Introduction to Gestural Similarity in Music. An Application of Category Theory to the Orchestra

Maria Mannone**
School of Music,
University of Minnesota, Minneapolis, USA
THIS IS NOT the latest version.
The latest version has been published
by the Journal of Mathematics and Music.

September 20, 2018

Abstract

Mathematics, and more generally computational sciences, intervene in several aspects of music. Mathematics describes the acoustics of the sounds giving formal tools to physics, and the matter of music itself in terms of compositional structures and strategies. Mathematics can also be applied to the entire *making* of music, from the score to the performance, connecting compositional structures to acoustical reality of sounds. Moreover, the precise concept of *gesture* has a decisive role in understanding musical performance. In this paper, we apply some concepts of category theory to compare gestures of orchestral musicians, and to investigate the relationship between orchestra and conductor, as well as between listeners and conductor/orchestra. To this aim, we will introduce the concept of *gestural similarity*. The mathematical tools used can be applied to gesture classification, and to interdisciplinary comparisons between music and visual arts.

Introduction

The topic of *musical gestures* in performance and composition is the object of an increasing interest from scholars [2, 17, 37, 53]. Interest in gestural comparison

---

** Email: manno012@umn.edu

is a natural consequence. New research about the mathematical description of gestures starts from piano, moves to orchestral conducting, and reaches singing as 'inner movement' [37, Chapter 37]. The topic of *musical similarity* is another important object of research [11, 21, 31].

In music, gestures are the intermediary between musical thinking and acoustics. Musical scores contain playing instructions and performers interact with their instruments via gestures to produce a specific sound result. Gestures are also involved in musical perception [48].

In this article, we focus on the comparison, in terms of *gestural similarity*, among gestures of the orchestral musicians, conductor's gestures, and specialization of the conductor's instructions into the movements of each musician, using the framework of category theory [28]. We will also briefly refer to visual arts, using the concept of gesture to connect them with music.

The interest of scholars in categories is motivated by their recent applications in physics and other sciences [51], and, more generally, in philosophy and diagrammatic thinking [1]. Mathematical music theory that uses category theory has already been developed [40], as well as differential calculus for musical analysis and musical gestures [33, 38, 40]. In conclusion, we refer to fuzzy logic for the definition of the degree of gestural similarity. The fuzzy logic had already been applied to music while investigating the relations between emotions and music [45].

The state of the art of the mathematical theory of musical gestures

What is a musical *gesture*? We can intuitively think of the movements of a dancer (gestures) while touching the floor at discrete points (notes), see Figure 1.

We can distinguish two different types of gestures: *symbolic* and *physical* gestures. The first are derived from the information contained in the score, and the second are the performer's real movements, see the top of Figure 1. Let us refer to the case of the pianist as a starting point. The symbolic gesture is derived from information such as the MIDI-like command "play this key, at this time, and with this loudness." The symbolic gesture contains straight lines and, in principle, it can require an infinite speed for the transition from pressed to released key, and from one note to another. The physical gesture is given by the real, physically-possible movements made by performers in non-zero time and finite speed, represented by the hands' smooth paths (their smooth curves) in space and time.

Symbolic gestures can be transformed into physical ones via an ideal connecting surface, the *world-sheet*. This formal tool, coming from string theory [57] in theoretical physics—in fact, we are comparing the continuous curves of gestures to strings—is interpreted here as a *hypergesture*.[1] See Section 1.4 for

---

[1]In string theory, point-like particles are substituted by vibrating strings. In [33] the paradigm of string theory is used to describe musical performance in terms of gestures, rather
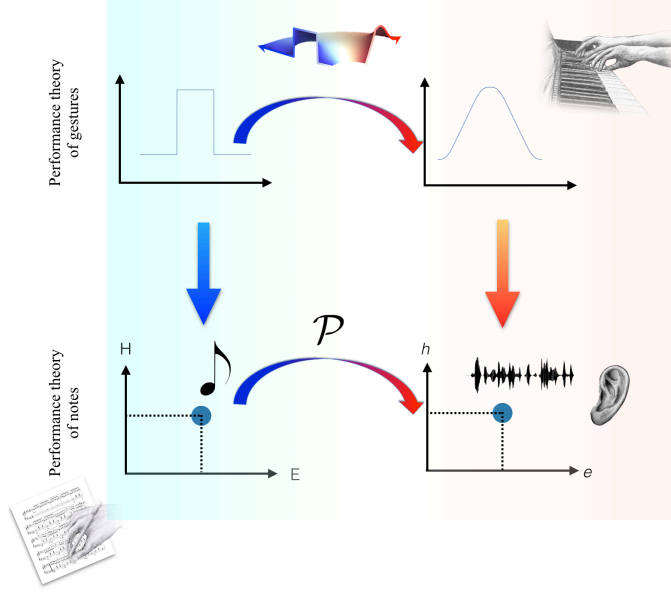
Figure 1: The general scheme of mathematical performance theory, from gestures (top) to notes (bottom). From the original German notation adopted in [40], H indicates symbolic pitch, h physical pitch, E symbolic onset, e physical onset. The other parameters are here omitted for graphical reasons. The graphs in the upper part of the image represent the gesture of the fingertip on piano keyboard, in the simplified case onset (in abscissa) and vertical position (in ordinate).

more details. However, because a gesture is a trajectory in a configuration space, we can formally study symbolic and physical gesture curves in the framework of Banach spaces and functional analysis. This would constitute a valid, and more precise, alternative approach to the string-theoretical one.

In this article, we refer more specifically to similarity between physical gestures. The concept of *similarity*, and even the term itself, had already been used in several musical contexts [21], for example while referring to notes and motives: we can think of "inner" comparisons inside the same musical composition [11], of motives' classification [10] and compositions' comparisons, including an adaptation of the formalism of physics to study memory [31, 32].

Similarity can also be introduced for gestures, while dealing with *classification*.[2] How can we *compare* a gesture with other gestures of the same kind

---

than notes and sounds as isolated events. The 'particles' correspond to the notes. To understand music as unfolding in time, rather than as isolated "points in time," we can study gestures to understand music. However, the "strings" here are not vibrating: thus, such a reference to string theory is just a general metaphor.

[2]For a precise but non-mathematical approach to gesture classification in the framework of mixed music (electroacoustic and with traditional instruments), see [3]; the term *similarity* is used to qualitatively describing imitational gestures, see [19].

(piano with piano), or other kinds (piano with violin)? We may also wonder how to connect musical gestures with other forms of artistic expression (for example, drawing), or with external references, as in expressive gesture performance studies [45].

In this article we compare gestures having the same (or a similar) generator. We will make implicit use of homotopic transformation (see Paragraph 6.1 in the Appendix), as also proposed in contexts of gesture following [2, 8]. Let us consider the following example. When the orchestral conductor signals a *forte* for *tutti*, every musician makes a gesture or a combination of the more appropriate gestures depending on the technique of his or her musical instrument, to obtain acoustical spectra inside the desired range of loudness and timbre. Their gestures are thus *similar* (in loudness). We can intuitively find examples of such a concept also from composition. For example, while thinking of a delicate gesture, even outside music, the composer can write down a combination of notes, dynamics, tempi, to suggest to the performer a *delicate* final result. The performer then recreates, by using his or her knowledge of gestural technique, a sound as close as possible to the one *thought* by the composer[3] and filtered by the sensibility of the performer.

An inverse mechanism affects listening and music perception. A listener judges a musical performance as 'expressive' if he or she *feels*, also beyond music, a gestural reference to be translated into mental images, a mental feeling. We can here refer to the concept of emotions and words viewed as hidden gestures [54], and refer to semiotics.[4] There can be cases where there is not any 'human' gestural generator, as for electronic music, but the feeling of a clear gesture is recreated in the mind of the listener, and other cases when this process is not possible. Some studies compare 'human' and electronic gestures, and describe analysis and creative developments of electronic music with respect to traditional musical instruments [3]. Even the *artistic inspiration* may be seen in light of gestural similarity, extending the language of artistic aesthetics to include the 'ineffable' inside the 'calculable' and understandable.

This article is structured as follows. After a list of preliminary mathematical concepts (Section 1), a first approach to gestural similarity is presented (Section 2), with a discussion of the role of conductor (Section 2.1), composer to conductor (Section 3), and some short references to music and visual arts (Section 4). Some final remarks about the possible role of fuzzy logic in classification problems, as elements of future research, and an appendix with some comments on homotopic transformation and more mathematical details on musical gesture theory, conclude the article.

---

[3] Of course, in the limit that such a thinking can be inferred from the score.

[4] We will not go deeper into detail with semiotics. The passage from a simple, instrumental movement to an *expressive* movement may raise issues about semiotics. For example, a "caressing" piano touch is not only finalized to get a sound, but a sound with a specific, soft timbre, that also carries *a meaning*, the meaning of a caressing gesture. Visualization of such a gesture is also relevant for the perception of the intended meaning. An accurate study of this field should require a separate and detailed description, as well as some perception experiments, to substantiate the connection between our mathematical approach with the musicological studies in the field [18, 19].
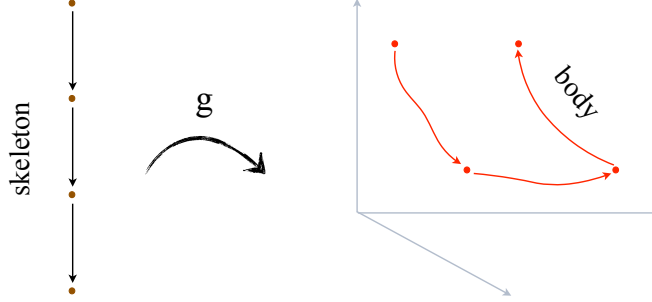
Figure 2: An example of a gesture, a mapping from an oriented graph to a system of continuous curves in a topological space, from [36]. We can compare this approach to a topic of neuroscience, about how to embed time-discrete symbolic processes into continuous time of neural processes [6].

1 Gestures and hypergestures

We need some preliminary mathematical tools before starting our analysis and investigation of gestural similarity.

1.1 Mathematical definition of gesture

Gestures have been mathematically defined as mappings from directed graphs to systems of continuous curves in topological spaces [36]. More precisely, we start from an abstract system of points and connecting arrows which is a directed graph Δ, as shown on the left of Figure 2. We map this directed graph to a system of continuous curves[5] in a space[6] X, with the same configuration as the directed graph Δ, as shown on the right of Figure 2. The directed graph Δ is called the gesture's *skeleton*, whereas the system of continuous curves constitutes the *body* of the gesture in $\vec{X}$. With $\vec{X}$ we indicate a set of mappings in X (when we add a topology, $\vec{X}$ becomes the space of curves in X). A gesture is then a mapping $g : \Delta \to \vec{X}$. We denote the space of all gestures from Δ to $\vec{X}$ as $\Delta@\vec{X}$, notation used in [36] for $Hom(\Delta, \vec{X})$. When we equip $\Delta@\vec{X}$ with a topology, we denote it as $\Delta\vec{@}X$. This is the space of gestures. A curve in the space of gestures (in $\Delta\vec{@}X$) is a gesture of gestures, and is called *hypergesture*.[7] Parametrized hypergestures are described in Section 6.3.1.

---

[5]A curve c in X is a continuous function $c : I \to X$, where $I = [0, 1]$ is the real unit interval.

[6]The space $\vec{X}$ is a topological space, e.g. spacetime.

[7]This is the name introduced in [36] and used in the related literature. However, as suggested by a reader of these works, we could perhaps use the term *metagesture*.

1.2 Gesture morphisms

A *morphism of two gestures* transforms one gesture into another. Let us refer to diagram @1, containing two gestures, $g : \Delta \to \vec{X}$ and $h : \Sigma \to \vec{Y}$. We define two functions, t and m. The function t is a transformation[8] from the skeleton Δ to the skeleton Σ, and the continuous function[9] m, from the topological space X to Y. We can also define $\vec{m} : \vec{X} \to \vec{Y}$ for the curves having points in X and Y respectively. These functions, jointly with the requirement of commutativity of diagram @1, define a morphism between gestures: in fact, a morphism of gestures is such a pair $f = (t, m)$ of morphisms of digraphs and topological spaces, respectively, such that $\vec{m} \circ g = h \circ t$ [38].

$$
\begin{array}{ccc}
\Delta & \xrightarrow{\ g\ } & \vec{X} \\
\downarrow{\scriptstyle t} & & \downarrow{\scriptstyle \vec{m}} \\
\Sigma & \xrightarrow{\ h\ } & \vec{Y}
\end{array}
\tag{1}
$$

Throughout the paper, we will frequently refer to morphisms of gestures while describing the deformation of gestures with particular characteristics, such as a gesture giving a specific musical dynamic (e.g. *piano*) into a gesture giving a different dynamic (e.g. *forte*). In general, we can define a category of gestures, with the defined morphisms as morphisms between gestures.[10] Such a formalism can be seen in terms of 2-categories [26]. For example, two points in the space are 0-cells; a gesture between them is a 1-cell, and a hypergesture between two gestures is a 2-cell. If we treat point as gestures, 1-cells are hypergestures, and 2-cells are hyper-hypergestures.

1.3 Homotopies of gestures

In the *gesture follower* system, the identification of gestures is based on recognition of classes of gestures, that we can consider as homotopy classes of gestures [2, 9, 12, 53]. The authors do not use a category theoretic framework.

A homotopy of two gestures connects them via an entire family of intermediary gestures—we have a progressive 'deformation' of a gesture into the other. To understand homotopies of gestures, we can intuitively think of examples such as anamorphosis frequently used in Escher's drawings. We consider a parameter λ, with $\lambda = 0$ for the first gesture, and $\lambda = 1$ for the second one. The progressive transformation of the first gesture into the second one is given by all the gestures labeled with values of λ ranging between 0 and 1. This means that for every $\lambda \in [0, 1]$, we have a gesture. A homotopic connection between

---

[8]For example, a transformation t^* modifies the skeleton $\Delta = . \to .$, that is an arrow between two points, into another skeleton $\Gamma = . \to . \to .$, with one more point and arrow.

[9]More precisely, m is a continuous functor of topological categories.

[10]The category of gestures can be described as a *comma category* $(id, \vec{\cdot})$, where id is the identity on directed graphs, and $\vec{\cdot}$, defined from spaces to directed graphs, is the functor that associates to a space X the digraph $\vec{X}$ with the paths in X.

gestures does not automatically imply also a functional relation between them: it is another result, not true *a priori*. The use of a parameter λ for gestural comparisons recalls fuzzy logic[11] [22], as discussed in the concluding Section 5.

Some more considerations about homotopy, with references to the physics of sound, are given in Paragraph 6.1 of the Appendix, with a "physical" statement about similar gestures (see the Euristic Conjecture 6.1).

1.4 Hypergestures

The concept of *hypergesture* generalizes that of homotopy. As explained before, in a (topological) gesture space, where each point is a gesture, a curve connecting points is a *gesture of gestures*, a hypergesture. The difference between hypergestures and morphisms is also stressed in [36]. In the case of morphisms, we verify the relations between skeleta and bodies that are compatible; in the second case, we build curves in the space of gestures. A particular case of hypergesture is the world-sheet, the ideal surface connecting symbolic and physical gestures, see [33] for a first example. However, in this article we will not use the formalism of world-sheets for our analysis.

The main structure gesture / hypergesture suggests the use of the 2-categories formalism. In principle, because we can build hypergestures of higher order, we may extend the formalism to $\infty-$categories. The topic of infinite categories and quasicategories was investigated extensively by [20] and [27]. Music can be analyzed through nested structures. For example, a gesture (0-cell) connects two points (an arrow between points); a *crescendo* is a hypergesture (1-cell) that connects less loud gestures with louder gestures (an arrow between two arrows); and an *accelerando* is a hyper-hypergesture (2-cell) that transforms a slower crescendo into a faster crescendo (an arrow between arrows connecting arrows). We can build N-cells (and ideally ∞-cells) by adding each time transformations and transformations of transformations.

2 Mathematics for similar gestures of orchestral musicians

Let us start from a simple pianistic gesture (a primitive up-down), and an equivalent movement for a percussive gesture, let it be on a vibraphone or a xylophone. In both cases there is a 'percussive' gesture/generation of the sound.

Let g_ϕ be the physical gesture of the pianist P, and h_ϕ the gesture of the percussionist P_c. Let us for simplicity choose the same skeleton for both. The two gestures are:
$$g_\phi : \Delta \to \vec{X}, \ \ h_\phi : \Delta \to \vec{Y}. \tag{2}$$

---

[11]We will not delve into details about fuzzy logic. Whereas Definition 2.1 assesses a criterion of gestural similarity, we can more precisely talk about *degrees* of gestural similarities, with infinite intermediate values between perfect similarity of two identical gestures and two completely unrelated ones.

The topological spaces for the two instruments are different, and so is the space of their curves. We start with the same skeleton Δ for both gestures, an arrow connecting two points. We use the formalism of category theory to characterize the action of the potential.[12] For a reference about world-sheet, force field, and potential see Paragraph 6.4 in the Appendix. Let g_ϕ be the physical gesture of the pianist, represented here as the result of the action of a function $f_P : g_\sigma \rightarrow g_\phi$, where P stands for the 'piano' instrument. The function f_P has, as its variable, symbolic gestures g_σ with unspecified dynamics (that we will indicate as a 0-potential). The physical pianistic gesture corresponding to a *forte* dynamic is here given as a result of the same function, with a non-vanishing potential F_P as second argument:

$$g_\phi = f_P(g_\sigma, 0), \quad g_\phi^F = f_P(g_\sigma, F_P). \tag{3}$$

Here, g_ϕ stands for g_ϕ^0. Equivalently, the percussionist's gesture is:

$$h_\phi = f_{Pc}(h_\sigma, 0), \quad h_\phi^F = f_{Pc}(h_\sigma, F_{Pc}). \tag{4}$$

We can also represent the same situation with the (commutative) diagram @5, where $g_\phi^F = F \circ g_\phi$. The operator F transforms 'generic' curves in $\vec{X}$ into *forte* curves in $\vec{X}$, that is, $(g_\phi \in)\vec{X} \rightarrow (g_\phi^F \in)\vec{X}$.

$$\begin{array}{ccc} \Delta & \xrightarrow{\ g_\phi^F\ } & \vec{X} \\ & \searrow{\scriptstyle g_\varnothing} & \uparrow{\scriptstyle F} \\ & & \vec{X} \end{array} \tag{5}$$

The digraph morphism $F : \vec{X} \rightarrow \vec{X}$ is an endomorphism of arrows of $\vec{X}$, because the gestural curves with unspecified dynamic (that we indicate here as with 0-potential) are in the same space of the curves with *forte* dynamic (F-potential for us). The digraph morphism $F : \vec{X} \rightarrow \vec{X}$ indicates a curve in the space of potentials, as already done in the case of voice [37]. We change the values, but we always are in the same space. There is an entire family of gestures generated by the choice of potential *forte*. There is a homotopy of connected (linked) gestures, with the second obtained from the first via a functional composition (and this, they are not independent).

We have been composing plain morphisms of graphs. To simplify the former and the following diagrams, as well as to give a description that is coherent with the 1-cell and the 2-cell definitions, we can use 2-cell notation. In a 2-categorical

---

[12]If restricted to the physical curve, the potential is related to the force as known in common physical situations. And, of course, the force strongly influences the touch and the final spectral result at the level of acoustics. A more general 'artistic force' determines the shape of the entire surface of the world-sheet.

context, diagram @5 can be modified as diagram @6.

$$\Delta \quad \overset{\overset{g_\phi}{\Downarrow F}}{\underset{g_\phi^F}{}} \quad \vec{X} \tag{6}$$

We are transforming the curves within the space $\vec{X}$, modifying the loudness coordinate via the velocity/acceleration. We may also describe such a musical situation in the space of phases with position and speed, as often done in physics. In such a space, a *crescendo* would be easily described as a hypergesture connecting gestures with different coordinates of loudness.

To summarize, we start from a physical gesture $g_\phi = f(g_\sigma, \circ)$ from a certain symbolic gesture g_σ and a potential $\circ$, giving the dynamic. We could think of $\circ$ as a generic potential V, that is then specialized when a particular dynamic (*forte*, *mezzoforte*) is indicated. A transformation $0 \to F$ with the *forte* induces a deformation of the gestural curve into another realizing the *forte* dynamic as result of the interaction between performer and musical instrument: we have $g_\phi \to g_\phi^F$, where $g_\phi^F = f(g_\sigma, F)$. The transformation of the potential $0 \to F$ can be described by a continuous function[13] $V(t)$, that is equal to 0 at $t = 0$, and to F at $t = 1$, the coordinate t being a generic parameter, or a time coordinate. We can describe a crescendo via diagram @7. We start from a gesture g with unspecified dynamic, we then identify the required transformation (i.e. gestural deformation) to get the *piano* and *forte* respectively, and *then* we define a *temporal* transformation PF that brings the *piano* gesture g_ϕ^P into the *forte* gesture g_ϕ^F, that *is* the crescendo. In terms of potentials, we have the transformation $F = PF \circ P$.

$$\begin{array}{ccc} g_\phi^P & \xrightarrow{\quad PF \quad} & g_\phi^F \\ & & \\ & \nwarrow_{P} \qquad \nearrow_{\mathit{f}} & \\ & g_\phi & \end{array} \tag{7}$$

In terms of 2-category formalism, diagram @7 can be re-drawn as diagram @8, using the vertical composition property.

$$\Delta \quad \begin{array}{c} \overset{g_\phi}{\Downarrow P} \\ \overset{g_\phi^P}{\Downarrow PF} \\ g_\phi^F \end{array} \quad \vec{X} \qquad = \qquad \Delta \quad \begin{array}{c} g_\phi \\ \Big\Downarrow F \\ g_\phi^F \end{array} \quad \vec{X} \tag{8}$$

---

[13]More generally, we can define a gesture via a function f containing such a $V(t)$ function.

$$f(g_\sigma, V(t)) =: g_\phi(t) = \left\{ \begin{array}{ll} g_\phi^0, & for\, t = 0 \\ g_\phi^F, & for\, t = 1 \end{array} \right\}$$

Let us now explicitly use functionality. We reach the *forte* curves in $\vec{X}$ via $f(\circ, F)$ for the *forte F*, and $f(\circ, 0)$ without any specified dynamic, where the symbol $\circ$ indicates the place reserved for the symbolic gesture. Here, we choose an elementary skeleton[14] $\Delta = \uparrow$, and a simple symbolic gesture that is the image of such an abstract curve, indicated here also with $\uparrow$. We have $f(\uparrow, F) = g_\phi^F$ and $f(\uparrow, 0) = g_\phi$, $F \circ f(\uparrow, 0) = f(\uparrow, F)$, and the diagram @9 commutes.

$$
\begin{array}{ccc}
\uparrow & \xrightarrow{\ f(\uparrow,F)\ } & \vec{X} \\
& \searrow{\scriptstyle f(\uparrow,0)} & \big\uparrow F \\
& & \vec{X}
\end{array}
\tag{9}
$$

Two main questions arise:

1. Is this homotopy?

2. Is the process $g_\phi \to g_\phi^F$ functional (well-defined)?

The answer to both questions is affirmative for the following reasons.

1. Homotopy. There is homotopy between the two gestures, because there is an entire family of intermediate gestures continuously connecting the first and the second one. The potentials themselves are connected by a family of potentials, going from 0 and reaching F.

2. Functionality. The process is also functional, for the reason expressed above: the gesture $f(\uparrow, F)$ is functionally obtained from $f(\uparrow, 0)$ via the composition with F. In fact, $g_\phi^F \equiv F \circ g_\phi$.

Let us now extend the previous discussion to include percussionists' gestures. We again use the label P for the piano, and the label Pc for the percussion. In fact, this time we distinguish between F_P and F_{Pc}, acting on different spaces. Diagram @10 shows the pianist's gesture (right) and the percussionist's gesture (left), having the same skeleton Δ.

$$
\begin{array}{ccccc}
\vec{Y} & \xleftarrow{\ h_\phi^F\ } & \Delta & \xrightarrow{\ g_\phi^F\ } & \vec{X} \\
\big\uparrow{\scriptstyle F_{Pc}} & \nwarrow{\scriptstyle h_\phi} & & \searrow{\scriptstyle g_\phi} & \big\downarrow{\scriptstyle F_P} \\
\vec{Y} & & & & \vec{X}
\end{array}
\tag{10}
$$

$$
\begin{array}{ccc}
& h_\phi \quad\quad g_\phi & \\
\vec{Y} \;\Big\Downarrow{\scriptstyle F_{Pc}}\Delta\;\; & \Big\Downarrow{\scriptstyle F_P} & \vec{X} \\
& g_\phi^F & \\
h_\phi^F & &
\end{array}
\tag{11}
$$

---

[14]We denote by $\uparrow$ the digraph having two vertices that are connected by one arrow.

Diagram @11 shows the same content of diagram @10 in a 2-categorical context.[15]

The two triangles in diagram @10 show functional transformations. How can we connect $\vec{Y}$ to $\vec{X}$ and compare the two gestures g and h? We think of g and h as being 'gesturally similar,' because we can transform the one into the other via an homotopic transformation, *and* they lead to similar timbres both modified into *forte*—see the Euristic Conjecture 6.1. We will now consider two different skeleta, and the change of skeleta $\Delta \to \Gamma$, as shown in diagram @12, provided we are given a continuous function m such that $m : X \to Y$.

$$
\begin{array}{c}
\text{(12)}
\end{array}
$$

The two triangles are commutative because $g_\phi^F = F_P \circ g_\phi$ and $h_\phi^F = F_{Pc} \circ h_\phi$, as well the square of diagram @13.

$$
\begin{array}{ccc}
\Delta & \xrightarrow{g_\phi^F} & \vec{X} \\
\downarrow{t} & & \downarrow{\vec{m}} \\
\Gamma & \xrightarrow{h_\phi^F} & \vec{Y}
\end{array}
\tag{13}
$$

This means that there is a morphism of gestures, as defined in Section 1.2. Diagram @12, with 2-category formalism, looks like diagram @14.

---

[15]We can observe that, in the framework of 2-category formalism, vertical and horizontal composition properties are intuitively verified. Vertical composition means, in our orchestral context, transformation of loudness: a *piano* gesture can be deformed into a *forte* gesture, that can be deformed into a *fortissimo* gesture. Horizontal composition, given the same skeleton, means here transition from the space of gestures for a musical instrument, for example piano, to the space of gestures of another instrument, as percussion, and then to another instrument, such as violin.

$$
\begin{array}{c}
\xymatrix{
& g_\phi \\
\Delta \ar[r]^{F_P} \ar[d]_{t} \ar@/^/[ur]^{g_\phi} & \vec{X} \ar[d]^{\vec{m}} \\
& g_\phi^F \\
& h_\phi^F \\
\Gamma \ar@/_/[dr]_{h_\phi} & F_{P_c}\ \vec{Y} \ar[u]_{F_{P_c}} \\
& h_\phi
}
\end{array}
\qquad (14)
$$

Under which conditions is diagram @12 commutative also with $\vec{X}$ and $\vec{Y}$? The similarity definition also includes the commutativity, in diagram @12, of the square shown in diagram @15.

$$
\begin{array}{ccc}
\vec{X} & \xleftarrow{\ F_P\ } & \vec{X} \\
\downarrow{\scriptstyle \vec{m}} & & \downarrow{\scriptstyle \vec{m}} \\
\vec{Y} & \xleftarrow{\ F_{Pc}\ } & \vec{Y}
\end{array}
\qquad (15)
$$

Similar gestural generators are supposed to provoke similar changes in the timbres, for example the creation of two *forte* gestures from gestures with an unspecified loudness. But, first of all, the changes have to happen in the appropriate spaces. Diagram @12 is not commutative when, for example, the generative process is similar but the resulting spectrum is not. If we strongly hit the flute with a piano hammer, it will not sound louder, and a strong use of the bow on a flute will not make a stronger flute sound. They are examples of *forte* gestures in non-appropriate spaces for the flute. This means that we cannot apply the same "generator" F_P of pianistic *forte*, to obtain a *forte* gesture on flute. We need another generator, a F_{Fl}, to transform generic curves of flute playing into *forte* curves (gestures having as result a *forte* sound). F_{Fl} does not act on $\vec{X}$ but on some $\vec{Z}$, where $\vec{Z}$ is the space of curves for the flute playing. The spaces $\vec{X}$ and $\vec{Z}$ can be connected by a suitable m function to transform the space of gestures of an instrument into the space of gestures of another instrument.

We can finally give a first definition of gestural similarity, as expressed in Definition 2.1.

Definition 2.1. Two gestures $g_\phi^F : \Delta \to \vec{X}$ and $h_\phi^F : \Gamma \to \vec{Y}$ are *similar*, and we write $g_\phi^F \sim h_\phi^F$, if it is possible to find:

1. A morphism from the first to the second, it means, two functions $t : \Delta \to \Gamma$ and $m : X \to Y$ such that diagram @13 is commutative;

2. Two homotopic transformations[16] $F_P : \vec{X} \to \vec{X}$ and $F_{Pc} : \vec{Y} \to \vec{Y}$, with the *common generators F_P and F_{Pc} acting in the same way*[17] in their spaces respectively, such that $F_P \circ g_\phi = g_\phi^F$ and $F_{Pc} \circ h_\phi = h_\phi^F$.

Let us see more details in Example 2.2.

Example 2.2. Let us suppose *ad absurdum* that two gestures g_1, g_2 for the same instrument (topological space $\vec{X}$), with unspecified dynamic but sharing the same skeleton, are similar. If the skeleta are the same, we have an identity map between them. We also have an identity for the topological space $X \to X$, being the two gestures for the same instrument. So we have a gesture morphism with identities in this case, and point (1) is satisfied. However, because the dynamic is unspecified, we also have unspecified generators. This implies that point (2) is not satisfied, and thus the hypothesis of similarity is not verified. In fact, two gestures cannot be similar if they are obtained via two generators acting in different ways in their respective spaces (the same in our case). We can consider a *forte* gesture and a *piano* gesture, both for the same musical instrument, for example percussion Pc, with the same skeleton. We have an identity for the topological space $Y \to Y$. The two generators $F_{Pc} : \vec{Y} \to \vec{Y}$ and $P_{Pc} : \vec{Y} \to \vec{Y}$ do not act in the same way: the first selects *forte* curves, while the second *piano* curves. We have $F_{Pc} \circ h_\phi = h_\phi^F$ and $P_{Pc} \circ h_\phi = h_\phi^P$, and point (2) is not satisfied.

See also the Heuristic Conjecture 6.1 in Paragraph 6.1 of the Appendix for a physical condition about acoustical spectra. The musical meaning of Definition 2.1 is the following. A *forte* gesture for the pianist and a *forte* gesture of the percussionist (in our example) are *similar* because:

1. It is possible to transfer (connect) the *forte* on the piano to the *forte* on a percussion;

2. The generator that transforms generic (neutral) gestures on piano into *forte* gestures on piano is the same generator that transforms generic gestures on a percussion into *forte* gestures on a percussion, acting in different spaces in the same way.

Both *forte* gestures[18] can be projected into the *forte* conducting gestures, using the categorical notion of *colimit* (Section 2.1). To decide if the two generators F_P and F_{Pc} act in the same way in their spaces, and, consequently, if two resulting gestures are similar, we can also use concepts from fuzzy logic. A gesture can belong to a category but also partially to another one. For example, a range of variability allows the distinction of several *forte* gestures, that are all

---

[16]Two *homotopic* curves can be continuously transformed the one into the other. A homotopic transformation continuously transforms curves in curves.

[17]Similar deformations in their respective spaces lead to similar effects in their resulting sound spectra. See Paragraph 6.1 in the Appendix for two graphic representations, Figures 3 and 4.

[18]We discussed loudness, but gestural similarity can also involve articulation, and in some cases also rhythmic-melodic profiles. Even specific harmonic sequences can suggest particular gestural solutions. We can think of a deceptive cadence, highlighted by the performer with a *fermata* or a *forte*. In fact, elements from musical analysis can act as weights for gestures.

equally well-working for the same musical passage. This opens another potential research in the application of fuzzy logic to category theory [55] for music.
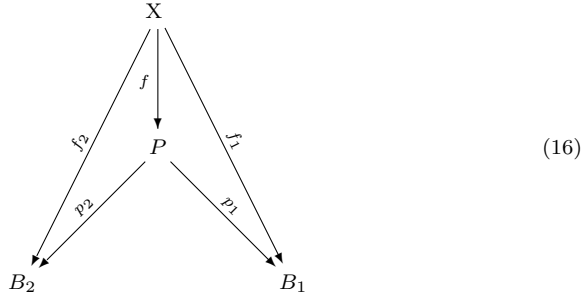
2.1 Conducting gestures

The conductor's gesture is specialized into gestures of all orchestral musicians, and it may be described as their *colimit*. For the listener, the conducting gesture can constitute a *limit*.[19] We can think of a person trying to get information about the orchestral sound by watching conducting gestures on a video with the audio turned off. We are considering limit and colimit because the listener (i. e., each of the listeners in the audience) refers to the orchestra, and all the orchestral instruments refer to the conductor. Moreover, the conductor, while studying the score, prepares his or her gesture depending on the orchestration, meaning what orchestral musicians are supposed to play. The listener creates his or her mental idea of the music depending on what orchestral performers are playing.
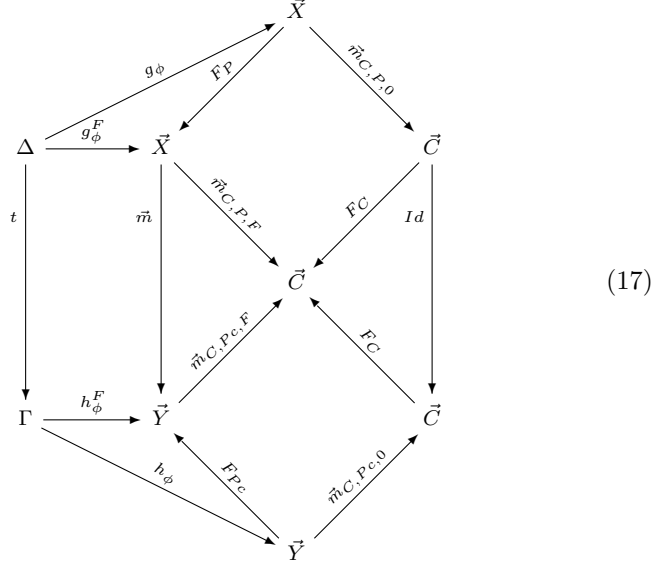
We indicate orchestral gestures as $\mathcal{D}$, conductor's gestures is their $\operatorname{colim}(\mathcal{D})$, while the audience is the $\lim(\mathcal{D})$.

We can draw diagram @17, including the contribution of the conductor (with $\vec{C}$), where F_C ($\vec{m}_{C,F}$), associated to the conductor's gesture, transforms conducting generic curves (with unspecified dynamic) into *forte* curves. In this diagram, to simplify the graphical representation, we omit the skeleton of the

---

[19]Intuitively, *limits* and *colimits* are generalizations of *products* and *coproducts*, respectively. The product is a special case of the limit, with discrete categories. The coproduct (also called *sum*) is the dual of the product, obtained by reversing the arrows. Given an object P and two maps $p_1 : P \to B_1$, $p_2 : P \to B_2$, P is a product of B_1, B_2 if for each object X and for each pair of arrows f_1, f_2 we have only and only one arrow $f : X \to P$ such that $f_1 = p_1 f$, $f_2 = p_2 f$ [25], see diagram @16.
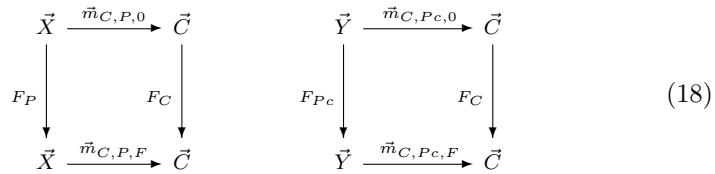

$$(16)$$

conducting gesture.

$$
\begin{array}{c}
\vec{X} \\
\nearrow \quad \searrow \\
g_\phi \quad \hat{F}p \quad \vec{m}_{C,P,0} \\
\Delta \xrightarrow{g_\phi^F} \vec{X} \qquad \vec{C} \\
t \quad \vec{m} \quad \vec{m}_{C,P,F} \quad \hat{F}_C \quad Id \\
\vec{C} \\
\vec{m}_{C,Pc,F} \quad F_C \\
\Gamma \xrightarrow{h_\phi^F} \vec{Y} \qquad \vec{C} \\
h_\phi \quad F_{Pc} \quad \vec{m}_{C,Pc,0} \\
\vec{Y}
\end{array}
\tag{17}
$$

In diagram @17, we distinguish the functions bringing the conductor's gesture to the pianist and to the percussionist. Summarizing, we have:

- gesture without specified dynamic of the pianist reflected into the basic metric gesture (no indication of dynamic) of the conductor: $\vec{m}_{C,P,0} : \vec{X} \to \vec{C}$,

- gesture without specified dynamic of the percussionist reflected into the basic metric gesture (no indication of dynamic) of the conductor: $\vec{m}_{C,Pc,0} : \vec{Y} \to \vec{C}$,

- gesture of *forte* for the pianist reflected into the conductor's *forte*: $\vec{m}_{C,P,F} : \vec{X} \to \vec{C}$,

- gesture of *forte* for the percussionist reflected into the conductor's *forte*: $\vec{m}_{C,Pc,F} : \vec{Y} \to \vec{C}$,

- basic metric gesture transformed into the *forte* one of the conductor: $F_C : \vec{C} \to \vec{C}$,

- identity Id transforming conductor's basic metric gesture into itself.

Diagrams @18 show the two commutative squares conductor-pianist and conductor-percussionist from diagram @17, respectively.

$$
\begin{array}{ccc}
\vec{X} \xrightarrow{\vec{m}_{C,P,0}} \vec{C} & \qquad & \vec{Y} \xrightarrow{\vec{m}_{C,Pc,0}} \vec{C} \\
\downarrow F_P \qquad \downarrow F_C & & \downarrow F_{Pc} \qquad \downarrow F_C \\
\vec{X} \xrightarrow{\vec{m}_{C,P,F}} \vec{C} & & \vec{Y} \xrightarrow{\vec{m}_{C,Pc,F}} \vec{C}
\end{array}
\tag{18}
$$

15

These diagrams commute because it is equivalently possible that:

- The pianist plays *forte* because the conducting basic metric gesture has been modified by the operator *forte*;

- The conductor gives a basic metric gesture and the pianist decides to play *forte* for reasons of expressivity or written indication.

This second case can happen, for example, for scores where rhythm and other specific indications are so complicated that the conductor needs to give simple metric indications. In formulas, we have, for the pianist, $F_C \circ \vec{m}_{C,P,0} = \vec{m}_{C,P,F} \circ F_P$, and the same for the percussionist, substituting P_c to P.

Diagram @19 shows the gestural similarity between the *forte* gesture of the pianist and the percussionist (via $\vec{m}$), and their reflection into the conducting *forte* gesture, for the pianist (via $\vec{m}_{C,P,F}$) and the percussionist (via $\vec{m}_{C,Pc,F}$).

$$
\begin{array}{c}
\vec{X} \\
\downarrow \vec{m} \quad \searrow^{\vec{m}_{C,P,F}} \\
\vec{C} \\
\nearrow_{\vec{m}_{C,Pc,F}} \\
\vec{Y}
\end{array}
\tag{19}
$$

Using 2-categories, we can schematize diagram @17 as shown by diagram @20, with functors transforming $\vec{m}_{C,P,0}$ in $\vec{m}_{C,P,F}$ and $\vec{m}_{C,P_c,0}$ in $\vec{m}_{C,P_c,F}$, respectively.[20]

$$\tag{20}$$

---

[20]This is not shown in the diagram.

We can apply such a diagram to all musicians of the orchestra, getting the *colimit*: orchestra $= \mathcal{D} \longrightarrow$ conductor $= \mathrm{colimit}(\mathcal{D})$. We can define the conductor-orchestra-listener via diagram @21 in terms of limits and colimits.

$$conductor = \mathrm{colim}(\mathcal{D})$$

$$orchestra = \mathcal{D} \tag{21}$$

$$listener = \lim(\mathcal{D})$$

If we schematically represent the orchestral gestures as in equation @22, we can define the limit as the listener, and the colimit as the conductor, see diagram @23. The choice of listener as a limit, even if considered as a metaphor, satisfies the universal property because all the "listening and perception activities" can be reduced to the listener, in contraposition to the sound-production activities. The conductor plays the opposite role: all the "sound production" gestural activities can be related to the conducting gesture, that is a pure gesture without any direct sound production.

$$\mathcal{D} = \begin{array}{c} X_\lambda \xrightarrow{f_{\lambda\mu}} X_\mu \\ f_{\kappa\lambda} \nwarrow \quad \nearrow f_{\kappa\mu} \\ X_\kappa \end{array} \tag{22}$$

We can also specify the colimit diagram as shown by diagram @24, where $\mathcal{D}$ is the orchestra, and the colimit *is* the conductor's gesture $\Delta_{conductor} \rightarrow$

$\vec{X}_{conductor}$.

$$
\begin{array}{ccc}
 & \Gamma \xrightarrow{\ \ g\ \ } \vec{Z} & \\
 & \exists! \nearrow \qquad \qquad \searrow \not\vdash & \\
\Delta_{conductor} \xrightarrow[g_{conductor}]{} \vec{X}_{conductor} \longrightarrow \mathcal{D} &
\end{array}
\tag{24}
$$

Let us now include the detail of (some) orchestral gestures. In diagram @25, for simplicity we represent just one arrow from one orchestral gesture, g_λ, to the conductor; however, *each* orchestral gesture, it means each g_i (g_λ, g_κ, g_μ) in the diagram @25, has an arrow to the conductor's gesture $g_{conductor}$, and also to the g gesture.

$$\tag{25}$$

We will not deal here with a description of the universal properties of these limits and colimits, preferring some more comments on their musical meaning and implications. The gestures of the orchestral musicians can be (mathematically) injected into the conductor's ones.

Intuitively, we would say the opposite: the gestures of the conductor are the initial source of movement for the orchestra, and the listener is the final target who sees the gestures, and listens to their sound result. This means that the conductor's movements are specified into the gestures of orchestral musicians, and collected as a whole result by the listener. This would imply an inverse order of the arrows in diagrams @22 and @23.

However, we can also say that the listener (emotionally) projects his or her thought inside the sound and the gestures, so the arrows come from the listener

18

and reach the orchestral gestures. We may see the orchestral gestures as contained "inside" the conductor's ones. Let us think of the violinist's, or pianist's gestures. They are much more complicated than the conducting ones. The complicated gesture of each performer refers to, and injects into the simpler gesture, i.e. the conducting one. This means that it projects into its *simplification*. We can interpret piano or violin gestures as something that can be envisaged into the conducting gestures. Orchestral gestures are the development of the conducting gestures. The listener may imagine the orchestral gestures—and the shape of music—by observing the conducting gestures, even without listening to the sound. In fact, the conductor's gesture is the *terminal* element: there is one and only one function going from it to the $g : \Gamma \to \vec{Z}$ having the property that *all* orchestral gestures are injecting into it. There is a morphism of each instrument to the conductor, and there is one and only one morphism of gestures of the conductor in Z that makes the diagram commutative. We can talk metaphorically about categorical adjunction between these two perspectives. A mediation between the two poles would involve the definition of exchanges and mutual feedback between performers, conductor, and listeners; the conductor is also a listener and the listener can influence the performance through applause or boos. An intermediate figure between who *gives indication to make* and who *listens to* the sounds may be a conductor that changes the gesture according to the performers' outcome, or, better, a performer of an electronic instrument that influences the sound production through the gesture itself (as a theremin player), and adjusts the gesture according to the sound.

If we opt for the second description or for the intuitive one, both listener and conductor are still at the opposite side, because:

- there is only one element (initial or terminal, depending on the chosen description) that collects all the orchestral gestures (for example, a "simplified version" of orchestral gestures with collected analogous movements), and there is one and only one morphism to the conductor's *motor* gesture;

- there is only one element (again, terminal or initial, depending on the chosen description) that reaches all the orchestral gestures (for example, a more prominent sound result given by the most relevant musical structures, elements, phrases, timbres), and there is one and only one morphism to the listener's *perceptive* gesture.

Both listener and conductor are silent. The conductor is making active gestures, the listener is interpreting the music heard via making similarities with his or her background, past listening experiences, musical knowledge, and personal sensibility. The conductor belongs to the *motor* world, while the listener to the *perceptive* world. If, through a software, the listener can manipulate an electronic conductor, he or she is also a conductor, and if the conductor changes his or her gestures depending on the *listened sound*, he or she is also a listener. In each case, listener and conductor can only be compared with initial and terminal objects in the world of the orchestral performance.[21] In the case of a

---

[21]This happens if we already have a musical score. If, in such a description, we include the

piano improviser who is listening to his or her own music, listener's, (unique) performer's, and composer's roles all coincide.

Within the formalism of 2-category, we can generalize the description of listener $\rightarrow$ conductor $\rightarrow$ orchestra in terms of 2-colimits and 2-limits. In this case, horizontal composition properties (among orchestral instruments, among different conductors, among different listeners) and vertical composition properties (among gestures leading to different dynamics, articulations, tempo changes) are still valid. This can be the topic of further research in itself.

3 Gestures from the composer to the conductor

We can schematize the gestures of the orchestral performance as a category, and the gestures hidden in the score as another category. We can define a *Performance Functor* connecting symbolic to physical gestures,[22] and a *Compositional Functor*, for the inverse movement, from the improvisation to the music transcription/composition. In diagram @26, we consider different skeleta, using labels σ, ϕ to denote symbolic and physical gestures, respectively. Let us use the concept of functor as morphism between categories. The Performance Functor (P. F.) connects symbolic gestures to physical gestures, for piano and percussion.

$$
\begin{array}{ccc}
\Delta_\sigma & \longrightarrow & \vec{X}_\sigma \\
\downarrow & & \downarrow \\
\Gamma_\sigma & \longrightarrow & \vec{Y}_\sigma \\
& & \\
& \Delta_\phi & \longrightarrow & \vec{X}_\phi \\
& \downarrow & & \downarrow \\
& \Gamma_\phi & \longrightarrow & \vec{Y}_\phi
\end{array}
\tag{26}
$$

(with diagonal arrows labelled $P.F._{P}$ and $P.F._{Pc}$)

---

composer of the orchestral score, then the colimit role may be envisaged into the composing activity. We may say that, in this extended description, the initial and final points are both in the mind: the mind of the composer, and the mind of the listener.

[22]The connection of single curves from (symbolic) systems of continuous curves to single curves from other (physical) systems of continuous curves, their skeleta being the same, has been investigated via branched graphs and branched world-sheets [33, 38].

4   Gestural similarity between music and visual arts

The concept of gestural similarity can be used not only in music, but also between music and visual arts.[23] In art history, there are several examples where visual artists got inspiration from music and vice versa: we can think of Boris Mussorgsky, Luigi Russolo, and Morton Feldman, as few names.[24] In movies, particular scenes are often emphasized by musical gestures: we can think of the correspondence between the frightening bow's strikes and the knife scene in *Psycho*.

Several studies highlight the correspondence between music and movement [56], especially with the applications in the field of soundtracks. In general, we can say that "music is a gesture," it is the "result of forces, impulses, experiences... both from individual and general" [49].

Finally, we can discuss examples of sonification (mapping of non-sound data into sound-data), and we can argue that they are more effective when there are gestural similarities: a rising shape may be described via a rising pitch sequence [46], or a "staccato" musical sequence may be compared with a collection of points on canvas. It is sufficient to see a drawing, for example, as the result of a drawing gesture. Studies in the field of psychology [56], crossmodal correspondences [7, 43, 50], iconicity in linguistics [34], and audio-visual objects and the theory of indispensable attributes [23] can support these ideas. Finally, we can re-read examples of sonification [29] in light of gesture theory, and we can compose new music [30].

We can use these ideas to investigate the connections between vocal gestures in speaking and singing, and visual shapes, framing within category theory classic perceptual experiments [41, 42, 52].

Further developments of such a research will also include perceptual experiments to test the expected degree of similarity between music and visuals.

5   Conclusion

In this paper we contextualized in a categorical framework the analogies between gestures of different musicians, the gestural communication between conductor and performers. Such a formalism can also be used to compare gestures in music and in visual arts.

We proposed a first mathematical definition of gestural similarity, to allow comparisons between gestures belonging to different spaces (such as piano's and percussion's spaces) having similar characteristics, for example all realizing

---

[23]Category theory includes functors: so, we could provocatively talk about *functorial aesthetics*.

[24]In future developments of gestural similarity analysis, we may try to find analogies between a painting and the music inspired from it. This may be part of a more general approach to artistic movements, finding the connections between music, visual art, and poetry within a specific movement, in terms of basic "shared gestures."

a *forte* sonority. Further research can address the more general question of the gesture classification. We can use concepts from *fuzzy logic*, to quantify the *degree of similarity* between two different gestures. Moreover, fuzzy logic can intervene in a more fundamental way, to connect symbolic and physical gestures. In fact, analogously to what we do when we think of "the" circle (as a Platonic idea) without being able to draw it, except in terms of a fuzzy circle 'not perfectly circular,' we can write a score and not be able to play *exactly what it is written*. Each performance can be described in fuzzy terms: 0 for the score, 1 for the "perfect" and complete musical performance, and values within $[0, 1]$ to denote the performance in progress. Transition from symbolic to physical gestures can be reformulated in terms of progressive action of a fuzzy function.

About category theory, further research may explore in detail the universal properties of 2-limits and 2-colimits between conductor, orchestra and listener, as well as theoretical extensions and artistic implications in the field of ∞-categories.

About the relations between gestural similarity and sound, future research can also involve collaboration with scholars in the field of neurosciences, more precisely about interactions between symbolic gestures and words [4], spoken language and arm gestures [16], gesture, sound, imitation in phoneme production [13–15].

In conclusion, future studies can deepen and consolidate the connections between physics, psychophysics, visual art, linguistics, and neuroscience, via a common mathematical model.

6 Appendix: More details on the mathematical theory of musical gestures

6.1 Homotopy and gestures

We will clarify the concept of homotopy and gestures used in this article. Let us suppose we have two gestures, one for the piano, and the other for the violin, with the same skeleton, see Figure 3. A *forte* gesture for the pianist is similar to the *forte* for the violinist. The *forte* is transformed into *piano* via a *diminuendo*, represented by the action of the functor FP in the 2-category (more precisely, we have FP_P acting on piano gestures, and FP_V acting on violin gestures). There is similarity between these gestures, because the articulations and the dynamics are the same. However, if we choose a *staccato* gesture for the violin, there is similarity only for the dynamic, but not for the articulation. The diagram involving non-similarity is not commutative. All transformations involved are homotopic, except the one required to pass from *legato* to *staccato* gestures. We can argue the Heuristic Conjecture 6.1, connected with Definition 2.1—its physical equivalent of Condition 2 of the Definition.

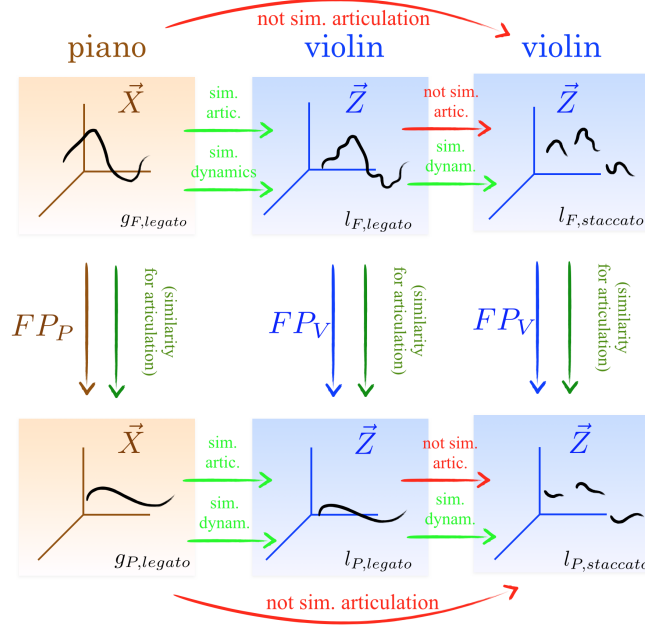**Conjecture 6.1.** *Heuristic conjecture. Two gestures, based on the same skele-*

Figure 3: Comparison between *forte* gestures on the piano and on the violin, *piano* gestures on both instruments, as well *legato* and *non legato*, as well transitions from forte to piano (*diminuendo*), and from *legato* to *staccato*.

ton, are similar[25] *if and only if they can be connected via a transformation:*

1. *that homotopically transforms a gesture into the other,*

2. *and that also leads to similar changes in their respective acoustical spectra.*

Homotopy is a necessary, but not sufficient, condition to get similar gestures.

Changes of loudness in orchestral playing lead also to changes in timbre, and so to changes in the spectra. The tridimensional diagram of Figure 4 shows how we can compare gestures and their acoustical results inside the same diagram. In this way, the analysis of the spectrograms can be compared with the analysis of the gestural curves in space and time. Moreover, we may define a *degree of similarity* considering a parameter λ that is equal to 1 for two identical gestures, and decreases to zero for gestures more and more different.

6.2   The category ∇

We describe now the ∇ category [38], used in Section 6.3. The notation ∇ denotes here an internal category in Top, generalizing the unitary interval $I = [0, 1]$ in the two-dimension plane. To define ∇, let us think of the triangular

---

[25]Two gestures are formally called *similar* if they satisfy the conditions of Definition 2.1.
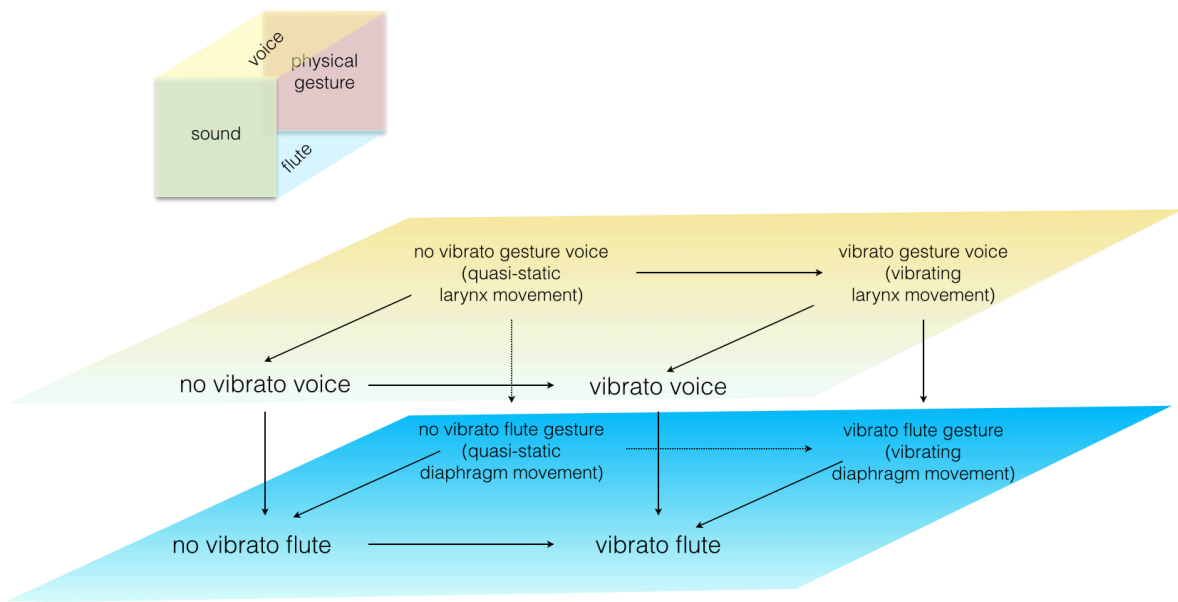
Figure 4: Tridimensional diagram that relates physical gestures with with their acoustical results [30], in the case of the voice versus flute, and *vibrato* versus *non vibrato*. This kind of diagrams is not always commutative: it is not commutative, for example, when a sound cannot be made by any human performer doing a musical gesture on an acoustic instrument.

region of space above the line $x = y$, consisting of points (x, y) with $0 \leq x < 1$ and $0 \leq y < 1$. The ∇ category has:

1. as objects, the points belonging to the diagonal, i.e. their pairs of coordinates that verify $x = y$, and

2. as morphisms, the points inside the upper triangle, i.e. their pairs of coordinates verifying $x \leq y$, the pairs (x, y) with $x \leq y$ of real numbers $x, y \in I = [0, 1]$.

The choice of the symbol ∇ is due to such a triangular form. We can use the symbol ∇ to distinguish the set of points in the space X and the set of arrows X^∇ having points in X, also indicated with $\vec{X}$.[26]

6.3 Parametrized gestures

The concept of *parametrized gestures* is introduced to formally take into account the influence of physical and physiological parameters on gestures.[27] For each choice α of parameters in a parameter space A, we have a corresponding gesture $g(\alpha)$ in $\Delta \vec{@} X$. Each choice of parameters implies a different "embodiment" of the given skeleton Δ of Figure 2. Moreover, we require that the mapping such as $g : A \to \Delta \vec{@} X$ is continuous, and we define the set of these A-parametrized curves[28] as $\Delta \vec{@}_A X$. Figure 5 shows two slightly different gestures, corresponding to a slightly different choice of parameters α_1 and α_2. More formally, let ∇ be the topological category of curve parameters, and A a topological category. For every choice of parameters, we can define a curve with values in the topological category X, given by $q : A \to \nabla @ X$. This is equivalent to $q : A \times \nabla \to X$. This means that, for $\alpha \in A$, we have a curve $C(\alpha) : \nabla \to X$. Let us consider the unit interval $I = [0, 1]$ and a topological category A. Going along I with a parameter t ranging from 0 to 1, we can build a parametric curve having s as parameter (for example: s for the variable of $(\sin(s), \cos(s))$). In the rectangle $A \times I$, we have a set of values of s ranging from 0 to 1 *for each choice* of parameter $\alpha \in A$. A straight line in $A \times I$ is *not yet* a skeleton: it can instead constitute one of

---

[26]Curves in $\vec{X}$ are mapped into points of X, for example with functions taking the tail and the head of an arrow. We can interpret these functions as directed graphs of the set of arrows and points. In fact, we indicate as $\vec{X}$ a directed graph, identified by the curves $c : \nabla \to X$, whose the *head* and *tail* functions, projecting to X the final and initial points of continuous curves in $\vec{X}$. This action of projecting a curve into points can be generalized not only for final and initial points, but also for every other point belonging to the curve.

[27]Parametrized gestures have been used to mathematically describe the mechanism of voice in singing [37, Chapter 37]. The dimension of the hand of the pianist, the anatomy of the vocal tract for the singer, and the values of diaphragm pressure and position of the larynx are examples of parameter choices.

[28]We can define an *A-parametrized gesture* as a continuous function from the Cartesian product of the parameter space ∇ with the category A, with values in the topological space X, that is, $A \times \nabla \to X$. According to what is called the *currying theorem* in informatics, for a category C we have $C(X \times Y, Z) \vec{\to} C(X, Z^Y)$, where Z^Y are the curves from Y to Z, and we can write that $q : A \times \nabla \to X \sim q : A \to \nabla @ X$. As described in [38], an *A-addressed gesture with skeleton digraph* Δ *and body* X is a digraph morphism $g : \Delta \to A @ \vec{X}$ into the spatial digraph of $A @ \vec{X}$.
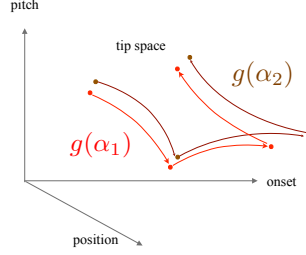
Figure 5: Two slightly different parametrized gestures, $g(\alpha_1)$ and $g(\alpha_2)$, corresponding to a bit different choice of parameters α_1 and α_2. The *tip space* is the space of the positions of a finger's tip of the hand of the pianist. A tip space indicated as position-onset-pitch has been used in [36].

its parts. These values can be mapped into a curve in the topological space X, exactly as it happens in the definition of gesture, from a skeleton of a digraph to a curve (system of curves) in a topological space.

A *morphism* $f : g \to h$ of two *addressed gestures* $g : \Gamma \to A@\vec{X}$ and $h : \Delta \to B@\vec{Y}$, is a triple $f = (t : \Gamma \to \Delta, a : B \to A, m : X \to Y)$, consisting of a digraph morphism t, an address change a, and a continuous functor of topological categories m, such that $\Gamma@_a Y \circ \Gamma@_A m(g) = t@_B Y(h)$. Finally, we have that an A-addressed gesture with skeleton digraph Γ and body X is a digraph morphism $g : \Gamma \to A@\vec{X}$, and we denote this set of gestures by $\Gamma@_A\vec{X}$. We can finally write the relations of equation 27 [38].

$$\Gamma@_a X : \Gamma@_A\vec{X} \to \Gamma@_B\vec{X}, \Gamma@_A m : \Gamma@_A\vec{X} \to \Gamma@_A\vec{Y}, t@_A X : \Gamma@_A\vec{X} \to \Delta@_A\vec{X}. \tag{27}$$

6.3.1 Parametrized hypergestures

The concept of parametrization can also be applied to hypergestures. Let us indicate with $\Delta@X$ the set of all the curves from Δ to X. Let g be a B-parametrized gesture, going from the skeleton Δ to the system of curves in the space X^B. Briefly, $g \in \Delta\vec{@}_B X$, where we consider the new set having the structure of topological space. In such a space, the gesture g is a point. As known, the morphisms in this new space, connecting points (gestures), are hypergestures. If we have a skeleton Γ, mapped via the hypergesture l into a system of A-parametrized curves connecting points g and h (both gestures $g, h \in \Delta\vec{@}_B X$), we say that l is an A-parametrized hypergesture, and we can write $l \in \Gamma@_A(\Delta\vec{@}_B X)$. If also this space has the structure of a topological space, we can write $\Gamma\vec{@}_A(\Delta\vec{@}_B X)$.

6.4 Musical forces

In the Introduction, we cited world-sheets. They can be framed in the topic of *musical forces* that shape their surface. World-sheets are obtained from the

minimization of the action for a Lagrangian including complex time[29] [33, 37]. The shape of the world-sheet is determined by an *(artistic) potential*, a force field.[30] Such a surface creates a tight relation between symbolic and physical gesture, representing "artistic" forces that transform thought into artistic expression. The concept of force is a classic metaphor in musical creativity [5, 24, 39, 40, 47]. The concept of such a potential is used in numerical calculations, and used to mean the characteristic combination of parameters that distinguish a gesture from another [38] but it has not been precisely defined yet. The role of the force field may be related to the non-trivial conversion of the score's mental reality to the performance's physical reality. This process depends upon several elements such as technical difficulties, training, technical ability/skills of the performer, and expressivity of the piece.

Disclosure statement

The author has no conflict of interest.

Acknowledgements

References

[1] Alunni, Charles. 2004. *Penser par le diagramme—de Gilles Deleuze à Gilles Châtelet*, chap. Diagrammes et Catégories comme prolégomènes à la question : Qu'est-ce que s'orienter diagrammatiquement dans la pensée ? Presse Universitaire de Vincennes.

[2] Antoniadis, Pavlos, and Frédéric Bevilacqua. 2016. "Processing of symbolic music notation via multimodal performance data: Brian Ferney-

---

[29]In [37] the *complex time* is introduced to extend the formalism of Physics to the mental reality of the score, where music flows in the imaginary time. The reality of the physical music performance flows in the real time [44].

[30]The question of such an *artistic* potential is currently under research. It has an analogy at the level of notes with the performance operators [35, 40], both mathematically and conceptually.

hough's Lemma-Icon-Epigram for solo piano, phase 1." *Proceedings of TENOR 2016* http://tenor2016.tenor-conference.org/papers/18_Antoniadis_tenor2016.pdf.

[3] Bachratá, Petra. 2011. "Modèles musicaux interactifs basés sur le geste pour l'analyse et la composition de musique mixte." *Revue Francophone d'Informatique Musicale, Maison des Science de l'Homme, Paris Nord* (1). http://revues.mshparisnord.org/rfim/index.php?id=123.

[4] Barbieri, Buonocore Antonio Dalla Volta Riccardo, Filippo, and Maurizio Gentilucci. 2009. "How symbolic gestures and words interact with each other." *Brain Lang* 110 (1).

[5] beim Graben, Peter, and Reinhard Blutner. 2016. "Toward a Gauge Theory of Musical Forces." In *Proceedings of the Quantum Interaction Conference 2016,* Lecture Notes in Computer Science.

[6] beim Graben, Peter, and Roland Potthast. 2009. "Inverse problems in dynamic cognitive modeling." *Chaos* 19 (1): 015103.

[7] Berlin, Brent. 2006. "The First Congress of Ethnozoological Nomenclature." *Journal of the Royal Anthropological Institute* 12 (s1): S23–S44.

[8] Bevilacqua, Frédéric, Norbert Schell, Nicolas Rasamimanana, Bruno Zamborlin, and Fabrice Guédy. 2011. *Musical Robots and Interactive Multimodal Systems*, chap. Online Gesture Analysis and Control of Audio Processing, 127–142. Springer-Verlag, Basel.

[9] Bevilacqua, Frédéric, Bruno Zamborlin, Anthony Sypniewski, Norbert Schnell, Fabrice Guédy, and Nicolas Rasamimanana. 2009. "Continuous Realtime Gesture Following and Recognition." In *Gesture in Embodied Communication and Human-Computer Interaction,* edited by Stefan Kopp and Ipke Wachsmuth, Lecture Notes in Computer Science. Springer.

[10] Buteau, Chantal, and Guerino Mazzola. 2000. "From Contour Similarity to Motivic Topology." *Musicae Scientiae* VI (2).

[11] Cont, Arshia, Shlomo Dubnov, and Gérard Assayag. 2011. "On the information geometry of audio streams with applications to similarity computing." *Audio, Speech, and Language Processing, IEEE Transactions on* 4: 837–846.

[12] Françoise, Jules, Baptiste Caramiaux, and Frédéric Bevilacqua. 2012. "A Hierarchical Approach for the Design of Gesture to Sound Mappings." In *Proceedings of the 9th Sound and Music Conference (SMC),* Copenhagen.

[13] Gentilucci, Dalla Volta Riccardo, Maurizio, and Claudia Gianelli. 2008. "When the hand speak." *The Journal of Physiology* 102: 21–30.

[14] Gentilucci, Maurizio, and Paolo Bernardis. 2006. "Speech and gesture share the same communication system." *Neurospychologia* 44 (178-190).

[15] Gentilucci, Maurizio, and Paolo Bernardis. 2007. "Imitation during phoneme production." *Neuropsychologia* 45 (3): 608–615.

[16] Gentilucci, Maurizio, and Riccardo Dalla Volta. 2008. "Spoken language and arm gesture are controlled by the same motor control system." *The Quarterly Journal of Experimental Psychology* 61 (6): 944–957.

[17] Godøy, Rolf, and Marc Leman, eds. 2009. *Musical Gestures: Sound, Movement and Meaning.* Routledge, New York.

[18] Henrotte, Gayle. 1992. "Music and Gesture: A Semiotic Inquiry." *American Journal of Semiotics* 9 (4): 103–113.

[19] Iazzetta, Fernando. 2000. *Trends in Gestural Control of Music*, chap. Meaning in Musical Gesture. Ircam, Centre Pompidou.

[20] Joyal, André. 2008. "The theory of quasi-categories and its applications." *Centre de Recerca Matematica Barcelona* 45, II. http://mat.uab.cat/~kock/crm/hocat/advanced-course/Quadern45-2.pdf.

[21] Knees, Peter, and Markus Schedl. 2016. *Music Similarity and Retrieval.* Springer, Heidelberg.

[22] Kosko, Bart. 1994. *Fuzzy Thinking: The New Science of Fuzzy Logic.* Hyperion, Westport.

[23] Kubovy, Michael, and Michael Schutz. 2010. "Audio-visual objects." *Review of Philosophy and Psychology* 1 (1): 41–61.

[24] Larson, Steve. 2012. *Musical Forces.* Musical Meaning and Interpretation. Indiana University Press, Bloomington.

[25] Lawvere, William, and Stephen Schanuel. 2009. *Conceptual Mathematics. A first introduction to categories.* Cambridge University Press.

[26] Leinster, Tom. 1998. "Basic Bicategories." *ArXiV* https://arxiv.org/pdf/math/9810017v1.pdf.

[27] Lurie, Jacob. 2008. "What is an ∞-Category?." *Notices of the AMS* 55 (8): 949–950.

[28] Mac Lane, Saunders. 1971. *Categories for the Working Mathematician.* Springer, Heidelberg.

[29] Mannone, Maria. 2011. *Dalla Musica all'Immagine, dall'Immagine alla Musica. Relazioni matematiche fra composizione musicale e arte figurativa (From Music to Images, from Images to Music. Mathematical relations between musical composition and figurative art).* Palermo: Edizioni Compostampa.

[30] Mannone, Maria. 2017. "Musical Gestures between Scores and Acoustics: A Creative Application to Orchestra." Ph.D. thesis, University of Minnesota.

[31] Mannone, Maria, and Giuseppe Compagno. 2014. "Characterization of the degree of Musical non-Markovianity." *ArXiV* http://arxiv.org/pdf/1306.0229v2.pdf.

[32] Mannone, Maria, Rosario Lo Franco, and Giuseppe Compagno. 2013. "Comparison of non-Markovianity criteria in a qubit system under random external fields." *Physica Scripta* T153: 014047.

[33] Mannone, Maria, and Guerino Mazzola. 2015. "Hypergestures in Complex Time: Creative Performance Between Symbolic and Physical Reality." In *Proceedings of the MCM 2015 Conference,* edited by T. Collins et al., 137–148. Springer, Heidelberg.

[34] Mayerthaler, Willi. 1987. "System-independent morphological naturalness." In *Leitmotifs in Natural Morphology,* Vol. 10 of *Studies in Language Companion Series* edited by W. U. Dressler, W. Mayerthaler, O. Panagl, and W. U. Wurzel, 25–58. Benjamins.

[35] Mazzola, Guerino. 2011. *Musical Performance.* Computational Music Science. Springer, Heidelberg.

[36] Mazzola, Guerino, and Moreno Andreatta. 2007. "Diagrams, Gestures and Formulae in Music." *Journal of Mathematics and Music* 1 (1): 23–46.

[37] Mazzola, Guerino, René Guitart, Jocelyn Ho, Alex Lubet, Maria Mannone, Matt Rahaim, and Florian Thalmann. To appear 2018. *The Topos of Music III: Gestures.* Computational Music Science. Springer, Heidelberg.

[38] Mazzola, Guerino, and Maria Mannone. 2016. "Global Functorial Hypergestures over General Skeleta for Musical Performance." *Journal of Mathematics and Music* http://dx.doi.org/10.1080/17459737.2016.1195456.

[39] Mazzola, Guerino, Joomi Park, and Florian Thalmann. 2011. *Musical Creativity.* Computational Music Science. Springer, Heidelberg.

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

[41] Nobile, Luca. 2013. *Iconicity in Language and Literature 10 : Semblance and Signification*, chap. Words in the mirror: Analysing the sensorimotor interface between phonetics and semantics in Italian, 101–131. Amsterdam/Philadelphia: John Benjamins.

[42] Nobile, Luca. 2015. "Phonemes as images: An experimental inquiry into shape-sound symbolism applied to the distinctive features of French." In *Iconicity: East meets West,* edited by Herlofsky William Shinohara Kazuko Hiraga, Masako and Kimi Akita, 71–91.

[43] Ohala, John. 1982. "The voice of dominance." *The Journal of the Acoustical Society of America* 72 (S66).

[44] Primas, Hans. 2007. "Non-Boolean descriptions for mind-matter problems." *Mind and Matter* 5 (1): 7–44.

[45] Rinman, Marie-Louise, Anders Friberg, Ivar Kjellmo, Antonio Camurri, Damien Cirotteau, Sofia Dahl, Barbara Mazzarino, Bendik Bendikesen, and Hugh McCarthy. 2003. "EPS - An Interactive Collaborative Game Using Non-Verbal Communication." In *Proceedings of the Stockholm Music Acoustics Conference,* Stockholm.

[46] Roffler, Suzanne, and Robert Butler. 1968. "Factors that influence the localization of sound in the vertical plane." *The Journal of the Acoustical Society of America* 43: 1255–1259.

[47] Schoenberg, A. 1983 edition. *Theory of Harmony.* University of California Press.

[48] Schutz, Michael, and Scott Lipscomb. 2007. "Hearing gestures, seeing music: Vision influences perceived tone duration." *Perception* 36: 888–897.

[49] Sessions, Roger, and Edward T. Cone. 1979. *Roger Sessions on Music, Collected Essays.* Princeton Legacy Library.

[50] Spence, Charles. 2011. "Crossmodal correspondences: A tutorial review." *Attention, Perception and Psychophysics* 73: 971–995.

[51] Spivak, David. 2013. *Category Theory for Scientists.* MIT.

[52] Uznadze, Dimitri. 1924. "Ein experimenteller Beitrag zum Problem der psychologischen Grundlagen der Namengebung (An experimental contribution to the problem of the psychological foundations of naming)." *Psychologische Forschung* 5 (1-2): 24–43.

[53] Visi, Federico, Rodrigo Schramm, and Eduardo Miranda. 2014. "Gesture in performance with traditional musical instruments and electronics: Use of embodied music cognition and multimodal motion capture to design gestural mapping strategies." In *Moco '14: proceedings of the 2014 international workshop on movement and computing,* Paris.

[54] Wieland, Renate, and Jürgen Uhde. 2002. *Forschendes Üben. Wege instrumentalen Lernens. Über den Interpreten und den Körper als Instrument der Musik.* Bärenreiter, Neumünster.

[55] Winter, Michael. 2004. "Goguen Categories." *Journal on Relational Methods in Computer Science* 1: 339–357.

[56] Zbikowski, Lawrence. 2002. *Conceptualizing Music.* AMS Studies in Music. Oxford University Press.

[57] Zwiebach, Barton. 2004. *A First Course in String Theory.* Cambridge University Press, Cambridge.