The 2-Categorical Position of Music Among Time-Value Art Forms

Carlo Perassi

June 8, 2026

Abstract

We promote the parameter pair $(\mathbf{B}, \mathcal{V})$ of the previous framework to a variable and organise the resulting time-value art forms into a 2-category $\mathfrak{A}$. The chromatic lattice, viewed as a finite directed graph, distinguishes a reference object $X_{\text{mus}} \in \mathfrak{A}$. The question *"is X music?"* is recast as a structural question about the hom-category $\mathfrak{A}(X, X_{\text{mus}})$: it admits faithful musicalisations, fully faithful ones, none. We classify several borderline cases — paired poetry, dialogue, Gregorian chant, classical Chinese verse, *Sprechstimme*, birdsong, ritual percussion, a calling flock — and read off, in each case, what the verdict depends on.

1 Introduction

This article treats music as a structural predicate rather than as an intuitive label. By encoding temporal organization and value systems as categorical data, it asks which mappings make an art form musical and which assumptions a verdict depends on. The goal is not to settle every aesthetic case, but to make the source of each disagreement visible in the formal model.

2 From parameters to a 2-category of frames

The previous framework parametrises rhythm, melody, and harmony by a pair $(\mathbf{B}, \mathcal{V})$ of small categories. It treats the pair as fixed. Here we let it vary.

Definition 2.1 (The 2-category of time-value frames)**.** The 2-category $\mathfrak{A}$ has:

(i) *Objects* (called *frames*): pairs $X = (\mathbf{B}_X, \mathcal{V}_X)$ of small categories.

(ii) *1-morphisms* $X \to Y$: pairs (β, ν) where $\beta : \mathbf{B}_X \to \mathbf{B}_Y$ and $\nu : \mathcal{V}_X \to \mathcal{V}_Y$ are functors.

(iii) *2-morphisms* $(\beta, \nu) \Rightarrow (\beta', \nu')$: pairs (σ, τ) of natural transformations $\sigma : \beta \Rightarrow \beta'$ and $\tau : \nu \Rightarrow \nu'$.

Composition and units are component-wise.

Construction 2.2 (Functoriality of rhythm, melody, harmony)**.** The constructions of the previous framework extend to 2-functors $\mathfrak{A} \to \mathbf{Cat}$:

$$\begin{aligned} \mathbf{Rhy} &: X \mapsto \mathbf{Rhy}(\mathbf{B}_X), \\ \mathbf{Mel} &: X \mapsto \mathbf{Mel}(\mathbf{B}_X, \mathcal{V}_X), \\ \mathbf{Harm}_{\mathcal{J}} &: X \mapsto \mathbf{Harm}_{\mathcal{J}}(\mathbf{B}_X, \mathcal{V}_X). \end{aligned}$$

A 1-morphism $(\beta, \nu) : X \to Y$ acts on rhythms by post-composition with β, on melodies by post-composition with (β, ν) in each leg of the span, and on harmonies component-wise on each voice.

Remark. This is the meta-theoretic move: every art form whose phenomenology is "events in some base, labelled by some values" is now an object of $\mathfrak{A}$, and the rhythm/melody/harmony machinery applies to all of them uniformly. What distinguishes *music* from its neighbours is not the machinery but the choice of frame.

3 The chromatic graph

The reference frame for music is built from a graph, not just a set.

Definition 3.1 (Chromatic graph)**.** The *chromatic graph* Γ has vertex set $\mathbb{Z}$ and a single directed edge $n \to n{+}1$ for each $n \in \mathbb{Z}$. Its *free category* is the chromatic lattice

$$\mathbf{Pitch} \;:=\; \mathrm{Free}(\Gamma).$$

Concretely, $\mathrm{Hom}_{\mathbf{Pitch}}(m, n) = \{*\}$ if $n \geq m$ and $\emptyset$ otherwise; composition is the unique map. Equivalently, $\mathbf{Pitch} \cong (\mathbb{Z}, \leq)$ as a thin category.

Remark (Why graph and not just set)*.* $\mathbf{Pitch}$ as a discrete set $\mathbf{Disc}(\mathbb{Z})$ records pitch identity only. $\mathbf{Pitch}$ as the free category on Γ additionally records *intervallic order*: the unique morphism $m \to n$ exists iff n lies above m, and the length of the shortest factorisation is $n - m$ semitones. Functors $\mathcal{V}_X \to \mathbf{Pitch}$ must therefore preserve whatever ordered structure $\mathcal{V}_X$ carries, not merely the underlying set of values. This distinction is the *entire content* of the question that follows.

Definition 3.2 (Finite ambitus, groupoidification)**.** For $a \leq b$ in $\mathbb{Z}$, $\mathbf{Pitch}_{[a,b]}$ is the full subcategory on $\{a, a{+}1, \ldots, b\}$. The *interval groupoid* $\mathbf{Pitch}^{\sim}$ is the localisation of $\mathbf{Pitch}$ at all morphisms; it has the same objects with $\mathrm{Hom}_{\mathbf{Pitch}^{\sim}}(m, n) = \{*\}$ for all m, n. Inversion symmetry — descending intervals as inverses of ascending ones — lives in $\mathbf{Pitch}^{\sim}$.

Definition 3.3 (Musical reference frame)**.** The *musical reference frame* is

$$X_{\mathrm{mus}} \;:=\; \big((\mathbb{R}_{\geq 0}, \leq),\ \mathbf{Pitch}\big) \;\in\; \mathfrak{A}.$$

4 Musicalisation and the musical predicate

Definition 4.1 (Musicalisation)**.** A *musicalisation* of a frame $X \in \mathfrak{A}$ is a 1-morphism

$$\Phi = (\beta, \nu) : X \longrightarrow X_{\mathrm{mus}}$$

in $\mathfrak{A}$. The *musicalisation category* of X is the hom-category $\mathfrak{A}(X, X_{\mathrm{mus}})$, whose objects are musicalisations and whose morphisms are 2-cells.

Definition 4.2 (The musical predicate)**.** A frame X is:

- *transcribable* if $\mathfrak{A}(X, X_{\mathrm{mus}})$ is non-empty;
- *musical* if it contains a 1-morphism (β, ν) with both β and ν faithful;
- *strictly musical* if it contains one with both β and ν fully faithful;

- *canonically musical* if the connected component of any faithful Φ is unique up to a chosen 2-isomorphism class;
- *amusical* if $\mathfrak{A}(X, X_{\text{mus}}) = \emptyset$.

Remark (The four-step ladder)*.* The conditions above form a ladder:

$$\text{amusical} \;<\; \text{transcribable} \;<\; \text{musical} \;<\; \text{strictly musical},$$

with *canonically musical* as a transverse property concerning the *number* of faithful musicalisations rather than their existence. Most everyday verdicts about "what is music" confuse these levels; the framework lets us separate them.

5 Worked examples

In every example below, the time-leg β is the identity inclusion $(\mathbb{R}_{\geq 0}, \leq) \hookrightarrow (\mathbb{R}_{\geq 0}, \leq)$. The verdict is therefore controlled by the value-leg $\nu : \mathcal{V}_X \to \mathbf{Pitch}$.

5.1 Two poets in alternation

Let $\mathcal{V}^{\flat}_{\text{phon}} := \mathbf{Disc}(P)$, the discrete category on the finite set P of phonemes of the poets' language. Any injection $\nu : P \hookrightarrow \mathbb{Z}$ — alphabetical, frequential, arbitrary — gives a fully faithful functor $\mathcal{V}^{\flat}_{\text{phon}} \to \mathbf{Pitch}$, because both source and target are discrete on their objects.

Verdict. Two poets are *strictly musical* but *not canonically musical*: the musicalisation exists, even fully faithfully, but no choice of ν is privileged.

The framework here corrects a common intuition: the familiar judgement *"poetry is not music"* does not rest on a structural failure but on the absence of a *distinguished* musicalisation. Cultures that fix one (e.g. liturgical psalmody, Vedic recitation) immediately upgrade the same raw structure to canonically musical.

5.2 Conversation

Now enrich the value category to $\mathcal{V}_{\text{conv}}$: same objects P, but with non-identity morphisms encoding minimal-pair relations ($/p/ \to /b/$ for voicing, $/n/ \to /m/$ for place, etc.). A functor $\nu : \mathcal{V}_{\text{conv}} \to \mathbf{Pitch}$ must send each minimal-pair morphism to a chromatic morphism, i.e., must respect a partial order on P compatible with the chromatic order on $\mathbb{Z}$. The minimal-pair graph is in general not a sub-poset of $(\mathbb{Z}, \leq)$ — it has cycles, incomparable pairs, and no consistent direction.

Verdict. Conversation is *transcribable* (drop the morphisms, fall back to the discrete case) but *not musical*: no faithful ν respects the linguistic structure.

The boundary between this and the poets case is exactly how richly one models $\mathcal{V}$: stripping the value category to its discrete shadow is the act of treating speech as music.

5.3 Gregorian chant

$\mathcal{V}_{\text{chant}} := \mathbf{Pitch}_{[a,b]}$ for some Dorian or Mixolydian ambitus. The inclusion $\mathcal{V}_{\text{chant}} \hookrightarrow \mathbf{Pitch}$ is fully faithful and even essentially surjective onto its image.

Verdict. Gregorian chant is *strictly musical* and *canonically musical*: the musicalisation is the inclusion, with no choice involved.

5.4 Classical Chinese verse

$\mathcal{V}_{\text{verse}} := \mathcal{V}^{\flat}_{\text{phon}} \times \mathbf{Disc}(\mathbb{Z}/4)$, where the second factor records lexical tone (level, rising, departing, entering). Two natural musicalisations present themselves:

$$\nu_{\text{tone}} : \mathcal{V}_{\text{verse}} \xrightarrow{\text{pr}_2} \mathbf{Disc}(\mathbb{Z}/4) \hookrightarrow \mathbf{Pitch}/12\mathbb{Z},$$
$$\nu_{\text{full}} : \mathcal{V}_{\text{verse}} \xrightarrow{\text{any injection}} \mathbf{Pitch}.$$

The first is canonical but not faithful (it forgets phonemes). The second is faithful but not canonical (any injection works).

Verdict. Classical Chinese verse is *musical* and admits a *canonical quotient-musicalisation* via tone, but not a canonical full musicalisation. The genre's lived ambiguity between "poetry" and "song" is exactly this split.

5.5 Sprechstimme

$\mathcal{V}_{\text{spr}} := \mathbf{Pitch} \times \mathbf{Disc}(\{\text{spoken}\})$. The first projection $\text{pr}_1 : \mathcal{V}_{\text{spr}} \to \mathbf{Pitch}$ is fully faithful (the second factor is terminal).

Verdict. Sprechstimme is *strictly musical* and *canonically musical.* The `spoken` marker is a stylistic decoration that the framework correctly recognises as inessential to the classification.

5.6 Birdsong

$\mathcal{V}_{\text{bird}} := (\mathbb{R}, \leq)$, continuous pitch contour with its order. A faithful functor $(\mathbb{R}, \leq) \to (\mathbb{Z}, \leq) \cong \mathbf{Pitch}$ would require an order-injection of an uncountable set into a countable one — impossible.

Verdict. Birdsong is *transcribable* (the rounding map $\mathbb{R} \to \mathbb{Z}$ is a non-faithful musicalisation) but *not musical.* Western notation of birdsong is precisely this quantisation, and the loss of information is structural, not artisanal.

5.7 Ritual percussion of arbitrary timbres

$\mathcal{V}_{\text{drum}} := \mathbf{Disc}(T)$ for a finite set T of timbres. This is the two-poets situation again: any injection $T \to \mathbb{Z}$ is a fully faithful musicalisation; none is canonical.

Verdict. Strictly musical, not canonically musical. The standard practice of assigning timbres to lines of a percussion staff is the social act of fixing ν by convention.

5.8 A flock of birds calling simultaneously

This is a harmony in $\mathbf{Harm}_{\mathcal{J}}(\mathbf{B}, \mathcal{V}_{\text{bird}})$ for $|\mathcal{J}|$ = the number of birds. Musicalisability is inherited from the voice components: a harmony is (strictly) musical iff each voice is.

Verdict. Amusical, on the same grounds as a single bird. The framework *cannot* confer musicality on an ensemble whose components lack it.

6 Summary table

Frame	**Verdict**	**Decided by**
Two poets	strictly, not canonically	freedom of ν
Conversation	transcribable, not musical	morphisms in $\mathcal{V}_{\mathrm{conv}}$
Gregorian chant	strictly + canonically	inclusion of ambitus
Classical Chinese verse	quotient-canonical	tone projection
Sprechstimme	strictly + canonically	terminal second factor
Birdsong	transcribable, not musical	cardinality
Ritual percussion	strictly, not canonically	freedom of ν
Calling flock	amusical	inherited

7 Conclusion: What the Framework Refuses to Decide

The framework converts *"is X music?"* into a precise question about which 1-morphisms exist in $\mathfrak{A}(X, X_{\mathrm{mus}})$, and it exposes the variables on which a verdict depends:

(i) how richly one models $\mathcal{V}_X$ (discrete shadow vs. full morphism structure);

(ii) whether one demands the musicalisation to be canonical or merely possible;

(iii) whether one accepts X_{mus} as fixed or allows alternative reference frames (microtonal $\mathbf{Pitch}_{\mathrm{micro}}$, just-intonation $\mathbf{Pitch}_{\mathrm{ji}}$, electroacoustic spectrum $\mathbf{Pitch}_{\mathrm{spec}}$, etc.).

What the framework *does not* decide is which of these choices is the right one: the choice is a substantive musicological or philosophical commitment, and different communities make it differently. The framework's contribution is to make the disagreement *locatable*: two interlocutors who reach opposite verdicts on "poetry as music" will, on inspection, disagree about whether to take $\mathcal{V}^{\flat}_{\mathrm{phon}}$ or $\mathcal{V}_{\mathrm{conv}}$, and that disagreement has a name.

Remark (Why this is a useful refactor). A philosophical question of the form *"is X really Y?"* is typically irresolvable as posed because Y is given by ostension rather than by definition. The categorical reframing replaces Y by a structured object (X_{mus}) and the predicate "really Y" by a class of structure-preserving maps into it. Whether or not one accepts the chromatic frame as the right reference for "music", one at least now knows what one is rejecting.