Postprint Version

This is a postprint version of the following published document:

Flieder, Drew. Towards a mathematical foundation for music theory and composition: A theory of structure. Journal of Mathematics and Music, 2024. https://doi.org/10.1080/17459737.2024.2379788

Submitted exclusively to the *Journal of Mathematics and Music*
Last compiled on September 3, 2024

Towards a mathematical foundation for music theory and composition: A theory of structure

Drew Flieder[a*]

[a]*Department of Music, University of California-Santa Barbara, Santa Barbara, USA*

()

This paper proposes mathematical foundations for music theory and composition. While mathematical methods have proven effective in music theory, a deficiency in rigorous mathematical foundations often leads to ad-hoc constructions and a reliance on intuitive notions with inexplicit definitions. The proposal introduces a comprehensive environment for encoding musical phenomena, as well as a theory of musical parameter.

Drawing from the methodologies of Mazzola (Mazzola, Guerino et al. 2002. *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance.* Basel: Birkhäuser Verlag), the proposed framework aims to extend his theory of forms to accommodate a wider class of musical structures.

Additionally, I advocate for a framework characterized by "comprehensive comprehension" (accommodating a broad class of musical objects), "explicit encoding" (capturing essential features of objects), and "limited ontological commitments" (minimizing posited ontological primitives). These three aspects aim to ensure the framework's generality, explicit representation of structure, and economic efficiency in theoretical constructions. Minimizing ontological commitments also has the benefit of facilitating easier comparisons between entities. This approach is motivated by the desire for a highly versatile framework that enables systematic and standardized construction methods without constraining the intended domain of discourse.

To achieve these goals, the paper introduces a mathematical theory of structure grounded in topos theory. By offering a unified and systematic approach, this work contributes to establishing a more rigorous and standardized basis for mathematical music theory. This, in turn, fosters a deeper understanding of musical phenomena and enables more robust theoretical constructions and compositional applications.

Keywords: category theory; topoi; structure; ontology; foundations; presheaves

2020 Mathematics Subject Classification: 00A65; 18A25; 18A30; 18C40; 18F20; 18B25

1. Introduction

This paper proposes new mathematical foundations for music theory and composition, the goal being to provide a systematic and highly general framework for theorizing about musical phenomena.

The effectiveness of rational constructions in understanding and conceptualizing phenomena relies heavily on the theoretical framework employed. For instance, the classification of chords (pitch-class sets) and their relationships would be more challenging without the tools provided by set theory and group theory. In contrast to this more rigorous approach of classifying chords based on invariance with respect to group transformations,

*Drew Flieder. Email: drew@drewflieder.xyz

classifying chords based on subjective criteria—such as e.g. "dissonant," "consonant," "pretty," etc.—lacks a solid foundation agreeable to a community. Assessing the "prettiness" of a chord, for example, is highly subjective and varies from person to person, whereas the objective fact that a chord can be transposed or inverted into another is uncontroversially accepted within a scientific community.

Additionally, performing more advanced constructions based on subjective classifications would be complex. In contrast, the latter set- and group-theoretic approaches prove to be more manageable and systematic, as evidenced by the rapid progress in pitch-class set theory since its formulation in the mid-20th century. Hence, having mathematical principles for music theory has proved beneficial, in that it provides the means to encode various entities according to scientific standards.

However, even much of mathematical music theory lacks a solid foundation in mathematics. In pitch-class set theory, for example, reliance on intuitive notions of sets is common, and more advanced constructions are frequently assembled in an ad-hoc manner.

In contrast, a unified foundational system in musical discourse offers the advantage of conducting constructions according to standard procedures. This, in turn, facilitates the comparison of theoretical entities, as they will all relate by means of a coherent underlying ontology. Such a system would allow the theorizing of diverse musical genres, such as e.g. the works of Milton Babbitt and Gagaku music, without requiring alterations to the foundational ontology. It is important to emphasize that advocating for a *single* ontology does not imply reductionism. Rather, it suggests that as long as the ontology comports with its intended class of objects, any object belonging to that class can be integrated without compromising its essential features.

This draws a parallel to a philosopher conceptualizing a system of categories. The philosopher's objective is not to narrow reality by reducing it to these categories, which would result in a more limited worldview; instead, the aim is to articulate a comprehensive set of conditions that any phenomenon necessarily exhibits. In the same way, our task is to conceive of the conditions under which music-theoretical entities exist. This research problem ultimately leads us to propose a broad yet mathematical definition of *structure*, asserting that any musical phenomenon can be effectively described in terms of structure.

The following is a roadmap of what we will cover. In Section 2, we will outline the theoretical challenges we must address. Broadly, our task here is to define our problem, aiming to provide a framework that (1) comprehensively comprehends the intended class of musical objects, (2) explicitly encodes their essential features, and (3) limits our ontological commitments.

In Section 3, we will contextualize our work within the framework of functor categories in music theory, delineating two distinct motivations for their application. We will demonstrate that our motivation parallels that of Mazzola in (Mazzola 2002), as our approach is meta-theoretical, offering a systematic method for concept building. However, a notable difference between our proposed framework and Mazzola's lies in our aim to formulate a broader class of structures than Mazzola's form framework in (Mazzola 2002).

Section 4 introduces our new theory of mathematical structure. Section 4.1 presents the initial proposal for defining structure. However, we will see that it is limited in certain respects, and in Section 4.2, we will address these limitations by transitioning to a more rigorous framework involving presheaves. This transition leads us to the formal definition of structure in Section 4.3.

Our theory of structure exhibits similarities to the definition of mathematical structures in model theory (Hodges 1997, page 2) and likewise to Bourbaki's concept of structure, as detailed in (Corry 1992, Section 4). While a comprehensive discussion of the motivation behind our presentation, in contrast to the methods of model theory or Bourbaki, exceeds the scope of this text, I will briefly outline the motivation for our definition as follows: Our framework offers the flexibility to encompass a broader class of structures than those achievable through traditional model theory. Moreover, compared to Bourbaki, our structures can be encoded with greater economy and intuition.

Section 5 presents initial applications to music theory, aiming to showcase the versatility of the framework rather than advancing a specific area of music theory. Below is a brief overview of the examples addressed in this section:

- Ordered sets and lists.
- Groups and twelve-tone theory.
- Gestures in Mazzola's gesture theory framework (Mazzola et al. 2018).
- Modules and score objects, as discussed in (Mazzola 2002, Chapter 6).
- Parameter spaces for synthesizers.

These examples serve to illustrate the broad applicability and potential of the framework in diverse musical contexts. We conclude this section with a brief discussion highlighting the advantages of employing the structure theory framework over ad-hoc models.

In Section 6, we will discuss the challenges of translating Mazzola's concepts of denotators, local compositions, and global compositions into our framework. While technically feasible to inherit the denotator methodology, we will discuss philosophical considerations that may question its appropriateness. Therefore, we propose an open problem: formulating an analog to denotators within our framework (Section 6.1). Additionally, we will briefly examine the concepts of local and global compositions in Mazzola's framework in Section 6.2. We suggest that once the denotator formalism is established in our theory, the subsequent challenge is to formulate and classify local and global objects within our framework.

Finally, we will present our conclusions in Section 7.

Before delving into the core content of this paper, I offer a word of caution to the reader. The mathematical formalisms presented herein are admittedly abstract, which may pose a challenge for some readers to grasp the true essence and potential benefits of the proposed framework. To provide a more practical perspective, I have included Appendix B, where I contextualize the framework in informal, relatable terms. Additionally, I share insights into how my own musical thinking has been enriched by this framework, drawing from a composition that heavily relies on its methodologies. I highly encourage readers who resonate with the spirit of this proposal to explore Appendix B for concrete insights that complement the abstract formalisms presented in the main text.

2. Foundational principles

In formulating a foundation for a research domain, I propose that the framework should possess the following features:

(1) **Comprehensive comprehension**: The framework must be capable of comprehending every entity considered a member of its domain of discourse.[1]
(2) **Explicit encoding**: The framework should encode the essential features of the ob-

[1] This can be particularly challenging in dynamic domains like music theory, where the class of musical objects is in constant flux due to the continuous creation of new music. Nonetheless, we can treat this principle as a matter of degree, implying that a more comprehensive framework is preferable to a less comprehensive one.

jects within the domain.

(3) **Minimal ontological commitments**: The framework should limit its ontological commitments as much as possible.

The rationale for these requirements is expounded as follows:

Regarding (1): The objective is to avoid constraining the variety of objects that the framework can account for. For instance, set theory serves as a robust foundation for mathematics, since any mathematical object or proposition can be described in terms of sets.

Regarding (2): When studying a phenomenon p within a framework F, the "essential features" of p should be inherently encoded by F. This encoding should be intrinsic to p rather than arbitrarily declared "by fiat." An instance illustrating the latter scenario is found in Forte's naive set-theoretical framework in (Forte 1973), where the distinction between ordered and unordered sets lacks an underlying systematic foundation that inherently encodes their features of being ordered or unordered.

The following (Forte 1973, page 3) demonstrates this situation:

> For a number of reasons it is important to distinguish between ordered and unordered pitch-class sets. If, for example, $[0, 2, 3]$ is regarded as the same as $[2, 3, 0]$ it is assumed that the difference in order does not render the sets distinct from one another; they are equivalent sets since both contain the same elements. In such case the sets are referred to as *unordered sets*. If, however, the two sets are regarded as distinct, it is evident that they are distinct on the basis of difference in order, in which case they are called *ordered sets*.

There are problems with this methodology. The lack of a systematic framework establishing the ordered versus unordered nature of a set results in instances where one can determine whether a set is ordered or unordered only if Forte declares it "by fiat." Contrastingly, a foundation in set theory allows one to unambiguously posit a set $X = \{0, 2, 3\}$ and an ordered set as a pair $(X, <)$ where $<$ is an ordering relation on X. This avoids declarations "by fiat" and ensures that the status of X as an unordered set and $(X, <)$ as an ordered set is encoded by virtue of the foundational system (namely set theory).

The inexplicit encoding of essential aspects of phenomena is not ideal, particularly in the context of performing complex theoretical constructions. Grounding constructions on inexplicit definitions leads to difficulties in systematically recovering their meaning, allowing erroneous thinking to go undetected. Explicit definitions resolve this problem by ensuring the systematic recoverability of the total meaning of such constructions.

Regarding (3): The motivation for limiting ontological commitments is not merely economic, but aims to facilitate a clearer understanding of relations between phenomena. With only one ontological commitment, for instance, every entity can be understood as embodying the same underlying principle. In contrast, asserting multiple kinds of entities as ontological primitives creates an insurmountable barrier between the members of these different classes.

For instance, consider Forte's distinction between ordered and unordered sets. The lack of a foundational system necessitated the positing of two fundamentally distinct ontological categories: ordered and unordered sets. However, non-naive set theory (such as ZFC) would have entailed fewer ontological commitments, requiring only a commitment to the concept of a set. It would have enabled the derivation of ordered and unordered sets as incarnations of the same kind of entity (namely a set). Consequently, there is no insurmountable barrier between ordered and unordered sets; instead, there is a rule (an ordering relation) enabling the transformation of an unordered set into an ordered set.

3. Motivating the use of functor categories in music theory

Given our framework's reliance on functor categories, we begin this section by situating our work within the context of functor category applications in music theory as explored by other theorists. Functor categories have been applied in various instances within music theory; notable examples include (Noll 2005), (Popoff, Andreatta, and Ehresmann 2018), and Mazzola's seminal text (Mazzola 2002). These works reveal two overarching motivations for employing functor categories. In the works of Noll and Popoff et al., functor categories are utilized to address specific, concrete topics within music theory. Conversely, Mazzola's application of functor categories is driven by broader, meta-theoretical considerations. For instance, in (Mazzola 2002), Mazzola develops form and denotator theories aimed at establishing a framework for conceptualizing music-theoretic entities. His approach aims to provide a foundational basis adaptable to formulating novel music-theoretic entities as they emerge. Noll distinguishes between the motivations underlying Mazzola's application of topos theory and his own investigation by stating that while Mazzola's approach is primarily motivated by arguments situated at the meta-level or within broader epistemological and semiological frameworks, Noll's investigation is driven by a direct effort to interpret topos theory within the context of specific music-theoretical issues involving harmony (Noll 2005, page 4).

Following Noll's distinction, our methodology in this paper closely aligns with Mazzola's approach, particularly inspired by his theory of forms and denotators as elucidated in (Mazzola 2002, Chapter 6). Like Mazzola, our approach delves into meta-theoretical considerations, aiming to establish a systematic method for concept building. While our work draws inspiration from Mazzola's form and denotator theories and shares deep methodological parallels with his framework, it diverges in motivation in some significant ways, as we shall soon see.

Mazzola's form and denotator framework can be roughly defined as a conceptual framework for encoding musical "formulae," which are algebraic in nature.[2] Specifically, his theories are rooted in mathematical modules—in particular, module presheaves—facilitating recursive construction methods crucial for systematic concept building in music theory. As evidenced by its extensive applications, this framework offers a robust methodology for discussing formulae in music.

In contrast to the module-theoretic approach, Mazzola's more recent topological framework[3] shifts focus towards musical "gestures." The ontological distinction between these two theories lies in their treatment of algebraic versus topological entities, respectively. For Mazzola, this distinction marked a significant insight, highlighting the differences between formulaic and gestural realities in music.

Our paper does not seek to dispute the significance of this distinction in Mazzola's work. Rather, we aim to construct a framework that integrates algebraic and topological entities, among others, within a unified theoretical context. *Our fundamental objective is to establish a universal framework for encoding structure, where there is no inherent ontological divide between formulae, gestures, or any other musical entities, irrespective of the kinds of structure they exhibit.* The motivation behind this endeavor lies in the recognition that, particularly in composition but also in theory, various phenomena of different ontological statuses often converge into a composite phenomenon. Therefore, to facilitate the synthesis of diverse objects according to a unified concept-building approach, a singular theory becomes imperative.[4]

[2]See the distinction between formulae and gestures in (Mazzola and Andreatta 2007).

[3]See texts such as (Mazzola and Andreatta 2007; Mazzola 2009; Mazzola et al. 2018).

[4]It prompts a philosophical inquiry to ponder whether the consolidation of objects that typically reside within

Thus, our goal is to provide a comprehensive framework for encoding structure in music. We articulate this motivation to offer readers insight into our overarching objective.

The framework presented in this paper is an original application of the structure theory framework initially introduced in (Flieder 2022, Chapter 2), extended into the realm of music theory. In its original conception, this framework serves as a general method for encoding structure, not confined solely to music theory.

4. Defining structure

The following Sections (4.1–4.3) assume a familiarity with category theory. The reader is referred to Appendix A for the necessary technical foundations.

4.1. *Elementary structures*

We present our theory of structure starting with the category **Rel**, where objects are sets and morphisms are binary relations between sets. A binary relation R between sets A and B is defined set theoretically as a subset $R \subseteq A \times B$ of their Cartesian product. However, we use the notation $R : A \to B$ to reflect that relations are morphisms in **Rel**.

Our initial focus lies in defining *elementary structures*. This notion provides an intuitive grasp on structure, although it is inherently limiting, leading us to a more rigorous definition in Section 4.3. This advanced format ensures generality and facilitates the synthesis of structures through the categorical constructions of limits, colimits, power objects, and function objects.

An *elementary structure* is a set X in **Rel**, along with a collection $R = \{R_i : A_i \to X\}_{i \in I}$ of relations between various sets A_i and X. This collection generates structure on X. To exemplify, we show how a group is encoded using this approach.[5]

Consider a group as a pair $\mathbf{G} = (G, A)$, where G is the group's underlying set and $A = \{+, -, e\}$ represents a set of functions,[6] expressed by the following diagram.

$$\begin{array}{ccccc} G \times G & \xrightarrow{\;+\;} & G & \xleftarrow{\;-\;} & G \\ & & \uparrow{\scriptstyle e} & & \\ & & 1 & & \end{array}$$

For those familiar with group axioms, the diagram signifies the following:

- The morphism $+ : G \times G \to G$ is the group's binary operation.
- The morphism $- : G \to G$ maps each element $g \in G$ to its inverse $-g \in G$.
- The morphism $e : 1 \to G$ maps the singleton set 1 to G, identifying the identity element.

---

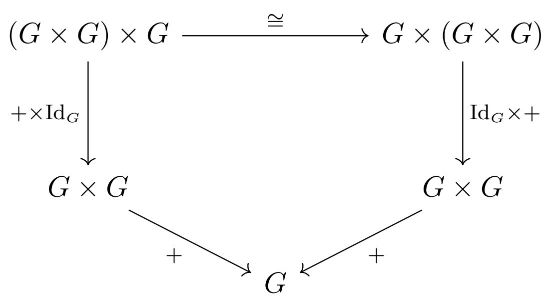
disparate mathematical categories—such as modules, topological spaces, ordered sets, and so forth—into a single mathematical category of structures (as proposed in this paper, Section 4) can be regarded as a form of *Aufhebung* in the Hegelian sense. Regardless of whether such an argument can be made, the central aim of this proposal remains consistent: to furnish a comprehensive framework capable of encoding diverse kinds of structures within a unified framework.

[5]The following construction is presented in (Awodey 2010) in the context of the category of sets.

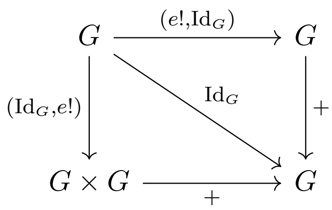
[6]Note that **Rel** contains **Set** as a full subcategory, and therefore any set function is a morphism in **Rel**.

One can check that the operations $+$, $-$, and e constitute a group on G, by verifying that they comply with group axioms. The group structure can be verified categorically through objects and morphisms in the following way:

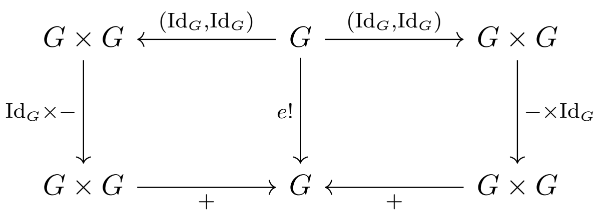
- Associativity of the binary operation is ensured through the commutativity of the following diagram.[7]



- The morphism $e : 1 \to G$ picks out the identity element, which means that the following diagram commutes, where $e! : G \to G$ sends every $g \in G$ to the identity element.



- The morphism $- : G \to G$ is an inverse with respect to $+$, expressed by the commutativity of the following diagram.



Satisfying these conditions confirms that the set $A = \{+, -, e\}$ indeed generates a group structure on G, allowing us to encode the group structure as the pair $\mathbf{G} = (G, A)$.

This example serves as an adequate introduction to our concept of an elementary structure. In (Flieder 2022), it is demonstrated how this methodology enables the encoding of topological spaces, ordered sets, modules, and various other mathematical structures.

[7] The symbol $\cong$ expresses the canonical isomorphism defined by $((a, b), c) \mapsto (a, (b, c))$.

4.2. *From elementary structures to presheaves*

In our structure theory, we will use the category of presheaves over **Rel**.[8] Inheriting Mazzola's notation in (Mazzola 2002) for presheaves, we will denote the representable presheaf of a set X by $@X$, the hom set $\mathrm{Hom}_{\mathbf{Rel}}(A, X)$ as $A@X$, and the presheaf category $[\mathbf{Rel}^{\mathrm{op}}, \mathbf{Set}]$ as $\mathbf{Rel}^{@}$.

Now consider our definition of an elementary structure $S = (X, R)$, where we choose a set X and a collection of relations $R = \{R_i : A_i \to X\}_{i \in I}$. This collection R induces a presheaf $P_R : \mathbf{Rel} \to \mathbf{Set}$.

To understand this, we first introduce the concept of a sieve.

Definition 4.1 (Sieve) Let $\mathcal{C}$ be a category and X an object in $\mathcal{C}$. A *sieve* C on X is a collection of morphisms with X as their codomain, closed under precomposition with morphisms in $\mathcal{C}$. Formally, C consists of morphisms such that whenever $(f : A \to X) \in C$ and $(g : B \to A)$ is a morphism in $\mathcal{C}$, then $(f \circ g : B \to X) \in C$.

In (Mac Lane and Moerdijk 2012, page 38), it is shown that the collection of sieves on X corresponds one-to-one with the collection of subfunctors of the representable presheaf $\mathrm{Hom}_{\mathcal{C}}(-, X)$. Thus, for every sieve C on X, there exists a unique subfunctor $\widehat{C} \subseteq \mathrm{Hom}_{\mathcal{C}}(-, X)$. In the context of an elementary structure $S = (X, R)$, the collection R induces a sieve on X, thereby generating the corresponding subfunctor of $@X$ in $\mathbf{Rel}^{@}$.

Since we often start by defining an elementary structure $S = (X, R)$ and derive its presheaf in $\mathbf{Rel}^{@}$, we define the following map.[9]

$$\mathrm{PSh} : \bigcup_{X \in \mathbf{Rel}_0} \mathcal{P}\left(\bigcup_{A \in \mathbf{Rel}_0} A@X \right) \to \mathbf{Rel}^{@}_0 \tag{1}$$

This map sends every elementary structure to its induced presheaf. The domain is constructed in the following way. For each set X, we union all hom sets with codomain X and form the powerset. This provides the collection of all subsets of relations into X, which is crucial for defining an elementary structure on X. This procedure is then repeated for every set X in **Rel**, and all such collections are unioned together, giving us the collection of all elementary structures. Then each elementary structure is mapped to its corresponding presheaf in $\mathbf{Rel}^{@}$.

The reason for transitioning to the presheaf framework lies in the insufficiency of elementary structures when synthesizing structures to create new ones—a frequent need in both mathematics and music. To illustrate, the creation of the product of sets A and B involves the following straightforward operation.

$$A \times B \coloneqq \{(a, b) \mid a \in A \wedge b \in B\}$$

In the realm of music, a structure might define a pitch space while another defines a duration space; their product structure would therefore contain both pitch and duration information. These crucial operations are absent in the framework of elementary structures, which hinders the potential for diverse constructions. For instance, in **Rel**, products and coproducts coincide due to self-duality, and they correspond to disjoint unions of sets. Hence, the product of sets A and B in **Rel** fails to yield their set-theoretic Cartesian product, thereby imposing constraints. For instance, as illustrated earlier in this para-

[8] See Appendix A for discussion about presheaves.

[9] The subscript "0" in $\mathbf{Rel}_0$ and $\mathbf{Rel}^{@}_0$ denotes their respective object classes.

graph, one may wish to form the Cartesian product of pitch and duration structures, generating pairs (p, d) where p represents pitch and d duration. The presheaf framework resolves this issue, as products in $\mathbf{Rel}^{@}$ align more closely with the desired set-theoretic Cartesian product. Specifically, the presheaf framework offers essential operations such as product, coproduct,[10] power, and function structures, facilitated by the topos properties inherent in the presheaf category.

An outline of these operations is as follows:

- *Product structures.* Given two structures, $S = (A, R)$ and $T = (B, Q)$, we want their product $S \times T$ to comprise the Cartesian product $A \times B$ of their underlying sets, where the A coordinate inherits the structure given by R, while the B coordinate inherits the structure given by Q. For example, if R generates a total order on A and Q generates a group on B, a pair $(a, b) \in A \times B$ gives an order position a *and* a group element b.
- *Coproduct structures.* Using structures S and T, we want their coproduct $S + T$ to consist of the disjoint union of A and B, and have the A elements of $A + B$ inherit the structure generated by R, and the B elements inherit the structure generated by Q. For instance, if S is a total order and T is a group, an element $k \in A + B$ represents *either* an order position *or* a group element.
- *Power structures.* We want a power structure $\Omega(S)$ of a structure S to consist of all substructures of S. Intuitively, one might consider $\Omega(S)$ akin to the powerset $\mathcal{P}(A)$ of the underlying set of S, where each subset $X \subseteq A$ embodies the structure of S restricted to the X portion of A. Although we can certainly encode such restrictions in the formal setup, the latter is, in fact, even more general.
- *Function structures.* We want the function structure T^S for structures S and T to consist of the set of functions $f : A \to B$ on their underlying sets. However, the concept involves mapping the structure of S to the structure of T. For instance, if S is a total order and T is a group, a mapping $\overline{f} : S \to T$ expresses a sequence of group elements from T, where the sequence is indexed by the elements of S. A category that accommodates such function types is said to contain *internal homs.*

In Section 5, we will see how the presheaf framework enables all of the aforementioned operations. Additionally, that section will elucidate why each of these operations is indispensable for music-theoretic purposes.

4.3. *Formal definition of structure*

We present the formal definition of a structure, following the format of Mazzola's definition of a *form* in (Mazzola 2002, pages 63–64).

Definition 4.2 (Structure) A *structure* S is a quadruple $S = (N, T, C, I)$, where:

(1) N serves as the *name* of S; it consists of a string of symbols from the free monoid $\mathfrak{N}$ over an alphabet[11] A. This alphabet consists of symbols from formal and informal languages to allow maximal freedom in naming. We denote the name of S as $\mathrm{Name}(S)$.

(2) T is the *type* of S, and is one of the following symbols:
   (a) **Simple**,

[10] Indeed, the presheaf framework extends beyond mere product and coproduct constructions to encompass general limits and colimits as well.

[11] The choice of the alphabet A is intentionally left ambiguous to avoid restricting naming possibilities. In a fully formal context, explicit specification of this alphabet as a set would be required for the free monoid construction. However, for our current purposes we tolerate this fuzziness in the definition.

(b) **Limit**,
(c) **Colimit**,
(d) **Power**,
(e) **Sub**,
(f) **Hom**.
We denote the type of S as Type(S).

(3) C is the *coordinator* of S, and it depends on the type T as follows:
(a) If T is **Simple**, then C is a set X.
(b) If T is **Limit** or **Colimit**, then C is a diagram[12] $\mathcal{D}$ of structures.
(c) If T is **Power** or **Sub**, then C is a structure.
(d) If T is **Hom**, then C is an ordered pair (C_1, C_2) of structures.
We denote the coordinator of S as Coordinator(S).

(4) I is the *identifier* of S, and is a monomorphism of functors $I : Fu \rightarrowtail A$ in $\mathbf{Rel}^{@}$. The codomain A is defined as follows:
(a) If T is **Simple**, then $A = @X$.
(b) If T is **Limit**, then $A = \lim \mathcal{D}$.
(c) If T is **Colimit**, then $A = \mathrm{colim}\ \mathcal{D}$.
(d) If T is **Power**, then $A = \Omega^{Fun(C)}$.[13]
(e) If T is **Sub**, then $A = Fun(C)$.
(f) If T is **Hom**, then $A = Fun(C_2)^{Fun(C_1)}$.
The identifier associates the functor Fu of S with the functor constructed by the type and coordinator. The role of the identifier $I : Fu \rightarrowtail A$ typically operates similarly to set comprehension in set theory, but in this context, we enumerate over "elements" of structures rather than elements of sets. The domain of the identifier, Fu, is the structure's functor—essentially, the component that provides the content of the structure. We denote the identifier of S as Identifier(S).

Before concluding this section, it is worth noting the connections of our new formal setup with the types of structures discussed at the end of Section 4.2.

(1) Structures of type **Simple** correspond to elementary structures.
(2) Structures of type **Limit** generalize product structures. (A product structure is a limit structure with a discrete[14] diagram as its coordinator.)
(3) Structures of type **Colimit** generalize coproduct structures. (A coproduct structure is a colimit structure with a discrete diagram as its coordinator.)
(4) Structures of type **Power** correspond to power structures.
(5) Structures of type **Sub** are used to define subobjects of the coordinator structure.
(6) Structures of type **Hom** correspond to function structures.

For brevity, we conclude the discussion here. A comprehensive elaboration of the above six kinds of types is available in (Flieder 2022).

5. Initial applications of structures to music theory

In this section, we demonstrate how our structure theory framework facilitates the encoding of musical phenomena. Mathematical music theory employs various structures,

[12] A *diagram* $\mathcal{D} : J \to \mathcal{C}$ is a functor, where J is a small index category. Essentially, a diagram identifies a set of objects and morphisms within the category $\mathcal{C}$.
[13] The notation $Fun(C)$ denotes the functor of C, i.e. an object in $\mathbf{Rel}^{@}$.
[14] A diagram $\mathcal{D} : J \to \mathcal{C}$ is called *discrete* when J contains no morphisms other than the identities.

including groups, sets, ordered sets, topological spaces,[15] and many more. An advantage of our structure theory is its ability to encode these objects within the same framework, namely as structures with functors in $\mathbf{Rel}^@$. Additionally, it facilitates the synthesis of structures using topos-theoretic constructions, such as limits, colimits, power objects, and internal homs, offering diverse methods for creating new structures from existing ones. The examples that follow offer a glimpse into the diverse applications facilitated by our structure theory.

Example 5.1 (Arbitrary ordered sets and lists) A totally ordered set is a set whose elements are arranged in a specific sequence. A list is akin to a totally ordered set, with the distinction that it permits the presence of duplicate elements. In our framework, a list can be constructed by defining a structure *Elems* containing elements that one wishes to construct a list with. For a list of length k, we first define the following structure.

$$K \underset{\phi:Fu \rightarrowtail @\{1,\dots,k\}}{\longrightarrow} \mathbf{Simple}(\{1,\dots,k\})$$

The structure K is a total order on the first k natural numbers. Such a structure is derived from an elementary structure $(\{1,\dots,k\},\{<\})$, where $<$ is a total ordering relation on $\{1,\dots,k\}$. We then derive its functor Fu via the presheaf map defined in (1).

$$\mathrm{PSh}\left(\{<\}\right) = Fu$$

A list is then represented by a morphism $l : K \to Elems$, where the value of $l(i) \in Elems$ provides the ith element of the list.[16]

Now, any morphism $f : X \to Y$ in $\mathbf{Rel}^@$ can be turned into an object in $\mathbf{Rel}^@$ by taking the limit of the following diagram.

$$X \xrightarrow{f} Y$$

Thus, the list l is encoded as the following structure.

$$List \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}\left(K \xrightarrow{l} Elems\right)$$

It consists of pairs[17] (i,e) where i is an order position in K, and e is an element in *Elems*.

Example 5.2 (Groups and twelve-tone theory) Groups, which are central to music theory, can be encoded as structures, as we saw in Section 4.1. For example, the set of twelve pitch classes is treated as the additive group $\mathbb{Z}_{12}$. The T/I group of transpositions and inversions of $\mathbb{Z}_{12}$ can be defined as a subset of the hom set $\mathrm{Hom}_{\mathbf{Rel}^@}(\mathbb{Z}_{12},\mathbb{Z}_{12})$. Since $\mathbf{Rel}^@$ has internal homs, we can encode the hom set as the following structure.

$$T/I \underset{\phi:Fu \rightarrowtail Fun(\mathbb{Z}_{12})^{Fun(\mathbb{Z}_{12})}}{\longrightarrow} \mathbf{Hom}(\mathbb{Z}_{12},\mathbb{Z}_{12})$$

[15] While topological spaces are less common in the context of music theory compared to other structures, refer to (Callender, Quinn, and Tymoczko 2008; Polansky 1996; Tymoczko 2010) for applications.

[16] Note that the notation $L(i) \in Elems$ is technically an abuse of set-theoretic notation.

[17] Again, this is an abuse of set-theoretic notation, but is a convenient way to express the resulting structure.

Here, the identifier $\phi : Fu \rightarrowtail Fun(\mathbb{Z}_{12})^{Fun(\mathbb{Z}_{12})}$ ensures that only maps constituting transpositions and inversions are included in T/I.

Using the list construction from the previous example, we can define a twelve-tone row as a structure of the following form.

$$ToneRow \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}\left(12 \xrightarrow{t} \mathbb{Z}_{12}\right)$$

The structure 12 is the totally ordered set of integers from 1 to 12, and t is epimonic.[18] In other words, a twelve-tone row is a twelve-element list of non-repeating elements from $\mathbb{Z}_{12}$.

A transformation of a tone row can be visualized via a commutative diagram expressing the fact that one tone row is a transposition or inversion of another.

$$\begin{array}{ccc} 12 & \xrightarrow{\mathrm{Id}_{12}} & 12 \\ \downarrow t & & \downarrow t' \\ \mathbb{Z}_{12} & \xrightarrow[\varphi]{} & \mathbb{Z}_{12} \end{array}$$

Assuming that $\varphi \in T/I$, the diagram expresses that t transforms into t' under a transposition or inversion operator.

Furthermore, the collection of all tone rows is encoded as the following structure.

$$AllRows \underset{\psi : Gu \rightarrowtail Fun(\mathbb{Z}_{12})^{Fun(12)}}{\longrightarrow} \mathbf{Hom}(12, \mathbb{Z}_{12})$$

Here, the identifier $\psi : Gu \rightarrowtail Fun(\mathbb{Z}_{12})^{Fun(12)}$ indicates that we want all and only those morphisms $t : 12 \to \mathbb{Z}_{12}$ that are epimonic.

All the results in this example generalize to equal-tempered scales of any size.

Example 5.3 (Gestures) In this example, we demonstrate the encoding of gestures from Mazzola's gesture theory framework, as detailed in (Mazzola et al. 2018). In Mazzola's framework, gestures are defined as a distinct type of morphism $\delta : \Delta \to \overrightarrow{X}$ within the category of directed graphs (digraphs). To encode these in our framework will require us to encode digraphs and topological spaces as structures. The encoding of topological spaces as structures has been previously achieved and discussed in (Flieder 2022, Chapter 2), which we recommend for further verification.

Prior to presenting the formalism for encoding gestures within our framework, we first offer Mazzola's definition of the concept in (Mazzola et al. 2018, page 914). To begin, we introduce the notion of a digraph. A *digraph* is given by a function $\Gamma : A \to V \times V$, where A is a set of arrows and V is a set of vertices. The function's first projection, $t = \mathrm{pr}_1 \circ \Gamma$, is referred to as the *tail* function, while the second projection, $h = \mathrm{pr}_2 \circ \Gamma$, is referred to as the *head* function. Therefore an arrow a can be symbolized as $t(a) \xrightarrow{a} h(a)$. For digraphs $\Gamma : A \to V \times V$ and $\Gamma' : A' \to V' \times V'$, a digraph morphism $f : \Gamma \to \Gamma'$ is a pair

[18] An epimonic map is the category-theoretic generalization of the set-theoretic notion of a *bijection*.

$f = (f_0, f_1)$ such that the following diagram commutes.

$$\begin{array}{ccc} A & \xrightarrow{\Gamma} & V \times V \\ \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_0 \times f_0} \\ A' & \xrightarrow[\Gamma']{} & V' \times V' \end{array}$$

Special types of digraphs, referred to as *spatial digraphs*, are fundamental in Mazzola's gesture theory. A spatial digraph is associated with a topological space X and denoted by $\overrightarrow{X}$. The arrow set $A_{\overrightarrow{X}}$ of a spatial digraph $\overrightarrow{X}$ is defined as the set of continuous curves $c : I \to X$ in X, where $I = [0,1]$ denotes the unit interval. The vertex set of $\overrightarrow{X}$ is $V_{\overrightarrow{X}} = X$. For any arrow $c \in A_{\overrightarrow{X}}$, it is required that $t(c) = c(0)$ and $h(c) = c(1)$. Consequently, Mazzola defines a gesture as follows: For a digraph Δ and a topological space X, a Δ-*gesture in* X is a digraph morphism $\delta : \Delta \to \overrightarrow{X}$. The intuition is that δ actualizes the abstract vertices and arrows in Δ within the topological space X.

To encode such gestures within our framework, we must first define digraphs and topological spaces. We start with topological spaces. For a set $\mathcal{X}$ in **Rel**, we define a topology on $\mathcal{X}$ as a structure of type **Simple**.[19]

$$X \underset{\tau_X : Fu_X \rightarrowtail @\mathcal{X}}{\longrightarrow} \mathbf{Simple}(\mathcal{X})$$

Similarly, we can encode the unit interval $[0,1]$ as a topological structure.

$$I \underset{\tau_I : Fu_I \rightarrowtail @[0,1]}{\longrightarrow} \mathbf{Simple}([0,1])$$

Consequently, the set of continuous curves in X, constituting the arrow structure of $\overrightarrow{X}$, can be encoded as a structure of **Hom** type.

$$A_{\overrightarrow{X}} \underset{curves : Gu \rightarrowtail Fun(X)^{Fun(I)}}{\longrightarrow} \mathbf{Hom}(I, X)$$

The identifier $curves : Gu \rightarrowtail Fun(X)^{Fun(I)}$ specifies that only morphisms from I to X that constitute *continuous* maps are permissible in the functor of $A_{\overrightarrow{X}}$.

To encode a spatial digraph, we define a structure morphism $\Gamma : A_{\overrightarrow{X}} \to X \times X$. This morphism, as mandated by Mazzola's definition, is such that for all $c \in A_{\overrightarrow{X}}$, the conditions $\mathrm{pr}_1 \circ \Gamma(c) = t(c) = c(0)$ and $\mathrm{pr}_2 \circ \Gamma(c) = h(c) = c(1)$ are satisfied.

To facilitate the definition of a digraph morphism as a morphism between structures, we encode the spatial digraph defined by Γ as a structure of **Limit** type.

$$\overrightarrow{X} \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}\left(A_{\overrightarrow{X}} \xrightarrow{\Gamma} X \times X\right)$$

Next, we define a digraph $D : A_D \to V_D \times V_D$, where A_D and V_D represent arrow and

[19] For detailed instructions regarding how to encode topological spaces using our structure theory, refer to (Flieder 2022, Chapter 2).

vertex structures, respectively. This digraph is then encoded as a structure.

$$\Delta \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}\Big(A_D \xrightarrow{D} V_D \times V_D\Big)$$

Finally, we define a Δ-gesture by specifying a structure morphism $\delta : \Delta \to \overrightarrow{X}$. The morphism δ is determined by a pair (δ_0, δ_1) of structure morphisms on the vertex and arrow structures, respectively, ensuring the commutativity of the following diagram.

$$\begin{array}{ccc} A_D & \xrightarrow{\quad D \quad} & V_D \times V_D \\ {\scriptstyle \delta_1}\Big\downarrow & & \Big\downarrow{\scriptstyle \delta_0 \times \delta_0} \\ A_{\overrightarrow{X}} & \xrightarrow[\quad \Gamma \quad]{} & X \times X \end{array}$$

The method outlined in this example therefore establishes the process for encoding gestures within our framework.

Example 5.4 (Scores) In this example, we demonstrate how to encode scores within our framework, following Mazzola's methodology (Mazzola 2002, Chapter 6).

We begin by encoding a structure that encompasses note event information, including onset, pitch, loudness, and duration coordinates. Each of these parameters can be defined as the same module structure over $\mathbb{R}$. For instance, we define an elementary structure $(\mathbb{R}, M)$ where the relations $M_i \in M$ collectively generate a module structure on $\mathbb{R}$.[20] Subsequently, all four structures can be encoded using the same functor $\mathrm{PSh}(M)$, differing only in name.

$$\begin{aligned} &\mathit{Onset} \underset{\phi:\mathrm{PSh}(M)\rightarrowtail @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R}) \\ &\mathit{Pitch} \underset{\phi:\mathrm{PSh}(M)\rightarrowtail @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R}) \\ &\mathit{Loudness} \underset{\phi:\mathrm{PSh}(M)\rightarrowtail @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R}) \\ &\mathit{Duration} \underset{\phi:\mathrm{PSh}(M)\rightarrowtail @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R}) \end{aligned}$$

The space consisting of note events is then the product of these four structures.

$$\mathit{Note} \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}(\mathit{Onset}, \mathit{Pitch}, \mathit{Loudness}, \mathit{Duration})$$

A single point in this space signifies a note event.

However, a score also incorporates rest information. Thus, an event in a score may consist of either a note or a rest. We can define the space of rests as comprising an onset coordinate and a duration coordinate.

$$\mathit{Rest} \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}(\mathit{Onset}, \mathit{Duration})$$

Consequently, the total musical event space encompasses all note and rest event pos-

[20] See (Flieder 2022, Chapter 2) regarding how to encode modules as elementary structures.

sibilities, forming a coproduct structure.

$$MusEvent \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Colimit}(Note, Rest)$$

Considering a score as a collection of such events, the space of possible scores constitutes a structure of type **Power**.

$$Score \underset{\psi:2^{Fun(MusEvent)} \rightarrowtail \Omega^{Fun(MusEvent)}}{\longrightarrow} \mathbf{Power}(MusEvent)$$

Thus, a single point in $Score$ encodes a musical score.

Example 5.5 (Parameter space of a synthesizer) The parameters of a synthesizer can be encoded as structures. Begin by encoding each parameter individually as a structure. For example, a frequency parameter can be conceived as a vector space over real numbers.

$$Freq \underset{V:Fu \rightarrowtail @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R})$$

The entire parameter space is then the product of all parameter structures.

$$SynthParams \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}(P_1, \ldots, P_n)$$

An element in $SynthParams$ denotes a specific configuration of parameter values for the synthesizer.

We can also model the process of improvisation on this synthesizer. Such an improvisation consists in varying its parameters over time. This can be modeled by introducing a temporal structure $Time$ with elements representing points in time. The improvisation itself is modeled as a mapping $I : Time \to SynthParams$ and is encoded as a structure.

$$Improv \underset{\mathrm{Id}}{\longrightarrow} \mathbf{Limit}\Big(Time \overset{I}{\to} SynthParams\Big)$$

The value of a pair $(t, s) \in Improv$ provides the state s of the synthesizer at time t.

The collection of all possible improvisations can likewise be encoded as a structure, as so.

$$PossibleImprovs \underset{\phi:Gu \rightarrowtail Fun(SynthParams)^{Fun(Time)}}{\longrightarrow} \mathbf{Hom}(Time, SynthParams)$$

Continuing this elaborate construction, one may then wish to define a metric structure on $PossibleImprovs$, in order to compare improvisations in terms of their degree of similarity. A *metric* on a space X is defined as a mapping of the form $d : X \times X \to \mathbb{R}$, satisfying the following requirements for every $u, v \in X$:

- The distance $d(u, u) = 0$.
- If $u \neq v$, then $d(u, v) > 0$.
- If $d(u, v) = r$, then $d(v, u) = r$.

Hence a metric on $PossibleImprovs$ requires specifying first the codomain structure.

$$Distances \underset{\psi:Hu \rightarrow @\mathbb{R}}{\longrightarrow} \mathbf{Simple}(\mathbb{R})$$

Then, one defines the distance function as a structure morphism satisfying the requirements of a metric.

$$d : PossibleImprovs \times PossibleImprovs \to Distances$$

These examples offer initial insights into how diverse music-theoretic entities can be encoded using the structure theory framework. They suggest the benefits of adopting a unified, systematic approach to addressing various music-theoretic topics, as opposed to relying on ad-hoc models. Importantly, the structure theory's capacity to incorporate new music-theoretical entities as they emerge enables the consolidation of new phenomena into a cohesive theoretical framework. This stands in contrast to situations where ad-hoc models are used for emerging objects of interest. In such cases, diverse music-theoretical entities exist across different categories with rigid boundaries between them, inhibiting systematic methods of comparison. A systematic framework for unifying diverse entities within the same context not only facilitates systematic construction but also enables systematic comparison. This significantly enhances the potential for theoretical advancement.

6. Open problems: Translation of denotators, local compositions, and global compositions into the structure theory framework

This section will be of particular interest to readers familiar with Mazzola's denotator theory as presented in (Mazzola 2002), along with his theories of local and global compositions developed therein. In the next subsection, we propose, as an open problem, the formulation of an analog of denotators in our theory. While we can, in principle, inherit the denotator methodology in our framework, there are philosophical reasons that suggest this may not be the appropriate methodology in our framework.

Furthermore, since local compositions are constructed based on the denotators of Mazzola's theory, and global compositions are, in turn, built on local compositions, we need to establish an analog to denotators in our structural framework. Once this formulation is achieved, we can systematically formulate local and global compositions in our framework. The eventual goal is to develop a method for calculating the isomorphism classes of both local and global compositions in our framework, which is one of the milestones Mazzola achieved in the module-theoretic framework of his earlier work.

Completing these tasks is, however, beyond the scope of this work. We present the topics here as open problems, in hope that future research can solve them.

6.1. *Denotators*

Mazzola's theory of forms and denotators is discussed in (Mazzola 2002). In a more recent work (Zheng and Mazzola 2023), the formalisms of forms and denotators have been extended to arbitrary presheaf categories. However, for the sake of clarity and simplicity, we confine our discussion to the presheaf category $\mathbf{Mod}^{@}$ outlined in the former work.

The motivation for addressing this topic here stems from the challenge of formulating an analog of the denotator format within our structure theory. This challenge remains as an open problem, primarily due to significant philosophical considerations. In this section, we will delve into these philosophical issues.

While a detailed examination of denotators in Mazzola's work is beyond the scope of this paper, we will provide a brief overview, directing interested readers to (Mazzola 2002) for a comprehensive treatment. Let's begin with the definition outlined in (Mazzola 2002, page 67).

Definition 6.1 (Denotator) Let M be a module (also called an *address*). An *M-addressed denotator* is a triple $D = (ND, FD, CD)$ where:

(1) ND is a string of ASCII characters; it is called the *name* of D, and denoted by $N(D)$.
(2) FD is a form; it is called the *form* of D, and denoted by $F(D)$.
(3) CD is an element of $M@Fun\,(F(D))$; it is called the *coordinates* of D, and denoted by $C(D)$.

For a denotator $D = (ND, FD, CD)$, we represent it using the following notation.

$$ND : M \rightsquigarrow FD(CD)$$

Let us illustrate the use of denotators with an example. Define $M = 0$ as the zero module over the coefficient ring $\mathbb{Z}$, and define the following pitch-class form.[21]

$$PiMod_{12} \xrightarrow[\text{Id}]{} \mathbf{Simple}(\mathbb{Z}_{12})$$

We can represent the pitch class $5 \in \mathbb{Z}_{12}$, for instance, as a 0-addressed denotator.

$$\text{pc-5} : 0 \rightsquigarrow PiMod_{12}(t^5)$$

The denotator represents $5 \in \mathbb{Z}_{12}$ through the affine map $t^5 : 0 \to \mathbb{Z}_{12}$ that sends the single element in 0 to 5.

There are no restrictions to defining denotators solely for 0-addressed points of the form's functor. For example, Mazzola illustrates how a dodecaphonic series (Mazzola 2002, page 149) can be encoded as a denotator of the following form, where the coordinates $Ser_i = Ser(e_i)$ are pairwise distinct.

$$Ser : \mathbb{Z}^{11} \rightsquigarrow PiMod_{12}(Ser_0, Ser_1, \ldots, Ser_{11})$$

The underlying intuition behind this formalism is that the zero vector in $\mathbb{Z}^{11}$ maps to the first pitch class of the series, while the unit vector for the ith coordinate in $\mathbb{Z}^{11}$ maps to the $(i+1)$th pitch class of the series. In essence, such a denotator articulates an ordering on the pitch classes of $\mathbb{Z}_{12}$.

An essential aspect of denotator theory lies in the significance attributed to such denotators *precisely because the objects within the source category* **Mod** *possess inherent structure*. For example, the denotator Ser establishes a serial ordering on $\mathbb{Z}_{12}$ by utilizing the structured nature of $\mathbb{Z}^{11}$. This structure functions as the module indexing the elements of $\mathbb{Z}_{12}$, thereby representing a series. Consequently, any A-addressed point of a module presheaf $\mathcal{M}$ in $\mathbf{Mod}^@$ represents a structured perspective constitutive of $\mathcal{M}$, where the perspective's structure is contingent upon the structure of A. Using the aforementioned example, a series morphism $s \in \mathbb{Z}^{11}@\mathbb{Z}_{12}$ delineates a series over elements

[21] Note that we are currently discussing Mazzola's *forms* rather than our *structures*. Therefore, the form $PiMod_{12}$ has an object in $\mathbf{Mod}^@$, not $\mathbf{Rel}^@$, for its functor.

from $\mathbb{Z}_{12}$ precisely because $\mathbb{Z}^{11}$ possesses a structure that determines the significance of the morphism as an "ordering perspective" of $\mathbb{Z}_{12}$. In essence, we can interpret any morphism $f : M \to N$ in **Mod** as, intuitively speaking, a *structured perspective* of N, where the perspective's structure is determined by the structure of M.

In our structure theory, the situation differs from Mazzola's module-theoretic framework. Objects within our source category **Rel** lack inherent structure; they are simply sets. Transitioning from **Rel** to $\mathbf{Rel}^{@}$ holds philosophical significance, characterized by the notion of *generating structure* on the sets in **Rel**. Here, every sieve C over any X in **Rel** initiates a process that *generates structure* on X. Consequently, the resulting induced presheaf $\mathrm{PSh}(C)$ in $\mathbf{Rel}^{@}$ emerges as a structured object, resulting from the relations in C that generate structure on X.

In Mazzola's framework, the motivation for transitioning from **Mod** to $\mathbf{Mod}^{@}$ lies in the latter's status as a topos, facilitating universal constructions such as limits, colimits, and power objects—features absent in **Mod** alone. Unlike our framework, where structure is built from scratch, Mazzola's framework sees modules M transitioning to their representable functor $@M$ under the premise of their equivalence. This transition is not driven by the necessity to construct structure on objects in **Mod**, as they inherently possess the requisite module structure for Mazzola's theoretical framework. Instead, the presheaf construction primarily aims to provide a category conducive to significant universal constructions lacking in **Mod** itself.

Conversely, in our structure theory, our perspective is that the transfer of objects in **Rel** to their presheaves (representable or otherwise) in $\mathbf{Rel}^{@}$ is a process through which the "bare" objects in **Rel** *acquire* structure. Thus, we cannot simply consider objects X in **Rel** as replaceable by their representable presheaves $@X$ in $\mathbf{Rel}^{@}$.

Hence, we can distinguish between morphisms in **Mod** and morphisms in **Rel** as follows: In **Mod**, a morphism $f : M \to N$ represents a *structured perspective of* N, while in **Rel**, a morphism $R : X \to Y$ denotes a *structuring act on* Y.

For example, a morphism $f : \mathbb{Z}^{11} \to \mathbb{Z}_{12}$ in **Mod** does not *generate* a total order structure on $\mathbb{Z}_{12}$. Instead, it presents $\mathbb{Z}_{12}$ in an "ordered way," by associating the basis vectors of $\mathbb{Z}^{11}$ with elements in $\mathbb{Z}_{12}$.

In contrast, a total ordering relation $< : X \to X$ in **Rel** does *generate* a total order structure on X. It is not merely a "viewing" of X in an "ordered way," since X lacks inherent order. Rather, the relation $< : X \to X$ is the very process through which X acquires its order structure; in other words, X *becomes* structured, rather than being "viewed in a structured way."

Once in $\mathbf{Rel}^{@}$ itself, however, the morphisms in $\mathbf{Rel}^{@}$ resemble "structured perspectives." This is exemplified in Example 5.2, where we considered the case of a dodecaphonic series. Unlike in Mazzola's framework, where a series is an element of the representable functor $@\mathbb{Z}_{12}$ at address $\mathbb{Z}^{11}$, in our framework, a series is encoded as a morphism in $\mathbf{Rel}^{@}$ itself. Specifically, it is a natural transformation of functors $s : Fun(12) \to Fun(\mathbb{Z}_{12})$. Therefore, if the idea of a denotator is, informally, to encode "structured perspectives" of recursively generated forms, then such structured perspectives in our framework would be the morphisms in $\mathbf{Rel}^{@}$. This is because the analog of "recursively generated forms" in our framework are the recursively generated structures in $\mathbf{Rel}^{@}$, and the structured perspectives in our framework are the morphisms in $\mathbf{Rel}^{@}$.

This situation leads me to believe, for the time being, that the analog of a denotator in our framework should be defined as a named morphism in $\mathbf{Rel}^{@}$. To reiterate, the distinction between this methodology and Mazzola's denotator methodology is that, in Mazzola's framework, a denotator is a point of a functor F in $\mathbf{Mod}^{@}$ at an address A. In contrast, the equivalent of a denotator in our framework corresponds to a natural transformation of functors in $\mathbf{Rel}^{@}$.

However, if we aim to preserve the same format as Mazzola's theory precisely, we could do so by iterating the presheaf construction over $\mathbf{Rel}^{@}$, resulting in the (second-order) presheaf category

$$\left(\mathbf{Rel}^{@}\right)^{@} := \left[\left(\mathbf{Rel}^{@}\right)^{\mathrm{op}}, \mathbf{Set}\right].$$

Yet, this seems unnecessary, since the primary motivation for moving to presheaf categories in Mazzola's form and denotator framework is to derive the topos structure initially lacking in the source category **Mod**. As $\mathbf{Rel}^{@}$ already possesses the desired topos structure, there is no need to transition to a second-order presheaf category.

Nevertheless, rather than hastily committing to a formalism, I consider it an open problem to determine the appropriate analog of a denotator in our framework, one that hopefully offers similar capabilities as Mazzola's denotator framework.

6.2. *Local and global compositions*

This section briefly discusses the challenge of integrating a theoretical framework for addressing local and global musical objects—akin to Mazzola's framework of *local* and *global compositions*—into our structure theory framework. Since Mazzola's theories of local and global compositions rely on his denotator framework, and we have not yet established a comparable denotator formalism in our framework, we currently lack the capacity to formulate such objects. Nonetheless, our aim is to eventually tackle the topic of local and global compositions and endeavor towards their formulation and classification within our framework.

In Mazzola's framework, local compositions are essentially subparts of forms.[22] For instance, a pitch-class set P can be conceived as a subset of $\mathbb{Z}_{12}$, corresponding to a subset $P \subset 0@\mathbb{Z}_{12}$ of zero-addressed points of $\mathbb{Z}_{12}$. Formally, local compositions in Mazzola's theory correspond to subfunctors of functors in $\mathbf{Mod}^{@}$. Specifically, a *local composition* is a denotator $D : A \rightsquigarrow F(x)$, where F is of type **Power**. For the coordinator form S of F, such a denotator corresponds (modulo its name) to a subfunctor $x \hookrightarrow @A \times Fun(S)$.

Classifying such local compositions involves situating them in a category and identifying methods for calculating their isomorphism classes, a task Mazzola has accomplished.

Global compositions, on the other hand, emerge from local compositions as the amalgamation or "gluing" of these local parts.[23] The idea is that these gluings form composite objects that exhibit varying degrees of complexity. While a formal discussion of global compositions in Mazzola's framework lies beyond the scope of this text, it should be noted that Mazzola has achieved a complete classification of global compositions within his framework.

Working towards a formulation and classification of such local and global objects in the context of our structure theory is an important task. However, settling on a formalism corresponding to Mazzola's denotators is crucial, as local compositions build upon denotators, and global compositions build upon local compositions. Once this formalism is established, we can then proceed to tackle the formulation and classification of local and global objects within our framework.

[22] See (Mazzola 2002, Chapter 7) for a formal treatment of local compositions.
[23] See (Mazzola 2002, Chapter 13) for a formal treatment of global compositions.

7. Conclusion

This paper introduces a theory of structure with the goal of establishing mathematical foundations for musical thinking. The core premise is that a rigorous definition of structure can serve as a foundational framework for musical thinking, particularly when we conceptualize musical phenomena as *structures in the context of sound and time*. This framework equips musicians, especially music theorists and composers, with a powerful set of tools. These tools not only facilitate the explicit formulation of thoughts but also enable the exploration of connections between theoretical constructions within the same mathematical context, namely the category $\mathbf{Rel}^{@}$.

A current challenge lies in grappling with the abstract notions of topos theory. While those familiar with Mazzola's work will find parallels in my framework, the specialized nature of this area poses a hurdle to widespread adoption within the musical research community. Despite this, the potential for significant musical developments remains substantial.

Reflecting on my own experience with the developed framework, it has proven immensely beneficial in both theoretical and compositional endeavors. The methodological approach offered by my structure theory enables the explicit articulation of various musical phenomena. Not only does it serve as a robust encoding tool, but it also suggests new ideas.

After acquiring the skill of working in $\mathbf{Rel}^{@}$, formulating musical ideas within this framework becomes simpler than less rigorous approaches. This aligns with the sentiments expressed in Section 2, where we discussed how firm foundations facilitate the systematic recovery of the meaning of certain constructions, alleviating the burden of mentally keeping track of every phase of construction. Explicit mathematical frameworks grounded in standard procedures are essential for achieving this goal.

In essence, my use of structure theory has significantly enhanced my ability to conceive new musical situations. This theory facilitates navigation through complex constructions within a precisely defined framework. By encoding the essential features of its intended objects, the formal framework alleviates the mental burden associated with tracking these features. Moreover, the reliable encoding of structural information not only alleviates cognitive overload, leading to a deeper theoretical understanding and advancements in both theory and composition, but also provides a formal basis for systematic comparisons between entities. This capability fosters the potential for significant theoretical progress.

Acknowledgements

I am grateful to Curtis Roads for offering constructive feedback on the initial draft of this paper. Furthermore, I extend my thanks to the editor, Darrell Conklin, and the two anonymous referees for their invaluable suggestions, which greatly enhanced the quality of the paper.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Do not change this. Production will take care of it if the paper is accepted.

References

Awodey, Steve. 2010. *Category Theory.* Oxford University Press.

Callender, Clifton, Ian Quinn, and Dmitri Tymoczko. 2008. "Generalized Voice-Leading Spaces." *Science* 320 (5874): 346–348.

Corry, Leo. 1992. "Nicolas Bourbaki and the Concept of Mathematical Structure." *Synthese* 92: 315–348.

Flieder, Drew. 2022. "Tools for Worldmaking: Universal Concepts for Mental Reality." `https://drewflieder.xyz/_writings/ToolsForWorldmaking.pdf`.

Forte, Allen. 1973. *The Structure of Atonal Music.* Vol. 304. Yale University Press.

Hodges, Wilfrid. 1997. *A Shorter Model Theory.* Cambridge university press.

Mac Lane, Saunders. 2013. *Categories for the Working Mathematician.* Vol. 5. Springer Science & Business Media.

Mac Lane, Saunders, and Ieke Moerdijk. 2012. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* Springer Science & Business Media.

Mazzola, Guerino. 2009. "Categorical gestures, the diamond conjecture, Lewin's question, and the Hammerklavier Sonata." *Journal of Mathematics and Music* 3 (1): 31–58.

Mazzola, Guerino, and Moreno Andreatta. 2007. "Diagrams, gestures and formulae in music." *Journal of Mathematics and Music* 1 (1): 23–46.

Mazzola, Guerino, René Guitart, Jocelyn Ho, Alex Lubet, Maria Mannone, Matt Rahaim, and Florian Thalmann. 2018. *The Topos of Music III: Gestures.* Springer.

Mazzola, Guerino et al. 2002. *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance.* Basel: Birkhäuser Verlag.

Noll, Thomas. 2005. "The Topos of Triads." *Grazer Mathematische Berichte* 347: 1–26.

Polansky, Larry. 1996. "Morphological Metrics." *Journal of New Music Research* 25 (4): 289–368.

Popoff, Alexandre, Moreno Andreatta, and Andrée Ehresmann. 2018. "Relational poly-Klumpenhouwer networks for transformational and voice-leading analysis." *Journal of Mathematics and Music* 12 (1): 35–55.

Tymoczko, Dmitri. 2010. *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice.* Oxford University Press.

Zheng, Linshujie, and Guerino Mazzola. 2023. *Classification of Musical Objects for Analysis and Composition.* Springer Nature.

Appendix A. Mathematical preliminaries

This appendix introduces concepts from category theory that are necessary for comprehending my structure theory. More thorough treatments can be found in texts such as (Awodey 2010) and (Mac Lane 2013). Definitions of *category*, *functor*, and *natural transformation* will be provided. Finally, we will delve into the Yoneda lemma, a crucial step in establishing the formal definition of structure in Section 4.3.

Informally, a category is a mathematical object comprised by two classes: (1) a collection C_0, referred to as *objects*, and (2) a collection C_1, referred to as *morphisms*, representing the mappings between objects in C_0. In essence, a category resembles a directed graph, with objects as nodes and morphisms as arrows between these nodes. The formal definition of a category is as follows.

Definition A.1 (Category) A *category* $\mathcal{C}$ consists of the following data:

- A class $\mathcal{C}_0$ of *objects*.
- A class $\mathcal{C}_1$ of *morphisms*.

- A binary operation $\circ : \mathcal{C}_1 \times \mathcal{C}_1 \to \mathcal{C}_1$ on the morphism class, referred to as *composition*. For objects X, Y, Z and morphisms $f : X \to Y$, $g : Y \to Z$, their composition is a unique morphism $g \circ f : X \to Z$ (often pronounced "g after f," since one first performs f, and then g). Furthermore, this binary operation must satisfy the following conditions:
  - **Identity**. For every object X, there exists a morphism $\mathrm{Id}_X : X \to X$. Such morphisms are specified by the fact that for any morphism $f : A \to B$, the following holds.

$$f \circ \mathrm{Id}_A = \mathrm{Id}_B \circ f$$

    Such a morphism Id_X is termed the *identity morphism* of X.
  - **Associativity**. For morphisms $f : A \to B$, $g : B \to C$, $h : C \to D$, we have the following equality.

$$h \circ (g \circ f) = (h \circ g) \circ f$$

Functors serve as morphisms between categories. Given that categories consist of collections of objects $\mathcal{C}_0$ and collections of morphisms $\mathcal{C}_1$, a morphism of categories (*functor*) $F : \mathcal{C} \to \mathcal{D}$ can be expressed as a pair of mappings $F_0 : \mathcal{C}_0 \to \mathcal{D}_0$ and $F_1 : \mathcal{C}_1 \to \mathcal{D}_1$ on the object and morphism classes, respectively. Moreover, a functor must preserve the structure of a category. Consequently, the formal definition of a functor is as follows.

Definition A.2 (Functor) Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A *functor* $F : \mathcal{C} \to \mathcal{D}$ constitutes a pair of mappings $F_0 : \mathcal{C}_0 \to \mathcal{D}_0$ and $F_1 : \mathcal{C}_1 \to \mathcal{D}_1$ satisfying the following conditions:

- For every identity morphism Id_X in $\mathcal{C}$, $F(\mathrm{Id}_X) = \mathrm{Id}_{F(X)}$.
- For morphisms $f : A \to B$, $g : B \to C$ in $\mathcal{C}_1$, it holds that $F(g \circ f) = F(g) \circ F(f)$.

A functor is called *covariant* if it preserves the direction of arrows and *contravariant* if it reverses them.

As functors are morphisms between categories, morphisms between functors are called *natural transformations*. As the structures we defined are specialized kinds of functors (presheaves), the morphisms of these structures arise as natural transformations of functors.

Definition A.3 (Natural transformation) Let $\mathcal{C}$ and $\mathcal{D}$ be categories, with $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{C} \to \mathcal{D}$ functors. A *natural transformation* $\eta : F \to G$ constitutes a class of morphisms satisfying the following conditions:

- For every object X in $\mathcal{C}$, there exists a morphism $\eta_X : F(X) \to G(X)$, called the *component* of η at X.
- The components must adhere to the rule that for every morphism $f : X \to Y$ in $\mathcal{C}$, the following equality holds.

$$\eta_Y \circ F(f) = G(f) \circ \eta_X$$

To express this categorically, we say that the following diagram *commutes*.

$$\begin{array}{ccc} F(X) & \xrightarrow{\eta_X} & G(X) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(Y) & \xrightarrow[\eta_Y]{} & G(Y) \end{array}$$

Lastly, we discuss the Yoneda lemma.[24] To formally present this idea, let us begin with its underlying intuition. In a category, we can conceptualize the objects as "entities" and the morphisms as "perspectives." For instance, a morphism $f : A \to B$ can be seen as a perspective providing insight into object B from the vantage point of object A. When focusing on an object X, a set of morphisms $\{f_i : A_i \to X\}_{i \in I}$ represents a collection of perspectives on X, each providing a distinct viewpoint.

Yoneda's insight lies in the revelation that an object X in a category $\mathcal{C}$ equates to all perspectives of X stemming from every other object in $\mathcal{C}$. This equivalence enables us to substitute objects X in $\mathcal{C}$ with entities derived from their perspectives. This process gives rise to a new category denoted by $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$. Consequently, instead of studying $\mathcal{C}$ we can study $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$, without any loss of information. While seemingly abstract, this substitution has proven its utility. Notably, by replacing objects with collections of perspectives, the new category inherits the property of being a topos, enabling significant constructions not always possible in the original category.

To begin the presentation of the Yoneda lemma, we first introduce some definitions. The first definition is that of a *hom functor*, which formally defines the replacement of X with its collection of perspectives. Once we establish hom functors, the concept of a *presheaf* emerges, which generalizes the concept of a hom functor.

Before we proceed with definitions, a brief technical note. A category $\mathcal{C}$ is considered *locally small* when, for any two objects X and Y in $\mathcal{C}$, the collection $\text{Hom}_{\mathcal{C}}(X, Y)$ of morphisms from X to Y forms a *set*.[25] (This detail primarily concerns the mathematically inclined reader and can be skipped without sacrificing an understanding of the essential concepts presented in this work.)

Now we proceed to define a concept that is pivotal for comprehending the Yoneda lemma.

Definition A.4 (Contravariant hom functor) Let $\mathcal{C}$ be a locally small category. A *contravariant hom functor* is a functor, notated in the following way.

$$\text{Hom}_{\mathcal{C}}(-, X) : \mathcal{C} \to \mathbf{Set}$$

It is defined as follows:

- Each object A in $\mathcal{C}$ maps to the hom set $\text{Hom}_{\mathcal{C}}(A, X)$.
- Each morphism $f : A \to B$ in $\mathcal{C}$ maps to the following set function.

$$\begin{array}{cccc} \text{Hom}_{\mathcal{C}}(f, X) : & \text{Hom}_{\mathcal{C}}(B, X) & \to & \text{Hom}_{\mathcal{C}}(A, X) \\ & g & \mapsto & g \circ f \end{array}$$

[24] See (Mac Lane 2013, pages 59–62) for a thorough treatment.

[25] This is a technical issue regarding size issues in set theory. A collection that is too big may not be a *set*, but a *proper class*.

A *presheaf* is a generalization of a contravariant hom functor, as it is defined simply as a contravariant functor into **Set**. In fact, contravariant hom functors are also termed *representable presheaves*. This "representability" concept essentially signifies the ability to "represent" an object X via the network of all perspectives of X. If, however, the functor contains only a subset of the perspectives, it is not representable, indicating it carries only partial information of X.

The category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ encompasses all presheaves over $\mathcal{C}$ for objects and natural transformations between presheaves for morphisms. Converting each object A in $\mathcal{C}$ to its representable presheaf $\mathrm{Hom}_{\mathcal{C}}(-, A)$ extends to a functor $Y : \mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. The Yoneda lemma implies that this functor is full and faithful,[26] embedding $\mathcal{C}$ within its category of presheaves. Consequently, replacing $\mathcal{C}$ with its category of presheaves preserves all information and often bestows the new category with advantageous features absent in $\mathcal{C}$ alone—such as limits, colimits, a subobject classifier, and internal homs.

Appendix B. Concrete applications and informal discussions of the $\mathbf{Rel}^{@}$ topos

This appendix provides concrete context and constructions for working within the $\mathbf{Rel}^{@}$ topos. It is intended for readers who appreciate the framework's goal of offering a universal structure-encoding system but are not yet comfortable with the topos-theoretic technical formalities required for full theoretical rigor. This appendix will demonstrate how the framework naturally aligns with musical thought.

The approach here will be as informal as possible, focusing on how I have personally employed the framework in my musical thinking.

In Section B.1, I will briefly discuss the connection between the categories $\mathbf{Rel}^{@}$ and **Set**, highlighting their similarities. Since readers are assumed to be familiar with basic set theory, this comparison should offer an intuitive entry point into the new framework.

In Section B.2, we will apply the framework to a concrete musical construction. Then, in Section B.2.1, I will explain how I used the framework in composing *Reverie* for solo cello. This concrete example aims to demonstrate how the framework has supported my musical endeavors, making it more relatable.

Finally, in Section B.3, I share my concluding thoughts, aiming to provide useful context for those interested in working within the proposed framework.

B.1. *Thinking in terms of sets*

An intuitive entry point into our framework is to consider the category **Set** of sets. The category $\mathbf{Rel}^{@}$ is quite similar to **Set**, with the crucial distinction that the objects in $\mathbf{Rel}^{@}$ have additional structure. While we will not delve into the complex relationship between **Set** and $\mathbf{Rel}^{@}$ here, it is important to note that morphisms between structured sets in $\mathbf{Rel}^{@}$ generally offer as much freedom as set maps in **Set**.

For example, consider sets X and Y. If we equip these sets with additional structure, resulting in objects $\overline{X}$ and $\overline{Y}$ in $\mathbf{Rel}^{@}$, then any set map $f : X \to Y$ can be translated into a corresponding map $\overline{f} : \overline{X} \to \overline{Y}$ on the structured sets. Therefore even though the sets in $\mathbf{Rel}^{@}$ have additional structure, the morphisms between these structured sets are not required to preserve their structure. For instance, $\overline{f} : \overline{X} \to \overline{Y}$ can be a mapping between two groups that is not a group homomorphism, or $\overline{X}$ and $\overline{Y}$ can possess entirely

[26] See (Mac Lane 2013) for definitions of *full* and *faithful* functors.

different kinds of structure. Thus, operations on sets in **Set** can be similarly performed on the structured sets in $\mathbf{Rel}^{@}$. The structure framework offers a method for encoding additional structure while maintaining the intuitive ease of working with sets.

In the subsequent subsections, I will treat the constructions as sets and declare "by fiat" when they incorporate additional structure. This approach aims to enhance intuitiveness and manageability for readers who have not yet mastered the abstract details of the formal setup. Readers are encouraged to revisit the formal framework presented in the main text to formalize these informal constructions.

B.2. *Applying the structure framework*

Before getting started, I would like to outline the early phases of my compositional process and how theorization is involved.

(1) In the germinal phases of a composition, I imagine certain kinds of musical phenomena.
(2) After conceiving these phenomena, I formulate general structural descriptions that explicate their properties of interest at a conceptual level.
(3) I begin again at Step 1, informed by the theoretical formulations from Step 2.

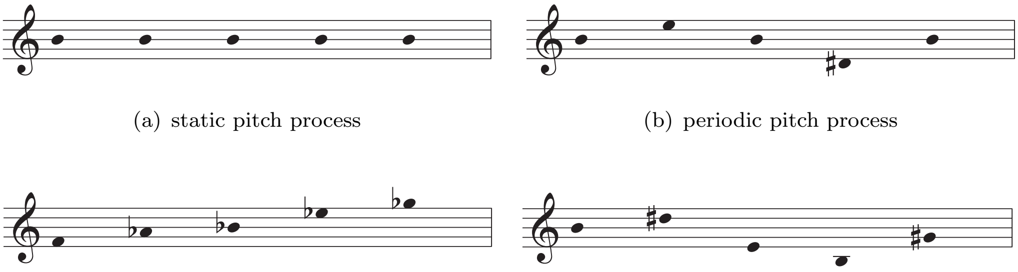
This process does not always lead to fully explicated musical archetypes. Often, I settle on a formulation that is simple and intelligible enough to work with effectively. For myself, and likely for many others, the compositional process often spawns theoretical ideas. These ideas are theoretically formulated enough to provide a framework for efficient work, while working out the full theoretical implications is postponed. Later, I can approach these ideas from a purely speculative standpoint, fully explicating the general principles formulated during composition. This theoretical explication, in turn, enhances my ability to conceive novel musical phenomena for future projects. Thus, there is a feedback loop between the spontaneous occurrences in the compositional phase and the theoretical explications that follow.

B.2.1. Composing Reverie

To illustrate the application of the structure theory framework in compositional thinking, I will walk the reader through the compositional process of a piece I composed in 2023 titled *Reverie* for solo cello. A distinctive feature of this piece is the concept of the "compositional unit." Each compositional unit represents a complete musical idea, and the progression of these units shapes the overall form of the piece. These units are defined by associating one of the following four process types to each of the four parameters of pitch, amplitude, timbre, and duration.

- **Static**: There is no change in the parameter value throughout the event.
- **Periodic**: The parameter value recurs at a regular interval, or there is repetition of a symmetric pattern.
- **Trajectorial**: There is a gradual transition from a starting value a to an ending value b.
- **Random**: The parameter values follow no discernible pattern.

For instance, a trajectorial pitch process might involve the pitch contour following a general ascending or descending trajectory over the course of a compositional unit. In the duration domain, a trajectorial process might manifest as an accelerando or ritardando, or simply a gradual increase or decrease in rhythmic density. The static process type is the



(a) static pitch process

(b) periodic pitch process

(c) trajectorial pitch process

(d) random pitch process

Figure B1. Examples of the four types of pitch processes

easiest to define, as it generally involves a fixed value over the course of a compositional unit; for example, if the pitch is static, the event will feature just one pitch. Figure B1 shows an example of each of the four process types manifested in the pitch domain.[27]

The question now is, how can these musical conceptions be formulated explicitly? My structure theory framework provides the technical means to explicitly formulate these concepts. Hence, I will guide the reader through an overview of how I expressed these musical processes using my structure theory. Section 5 of the main paper provides examples of the systematic and abstract approach to working with this framework. However, to keep things more concrete, I will adhere less to strict notational rigor and aim to ground the construction steps in a more immediately recognizable way for those less acquainted with the topos-theoretic specifics of the formal setup.

For starters, let us consider how to explicate what is meant by a “process,” as mentioned above. A process is a time-indexed selection of parameter values from some parameter space X, in other words, a sequence of values from X. It is natural to define such a phenomenon by a mapping $\phi : N \to X$, where N is the collection of natural numbers from 1 to N (inclusive). We may call such a map $\phi : N \to X$ an X-*valued process*.

It is then natural to ask, what are the structural requirements of the objects N and X that will enable one to define such a process? Let us start with N. The structure we require for N is simply a total order structure, since an X-valued process is a sequence of values from X. Therefore, the value $\phi(i) \in X$ of the ith ordinal in N gives the ith event of the process.

To make this concrete, let us consider an example where $X = \mathbb{Z}$, the collection of pitches in twelve-tone equal temperament (setting 0 to middle-C, 1 to middle-C♯, etc.), and let 4 denote the totally ordered set of integers from 1 to 4. Then an X-valued process $\phi : 4 \to X$ is a process of selecting four pitches (see Figure B2).

At this juncture, we have addressed the necessary structural requirements for the domain structure N of a process $\phi : N \to X$. We have established that all we need is for N to be totally ordered. It is worth briefly noting that to accommodate *continuous* processes, we would need N to possess some topological structure as well. For instance, if we desire a process to represent a continuous glissando curve rather than discrete pitch

[27]This approach to composition may remind one of the methods prevalent in the 1950s and 1960s, with prominent composers such as Stockhausen, Xenakis, Cage, and Babbitt. These figures are part of my musical heritage, and much of my musical thinking has evolved towards generalizing across their diverse approaches to extract musical principles that lead to complex and intelligible forms of musical organization. Although many of these radical composers' initial attempts were often, in my view, rife with error, their visionary attitudes remain influential. My desire is to continue this musical attitude, even if practiced only by a modest group of dedicated artists. The attitude expressed by these radical composers was inherently self-critical, suggesting a potential marriage of art and science.

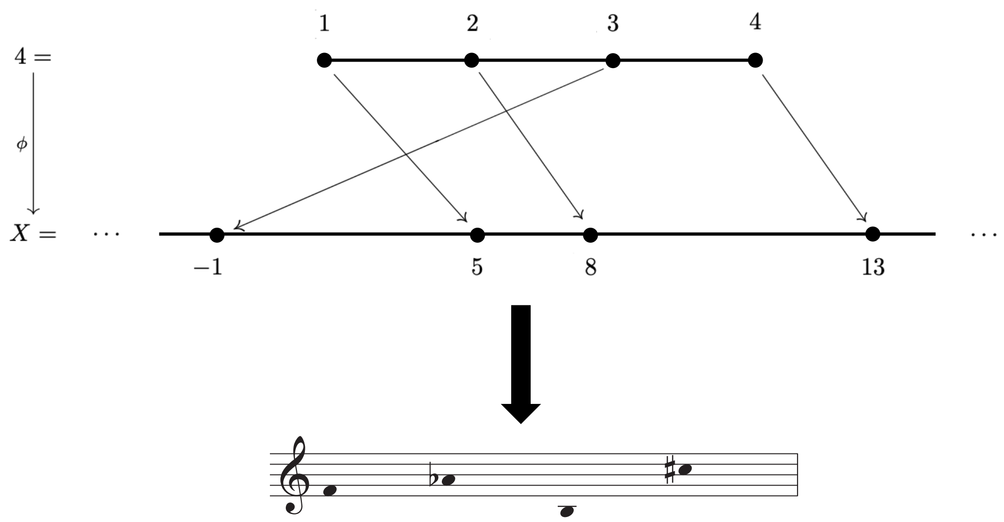


Figure B2. This figure demonstrates how an X-valued process $\phi : 4 \to X$ corresponds to an ordered selection of pitches.

selections, we would require N to be something akin to the unit interval $[0, 1]$ equipped with topological structure. However, for the present, we set aside such generalizations.

The next inquiry we must address is the structural requirements of the codomain structure X—the parameter space of the process—to determine when a process is static, periodic, trajectorial, or random. Let us begin by considering the prerequisites for X to ascertain whether a process $\phi : N \to X$ is trajectorial. This entails a continual transition from one point $x \in X$ to another point $y \in X$ over the course of the process. It is evident that X must be endowed with structural information to establish a notion of "closeness," allowing us to progressively approach y with each step in the process. When discussing the phenomenon of "closeness" in mathematics, this typically involves a topological context, necessitating that X is equipped with an appropriate topological structure.

For instance, when $X = \mathbb{Z}$ represents pitch space, we can achieve the required topological structure through the metric defined by $d(x, y) = y - x$, representing the pitch interval between two pitches. To further specify the structure of X, we may also equip it with group structure, as commonly employed in musical analysis and composition. However, such group structure is not strictly necessary for defining a trajectorial process.

The topological structure of the codomain suffices to define the remaining process types. In brief, we can delineate each process type as follows:[28]

(1) **Static**: A static process $\phi : N \to X$ is characterized by every $i \in N$ mapping to the same value $x \in X$. In other words, ϕ constitutes a constant map.
(2) **Periodic**: In a periodic process $\phi : N \to X$, a value $x \in X$ recurs periodically. Specifically, for a fixed integer k and positive integer $n < N$, there exists an $x \in X$ such that

$$\phi(i) = \phi(i + k) = \cdots = \phi(i + nk) = x.$$

Here, k denotes the period of the process, and i represents its initial onset. Another

[28]The following descriptions are simplifications. In composing the piece, numerous variations of each process type were used throughout. However, detailing each variation would be time-consuming. Thus, we prioritize brevity to offer the reader a basic understanding.

kind of periodic process $\phi : N \to X$ displays symmetrical behavior, such as when the values of the second half are the reverse of those in the first half.

(3) **Trajectorial**: A trajectorial process $\phi : N \to X$ begins with a start value $x \in X$ and concludes with an end value $y \in X$, with each subsequent value drawing nearer to y. This trajectory is determined by a distance function $d : X \times X \to X$, where the condition $|d(\phi(1), y)| > |d(\phi(2), y)| > \cdots > |d(\phi(N-1), y)|$ ensures convergence towards y. While this requirement specifies a strictly converging sequence, it can be relaxed by stipulating a general trajectory towards y, rather than mandating that every subsequent value of the process approaches closer to y.

(4) **Random**: A random process $\phi : N \to X$ is one that is neither static, periodic, nor trajectorial.

Our essential components for formulating the four process types are therefore: (1) ensuring the domain structure of a process $\phi : N \to X$ is totally ordered, and (2) equipping the codomain with topological structure. Once these are established, specifying values for N and X defines a particular process.

Given that a generic process may consist of any number $N \in \mathbb{N}$ of events, our aim is to devise a general scheme for a process. To accomplish this, we take the collection of mappings $\mathrm{Hom}(N, X)$ for each $N \in \mathbb{N}$. Then, we take the coproduct (disjoint union) of all such processes, resulting in the following collection.

$$\mathcal{X} = \coprod_{N \in \mathbb{N}} \mathrm{Hom}(N, X)$$

As "static," "periodic," "trajectorial," and "random" are properties of processes, and $\mathcal{X}$ represents the collection of all X-valued processes, each of the aforementioned processes defines a subset of $\mathcal{X}$—specifically, the subset consisting of processes that exhibit said property. Therefore, we can regard each of the four properties as a subset of $\mathcal{X}$.

Now let us define the parameter spaces of pitch, amplitude, timbre, and rhythm used in *Reverie*. As mentioned, any of the process types can occur in the pitch, amplitude, timbral, and duration domains.

- **Pitch**: The pitch domain is $\mathrm{P} = \mathbb{Z}$, with 0 denoting middle-C, 1 denoting middle C-$\sharp$, and so on.
- **Amplitude and Duration**: Both the amplitude and duration domains are $\mathrm{A} = \mathrm{D} = \mathbb{R}$, where a real number in the amplitude domain represents decibels, and a real number in the duration domain represents seconds.
- **Timbre**: The timbre parameter T is more complex and less straightforward to define structurally. I conceptualized it as a "degree of noise" parameter, with different playing techniques on the cello corresponding to varying degrees of noise. For example, "arco ordinario" is considered the least noisy, "sul ponticello" has a medium level of noise, and "pizzicato" is the noisiest.

Hence, the collections of processes over each parameter are given by $\coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{P})$, $\coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{A})$, $\coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{T})$, and $\coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{D})$. Each collection represents sequences of pitches, amplitudes, timbres, and durations, respectively. To denote static, periodic, trajectorial, and random processes, I will use the following symbols: $\cdot$, $\sim$, $\to$, $\#$. For example, $\mathrm{P}(\cdot) \subset \coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{P})$ denotes the subcollection of static pitch processes, $\mathrm{P}(\sim) \subset \coprod_{N\in\mathbb{N}} \mathrm{Hom}(N, \mathrm{P})$ denotes the subcollection of periodic pitch processes, and so on.

A compositional unit in my piece *Reverie* is determined by a quadruple

$(\mathrm{P}(\phi_1), \mathrm{A}(\phi_2), \mathrm{T}(\phi_3), \mathrm{D}(\phi_4))$, where each ϕ_i is either $\cdot$, $\sim$, $\rightarrow$, or #. The compositional unit occurs at the measure level, so each new measure features a new quadruple. For instance, the first measure of the piece has the compositional unit $(\mathrm{P}(\cdot), \mathrm{A}(\cdot), \mathrm{T}(\cdot), \mathrm{D}(\cdot))$, meaning that every parameter value deploys a static process.

The final topic I would like to discuss in this section is how compositional units were organized into the large-scale formal structure of the piece (see Figure B3 for an excerpt of the formal scheme used during composition). Often, it is desirable to control the evolution of each parameter according to its own compositional strategy. For instance, the evolution of pitch processes might follow a different scheme than that of amplitude processes. Before composing the sequence of quadruples, each parameter's evolution is composed individually. Once each evolution is determined, they are "merged" into quadruples that define what happens at specific moments in the piece. This method can be naturally encoded in our structure framework, as we will now see.

In *Reverie*, there are 133 compositional units. Thus, we define a totally ordered set labeled as 133 containing 133 elements and construct the sequence of pitch processes for each compositional unit. We employ the coproduct construction to obtain the disjoint union of all pitch processes.

$$\mathbf{P} \coloneqq \mathrm{P}(\cdot) \amalg \mathrm{P}(\sim) \amalg \mathrm{P}(\rightarrow) \amalg \mathrm{P}(\#)$$

An element of this collection is either a static, periodic, trajectorial, or random pitch process. We then define a mapping that provides the sequence of pitch processes for each compositional unit.

$$\Phi_{\mathrm{P}} : 133 \rightarrow \mathbf{P}$$

(A perceptive reader might notice that Φ itself can be conceived as a $\mathbf{P}$-valued process, enabling the definition of novel kinds of processes over it. Indeed, such strategies were employed in the composition of the piece.)

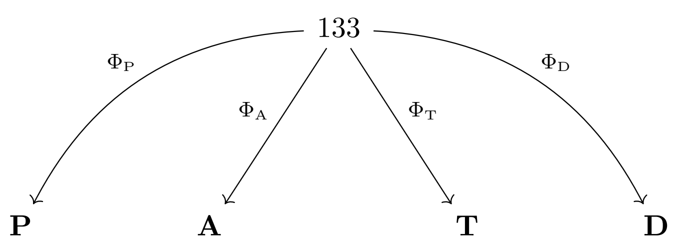
We repeat these steps to specify the sequence of amplitude, timbre, and duration process types for each compositional unit:

$$\Phi_{\mathrm{A}} : 133 \rightarrow \mathbf{A},$$

$$\Phi_{\mathrm{T}} : 133 \rightarrow \mathbf{T},$$

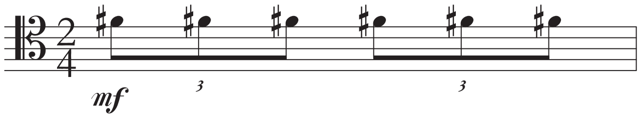
$$\Phi_{\mathrm{D}} : 133 \rightarrow \mathbf{D}.$$

To merge these independent parameter evolutions into quadruples that define each compositional unit, we take the limit of the following diagram.

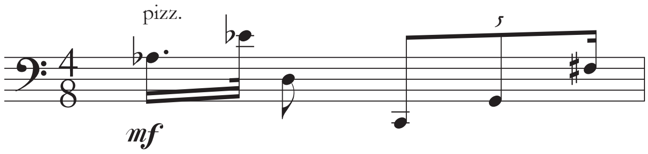


		Section Name		F				
Global Information		System Duration		36♩				
		Rhythmic Scheme: System Level		Trajectorial				
		System		5				
		Rhythmic Scheme: Phrase Level		Trajectorial				
		Phrase Number		11	12	13	14	15
		Phrase Length		4.5	6.75	6.75	9	9
	Realization Types	Options		Systems & Phrases				
		Type		System				
Cello Information	Process Parameters	Process Type Evolution		Periodic				
		Cello Parameters	Pitch	P(#)	P(#)	P(~)	P(~)	P(~)
			Amplitude	A(#)	A(.)	A(~)	A(.)	A(->)
			Timbre	T(#)	T(#)	T(~)	T(#)	T(#)
			Duration	D(#)	D(.)	D(~)	D(.)	D(.)
	Phrase Parameters	Pitch	Quantity	2				
			High					
			Mid	P(=)	P(=)	P(=)	P(=)	P(=)
			Low	P(−)	P(−)	P(−)	P(−)	P(−)
		Amplitude	Quantity	1				
			Loud					
			Medium					
			Quiet	A(−)				
		Timbre	Quantity	3				
			Noisy	T(+)	T(+)	T(+)	T(+)	T(+)
			Medium	T(=)	T(=)	T(=)	T(=)	T(=)
			Pure	T(−)	T(−)	T(−)	T(−)	T(−)
		Duration	Quantity	1				
			Short (Dense)					
			Medium					
			Long (Sparse)	D(−)				

Figure B3. The chart illustrates the formal structure of the final five sections of *Reverie*, out of a total of 133 sections. The compositional units are highlighted in yellow.

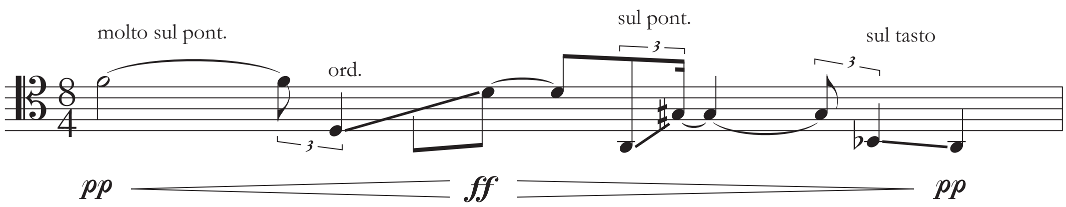


(a) first compositional unit: $(\mathrm{P}(\cdot), \mathrm{A}(\cdot), \mathrm{T}(\cdot), \mathrm{D}(\cdot))$

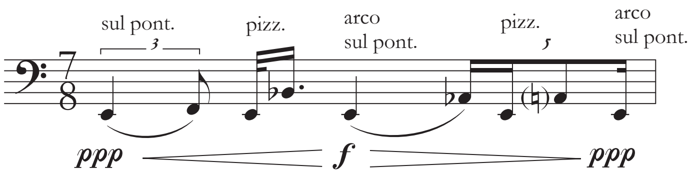


(b) seventh compositional unit: $(\mathrm{P}(\#), \mathrm{A}(\cdot), \mathrm{T}(\cdot), \mathrm{D}(\#))$

(c) forty-first compositional unit: $(\mathrm{P}(\rightarrow), \mathrm{A}(\rightarrow), \mathrm{T}(\rightarrow), \mathrm{D}(\rightarrow))$



(d) seventy-eighth compositional unit: $(\mathrm{P}(\rightarrow), \mathrm{A}(\sim), \mathrm{T}(\#), \mathrm{D}(\#))$



(e) eighty-first compositional unit: $(\mathrm{P}(\sim), \mathrm{A}(\sim), \mathrm{T}(\sim), \mathrm{D}(\#))$

Figure B4. The above excerpts present realizations of five of the 133 compositional units of *Reverie*.

The limit of this diagram encodes the sequence of compositional units that govern the form of the piece. Specifically, it consists of quadruples of the following form.

$$(i, \Phi_{\mathrm{P}}(i), \Phi_{\mathrm{A}}(i), \Phi_{\mathrm{T}}(i), \Phi_{\mathrm{D}}(i))$$

Here, the first coordinate i specifies the ith compositional unit, and the values $\Phi_{\mathrm{P}}(i)$, $\Phi_{\mathrm{A}}(i)$, $\Phi_{\mathrm{T}}(i)$, and $\Phi_{\mathrm{D}}(i)$ provide the ith process type over the domains of pitch, amplitude, timbre, and duration, respectively. Examples of excerpts of music corresponding to some compositional units of *Reverie* are provided in Figure B4.

I have illustrated how the structural approach informed my composition of *Reverie* for solo cello. However, the formulations discussed here do not cover all those employed in the piece—numerous other methods incorporating the structural approach were utilized. Due to constraints, I cannot delve into all of them here.

B.3. *Final thoughts*

I should note that my compositional thinking has followed a structural approach for years, even before I was well-versed enough in mathematics to formulate my ideas explicitly. The structural approach presented in this paper should not be seen as an attempt to constrain the spontaneity of musical thought into rigid procedures, but rather as a way to lend explicit form to the conceptions that arise from the creative activity of music composition and theoretical contemplation. When a musical object can be formulated mathematically, it enters a layer of imagination that is more concrete than the immediate intuitions that inspired the initial conceptions. This increased objectivity allows for more sophisticated musical treatment.

I wish I could devote pages to discussing the development of musical ideas and how this evolution is rooted in formal procedures. However, for the sake of brevity, I encourage readers to explore this relationship through their own imagination. The idea is that formal sign systems offer systematic procedures that facilitate the translation of general conceptions into abstract objects. The objectivity of these abstract objects, in contrast to the less objective nature of immediate conceptions, fosters a scientific approach that promotes the growth of thought in general, and musical thought in particular.

For instance, in years past, I might have considered the concept of a "trajectorial process," involving movement from point a to point b. However, I would not have had the depth of understanding to articulate the structural conditions inherent in this process, especially not in the clear and concise terms outlined in Section B.2.1. It is doubtful that I would have explicitly recognized it as a mapping from a total order to another object with topological structure.

Within this structural framework, once a concept is explicitly formulated, it determines an abstract object (i.e., an object in $\mathbf{Rel}^{@}$). This object can then be explored and experimented with in ways that less formalized concepts cannot.

I hope this appendix has been beneficial to the reader. Those inclined to contribute to this nascent formal framework are warmly encouraged to do so.