$
$$v_s = <35314>$$
$$d_s = <13113>$$

Figure 9: *Necklace representation of the rhythm of Figure 7.*

Before continuing with the encoding of R, there are some new concepts and conventions to be introduced. Define the *phase* (denoted by p) of a rhythm as the distance (measured in IUs) between the starting point of the cycle/timeline (always indicated by the zero "hour" in the necklace) and the position of the first onset. In the case of our example, the rhythm has phase $p_s = 2$. The *intervallic vector* v proposes a more compact representation through the depiction of the sequence of differences (measured in IUs or pulses) between the onsets.[13] In the present case, $v_s = <35314>$.[14] Define also the *dynamic vector* d as the sequence of dynamic levels in a one-to-one relation with the onsets of R ($d_s = <13113>$ in the exemplified case).

The algorithm that encodes R, called *TranslateRhythm*, uses as input vectors w, v, d, and phase p. It functions basically as follows: Firstly, it forms a sequence of prime bases P, whose length equals

---

[13]In other words, v denotes a sequence of IOIs of R.

[14]Observe that these differences are proportional to the lengths of the edges in the polygon of Figure 9. Notice also that it is necessary to include in v the interval between the last and the initial onsets of the rhythm (4 pulses in the example), closing the circuit, which is perfectly consistent with the cyclic nature of a timeline.

the cardinality of the timeline. Considering that, in our example, the snare pattern has 16 pulses, then P = <2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53>. After this, the members of P that correspond to non-zero entries of w are copied into another vector, P_R, as shown in Figure 10.

W	0	0	1	0	0	3	0	0	0	0	1	0	0	1	3	0
P_R	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53

Figure 10: *Formation of vector P_R, using as example the rhythm of Figure 7.*
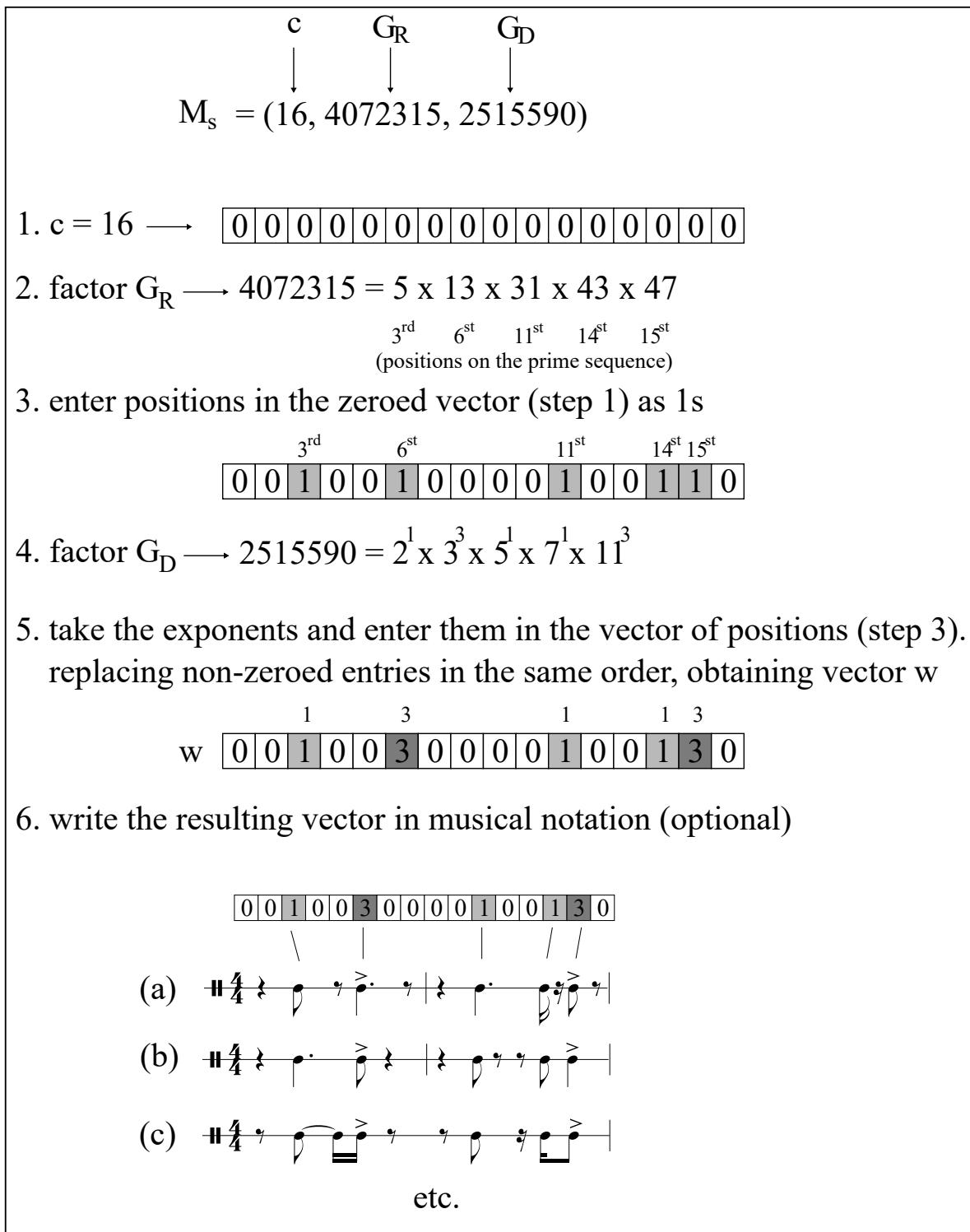
The integer resulting from the product of these prime numbers is the code referred to R. Let us label this integer as G_R. Considering the snare rhythm, we have $G_{R_S} = 5 \times 13 \times 31 \times 43 \times 47 = 4072315$.

Let now G_D be the integer that encodes the dynamic pattern of R. With the goal of minimizing the magnitude of this number, a different strategy is used in this case. Since there is a one-to-one mapping between the dynamic indications of vector d and the respective onsets of vector v, it is not necessary to locate their positions in the cycle (as it was done with the onsets). Therefore, it is enough to associate the n elements of d ($d_1, d_2, \ldots, d_n$) as exponents of the first n prime numbers ($p_1, p_2, \ldots, p_n$). Thus, G_D will be the product of these powered bases. For the snare timeline, we have: $G_{D_S} = 2^1 \times 3^3 \times 5^1 \times 7^1 \times 11^3 = 2515590$.

Finally, the code M of R, is represented by the triplet M = (c, G_R, G_D), where c is the length of the timeline (measured in IUs), G_R is the Gödel number related to R, and G_D is the Gödel number referred to the particular dynamic pattern of R. Thus, in our example, M_S = (16, 4072315, 2515590).

The reversal process, that is, for obtaining the original rhythm from the code, is considerably simple, through factoring both integers G_R and G_D, considering also the context (i.e., the cycle length). The decoding is operated by another algorithm, called *TranslateCode*, whose input is M. Taking the case of our example, the algorithm proceeds as depicted in Figure 11. However, as suggested in the sixth step of the algorithm, it is not possible to recover *exactly* the original musical notation, since the whole process deals with onsets (and IOIs), rather than properly durations.[15] Put another way, the three possible interpretations depicted at the bottom of Figure 11, or any other that obeys the positions of the onsets in the same 16-pulse context, will be considered perfectly equivalent with the respect to onset/dynamic patterns. Actually, since we are working with drums, the determination of the durations of the attacks is entirely immaterial. Given this, from now on, we will adopt as the preferential the simplest and clearest (in the drummer's perspective) possible notation for representing both the original and recovered rhythm. The next section examines a more complex situation, involving superimposition of rhythms, as it normally occurs in common drum grooves.

---

[15]Evidently, there are some special situations in which drum-set durations matter, as in snare rolls or in the reverberation of attacks on crash or splash cymbals. For simplicity, we will not consider such situations in the present model.

$$\begin{array}{ccc} \text{c} & G_R & G_D \\ \downarrow & \downarrow & \downarrow \end{array}$$

$$M_s = (16, 4072315, 2515590)$$

1. $c = 16 \longrightarrow$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

2. factor $G_R \longrightarrow 4072315 = 5 \times 13 \times 31 \times 43 \times 47$

$3^{rd} \quad 6^{st} \quad 11^{st} \quad 14^{st} \quad 15^{st}$
(positions on the prime sequence)

3. enter positions in the zeroed vector (step 1) as 1s

$3^{rd} \qquad 6^{st} \qquad\qquad 11^{st} \qquad 14^{st} \; 15^{st}$

| 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |

4. factor $G_D \longrightarrow 2515590 = 2^1 \times 3^3 \times 5^1 \times 7^1 \times 11^3$

5. take the exponents and enter them in the vector of positions (step 3). replacing non-zeroed entries in the same order, obtaining vector w

$\qquad\qquad 1 \qquad\quad 3 \qquad\qquad\quad 1 \qquad\quad 1 \; 3$

w | 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 3 | 0 |

6. write the resulting vector in musical notation (optional)

| 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 3 | 0 |

(a)

(b)

(c)

etc.

Figure 11: *Basic algorithm of TranslateCode for recovering an isolated rhythm R from a M code.*
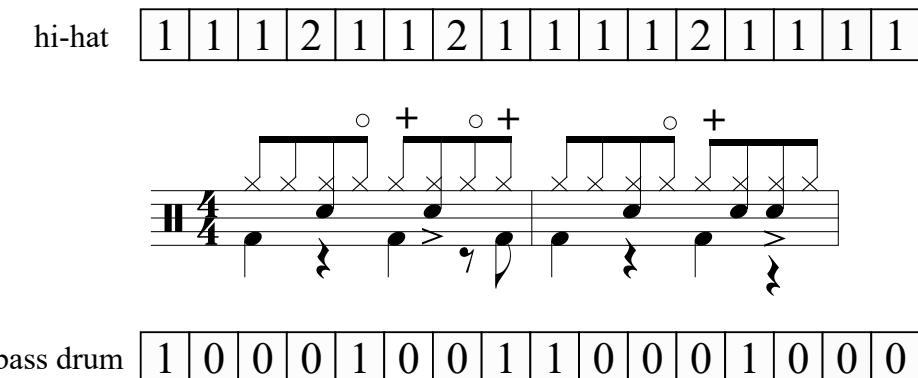
ii.   Drum grooves

Speculating about what was suggested in the third section of this article, it is possible to consider other additional "dimensions" for the onsets of a rhythm, besides dynamics. Information concerning timbre, especially, is a very useful feature to be added with respect to drums. Provided that clear conventions are previously established by the analyst, even complex drum patterns can be encoded with this technology. Consider as an illustration of this possibility the following conventions (Table 1). They can freely combine dynamic and timbral information (as in the examples) in order to adapt to genre (heavy rock, blues, bossa, salsa, etc.) or other contextual conditions.

Table 1: *Two possible convention charts for drum-set encoding (shaded cells refer to timbral features).*

streams	1	1	3	4
		(case a)		
1	bass drum	accented onset	–	–
2	snare drum	ghost note	accented onset	rim shot
3	closed hi-hat	opened hi-hat	–	–
		(case b)		
1	bass drum	accented onset	–	–
2	snare drum	low tom	mid tom	high tom
3	closed hi-hat	opened hi-hat	ride cymbal	attack cymbal

Let us adopt the simpler case (a) of Table 1 for addressing the superimposition of rhythms in a three-stream groove. Figure 12 adds bass-drum and hi-hat lines to the snare pattern of the previous section, forming a rock-like simple groove. The respective positional vectors can then be accordingly written. For visual clarity, no accents were added to the bass-drum rhythm.
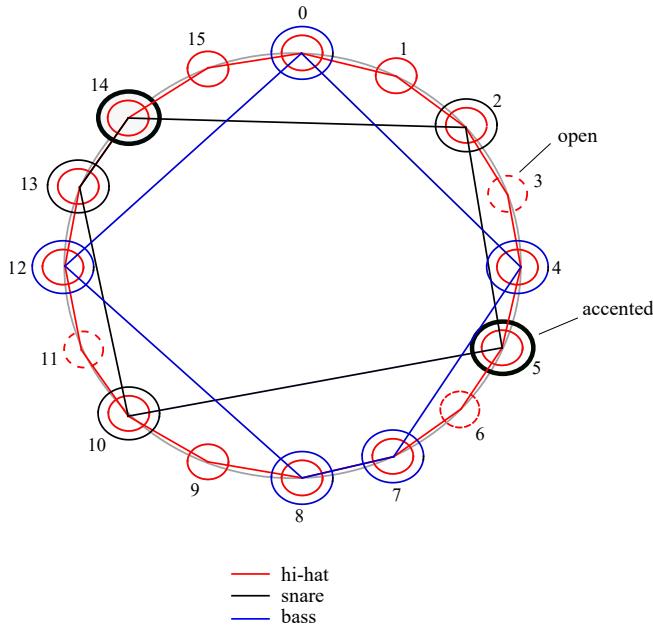
Figure 12: *Drum groove obtained by superimposition of bass-drum and hi-hat rhythms to the snare pattern of Figure 7.*

Since the three weight vectors (w_b, w_s, w_h) share, evidently, the same cycle of 16 pulses, they can be reunited in a unique three-row matrix (denoted as T), reproducing from top to the bottom the order of the pieces in the score:[16]

---

[16]Hoesl and Senn [14, p. 7] also use a 3-row matrix to notate onset positions of the drum pieces. However, since they are interested in syncopation, their matrix depicts rhythmic information related only by snare and bass drums (rows 1 and 2).

$$T = \begin{vmatrix} 1 & 1 & 1 & 2 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 3 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{vmatrix}$$

Matrix T is then input in the program *PlotRhythm*, which returns the graphic representation depicted in Figure 13.

Figure 13: *Necklace representation of the drum groove of Figure 12.*

Due to the increase of complexity, colors are used for discriminating the pieces' rhythms. As in the simple mode, the number of edges in each polygon represents the number of onsets of the rhythms. Concentric, superimposed circles indicate coincidence of onsets involving two pieces (in the exemplified groove there are no triple coincidence). The dashed-line red nodes refer to attacks on the opened hi-hat.

The next step proceeds as it was done with the snare line. The individual weight vectors of bass drum and hi-hat are input, one at a time, in the *TranslateRhythm* algorithm. In the case of the bass drum, there is nothing new: the algorithm returns $M_b = (16, 394174, 1)$. Observe that $G_D = 1$, which is consistent with the fact that the bass-drum rhythm is completely flat with respect to the dynamic level. On the other hand, the calculation of M_h gives rise to a problematic issue. Because its rhythm articulates all the pulses of the cycle, this implies a very huge value for GR (approximately 3.26×10^{19}). Apart from the awkwardness of representing an integer of such magnitude in the code, the real greatest difficult in dealing with it lies actually in the decoding process, due to the computational limitations for factoring mammoth integers like this.[17]

In the following, this paper will share some strategies on how to reduce the potential size of the integers involved within the structure of code M (for computational purposes), specifically for
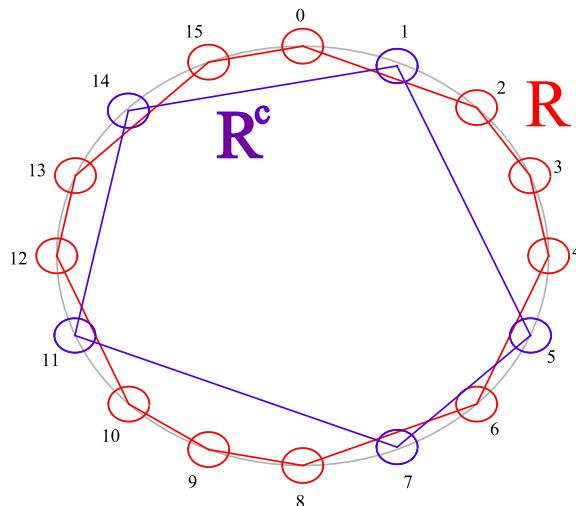
The third row is reserved for what they call *metric weight*, associated with distinct values of syncopation, from 0 to -3 (it is important to state, as it will be clear ahead, that this meaning for "weight" is completely different from which is here adopted.

[17]Factoring becomes complicated with integers greater than 10^{15}.

the cases of rhythms that contain a lot of information. The presentation of those strategies could suggest to the reader that the intent behind this encoding technique is to generate integers able to uniquely represent a rhythm. If that were simply the case, the coordinates of the informational vectors themselves would easily fulfill that wish. The hi-hat line of Figure 12, for instance, could be naively encoded by <1112112111121111>. But that would only provide a *source of representation* that is null in terms of providing arithmetical feedbacks useful to any sort of analysis. The main concern of the encoding technique here presented is to uniquely *link* an arbitrary musical representation to numerical informational vectors, *in such a way that the analysis of rhythmic attributes can be carried out within a robust arithmetical framework*, in the sense previously described. Rhythmic superimpositions/decompositions, displacements, dynamic changes and even modulations can be easily achieved as results of the use of simple arithmetic´s operations such as multiplication and division. Metaphorically, it can be said that Toussaint´s necklace notation is to geometry as the encoding technique here presented is to arithmetic.

The first strategy considered was to "break" G_R into three factors, such that $G_R = g_1 \times g_2 \times g_3$.[18] Although this solves perfectly the problem, it is an unsatisfactory solution, since it would force us to restructure code M, changing it from a triple to a quintuple, in order to eventually accommodate two new entries. A quite more elegant alternative is to consider the idea of *complement of a rhythm* R, denoted as R^R. This corresponds to the set of onsets that are not articulated by the rhythm, in the context of the cycle it is inserted in. In the case of our hi-hat example, since all pulses are articulated, its complement R^R corresponds to no articulation at all, being arithmetically represented by the product of the sixteen primes, all of them powered by 0, which equals 1. Evidently, the use of R^R is destined not only to express, so to speak, all-pulse rhythms. As a matter of fact, a rhythm with a number of onsets greater than half of the number of pulses can be "negatively" represented by its complement.

As an illustration of this point, consider, for instance, a hypothetical hi-hat variation of the groove, with an intermittent onset pattern denoted by the weight vector w = <1011101011101101>. Then, its complementary weight vector w_c will be expressed as w_c = <0100010100010010>. Figure 14 provides a necklace representation of both rhythms.

Figure 14: *Necklace representation of a hypothetical rhythm R and its complement R_c.*

---

[18]In the present case, the values are: g_1 = 2310; g_2 = 2800733; and g_3 = 5037203051.

Now, we calculate G_R for both rhythms and compare the results. As done in ordinary cases, G_R is the product of prime numbers corresponding to non-zeroed entries of the weight vector. The lowest integer calculated will then be chosen for encoding the rhythm. As shown in Table 2 the comparison of both alternatives favors the choice of the complement as the representative of the rhythm.

Table 2: *Comparison between R and R_c referring to the case of Figure 14.*

Pulses	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	G_R
Primes	2	3	6	7	11	13	17	19	23	29	31	37	41	43	47	53	
w	1	0	1	1	1	0	1	0	1	1	1	0	1	1	0	1	2.53×10^{13}
w^c	0	1	0	0	0	1	0	1	0	0	0	1	0	0	1	0	**1288599**

The indication in the code M that a rhythm is represented by its complement is accomplished through a very simple artifice: the multiplication of the respective G_R by -1.

On the other hand, if the complement is chosen for representing a given rhythm, its G_D shall be calculated in a different way from the ordinary case. Actually, the algorithm is very simple and can be described in two steps: (a) subtract 1 from the entries of original vector w, an action that we will represent by w -– 1; (b) power the primes with the resulting vector (as it was done in the calculating of G_R) and calculate their product. In the case of the hypothetical example of Figure 14, we have then the subtraction $w^c - 1 = {<}0000000000000000{>}$, implying then $G_D = 1$ (a not surprising result, since w^c is perfectly flat). We can then encode this rhythm as M = (16, -1288599, 1).
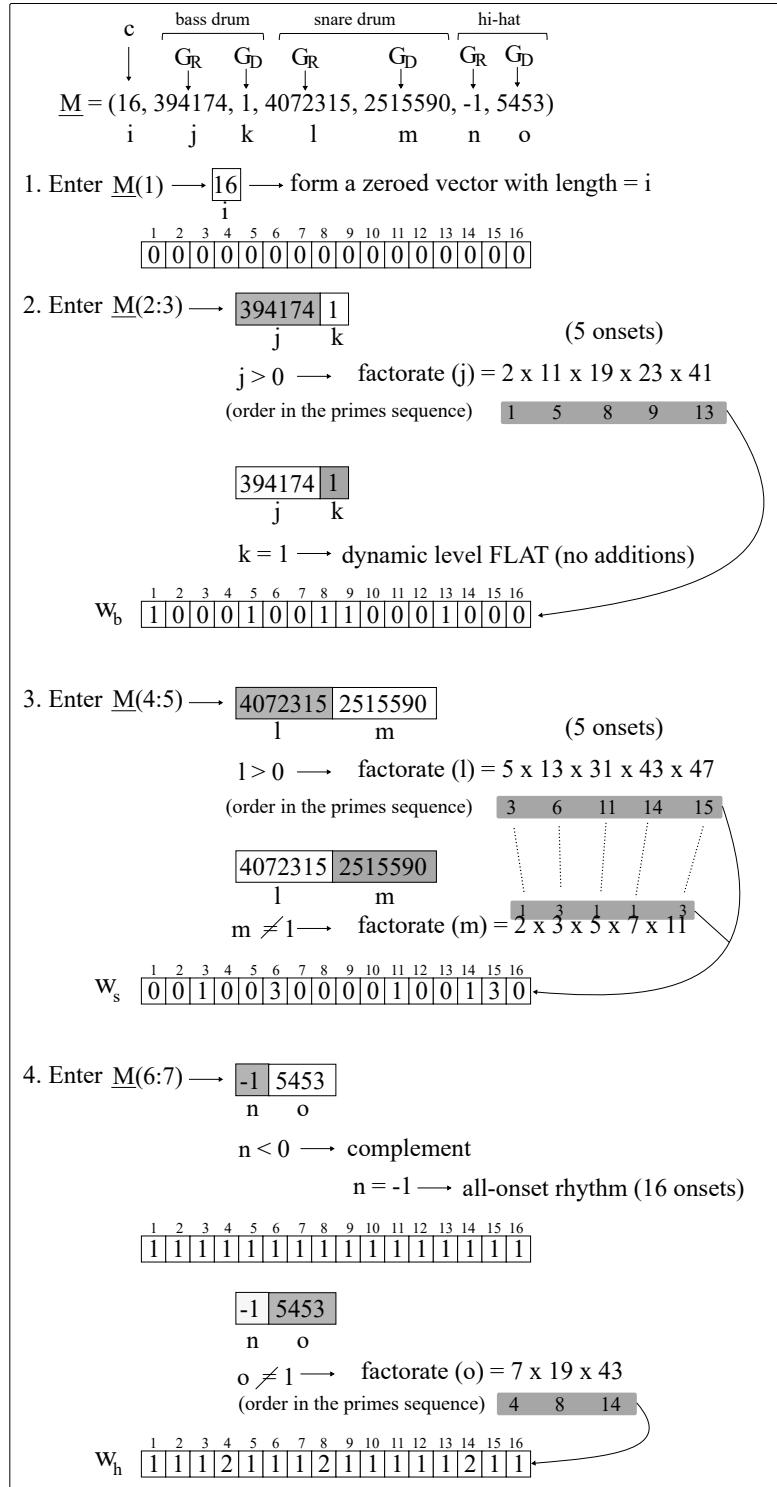
We can now return to the original hi-hat pattern and encode it properly. Recall that its weight vector (that includes some opened onsets) is expressed as w = ${<}1112112111121111{>}$. Since the rhythm saturates the onset space, $G_R = -1$. For calculating G_D, we take w — 1, obtaining the exponent sequence ${<}0001001000010000{>}$. As we know, to power this vector to the primes sequence is the same as to pick the primes corresponding to the non-zero entries and multiply them. Therefore, the integer referred to in the weight vector (in complementary mode) is $G_D = 7 \times 19 \times 41 = 5453$. Given this, we can express the hi-hat code in the form $M_h = (16, -1, 5453)$.

Finally, the codes of the individual pieces can be combined into a unique structure (denoted as <u>M</u>) representing the complete groove. Ultimately, <u>M</u> will be expressed as the septuplet <u>M</u> = (c, M_b, M_s, M_h) = (c, $[G_R, G_D]_b$, $[G_R, G_D]_s$, $[[G_R, G_D]_h$). Considering our example, M = (16, 394174, 4072315, 2515590, -1, 5453).

The retrieval of a three-piece drum groove, given a code M, is achieved through the application of the same algorithm *TranslateCode*, used for isolated rhythms (see Figure 11). However, due to the compound structure of the groove-code, some new elements have to be incorporated in the definitive implementation of the algorithm. Firstly, the algorithm shall be able to go through M in a two-step window from the second entry on, in order to adequately process the three pairs of information, respectively related to any of the groove's streams. Furthermore, it is also necessary to provide the algorithm with the necessary conditions for dealing with eventual complementary rhythms, indicated by negative entries.

Before continuing, especially envisaging further improvement, let us assign variables to denote the seven entries of <u>M</u> (we arbitrarily chose for this task the sequence of the alphabet between letters *i* and *o*). Thus, from now on a generic compound code will be denoted as <u>M</u> = (i, j, k, l, m, n, o). Figure 15 illustrates the functioning of the definitive version of *TranslateCode* taking the code

<u>M</u> as input.

c   bass drum   snare drum   hi-hat
        G_R G_D G_R G_D G_R G_D
<u>M</u> = (16, 394174, 1, 4072315, 2515590, -1, 5453)
        i   j   k   l   m   n   o

1. Enter <u>M</u>(1) ⟶ 16 ⟶ form a zeroed vector with length = i
                      i

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2. Enter <u>M</u>(2:3) ⟶ 394174 1   (5 onsets)
                        j   k

j > 0 ⟶ factorate (j) = 2 x 11 x 19 x 23 x 41

(order in the primes sequence)   1   5   8   9   13

394174 1
   j   k

k = 1 ⟶ dynamic level FLAT (no additions)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
W_b | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |

3. Enter <u>M</u>(4:5) ⟶ 4072315 2515590   (5 onsets)
                        l   m

l > 0 ⟶ factorate (l) = 5 x 13 x 31 x 43 x 47

(order in the primes sequence)   3   6   11   14   15

4072315 2515590
   l   m

m ≠ 1 ⟶ factorate (m) = 2 x 3 x 5 x 7 x 11
                         1   3   1   1   3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
W_s | 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 3 | 0 |

4. Enter <u>M</u>(6:7) ⟶ -1 5453
                        n   o

n < 0 ⟶ complement

n = -1 ⟶ all-onset rhythm (16 onsets)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

-1 5453
 n   o

o ≠ 1 ⟶ factorate (o) = 7 x 19 x 43
(order in the primes sequence)   4   8   14

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
W_h | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |

Figure 15: *Basic algorithm of TranslateCode for recovering a three-stream groove rhythm R from a <u>M</u> code.*

V. Encoding rhythmic variations

Maybe the most attractive application of the encoding technology addresses the aspect of rhythmic variations, both with analytical and compositional purposes. For this, instead of dealing with the content of a given referential <u>M</u> expressed with actual numbers, we need only to use its algebraic version (i.e., denoted by variables i-o).[19]

Firstly, let us consider the entries associated uniquely with onset configurations (variables j, l, and n). There are three manners of transforming an onset pattern of a rhythm R:

1. By adding a new onset to R — this is equivalent to multiplying the variable in question by a prime number of the sequence that is not a factor of R;

2. By deleting an existent onset — this is equivalent to dividing the variable in question by a prime number of the sequence that is already a factor of R;

3. By replacing an existent onset by another one -– this is equivalent to multiplying the variable in question by a fraction such that the numerator is a new prime *and* the denominator is an already existing one, which is intended to be replaced.

Figure 16: *A possible variation of Figure 13.*

In the case of the additional information of a groove (dynamics or timbre, represented by variables k, m, and o), transformations are referred to a unique category, namely the change of a dynamic level or a timbre of a specific onset, which is achieved by multiplying or dividing the

---

[19]Notice that if we consider <u>M</u> algebraically, it is not necessary any more to replace an eventual very intense rhythm by its complement. In this manner, considering our example, variable *n* (that denotes the hi-hat onset configuration) represents the original (very huge) integer which corresponds to the product of the sixteen first prime numbers, rather than its complement (-1).

variable in question by a prime factor related to the original configuration (such procedure affects the magnitude of the exponents, by adding or subtracting units). Finally, the cycle itself (variable i) can also be transformed, by expansion or contraction, by multiplication and/or division.[20]

As an illustration of how these procedures can be applied in an analytical situation, consider the variation of our original groove shown in Figure 16. As above stated, let the groove be expressed as $\underline{M}$ = (i, j, k, l, m, n, o). Given this, it is possible analyze the variant, whose code we will denote as $\underline{M}_V$, in terms of algebraic transformations of the content of $\underline{M}$.

For clarity, Table 3 details the transformational process, considering the individual vector differences between both compound rhythms. Bold-face numbers indicate affected onsets or dynamic/timbral elements.

Table 3: *Transformations involved in the variation of Figure 16*

		groove	variation	divide by	multiply by
cycle	i	16 (scalar)	16 (scalar)	1	1
bass	j	**1000100110001000**	**0000**100**010001001**	19	53
drum	k	1111**1**	1111**3**	1	11^2
snare	l	0010**0100000100110**	0010**001000100100**	13×47	17
drum	m	131**13**	131**30**	11^3	7^2
hi-hat	n	111111111111**1111**	1111111111111**100**	47×53	1
	o	111**2**112112**11**121**111**	11**21**112112**111100**	$7 \times 37 \times 47 \times 53$	5×23

This allows us to propose the following algebraic representation of the variation:

$$\underline{M}_V = (i, \tfrac{19}{53}j, \tfrac{1}{11^2}k, \tfrac{611}{17}l, \tfrac{11^3}{7^2}m, 2491n, \tfrac{645169}{115}o)$$

Let us now examine another very useful application. Consider an archetypal samba drum groove, which is characterized by a fixed bass-drum pattern (normally doubling rhythmically the bass line) over which cross-stick attacks and ride accents are added in almost random sixteenth-note onsets. In a sense, most of samba swing lies precisely in the low expectancy of these accentuation patterns. Normally, the decisions about where and when an onset shall be stressed depends on the drummer's intuition and, of course, expertise in playing this genre of music, in sum, of his/her improvisatory capacity in this kind of music. Now, suppose that we intend to model a samba-drum performance, as in the excerpt shown in Figure 17. Instead of treating the whole passage as a 40-pulse cycle, which would turn the underlying computation highly complex and costly, an alternative, much simpler strategy is to see it as a set of nine variations of a very small basic "theme", at the first beat of the groove. According to this approach, it can be seen as the generator of the nine subsequent transformed versions of itself. Notice that now generator and variants are quite small cycles (four pulses), which implies few and small primes (2, 3, 5, and 7), and considerably simple codes. Observe also that five out seven structures are kept invariant along the groove: cycle length (i), bass-drum onset and dynamic/timbral configurations (j-k), snare dynamic/timbral pattern (m), and ride onset positions (n). As a matter of fact, cross-stick positions (l) and ride accentuations (o) are the only parameters subject to variation.

Table 4 depicts the algebraic structure of the nine variations in relation to that of the generator (invariant elements are let shaded). Observe that, although variation 4 incidentally replicates the

---

[20]For example, the transformation of a 16-pulse into a 12-pulse cycle is achieved through the multiplication of i by $\tfrac{3}{4}$

conventions	onset	accent.
ride	1	2
cross stick	1	
bass drum	1	

generator

$\underline{M} = (i, j, k, l, m, n, o) = (4, 14, 1, 5, 1, 210, 1)$
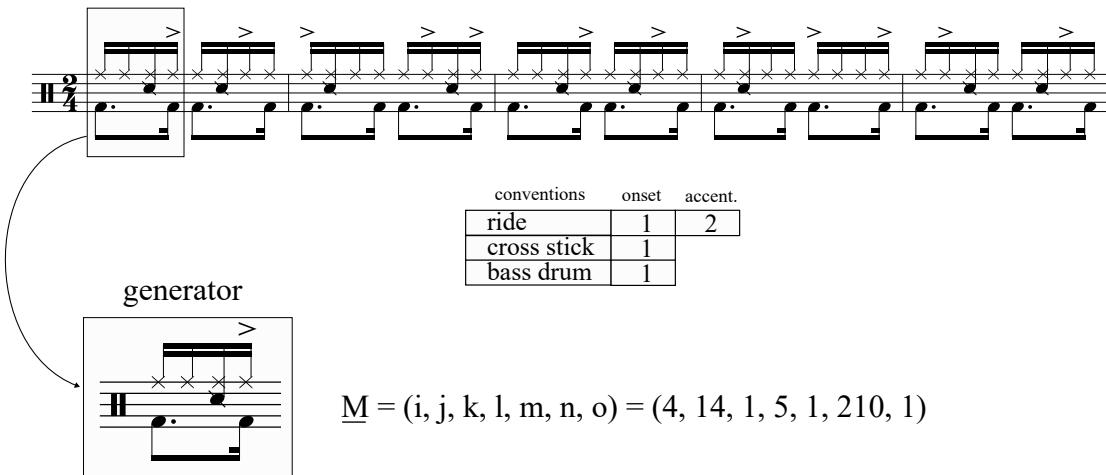
Figure 17: *Samba variations by transformations applied to a basic generator.*

generator (which is perfectly normal in real-music situation), there is a considerable variety of patterns among the remaining variants, due to combinatorial possibilities.

Table 4: *Transformations involved in the variation of Figure 16*

m. 1	generator	i	j	k	l	m	n	o
	variant 1	1	1	1	$\frac{3}{5}$1	1	1	$\frac{5}{7}$o
m. 2	variant 2	1	1	1	$\frac{3}{5}$1	1	1	$\frac{2}{7}$o
	variant 3	1	1	1	1	1	1	3o
m. 3	variant 4	1	1	1	1	1	1	1
	variant 5	1	1	1	$\frac{3}{5}$1	1	1	$\frac{5}{7}$o
m. 4	variant 6	1	1	1	$\frac{3}{5}$1	1	1	$\frac{3}{7}$o
	variant 5	1	1	1	$\frac{1}{5}$1	1	1	2o
m. 5	variant 7	1	1	1	1	1	1	$\frac{3}{7}$o
	variant 8	1	1	1	$\frac{3}{5}$1	1	1	$\frac{5}{7}$o

Lastly, we can express the whole passage as a 10×7 matrix, in which any row represents a beat (the generator and the nine variants) and the columns correspond to the elements of the respective M codes.

$$
\begin{vmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 3/5 & 1 & 1 & 5/7 \\
1 & 1 & 1 & 3/5 & 1 & 1 & 2/7 \\
1 & 1 & 1 & 1 & 1 & 1 & 3 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 3/5 & 1 & 1 & 5/7 \\
1 & 1 & 1 & 3/5 & 1 & 1 & 3/7 \\
1 & 1 & 1 & 1/5 & 1 & 1 & 2 \\
1 & 1 & 1 & 1 & 1 & 1 & 3/7 \\
1 & 1 & 1 & 3/5 & 1 & 1 & 5/7
\end{vmatrix}
$$

VI. Conclusion

This article introduced an original theoretical-methodological proposal intended to address drum-set rhythms in a formal manner. The adaptation and combination of concepts, formulations, and conventions coined by Gauss, Gödel, and Toussaint that resulted in the technology which is here described allows complex three-stream drum grooves to be represented both geometrically (as superimposed polygons) and algebraically (as tuples of codes, vectors, and matrices). By exploring the inherent property of prime numbers (expressed in the FTA), some simple algorithms are used for plotting and encoding rhythms of the individual pieces of a groove, as well as for the reversal process, retrieving the original timelines by factoring their codes. Our proposal includes also the possibility of extending the methodology to encompass auxiliary parameters, as dynamics and timbre, turning the model more flexible and comprehensive. Besides the primary advantages of encoding relatively complex musical information as retrievable tuples (compactness, especially), the practical applications that are here envisaged – both in analysis and composition – open a potentially wide avenue to be explored in future work. In this sense, the idea of using this technology for creating subtle variations of a groove-seed (expressed as an abstract generator), as it was done in the samba example, seems to be the most promising.

References

[1] Almada, C. 2017. Gödel-vector and Gödel-address as Tools for Genealogical Determination of Genetically-Produced Musical Variants. In: Pareyon, D.; Pina-Romero, S., Agustín-Aquino, G.; Lluis-Plueba, E. (eds.) *The Musical-Mathematical Mind*: Patterns and Transformations. Cham (Switzerland): Springer Verlag, pp. 9–16.

[2] Cohn, R. 1992a. The Dramatization of Hypermetric Conflicts in the Scherzo of Beethoven's Ninth Symphony. *19th-Century Music*, v. 15, n. 3, pp. 188—206.

[3] Cohn, R. 1992b. Metric and Hypermetric Dissonance in the Menuetto of Mozart's G-Minor Symphony, K.550. *Intégral* v. 6, pp. 1—33.

[4] Cohn, R. 2001. Complex Hemiolas, Ski-Hill Graphs and Metric Spaces. *Music Analysis*, v. 20, n. 3, pp. 295—326.

[5] Cohn, R. 2016. A Platonic Model for Funky Rhythms. *Music Theory Online*, v. 22, n. 2. Available at: https://mtosmt.org/issues/mto.16.22.2/mto.16.22.2.cohn.php.

[6] Figueiredo, A.; Wolf, P. S. A. 2009. Assortative Pairing and Life History Strategy – A Cross-Cultural Study. *Human Nature*, v. 20, pp. 317–330.

[7] Gotham, M. 2005. Attractor Tempos for Metrical Structures. *Journal of Mathematics and Music*, v. 9, n. 1, pp. 23-–44.

[8] Gómez-Martín, F.; Taslakian, P.; Toussaint, G. 2009. Interlocking and Euclidean Rhythms. *Journal of Mathematics and Music*, v. 3, n. 1, pp. 15-–30.

[9] Guastavino, C.; Gómez F.; Toussaint, G.; Marandola, F.; Gómez, E. 2009. Measuring Similarity between Flamenco Rhythmic Patterns. *Journal of New Music Research*, v. 38, n. 2, pp. 129-–38.

[10] Guerra, S. 2018. *Expanded Meter and Hemiola in Baden Powell's Samba-Jazz*. Dissertation (PhD in Music), Yale University.

[11] Guerra, S. 2018. Toward a Theory of Structuring Rhythm in Improvisation in Timeline-Based Musics. *MusMat*: Brazilian Journal of Music and Mathematics, v. 3, n. 1, pp. 44-–60.

[12] Hardesty, J. 2016. A Self-Similar Map of Rhythmic Components. *Journal of Mathematics and Music*, v. 10, n. 1, pp. 36-–58.

[13] Herrlein, J. 2018. *Das alturas ao ritmo*: teoria dos conjuntos rítmicos como ferramenta composicional. Dissertation (PhD in Music), Federal University of Rio Grande do Sul.

[14] Hoesl, F.; Senn, O. 2018. Modelling Perceived Syncopation in Popular Music Drum Patterns: A Preliminary Study. *Music & Science*, v. 1, pp. 1-–15.

[15] London, J. 2002. Some Non-Isomorphisms between Pitch and Time. *Journal of Music Theory* v. 46, n. 1–2, pp. 127-–51.

[16] Murphy, S. 2016. Cohn's Platonic Model and the Regular Irregularities of Recent Popular Multimedia. *Music Theory Online*, v. 22, n. 3. Available at: https://mtosmt.org/issues/mto.16.22.3/mto.16.22.3.murphy.html

[17] Nagel, E.; Newman, J. 2001. *Gödel ́s Proof*. New York: New York University Press.

[18] Temperley, D. 2001. *The Cognition of Basic Musical Structures*. Cambridge: The MIT Press.

[19] Toussaint, G. 2013. *The Geometry of Musical Rhythm*: What Makes a "Good" Rhythm Good? Boca Raton: CRC Press.

Feathered Beams

Paul Lombardi

University of South Dakota

paul.lombardi@usd.edu

Orcid: 0000-0001-8220-6376

Abstract: In music, durations are quantized to subdivisions of time in the form of fractions of the inverse powers of 2 (e.g., $1/2^0$, $1/2^1$, $1/2^2$, etc.). All durations that are not involved in tuplets can be represented by sums of these fractions. The gradual transition from one note duration to another through a specified number of intermediate note values requires an accelerando/ritardando beam (i.e., feathered beam). This notation, however, does not indicate exactly how the gradual transition through the intermediate note values is to occur. The various details may be so contradictory that feathered beams may be impossible to realize. Thus, the notation is inherently indeterminate, although it is not often regarded as such. This paper examines these concepts and combines rhythmic nomenclature with a graphing system to deconstruct feathered beams using examples from George Crumb's Night Music I.

Keywords: Feathered Beams. Accelerando/ritardando Beam. Duration. Rhythm. Graph. George Crumb.

1. Feathered Beams

Accelerando/ritardando beams notate the gradual transition from one note duration to another. They are nicknamed feathered beams because in music notation the beams flare out like feathers (see Figures 4, 5, and 16 below for examples from the standard repertoire). This notation specifies note values at the beginning and ending of the beamed group, a gradual transition between those note values through a certain number of intermediate values, and in some cases, a precisely specified total duration. This notation, however, does not indicate exactly how to execute the gradual transition through the intermediate note values. This lack of information means that the notation is inherently indeterminate. However, the specificity of the notation disguises this indeterminacy. Furthermore, in some cases, the various specific requirements of a feathered beam group may be so contradictory that they are impossible to realize in practice. Although this paper examines feathered beams precisely, some composers do not consider feathered beams to be precise, but rather merely mean, for example, from something fast to something slow.

This paper combines rhythmic nomenclature (section II) with a graphing system (section III) to deconstruct feathered beams, and to compare the information inherent in the notation, the composer's intention, and the practical aspects of execution. Examples from George Crumb's *Night Music I* are plotted to explain the feathered beam technique and the theory behind it. First, this paper examines the more general case for when the total duration of the beamed group is not

specified, and some questions are considered with respect to performance practice and cognition (section IV). Next, a more constrained case is examined, one in which the total duration is specified (section V). In this scenario, however, some problems arise, most notably the impossibility of accurately realizing the notation. Finally, feathered beams are generalized to calculate the full duration of any linear or quadratic feathered beam (section VI).

This paper's focus is intentionally limited in order to demonstrate particular problems with feathered beams, mostly in the abstract. Much work remains to be done to fully understand how feathered beams are executed and perceived within a greater context of the music and phrase, but some ramifications are considered with respect to research that has been done on traditional rhythmic scenarios. Studying the examples in isolation reveals the limitations and complexities of this notation.

2. Definitions

2.1. Duration

In Western musical notation, the unit of duration is the whole note (w) and durations are given as subdivisions of it. The system is organized around note values based on the inverse powers of 2 where $(\frac{1}{2^p}), p \in \mathbb{Z}^+$:

$$\text{𝅝} = (\tfrac{1}{2^0})w = w$$

$$\text{𝅗𝅥} = (\tfrac{1}{2^1})w = \tfrac{w}{2}$$

$$\text{𝅘𝅥} = (\tfrac{1}{2^2})w = \tfrac{w}{4}$$

$$\text{𝅘𝅥𝅮} = (\tfrac{1}{2^3})w = \tfrac{w}{8}$$

$$\text{𝅘𝅥𝅯} = (\tfrac{1}{2^4})w = \tfrac{w}{16}$$

$$\text{𝅘𝅥𝅰} = (\tfrac{1}{2^5})w = \tfrac{w}{32}$$

etc.

Durations are subject to the normal operative rules of addition, subtraction, multiplication, and division; the commutative and transitive properties hold true with duration operations.

2.2. Augmentation Dots

A single augmentation dot adds 50% to the duration of the note that it follows:

$$\text{𝅗𝅥.} = \text{𝅗𝅥} \smile \text{𝅘𝅥} = \tfrac{w}{2} + \tfrac{w}{4} = \tfrac{3w}{4}$$

$$\text{𝅘𝅥.} = \text{𝅘𝅥} \smile \text{𝅘𝅥𝅮} = \tfrac{w}{4} + \tfrac{w}{8} = \tfrac{3w}{8}$$

$$\text{♪.} = \text{♪⌣♪} = \frac{w}{8} + \frac{w}{16} = \frac{3w}{16}$$

etc.

Consecutive augmentation dots are given by the summation:

$$\sum_{i=p}^{d+p} \frac{w}{2^i} = \frac{(2^{d+1}-1)w}{2^{d+p}}, \tag{1}$$

where p is the exponent of the power of 2 of the corresponding note without the augmentation dots with respect to w (e.g., whole note when $p = 0$, half note when $p = 1$, quarter note when $p = 2$, eighth note when $p = 3$, etc.) and d equals the number of augmentation dots ($p, d \in \mathbb{Z}^+$).

$$\text{♩.} = \text{♩⌣♪} = \frac{(2^{1+1}-1)w}{2^{1+2}} = \frac{3w}{8}$$

$$\text{♩..} = \text{♩⌣♪⌣♪} = \frac{(2^{2+1}-1)w}{2^{2+2}} = \frac{7w}{16}$$

$$\text{♩...} = \text{♩⌣♪⌣♪⌣♪} = \frac{(2^{3+1}-1)w}{2^{3+2}} = \frac{15w}{32}$$

etc.

When continually applying augmentation dots, the duration limits to 2 times the duration of the note they are on.

2.3. Tuplets
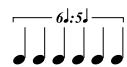
The note values presented so far in this paper can indicate any duration that is evenly divisible by $\frac{1}{2^p}$. In tuplets, however, note values are compressed according to ratios, such as in a traditional triplet of eighth notes that has 3 eighth notes compressed into the duration of 2 eighth notes. The 2 eighth notes in the example that follows are not compressed; their ratio is $2♪ : 2♪ = 1$. Below that, the 3 eighth notes are compressed into the duration of 2 eighth notes, yielding a ratio of $3♪ : 2♪ = \frac{3}{2}$. Each of these individual eighth notes have the duration $(\frac{2}{3})(\frac{w}{8}) = \frac{w}{12}$, which is one-third of a quarter note (i.e., $(w/4)/3$). (Traditional triplets are simply indicated by 3 in a bracket or next to a beam, but this paper is more specific in that it uses the corresponding note values along with their ratio (e.g., $3♪ : 2♪$) which is especially applicable to more complicated tuplets.)

$$\frac{w}{8} + \frac{w}{8} = \frac{w}{4} \qquad\qquad 2♪ : 2♪ = 1$$

$$\left(\frac{2}{3}\right)\left(\frac{w}{8} + \frac{w}{8} + \frac{w}{8}\right) = \frac{w}{4} \quad 3♪ : 2♪ = \frac{3}{2}$$

Some more examples follow:

$$\left(\frac{4}{5}\right)\left(\frac{5w}{16}\right) = \frac{w}{4} \quad 5♪ : 4♪ \text{ or } 5♪ : ♩ = \frac{5}{4}$$

67
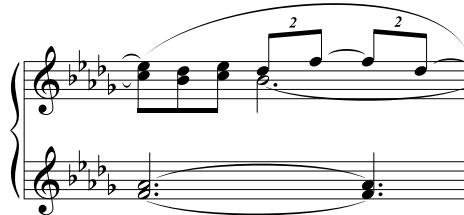
$$\left(\frac{5}{6}\right)\left(\frac{6w}{4}\right) = \frac{5w}{4} \qquad 6\,\flat : 5\,\flat = \frac{6}{5}$$

In the above examples and throughout this paper, tuplets have ratios greater than 1 and less than 2 indicating that the note durations have been compressed; however, some composers find it acceptable to use tuplets with ratios of less than 1 augmenting the note durations, such as in Debussy's *Suite bergamasque*, 3. "Clair de Lune", m. 3 (Figure 1).

Figure 1: *Debussy,* Suite bergamasque, *3. "Clair de Lune", m. 3.*

Here, the compound meter has dotted quarter-note beats, and the simple subdivisions of the beat are notated with eighth-note duplets. The ratio is $2\flat:3\flat= \frac{2}{3}$, which, according to common notation practice described in many style manuals ([5]), is usually not preferred because $\frac{2}{3} < 1$. Preferably, this example should use 2 quarter notes instead of 2 eighth notes in this duplet, because it makes the ratio $2\flat:3\flat= 4\flat:3\flat= \frac{4}{3}$, which is $1 < \frac{4}{3} < 2$.

Situations similar to this duplet can be visually misleading in some circumstances. Usually there are simpler ways of notating tuplets and partial tuplets. For instance, the above duplet can be notated more simply as 2 dotted-eighth notes.
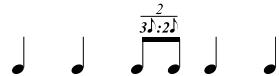
2.4. Partial tuplets

The individual notes within tuplets can represent many rational durations. For example, the first two notes in a traditional triplet of eighth notes equal the duration $\left(\frac{2}{3}\right)\left(\frac{w}{4}\right)$. Consider a musical context where these two notes occur without the triplet's remaining eighth note. According to traditional musical notation, the remaining eighth note should be included in the score enclosed in brackets as shown in the following rhythm.[8]

The note in the brackets is not played and its value does not occur in the time of the music. In other words, if the quarter note gets the beat, then the last two quarter notes in the rhythm are

shifted a third of a beat forward in time. Only the first two-thirds of the triplet are to be played indicating a partial amount of the tuplet. In this paper, the following notation will be used instead, where the number in the numerator specifies the partial amount of the tuplet.

The tuplet is given as $\frac{z}{x:y}$, which indicates z parts of x notes in the span of y notes where z, x, and y are positive integers such that $x > y \geq z \geq 1$.[1] Partial tuplets can usually be avoided by using other notation that can express a similar or equivalent musical result. An equivalent rhythm to the one above, but with compound beats and without the partial tuplet, is given below.

Any rational duration even if it is not evenly divisible by $\frac{1}{2^p}$ can be notated with tuplets or partial tuplets. For example, although musically impractical, the duration $\frac{37w}{41}$ can be notated as follows:

$$= \left(\frac{37}{41}\right)\left(\frac{32w}{32}\right) = \frac{37w}{41} \qquad 41\flat : 37\flat = \frac{41}{37}$$

2.5. Irrational durations
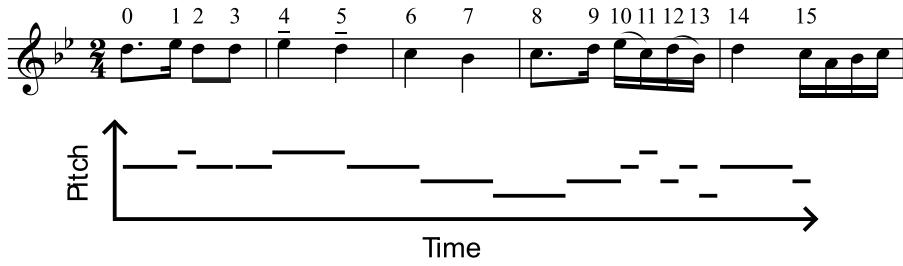
All the durations presented in this paper to this point are rational, but sometimes it is necessary to indicate irrational durations. Irrational durations cannot be represented as previously shown in this paper. So, for the purpose of analysis, they will be given as solid stemless noteheads with the durations in whole-note units written above them. Some examples are as follows:

$$\sqrt{2} \qquad\qquad \pi \qquad\qquad 1.367274...$$

3. Duration Graphs
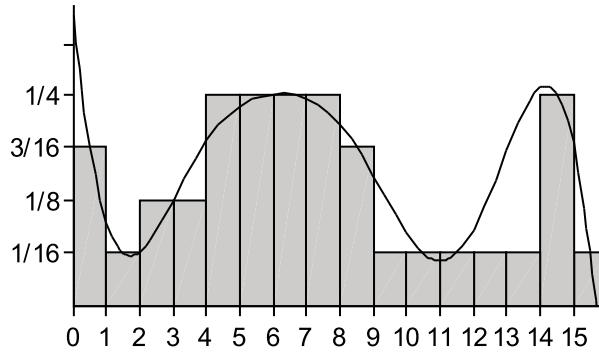
A musical score is like a graph where time is given on the horizontal axis and pitch is given on the vertical axis. Consider the first phrase of Brahms, Variations on a Theme by Haydn, Op. 56 (Figure 2).

---

[1]This notation may not be conducive to sight reading, but accurately describes the rhythmic phenomena and is useful for the purposes of analysis.

Figure 2: *Brahms,* Variations on a Theme by Haydn, *Op. 56, mm. 1–5.*

The notes are indexed beginning with zero. The horizontal bars below the staff are proportional to the durations of the notes. In Figure 3, all those horizontal bars are arranged vertically. The notes are given in temporal order on the horizontal axis according to their indices, and their durations are given on the vertical axis. This way, changes in duration can be plotted.(N.b., the curve fitting in the illustration merely shows that changes in duration can follow a curve.)

Figure 3: *Duration graph of Figure 2.*

In this paper, durations in feathered beams are given as a series of (i, r) coordinates, where i is the index and r is the duration ($r \in \mathbb{Q}$). For the Brahms example above, this is:

$$\left\langle \left(0, \frac{3w}{16}\right), \left(1, \frac{w}{16}\right), \cdots, \left(15, \frac{w}{16}\right) \right\rangle.$$

Changes in duration can be plotted as lines and curves on these graphs, and can be given with standard algebra. Lines are given in the form

$$f(x) = ax + b$$

and quadratic curves are given in the form

$$f(x) = ax^2 + bx + c,$$

where $a, b, c, x \in \mathbb{R}$.

When a series of durations does not correspond to an easily determined equation, a curve may be plotted through the coordinates using splines (piecewise polynomials that match their values and derivatives between the pieces). The first derivative of the lines and curves give the

rate of change, and the second derivative gives the acceleration.The integral of the lines and curves approximates the sum of the durations of all the notes. These concepts will be explained more fully in the following section.

4. George Crumb, *Night Music* I, "Notturno I", rehearsal 5

In this section of the paper, the following excerpt from the 1976 revised version of George Crumb's *Night Music I*, "Notturno I" is exploited to demonstrate the details of the note-duration graphing system. I chose this excerpt after searching extensively for examples of feathered beams. I found a plethora of examples where the feathered beams do not even closely fit within their corresponding spans of music and/or be approximately performed as notated. In light of these observations, examples in George Crumb's *Night Music I* best suit my needs to illustrate the points in this paper. In "Notturno I" at rehearsal 5, ten notes under a feathered beam group gradually transition from a thirty-second to eighth note [2] (Figure 4).

Figure 4: *George Crumb,* Night Music I, *"Notturno I", rehearsal 5.*

The original 1967 version [1] of the score incorporated various improvisatory elements, most of which are notated more precisely in the 1976 revision. Throughout the revised version, there are ritardando or accelerando indications written above the feathered beams that are not in the original version. Other examples, such as the one below from "Notturno III",suggest that the feathered beams should not be further exaggerated by tempo changes, but that the rit. and accel. tempo indications merely reinforce the rhythmic gestures (Figure 5). Thus, the rit. indication in the above example is taken as a redundant instruction.
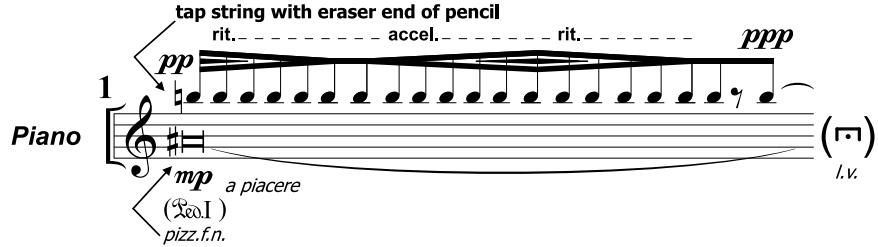
Figure 5: *George Crumb,* Night Music I, *"Notturno III".*

The 10-note feathered beam from the "Notturno I" example is in an unmetered context and there are no simultaneous figures to indicate that it spans a specific duration. The ten notes in the feathered beam are identified in temporal order as $n_0, n_1, \cdots, n_9$. The difference between a thirty-second and an eighth note is:

$$\frac{w}{8} - \frac{w}{32} = \frac{3w}{32}$$

and the difference between the note numbers n_0 and n_9 is

$$9 - 0 = 9.$$
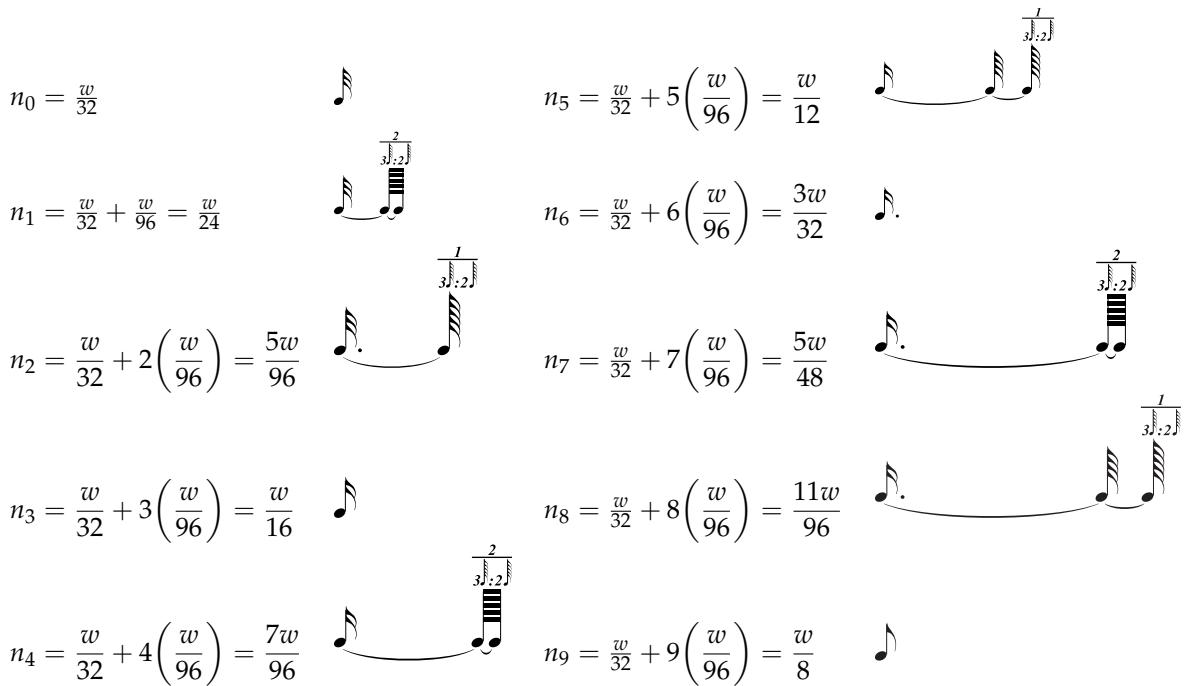
So, assuming that change in duration is linear, each note in succession is

$$\frac{\frac{3w}{32}}{9} = \frac{w}{96}$$

longer than the previous one, such that

$$n_i = n_{i-1} + \frac{w}{96}.$$

Therefore:

$$n_0 = \frac{w}{32}$$

$$n_1 = \frac{w}{32} + \frac{w}{96} = \frac{w}{24}$$

$$n_2 = \frac{w}{32} + 2\left(\frac{w}{96}\right) = \frac{5w}{96}$$

$$n_3 = \frac{w}{32} + 3\left(\frac{w}{96}\right) = \frac{w}{16}$$

$$n_4 = \frac{w}{32} + 4\left(\frac{w}{96}\right) = \frac{7w}{96}$$

$$n_5 = \frac{w}{32} + 5\left(\frac{w}{96}\right) = \frac{w}{12}$$

$$n_6 = \frac{w}{32} + 6\left(\frac{w}{96}\right) = \frac{3w}{32}$$

$$n_7 = \frac{w}{32} + 7\left(\frac{w}{96}\right) = \frac{5w}{48}$$

$$n_8 = \frac{w}{32} + 8\left(\frac{w}{96}\right) = \frac{11w}{96}$$

$$n_9 = \frac{w}{32} + 9\left(\frac{w}{96}\right) = \frac{w}{8}$$
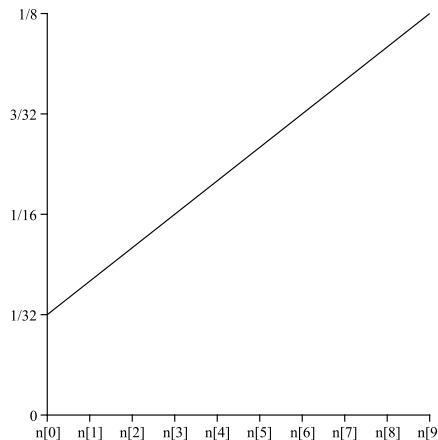
The addition of a constant value to each note in succession gives the linear equation

$$f(x) = \frac{w}{96}x + \frac{w}{32},$$
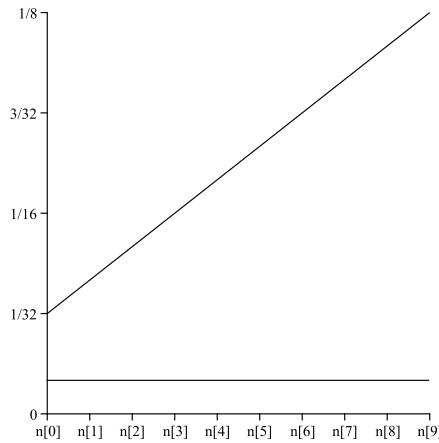
which corresponds to the following graph (Figure 6):

Figure 6: *George Crumb, Linear duration graph of "Notturno I", rehearsal 5.*

The linear graph has a constant rate of change or slope, so the first derivative yields a horizontal line. Since the rate of change is constant, there is no acceleration, so the second derivative is zero (Figure 7).

$$f(x) = \frac{w}{96}x + \frac{w}{32}, \quad f'(x) = \frac{w}{96}x, \quad f''(x) = 0.$$

73

Figure 7: $f(x)$, $f'(x)$, and $f''(x)$ of the linear duration graph of "Notturno I", rehearsal 5.

The sum of the durations of all 10 notes is
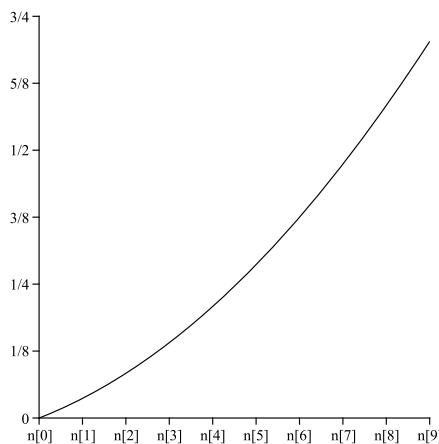
$$f(n_0) + \cdots + f(n_9) = \frac{25w}{32}.$$

The integral of $f(x)$ is

$$\int f(x)dx = \frac{w}{192}x^2 + \frac{w}{32}x,$$

which gives the sum of the durations of the notes as they occur, and corresponds to the graph (Figure 8)
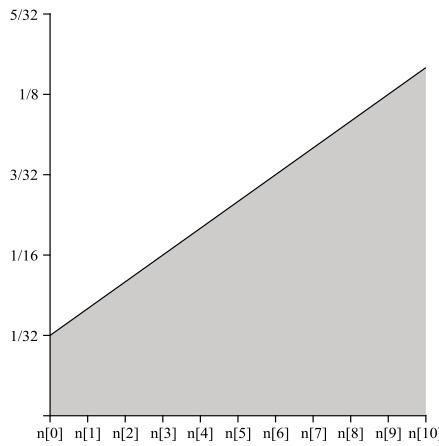
Figure 8: Integral of the linear duration graph of "Notturno I", rehearsal 5.

Solving the integral from 0 to 10 gives

$$\int_0^{10} f(x)dx = \frac{5w}{6},$$
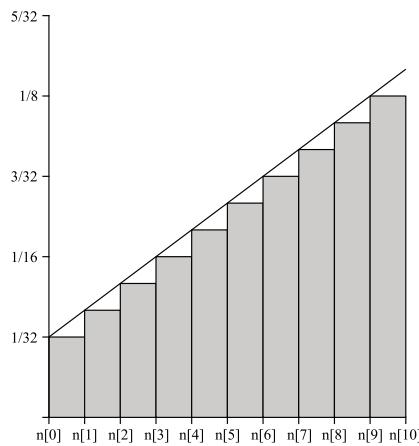
which is approximately the duration of all 10 notes in the feathered beam. The value is an approximation because $\int_0^{10} f(x)dx$ is equal to the entire area underneath $f(x)$ while the actual duration of all 10 notes is equal to the Riemann sums of $f(x)$.

$$\int_0^{10} f(x)dx = \frac{5w}{6},$$

Figure 9: *Entire area underneath $f(x)$.*

$$f(n_0) + \cdots + f(n_9) = \frac{25w}{32}.$$

Figure 10: *Riemann sums of $f(x)$.*

The linear interpretation, as shown above, is only one of several ways to realize feathered beams. Also useful is the quadratic equation which can lead to multiple interpretations. Ulf

Kronman and Johan Sundberg show that performers' ritardandi at the ends of compositions tend to follow quadratic curves, and Neil Todd uses parabolas to model gestures at the ends of phrases [7], [21],[22]. Furthermore, Bruno Repp suggests that the parabolic curve might "indeed represent a 'natural' way of changing tempo, including both accelerando and ritardando" [11]. For the "Notturno I" feathered beam above, the two endpoints $(0, \frac{w}{32})$ and $(9, \frac{w}{8})$ in conjunction with a third symmetrical coordinate $(-9, \frac{w}{8})$ can be used to solve the quadratic equation

$$g(x) = \frac{w}{864}x^2 + \frac{w}{32}.$$

The third coordinate may be chosen differently, and depending on its value, a wide variety of curves can be made for $g(x)$.[2] The graphs of $g(x)$ and its first and second derivatives are

$$g(x) = \frac{w}{864}x^2 + \frac{w}{32}, \quad g(x) = \frac{w}{432}x, \quad g''(x) = \frac{w}{432},$$
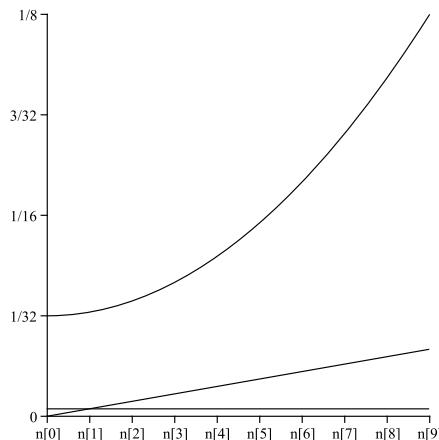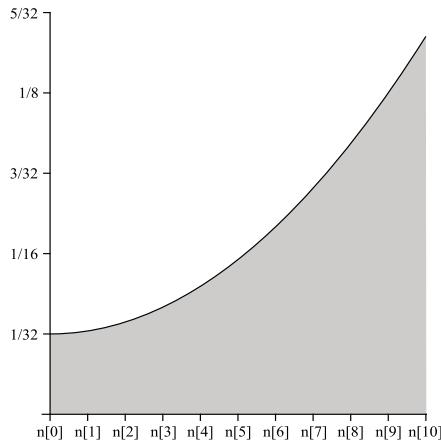
Figure 11: $g(x)$, $g'(x)$, and $g''(x)$ of the quadratic duration graph of "Notturno I", rehearsal 5.

which give the note values, rate of change, and the acceleration respectively. The integral of $g(x)$ from 0 to 10 is approximately equal to the sum of all 10 notes.
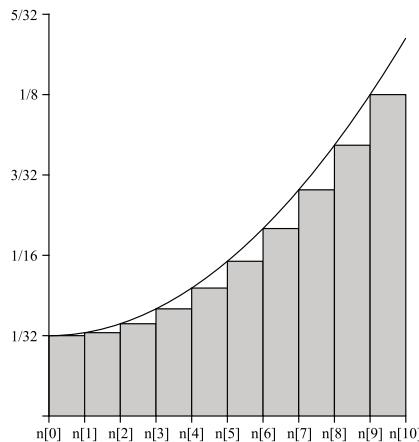
$$\int_0^{10} g(x)dx = \frac{905w}{1296}.$$

---

[2]Honing ([6]) points out that "the mere fact that the overall shape (e.g. of a square root function) can be predicted by the rules that come with human motion is not enough evidence for an underlying physical model of expressive timing in music performance, however attractive such a model might be." Thus, it is understood that this curve is taken arbitrarily.

Figure 12: *Entire area underneath $g(x)$.*

$$g(n_0) + \cdots + g(n_9) = \frac{185w}{288}$$
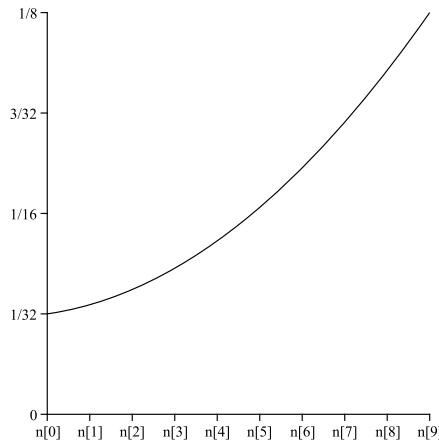
Figure 13: *Riemann sums of $g(x)$.*

Just like before, the discrepancy in the durations is due to the difference between the continuous area underneath the curve and the area quantized by the note values.

There are many ways that the notes in the feathered beam can be quadratically interpreted other than the way $g(x)$ shows. For example, consider the quadratic equation solved from the two endpoints $\left(0, \frac{w}{32}\right)$ and $\left(9, \frac{w}{8}\right)$ in conjunction with a third coordinate $\left(-11, \frac{w}{8}\right)$.

$$h(x) = \frac{w}{1056}x^2 + \frac{w}{528}x + \frac{w}{32}$$

Figure 14: *Another possible quadratic duration graph of "Notturno I", rehearsal 5.*

One notable difference between the two quadratic curves is that $g(x)$ begins with a slope of $g'(0) = 0$, while $h(x)$ begins with a greater slope $h'(0) = \frac{w}{528}$.

All of the examples so far have dealt with purely abstract concepts to conceptualize feathered beams. These concepts will now be applied to a real performance. Time points from a sound wave of the *Night Music I*, "Notturno I" excerpt as performed by Speculum Musicae are given below [3]. (N.b., the time points are measured from the highest relative amplitude for each note to 3 significant digits.) The time index begins at 0 seconds, and the times of the grace note and 10 notes in the feathered beam are

$$
\begin{array}{lll}
g.n. = 0.510s & n_3 = 1.020s & n_7 = 1.785s \\
n_0 = 0.544s & n_4 = 1.195s & n_8 = 2.071s \\
n_1 = 0.677s & n_5 = 1.374s & n_9 = 2.398s \\
n_2 = 0.837s & n_6 = 1.552s.
\end{array}
$$

Since the feathered beam begins with a thirty-second note, the duration of a thirty-second note in seconds can be calculated from the recording by subtracting n_0 from n_1 ($0.677s - 0.544s = 0.133s$). Furthermore, the tempo is *senza misura, quasi improvisando* ($\quarternote$ = ca.40, but very free). At $\quarternote = 40$, a thirty-second note is $0.1875s$, which differs significantly from the $n_1 - n_0$ calculation. Admittedly problematic, discrepancies such as this one will be discussed later in this paper. For now we will consider a thirty-second note to be $0.133s$, in order to purely demonstrate the application of splines to feathered beams. With a thirty-second note corresponding to $0.133s$, the notes under the feathered beam have the following durations (in whole-note units to 5 significant digits):
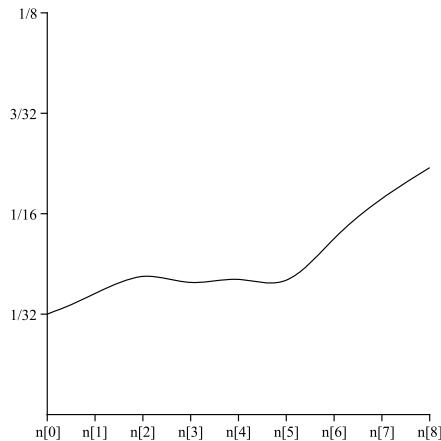
$$
\left\langle
\begin{array}{l}
\left(n_0, \dfrac{w}{32}\right), (n_1, 0.03759w), (n_2, 0.043w), (n_3, 0.04112w), (n_4, 0.04206w), \\
(n_5, 0.04182w), (n_6, 0.05475w), (n_7, 0.0672w), (n_8, 0.07683w))
\end{array}
\right\rangle
$$

The duration of the final note n_9 is not shown in the series of durations because there is no immediate onset after it for which to calculate its duration. Furthermore, n_9 has resonance slurs indicating that it is to be sustained. With the 9 coordinates, the spline $j(x)$ is solved as [3]

$$j(x) = \begin{cases} 0.03125w + 0.00599wx + 0.00035wx^3 & x < 1 \\ 0.03054w + 0.00705wx + 0.00106w(x-1)^2 - 0.00269w(x-1)^3 & 1 < x < 2 \\ 0.04084w + 0.00108wx - 0.00702w(x-2)^2 + 0.00406w(x-2)^3 & 2 < x < 3 \\ 0.04346w - 0.00078wx + 0.00516w(x-3)^2 - 0.00344w(x-3)^3 & 3 < x < 4 \\ 0.04516w - 0.00078wx - 0.00515w(x-4)^2 + 0.00569w(x-4)^3 & 4 < x < 5 \\ 0.01189w + 0.00599wx + 0.01192w(x-5)^2 - 0.00497w(x-5)^3 & 5 < x < 6 \\ -0.03467w + 0.0149wx - 0.003w(x-6)^2 + 0.00054w(x-6)^3 & 6 < x < 7 \\ -0.00658w + 0.01054wx - 0.00137w(x-7)^2 + 0.00046w(x-7)^3 & 7 < x \end{cases}$$

and corresponds to the graph (Figure 15):

Figure 15: *Duration graph of Night Music I, "Notturno I",rehearsal 5 as performed by Speculum Musicae.*

The notes n_3 and n_5 are shorter than the notes that immediately precede them. Furthermore, with the duration of a thirty-second note being calculated by the time difference between n_0 and n_1(i.e., $n_1 - n_0 = 0.133s = w/32$),the corresponding eighth note at n_9 would be significantly longer than n_8 according to the graph, contradicting what the feathered beam suggests.

The linear $f(x)$, quadratic $g(x)$, and spline $j(x)$ are three of many possible representations of the example from *Night Music I*, "Notturno I". The purpose of these three equations given in this paper is to demonstrate some technical aspects behind the feathered-beam technique. In the linear and quadratic representations, the note n_9 was assigned the duration of an eighth note even though its resonance slurs indicate that it is to be sustained. In the recording, n_9 was not measured because there is no event immediately following for which to calculate its duration. These seemingly arbitrary decisions were made to best compare different representations of the feathered beam excerpt. As one scours the repertoire for examples well suited to the analysis

---

[3]I would like to thank Michael Wester for his assistance in calculating the splines. It is not necessary to have full understanding of splines in order to grasp their relevance to this paper, but the reader may find further details about them in [4].

techniques demonstrated above, it is difficult to find feathered beams that do not require some arbitrary adjustment. The feathered beam from "Notturno I" is taken out of context to demonstrate technical aspects of feathered beams (as discussed in the previous few paragraphs) because the greater musical context introduces many more variables. Several questions remain to be pursued, especially about the complex ramifications for perception and performance.

The deviation from an exact representation of a feathered beam to a real performance, such as the one by Speculum Musicae, are known as "expressive microstructure." Deviations may not even be perceptible if they are not very severe. Under controlled circumstances with regularly repeating rhythms, deviations need to be at least 10 msec before they are perceptible, but deviations this small may not even be perceptible in a feathered beam because of the irregularity of the rhythm. It is difficult (or impossible) to generalize these deviations because, according to Bruno Repp, "the expressive microstructure of a musical performance reflects general, composer-specific, performer-specific, and piece-specific factors." In addition, "actual performances are likely to contain 'noise' in the form of random and planned deviations. . . , as one should expect from a human performer." "Musicians are usually only dimly aware of these variations, which they control intuitively rather than deliberately. Similarly, listeners perceive the structure and expression conveyed by these variations without being aware of the microstructure as such."[11, p. 222]

A feathered beam would likely be heard as a single expressive unit. The performance timing pattern of a feathered beam gesture probably needs to be compatible with the larger temporal organization [11, p. 225], which further complicates the interpretive restrictions placed upon it. According to Repp, the interpretation of expressive timing in general is due to cognitive analysis and mental representation of musical structure, and to feeling and expressive characterization, so to fully understand a feathered beam would require it to be considered within its context and not just as an isolated rhythm. In the greater context, tempo significantly affects the execution and perception of timing of rhythms [15, p. 389], especially when they comprise more than two interval durations [14, p. 590], such as with feathered beams. Much of the work that has gone into understanding how we execute and perceive rhythms has been done on interval durations based on simple ratios,[17, p. 63] but not much is known about how this happens with rhythms that rely on gradual changes of duration. Generalizations are hard to make even with simple traditional rhythms because, according to Bruno Repp, Justin London, and Peter Keller, "simple interval structure of rhythms seems to have an influence on the nature and degree of ratio distortion, and there are considerable individual differences as well." [17, p. 74] In regards to research on rhythm that has already been performed, the same authors point out that "It remains to be seen whether this result will hold up for sequences that have more complex (nonmetrical) interval ratios." The way in which we hear feathered beams is intrinsically complicated, and little research has been done to understand it. There cannot be a universal way to hear them because, according to Bruno Repp, John Iversen, and Aniruddh Patel, there is no universal way to hear rhythm in general:

> Although it is not uncommon to find that different listeners (or even the same listeners at different times) arrive at different metrical interpretations of the same rhythm, for any given listener at any given time his or her interpretation is considered to constitute a single, momentarily optimal solution to the informational jigsaw puzzle, with other possible solutions being discarded along the way or never even being considered. [19]

The anomalies in the splines example above may be perceptually insignificant, and the human element must certainly factor in, but there must be a point at which discrepancies in general can be considered errors. Another recording of this same passage or an investigation of other

feathered beams in the same recording would likely produce drastically different results, and further study of more passages might reveal further conclusions. In general, much more work remains to be done to fully understand these implications. Feathered beams are notation given by a composer, which are to be interpreted by a performer, and then heard by a listener and interpreted based on that hearing. In light of this observation, several questions come to mind. When a composer writes a feathered beam, how exact are they being with respect to the sound that they are hearing or wants to occur? When a performer reads a feathered beam, how exact are they being according to what the composer wants, and what the performer is trying to do? How consistent is the performer across performances? And when listeners hear feathered beams performed, how sensitive are they to variations in them?[4] For some composers, the preciseness of a feathered beam is irrelevant because they may merely mean from something fast to something slow.

The most notable discrepancy between the above linear, quadratic, and spline representations is their total durations. Although there are complications with exact comparison (e.g., how to handle n_9), the durations of $f(x)$, $g(x)$, and $j(x)$ from n_0 to n_8 as given in this paper are

$$f(0) + f(1) + \cdots + f(8) = \frac{21w}{32} = 0.65625w$$

$$g(0) + g(1) + \cdots + g(8) = \frac{191w}{352} = 0.542613\overline{6}w$$

$$j(0) + j(1) + \cdots + j(8) = \frac{27,781w}{50,000} = 0.43562w$$

The differences between the full durations show that problems arise when fitting a feathered beam group within a specific duration of metered music. This dilemma will be tackled in the next section of this paper.

5. The Full Duration of Feathered Beams

There exist feathered beams that exactly fit within a specific number of beats in commonly used meters. For example, consider a feathered beam group with 16 notes that linearly transitions from an eighth to a thirty-second note.

$$f(x) = \frac{w}{160} + \frac{w}{8}$$

$$f(0) + f(1) + \cdots + f(15) = \frac{5w}{4}$$

The sum of all 16 notes from n_0 to n_{15} is equal to 5 quarter notes. Feathered beams that work out this nicely are uncommon. When clarification for the full duration of a feather beam is needed, particularly in a metered context, Kurt Stone recommends placing a horizontal bracket over the feathered beam with a notehead indicating the full duration ([20, p. 124]). This section of the paper primarily focuses on the discrepancy between the sum of the note durations and the specified

---

[4]I would like to thank David Bashwiner for helping me arrive at these questions.

metered duration they span. Linear representations of the feathered beams are given in all cases, although the results can be duplicated for quadratic, spline, or some other representation of them. The conclusions drawn rely on the general concepts of feathered beams rather than exact calculations, so similar results can be duplicated with representations other than the linear case.

The following excerpt from the 1963 version of *Night Music I*, "Notturno V" has two feathered beams; each has 5 notes that gradually transition from a thirty-second to eighth note in the duration of 1 metered quarter note (Figure 16).

Figure 16: Night Music I, *"Notturno V"*, mm.13–15

The 5 notes are indexed n_0 to n_4. Subtracting the first $(0, \frac{w}{32})$ and last $(4, \frac{w}{82})$ notes from the total duration equals

$$\frac{w}{4} - (\frac{w}{8} + \frac{w}{32}) = \frac{3w}{32}$$

The feathered beam suggests that the remaining 3 notes between n_0 and n_4 are longer than a thirty-second note and shorter than an eighth note $(\frac{w}{32} < n_1, n_2, n_3 < \frac{w}{8})$; however, the allotted space they have is not long enough $f(n_1) + f(n_2) + f(n_3) > \frac{3w}{32})$. Given the slope

$$\frac{\frac{3w}{32}}{4} = \frac{3w}{128},$$

the linear equation that corresponds to the feathered beam is

$$f(x) = \frac{3w}{128}x + \frac{w}{32}$$

and fits in the duration of

$$f(0) + f(1) + \cdots + f(4) = \frac{25w}{64}$$

The total duration $\frac{25w}{64}$ is $\frac{9w}{64}$ longer than the $\frac{w}{4}$ space allotted. Therefore, the notes must be shortened, and there are many ways this can be done. Three ways are given below that act uniformly on the linearity.

$g(x)$: n_0 remains $\frac{w}{32}$ in duration, and the following 4 notes are shortened proportionally to decrease the slope.

$h(x)$: n_4 remains $\frac{w}{8}$ in duration, and the preceding 4 notes are shortened proportionally to increase the slope.

$j(x)$: All 5 notes are shortened by the same amount, and the slope remains unchanged.

The linear equations for $g(x)$, $h(x)$, and $j(x)$ are

$$g(x) = \frac{3w}{320}x + \frac{w}{32}$$
$$h(x) = \frac{3w}{80}x - \frac{w}{40}$$
$$j(x) = \frac{3w}{128}x + \frac{w}{320}.$$
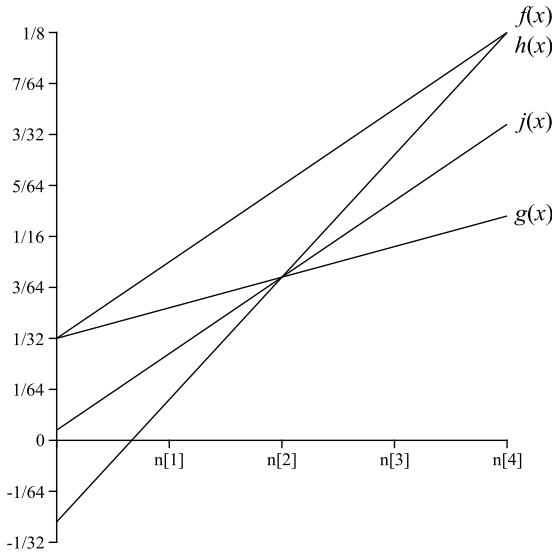
If desired, the reader may verify these equations by substituting any known values for n_0, n_4, or a into the sum of the linear equation $(ax + b)$ from $x = 0 \cdots 4$.

$$\frac{w}{4} = (0a + b) + (1a + b) + \cdots + (4a + b)$$
$$n_0 = 0a + b = b$$
$$n_4 = 4a + b$$
$$a = \frac{n_4 - n_0}{4}$$

The total durations of all three linear representations equals $\frac{w}{4}$ (this does not take into account that a performer may make the feathered beam longer than a quarter note).

$$g(0) + g(1) + \cdots + g(4) = \frac{w}{4}$$
$$h(0) + h(1) + \cdots + h(4) = \frac{w}{4}$$
$$j(0) + j(1) + \cdots + j(4) = \frac{w}{4}$$
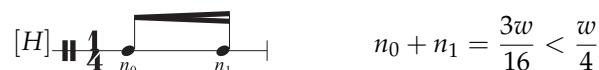
$f(x)$, $g(x)$, $h(x)$, and $j(x)$ are plotted on the graph in Figure 17:

Figure 17: *Fitting the feathered beam into $w/4$ by leaving the first note unchanged $(g(x))$, the last note unchanged $(h(x))$, and the slope unchanged $(j(x))$.*

When n_0 remains unchanged as in $g(x)$, n_4 is nearly a sixteenth note, which contradicts the eighth note in the score. When n_4 remains unchanged as in $h(x)$, $\frac{w}{4}$ is not long enough for the remaining notes to fit linearly, thus n_0 has a negative duration. Each note in a traditional tuplet is compressed by the same amount, so shortening each note in the feathered beam by the same amount seems like the intuitive choice, as in $j(x)$ but, with this choice, n_0 is nearly a two-hundred-fifty-sixth note, which is extremely fast at the tempo. Of the three lines whose durations sum to $\frac{w}{4}$, only $g(x)$ is performable, but it does not represent the final note duration in the feathered beam. A performance of this feathered beam would require something different than what is notated, indicating the subjectivity necessary to execute the rhythm. In the method that follows, this subjectivity is examined more definitively in several feathered-beam examples.

All of the feathered beams that follow are encapsulated within $\frac{1}{4}$ measures. For the first scenario, consider a feathered beam with only two notes.

$$[H] \qquad n_0 + n_1 = \frac{3w}{16} < \frac{w}{4}$$

There are no intervening notes to make the change in duration from the first to last notes gradual, thus it does not suffice for a feathered beam. If it is taken as a feathered beam, however, its duration needs to be increased to fill out the measure. Lengthening only the first note gives the rhythm:

Lengthening only the last notes gives the rhythm:

84

And lengthening both notes by the same amount gives the rhythm:
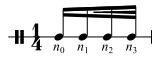
Thus, the feathered beam notation is superfluous in this case. The following table shows 8 more feathered-beam scenarios labeled a–h that in their literal linear interpretations most approximate the duration $\frac{w}{4}$.

a) $n_0 + ... + n_2 = \frac{9w}{32} > \frac{w}{4}$ e) $n_0 + ... + n_3 = \frac{3w}{16} < \frac{w}{4}$

b) $n_0 + ... + n_2 = \frac{15w}{64} < \frac{w}{4}$ f) $n_0 + ... + n_4 = \frac{15w}{64} < \frac{w}{4}$

c) $n_0 + ... + n_3 = \frac{5w}{16} > \frac{w}{4}$ g) $n_0 + ... + n_5 = \frac{9w}{32} > \frac{w}{4}$

d) $n_0 + ... + n_4 = \frac{25w}{64} > \frac{w}{4}$ h) $n_0 + ... + n_6 = \frac{21w}{64} > \frac{w}{4}$
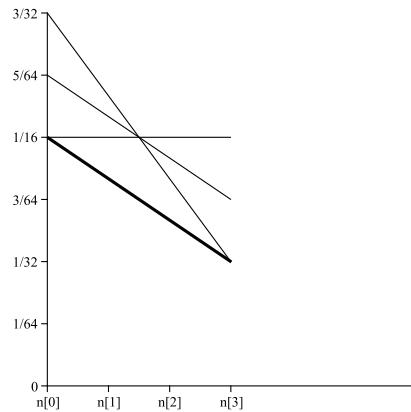
With 1 intervening note between the eighth and sixteenth notes, such as in *scenario a*, the total duration of the measure is exceeded by $\frac{w}{32}$ and the notes must be shortened to fit within the measure. Similarly in the other scenarios, the durations of the feathered beams are shorter or longer than the $\frac{w}{4}$ duration allotted. *Scenarios b, c,* and *d* transition from the durations of an eighth to a thirty-second over 3, 4, and 5 notes respectively. *Scenario b* is shorter than $\frac{w}{4}$ by $\frac{w}{64}$, and *scenarios c* and *d* are longer than $\frac{w}{4}$ by $\frac{w}{16}$ and $\frac{9w}{16}$ respectively. *Scenarios e, f, g,* and *h* transition from the durations of a sixteenth to a thirty-second over 3, 4, 5, and 6 notes respectively. *Scenarios e* and *f* are shorter than $\frac{w}{4}$ by $\frac{3w}{16}$ and $\frac{15w}{64}$ respectively, and *scenarios g* and *h* are longer than $\frac{w}{4}$ by $\frac{9w}{32}$ and $\frac{21w}{64}$ respectively.

For the eighth/thirty-second note feathered beams, *scenarios b* and *c* straddle the $\frac{w}{4}$ duration, and for the sixteenth/thirty-second note feathered beams, *scenarios f* and *g* straddle the $\frac{w}{4}$ duration. The other scenarios differ from $\frac{w}{4}$ by a greater amount. *Scenarios b* and *c* are as close to $\frac{w}{4}$ as the eighth/thirty-second note feathered beams can be, while *scenarios f* and *g* are as close to $\frac{w}{4}$ the sixteenth/thirty-second note feathered beams can be. Figures 18, 19, 20, and 21 shows, respectively, the graphs of *scenarios e, f, g,* and *h* for $f(x)$. The other lines modify the note durations in $f(x)$ so that the feathered beams precisely fit within the $\frac{1}{4}$ measures: $g(x)$ leaves the first note unchanged, $h(x)$ leaves the last note unchanged, and $j(x)$ leaves the slope unchanged. (In each graph, $f(x)$ is bolder than the other lines.)

Scenario e	Scenario f	Scenario g	Scenario h
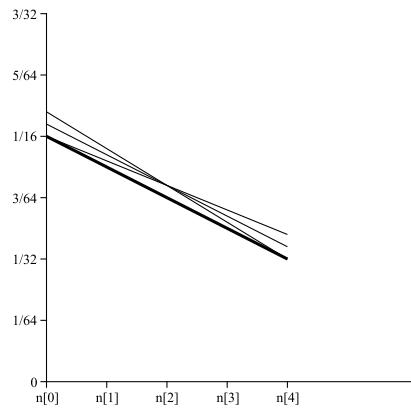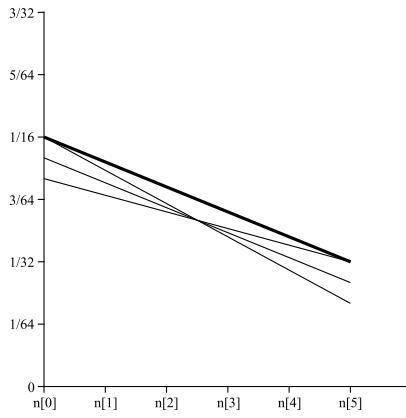
$$f(x) = -\frac{w}{96}x + \frac{w}{16} \qquad f(x) = -\frac{w}{128}x + \frac{w}{16} \qquad f(x) = -\frac{w}{160}x + \frac{w}{16} \qquad f(x) = -\frac{w}{192}x + \frac{w}{16}$$

$$g(x) = 0x + \frac{w}{16} \qquad g(x) = -\frac{w}{160}x + \frac{w}{16} \qquad g(x) = -\frac{w}{120}x + \frac{w}{16} \qquad g(x) = -\frac{w}{112}x + \frac{w}{16}$$

$$h(x) = -\frac{w}{48}x + \frac{3w}{32} \qquad h(x) = -\frac{3w}{320}x + \frac{11w}{160} \qquad h(x) = -\frac{w}{240}x + \frac{5w}{96} \qquad h(x) = -\frac{w}{672}x + \frac{9w}{224}$$

$$j(x) = -\frac{w}{96}x + \frac{5w}{64} \qquad j(x) = -\frac{w}{128}x + \frac{21w}{320} \qquad j(x) = -\frac{w}{160}x + \frac{11w}{192} \qquad j(x) = -\frac{w}{192}x + \frac{23w}{448}$$
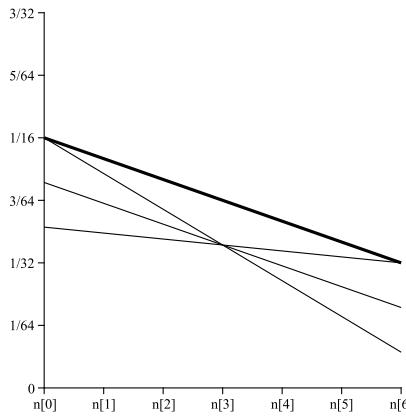
Figure 18: *Duration graph of* scenario e.

Figure 19: *Duration graph of* scenario f.

Figure 20: *Duration graph of* scenario g

Figure 21: *Duration graph of* scenario h.

In *scenarios e* and *f*, the feathered beams are shorter than $\frac{w}{4}$, so $g(x)$, $h(x)$, and $j(x)$ are above $f(x)$. In *scenarios g* and *h*, the feathered beams are longer than $\frac{w}{4}$, so $g(x)$, $h(x)$, and $j(x)$ are below $f(x)$. Of these 4 situations, *scenario f* is the closest in duration to $\frac{w}{4}$, so its notes need to be altered less than the other scenarios to fit within the measure. As the difference between the total duration of the feathered beam and $\frac{w}{4}$ grows, the notes need more alteration to fit within the measure. This can visually be seen in the graphs: $g(x)$, $h(x)$, and $j(x)$ move further from $f(x)$ as the difference increases. It is obvious that a greater discrepancy requires a greater modification. However, the graphs clearly provide specific information about—as well as a general visual representation of—the amount of compression or expansion involved in the performance modification.

6. Feathered Beams Generalized

Although the notation is imprecise, the full duration of a feathered beam can give a general idea about its precision. Of the numerous ways they can be interpreted, the general linear case is examined in this section of the paper because it is the simplest and most straight forward case. Although similar results can be duplicated with other representations of feathered beams, the linear case is taken here as mathematically representative.

Given the number of notes in the feathered beam group and the durations of the first and last notes, the equation for the full duration of all linear feathered beams is derived as follows. With the linear equation $f(x) = ax + b$, the full duration of the feathered beam is the sum of all the individual note durations

$$f(n_0) + f(n_1) + \cdots + f(n_i),$$

which expands to

$$(0a + b) + (1a + b) + \cdots + (ia + b),$$

and corresponds to the summation

$$\sum_{x=0}^{i} ax + b = \frac{i(i+1)a}{2} + (i+1)b. \tag{2}$$

The number of notes in the feathered beam is $(i+1)$. So that they are eighth notes or shorter, the first and last notes are in the form $\frac{1}{2^p}$ for $p > 2$, and they are

$$f(n_0) = 0a + b = b$$
$$f(n_i) = ia + b$$
$$f(n_0) \neq f(n_i).$$

Let the last note equal c so that the slope a is

$$a = \frac{c - b}{i}.$$

Substituting a into the summation gives

$$[H] \frac{i(i+1)(c-b)}{2i} + (i+1)b.$$

Given the durations of the first and last notes (b and c respectively) and the number of notes in the feathered beam ($i+1$), the equation can be expressed as a function of b, c, and i:

$$[h]f(b,c,i) = \frac{i(i+1)(c-b)}{2i} + (i+1)b,$$

which gives the full duration of linear feathered beams for all cases. In the above function, substituting values for the first note duration, last note duration, and number of notes in the feathered beam (b, c, and i respectively) simply gives the full linear duration. In application, for example, the full duration of the "Notturno V" excerpt earlier in this paper is

$$[H]f\left(\frac{w}{32}, \frac{w}{8}, 4\right) = \frac{4(4+1)\left(\frac{w}{8} - \frac{w}{32}\right)}{2 \cdot 4} + (4+1)\frac{w}{32} = \frac{25w}{64}$$

The reader may choose to derive the full length of the quadratic feathered beam. It is given as follows, assuming a third symmetrical coordinate.

$$[H]f(b,c,i) = -\frac{b(i+1)^3}{3i^2} + \frac{b(i+1)^2}{2i^2} - \frac{b(i+1)}{6i^2} + b(i+1)$$

The citations earlier in this paper suggest that the quadratic representation may be most useful because it most naturally reflects the human ritardando or accelerando. The general case is useful because it calculates the full duration, which can be compared to the allotted space in the score. Furthermore, the Java executable jar that can be accessed in the link `https://musmat.org/wp-content/uploads/2021/12/FeatheredBeamCalculator.zip` also quickly provides the full duration, in addition to providing other valuable information. When the difference between the full duration and the allotted space in the score is relatively small, the rhythm can be executed similarly to what the feathered beam suggests. The greater the difference, the less the notation reflects what can actually be played. Composers may find this information useful when implementing feathered beams.

7. Conclusion

As musicians, we are trained to precisely execute traditional triplets. Tuplets with more complicated ratios are usually more difficult to execute. Although the one in the measure below may be harder to negotiate than a traditional triplet, its rhythm exactly corresponds to the notation.

In tuplets, the usual circumstance is to compress all the notes according to exact ratios rather than augmenting them. To make the execution of feathered beams possible, there seems to be no preference for compressing the notes over augmenting them, and that individual circumstances dictate the choice. Bruno Repp found that in expressive microstructure variability in traditional rhythms, "lengthening is a common strategy, but accelerations beyond a maximal local tempo serve no expressive purpose and suggest poor timing control,"[12, p. 288] which implies that whether a feathered beam is compressed or augmented may have profound implications on how it is executed and perceived. The *scenarios a* through *h* earlier in this paper show a few ways the notes within feathered beams can be compressed or augmented to fit them within a prescribed duration of music. Furthermore, the examples show that some methods will not work in all circumstances, such as $g(x)$ in *scenario e* where there is no change in duration at all and $h(x)$ in the "Notturno V" exaLmple where there is a negative duration. Feathered beams in nearly all cases require some modification. The circumstances vary, and there is usually no single way in which they can be executed.The possibilities are bountiful, which is perhaps why Kurt Stone writes, "Besides, the gradual increase or decrease in the number of beams makes exact indications of beat units impossible."[20, p. 124] Even though the notation gives precise durations for the first and last notes in feathered beams, their execution is imprecise.

References

[1] Crumb, G. (1967). *Night Music I*. New York: Mills Music.

[2] Crumb, G. (1976). *Night Music I*. New York: EMI Mills, Inc.

[3] Crumb, G. (1996). *Night Music I*. Quest, Bridge Records 9069.

[4] Forsythe, G., Malcolm, M., Moler, C. (1977). *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

[5] Gould, E. (2011). *Behind Bars*: The Definitive Guide to Music Notation. London: Faber.

[6] Honing, H. (2003). Some Comments on the Relation Between Music and Motion. *Music Theory Online*, v. 9, n. 1, `https://www.mtosmt.org/issues/mto.03.9.1/mto.03.9.1.honing_frames.html`.

[7] Kronman, U.; Sundberg, J. (1989). Is the Musical Ritard an Illusion to Physical Motion? In: Gabrielsson, A. (ed.). *Action and Perception in Rhythm and Music*. Stockholm: Royal Swedish Academy of Music.

[8] Read, G. (1978). *Modern Rhythmic Notation*. Bloomington, IN: Indiana University Press.

[9] Repp, B. (1990). Composers' Pulses: Science or Art? *Music Perception*, v. 7, n. 4, pp. 423–434.

[10] Repp, B. (1989). Expressive Microstructure in Music: A Preliminary Perceptual Assessment of Four Composers' "Pulses". *Music Perception*, v. 6, n. 3, pp. 243–273.

[11] Repp, B. (1992). A Constraint on the Expressive Timing of a Melodic Gesture: Evidence from Performance and Aesthetic Judgment. *Music Perception*, v. 10, n. 2, pp. 221–241.

[12] Repp, B. (1998). The Detectability of Local Deviations from a Typical Expressive Timing Pattern. *Music Perception*, v. 15, n. 3, pp. 265–289.

[13] Repp, B. (2000). Pattern Typicality and Dimensional Interactions in Pianists' Imitation of Expressive Timing and Dynamics. *Music Perception*, v. 18, n. 2, pp. 173–211.

[14] Repp, B. (2002). Effects of Tempo on the Timing of Simple Musical Rhythms. *Music Perception*, v. 19, n. 4, pp. 565–593.

[15] Repp, B. (2010). Self-Generated Interval Subdivision Reduces Variability of Synchronization with a Very Slow Metronome. *Music Perception*, v. 27, n. 5, pp. 389–397.

[16] Repp, B.; Iversen, J.; Patel, A. (2008). Tracking an Imposed Beat within a Metrical Grid. *Music Perception*, v. 26, n. 2, pp. 1–18.

[17] Repp, B.; London, J.; Keller, P. (2005). Production and Synchronization of Uneven Rhythms at Fast Tempi. *Music Perception*, v. 23, n. 1, pp. 61–78.

[18] Repp, B.; London, J.; Keller, P. (2008). Phase Correction in Sensorimotor Synchronization with Nonisochronous Sequences. *Music Perception*, v. 26, n. 2, pp. 171–175.

[19] Repp, B.; Iversen, J.; Patel, A. (2008). Tracking an Imposed Beat within a Metrical Grid. *Music Perception*, v. 26, n. 2, pp. 1–18.

[20] Stone, K. (1980). *Music Notation in the Twentieth Century*. New York: W. W. Norton and Company.

[21] Todd, N. (1985) A Model of Expressive Timing in Tonal Music. *Music Perception*, n. 3, v. 1, pp. 33–57.

[22] Todd, N. (1989). A Computational Model of Rubato. *Contemporary Music Review*, n. 3, pp. 69–88.

Toward a Probabilistic Fourier Analysis on Audio Signals

Hugo Tremonte de Carvalho

Federal University of Rio de Janeiro (UFRJ)

Institute of Mathematics (IM)

Department of Statistical Methods (DME)

hugo@dme.ufrj.br

Orcid: 0000-0003-0776-0400

DOI: 10.46926/musmat.2021v5n2.91-102

Abstract: *In audio signal processing there are several tools employed to perform time-frequency analysis, being the spectrogram one of the most widely used. In a nutshell, it can be understood as a visual representation of the frequency content of an audio signal as it varies with time. Here we propose a probabilistic alternative to the spectrogram, which can be roughly interpreted as the most likely frequency to be present within an audio signal, also along time. This will be achieved by computing a specific posterior distribution in a Bayesian context. Preliminary experiments indicating the suitability of this object are presented, and potential applications to audio signal processing are outlined.*

Keywords: *Audio Signal Processing. Statistical Signal Processing. Bayesian Inference. Fourier Analysis. Spectrogram.*

I. Introduction

The main goal of this short note is to provide an answer to the following question: "how likely is some particular frequency to be present in some audio signal, regardless of its amplitude and phase, and the variance of any existing superimposed noise?". In fact, the answer here is not new, being firstly proposed in [2] for more general discrete-time signals. The novelty in this paper is the application of this general framework to audio signals, allowing then for probabilistic counterparts of well-known objects in audio signal processing, such as the spectrogram and the chromagram, and opening new possibilities to time-frequency related tasks, such as fundamental frequency estimation [6]. The paper is organized as follows: Section II introduces some basic notation to be used throughout the exposition, recalls the theory proposed in [2, Chap. 2] with more explained details in the analytical computations, and also extends the discussion on interpretations and suitability of the obtained quantities; in Section III, based on the previous discussions, we propose the probabilistic spectrogram and qualitatevely compares it with its classical counterpart; and finally, in Section IV we draw some conclusions and point out future works in this direction.

II. Bayesian spectrum estimation

Recall the formerly presented question: "how likely is some particular frequency to be present in the signal $x(t)$, regardless of its amplitude and phase, and the variance of any existing superimposed noise?". In this Section, we follow the discussion proposed on [2, Chap. 2], with more detail on some analytical computations and interpretations. It is important to remark that the exposed answer is not at its greatest generality, which can be found on [2, Chap. 3]; but we choose to maintain a simpler discussion for the sake of clarity. Moreover, preliminary computational experiments demonstrated only a negligible improvement by applying the more general framework.

i. The model

Let $x[n]$, for $n = 1, \ldots, N$, be a digital audio signal, derived from an analogical audio signal, denoted by $x(t)$, via uniform sampling with known frequency f_s, usually 44,100 Hz for an audio signal with the typical CD quality. More specifically, $x[n]$ is the n-th time sample of $x(t)$ and is given by $x(n/f_s)$. For answering the question in the beginning of Section II, we will assume the following model for $x(t)$:

$$x(t) = f(t) + \varepsilon(t), \tag{1}$$

where $f(t) = B_1 \cos(2\pi \omega t) + B_2 \sin(2\pi \omega t)$, $\varepsilon(t) \sim \mathcal{N}(0, \sigma^2)$ are observations of independent white noise, and $\omega > 0$ is measured in Hz.

Notice that $f(t)$, the systematic part of the signal $x(t)$, is equivalent to the more intuitive function $B \cos(2\pi(\omega t + \varphi))$, for adequate choices of B and φ. However, we will keep the former choice, since the latter will imply in more complicated computations. Still, two points are still unclear: 1) why only to use a single frequency to model the signal, since a musical signal will be much more complex than a simple sinusoid? and 2) why is the hypothesis of Gaussian white noise reasonable?

Essentially, we are looking for a compromise between simplicity and accuracy of our modeling. If we try to superimpose more and more frequencies in function $f(t)$, we will be approaching a perfect representation of the musical signal $x(t)$ by means of the Fourier Transform; and if we try to perfectly model how the observed signal $x(t)$ deviates from a single frequency and then incorporate it in the noise $\varepsilon(t)$, it will become extremely complex and thus intractable. The gaussianity of $\varepsilon(t)$ still could be questioned, but we justify it intuitively via the Maximum Entropy Principle [5]: the additive noise comprises several distinct deviations of the observed signal $x(t)$ from the model $f(t)$, therefore, the "most non-informative" probability distribution with finite mean and variance and supported in $\mathbb{R}$ is the Gaussian distribution. Moreover, it is still possible to argue that our goal is not to achieve a perfect substitute for the observed signal $x(t)$ by means of the model $f(t)$, but to answer probabilistic questions about the presence of a particular frequency on the observed signal $x(t)$.

ii. Formulating the problem

Given the remarks at the end of the last section, we may proceed with a few steps of the inference procedure. After the time sampling procedure, we have a sequence of observations given by

$$x[n] = f[n] + \varepsilon[n], \text{ for } 1 \leq n \leq N, \tag{2}$$

where $f[n] = f(n/f_s)$, and the $\varepsilon[n] \sim \mathcal{N}(0, \sigma^2)$ are independent. This hypothesis of white Gaussian noise implies the following distribution for the random vector $\varepsilon = [\varepsilon[1], \ldots, \varepsilon[N]]^T$:

$$p(\varepsilon|\sigma) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\varepsilon[n]^2}{2\sigma^2}\right\} \propto \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{n=1}^{N}\varepsilon[n]^2\right\}. \tag{3}$$

Note that $\varepsilon[n] = x[n] - f[n]$, for $n = 1, \ldots, N$, is a change of variables from vector ε to $\mathbf{x} = [x[1], \ldots, x[n]]^T$ with unitary Jacobian, that leads to the following distribution for $\mathbf{x}$:

$$p(\mathbf{x}|B_1, B_2, \omega, \sigma) \propto \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{n=1}^{N}(x[n] - f[n])^2\right\}$$

$$= \sigma^{-N} \exp\left\{-\frac{1}{2\sigma^2}\sum_{n=1}^{N}\{x[n] - [B_1\cos(2\pi\omega n/f_s) + B_2\sin(2\pi\omega n/f_s)]\}^2\right\}. \tag{4}$$

When considered as a function of $\{B_1, B_2, \omega, \sigma\}$, the expression in Equation 4 is called the *likelihood* and it quantifies the probability of observing the signal $\mathbf{x}$ if the parameters are given by some particular value of $\{B_1, B_2, \omega, \sigma\}$. Maximizing this function in variables $\{B_1, B_2, \omega, \sigma\}$ leads to the *maximum likelihood estimator* (MLE), interpreted as as the values of $\{B_1, B_2, \omega, \sigma\}$ for which the observed data $\mathbf{x}$ is the most probable.

But notice that this estimate has a severe problem: it does not answer our question! In order to convince ourselves, let us compare the question with the interpretation given above:

- **QUESTION**: "How likely is some particular frequency (ω) to be present in the signal $x(t)$, regardless of its amplitude and phase (both encoded in B_1 and B_2), and the variance of any existing superimposed noise (σ^2)?"
- **INTERPRETATION OF THE MLE**: "The values of B_1, B_2 (encoding amplitude and phase), ω (frequency), σ (standard deviation – square root of the variance – of the superimposed noise) for which the observed data $\mathbf{x}$ is the most probable."

Notice that in the interpretation of the MLE the probability is with respect to $\mathbf{x}$, where in the question it should be considered over the frequency ω. Moreover, the question is only about the frequency disregarding all the other parameters, but the interpretation considers all of them in the optimization process. The key to address these issues is Bayesian inference, which will be briefly discussed in the next subsection, in a more general context.

iii. Interlude – A glimpse of Bayesian inference

Essentially, Bayesian inference is a method of statistical inference where the information contained in observed data can be used to update the knowledge about parameters of interest. More precisely, let Z be a random variable which probability function of probability density function is denoted by $p(z|\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^k$ is a vector of parameters. We observe $\mathbf{z} = [z_1, \ldots, z_N]^T$ independent samples from Z and we want to estimate vector $\boldsymbol{\theta}$ from this data. The likelihood function is given by

$$\mathcal{L}(\boldsymbol{\theta}) = p(\mathbf{z}|\boldsymbol{\theta}) = \prod_{n=1}^{N} p(z_n|\boldsymbol{\theta}), \tag{5}$$

and quantify the probability of observing data $\mathbf{z}$ if the parameter vector is $\boldsymbol{\theta}$. Maximizing this function with respect to $\boldsymbol{\theta}$ will lead to its *maximum likelihood estimation*, which can be interpreted as the value of $\boldsymbol{\theta}$ for which the observed data is the most probable.

If we have prior information about $\boldsymbol{\theta}$, encoded in the so-called *prior distribution* $p(\boldsymbol{\theta})$, we can use Bayes theorem and obtain the *posterior distribution* for the parameters:

$$p(\boldsymbol{\theta}|\mathbf{z}) = \frac{\mathcal{L}(\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{z})} \propto \mathcal{L}(\boldsymbol{\theta})p(\boldsymbol{\theta}), \tag{6}$$

since $p(\mathbf{z})$ is a normalization term which in our application can be ignored. Therefore, maximizing the posterior distribution will lead to the *maximum a posteriori* (MAP) estimate for $\boldsymbol{\theta}$, representing the most probable parameters for the corresponding set of observation, a much easily interpretable quantity.

For more details on statistical and Bayesian inference, we refer the reader to [3, 4], respectively.

iv. The (Bayesian) inference procedure

By employing the Bayes' theorem, we are able to invert the conditional probability $p(\mathbf{x}|B_1, B_2, \omega, \sigma)$ and obtain

$$p(B_1, B_2, \omega, \sigma|\mathbf{x}) \propto p(\mathbf{x}|B_1, B_2, \omega, \sigma)p(B_1, B_2, \omega, \sigma). \tag{7}$$

The new term $p(B_1, B_2, \omega, \sigma)$ can encode previous knowledge about these variables, but we will adopt a conservative approach and use *non-informative priors*, that is, prior distributions that describe the most vague knowledge about them. For the parameters that can assume any real value (in this case B_1, B_2, ω)[1], we consider an improper prior distribution $p(B_1, B_2, \omega) \propto 1$, and for the scale parameter σ, which must be strictly positive, we impose the Jeffreys prior, given by $p(\sigma) \propto 1/\sigma$. By assuming prior independence between B_1, B_2, ω, σ, we have that

$$p(B_1, B_2, \omega, \sigma|\mathbf{x}) \propto p(\mathbf{x}|B_1, B_2, \omega, \sigma)p(B_1, B_2, \omega, \sigma)$$
$$\propto \sigma^{-(N+1)} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{n=1}^{N} \{x[n] - [B_1 \cos(2\pi\omega n/f_s) + B_2 \sin(2\pi\omega n/f_s)]\}^2 \right\}. \tag{8}$$

Firstly, let us expand the summation of squares inside the exponential to obtain:

$$\sum_{n=1}^{N} \{x[n] - [B_1 \cos(2\pi\omega n/f_s) + B_2 \sin(2\pi\omega n/f_s)]\}^2 =$$

$$= \sum_{n=1}^{N} x[n]^2 - 2B_1 \sum_{n=1}^{N} x[n] \cos(2\pi\omega n/f_s) - 2B_2 \sum_{n=1}^{N} x[n] \sin(2\pi\omega n/f_s)$$

$$+ B_1^2 \sum_{n=1}^{N} \cos^2(2\pi\omega n/f_s) + B_2^2 \sum_{n=1}^{N} \sin^2(2\pi\omega n/f_s)$$

$$+ 2B_1 B_2 \sum_{n=1}^{N} \cos(2\pi\omega n/f_s) \sin(2\pi\omega n/f_s). \tag{9}$$

Equation 9 can be simplified (exactly, without approximations) by using trigonometric identities, but this leads to tedious computations to prove such identities, or to evocations of some aspects from Linear Algebra ([2, Chap. 3] presents a discussion on this direction). We will not follow this path here and opt to employ some approximations, since we want to keep this discussion simple and easy to follow. Moreover, preliminary computational tests indicate that the gain in the

---

[1]In the beginning of Section II we constrained ω to be strictly positive for obvious physical reasons, but regarding the periodicity of trigonometric functions negative values of ω will also be allowed, from the mathematical viewpoint, to simplify the computations.

accuracy of the frequency estimation with the exact procedure is mostly negligible. In the light of these aspects we choose to approximate some of the terms in Equation 9 as follows:

- The term $\sum_{n=1}^{N} x[n]^2$ is constant (as a function of $\{B_1, B_2, \omega, \sigma\}$), and it will be denoted as $N\overline{x^2}$, where $\overline{x^2} = \dfrac{1}{N} \sum_{n=1}^{N} x[n]^2$.

- Both terms $\sum_{n=1}^{N} x[n]\cos(2\pi\omega n/f_s)$ and $\sum_{n=1}^{N} x[n]\sin(2\pi\omega n/f_s)$ are related to the real and imaginary parts of the Discrete Fourier Transform, and will be denoted, respectively, by $R(\omega)$ and $I(\omega)$.

- Both expressions $\sum_{n=1}^{N} \cos^2(2\pi\omega n/f_s)$ and $\sum_{n=1}^{N} \sin^2(2\pi\omega n/f_s)$ can be rewritten by using the simple trigonometric identities[2] $\sin^2(\alpha) = \dfrac{1 - \cos(2\alpha)}{2}$ and $\cos^2(\alpha) = \dfrac{1 + \cos(2\alpha)}{2}$ to obtain:

$$\sum_{n=1}^{N} \cos^2(2\pi\omega n/f_s) = \frac{N}{2} + \frac{1}{2}\sum_{n=1}^{N} \cos(4\pi\omega n/f_s), \tag{10}$$

$$\sum_{n=1}^{N} \sin^2(2\pi\omega n/f_s) = \frac{N}{2} - \frac{1}{2}\sum_{n=1}^{N} \cos(4\pi\omega n/f_s). \tag{11}$$

Analogously, the term $\sum_{n=1}^{N} \cos(2\pi\omega n/f_s)\sin(2\pi\omega n/f_s)$ can be rewritten, using the formula for the sine of a sum of angles, to obtain:

$$\sum_{n=1}^{N} \cos(2\pi\omega n/f_s)\sin(2\pi\omega n/f_s) = \frac{1}{2}\sum_{n=1}^{N} \sin(4\pi\omega n/f_s). \tag{12}$$

In [2, p. 17] it is claimed that these three trigonometric sums can be neglected, when compared to $N/2$, the dominant term appearing in the last two lines of Equation 9, assuming that the rather vague conditions $N \gg 1$ and $\omega N/f_s \ll 1$ are valid. Indeed, computational tests verified that for the human hearing frequency range (about 20 Hz to 20,000 Hz), the usual sampling rate of a CD-quality audio signal ($f_s = 44,100$ Hz) and the window length here employed ($N = 4,096$ – see Sec. III for more details), these neglected terms are two to three orders of magnitude smaller than $N/2$, validating the claims above in our context.

Given the discussion above, we can rewrite Equation 9 as[3]:

$$\sum_{n=1}^{N} \{x[n] - [B_1\cos(2\pi\omega n/f_s) + B_2\sin(2\pi\omega n/f_s)]\}^2 =$$

$$= N\overline{x^2} - 2B_1 R(\omega) - 2B_2 I(\omega) + B_1^2\frac{N}{2} + B_2^2\frac{N}{2} \tag{13}$$

$$= N\left[\overline{x^2} - \frac{2}{N}[B_1 R(\omega) + B_2 I(\omega)] + \frac{1}{2}(B_1^2 + B_2^2)\right].$$

---

[2] Both these identities follow from the formulas for the sine and cosine of a sum of angles and from the fundamental trigonometric relation $\sin^2(\alpha) + \cos^2(\alpha) = 1$.

[3] The first symbol "=" is not accurate in Equation 13, being the "≈" preferable. However, we will go ahead with the slight abuse of notation, in order to simplify the notation.

By substituting Equation 13 in Equation 8, we have that:

$$p(B_1, B_2, \omega, \sigma | \mathbf{x}) \propto p(\mathbf{x} | B_1, B_2, \omega, \sigma) p(B_1, B_2, \omega, \sigma)$$

$$\propto \sigma^{-(N+1)} \exp \left\{ -\frac{N}{2\sigma^2} \left[\overline{x^2} - \frac{2}{N} [B_1 R(\omega) + B_2 I(\omega)] + \frac{1}{2} (B_1^2 + B_2^2) \right] \right\}. \tag{14}$$

v. Marginalization of nuisance parameters

Now, with the simplified expression for the posterior distribution, given in Equation 14, we are able to marginalize the undesired parameters and obtain only $p(\omega | \mathbf{x})$, given by

$$p(\omega | \mathbf{x}) \propto \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(B_1, B_2, \omega, \sigma | \mathbf{x}) \, dB_1 dB_2 d\sigma. \tag{15}$$

To perform the integrals in B_1 and B_2 (and then obtaining $p(\omega, \sigma | \mathbf{x})$) we must resort to the following Lemma:

Lemma 1. *Let $a \in \mathbb{R}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{C} \in \mathbb{R}^{m \times m}$ be numerical, vectorial, and matricial constants, respectively. Assume also that matrix $\mathbf{C}$ is invertible. Then, the following equality holds:*

$$\int_{\mathbb{R}^m} \exp \left\{ -\frac{1}{2} (a + \mathbf{b}^T \mathbf{y} + \mathbf{y}^T \mathbf{C} \mathbf{y}) \right\} \, d\mathbf{y} = \frac{(2\pi)^{m/2}}{|\det(\mathbf{C})|^{1/2}} \exp \left\{ -\frac{1}{2} \left[a - \frac{\mathbf{b}^T \mathbf{C}^{-1} \mathbf{b}}{4} \right] \right\}. \tag{16}$$

The proof of this lemma, although simple, is quite tedious, since it involves completing the squares in the argument of the left exponential, identifying the kernel of a multivariate normal distribution, and rearranging the terms to obtain the desired result. Here we simply employ it and refer the interested reader to [1] for the proof.

As can be verified by a simple computation, by defining $a = \frac{N \overline{x^2}}{\sigma^2}$, $\mathbf{b} = -\frac{2}{\sigma^2} \begin{bmatrix} R(\omega) \\ I(\omega) \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} N/2\sigma^2 & 0 \\ 0 & N/2\sigma^2 \end{bmatrix}$, and substituting these values in the left-hand side of Equation 16 we recover exactly Equation 14. Therefore, by the direct application of Lemma 1, we have that

$$p(\omega, \sigma | \mathbf{x}) \propto \sigma^{-(N-1)} \exp \left\{ -\frac{N}{2\sigma^2} \left[\overline{x^2} - \frac{2}{N} C(\omega) \right] \right\}, \tag{17}$$

where $C(\omega)$ is defined as $[R(\omega)^2 + I(\omega)^2]/N$. This quantity is called the *periodogram* in [2, p. 7], and is a natural appearance in this probabilistic context of the squared-magnitude of the Discrete Fourier Transform.

Finally, in order to integrate Equation 17 with respect to σ, we can make a change of variables and use the kernel of an Inverse-Gamma distribution, a more classical approach – perhaps the one employed by Bretthorst in the computations omitted in [2, pp. 18–20]. To avoid extensive computations, we refer to the more recent work [7], where the *Inverse-Nakagami distribution* is proposed, which probability density function, in the variable $\sigma > 0$, is given by

$$\frac{2}{\Gamma(\lambda)} \left(\frac{\lambda}{\xi} \right)^\lambda \sigma^{-2\lambda - 1} \exp \left(-\frac{\lambda}{\xi \sigma^2} \right), \tag{18}$$

being λ and ξ strictly positive real parameters. By using the fact that every probability density function must integrate 1 over its support, we can use the normalizing constant of this distribution to obtain the result

$$\int_0^{+\infty} \sigma^{-2\lambda - 1} \exp \left(-\frac{\lambda}{\xi \sigma^2} \right) \, d\sigma = \frac{\Gamma(\lambda)}{2} \left(\frac{\xi}{\lambda} \right)^\lambda. \tag{19}$$

By comparing the term being integrated in Equation 19 with the joint density of $\{\omega, \sigma\}$ in Equation 17, we obtain that the suitable values of λ and ξ are given by

$$\lambda = \frac{N-2}{2}, \tag{20}$$

$$\xi = \frac{N-2}{2}\left[\overline{x^2} - \frac{2}{N}C(\omega)\right]^{-1}. \tag{21}$$

Therefore, by using the result of Equation 19 and discarding the multiplicative terms that does not depend on ω, we have that that

$$p(\omega|\mathbf{x}) \propto \left[1 - \frac{2C(\omega)}{N\overline{x^2}}\right]^{\frac{2-N}{2}}, \tag{22}$$

an expression for the distribution we are looking for.

Preliminary computational experiments indicated that trying to compute directly this quantity may cause a numerical overflow. Therefore, we opt to compute its logarithm[4], given by

$$\log_{10} p(\omega|\mathbf{x}) = \frac{2-N}{2}\log_{10}\left[1 - \frac{2C(\omega)}{N\overline{x^2}}\right] + \text{constant terms}, \tag{23}$$

where the constant terms are the logarithm of the multiplicative constants unconsidered along the computations. Recall that considering a strictly positive function or its logarithm is equivalent for finding its local maxima, because the logarithm is a strictly crescent function.

We now present an application of this distribution to audio signal processing.

III.   The probabilistic spectrogram

The *spectrogram* is a fundamental tool in time-frequency analysis of audio signals, and it can be interpreted as the frequency content of an audio signal along time. In order to understand it more precisely, recall that [8, Sec. 2.5] defines the Short-Time Fourier Transform as

$$\mathcal{Y}[m,k] = \sum_{n=0}^{N-1} y[n+mH]w[n]e^{-2\pi ikn/N}, \tag{24}$$

for $m \in \mathbb{Z}$[5] and $k = 0, \ldots, \lfloor N/2 \rfloor$. The quantity H is the hop size, and it is related to the overlap between two consecutive windows; a common value to adopt is $H = \lfloor N/2 \rfloor$, indicating 50% of overlap.

Essentially, Equation 24 computes the Discrete Fourier Transform of the signal

$$x_m[n] = y[n+mH]w[n], \text{ for } n = 0, \ldots, N-1, \tag{25}$$

that is, an excerpt of length N from an audio signal y, properly smoothed on the edges by some window function w, with support of size N. The spectrogram is then defined as the squared-magnitude of each value of the Short-Time Fourier Transform, that is, $|\mathcal{Y}[m,k]|^2$. For more details on the spectrogram and on the choice of windowing functions, see [8, Sec. 2.5].

---

[4]The base of the logarithm is of minor importance here, but we will adopt 10, since it is the standard base used to transform magnitude to dB, an usual unit in Signal Processing.

[5]Note that m is in practice restricted to a finite set $\{0, \ldots, M\}$ such that these successive hops of size H cover the entire signal y. Since this specific range will not be directly used here, we avoided its explicit introduction, for the sake of clarity.

Since the probability density function on Equation 22 is related to the presence of frequencies ω in signal $\mathbf{x}$, instead of computing the Discrete Fourier Transform of the windowed signal $\mathbf{x}_m$, we can compute the probability distribution $p(\omega|\mathbf{x}_m)$, for each $m \in \mathbb{Z}$. The collection of all these probability distributions for $m \in \mathbb{Z}$ is called the *probabilistic spectrogram* of the signal y. This representation is expected to be more sparse than the spectrogram, since it computes the most likely frequencies at each time-frame and may disregard high-order harmonics of the instrument. Moreover, for each $m \in \mathbb{Z}$, we can compute the maximum a posteriori estimate for ω, that is, argmax $p(\omega|\mathbf{x}_m)$. This sequence of values may provide an interesting low-dimensional feature useful in Music Information Retrieval[6].

In order to illustrate this concept, we consider the subject of *Ricercar a 6*, the six-voice fugue that is commonly regarded as the high point of *The Musical Offering* (BWV 1079), by Johann Sebastian Bach, interpreted in two instruments: the fortepiano and the harpsichord[7]. Both audio signals were downloaded from YouTube in MP3 format with 320 Kbps, manually trimmed to contain only the subject, exported in WAV format and reduced to a monophonic signal. The sampling frequency is 44,100 Hz, the length of the analysis window was taken to be $N = 4{,}096$ time-samples with hop size of $H = 2{,}048$, implying in 50% of overlap between adjacent windows.

Figures 1 and 2 display the spectrogram of the subject of *Ricercar a 6* played on the fortepiano and the harpsichord, respectively. Note that the harpsichord signal contains much more harmonics than the fortepiano one, as it is naturally expected for both instruments.

In Figures 3 and 4 are displayed the logarithm of the probabilistic spectrogram of the subject of *Ricercar a 6*, played on the pianoforte and harpsichord, respectively. Blue color indicates negligible probability, and colors close to yellow indicate more likely ones. The y axis is restricted from 0 Hz to 2,500 Hz, since there are no likely frequencies above this value. Notice that these representations also capture higher harmonics of the harpsichord, but are cleaner than the spectrogram.

Indeed, Figures 5 and 6 display both the probabilistic spectrograms as in Figures 3 and 4, but the red dots indicate the most likely frequencies at each time frame. Notice that, although visually distinct, they look more similar than the spectrograms in Figures 1 and 2. This fact indicates that these sequences of frequencies can be regarded as an interesting audio feature, and a refinement of this quantity may lead to advances in algorithms of fundamental frequency detection.

IV. Conclusion and future works
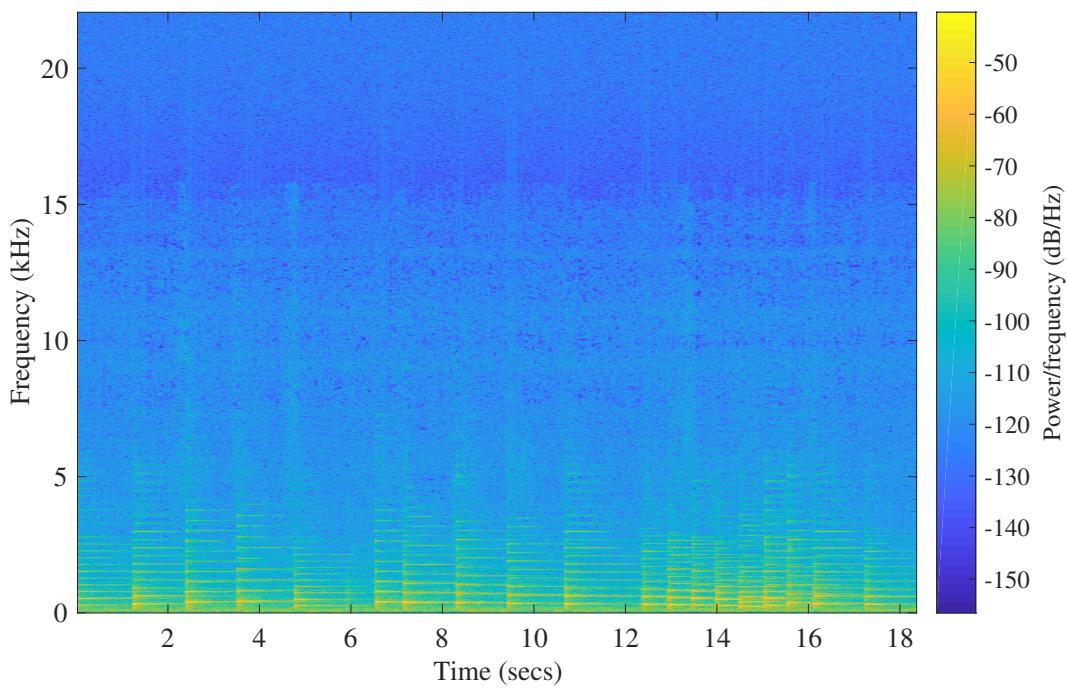
In this paper we recalled and detailed the computations done in [2, Cap. 2], in order to study the posterior distribution of frequencies present within an audio signal, disregarding its amplitude and phase, and also the variance of any existing superimposed noise. This probability distribution was used to propose a new concept in time-frequency analysis of audio signals, the *probabilistic spectrogram*, illustrated by two musical examples.

Althoguh not quantitatively analyzed, the proposed quantity seems promising to be further investigated, specially its usage in Music Information Retrieval tasks, mainly because of its sparseness and high interpretability. This is addressed as a future work, together with the development of the *probabilistic chroma features*, corresponding to the probability allocated in the frequencies bands related to the twelve pitch classes[8].

---
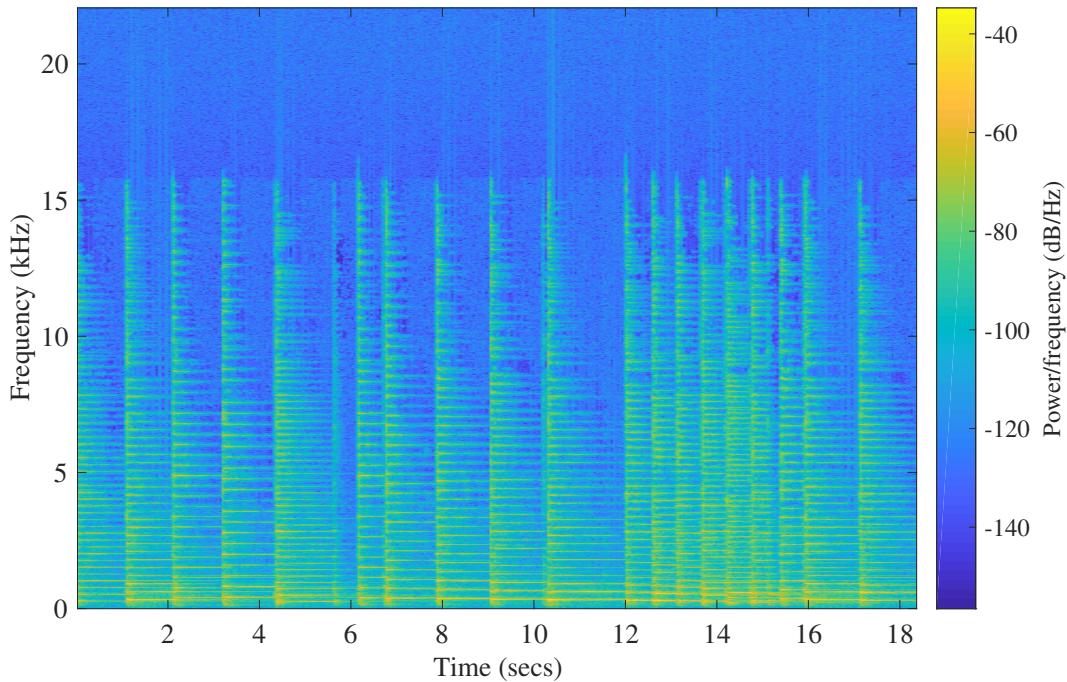
[6]In [8, Sec. 7.1.2] it is described that peaks of the spectrogram provides an useful fingerprint for the taks of content-based audio retrieval.

[7]Both interpretations are available on YouTube. The fortepiano version is played by Leo van Doeselaar (https://www.youtube.com/watch?v=fjxKy3pP41w) and the harpsichord one is interpreted by Iain Simcock (https://www.youtube.com/watch?v=AYw2E5F930M).
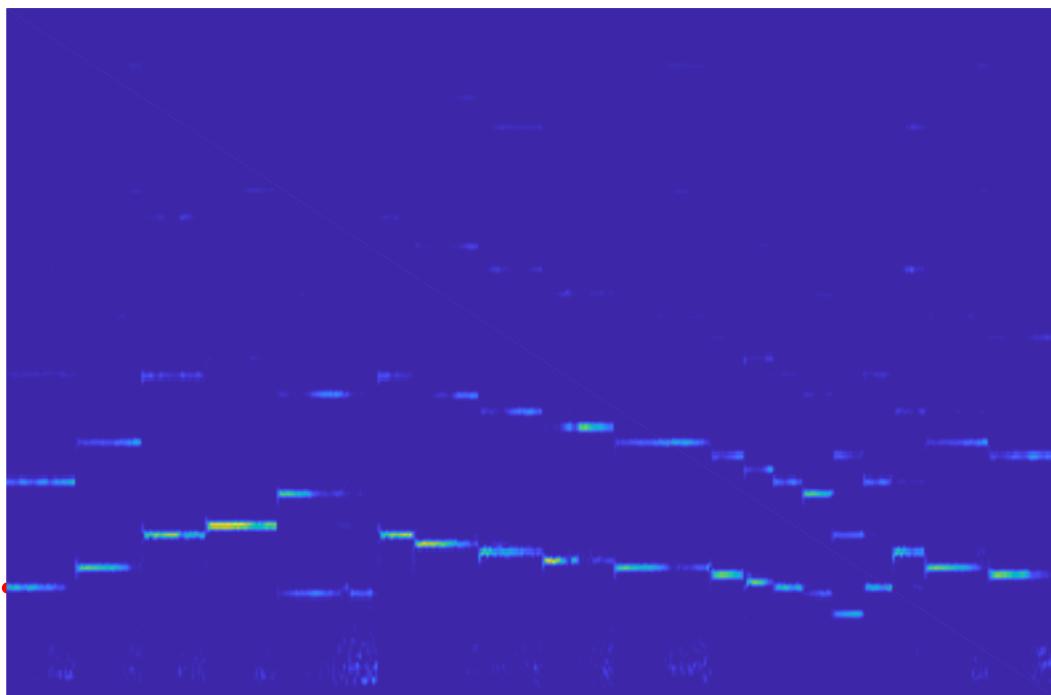
[8]Assuming an equal emperament tuning with fundamental frequency of A_4 defined as 440 Hz.
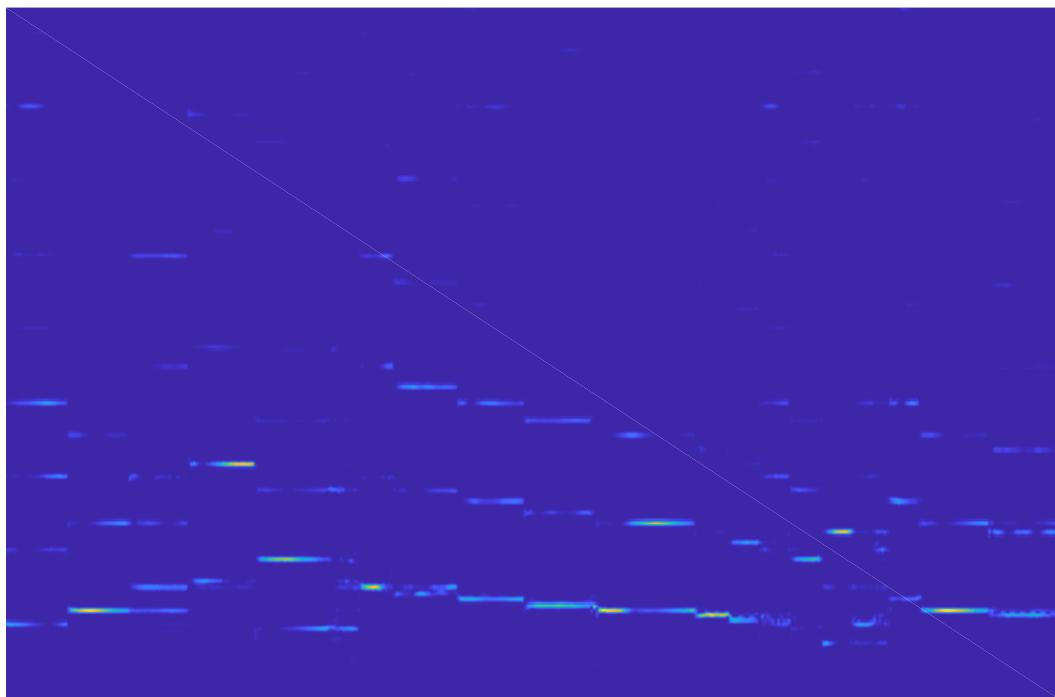
Figure 1: *Spectrogram of the subject of* Ricercar a 6, *played on the fortepiano.*
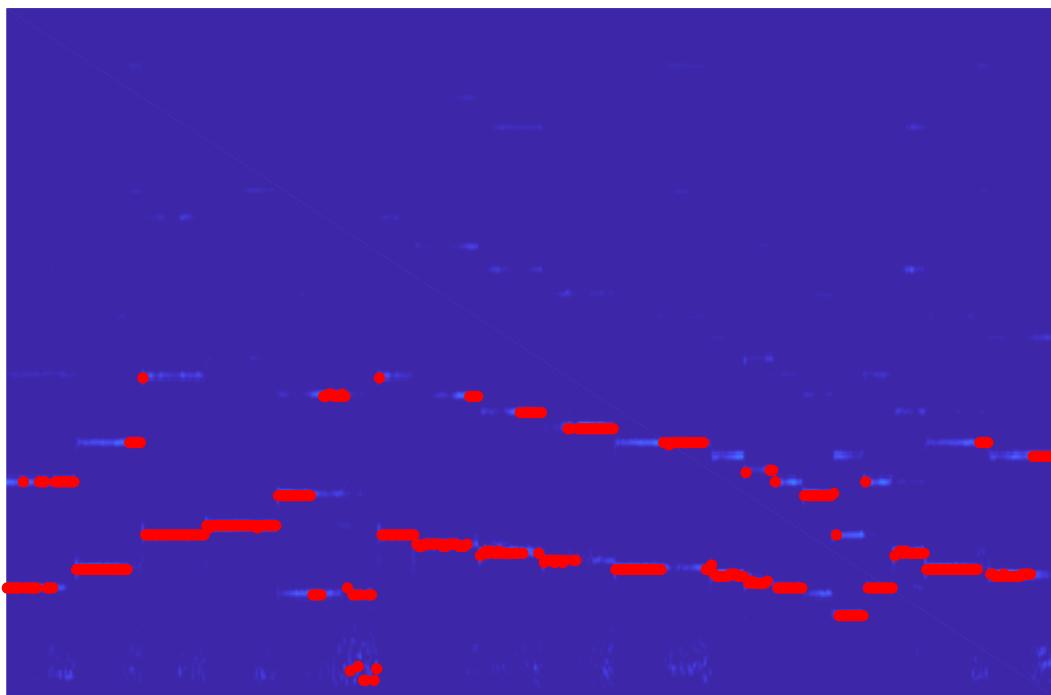
Figure 2: *Spectrogram of the subject of* Ricercar a 6, *played on the harpsichord.*
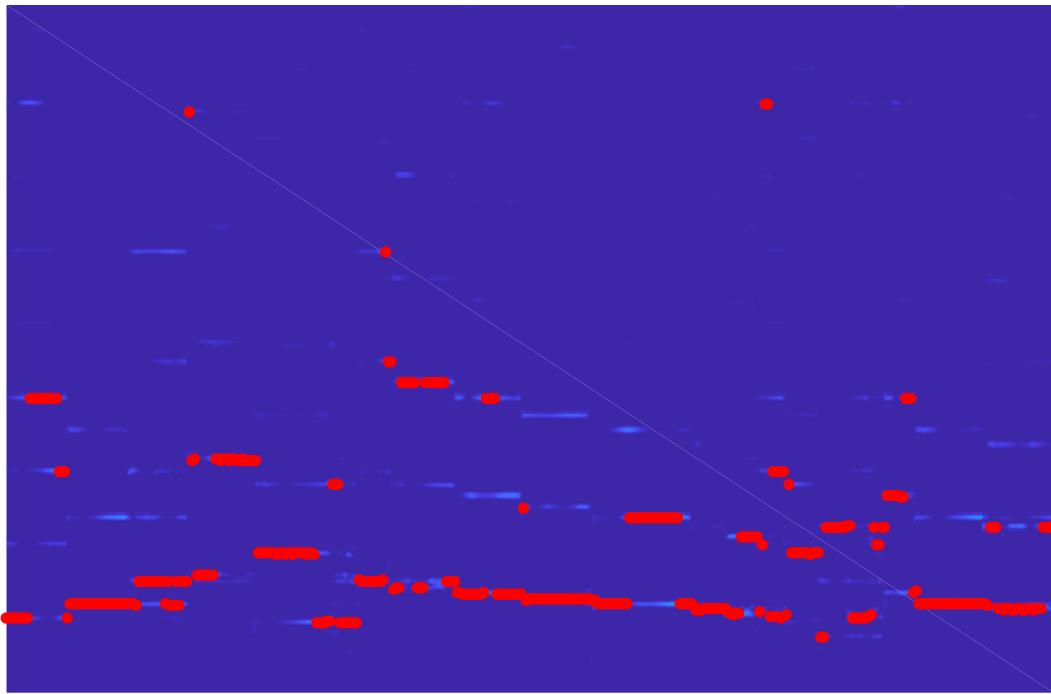
Figure 3: *Probabilistic spectrogram (logarithm; y axis restricted from 0 Hz to 2,500 Hz) of the subject of* Ricercar a 6, *played on the fortepiano.*

Figure 4: *Probabilistic spectrogram (logarithm; y axis restricted from 0 Hz to 2,500 Hz) of the subject of* Ricercar a 6, *played on the harpsichord.*

Figure 5: *Probabilistic spectrogram (logarithm; y axis restricted from 0 Hz to 2,500 Hz) of the subject of* Ricercar a 6, *played on the fortepiano. The red dots indicate most likely frequency at each time-frame.*

Figure 6: *Probabilistic spectrogram (logarithm; y axis restricted from 0 Hz to 2,500 Hz) of the subject of* Ricercar a 6, *played on the harpsichord. The red dots indicate most likely frequency at each time-frame.*

References

[1] Bishop, C. (2006) *Pattern Recognition and Machine Learning*. New York: Springer.

[2] Bretthorst, G. (1988) *Bayesian Spectrum Analysis and Parameter Estimation*. New York: Springer.

[3] Casella, G.; Berger, R. (2001) *Statistical Inference*. Boston: Cengage Learning.

[4] Gelman, A; Carlin, J.; Stern, H.; Dunson, D.; Vehtari, A.; Rubin, D. (2013) *Bayesian Data Analysis*. Boca Raton: CRC Press.

[5] Jaynes, E. (1982) On The Rationale of Maximum-Entropy Methods. *Proceedings of the IEEE*, 70/9, pp. 939–952.

[6] Klapuri, A. (ed.); Davy, M. (ed.) (2006) *Signal Processing Methods for Music Transcription*. New York: Springer.

[7] Louzada, F.; Ramos, P.; Nascimento, D. (2018) The Inverse Nakagami-m Distribution: A Novel Approach in Reliability. *IEEE Transactions on Reliability*, 67/3, pp. 1030–1042.

[8] Müller, M. (2015) *Fundamentals of Music Processing*: Audio, Analysis, Algorithms, Applications. New York: Springer.

Gestural Presheaves: From Yoneda to Sheaves

*Juan Sebastián Arias-Valero
Universidad Nacional Autónoma de México
jsariasv1@gmail.com
Orcid: 0000-0001-6812-7639

Emilio Lluis-Puebla
Universidad Nacional Autónoma de México
lluisp@unam.mx
Orcid: 0000-0002-2633-5955

Abstract: *We study gestural presheaves in the language of abstract gestures and, in particular, a general gestural version of the Yoneda embedding, based on a Mazzola's idea, and gestural sheaves. We apply the latter to Mozart and Beethoven.*

Keywords: *Gestures. Melodic contour. Category theory. Yoneda lemma. Sheaves.*

I. Introduction

Mathematical gestures are present in mathematical music theory in several forms. First, they intend to model the configuration and movement of the performer's body when playing an instrument (topological gestures). Second, melodic contours are also topological gestures on the score. Third, all diagrams of transformational theory, as any diagram, are abstract gestures in categories. Introductions to gesture theory are [5, 3, 15]. More general advances include Mazzola's [16], Mannone's [13, 12], and Arias' [2].

This article explores the gestural presheaf construction (Section II.i) in the general language of abstract gestures (Sections II and III), especially the relation thereof to two main developments of presheaves in category theory: the Yoneda embedding and sheaves.

Regarding the first, which represents any category as a subcategory of its associated category of presheaves, it is also relevant for gesture theory since we would like to have a *gestural Yoneda embedding* [14, p. 33] that now represents the given category in a suitable category related to

gestures and enables the recovery of a certain gestural intuition behind a morphism. For example, this embedding should help recover the gesture of a linear transformation associated with a rotation matrix, which only takes into account an argument and its image but not the intuitive movement between them. Similarly, following Lewin [9, p. 159], it should help to find the gestures behind musical transformations. We locate conceptually and generalize (Section V) Mazzola's gestural version of Yoneda lemma, which was originally formulated in the category of topological categories, to more arbitrary categories. In particular, we refine these results to obtain *an embedding of a category into a category of gestural presheaves, under certain conditions*. A previous version of the gestural Yoneda embedding was given in [14, 3.4], however it was not full.

With respect to sheaves, in this article we motivate from a musical perspective their introduction in gesture theory. The motivation is essentially based on the melodic contour gestures and the fact that in some musical works these gestures glue together to form new ones. These procedures could explain some passages obtained by variation in Mozart and Beethoven (Section VIII). The gestural sheaf notion (especially the global section one) is related to that of global gesture in [16, Section 66.5], which was formulated in the category of topological categories and resembles the global composition concept. However, sheaves are simpler and still useful. We provide the *gestural sheaf* notion for more general categories in Section VII.

By the way, the following results emerged in the present research. 1. Determination of left adjoints to gestural presheaves (Section II.i), which helps to characterize the (generalized) elements of objects of gestures. 2. Determination of left adjoints to the covariant gesture functors (Section VI, based on Section IV). 3. Characterization of sheaves in terms of cotensor products, suggesting a close relation between gestures and sheaves (Section VII).

We include a Glossary of specialized terms that the reader can access by clicking on a red term like sheaf. Also, these hyperlinks direct to some term in boldface if it was defined in the article. However, it is desirable that the reader has a minimum acquaintance with category theory [11, pp. 10-23]. On the other hand, we end each example with the symbol ♣ for organization.

II. Abstract gestures

Throughout this paper all examples are focused on topological gestures whose skeleta are digraphs, for simplicity. However, we will use the general language of *abstract gestures*, which includes topological gestures as a particular case, because it include other instances[1] of gesture theory and allows the relation with several valuable categorical concepts, like cotensor products, presheaves, and sheaves.

Consider the following data:

- A small category $\mathcal{D}$. In gesture theory, we often explore the case when $\mathcal{D}$ is the category pictured in the following diagram.

$$id \; \circlearrowright \; [0] \; \underset{\epsilon_0}{\overset{\epsilon_1}{\Longrightarrow}} \; [1] \; \circlearrowleft \; id \; .$$

  We use this instance for all examples, although *all theoretical results in this paper are valid for an arbitrary $\mathcal{D}$*.
- A category $\mathcal{E}$ with small hom-sets and all small limits. Usually, we work with the case when $\mathcal{E}$ is a category of (generalized) spaces like[2] **Top** (topological spaces), **Loc** (locales), **Cat**(**Top**) (topological categories), and **Cat** (small categories).

---

[1]See [5, 4] for discussions on these variations.

[2]A discussion of gestures in these categories, for the case of digraphs, is [1].

- A functor[3] $S : \mathcal{D}^{op} \longrightarrow \mathcal{E}$.

Let $\widehat{\mathcal{D}}$ be the category of presheaves on $\mathcal{D}$. Given a presheaf P of $\widehat{\mathcal{D}}$, **the object of P-gestures with respect to S** is the **cotensor product** $P \pitchfork S$, defined as the limit of the composite functor

$$\left(\int P \right)^{op} \xrightarrow{\pi^{op}} \mathcal{D}^{op} \xrightarrow{S} \mathcal{E}, \tag{1}$$

where $\int P$ denotes the category of elements of P.

Example II.1 (Topological gestures for digraphs). Let us suppose that $\mathcal{E} = \mathbf{Top}$. A functor $S : \mathcal{D}^{op} \longrightarrow \mathbf{Top}$ can be identified with a cuadruple (S_1, S_0, d_0, d_1), where S_1 and S_0 are topological spaces and d_0 and d_1 are continuous maps, by defining $S_1 = S([1])$, $S_0 = S([0])$, $d_0 = S(\epsilon_0)$, and $d_1 = S(\epsilon_1)$. We say that (S_1, S_0, d_0, d_1) is a **topological digraph**. Similarly, a presheaf Γ on $\mathcal{D}$ is just a *digraph* (A, V, t, h).

The cotensor product $\Gamma \pitchfork S$ is the limit in **Top** of the diagram consisting of for each $a \in A$ a copy of S_1, for each $z \in V$ a copy of S_0, a copy of d_0 whenever $z = t(a)$, and a copy of d_1 whenever $z = h(a)$. We compute this limit explicitly as the subspace of the product (with the Tychonoff topology)

$$\left(\prod_{a \in A} S_1 \right) \times \left(\prod_{z \in V} S_0 \right)$$

of all sequences $((c_a)_{a \in A}, (x_z)_{z \in V})$ satisfying $d_0(c_a) = x_{t(a)}$ and $d_1(c_a) = x_{h(a)}$ for each $a \in A$. We call these sequences *gestures*. We discuss the particular case of Mazzola's gestures and more examples in Section III. ♣

II.i. The cotensor adjunction and gestural presheaves

By dualizing[4] [11, Theorem I.5.2], the cotensor construction gives rise to an adjunction (left adjoint on the left)

$$\mathcal{E}(-, S) : \mathcal{E} \xrightleftharpoons{\hspace{1cm}} \left(\widehat{\mathcal{D}} \right)^{op} : - \pitchfork S \,,$$

with associated bijection

$$\mathcal{E}(E, P \pitchfork S) \cong \widehat{\mathcal{D}}(P, \mathcal{E}(E, S(-))), \tag{2}$$

natural in P and E. This bijection sends a natural transformation on the right-hand term, which can be identified with a cone on the functor in Equation (1) with vertex E, to the morphism $E \longrightarrow P \pitchfork S$ given by the universal property of the limit $P \pitchfork S$ of this functor.

We call a functor of the form $- \pitchfork S : \left(\widehat{\mathcal{D}} \right)^{op} \longrightarrow \mathcal{E}$ **gestural presheaf**.

Example II.2. Let us compute the action of a gestural presheaf on a morphism in the case of Example II.1. Consider two digraphs (presheaves) Γ and Γ' with $\Gamma = (A, V, t, h)$ and $\Gamma' = (A', V', t', h')$. A morphism of presheaves $\tau : \Gamma \longrightarrow \Gamma'$ can be identified with the *digraph morphism* (u, v), where $u = \tau_{[1]}$ and $v = \tau_{[0]}$. In particular, note that the category $\widehat{\mathcal{D}}$ is just the category of digraphs. Following the explicit presentation of the space of gestures in Example II.1, the continuous map $\tau \pitchfork S : \Gamma' \pitchfork S \longrightarrow \Gamma \pitchfork S$ sends the sequence $((c_a)_{a \in A'}, (x_z)_{z \in V'})$ to $((c_{u(a)})_{a \in A}, (x_{v(z)})_{z \in V})$. ♣

---

[3]We can also call it *presheaf on $\mathcal{D}$ with values in $\mathcal{E}$*.
[4]See [2, Section 3.4.1] and [2, Section 3.4.4] for more details.

Example II.3 (Individual gestures). The adjunction is useful to compute the elements of the object of gestures. Suppose we are in the situation of Example II.1. If we take E to be the point space $\{*\}$ and P as a digraph Γ, then the bijection in Equation (2) just says that there is a correspondence between the set $\Gamma \pitchfork S$ and the set of digraph morphisms from Γ to S regarded as a digraph (forget the topological structure). By definition, such a morphism (c, x) sends $a \in A$ to $c_a \in S_1$ and $z \in V$ to $x_z \in S_0$, and satisfies $d_0(c_a) = x_{t(a)}$ and $d_1(c_a) = x_{h(a)}$. This means that it is just a gesture sequence as in Example II.1. ♣

II.ii. Kan extension characterization

Let $\mathbf{y} : \mathcal{D} \longrightarrow \widehat{\mathcal{D}}$ be the Yoneda embedding. The restriction functor between functor categories

$$- \circ \mathbf{y}^{op} : \mathcal{E}^{(\widehat{\mathcal{D}})^{op}} \longrightarrow \mathcal{E}^{\mathcal{D}^{op}}$$

has a right adjoint R, with $R(S) = - \pitchfork S$, since the limit in Equation (1) exists for each S and P [10, Corollary X.3.2]. This means, by definition, that $- \pitchfork S$ *is the right Kan extension of S along* $\mathbf{y}^{op}$. Consequently, we have a bijection

$$\mathcal{E}^{(\widehat{\mathcal{D}})^{op}}(F, - \pitchfork S) \cong \mathcal{E}^{\mathcal{D}^{op}}(F \circ \mathbf{y}^{op}, S) \tag{3}$$

natural in F and S. The functor $\mathbf{y}^{op}_{(-)} \pitchfork S$ can be assumed to be S, so the counit component $\epsilon : \mathbf{y}^{op}_{(-)} \pitchfork S \longrightarrow S$ is the identity [10, Corollary X.3.3]. In this way, the bijection (3) sends a natural transformation $\sigma : F \longrightarrow - \pitchfork S$ to $\sigma_{\mathbf{y}^{op}} : F\mathbf{y}^{op} \longrightarrow \mathbf{y}^{op}_{(-)} \pitchfork S = S$. Conversely, given $\alpha : F \circ \mathbf{y}^{op} \longrightarrow S$, for each presheaf P in $\widehat{\mathcal{D}}$ it induces a cone[5] $\{\alpha_D F(p) : F(P) \longrightarrow S(D) \mid (D, p) \in \int P\}$ on the functor in Equation (1) and hence the component $\sigma_P : F(P) \longrightarrow P \pitchfork S$ of a natural transformation $\sigma : F \longrightarrow - \pitchfork S$.

   We show in Example V.1 that Equation (3) generalizes Mazzola's gestural Yoneda lemma [16, Theorem 39, p. 962]. Before, we need to define Mazzola's gestures on topological categories.

III. GENERALIZED MAZZOLA'S GESTURES

III.i. The functor of an object

Given an object C of $\mathcal{E}$ and a functor $T : \mathcal{D} \longrightarrow \mathcal{E}$ with all its images exponentiable in $\mathcal{E}$, we can construct the functor

$$C^T : \mathcal{D}^{op} \longrightarrow \mathcal{E} : D \xrightarrow{g} D' \mapsto C^{T(D')} \xrightarrow{C^{T(g)}} C^{T(D)},$$

which we denote by S_C. This construction generalizes Mazzola's topological digraph of a topological space and the categorical digraph of a topological category, as shown in the following examples.

Example III.1. Take $\mathcal{E} = \mathbf{Top}$. Consider the continuous endpoint inclusions $i_0, i_1 : \{*\} \longrightarrow I$ of the unit interval I in $\mathbb{R}$. They correspond to the functor $T : \mathcal{D} \longrightarrow \mathbf{Top}$ defined by $T([0]) = \{*\}$, $T([1]) = I$, $T(\epsilon_0) = i_0$, and $T(\epsilon_1) = i_1$. The images of T are (locally) compact Hausdorff and hence exponentiable in $\mathbf{Top}$. Given a topological space X, $S_X : \mathcal{D}^{op} \longrightarrow \mathbf{Top}$ corresponds to the *topological digraph* (X^I, X, e_0, e_1) *of X*, where the exponential X^I is the *function space* $\mathbf{Top}(I, X)$ of continuous paths in X, equipped with the compact-open topology [7, p. 558], X is isomorphic to the exponential $X^{\{*\}}$, and e_0 and e_1 are the continuous evaluations at 0 and 1. ♣

---

[5]Here we identify $p \in P(D)$ with its natural transformation $p : \mathbf{y}_D = \mathcal{D}(-, D) \longrightarrow P$ given by the Yoneda lemma.

Example III.2. Take $\mathcal{E}$ as the category of topological categories **Cat**(**Top**). There are topological functors $\mathbf{i}, \mathbf{j} : \mathbf{1} \longrightarrow \mathbb{I}$ from the final category to the topological category $\mathbb{I}$ of I, defined by $\mathbf{i}_0(*) = 0$, $\mathbf{j}_0(*) = 1$, $\mathbf{i}_1(*) = (0, 0)$, and $\mathbf{j}_1(*) = (1, 1)$. Once again, they amount to the functor $T : \mathcal{D} \longrightarrow$ **Top** defined by $T([0]) = \mathbf{1}$, $T([1]) = \mathbb{I}$, $T(\epsilon_0) = \mathbf{i}$, and $T(\epsilon_1) = \mathbf{j}$. In this case, $\mathbb{I}$ is exponentiable in **Cat**(**Top**) because its spaces of objects I, morphisms ∇, and composable morphisms are locally compact Hausdorff, that is, exponentiable in **Top**; see [2, Theorem 5.3.2].

If $\mathbb{K}$ is a topological category with spaces of objects and morphisms C_1 and C_0 respectively, then $S_{\mathbb{K}} : \mathcal{D}^{op} \longrightarrow$ **Cat**(**Top**) corresponds to the *categorical digraph* $(\mathbb{K}^{\mathbb{I}}, \mathbb{K}, \mathbf{e_0}, \mathbf{e_1})$ *of* $\mathbb{K}$, where $\mathbb{K}^{\mathbb{I}}$ is the category of all topological functors from $\mathbb{I}$ to $\mathbb{K}$ with its set of objects P_0 (that is, of topological functors) topologized as a subspace of $C_1^{\nabla} \times C_0^I$ and its set of morphisms P_1 (that is, of natural transformations) topologized as a subspace of $P_0 \times P_0 \times C_1^I$, and

$$\mathbf{e}_i : \mathbb{K}^{\mathbb{I}} \longrightarrow \mathbb{K} : F \xrightarrow{\tau} G \mapsto F(i) \xrightarrow{\tau_i} G(i)$$

for $i = 0, 1$. ♣

Example III.3 (General addresses)**.** Take $\mathcal{E}$ as the category of contravariant functors from **Cat**(**Top**) to itself. Given a functor $S : \mathcal{D}^{op} \longrightarrow$ **Cat**(**Top**) and a topological category $\mathbb{A}$, called *address*, we have a functor **Cat**(**Top**)$(\mathbb{A}, S(-)) : \mathcal{D}^{op} \longrightarrow$ **Cat**(**Top**). In fact, each image, which is of the form **Cat**(**Top**)$(\mathbb{A}, \mathbb{D})$, can be enriched with a topological category structure similar to that of the exponential $\mathbb{K}^{\mathbb{I}}$ in Example III.2, although it need not be an exponential in **Cat**(**Top**). By ranging $\mathbb{A}$ over all topological categories we obtain a functor

$$\textbf{Cat}(\textbf{Top})(-, S(-)) : \textbf{Cat}(\textbf{Top})^{op} \times \mathcal{D}^{op} \longrightarrow \textbf{Cat}(\textbf{Top})$$

or equivalently **Cat**(**Top**)$(-, S(-)) : \mathcal{D}^{op} \longrightarrow \mathcal{E}$. In the case when S is a categorical digraph $S_{\mathbb{K}}$ we denote this functor by $@S_{\mathbb{K}}$ and call $\mathbb{A}@S_{\mathbb{K}}$ *the* $\mathbb{A}$-*addressed categorical digraph of* $\mathbb{K}$, according to [16, p. 961]. ♣

III.ii. Gestures

Under the hypotheses of Section III.i, given an object C of $\mathcal{E}$ and a presheaf P on $\mathcal{D}$, we define the *object of P-gestures with body in C* as the cotensor product $P \pitchfork S_C$ in Section II. Though this limit can be hard to compute, Equation (2) gives us a simple characterization of its generalized elements, and in particular its points. In fact, for each object E of $\mathcal{E}$ we have the natural bijections (the second one given by the exponential adjunction)

$$\mathcal{E}(E, P \pitchfork S_C) \cong \widehat{\mathcal{D}}(P, \mathcal{E}(E, C^{T(-)})) \cong \widehat{\mathcal{D}}(P, \mathcal{E}(E \times T(-), C)),$$

which mean that *E-addressed elements* of $P \pitchfork S_C$ correspond to natural transformations from P to $\mathcal{E}(E \times T(-), C)$. In particular, if E is the final object of $\mathcal{E}$, then we obtain that the *points* of $P \pitchfork S_C$ correspond to natural transformations from P to $\mathcal{E}(T(-), C)$, the latter being the *underlying presheaf of* S_C. This leads us to define *a P-gesture with body in C* as a natural transformation $P \longrightarrow \mathcal{E}(T(-), C)$.

Example III.4. Let X be a topological space, T as in Example III.1, and Γ a digraph (A, V, t, h). The topological space of Γ-gestures with body in X coincides with the original definition $\Gamma@X$ in [15]. In fact, the limit defining the cotensor $\Gamma \pitchfork S_X$ in Example II.1 for the case of the topological digraph S_X, which is identified with the tuple (X^I, X, e_0, e_1) in Example III.1, is just that in [15, p. 31].
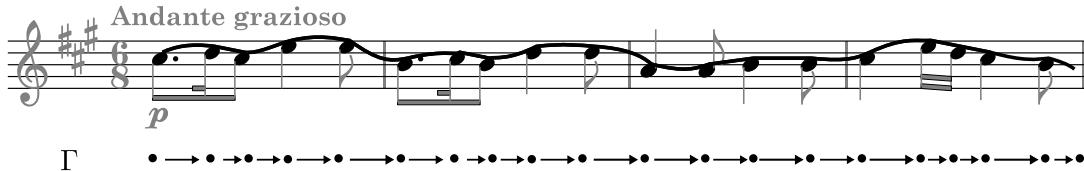
Figure 1: *The first phrase in Mozart's K. 331 as a gesture.*

On the other hand, the definition of an individual gesture as a natural transformation from Γ to $\mathbf{Top}(T(-), X)$ coincides with a digraph morphism from Γ to the underlying digraph of (X^I, X, e_0, e_1) in Example II.3 and [15, p. 29], and a sequence $((c_a)_{a \in A}, (x_z)_{z \in V})$ satisfying $c_a(0) = x_{t(a)}$ and $c_a(1) = x_{h(a)}$, as in Example II.1, where $c_a : I \longrightarrow X$ is a continuous map and $x_z \in X$. ♣

Example III.5 (The score space)**.** We can define gestures on musically meaningful spaces. Let us interpret the Euclidean space $\mathbb{R}^2$ as the *score space*, where a pair (t, p) denotes a sound event with pitch p that occurs at the time t. We choose the unities according to the situation. Here we use the quarter duration as time unity and identify the subset $\mathbb{Z}$ of $\mathbb{R}$ with the diatonic scale indicated by the key involved. For instance, the pair $(1/2, 0)$ denotes the pitch A4 occurring after an eight duration.

Consider the first phrase of Mozart's Piano Sonata[6] K. 331. The melodic contour can be regarded as a gesture in $\mathbb{R}^2$, see Figure 1. As we explain in Section 3, the melodic contour plays an important role in Mozart's variations of this phrase since they can be regarded as the result of transforming the original gesture with homotopies and a sheaf. ♣

Example III.6. Akin to Example III.4, the topological category of Γ-gestures with body in a topological category $\mathbb{K}$, based on the categorical digraph from Example III.2, coincides with $\Gamma@\mathbb{K}$, as defined in [14, Section 2.2].

In this case, an individual gesture is a digraph morphism from Γ to the underlying digraph of $(\mathbb{K}^{\mathbb{I}}, \mathbb{K}, \mathbf{e_0}, \mathbf{e_1})$, which only takes into account the objects[7] of $\mathbb{K}^{\mathbb{I}}$ (functors) and $\mathbb{K}$. They can be identified with sequences

$$((F_a)_{a \in A}, (C_z)_{z \in V})$$

satisfying $F_a(0) = C_{t(a)}$ and $F_a(1) = C_{h(a)}$, where $F_a : \mathbb{I} \longrightarrow \mathbb{K}$ is a topological functor and C_z is an object of $\mathbb{K}$. ♣
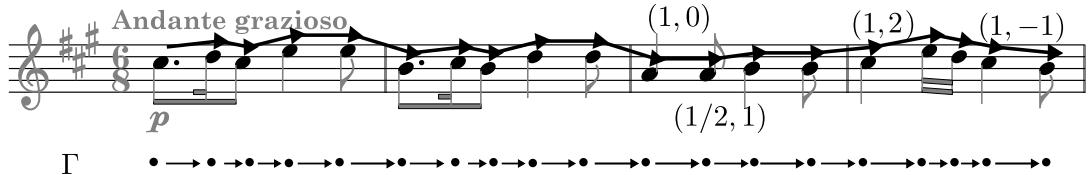
Example III.7. However, there are more abstract gestural notions for topological spaces with musical meaning, other than Mazzola's one. Consider the functor $S : \mathcal{D}^{op} \longrightarrow \mathbf{Top}$ corresponding to the *topological digraph of intervals of the score space* $\mathbb{R}^2$. The space of vertices is the score space $\mathbb{R}^2$, see Example III.5. The space of arrows is the subspace of $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ consisting of triples of the form $(x, y, y - x)$, where $y - x$ is the vector difference, that is, *the interval between x and y*. The tail and head are just the first and second product projections. In this case, a gesture is a sequence

$$\left((x_{t(a)}, x_{h(a)}, x_{h(a)} - x_{t(a)})_{a \in A}, (x_z)_{z \in V} \right),$$

where x_z is a sound event. A gesture of intervals between the sound events of the first phrase in Mozart's K. 331 corresponds to Figure 2. ♣

---

[6]The authors prepared the excerpts from [17] (theme) and the original manuscript (variation 4) at `https://mozart.oszk.hu/`.

[7]Note that, in this case, points are functors with domain the final category **1** and correspond to objects of the codomains.

Figure 2: *The first phrase in Mozart's K. 331 as a gesture of intervals. We only consider the third component (interval) of some arrows for simplicity.*

IV. THE FUNCTOR OF FUNCTORS OF AN OBJECT AND ITS LEFT ADJOINT

The construction in Section III.i induces a functor

$$S_{(-)} : \mathcal{E} \longrightarrow \mathcal{E}^{\mathcal{D}^{op}} : C \xrightarrow{f} C' \mapsto \left(f^{T(D)} : C^{T(D)} \longrightarrow C'^{T(D)} \right)_{D \in \mathcal{D}}. \tag{4}$$

We claim that this functor has a left adjoint. To prove this, first consider a natural transformation $\tau : S \longrightarrow S_C$, where $S, S_C : \mathcal{D}^{op} \longrightarrow \mathcal{E}$ are functors. This natural transformation τ is a family of morphisms $\left(\tau_D : S(D) \longrightarrow C^{T(D)} \right)_{D \in \mathcal{D}}$ making the following diagram commute for each morphism $g : D \longrightarrow D'$ of $\mathcal{D}$.

$$
\begin{array}{ccc}
S(D) & \xrightarrow{\tau_D} & C^{T(D)} \\
{\scriptstyle S(g)} \uparrow & & \uparrow {\scriptstyle C^{T(g)}} \\
S(D') & \xrightarrow{\tau_{D'}} & C^{T(D')}
\end{array}
$$

By transposing both composites across the exponential adjunction, the previous commutative diagram is equivalent to the following one, where $\widetilde{\tau_D}$ denotes the transpose of τ_D.

$$
\begin{array}{ccc}
S(D) \times T(D) & \xrightarrow{\widetilde{\tau_D}} & C \\
{\scriptstyle S(g) \times T(D)} \uparrow & & \uparrow {\scriptstyle \widetilde{\tau_{D'}}} \\
S(D') \times T(D) & \xrightarrow{S(D') \times T(g)} & S(D') \times T(D')
\end{array}
$$

But this is just a *wedge* from $S \times T$ to C. Thus, we have a bijection

$$\mathcal{E}^{\mathcal{D}^{op}}(S, S_C) \cong \text{Wedge}(S \times T, C). \tag{5}$$

In the case when $\mathcal{E}$ is small-cocomplete all coends exist [10, p. 224] and the definition of coend establishes a bijection

$$\text{Wedge}(S \times T, C) \cong \mathcal{E}\left(\int^D S(D) \times T(D), C \right). \tag{6}$$

To sum up, the bijections in Equations (5) and (6) amount to

$$\mathcal{E}^{\mathcal{D}^{op}}(S, S_C) \cong \mathcal{E}\left(\int^D S(D) \times T(D), C \right),$$

which can be shown to be natural in C and hence it determines an adjunction [10, Corollary IV.1.2]. We thus have the following theorem.

Theorem IV.1. *If $\mathcal{E}$ is small-cocomplete, then $S_{(-)}$ has as left adjoint the functor*

$$\int^D (-)(D) \times T(D) : \mathcal{E}^{\mathcal{D}^{op}} \longrightarrow \mathcal{E},$$

sending S to the coend $\int^D S(D) \times T(D)$.

We apply this theorem to gestures on the score space in Example VI.1.

V. A GESTURAL EMBEDDING

The Kan adjunction $- \circ \mathbf{y}^{op} \dashv R$ from Section II.ii restricts to an equivalence between $\mathcal{E}^{\mathcal{D}^{op}}$ and the full subcategory of $\mathcal{E}^{\widehat{\mathcal{D}}^{op}}$ of all functors $F : \left(\widehat{\mathcal{D}}\right)^{op} \longrightarrow \mathcal{E}$ that are naturally isomorphic to $- \pitchfork S$ for some S. This follows from [11, II.6.4], since the counit $\epsilon : R(-) \circ \mathbf{y}^{op} \longrightarrow id$ is the identity (Section II.ii), and for each S, $\eta_{-\pitchfork S} : - \pitchfork S \longrightarrow - \pitchfork (\mathbf{y}^{op}_{(-)} \pitchfork S)$ is the identity (check).

In particular, *the functor R is an embedding*. This means that for each pair $S, S' : \mathcal{E}^{\mathcal{D}^{op}} \longrightarrow \mathcal{E}$, R restricts to a bijection

$$\mathcal{E}^{\mathcal{D}^{op}}(S', S) \cong \mathcal{E}^{\widehat{\mathcal{D}}^{op}}(- \pitchfork S', - \pitchfork S). \tag{7}$$

Example V.1 (Gestural Yoneda for topological categories). Let $\mathcal{E}$, $\mathbb{K}$, and $@S_\mathbb{K}$ be as in Example III.3, where $S_\mathbb{K}$ is as in Example III.2. Let us denote by $@\mathbb{K}$ the gestural presheaf $- \pitchfork @S_\mathbb{K}$. By taking F as a functor naturally isomorphic to $- \pitchfork S$ in Equation (7), we obtain

$$\mathcal{E}^{\widehat{\mathcal{D}}^{op}}(@\mathbb{K}, F) \cong \mathcal{E}^{\widehat{\mathcal{D}}^{op}}(@\mathbb{K}, - \pitchfork S) \cong \mathcal{E}^{\mathcal{D}^{op}}(@S_\mathbb{K}, S).$$

Following Mazzola's terminology, $@S_\mathbb{K} = \overrightarrow{\mathbb{K}}$, F is a *limiting functor*, and $S = \overrightarrow{F}$. This is just the Yoneda lemma for gestures on topological categories [16, Theorem 39, p. 962]. ♣

Example V.2. Suppose that $\mathcal{E} = \textbf{Top}$. Given a natural transformation $\tau : S' \longrightarrow S$ of topological digraphs, let us compute the associated natural transformation $R(\tau) : - \pitchfork S' \longrightarrow - \pitchfork S$. If Γ is a digraph (A, V, t, h), then the continuous map $R(\tau)_\Gamma : \Gamma \pitchfork S' \longrightarrow \Gamma \pitchfork S$ sends a gesture sequence (Example II.1)

$$((c_a)_{a \in A}, (x_z)_{z \in V})$$

to

$$((\tau_{[1]}(c_a))_{a \in A}, (\tau_{[0]}(x_z))_{z \in V}).$$

♣

Example V.3 (Interplay between intervals (Lewin) and gestures (Mazzola) in the score). Let us assume the situation of Example V.2. Take S' as the topological digraph of intervals of the score (Example III.7) and S as the topological digraph of $\mathbb{R}^2$ (Examples III.4 and III.5).

First, we define a natural transformation τ from S' to S. The correspondence on arrows $S'([1]) \longrightarrow (\mathbb{R}^2)^I$ is the exponential transpose of the continuous map

$$S'([1]) \times I \longrightarrow \mathbb{R}^2 : ((x, y, y - x), s) \mapsto x + s(y - x)$$

and the correspondence on vertices is the identity. Hence the correspondence on arrows sends $(x, y, y - x)$ to the linear path with parametrization $\alpha(s) = x + s(y - x)$. This defines a natural transformation since $\alpha(0) = x = \pi_1(x, y, y - x)$ and $\alpha(1) = y = \pi_2(x, y, y - x)$. This ensures the

existence (Example V.2) of a natural transformation $R(\tau) : - \Cap S' \longrightarrow - \Cap S$ that *incarnates* every gesture of intervals as a gesture of line segments.

Conversely, we can define a natural transformation μ from S to S'. The correspondence on arrows sends a path c in $(\mathbb{R}^2)^I$ to $(c(0), c(1), c(1) - c(0))$. This is just the pointwise definition of the continuous map $(e_0, e_1, e_1 - e_0) : (\mathbb{R}^2)^I \longrightarrow S'([1])$, where e_0 and e_1 are the continuous evaluation maps (Example III.1). The correspondence on vertices is the identity again. We thus have $R(\mu)$, which transforms gestures into abstract intervals.

Note that $\mu\tau = id$ and hence $R(\mu)R(\tau) = id$, that is, $\Gamma \Cap S'$ is a *retract* of $\Gamma \Cap \mathbb{R}^2$ for each digraph Γ. ♣

Now, assume the hypotheses of Section III.i. By composing R with the functor $S_{(-)}$ from Equation (4), we obtain the functor

$$\mathcal{E} \xrightarrow{S_{(-)}} \mathcal{E}^{\mathcal{D}^{op}} \xrightarrow{R} \mathcal{E}^{\widehat{\mathcal{D}}^{op}}.$$

In the case when $T([0])$ is the final object,[8] $S_{(-)}$ is faithful because if $S_f = S_g$, where $f, g : C \longrightarrow C'$ are continuous maps, then in particular $f = S_f([0]) = S_g([0]) = g$. However, the functor $S_{(-)}$ need not be full as shown in the following example.

Example V.4. Let us assume the situation of Example III.4 and that $X = \mathbb{R}^2$. The functor $S_{\mathbb{R}^2}$ is just the topological digraph $((\mathbb{R}^2)^I, \mathbb{R}^2, e_0, e_1)$. We have an endomorphism (u, v) of this topological digraph where v is the constant map with value $(0, 0)$ and u is the constant map with value the circular loop with parametrization $(\cos(2\pi t) - 1, \sin(2\pi t))$ for $t \in I$. This endomorphism is not of the form S_f, otherwise $v = f$ and $u = f^I$, that is, u is the constant map with value the constant path on $(0, 0)$; a contradiction. ♣

In the case when $T([0])$ is the final object, we can correct this drawback by just restricting the morphisms of $\mathcal{E}^{\mathcal{D}^{op}}$ to the image of $S_{(-)}$ and then those of $\mathcal{E}^{\widehat{\mathcal{D}}^{op}}$ to the image of $R \circ S_{(-)}$. This implies that *if $T([0]) = \mathbf{1}$, then $R \circ S_{(-)}$ is an embedding of $\mathcal{E}$ into a subcategory of the category of gestural presheaves $\mathcal{E}^{\widehat{\mathcal{D}}^{op}}$.* This result applies to topological and categorical gestures (Examples III.4 and III.6).

VI. THE COVARIANT GESTURE FUNCTOR AND ITS LEFT ADJOINT

Let us suppose that $\mathcal{E}$ is small-cocomplete. The *covariant gesture functor $P \Cap -$* is the composite

$$\mathcal{E}^{\mathcal{D}^{op}} \xrightarrow{\mathcal{E}^{\pi_P}} \mathcal{E}^{(\int P)^{op}} \xrightarrow{\text{Lim}} \mathcal{E}.$$

We claim that both functors in this composite have a left adjoint, so the composite of these adjoints is the desired one. Certainly, the diagonal functor from [11, p. 21] is the left adjoint to Lim and, on the other hand, by [10, Corollary X.3.2], $\mathcal{E}^{\pi_P}$ has a left adjoint since $\mathcal{E}$ is cocomplete and both $\mathcal{D}$ and $\int P$ are small categories.

In turn, *Mazzola's gesture functor $P \Cap S_{(-)}$* is the composite

$$\mathcal{E} \xrightarrow{S_{(-)}} \mathcal{E}^{\mathcal{D}^{op}} \xrightarrow{P \Cap -} \mathcal{E}$$

and has a (composite) left adjoint since both functors have left adjoints.

In particular, $P \Cap S_{(-)}$ preserves limits. Next, we provide an example illustrating the situation.

---

[8] In a more general language, this means that T preserves the final object.

Example VI.1. Let us assume the situation of Example III.4. By preservation of the product $\mathbb{R} \times \mathbb{R}$, the topological space of gestures $\Gamma \pitchfork S_{\mathbb{R}^2}$ with skeleton Γ and body in $\mathbb{R}^2$ is homeomorphic to the product $(\Gamma \pitchfork S_{\mathbb{R}}) \times (\Gamma \pitchfork S_{\mathbb{R}})$. In Mazzola's notation: $\Gamma@\mathbb{R}^2 \cong (\Gamma@\mathbb{R}) \times (\Gamma@\mathbb{R})$.

In terms of the interpretation of $\mathbb{R}^2$ as the score space (Example III.5), the homeomorphism corresponds to the decomposition of gestures on the score as pairs of gestures where the first component is a time gesture and the second one a pitch gesture.   ♣

VII.   Gestural sheaves

Now we discuss the relation between gestures and sheaves in Grothendieck's sense. This section intends to show that given a functor $S : \mathcal{D}^{op} \longrightarrow \mathcal{E}$, the gestural presheaf $- \pitchfork S : \left(\widehat{\mathcal{D}}\right)^{op} \longrightarrow \mathcal{E}$ is a sheaf with values in $\mathcal{E}$ with respect to the canonical Grothendieck topology on $\widehat{\mathcal{D}}$.

First, we rewrite the characterization of sheaves in terms of equalizers within our language of cotensor products and gestures. Let $(\mathcal{C}, J)$ be a site. We say that a *presheaf $F : \mathcal{C}^{op} \longrightarrow \mathcal{E}$ with values in $\mathcal{E}$ is a sheaf* if and only if for each object C of $\mathcal{C}$ and each covering sieve R in $J(C)$ we have the identity

$$R \pitchfork F = F(C),$$

where the *cotensor product* $R \pitchfork F$ is defined as

$$\mathrm{Lim} \left(\left(\int R\right)^{op} \xrightarrow{\pi_R^{op}} \mathcal{C}^{op} \xrightarrow{F} \mathcal{E} \right).$$

By using the presentation of this limit as an equalizer [10, p. 113], the identity $R \pitchfork F = F(D)$ means that the following diagram, with appropriate arrows, is an equalizer, namely *that defining a sheaf with values in $\mathcal{E}$*; compare with [11, p. 122].

$$F(C) \longrightarrow \prod_{(f:D \to C) \in R} F(D) \rightrightarrows \prod_{\substack{(f:D \to C) \in R \\ m:D' \to D}} F(D')$$

Note also that the condition $R \pitchfork F = F(C)$ (for C ranging over $\mathcal{C}$) above just says that, for each object A of $\mathcal{E}$, the presheaf $\mathcal{E}(A, F(-)) : \mathcal{C}^{op} \longrightarrow$ **Set** is a sheaf. In other words, every matching family

$$\{x_f : A \longrightarrow F(D) \mid (f : D \to C) \in R\}$$

for R of generalized elements of F has a unique amalgamation $x : A \longrightarrow F(C)$; compare with [11, p. 121-122].

Second, we are interested in the case when $\mathcal{C}$ is the category of presheaves $\widehat{\mathcal{D}}$, so we need an appropriate Grothendieck topology on it.[9] The topology that we will use is the canonical one. Recall that [11, p. 126] a Grothendieck topology on $\mathcal{C}$ is *subcanonical* if all representable presheaves on $\mathcal{C}$ are sheaves, and that [6, II.2.5] the *canonical topology* on a category $\mathcal{C}$ is the greatest subcanonical topology. We can rewrite this definition by using the category of elements as follows. A topology J on $\mathcal{C}$ is subcanonical if and only if for each object C of $\mathcal{C}$ and each covering sieve R in $J(C)$ the identity

$$\mathrm{Colim} \left(\int R \xrightarrow{\pi_R} \mathcal{C} \right) = C$$

---

[9]There is a subtlety here: the category $\widehat{\mathcal{D}}$ is not small, so the definition of topology and site given in [11, p. 110] cannot be used. But this is easily corrected by *changing the universe*, as suggested in [6].

holds. In fact, this equality just says that, for each object C of $\mathcal{C}$, every matching family for R of elements of $\mathcal{C}(-, C)$ has a unique amalgamation. In the case when $\mathcal{C}$ is a category of presheaves, the canonical topology on $\mathcal{C}$ is that having as covering sieves all epimorphic ones. To prove this, note that epimorphic sieves define a topology in every category of presheaves (check), that every sieve of the canonical topology is epimorphic, and recall that, conversely, in a category of presheaves every epimorphic sieve is a sieve of the canonical topology by part 2) of [6, Proposition II.4.3]. In this way, in the case when $\mathcal{C} = \widehat{\mathcal{D}}$, the canonical topology is given by all sieves that are epimorphic; concretely, a sieve R on a presheaf P is an epimorphic family if and only if for each $D \in \mathcal{D}$ and each $a \in P(D)$ there is a natural transformation $\tau : P' \longrightarrow P$ in R and an element $x \in P'(D)$ such that $\tau_D(x) = a$.

Now we can prove the main result of this section.

Theorem VII.1. *Each gestural presheaf $- \pitchfork S : \left(\widehat{\mathcal{D}} \right)^{op} \longrightarrow \mathcal{E}$ is a sheaf with values in $\mathcal{E}$ for any subcanonical topology on $\widehat{\mathcal{D}}$, in particular for the canonical one consisting of all epimorphic sieves.*

Proof. Consider a subcanonical topology J on $\widehat{\mathcal{D}}$. According to the characterization of subcanonical topologies above, for each object P of $\widehat{\mathcal{D}}$ and each sieve R in $J(P)$ we have the identity

$$\text{Colim} \left(\int R \xrightarrow{\pi_R} \widehat{\mathcal{D}} \right) = P.$$

Moreover, since the gestural presheaf, as a right adjoint (Section II.i), transforms colimits in $\widehat{\mathcal{D}}$ into limits in $\mathcal{E}$, by applying $- \pitchfork S$ to the identity above, we obtain that

$$\text{Lim} \left(\left(\int R \right)^{op} \xrightarrow{\pi_R^{op}} \left(\widehat{\mathcal{D}} \right)^{op} \xrightarrow{-\pitchfork S} \mathcal{E} \right) = P \pitchfork S.$$

According to the characterization of sheaves above, this means that $- \pitchfork S$ is a sheaf. $\square$

VIII. The sheaf of topological gestures in Mozart and Beethoven

In particular, Theorem VII.1 says that the gestural presheaf

$$- \pitchfork S_X = -@X : \mathcal{D}^{op} \longrightarrow \textbf{Top}$$
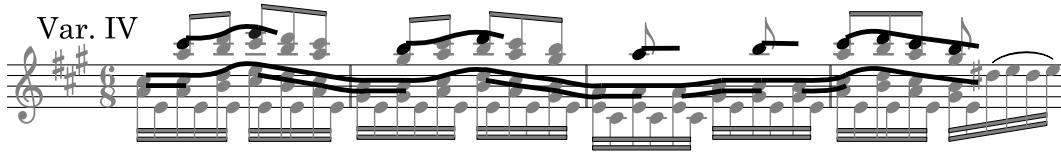
that results from the situation in Example III.4 is a sheaf.

Recall that (Example II.2) given a digraph morphism

$$\tau = (u, v) : \Gamma = (A, V, t, h) \longrightarrow \Gamma' = (A', V', t', h')$$

the sheaf $\tau@X$ sends a gesture sequence $((c_a)_{a \in A'}, (x_z)_{z \in V'})$ to $((c_{u(a)})_{a \in A}, (x_{v(z)})_{z \in V})$. In particular, if τ is a digraph inclusion (both u and v inclusions), then $\tau@X$ sends a Γ'-gesture to its *restriction* to Γ.

Let us consider the canonical topology on digraphs. An important example of epimorphic sieve on a digraph Γ, with $\Gamma = (A, V, t, h)$, is that generated by a *cover* $\{(A_i, V_i, t_i, h_i) \mid i \in \mathcal{I}\}$ of Γ by subdigraphs, which means that the inclusion pair is a digraph morphism from (A_i, V_i, t, h) to Γ for each $i \in \mathcal{I}$, $\bigcup_{i \in \mathcal{I}} A_i = A$, and $\bigcup_{i \in \mathcal{I}} V_i = V$. A coherent family of topological gestures for this sieve consists of a Σ-gesture g_Σ in X for each member Σ of the sieve. The coherence condition can be reduced to saying that g_{Γ_i} and g_{Γ_j} coincide on $(A_i \cap A_j, V_i \cap V_j, t, h)$, whenever $\Gamma_i = (A_i, V_i, t, h)$

Figure 3: *A global gesture in the fourth variation in Mozart's K. 331 (middle) from local melodic gestures (top and bottom). The harmonic tones are the black dots. We omit the digraphs, which are of the form $\bullet \to \bullet \cdots \bullet \to \bullet$, for simplicity.*

Figure 4: *Harmonic tones in the theme of Mozart's K. 331.*

and $\Gamma_j = (A_j, V_j, t, h)$. The sheaf property ensures that there is a unique *global* Γ-gesture g whose restriction to Γ_i is just the *local* gesture g_{Γ_i}.

In the following two musical examples we consider the case when Γ is a path digraph of the form $\bullet \to \bullet \cdots \bullet \to \bullet$ and each member of the cover is a path subdigraph of the same form. This kind of cover helps to reconstruct melodic contours from smaller fragments. Also, we assume that X is the score space $\mathbb{R}^2$ (Example III.5).

VIII.i. The fourth variation in Mozart's Piano Sonata K. 331
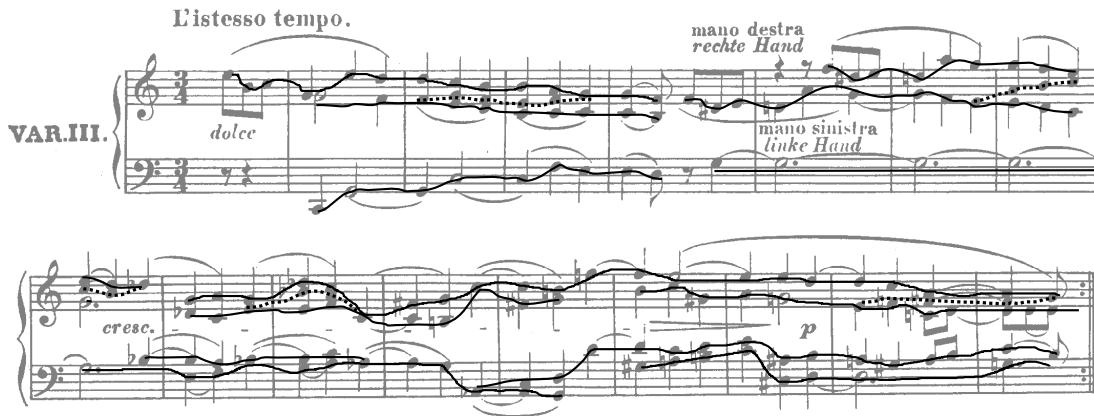
The construction of a global gesture from local ones can be observed in the fourth variation in Mozart's K. 331; see Figure 3. The melody can be reconstructed by following several steps. First, consider the original theme (Example III.5 and Figure 1) and take the harmonic tones as shown in Figure 4. Second, construct melodic fragments in each measure by joining the tones stepwise if necessary, following the contour gestures between harmonic tones, as shown in the top of Figure 3. Finally we coherently paste these fragments (transposed by an octave) by adding new melodic stepwise gestures and following a uniform rhythm (successive eights), as shown in the bottom of Figure 3. We thus obtain a global gesture (middle of Figure 3) that accompanies the melody by *continuation* of the original fragments.

This kind of gestural analysis could be useful to understand underlying topological processes in the theme with variations form, beyond transformational approaches. In this way, we highlight the plasticity of musical thinking, which we access mathematically thanks to the continuity of melodic contours obtained by continuation in the sheaf of gestures in the score. For instance, there is no symmetry (translation or reflection) between the melody in the first bar of the original theme (Figure 1) and the corresponding variation in Figure 3. However, *homotopies*[10] between paths, *preserving the harmonic tones*, seem to provide a valid explanation for the transit from the original fragment to its variation; see Figure 5. The preservation of some musical attribute, such as harmonic tones, seems to be essential for using this kind of explanation since in the score space $\mathbb{R}^2$ any two paths α, β can be transformed into each other by means of the homotopy defined by $(1-t)\alpha + t\beta$ for $t \in [0,1]$.

---

[10]A homotopy between two paths α, β in X is a path h in $X^{[0,1]}$ such that $h(0) = \alpha$ and $h(1) = \beta$. Similarly we can define homotopies between Γ-gestures in X as paths in the space $\Gamma \pitchfork S_X$.

Figure 5: *Transformation of the theme (first bar) in Mozart's K. 331 into the fourth variation (first bar) by means of homotopies between paths that preserve the harmonic tones. The black dots are the harmonic tones.*

Figure 6: *The third Diabelli variation in Beethoven's Op. 120; taken from [18]. The ramification, fusion, and connection of voices can be explained within the sheaf of topological gestures, contradicting the variety principle in counterpoint. We omit the digraphs again, which are of the form • → • ⋯ • → •, for simplicity.*

VIII.ii. The third Diabelli variation in Beethoven's Op. 120

The third variation in Beethoven's Op. 120, edition [18], is a very explicit example where the structure of a sheaf seems to be used intuitively. The following discussion refers to Figure 6. We analyze melodic contours by using topological gestures on the score space from Example III.5.

The two upper voices in measures 1-5 are intertwined by means of an intermediate voice, corresponding to the pointed line in measures 3-4. We thus have a global section connecting the two voices, which consists of the contour gesture ranging the middle voice in measure 2, the pointed line in measures 3-4, and the upper voice in measure 5. A similar phenomenon occurs in measures 7-9 and 16-17. In measures 10-11 the alto seems to split into two for a few quarters to fuse again.

The conceptual value of the sheaf involved seems to rely on the fact that it offers an explanation of a phenomenon that does not follow counterpoint rules: the fusion and connection of voices, which directly defies the *variety* principle in counterpoint [8, p. 21].

These contrapuntal ramification and fusion processes are considerably combined with symmetries in this case. The soprano theme in bars 1-5 reappears transposed in the alto in bars 5-9 with some intervallic variations. This imitation can also be regarded as homotopy equivalent to a literal transposition of the soprano theme. Then, in bars 6-9 the soprano forms an octave canon with the melody of the alto, which implies translation symmetries. In this context the voice fusion occurs (bars 7-9).

The topological tools considered (sheaves and homotopies) in the previous examples could be used to compose new music.

Glossary

In what follows $\mathcal{C}$ and $\mathcal{E}$ denote arbitrary categories.

Digraph A *digraph* Γ is a quadruple (A, V, t, h) such that A (*arrows*) and V (*vertices*) are sets and $t, h : A \longrightarrow V$ are functions called *tail* and *head*, respectively. A **digraph morphism** from (A, v, t, h) to (A', V', t', h') is a pair (u, v) of functions $u : A \longrightarrow A'$ and $v : V \longrightarrow V'$ such that $t'u = vt$ and $h'u = vh$.

Presheaf A *presheaf on a category* $\mathcal{C}$ is a contravariant functor from $\mathcal{C}$ to a suitable[11] category of sets. All presheaves on $\mathcal{C}$ form a category $\widehat{\mathcal{C}}$ whose morphisms are natural transformations between presheaves. *Examples*: 1. The **representable functors**, which are of the form $\mathcal{C}(-, C)$ for C object of $\mathcal{C}$, are presheaves on $\mathcal{C}$.

Category of elements Let P be a presheaf on $\mathcal{C}$. We denote by $\int P$ the *category of elements* of P. Its objects are of the form (C, p) where $p \in P(C)$ and C is an object of $\mathcal{C}$. A morphism from (C, p) to (C', p') is a morphism f of $\mathcal{C}$, from C to C', such that $P(f)(p') = p$. We denote by $\pi : \int P \longrightarrow \mathcal{C}$ the natural projection functor. This category is isomorphic to the comma category $\mathbf{y} \downarrow P$. Also, its opposite $\left(\int P \right)^{op}$ is isomorphic to $P \downarrow \mathbf{y}^{op}$.

Generalized element Given two objects C and D in a category $\mathcal{C}$, a C-**addressed element** of D is, by definition, a morphism $C \longrightarrow D$ of $\mathcal{C}$. This definition expresses a certain relativity of mathematical objects by means of the introduction of *observers* C for an object D.

Locally compact space A topological space X is *locally compact* if for each point $x \in X$ and each open neighborhood $U \ni x$, there is a compact neighborhood of x contained in U. In the case when X is a Hausdorff space, this definition is equivalent to saying that each point in X has a compact neighborhood. In this way, every compact Hausdorff space is locally compact.

Compact-open topology Let X and Y be topological spaces. The subbasic opens of the *compact-open topology* on the function space X^Y are those of the form $\{c : Y \longrightarrow X \text{ continuous} \mid c(K) \subseteq U\}$, where K is compact in Y and U is open in X. If Y is locally compact Hausdorff, then this makes X^Y an *exponential* in the category of topological spaces [7, p. 558].

Topological category A small category such that its sets of morphisms C_1 and objects C_0 are topological spaces and the identity $e : C_0 \longrightarrow C_1$, domain $d : C_1 \longrightarrow C_0$, codomain $c : C_1 \longrightarrow C_0$, and composition $m : E_2 = E_1 \times_{E_0} E_1 \longrightarrow E_1$ are continuous maps. It is usually written as the tuple (C_1, C_0, e, d, c, m). A **topological functor** $F : \mathbb{K} \longrightarrow \mathbb{D}$ between topological categories is a functor between the underlying categories such that the correspondences on objects and morphisms, namely F_0 and F_1, are continuous. It is usually written as the pair (F_1, F_0).

Topological category of the real unit interval[12] We define it by $\mathbb{I} = (\nabla, I, e', d', c', m')$, where

- $\nabla = \{(x, y) \in I \times I \mid x \leq y \text{ in } I\}$;
- $e' : I \longrightarrow \nabla$ is the diagonal, that is, $e'(x) = (x, x)$;
- $d', c' : \nabla \longrightarrow I$ are the first and second projections respectively;

---

[11]That is, the sets are the members of a Grothendieck universe [6, Exposé I].
[12]This category was introduced in [14].

- $E_2 = \nabla \times_I \nabla = \{((z,w),(x,y)) \in I^2 \times I^2 \mid x \leq y = z \leq w\}$, and $m' : E_2 \longrightarrow \nabla$ is defined by $m'((z,w),(x,y)) = (x,w)$; and
- we place the usual topology on I, ∇ is a subspace of $I \times I$ (product topology), and E_2 is a subspace of I^4, so e' (diagonal), d', c', and m' (projections) are continuous.

Yoneda lemma Given a presheaf P on $\mathcal{C}$ and an object of $\mathcal{C}$, it establishes a natural bijection between the sets of natural transformations from $\mathcal{C}(-,C)$ to P and $P(C)$. Explicitly, the bijection sends such a natural transformation τ to $\tau_C(id_C)$.

Yoneda embedding The functor $\mathbf{y} : \mathcal{C} \longrightarrow \widehat{\mathcal{C}}$ sending an object C to the representable functor $\mathcal{C}(-,C)$ and a morphism $f : C \longrightarrow D$ to the natural transformation defined by composition with f in each component. This functor is full and faithful by the Yoneda lemma.

Wedge Let $F : \mathcal{C}^{op} \times \mathcal{C} \longrightarrow \mathcal{E}$ be a functor and E an object of $\mathcal{E}$. A *wedge from F to E* consists of a family of morphisms (α_C) indexed by the objects of $\mathcal{C}$ such that the following diagram commutes for each morphism $f : C \longrightarrow C'$ of $\mathcal{C}$.

$$
\begin{array}{ccc}
F(C,C) & \xrightarrow{\;\alpha_C\;} & E \\
{\scriptstyle F(f,id)}\uparrow & & \uparrow{\scriptstyle \alpha_{C'}} \\
F(C',C) & \xrightarrow[F(id,f)]{} & F(C',C')
\end{array}
$$

Coend Let $F : \mathcal{C}^{op} \times \mathcal{C} \longrightarrow \mathcal{E}$ be a functor and E an object of $\mathcal{E}$. A *coend of F* is a pair (E,α), where E is an object of $\mathcal{E}$ and α is a wedge from F to E such that for every wedge β from F to E' there is a unique morphism $h : E \longrightarrow E'$ with $h\alpha_C = \beta_C$ for each object C of $\mathcal{C}$. We denote E by $\int^C F(C,C)$.

Sieve Let C be an object of a category $\mathcal{C}$. A *sieve* on C is a set of morphisms with codomain C that is closed under right composition. *Examples*: 1. The **maximal sieve** $t(C)$ consisting of all morphisms with codomain C. 2. The **sieve generated by a set** X of morphisms with codomain C, defined as the closure of X under right composition. 3. The **restriction sieve** $h^*(S)$ of a sieve S on C along a morphism $h : D \longrightarrow C$, defined as the set of all morphisms f with codomain D such that hf is in S.

Grothendieck topology A *Grothendieck topology* J on a category $\mathcal{C}$ consists of for each object C a set of *covering sieves* $J(C)$ such that i) the maximal sieve $t(C)$ is in $J(C)$, ii) if S is in $J(C)$, then all possible restriction sieves of S are covering sieves, and iii) if all possible restriction sieves of a given one S are covering sieves, then S is a covering sieve. *Examples*: 1. Let T be a topology (of a topological space) regarded as a category (category of a poset). The sieves generated by open coverings of opens in T are the covering sieves of a Grothendieck topology on T. 2. Consider the category $\widehat{\mathcal{C}}$ of presheaves on $\mathcal{C}$. The **epimorphic sieves** of a presheaf P are the sieves S on P such that for each object C of $\mathcal{C}$ the set of images $\{\mathrm{Im}(\tau_C) \mid \tau \in S\}$ covers $P(C)$.

Site Category with a Grothendieck topology $(\mathcal{C}, J)$.

Sheaf A presheaf F on a site $(\mathcal{C}, J)$ is a *sheaf* if for each object C of $\mathcal{C}$ and each covering sieve S in $J(C)$, given a family of *local sections* $\{x_f \mid f$ in S and $x_f \in F(\mathrm{dom}(f))\}$ such that $F(h)(x_f) = x_{fh}$ whenever the composite fh exists, there is a *unique* $x \in F(C)$ such that $F(f)(x) = x_f$ for each f in S.

In such a case we say that x is a *global section*. In words, we also express the sheaf condition above by saying that every *matching family* (x_f) of elements of F has a unique *amalgamation x*. *Examples*: given the site of the usual topology of $\mathbb{R}$ (respectively $\mathbb{C}$), the presheaf with $P(U)$ defined as the set of all (continuous or differentiable) functions defined on U with values in $\mathbb{R}$ (respectively $\mathbb{C}$) is a sheaf.

References

[1] Arias, J. 2017. Abstract gestures: a unifying concept in mathematical music theory. In: Mathematics and Computation in Music. Cham, Switzerland: Springer International Publishing, pp. 183–200.

[2] Arias, J. 2018. *Gesture Theory: Topos-Theoretic Perspectives and Philosophical Framework*. Dissertation (PhD in Mathematics), Universidad Nacional de Colombia. Available at: `https://repositorio.unal.edu.co/bitstream/handle/unal/63339/Thesisfinal.pdf?sequence=1&isAllowed=y`.

[3] Arias, J. 2018. Spaces of Gestures are Function Spaces. *Journal of Mathematics and Music*, 12/2, pp. 89–105. Available at: `https://doi.org/10.1080/17459737.2018.1496489`.

[4] Arias-Valero, J.; Lluis-Puebla, E. 2020. Simplicial Sets and Gestures: Mathematical Music Theory, Infinity-Categories, Homotopy, and Homology. *Preprint*, Available at: `https://jusearva.files.wordpress.com/2021/12/simpgest.pdf`.

[5] Arias-Valero, J.; Lluis-Puebla, E. 2021. A Conceptual Note on Gesture Theory. *Brazilian Journal of Music and Mathematics*, Available at: `https://musmat.org/en/musmat-journal/`.

[6] Artin, M.; Grothendieck, A.; Verdier, J-L. 1972. *Théorie des Topos et Cohomologie Étale des Schémas (SGA 4), Tome 1*. Berlin: Springer-Verlag.

[7] Escardó, M.; Heckmann, R.. 2001-2002. Topologies on spaces of continuous functions. *Topology Proceedings*, 26/2, 2001-2002, pp. 545-564. Available at: `http://topology.nipissingu.ca/tp/reprints/v26/tp26209.pdf`.

[8] Fux, J.; Mann, A.; Edmunds, J. 1965. *The Study of Counterpoint from Johann Joseph Fux's Gradus ad Parnassum*. New York: W. W. Norton.

[9] Lewin, D. 2007. *Generalized Musical Intervals and Transformations*. New York: Oxford University Press.

[10] Mac Lane, S. 1998. *Categories for the Working Mathematician*. New York: Springer-Verlag.

[11] Mac Lane, S.; Moerdijk, I. 1992. *Sheaves in Geometry and Logic*. New York: Springer-Verlag.

[12] Mannone, M. 2018. Knots, Music and DNA. *Journal of Creative Music Systems*, 2/2.

[13] Mannone, M.; Turchet, L. 2019. Shall we (math and) dance? In: Mathematics and Computation in Music. Cham, Switzerland: Springer International Publishing, pp. 84–97.

[14] Mazzola, G. 2009. Categorical Gestures, the Diamond Conjecture, Lewin's Question, and the Hammerklavier Sonata. *Journal of Mathematics and Music*, 3/1, pp. 31-58. Available at: `https://doi.org/10.1080/17459730902910480`.

[15] Mazzola, G.; Andreatta, M. 2007. Diagrams, Gestures and Formulae in Music. *Journal of Mathematics and Music*, 1/1, pp. 23-46. Available at: `http://dx.doi.org/10.1080/17459730601137716`.

[16] Mazzola, G. et al. 2018. *The Topos of Music III: Gestures*. Cham, Switzerland: Springer International Publishing.

[17] Mozart, W.; Broder, N. (ed.). 1960. *Sonatas and Fantasies for the Piano*. Bryn Mawr, PA: Theodore Presser Company.

[18] van Beethoven, L. 1862-1865. *Ludwig van Beethoven's Werke, Serie 17*. Leipzig: Breitkopf & Härtel.

Interview with Severine Neff

MusMat Research Group

28th June, 2021, Goleta, California, USA

I. Can you tell us about your research and its impact on your academic/theoretical work?

My musical–scholarly life has been devoted to a comprehensive study of the works of Arnold Schoenberg, that Janus-faced Modernist with deep roots in the past and open to a postmodernism to come. When I began my study in the late 1960s, North American music-theoretical thought about Schoenberg's music had become bifurcated into two traditions: an "historical" tradition grounded in the writings of Schoenberg's students and personal supporters, and a "compositional" tradition established by the then-cutting-edge theories of Milton Babbitt's "Princeton School" (see Figures 1–2). The "historical" tradition acknowledged Schoenberg as a music theorist, the "compositional" did not. I experienced both firsthand. As an undergraduate student at Columbia University in New York, I studied with Schoenberg's student, Patricia Carpenter. Several years later, after earning a master's degree in music theory at Yale University, I received a doctorate in music theory from Princeton University, having worked with Babbitt, Claudio Spies, and Peter Westergaard.

The Princeton School had produced their music–theoretical vision of Schoenberg's twelve–tone music by elucidating the taxonomy and properties of sets and the systematic consideration of surface groupings, invariance, and symmetry. Although Babbitt and his students—above all, David Lewin—extolled Schoenberg's music as an apogee of western musical thought, they dismissed his then-available body of theoretical writings as untenable. They held that unlike, for example, Schenker's theory of tonal music, Schoenberg's own theoretical teachings—writings based largely on tonal music—could not be formalized through mathematical models as shown most directly by Babbitt's student Michael Kassler. Equally important, Schoenberg's theories could not be explicated from the logic of analytic philosophy as conceived by Rudolf Carnap and others.

Schoenberg would have agreed with Babbitt and his circle. He had consistently stated, "I am more a composer than a theorist." However, he also held that composition was not in any deep sense related to science—it was an art. In Schoenberg's words, "While science has to demonstrate its problems perfectly and completely without omission and from every point of view, and therefore has to proceed systematically, logically ... art only presents a number of interesting cases and strives for perfection by the manner of their presentation."

In contrast to the Princeton School, the North American "historical" tradition did recognize Schoenberg as a theorist as early as the 1930s with articles by Schoenberg's students and colleagues. For example, the composer Henry Cowell integrated Schoenberg's theoretical views of the overtone series into his own concept of dissonance in New Music Resources; the bassoonist Adolph Weiss wrote about Schoenberg's teaching of traditional craft; translations of analytic essays on the earliest serial works by Erwin Stein appeared in Modern Music. In the 1940s Ernst Krenek penned his

text on 12-tone counterpoint. In the 1950s the composer Roger Sessions included aspects of Schoenberg's transformation chords in his Harmonic Practice, while Schoenberg's student Dika Newlin offered translations and an edition of Schoenberg's writings in *Style and Idea*. However, it was not until the 1980s, that Walter Frisch would refer to the theoretical work of Schoenberg in his *Brahms and the Principle of Developing Variation*. Martha Hyde would analyze Schoenberg's twelve-tone works through his working materials in Schoenberg's *Twelve–Tone Harmony: The Suite Op. 29 and the Compositional Sketches*.

Only Schoenberg's student and personal assistant, Patricia Carpenter, was willing to address if Schoenberg's thought could be regarded as a theory. She utilized not scientific criteria but those of aesthetics. She believed that Schoenberg's thought about music is characterized by a coherence and consistency claiming attention as a theory. His was primarily a theory of art and the work of art, and in that sense, an aesthetic theory. On that basis, she understood his specifically technical theory for the musical work.

As a young scholar, having published in 1981, in *Perspectives of New Music* on symmetrical relations in Schoenberg's Second String Quartet, Op. 10, and in 1984, in *Theory and Practice* on the implications of Schoenberg's notion of *Grundgestalt* in the First String Quartet, I began to be convinced that the "historical" and "compositional" camps of Schoenberg study could both require a complete collection of Schoenberg's music–theoretical writings in their German original and in English translation. I believed that certain passages in these writings might be suggestive both to those seeking to formalize structural aspects of Schoenberg's music or to those studying his work in an historical/philosophical context.

With this plan in mind, I ordered xerox copies of his theoretical teachings and writings from the Arnold Schoenberg Institute (then in Los Angeles) and studied his two main theoretical treatises, one written in Vienna, Austria, entitled *Coherence, Counterpoint, Instrumentation, Instruction in Form*, the other written in New York City, Chautauqua, New York, and Los Angeles, entitled *The Musical Idea and the Logic, Technique, and Art of Its Presentation*.

The earlier manuscript consisted of outlines, reworkings of preliminary definitions, brief discursive commentaries, and several series of questions left largely unanswered. Yet despite its fragmentary nature, I discovered that it was a treasure trove of information: it included Schoenberg's first extended discussion of his technique of developing variation and his first explanation of his principle of orchestration based on a composition's structure, which he called "inventing for the orchestra." His brief statements for *Counterpoint*, turned Fuxian rules on their head: after first writing a traditionally conceived cantus in whole notes and adding another line in half notes, he goes on to reverse the process—an original line in half notes with added whole notes. There was no doubt in my mind these texts had to appear in print.

II. Can you tell us about your work with Patricia Carpenter?

As I worked on the text of *Coherence, Counterpoint, Instrumentation, Instruction in Form*, I unexpectedly met a friend from my undergraduate years at Columbia University, William Germano, then Editor–in–Chief of Columbia University Press. We were on a New York City bus, and we began a lively conversation. He had only recently heard about the existence of Schoenberg's large manuscript *The Musical Idea and the Logic, Technique, and Art of Its Presentation*. As Bill got off the bus in front of the Ansonia Hotel at West 72nd Street, he shouted in a rather loud voice, "I want to give you a contract to do an edition of Schoenberg's manuscript! I'll be in touch." When I told Professor Carpenter about this meeting, she was curious where Schoenberg had written the manuscript. Years later we learned that Schoenberg had penned three-quarters of it at the Ansonia Hotel on West 72nd Street!

Given her profound knowledge of Schoenberg's compositional work, I immediately invited Professor Carpenter to join me on the project. She had studied with Schoenberg from 1942 to 1944 as an undergraduate at UCLA and then privately from 1945 to 1949. From 1947 to the spring of 1949, she acted as a personal assistant helping him edit the English text of his counterpoint book, typing out certain of its texts from dictations he had made on his Webster wire-recorder.[1] After moving from California via Boston to New York in 1949, Carpenter was accepted into a master's program in music composition at Columbia University. However, she turned to studying philosophy with the aesthetician Albert Hofstader and musicology with Paul Henry Lang. In 1971 she received a doctorate from the Faculty of Philosophy at Columbia with a thesis entitled "The Janus–Aspect of Fugue: An Essay in the Phenomenology of Music Form." Her major works dealt with aesthetic topics such as the philosophical nature and history of the "musical work" and with music-theoretical subjects such as Schoenberg's study of tonal music (see Figure 3 for her bibliography).

The Musical Idea and the Logic, Technique and Art of Its Presentation was a more philosophically oriented work than the earlier *Coherence, Counterpoint, Instrumentation, Instruction in Form*. Schoenberg's "musical idea," was a theoretical concept encompassing multiple aspects of the musical composition: first it designated the "first thought" of a musical work: in Schoenberg's words, "an indefinable space, resounding and in motion; a form shaped by its own characteristic relationships; a sense of masses in motion, their design as ineffable as it is incomparable."[2] In this sense the "idea" concerned something intuitively perceived but not rationally comprehended, not yet expressed to the outside world as a phenomenon, but rather known to the composer alone.

As soon as clear rhythms and articulations could be identified, Schoenberg understood the composition as consciously perceived in time, containing both stable and contradictory, unstable elements creating unrest, setting the work into temporal motion. For example, the first cello C♯ in Beethoven's "Eroica" contradicts the opening E♭ triad. Unlike most theorists, Schoenberg would consider such a single note to be a motive. He believed that the motion generated by such a "problem" could result in a work analogous to a "living being," a tradition of thought deriving from Aristotle and Plotinus and re-interpreted by Goethe. In my 1993 article, "Schoenberg and Goethe: Organicism and Analysis," I explore this idea, leading to an understanding of the of the tonal work as a field of contradictory forces, eventually fusing at a work's conclusion.

In a post-tonal work, the field of such forces are often vaguer—the wholeness must be discovered by the listener who both consciously and unconsciously comprehends the related sounds as they unfold in time. Schoenberg sees this mental synthesis as one related to a common experience: "Music is only understood when one goes away singing it and only loved when one falls asleep with it in one's head and finds it still there on waking up the next morning."[3]

Schoenberg's texts *Coherence, Counterpoint, Instrumentation, Instruction in Form* and *The Music Idea and the Logic, Technique, and Art of Its Presentation* have clearly impacted American music theory. Recently, in 2014, Jack Boss has reinterpreted "problems" of tonal works in a twelve-tone context in his now–classic text, *Schoenberg's Twelve-tone Music: Symmetry and the Musical Idea*. Matthew Arndt's 2019 article "Form-Function-Content" in Music Theory Spectrum critiqued William Caplin's work on musical form in light of Schoenberg's concept of tonal functions. Zachary Bernstein has just published a book entitled *Thinking In and About Music: Analytical Reflections on Milton Babbitt's Music and Thought*, showing the influence of Schoenberg's Goethean sense of organicism on the music of Milton Babbitt. For me, both Boss's and in Bernstein's books bring

---

[1]For a demonstration of the 1940s Webster wire-recorder, see https://www.youtube.com/watch?v=7Y6XLETWbqM.

[2]See Willi Reich, *Arnold Schönberg, Der konservative Revolutionär* (Vienna: Fritz Molden Verlag, 1968), p. 302.

[3]*Style and Idea: The Selected Writings of Arnold Schoenberg*, Leonard Stein, ed., Leo Black, trans.(New York: St. Martin's Press, 1975), p. 180.

my work full–circle—the tenets of the "historical" tradition of Schoenberg research have now connected with those of the "compositional."

After completing these books, my research has been devoted to an ongoing effort to understand the bridge between the extended tonal language of Schoenberg's youth and his atonal paradigms of sound. In 2006 I published a book entitled *Arnold Schoenberg's Second String Quartet in F-sharp Minor, Opus 10: A Norton Critical Score*, engaging the work through the paradox of reading its movement as fluctuating between the tonic key, F–sharp minor and the key of the flat–one, F major, a tonic that is not a tonic. Schoenberg was the first to mention such a problem but never wrote extensively about it.[4] Similarly, I engaged the paradoxical issue of how a quotation of popular music can be a structural focal point of a complex, chromatic work, and how a final movement freed from any tonal constraints, can end on an emphatic triad. In addition to substantial theoretical material, this book also contains biographical and cultural information surrounding the Second Quartet, which recently had significant influence on a documentary film on the Second Quartet, "Through the Darkness," produced and written by Hélan Warshaw, will appearing on public television in Austria, Sweden, and Finland during September 2021.

In addition, my published book chapters and articles have dealt with the impact of Schoenberg's theories on the work of his American Experimentalist students—as in John Cage's study's use of fugue in Second Construction in Metal and the relation of Lou Harrison's *Schoenbergiana* to Schoenberg's *Ode to Napoleon*. I have scrutinized Schoenberg's later handling of tonality in Second Chamber Symphony, Op. 38, a piece begun in 1906 and finished in 1939. In *Schenker Traditions: A Viennese School of Music Theory and its International Dissemination*, I have addressed differences between Schoenberg and Schenker's approach to the organic artwork. Another essay discusses Schoenberg's theories in relation to René Leibowitz's recorded performance of *The Rite of Spring*; it appeared in *The Rite of Spring at 100*, a book I co-edited with Professors Maureen Carr and Gretchen Horlacher, which was given the Ruth Solie Award of the American Musicological Society.

While working in the archive at the Arnold Schönberg Center in Vienna, Austria, I further found two previously unknown pieces written by Schoenberg: one an untitled, incomplete fugue expressing his emotions about the horrors of *Kristallnacht*. The work was played in New York at YIVO, The Center for Jewish Historical and Cultural Studies, by the Grammy-award honoree, pianist David Holzman. The other, "My Horses Ain't Hungry," was an incomplete arrangement of an Appalachian folksong, completed by the composer Allen Anderson and performed by Professor Susan Klebanow and the Chamber Singers at the University of North Carolina at Chapel Hill.

III. What is Schoenberg's role in your current work?

At the present time I am General Editor of the Oxford University Press, nine- volume series, "Schoenberg in Words" with the Schoenberg scholar Professor Sabine Feisst. So far, we have published *Schoenberg's Program Notes and Analyses* edited by J. Daniel Jenkins; a new edition of *Models for Beginners in Composition* by Gordon Root; *Schoenberg's Early Correspondence* edited by Ethan Haimo and Sabine Feisst; *Correspondence with American Composers*, edited by Sabine Feisst; and *Schoenberg's Correspondence with Alma Mahler* edited by Elizabeth Keathley and Marilyn McCoy.

My contribution to the series, entitled *Schoenberg on Counterpoint*, will offer a revealing explication of Schoenberg's understanding of techniques and forms associated with contrapuntal craft. The manuscripts in the text are formally and topically diverse, ranging from aphorisms to a 130–page book draft; disparate in subject matter, reaching from definitions of counterpoint to

---

[4]See Schoenberg's chart in Arnold Schoenberg, *Structural Functions of Harmony*, rev. ed. (New York: W. W. Norton & Co., 1969), p. 38.

philosophical musings on beauty; and distinct in presentation, from student-copied class handouts to Schoenberg's hand-copied scores of school compositions.

Part I of this edition presents a lengthy introduction positioning these teachings and writings within the framework of Austro-German contrapuntal study in the late-eighteenth, nineteenth- and early twentieth centuries and contrasting their content with Schoenberg's self-proclaimed "new method" of teaching, founded practically on his belief in the intrinsic unity of the subdisciplines of compositional craft and theoretically on the premise that even the tiniest example of species counterpoint or a complex school fugue must be regarded as a "little composition" emerging organically from the materials of a basic configuration, whether a cantus firmus without motives, an original, motivic "independent voice," a chorale melody, or the opening contrapuntal combination/*Grundgestalt* of complex canons or fugues. Schoenberg maintained that his recommended methods of study opened a gateway for the individual expression of ideas in music for composers, performers, theorists, or musicologists.

Part II chronologically presents the musical examples destined for inclusion in Schoenberg's eight attempted book projects. the contents of the first one, *Composing with Independent Voices*, spilled over into the second project, simply entitled *Counterpoint*. Three more projects led to his final, most extended, work on counterpoint, entitled *Preliminary Exercises* (1942–50). Here it is offered for the first time in its previously unpublished, second draft (1943–50)—the last version on which Schoenberg personally worked. *Preliminary Exercises* (1943–50) was to be the first of a three-volume set called *Counterpoint: Preliminary Exercises*, *Contrapuntal Composition*, and *Counterpoint in Homophonic Music*. Ca. 1947 Schoenberg proposed a final Book Project entitled *Bach's Counterpoint* in outline form, but it is filled with a multitude of musical examples, summarizing Schoenberg's hearing of Bach.

Part III of this volume offers additional commentaries not specifically slated for inclusion in a book project. Here they are divided into six topical areas: definitions and descriptions of "counterpoint;" the "musical idea" as understood in contrapuntal contexts; canon; traditional versus contemporary counterpoint; commentary on counterpoint in the work of other composers, scholars, and journalists; and observations on techniques in the works of Johann Sebastian Bach.

The book is accompanied by a website containing scores and midi files of several school compositions by Schoenberg and ca. 500 musical examples of various forms of counterpoint. I expect it to be done by the 150th birthday celebration for Arnold Schoenberg in 2024.

IV. What did you think of the MusMat conference on Schoenberg and mathematics held this year?

Your conference, attractively advertised with a logo having multi-colored squares forming hexachords and twelve-tone sets, was a valuable one. In general, I found it fascinating that Schoenberg's working out of a motive in developing variation could be translated into mathematics in so many ways (i.e., see the presentations of Carlos Almada, Edgardo Rodriguez and Alejandro Martinez, and Cecilia Saraiva). I also was pleased to learn more history and theory concerning Josef Hauer's music and thought (i.e., in = the lectures of Julio Herrlein and Dominik Sedivy). The concerts featuring Brazilian compositions were engaging for we rarely hear this music in the United States——thus, I appreciated the works of Carlos Amada, Vinicius Ramos Braga, Rodrigo Marconi, and Liduino Pitombeira.

Currently, much American music theory focuses on the cognitive or pedagogical aspects of tonal (or modal) music. I believe that the post-tonal sounds need more mathematical attention; for example, the study of algorithmic composition is important for the field of theory. As editor–in-chief of *Music Theory Spectrum*, I published Robert Wannamaker's "Rhythmicon Relationships,

Farey Sequences, and James Tenney's Spectral CANON for CONLON Nancarrow (1974)." The article was well received. In Schoenberg studies, I also would like to see more work on young composers like the New Yorker Christopher Cerrone, who sees Schoenberg's music as closely related to aspects of his own twenty–first-century compositions.

V. In your opinion, which is Schoenberg's most important contribution to music theory?

Schoenberg's signature theoretical concepts are intimately related around his personal conception of the organic artwork. Monotonality and *Grundgestalt* are central, but they are inseparable from his understanding of the motive and its atomization into elements and features, sentence forms, liquidation, neutralization, the compositional problem, and developing variation in homophonic music versus unfolding or unraveling [*Abwicklung*] in forms of complex canon and fugue.[5]

Interestingly, Schoenberg believed that re-invented forms of these concepts were especially valuable "in reading the future from the past."[6] Thus, his general description of term *Grundgestalt*—"that to which all is traced back"—easily morphs the description of a tonal theme into one of a twelve-tone set; analogously "developing variation and "unraveling," sentence forms, liquidation, and "compositional problems" used in interpreting the tonal works of Bach, Beethoven, and Brahms, can also be employed in novel ways to understand aspects of atonal or twelve-tone music. This transference of a single vocabulary from one music to another is virtually unique, and at the crux of Schoenberg's contributions.

VI. Which is your favorite piece among his compositions?

It is a tossup between the String Quartet No. 2 in F–Sharp Minor, Op. 10 (1907–08), and the unfinished oratorio *Die Jakobsleiter* [Jacob's Ladder] (1917–22, 1944). At the Quartet's close, a triad literally frees itself from tonality as it ascends into the atmosphere of "other planets." At the end of Part I of *Die Jakobsleiter*, two souls freed from life sing higher and higher until they reach the stratosphere, thus completing their journey to the heavenly sphere of God and the angels.

In this age of Covid confinement, it is especially wonderful to experience the freedom of finding new realms.

---

[5]For references to all these concepts in Schoenberg's writings, see the "Concordance of Terms" in Arnold Schoenberg, *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, Patricia Carpenter and Severine Neff, eds. With a New Foreword by Walter Frisch.

[6]See Arnold Schoenberg, *Theory of Harmony*, Roy E. Carter (Berkeley, California: University of California Press, 1978), 29.

Year	The American Historical Tradition: Selected Teachings and Writings of Schoenberg and his students (asterisked) and their disciples; Schoenberg manuscript study; Patricia Carpenter's aesthetic theory	The American Theoretical/Analytic Tradition: Selected Studies of Schoenberg's Music by Milton Babbitt, David Lewin, Allen Forte, and their disciples
1920s 1930s	1928 Hugo Leichtentritt, "Schönberg and Tonality," Modern Music 5/4 1930 *Erwin Stein, "Schoenberg's New Structural Form," Modern Music 7/4 translated by Hans Keller 1930 *Henry Cowell, New Music Resources 1932 *Adolph Weiss, "The Lyceum of Schoenberg," Modern Music 9/3 1934 Arnold Schoenberg, "Problems of Harmony," Modern Music translated by *Adolph Weiss 1936 Richard Hill, "Schoenberg's Tone-Rows and Music of the Future," The Musical Quarterly XXII/I	
1940s	1940 Ernst Krenek, Studies in Counterpoint: Based on the Twelve-Tone Technique 1945 Karl Eschman, Changing Forms in Modern Music 1947 Marion Bauer, Twentieth-Century Music 1948 Arnold Schoenberg, Theory of Harmony (abridged version) translated by Robert Adams 1949 René Leibowitz, Schönberg and His School translated by *Dika Newlin	1946 Milton Babbitt "The Function of Set Structure in the Twelve-Tone System," PhD diss., Princeton University (degree given in 1990)
1950s	1950 Arnold Schoenberg, Style and Idea (abridged version) translated by *Dika Newlin 1951 Rudolf Réti, The Thematic Process in Music 1951 Rogers Sessions, Harmonic Practice 1952 Roger Sessions, "Some Notes on Schoenberg and the 'Method of Composing with Twelve Tones," The Score and I.M.A. Magazine 6 1952 *Alban Berg, "Why is Schoenberg's Music So Hard to Understand?" translated by Anton Swarkowsky and Josef Lederer, Music Review 13 1953 *Erwin Stein, Orpheus in New Guises translated by Hans Keller 1954 *Josef Rufer, Composition with Twelve Tones translated by Humphrey Searle 1954 Arnold Schoenberg, Structural Functions of Harmony edited by Humphrey Searle 1955 George Rochberg, The Hexachord and its Relation to the Twelve-Tone Row 1958 Rudolf Réti, Tonality, Atonality, Pantonality 1959 *Josef Rufer, The Works of Arnold Schoenberg translated by *Dika Newlin	1950 Milton Babbitt, "Review of René Leibowitz, Schönberg and His School translated by *Dika Newlin" Journal of the American Musicological Society 3/1 1955 Milton Babbitt, "Some Aspects of Twelve-Tone Composition," The Score and I.M.A. Magazine 12

Figure 1: *Two traditions of studying Schoenberg in the United States*

Year	The American Historical Tradition: Selected Teachings and Writings of Schoenberg and his students (asterisked) and disciples	The American Theoretical/ Analytic Tradition: Selected Studies of Schoenberg's Music largely influenced by Milton Babbitt, David Lewin, Allen Forte, and their disciples
1960s	**1960** Arnold Schoenberg, "The Orchestral Variations, Op. 31: A Radio Talk," *The Score and I.M.A. Magazine* 27 **1962** *Patricia Carpenter, "The Piano Music of Arnold Schoenberg," *Piano Quarterly* 41 **1963** Arnold Schoenberg, *Preliminary Exercises in Counterpoint,* edited by *Leonard Stein [**1966** Arnold Schoenberg, *Sämtliche Werke {Complete Works},* edited by *Josef Rufer et al.] **1967** Arnold Schoenberg, *Fundamentals of Musical Composition,* edited by *Gerald Strang **1967.** Réti, Rudolf, *Thematic Patterns in Sonatas of Beethoven,* edited by Deryck Cooke.	**1960** David Lewin, "The Intervallic Content of a Collection of Notes: An Application to Schoenberg's Hexachordal Pieces," *Journal of Music Theory* 4 **1960** Milton Babbitt, "Twelve-Tone Invariants as Compositional Determinants," *The Musical Quarterly* 46 **1961** Milton Babbitt, "Set Structure as a Compositional Determinant," *Journal of Music Theory* 5/1 **1962** George Perle, *Serial Composition and Atonality* **1962/63** David Lewin, "A Theory of Segmental Association in Twelve-tone Music," *Perspectives of New Music* 1, No.1 **1963** George Perle, "Babbitt, Lewin, Schoenberg: A Critique," *Perspectives of New Music* 2/1 **1963** Milton Babbitt, "Reply to George Perle," *Perspectives of New Music* 2/2 **1965** Milton Babbitt, The Structure and Function of Music Theory," *College Music Symposium* 5 **1967** Michael Kassler, "Towards a Theory That is the Twelve-Tone Class System," *Perspectives of New Music* 5/2 **1967** David Lewin, "A Study of Hexachordal Levels in Schoenberg's Violin Fantasy," *Perspectives of New Music* 6/1 **1967** David Lewin, "*Moses und Aron*: Some General Remarks and Analytic Notes for Act I, Scene 1," *Perspectives of New Music* 6/1 **1968** David Lewin, "Inversional Balance as an Organizing Force in Schoenberg's Music and Thought." *Perspectives of New Music* 6/2

Figure 2: *Two traditions of studying Schoenberg in the United States (cont.).*

1962 "The Piano Music of Arnold Schoenberg I & II," *The Piano Quarterly* 41–42: 26–30, 24–29.

1965 "Musical Form Regained," *Journal of Philosophy* 62, no.2: 36–48.

1966 "On the Meaning of Music," In *Art and Philosophy: A Symposium*, ed. Sydney Hook, New York: New York University Press, 289–306.

1967 "The Musical Object," *Current Musicology* 5: 56–87.

1972 "The Janus–Aspect of Fugue: An Essay in the Phenomenolgy of Musical Form," PhD. diss., Columbia University.

1973 "Tonal Coherence in a Motet of Dufay," *Journal of Music Theory* 17, no. 1: 2–65.

1983 "*Grundgestalt* as Tonal Function," *Music Theory Spectrum* 5: 15–38.

1984 "Musical Form and Musical Idea: Reflections on a Theme of Schoenberg, Hanslick, and Kant," In *Music and Civilization: Essays in Honor of Paul Henry Lang*, eds. Edmond Strainchamps and Maria Rika Maniates, in collaboration with Christopher Hatch, New York: W. W. Norton & Co., 394–427.

1988 "A Problem in Organic Form: Schoenberg's Tonal Body," *Theory and Practice* 13: 31–63.

1988 "Aspects of Musical Space," *Explorations in Music, the Arts and Ideas, Essays in Honor of Leonard B. Meyer*, eds. Ruth Solie and Eugene Narmour, Stuyvesant, New York: Pendragon Press, 341–74.

1991 "Music Theory and Aesthetic Form," *Studies in Music from the University of Western Ontario*: 21–47.

1993 "Arnheim and the Teaching of Music," *Journal of Aesthetic Education* 27, no. 4: 105–114.

1993 "Review: William Thomson, *Schoenberg's Error*," *Music Theory Spectrum* 15: 286–99.

1995 With Severine Neff, *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, New York: Columbia University Press.

1997 With Severine Neff, "Schoenberg's Philosophy of Composition: Thoughts on the Musical Idea and Its Presentation," In *Constructive Dissonance: Schoenberg and Transformations of Twentieth Century Culture*, eds. Julianne Brand and Christopher Hailey, Berkeley: University of California Press, 146–59.

1997 "Tonality: A Conflict of Forces," In *Music Theory in Concept and Practice*, eds. James Baker, David Beach, Jonathan Bernard. Rochester: University of Rochester Press, 97–129.

1998 "Schoenberg's Philosophy of Composition" In *The Schoenberg Companion*, ed. Walter B. Bailey, Westport, CT: Greenwood Publishing Group, 209–22.

2006 With Severine Neff, *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, Paperback ed. with an Introduction by Walter Frisch, Bloomington, Indiana: Indiana University Press.

2015 With Severine Neff, *The Musical Idea and the Logic, Technique, and Art of Its Presentation*, Chinese trans. by Ping Jin, Beijing: Central Conservatory of Beijing Press.

Figure 3: *The publications of Patricia Carpenter (1923–2000).*

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

David Alves
Associate Director of Undergraduate Education

Marcelo Jardim
Director of the Cultural Artistic Sector

Maria José Bernardes Di Cavalcanti
Director of Public Outreach
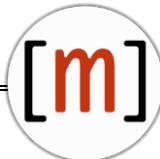
João Vidal
Coordinator of the Graduate Program in Music

Aloysio Fagerlande
Coordinator of the Graduate Program Professional in Music

MusMat Research Group

COORDINATORS

Carlos Almada (leader)
Daniel Moreira
Hugo T. de Carvalho
Liduino Pitombeira
Carlos Mathias
Cecília Saraiva

CURRENT MEMBERS

Ana Miccolis (Doctoral student)
Ariane Petri (Doctoral student)
Claudia Usai (Master student)
Eduardo Cabral (Undergraduate Research Project)
Érico Bonfim (Doctoral student)
Filipe de Matos Rocha (Doctoral student)
Igor Chagas (Master student)
Mariane Batista dos Santos (Undergraduate)
Max Kühn (Master student)
Natanael Luciano de Matos (Undergraduate Research Project)
Pedro Proença (Master student)
Pedro Zizels Ramos (Doctoral student)
Roberto Macedo (Doctoral student)
Rodrigo Furman (Master student)
Vinicius Braga (Undergraduate Research Project)

ALUMNI

Adriel Viturino (Undergraduate Research Project)
Alexandre Avellar (Undergraduate Research Project)
Alexandre Ferreira (Doctoral student)
André Codeço (Doctoral Student)
André Pinto (Undergraduate Research Project)
Desirée Mayr (Doctoral Student)
Fabio Monteiro (Master student)
Gabriel Mesquita (Master student)
Helder Oliveira (Doctoral student)

João Penchel (Undergraduate Research Project)
Jorge Santos (Master student)
Leandro Chrispim (Master student)
Marcel Castro Lima (Master student)
Marco Antônio Ramos Feitosa (Post-doc)
Rafael Fortes (Master student)
Rafael Soares Bezerra (Undergraduate Research Project)
Rodrigo Pascale (Undergraduate Research Project)

COLLABORATORS

Alexandre Schubert (Prof. UFRJ)
Fábio Adour (Prof. UFRJ)
Francisco Erivelton Aragão (Prof. UFC)
Pauxy Gentil-Nunes (Prof. UFRJ)
Nei Rocha (Prof. UFRJ)
Petrucio Viana (Prof. UFF)
Stefanella Boatto (Prof. UFRJ)
Raphael Sousa Santos

FRONT COVER ART

Title: *gold golden fractal zooming*
Author: Piotr Siedlecki / publicdomainpictures.net
https://www.freeimg.net/photo/1472033/gold-golden-fractal-zooming