Tempo Curves Revisited: Hierarchies of Performance Fields

Author(s): Guerino Mazzola and Oliver Zahorka

Source: *Computer Music Journal*, Spring, 1994, Vol. 18, No. 1 (Spring, 1994), pp. 40-52
Published by: The MIT Press
Stable URL: https://www.jstor.org/stable/3680521

Guerino Mazzola and Oliver Zahorka
MultiMedia Lab
Institut für Informatik
Universität Zürich
Winterthurerstr. 190
CH-8057 Zürich, Switzerland
gbm@presto.pr.net.ch, zahorka@ifi.unizh.ch

Tempo Curves Revisited: Hierarchies of Performance Fields

In this article, we present a new view of the still-controversial phenomenon of musical tempo. Our perspective is guided by the ongoing development of a general theory of performance, together with its implementation as a performance workstation on the NeXTSTEP programming environment (Mazzola 1993). The main result of this approach is a formalism for the description of musical performance based on local and global hierarchies of particular vector fields. These *performance fields* are superimposed on the given musical score as a separate *performance score* and describe the guiding structure of the physical performance. As a special and concrete application of this general result, we expose the stratification of tempo into hierarchies of local tempo curves, connected to each other by systematic synchroneity relations. The theoretical material has been implemented in the MIDI software Presto version 2.0. This performance and composition tool enabled us to experiment with and to verify the practical use of the theoretical constructs. It was written in ANSI C by M. Waldvogel et al. and has a source code length of 110,000 words. It is available on Atari and Macintosh computers.

What Is Tempo?

The problem of tempo has many facets, depending on the point of view and interest of the investigator. Relating to history, we observe two facets. On one hand, the concept of tempo, as it is understood in classical musicology and as it has been made precise and operational by Johannes Mälzel's metronome, seems to be a simple matter. It suffices to look up the discourse about tempo and agogics in the pretentious and voluminous *Neues Handbuch der Musikwissenschaft* (Danuser 1992). It is based on Carl Czerny's *Pianoforte Schule* (Czerny 1846), which was written more than 150 years ago!

On the other hand, the use of this concept in contemporary performance theory, music psychology, or music technology has proved to be something less than straightforward. In fact, our verification of Czerny's tempo propositions, if taken literally, yields a poor, quasi-mechanical performance.

Desain and Honing (1992b) conclude from their experiments with tempo deviations and tempo curves that the tempo curve "is a dangerous notion, despite its widespread use and comfortable description, because it lulls the users into the false impression that it has a musical or psychological reality. There is no abstract tempo curve in the music nor is there a mental tempo curve in the head of a performer or listener." This conclusion must be emphatically contradicted. The problem is not the a priori concept of tempo curves, but rather its elaboration for realistic applications, as demonstrated in this paper.

In the reductionist, physicalistic perspective of traditional electronic music, as documented by Eimert and Humpert (1973), tempo is even eliminated in favor of a unique time axis measured in milliseconds: "Electronic music neither knows tempo nor metronome marks, but it documents its connections to the phenomenon of time by the most precise time indications which exist in music." We should stress that music has its own symbolic reality beyond physics—an essential point in understanding tempo.

Musicological Definitions of Tempo

As historically founded and widely adapted by musicians and musicologists, the tempo or musical speed "is equal to the number of [musical] beats in the chosen [physical] time unit (duration)" (Albersheim 1974). Since the invention of the metronome by Mälzel in 1813 (Rien 1979), tempo has been measured by beats per minute, but it is up to the com-

Computer Music Journal, 18:1, pp. 40–52, Spring 1994

poser or performer to decide on the note value represented by the beat. In general, the beat of the metronome is the denominator of the time signature.

Here, as with other musicological definitions, we are confronted with a frequent phenomenon of musicological signs: they lack precision, and their meaning changes according to their current context. For example, Albersheim categorically denies the influence of a rubato effect on tempo; he views local agogics as a psychological fact while tempo remains constant!

We can also observe an unreflected mixture of symbolic score time and physical reality. For instance, the famous *Riemann Musik Lexikon* defines tempo as "absolute duration of note values." This approach to tempo has been adapted by Friberg (1991). The tempo problem is circumvented by reduction to the sequence of note durations, a simplification that cannot take the onset relations into account, in particular for polyphonic music. The tempo is reduced to a function of durations and is hence completely abandoned as an autonomous category.

Such a conceptual ambiguity is only acceptable as long as precise signs and relations are not required for adequate computerized data processing or for analysis in mathematical music theory. Despite their inaccurateness, we can learn from the above definitions that tempo is associated with the relation of score time and performance time. On a more formal level, tempo can be interpreted as the transformation of score time into performance time.

The Lack of Theory and Software

General representation languages and theories of musical objects—including a theory of time and performance—are still under development and should provide a basis for broad acceptance and use in both practice and research. As Desain and Honing (1992a) state on the representation of musical time and temporal structure, "We still lack a general theory of representation." From their experiments, these authors conclude that there is no chance for any formerly known abstract model of time structuring in music to gain acceptance.

A brief look at the tempo handling capabilities of today's commercial sequencing software reveals that appropriate software for tempo control is also lacking. Although their terminology is taken from classical musicology, time control is mostly based on discrete tempo changes. To our knowledge, the editing of continuous tempo curves is only available with two software products: Presto (by SToA Music) and Live (by Soft Arts).

Although discrete tempo control seems to be commonly used and accepted by certain practicing musicians, the tempo shaping of these sequencers is not satisfactory at all. If the tempo is subjected to discrete changes, the flow of time, as perceived by the listener, is flurried. The analogous perception of human beings is directly founded in nature: the surroundings of humans are, in general, not capable of changing states in a discrete manner. From this point of view, it doesn't make sense to treat musical time as discrete data. As Franz Liszt's teacher states, "A sudden slowdown or acceleration during a single note wastes the whole issue in this case" (Czerny 1840)—a remark that is easily verified by computer simulation.

Elementary Tempo Curves

To get off the ground with our formal discussion, let us call the symbolic, musical time E. This is a real number measured in units that we will call beats to fix the idea. For any concrete performance, we associate with the symbolic time E a physical time e, also a real number that is measured in seconds, say. Let us suppose that the *performance transformation* $\wp$

$$E \to e \tag{1}$$

is an invertible C^1 (continuously differentiable) function $e = \wp(E)$. According to the intuitive definition of tempo, we define the *tempo curve* associated with the transformation $\wp$ to be the derivative

$$T = \left(\frac{d\wp}{dE}\right)^{-1} [\text{Beats/sec}]\,. \tag{2}$$

Suppose that we know the physical time of the start time

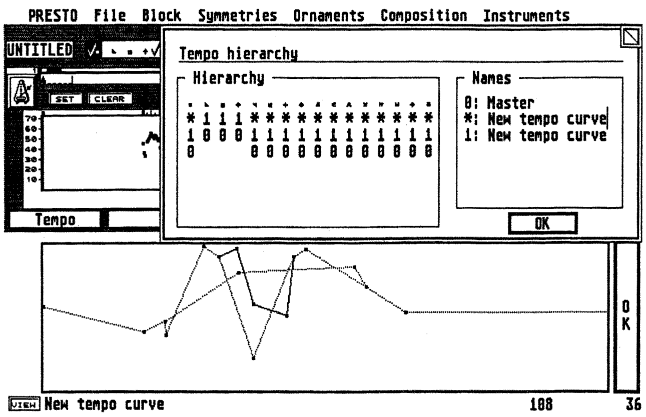


Figure 1. Graphical user interface for editing tempo hierarchies in Presto.

$$e_0 = \wp(E_0)\,. \tag{3}$$

Then the transformation $\wp$ is the well-known integral

$$\wp(E) = e_0 \int_{E_0}^{E} 1/T. \tag{4}$$

This easy formula allows any choice of the tempo curve as long as it is a positive, continuous function. We are not dealing with noncontinuous tempi because a more general concept of tempo will allow us to circumvent this situation, which we call the case of c*ellular tempo.*

For the standard assumption of a piecewise linear (i.e., polygonal) tempo, Equation 4 yields a sum of logarithmic functions of symbolic time. This type of tempo curve is implemented in the Presto software. The user is allowed to define and edit polygonal tempo graphically (Figure 1).

Performance Fields

To make the link with the general approach of our theory, we next consider the analogous situation for pitch. In this case, we are also given a symbolic pitch coordinate *H,* which will be measured in, for example, half-tone steps. With respect to the widespread MIDI codex, we give this unit the name *Key.* However, we shall admit real numbers as symbolic pitch values. In fact, musical thinking includes the infinite division of half-tone steps. Again, for a given performance, each symbolic pitch *H* is transformed into a physical pitch (i.e., the logarithm *h* of frequency) to be measured in Cents, for example.

The transformation $\wp$

$$H \to h$$

is again supposed to be induced by an invertible C^1 function $h = \wp(H)$, and the intonation curve *S* will be defined by

$$S = \left(\frac{d\wp}{dH}\right)^{-1} \left[\text{Key/Cent}\right]. \tag{5}$$

If we know the physical pitch $h_0 = \wp(H_0)$ of symbolic reference H_0, we have

$$\wp(H) = h_0 \int_{H_0}^{H} 1/S. \tag{6}$$

For example, the well-tempered tuning is the constant intonation curve of value $S = 10^{-2}$ (Key/Cent). Usually, the intonation curve is a periodic function with period 12.

If we put together tempo and intonation curves, we get the *tempo-intonation* vector field *TS* on the real *E-H*-plane $\mathbf{R}^2$ $\{E, H\}$

$$TS(E,H) = \big(T(E), S(H)\big)\,, \tag{7}$$

a continuous vector field on the symbolic plane of onset and pitch. By the fundamental theorem of ordinary differential equations (see Loomis and Sternberg 1968), there is a unique maximal integral curve

$$\int_X TS \tag{8}$$

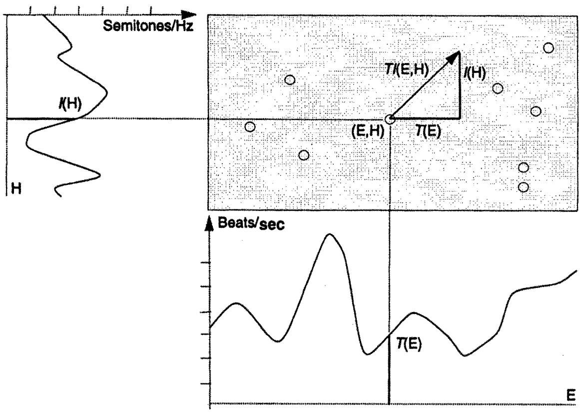
through a given point $X = (E, H)$. Suppose that this curve hits an initial point $X_0 = (E_0, H_0)$, where its physical value $x_0 = (e_0, h_0)$ in the *e-h*-plane is given. If this happens in the parameter value $-t_0$ of the integral curve, then we have

$$\wp(X) = x_0 + t_0 \cdot \Delta \tag{9}$$

where $\Delta = (1, 1)$ is the basic diagonal vector. This means that we may calculate the performance transformation $\wp(X)$ by direct integration of the tempo-intonation vector field *TS,* as shown in Figure 2.

For the following discourse, we suppose that the

Figure 2. The combination of tempo and intonation curves defines a two-dimensional performance field.



typical parameter set of onset, pitch, loudness, and duration for piano music is selected, though more general settings are at our disposal in our theory. We denote symbolic parameters by capital letters, and physical parameters by corresponding small letters. E stands for symbolic onset, e for the physical onset, and so on. All parameters are real numbers, and the corresponding real vector spaces are denoted by $\mathbf{R}\{E, H, L, D\}$ in the symbolic case and by $\mathbf{R}\{e, h, l, d\}$ in the physical case, respectively. The notation of particular subspaces, such as $\mathbf{R}\{E, H\}$, is evident. For a general set Π of symbolic parameters, we write $\mathbf{R}\Pi$. For the corresponding physical set π, we write $\mathbf{R}\pi$.

From the above analysis of tempo and intonation, a more general approach becomes feasible. We shall see below that this generality is by no means an academic game.

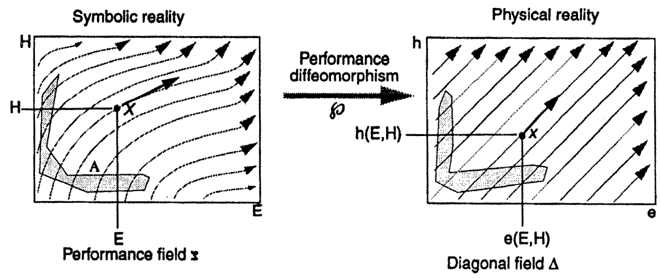
Suppose that we are given the *composition*, a finite set K of sound events in $\mathbf{R}\{E, H, L, D\}$. Each element $X \in K$ is transformed into a physical sound event $x = \wp(X)$ by means of a performance transformation

$$\wp : \mathbf{R}\prod \rightarrow \mathbf{R}\pi. \tag{10}$$

We now suppose that this transformation is defined on an open neighborhood U of K and is a C^1 diffeomorphism—that is, an invertible, continuously differentiable map onto an open neighborhood V of $\wp(K)$.

Performance fields are special vector fields that describe performance transformations. They perfectly generalize the situation studied above for tempo and intonation. We consider the diagonal constant vector field Δ on V

Figure 3. The performance transformation from symbolic to physical reality may be described by the inverse vector field of the diagonal field.



$$\Delta(x) = \Delta = (1,\ldots,1) \text{ for all } x \in V.$$

By a general technique of differential geometry, one may consider the so-called inverse image of the Δ-field, as in Figure 3. This is the C^1 performance field צ (Hebrew "TSadeh" for German *Tempo-Stimmung*) on U defined by the formula

$$\text{צ}(X) = J(\wp)^{-1}(X)(\Delta)\,, \tag{11}$$

where $J(\wp)$ is the Jacobian matrix at X

$$J(\wp) = \left(\frac{\partial x_i}{\partial X_j}\right)\Bigg|^{x_i = e,h,l,d}_{X_j = E,H,L,D}(X)\,. \tag{12}$$

By the fundamental theorem of ordinary differential equations, there is a unique maximal integral curve

$$\int_X \text{צ} \tag{13}$$

through every symbolic sound event X in K. If t is the curve parameter, we have

$$\frac{d}{dt}\int_X \text{צ}(t) = \text{צ}\left(\int_X \text{צ}(t)\right) \tag{14}$$

with the initial value

$$\int_X \text{צ}(0) = X. \tag{15}$$

Suppose now that the curve hits a symbolic point A for the parameter value $-t_0$. Then we have the straightforward equation

$$\wp(X) = \wp(A) + t_0 \cdot \Delta. \tag{16}$$

If we know the value of A under the performance transformation, the value of X is also known by Equation 16. Hence, it suffices to know the performance field and the performance transformation of special points to calculate the values for all points of the composition K.

We now make the assumption that there is a set I of symbolic points in U where the *initial performance transformation*

$$\wp_I = \wp\big|_I : I \to V \qquad (17)$$

is known and such that, for every point X in K, the integral curve of Equation 13 hits I. We further make the *compatibility hypothesis* that for any two points A and B in

$$I \cap \int_X \Xi, \qquad (18)$$

we have

$$\wp_I(B) - \wp_I(A) = (t_B - t_A)\cdot\Delta \qquad (19)$$

for the parameter values t_B and t_A of the curve where it hits B and A. This prevents ambiguities in the choice of reference values in the initial transformation.

A priori, there are different choices of a reference field in V, but the candidate Δ is a good one because it is suggested by the special cases of products of independent fields, as with tempo and intonation (see above), and by the simple symmetric expression in Equation 16 with respect to all parameters.

Articulation: The Two-Dimensional Case

Before we explore the general setting, we want to calculate an elementary default performance field with respect to onset and duration. We do not yet impose any conditions on articulation, such as legato or staccato. We only assume that a tempo field is present. This performance transformation

$$\wp : \mathbf{R}\{E,D\} \to \mathbf{R}\{e,d\} \qquad (20)$$

is defined by the two functions

$$e = \wp(E) = e(E) \qquad (21)$$

and

$$d = \wp(E,D) = d(E,D) = e(E+D) - e(E)\ . \qquad (22)$$

The Jacobian matrix looks like this:

$$\begin{bmatrix} \frac{\partial e}{\partial E} & \frac{\partial e}{\partial D} \\ \frac{\partial d}{\partial E} & \frac{\partial d}{\partial D} \end{bmatrix} = \begin{bmatrix} \frac{1}{T(E)} & 0 \\ \frac{1}{T(E+D)} - \frac{1}{T(E)} & \frac{1}{T(E+D)} \end{bmatrix}$$

and the corresponding performance field is

$$\Xi(E,D) = \big(T(E), 2T(E+D) - T(E)\big)\ . \qquad (23)$$

For a constant tempo, this is the diagonal tempo field, but for a nonconstant tempo, the D-component of this field is no longer independent of the E-coordinate; this means that the default E-D-performance field is not a product of two one-dimensional fields like the above field $TS = T \times S$.

However, there is a remarkable feature in this situation—it reveals that performance fields tend to build *hierarchies* in the following sense. We have a projection

$$\Xi(E,D) \Rightarrow T(E) \qquad (24)$$

of vector fields. This reminds us of the case of tempo-intonation fields. But we do not have a projection onto the second component; a D-field does not exist. In general, one will have an entire system of such projections, and this is essential in understanding the nature of performance and of its defining parameters. We call such a system of projections of performance fields a *performance hierarchy*.

To get an idea of how articulation acts on the default hierarchy in Equation 24, suppose that an articulation effect is required by stretching or shortening the durations by a factor of α percent. We get an articulation field Ξ_α that relates to the default field by a matrix product as follows:

$$\Xi_\alpha(E,D) = Q_\alpha(E,D)\Xi(E,D)\ , \qquad (25)$$

where

$$Q_a(E,D) = \begin{bmatrix} 1 & 0 \\ \frac{(1-a)}{a}\cdot\frac{T(E+D)}{T(E)} & 1 \end{bmatrix} \qquad (26)$$

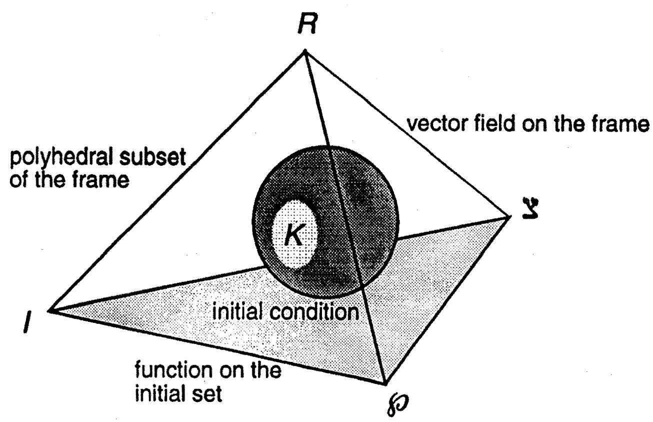


Figure 4. The local data of a musical performance are put together in an object called a performance cell. Its kernel consists of the local score events to be performed.

is a 2-by-2 matrix function of the articulation factor α, the tempo field T, and the coordinates E and D; in other words, the articulation is related to the default field by an automorphism of the tangent bundle of $\mathbf{R}\{E, D\}$, which locally looks like the matrix Equation 26.

In particular, the hierarchy of Equation 24 is preserved by the "articulation deformation" induced by the automorphism Equation 26

$$\mathfrak{Z}_\alpha(E,D) \Rightarrow T(E)\,. \tag{27}$$

This setting generalizes to all the articulation conditions, as they were proposed, for example, by Friberg (1991) in collaboration with Sundberg and Frydén (see Mazzola 1993 for a detailed study of these performance fields).

Performance Cells

The above formalism suggests a way to consider a basic, local structure that defines a performance transformation on a given set K of sound events in the n-dimensional parameter space $\mathbf{R}\Pi$. We call such a structure a *performance cell* for K. It consists of four ingredients: (1) an n-dimensional rectangle R, the frame of the performance cell, by definition,

$$K \subset \mathrm{R}; \tag{28}$$

(2) a continuous vector field $\mathfrak{Z}$ on R, the performance field of the cell; (3) a subset $\mathrm{I} \subset \mathrm{R}$ of initial events subjected to the conditions that every maximal integral curve that hits I verifies the compatibility hypothesis of Equation 19 relating to the following paragraph 4 and that every maximal integral curve through a point of K hits I; and (4) an initial transformation $\wp_I : I \rightarrow \mathbf{R}\pi$.

Notice that formula Equation 16 remains valid by definition in the sense that it defines a performance transformation on K as soon as this composition lives in a performance cell.

We visualize such a cell by a tetrahedron containing K as its "kernel" and such that the four vertices are the four elements R, $\mathfrak{Z}$, I and $\wp_\mathrm{I}$. Its edges or faces represent the relations expressed above, as in Figure 4. We formally write the symbol

$$\frac{K}{\left(\mathrm{R}, \mathfrak{Z}, I, \wp_I\right)} \tag{29}$$

to denote this cell.

The Performance Score

As we have seen, performance cells may build up a hierarchy (a *cellular hierarchy*). We will not go into details in this paper (refer to Mazzola 1993), but there is one thing that we should discuss in more detail in view of the structure of musical time: the local-global nature of performance.

From the classical notation of tempo, we recognize that a global continuous tempo curve is unreasonable—if not impossible—for complex scores. For instance, we may have a time segment that starts from a precise Mälzel metronome (M.M.) and includes a sequence of accelerandi and rallentandi. In classical notation, the appearance of the sign istesso tempo means that from this moment on, we should return to the initial value of the M.M. sign. This effect is a discontinuity in tempo.

But the reality is still more involved; consider an ornament like a trill or a grace note. It is evident that such a microstructure cannot have the same tempo as its context. The tempo of an ornament is a local one.

It is an easy task to imagine many other situations in which a single performance cell is not sufficient to describe the real performance. Think of the intonation specifications, as described by Friberg (1991), which depend on whether we are dealing with an ascending or descending melodic line. The result of this reflection is that we have to introduce a patchwork of performance cells. The definition runs as follows:

Let K be a set of sound events in $\mathbf{R}\Pi$—for example, the sounds of a classical (piano) score. A performance score for K is a partition

$$K = \bigcup_{i}^{\text{disjoint}} K_i \qquad (30)$$

together with a family of performance cells

$$\frac{K_i}{\left(\mathrm{R}_i, \mathfrak{X}_i, I_i, \left(\wp_I\right)_i\right)} \qquad (31)$$

for the *Ki*.

Together with the score, as it is represented in *K*, *the performance score constitutes the generating structural background to define a concrete performance.* Note that, in general, the performance score cannot be retraced from the performance by obvious ambiguities. In particular, the partition of Equations 30 and 31 is completely encrypted while performing because it is mainly rooted in the analysis of the musical work preceding the performance. (See Mazzola 1992 for an overview of the methodology of analytical tools and their role in performance.

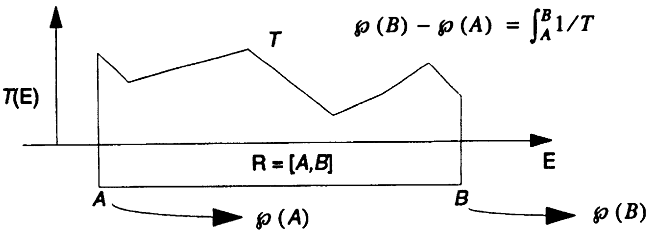
The Local-Global Principle

We now want to discuss the general principle of patching local performance cells to global hierarchies on the level of tempo. As a concrete realization of this approach, we describe some of the features of the sequencer and composition software Presto on the level of interactive graphical tempo editing.

Local Tempi

For the remainder of this paper, we work mainly on the onset space, that is, on $\mathbf{R}(E)$ and, correspondingly,

Figure 5. The performance cell for onset time reduces to a simple tempo cell. It consists of the tempo curve on an interval from onset A to onset B together with the initial values for the bordering points of the interval.



on $\mathbf{R}(e)$; these spaces will be identified with the set $\mathbf{R}$ of reals if no confusion is possible. To explain the concept of performance cells for the domain of onsets, we require the following data. They define a performance cell of onsets or *tempo cell* (Figure 5):

The frame of a tempo curve is a real interval $\mathrm{R} = [A, B]$, $A < B$, and the onset collection K is assumed to live in $[A, B]$.
The performance field $\mathfrak{X}$ is a continuous curve T, $T(E) > 0$ for all E, defined on R.
The initial set I is the set $\{A, B\}$ of edge points in $\mathbf{R}$. This restriction is made in view of the practical scope of the present paper.
The initial transformation is defined by the sequence $(\wp(A), \wp(B))$.

We assume that the compatibility condition

$$\wp(B) - \wp(A) = \int_A^B 1/T \qquad (32)$$

is fulfilled.

The physical onset of an element X in K is given by

$$\Delta(X) = \int_A^X 1/T\,. \qquad (33)$$

and then setting

$$\wp(X) = \wp(A) + \Delta(X)\,. \qquad (34)$$

Tempo cells are the local framework for tempo. The next step is to construct a procedure to generate tempo *scores*—that is, performance scores on the onset level. This means creating a global system of mutually related local tempi.

Splitting Performance Cells

A priori, we now know what a tempo score looks like; we give ourselves a partition of the onsets and assign a tempo cell to each of the members of that partition. However, this is a fairly abstract procedure because, in reality, nobody is willing to produce completely independent tempo cells. There are many relations governing the creation of a tempo score. To explain the general idea, let us briefly digress on the so-called splitting technique for performance cells.

The idea of splitting a performance cell

$$\frac{K}{\left(\mathrm{R}, \aleph, I, \wp_I\right)}$$

for a set K of sound events runs as follows. We first split the kernel K into the disjoint union of two subsets

$$K = K_1 \cup K_2 \quad (35)$$

which is given for some analytical reason. This allows for the second step: creating a virtual copy of the vertices [symbols] of the performance cell

$$\frac{K}{\left(\mathrm{R}, \aleph, I, \wp_I\right)} \Rightarrow \frac{K_1}{\left(\mathrm{R}, \aleph, I, \wp_I\right)}, \frac{K_2}{\left(\mathrm{R}, \aleph, I, \wp_I\right)} \quad (36)$$

Third, the first split cell for K_1 is left as a *mother cell,* whereas the second cell has changed its parameters such that it becomes a *daughter* of the mother cell. This means that we perform operations

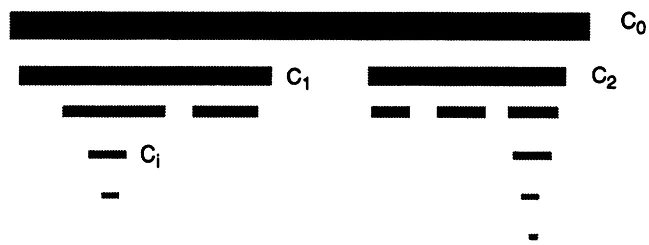
$$\left(\mathrm{R}, \aleph, I, \wp_I\right) \Rightarrow \psi\left(\mathrm{R}, \aleph, I, \wp_I\right) = \left(\mathrm{R}', \aleph', I', \left(\wp_I\right)'\right)$$

on the given cell data. This relation remains alive for the sequel; that is, if we change the mother's data, the daughter's data will change accordingly under the Ψ operator.

Tempo Hierarchies

Before introducing the Ψ operator for tempo, we must define the tempo hierarchies, starting from a *master tempo* cell C_0 for K_0

Figure 6. The tempo hierarchy consists of tempo cells with frames that are either disjoint or subframes of one other.



$$\frac{K_0}{\left(\left[A_0, B_0\right], T_0, \wp_0\left(A_0\right), \wp_0\left(B_0\right)\right)} \quad (37)$$

where we ignore the initial set $\{A_0, B_0\}$ because it is obvious. A daughter C_1 of this cell is, by definition, a restriction of this cell to a proper subinterval $[A_1, B_1]$ of $[A_0, B_0]$, together with the subcomposition

$$K_1 = K_0 \cap \left[A_1, B_1\right] \quad (38)$$

and the new initial data

$$\wp_1\left(A_1\right) = \wp_0\left(A_1\right) \text{ and } \wp_1\left(B_1\right) = \wp_0\left(B_1\right)$$

as they are given by the master tempo. The tempo of the daughter is

$$T_1 = T_0\big|\left[A_1, B_1\right]. \quad (39)$$

Equation 40 defines an inclusion of tempo cells

$$\frac{K_0}{\left(\left[A_0, B_0\right], T_0, \wp_0\left(A_0\right), \wp_0\left(B_0\right)\right)} \cup \frac{K_1}{\left(\left[A_1, B_1\right], T_1, \wp_1\left(A_1\right), \wp_1\left(B_1\right)\right)}. \quad (40)$$

The next tempo cell deduced from the master may either be a *sister* or a daughter C_2 of C_1; in the latter case, a *granddaughter* of the master C_0.

If we produce a sister C_2, her frame

$$[A_0, B_0] \supset [A_2, B_2] \quad (41)$$

has to be disjoint from the frame of her sister C_1

$$[A_1, B_1] \cap [A_2, B_2] = \varnothing. \quad (42)$$

This condition guarantees that the kernels of these sisters have well-defined *home frames* and therefore performance transformations.

We again take the restriction field

$$T_0\big|_{[A_2,B_2]} \tag{43}$$

and the initial values

$$(\wp_2(A_2),\wp_2(B_2)) = (\wp_0(A_2),\wp_0(B_2)) \tag{44}$$

for the sister C_2. The kernel of the sister is

$$K_2 = K_0 \cap [A_2, B_2]. \tag{45}$$

If a daughter C_2 of C_1 is produced, we require a proper inclusion

$$[A_1, B_1] \supset [A_2, B_2] \tag{46}$$

and

$$(\wp_2(A_2),\wp_2(B_2)) = (\wp_1(A_2),\wp_1(B_2)),$$

as well as

$$T_2 = T_1\big|_{[A_2,B_2]}. \tag{47}$$

This method is repeated to produce an entire hierarchy of local tempo cells

$$\left(\frac{K_i}{\left(\left[A_i,B_i\right],T_i,\wp_i\left(A_i\right),\wp_i\left(B_i\right)\right)}\right)_{i=0,\ldots,N} \tag{48}$$

We impose the above mutually exclusive conditions of Equations 42 and 46, respectively, and the accompanying conditions for each couple of cells of this hierarchy. This means that either two frames are disjoint or one is a subframe of the other, as illustrated in Figure 6.

Hence, every point X of the original composition K_0 is contained in exactly one smallest *home cell* with *home index* $i(X)$ of X.

This construction partitions K_0 into an evident disjoint union

$$K_0 = \bigcup_{j=\text{home index}}^{\text{disjoint}} L_j \tag{49}$$

of subcompositions according to home cells. Hence, every X in K_0 has its uniquely determined home cell

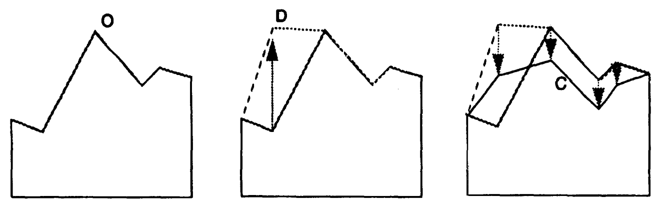


Figure 7. In Presto, a given original curve (O) is deformed to a new curve (D) by dragging a handle. The program corrects the defect in duration that is thus produced by preserving the shape (C).

with home index $i = i(X)$

$$\frac{L_i}{\left(\left[A_i,B_i\right],T_i,\wp_i\left(A_i\right),\wp_i\left(B_i\right)\right)} \tag{50}$$

and we may calculate unambiguously the physical time $e(X)$ according to the home cell.

Operating on Cells of a Tempo Hierarchy

For the cell of the tempo score (Equation 50), we can now do operations Ψ of the following types:

1. The frame $[A_i, B_i]$, the kernel L_i, and the initial values $\wp_i(A_i)$, $\wp_i(B_i)$ remain unaltered (always as a function of the transformation values of the immediately dominating mother cell!).
2. The field T_i may be changed into a new field $\Psi(T_i)$ with the condition

$$\int_{A_i}^{B_i} 1/\psi\left(T_i\right) = \int_{A_i}^{B_i} 1/T_i \, . \tag{51}$$

3. The field values at the initial points remain unaltered

$$\begin{aligned} T_i(A_i) &= \psi(T_i)(A_i) \\ T_i(B_i) &= \psi(T_i)(B_i). \end{aligned} \tag{52}$$

This means that the total duration within a cell is kept constant, and the tempo curve fits with its contiguous curves at the border points.

Conditions 1 and 2 are evident synchronization conditions for a daughter with respect to her mother. Hence, the operations Ψ cannot affect the time frame given by the mother. This is of great importance because it allows for a perfect fitting of the total duration of a small, local process—say, a trill or a rubato—with respect to the regular pulse of the mother tempo. The third condition simply guaran-

Figure 8. The original score of Czerny's exercise in agogics following his famous Pianoforte-Schule.

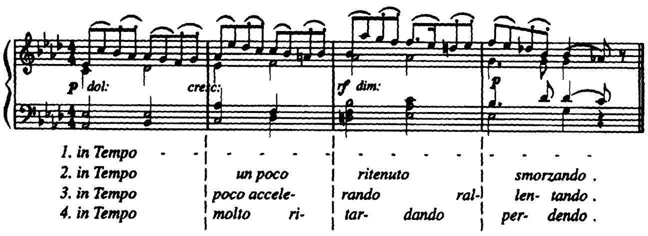
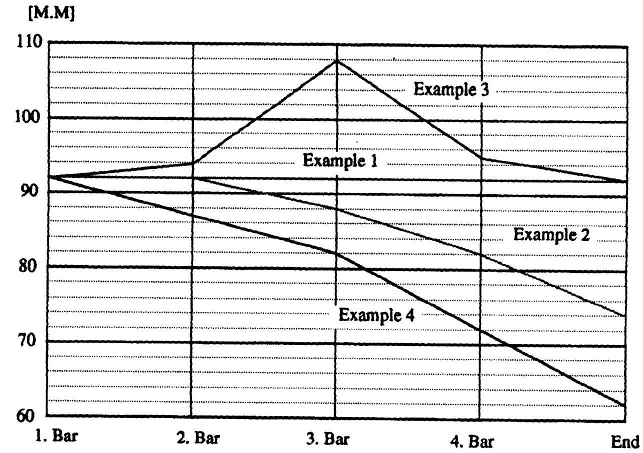


Figure 9. Translation of the four proposed agogical variants of Czerny's exercise into tempo curves (M.M. = quarters per minute).



tees that local tempi do not jump on the borders of their frames.

It is important to realize that every Ψ-operation on a cell immediately changes the parameters of its daughters, granddaughters, and so on. Splitting is really a method to organize the influence of an operator Ψ on the entire family of the actual cell!

Operations for Presto

In principle, a wide variety of curves and operations on curves may be considered (see Mazzola 1993 for a broader discussion, including splines). In view of the following concrete examples, we want to restrict to the Ψ-operations for Presto. In Presto, tempo has the shape of positive, polygonal curves. Suppose that we are given a cell

$$\frac{L_i}{\left(\left[A_i, B_i\right], T_i, \wp_i(A_i), \wp_i(B_i)\right)}. \qquad (53)$$

There are two mouse-driven methods for influencing the tempo curve (Figure 7). The user can introduce a new knot by clicking anywhere except on the abscissas of existing knots. The neighboring knots will automatically be connected to the new knot. Alternatively, the user can drag an existing knot of the polygon to another position. The connecting straight lines will follow the movement like rubber bands.

In general, such a new curve will not fulfill condition 2 above. The program now tries to deform this nonsynchronous curve in order to recover the original time conditions in a three-step approximation procedure. It is essential that the shape of the new curve is conserved as far as possible. This supports the user's intention to deform the old curve in a certain direction—though ignoring the defects of synchroneity that the user will produce as a side effect.

Of course, this type of interactive definition of the operator Ψ is very useful for the working musician, but it lacks theoretical background in the sense of performance rules, as they are discussed in Friberg 1991. However, our primary goal was to build an experimental tool for testing the efficiency of the theoretical machinery in a classical situation.

Two Examples: Czerny and Chopin

In this section, we present and discuss two concrete musical examples from classical piano literature. The first is taken from the famous *Pianoforte-Schule* by Carl Czerny; the second is a portion of two bars from the *Impromptu* op. 29 in A-flat major by Frédéric Chopin.

Technically, Presto allows the user to enter musical data via common MIDI sequencing or graphical input. The material is saved in Presto format or in standard MIDI-file format. Presto format stores symbolic onset, duration, loudness, and pitch as well as MIDI channel, program change, and MIDI controllers, together with the complete data of tempo hierarchies. Tempo hierarchies may be defined independently for arbitrary sets of voices. In contrast to the MIDI format, this format represents simultaneously the classical score events (symbolic parameters) and the performance-oriented tempo score.

Figure 10. Tempo hierarchy of more human agogics for the Czerny example (Figure 8).

Figure 11. Two bars of Chopin's Impromptu *op. 29, containing trills and arpeggios.*

Figure 12. Tempo hierarchy for the Chopin example (Figure 11).

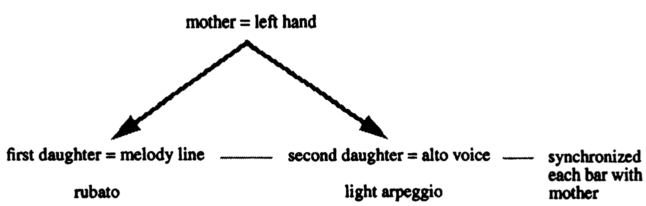
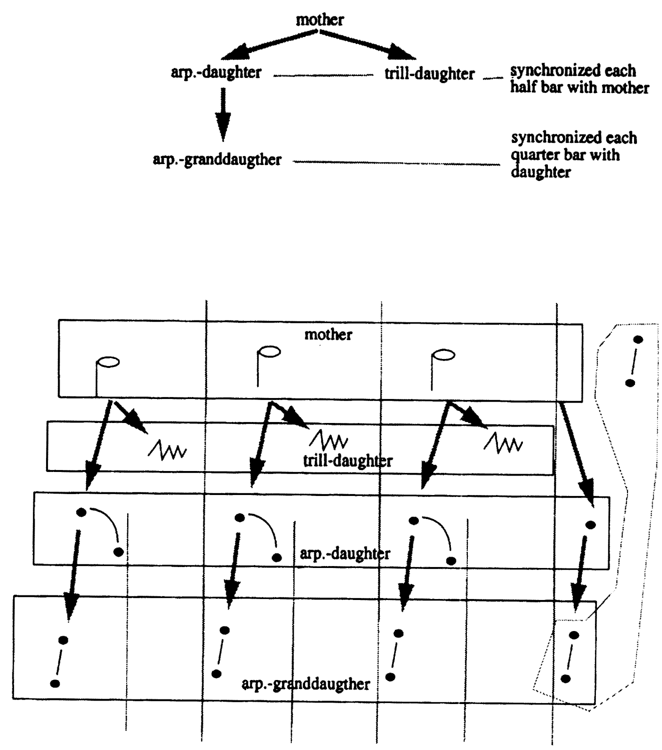


Figure 10.

Figure 11.



Czerny: Verification of Classical Tempo Rules via Computer Software

Czerny's *Pianoforte-Schule,* written about 1840, seems to be an excellent example for applying tempo curves in a classical context. Still used in the education of professional pianists and respected by musicologists, Czerny's pedagogical works are believed to be most precise in terms of performance rules, at least with respect to the composition style of his time.

In regard to performance fields, we consider this example as a verification of classical tempo rules by use of adequate software.

To get an impression of Czerny's ideas of ideal performance, we have translated the score and its comments from the example discussed in Danuser (1992, p. 295), shown in Figure 8, as faithfully as possible into computer data in Presto format. While accomplishing this task, we encountered some difficulties that reveal the vagueness of the author's descriptions. For instance, Czerny doesn't write down initial tempo in terms of Mälzel's metronome (although he could have done so), but instead requires the amount of tempo change to be "less than 1/6 to 1/5."

Following Czerny's instructions, we got four different tempo curves, as shown in Figure 9. The first is a mechanical performance with the advice to play the phrase with constant tempo. The second is a ritardando that begins at the end of the first bar. In the third example, Czerny proposes an accelerando in the second bar, followed by a decreasing ritardando in the third and fourth bars. In the last performance suggestion, there is an increasing ritardando from the second bar, leading to a perdendo effect.

According to Czerny, the third performance proposition is the most suitable for the phrase at hand. After listening to the four versions, we came to the same conclusion as Czerny, but we also noticed a frightening sterility in all four examples. As long as all voices of the piece are tied to the same performance cell, its oversimplified tempo structure doesn't relieve computer performance from the feeling of a machine at work.

What we need here are ramified hierarchies of tempo curves to enable different voices—especially for the left and right hand in piano music—to have their own tempo shaping. Although Czerny was aware of the possibilities given by the independency of left- and right-hand tempo, he did not refer to it in the above examples.

By introducing a simple tempo relation of a mother curve for the chords in the left hand as a conductor of two independent daughter curves—one for the melody line in the right hand and one for the alto voice in the left hand—we gain flexibility in tempo shaping. This flexibility is used to amplify the exposition and articulation of the musical structure.

The conditional temporal independency of the melody voice enables us to create a local tempo curve simulating such characteristics of a human performer as "leap-gap" and "faster-uphill"—characteristics shown in the psychoacoustic studies of Friberg, Sundberg, and collaborators—leading to a musically more satisfying result. The melody daughter curve of the right hand and its conducting mother curve are shown in Figure 10.

Chopin: Investigation of Tempo Hierarchies

To further explore the possibilities of multiple tempo curves, we decided to use a two-bar excerpt of Chopin's *Impromptu* op. 29 as a playground. This short phrase has a rather implicit notation and uses many productive signs—that is, signs that invoke improvisation, such as trills. The score is shown in Figure 11.

All of these musical signs call for a deformation of the generic performance score with respect to time. Besides the trills, which always have their own local tempo curve, the text contains arpeggios and grace notes that introduce local variations of the global tempo curve.

To achieve a more natural-sounding performance, we used a twofold branched hierarchy, as shown in Figure 12.

The mother curve can be understood as the conductor of the right hand trill daughter and the left hand arpeggio daughter, which itself is the conductor of its subordinate apreggio granddaughter.

The description of the relations among tempo curves in terms of family relations is used by Albersheim (1974). He compares the musical time variation called bound rubato to a walk of father and son. Compared to the father, the son has a smaller stride and less constant speed; now and then, he has to catch up with his father's faster pace.

With this rather simple, yet new, possibility of tempo curve hierarchies, a wide range of time structures becomes attainable. Highly complex time displacements can be achieved even by a tempo hierarchy that is not greatly branched, like the one used, without touching one single note's structural parameters, that is, its onset time and duration.

Conclusions

The discrimination of the reality of symbolic score data from the physical reality is crucial to capture and properly understand the complexity and the process-like nature of musical performance.

Performance fields are neither a mathematical game nor a mere academic approach to the problem of performance. Given the possibility of local and global stratification, they operate as a performance score and form an adequate representation of the formal structure underlying an actual performance.

The wide variation of individually perceived time is a well-known fact. Time layers are a quite natural and common phenomenon in the human perception of the world. The splitting and layering of time in the form of tempo hierarchies prove to be equivalent to the thinking of performers.

A software package for the realization of and experimentation with such tempo hierarchies is available.

Demo Version and Music Examples

The aforementioned musical examples are available via NeXT-mail from gbm@presto.pr.net.ch along with a demo version of Presto to allow interested individuals to experiment with the abilities of tempo curve hierarchies and to examine the presented material.

Acknowledgments

This work is supported by Swiss National Science Foundation Grant 21-33651.92. We are grateful for continuous encouragement from Gerald Bennett and Ernst Lichtenhahn.

References

Albersheim, G. 1974. *Zur Musikpsychologie.* Wilhelmshaven: Heinrichshofen Verlag.

Czerny, C. ca. 1840. *Vollständige theoretisch-praktische Pianoforte-Schule, von dem ersten Anfange bis zur höchsten Ausbildung fotschreitend, und mit allen nötigen, zu diesem Zwecke eigens komponierten zahlreichen Beispielen, in vier Teilen, verfaßt von Carl Czerny,* op. 500 Band 3. Wien.

Czerny, C. ca. 1846. *Die Kunst des Vortrags der ältern und neuen Claviercompositionen oder: Die Fortschritte bis zur neuesten Zeit. Supplement (oder 4. Theil) zur großen Pianoforte-Schule,* op. 500. Wien..

Danuser, H., et al. 1992. *Musikalische Interpretation* Bd. 11. Edited by H. Danuser. Wiesbaden (Neues Handbuch der Musikwissenschaft).

Desain, P., and H. Honing. 1992a. *Music, Mind and Machine: Studies in Computer Music, Music Cognition and Artificial Intelligence.* Amsterdam: Thesis Publishers.

Desain, P., and H. Honing. 1992b. "Tempo Curves Considered Harmful." In *Music, Mind and Machine: Studies in Computer Music, Music Cognition and Artificial Intelligence.* Amsterdam: Thesis Publishers.

Eimert, H., and H. U. Humpert. 1973. *Das Lexikon der elektronischen Musik.* Regensburg: Bosse Verlag.

Friberg, A. 1991. "Generative Rules for Music Performance: A Formal Description of a Rule System." *Computer Music Journal* 15(2): 56–71.

Loomis, L. H., and S. Sternberg. 1968. *Advanced Calculus.* Reading, Massachusetts: Addison-Wesley.

Mazzola, G. 1992. "Musical Performance and Vector Fields: Mathematical Music Theory between Physics and Esthetics." To appear in *Bericht 1993 Interdisziplinärer Arbeitskreis "Musik- und Kunstinformatik."* Johannes-Gutenberg-Univ. Mainz..

Mazzola, G. 1993. *Geometry and Logic of Musical Performance.* Swiss National Science Foundation Report. University of Zürich.

Riemann Musik Lexikon. 1882–1967. Edited by H. H. Eggebrecht et. al. Schott, Mainz.

Rien, R. 1979. "Beethovens Verhältnis zum Metronom." *Musik-Konzepte,* Heft 8.