Formalizing the Concept of Sound[1]

Hans G. Kaper
Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439

Sever Tipei
Computer Music Project, University of Illinois at Urbana-Champaign,
114 W. Nevada St., Urbana, IL 61801

Abstract

The notion of formalized music implies that a musical composition can be described in mathematical terms. In this article we explore some formal aspects of music and propose a framework for an abstract approach.

1 Introduction

Sounds and their attributes are traditionally defined at a relatively low level of abstraction. The terminology and basic concepts underlying the notion of a sound are tailored to Western music of the past three centuries and lack the level of abstraction needed, for example, for a significant body of recent works and some non-Western music. In this article we explore some formal aspects of music and propose a framework for a more abstract approach.

As Western musicians, we are trained to think in terms of the notated score. Accordingly, we view a piece of music as a collection of individual sounds represented by dots and ovals and characterized mainly by start time, duration, and pitch. The music notation system we use today is indeed precise in defining time and pitch, while also allowing for the relatively accurate transcription of dynamics and articulation.

Software synthesis offers examples in which the concept of "sound" becomes slippery, if not entirely obsolete. The often-quoted beginning passage of Risset's *Mutations I* [Risset 79] involves the transformation of a chord into a timbre: individual sounds dissolve into one sonority. When the spectrum of a sound changes dynamically, some partials may start long after or end well before the other partials, so the sound's beginning may not have much in common with its end. "Morphing," gradually moving partials from one sound to another, radically affects our perception of a sound.

We conclude that the notion of sound, however crucial to traditional music, has no universal meaning and is not a necessary part of the description of musical events. In this article we propose a formal definition of a sound that is precise and avoids many of the pitfalls indicated above. It is based on the observation that a sound is the manifestation of a complex audio wave. The audio wave is the universal object in the space of aural events. It has two aspects, one physical (the variation of the ambient air pressure, which makes the eardrums vibrate), the other psychophysical (the process that translates these vibrations into a perception of the sound). The definition we propose is sufficiently abstract that both aspects are accommodated.

2 Basic Formalism

The universal object in the space of aural events is the audio wave. Special cases are partial waves corresponding to pure tones, sound waves corresponding to sounds, and complex audio waves corresponding to entire musical compositions. Partial and sound waves are like threads floating in the space of aural events, from which the composer weaves the trajectory of a musical piece. This image suggests how we may formalize a musical composition.

2.1 Sound Space

Following Xenakis [Xenakis 92], we view an audio wave as a dynamic event that evolves in a multidimensional vector space. We call this vector space the *sound space*. Unlike Xenakis, however, we view time

[1] This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

as an independent variable, not as a degree of freedom in sound space. Sound is thus a vector-valued map from the time domain to sound space, and the description of sound space is an integral part of the definition of a sound.

A sound space is spanned by a set of independent vectors, each of which is associated with a degree of freedom of a sound. The vectors constitute a *basis* in the sound space. A basis is not unique, but a representation of a sound with respect to a given basis is. In the process of composing, the composer selects the basis vectors in the sound space (thus deciding which aspects of the sounds are going to be taken into account in the musical composition) and assigns values to the coordinates (thus deciding the actual nature of the composition). As composers of computer music, we have complete control over both these aspects.

2.2 Objects and Their Attributes

The object of a musical composition is a *complex audio wave*, which we denote generically by the symbol W. Two of its attributes are its starting time, $T_{w,0}$, and its duration, T_w. (The subscript w stands for "wave.") Thus, a musical composition (or its representation, the complex audio wave) is described by the set of all values $W(t)$ on an interval of length T_w beginning at $T_{w,0}$ and ending at $T_{w,1}$, where $T_{w,1} = T_{w,0} + T_w$,

$$W = \{W(t) : t \in [T_{w,0}, T_{w,1}]\}. \tag{1}$$

Note the difference between W and $W(t)$: W is a trajectory (a set of points) in sound space, whereas $W(t)$ is a single point in sound space, namely, the point on W associated with a particular value of time, t.

This description of a composition as a complex audio wave is independent of the time the piece actually starts or ends: both $T_{w,0}$ and T_w are attributes (degrees of freedom), to which we assign a value when we realize the piece. Since both are independent of time, they are *static* attributes.

The complex audio wave itself is the superposition of its constituent sounds. Hence, its value at any moment t in the interval $[T_{w,0}, T_{w,1}]$ is given by an expression of the form

$$W(t) = \sum_{i \in I_w(t)} S_i(t), \quad t \in [T_{w,0}, T_{w,1}]. \tag{2}$$

Here we encounter another attribute of the object W, namely I_w, the set of indices of all sounds in the audio wave; $I_w(t)$ is its value at time t, and the sum extends over all sounds that are "active" at time t. The ith sound contributes a value $S_i(t)$ to $W(t)$. The sound S_i may be a single partial or, more generally, a superposition of partials. Note that I_w is a *dynamic* attribute of the wave; its value may vary with time. In general, this variation occurs on a time scale that is characteristic for the composition.

We realize the composition by assigning values to its attributes. The values are real numbers in the case of static attributes and functions in the case of dynamic attributes. In the latter case, we specify the attribute's shape (envelope function) and size (maximum value).

The ith sound S_i in Equation (2) is an instantiation of the class of sounds. The definition of a sound is analogous to that of a composition. A sound S of duration T_s is the set of all its values $S(t)$ on an interval of length T_s beginning at $T_{s,0}$ and ending at $T_{s,1}$, where $T_{s,1} = T_{s,0} + T_s$,

$$S = \{S(t) : t \in [T_{s,0}, T_{s,1}]\}. \tag{3}$$

Here, $T_{s,0}$ and T_s are (static) attributes of the sound object, to which values are assigned when the piece is realized.

Just as a composition is the superposition of its constituent sounds, a sound is the superposition of its constituent partials. Hence, the value of a sound S at any moment t in the interval $[T_{s,0}, T_{s,1}]$ is given by an expression of the form

$$S(t) = \sum_{j \in I_s(t)} P_j(t), \quad t \in [T_{s,0}, T_{s,1}]. \tag{4}$$

The symbol I_s denotes the set of indices of all partials in the sound S; $I_s(t)$ is its value at time t, and the sum extends over all partials that "actively" contribute to the sound. The jth partial contributes a value $P_j(t)$ to $S(t)$. The index set I_s is a dynamic attribute of S; it varies in time, but the variation occurs generally on a time scale that is characteristic for the sound.

Finally, the jth partial P_j in Equation (4) is an instantiation of the class of partials. A partial P of duration T_p is again the set of all its values $P(t)$ on an interval of length T_p beginning at $T_{p,0}$ and ending at $T_{p,1}$, where $T_{p,1} = T_{p,0} + T_p$,

$$P = \{P(t) : t \in [T_{p,0}, T_{p,1}]\}. \tag{5}$$

We identify a partial—the elementary object from which the other objects (sound waves, complex audio waves) are constructed—with a sinusoidal wave with amplitude a, frequency f, and phase ϕ,

$$P(t) = a(t)\sin(2\pi f(t)t + \phi(t)), \quad t \in [T_{p,0}, T_{p,1}]. \tag{6}$$

When the amplitude, frequency, and phase are constant in time, Equation (6) represents a segment of a pure tone. In practice, at least the amplitude will vary with time, because the suppport of the partial (that is, the closure of the set of t for which $P(t) \neq 0$) must fit in the interval $[T_{p,0}, T_{p,1}]$. But in principle, all three variables (amplitude, frequency, and phase) represent dynamic attributes, which may vary with time.

2.3 Modifiers

Most musical and environmental sounds are actually quite complicated. They behave differently during the attack and decay phases of their active life, and the frequencies and amplitudes of their constituent partials seldom remain constant over the duration of a sound. Variations can range from tiny and slow oscillations affecting the frequency or amplitude of a single partial in a sound to the collective modulation of the frequencies or amplitudes of all the partials in a sound (vibrato, tremolo). They can also include effects on a larger time scale, as in "sound bends," glissandi, crescendo and descrescendo.

Because the parameters defining the various partials in a sound are completely independent, the mathematical description of an audio wave proposed here allows for the implementation of these effects by a corresponding modification of the parameters. For example, the formalism enables us to implement transients by introducing brief random disturbances in frequencies or amplitudes, whose occurrence can be controlled both in magnitude and in frequency. In fact, the list of modifiers can be extended much further to incorporate acoustic and psychoacoustic effects. The location of a sound source in space can be simulated by appropriate modulation of the phases; and reverberation effects, reflecting hall size, wall coverings, and mix of direct and reverberated sounds, can be accounted for in a consistent manner by introducing various delays and attenuation factors.

2.4 Time Scales

The expression (6) involves time *explicitly* (t multiplying the frequency f in the argument of the sine function), as well as *implicitly* (t as the argument of the amplitude a, the frequency f, and the phase ϕ). The explicit time variable defines the basic time scale of the partial; its unit is of the order of one period of a pure tone in the audio range. The implicit time, on the other hand, is associated with the action of the various modifiers of amplitude, frequency, and phase discussed in the preceding section. It is also the time associated with the evolution of the index set I_s in the expression (4) for the sound. The action of the modifiers occurs generally on a slower time scale than the one defined by the explicit time variable. The basic unit—typically a period of a modulating wave, the duration of a glissando, and so on—exceeds the unit of the basic time scale of the partial by several orders of magnitude and is more appropriately associated with a sound.

A similar phenomenon occurs at the level of the complex audio wave. Consider the expression (2), where time enters the audio wave again explicitly, as an argument of the constituent sounds S, and implicitly, as an argument of the index set I_w. The explicit time is associated with the evolution of the sound, while the implicit time is more characteristic of the entire composition. The unit of time for the latter exceeds that of the former by several orders of magnitude.

These observations suggest the existence of a hierarchy of time scales associated with the partial, the sound, and the complex audio wave. The hierarchy is indeed fundamental to the proposed formalism. It not only suggests a natural inheritance scheme of attributes and operations, but also offers a unifying structure for music composition.

3 Implications for Music Composition

Considering musical events as complex waves facilitates the view of composition as a homogeneous process that includes similar operations both at the level of sound synthesis and at the level of the entire piece. It also suggests that, at least in the case of electro-acoustic music, the main building blocks are waves (partials), not sounds.

The idea of time scales can be expanded to encompass cells, motives, phrases, themes, sections, and movements (to use the old terminology). Such a hierarchy would be based on "time intervals" of increasing "magnitude," to paraphrase Stockhausen [Stockhausen 57].

Seen from the higher vantage point of abstraction, the process of composition, including sound synthesis, can be described in terms of (a) a set of elements, and (b) an operation that "associates" these elements in more complex objects. The process repeats itself on various time scales, producing objects with a coarser and coarser granularity at increasingly higher levels of abstraction: partials congregating into sounds, sounds congregating into chords and/or melodies, melodies entering into contrapuntal relations, creating heterophony and motive transformations such as augmentation, distortion, and diminution, and so on.

We have been trained to think of sounds as defined by pitch and duration. But why not consider entities defined on other bases of the vector space and operations acting on time scales other than that determined by the frequencies in the audio range? The complex wave formula enables such nontraditional thinking. It suggests nonsequential associations that bridge large intervals of time: some modifiers may create a "timbre" at the beginning of a piece that relates to the product of other modifiers acting toward the end of the same piece, and both determine a third group of modifiers in another time regime of the complex wave.

Ultimately, the wave formula and the vector space point toward the possibility of creating non-narrative constructs that are free from the "words" (cells, motives, etc.) and phrases made out of pitches. Music could be more than an analogy for verbal discourse.

References

[Risset 79] Risset, J.-C., *Mutations I*, INA-GRM Recording (AM 56409) (1979)

[Xenakis 92] Xenakis, I., *Formalized Music, Thought and Mathematics in Music*, revised edition, Pendragon Press (1992)

[Stockhausen 57] Stockhausen, K., "*... how time passes ...*", Die Reihe, Vol. 3, Theodore Presser Co. (1957)