23

Statistics

O sancta simplicitas!
Jan Hus (1370–1415)

This chapter deals with a special type of experiment in performance theory, since experiments are necessary to test the relevance of a theoretical approach to real performance. How can we know that such an approach is explaining what we are experiencing in performance? The question is quite hard, because it is difficult in music to distinguish the creative subjective aspect from the scientific objective one. There are essentially two types of experiments with an expressive performance theory:

- Construct synthetic performances and test their quality by psychometrical methods, as done by the KTH school, for example. This is the psychological approach. It is important, but it does not tell us how to construct performance tools except by trial and error. It just takes the subjectivity of the listeners as a variable and ponders it against the output of a performance machine.
- Take human performances and investigate their fitting quality with rationales of the theory, e.g. with analytical, gestural, or emotional rationales. This one also refers to the aesthetic human individual dimension, but it realizes it in the realm of performers—if possible even distinguished performers, such as Horowitz, Brendel, or Pollini. The comparison is not with synthetic performances but with rationales of performance. This is completely logical, since the performance's expressivity refers to those rationales. Therefore, these experiments should reveal correlations between performance and some rationale(s), and—in the limit—provide us with suggestions about the functional relation supporting such correlations.

In this chapter, we focus on the second method. This research was done in collaboration with statistician Jan Beran. The musical material we con-

G. Mazzola, *Musical Performance*, Computational Music Science,
DOI 10.1007/978-3-642-11838-8_23, © Springer-Verlag Berlin Heidelberg 2011

sidered was Schumann's *Träumerei* op.15/7, Webern's *Variationen für Klavier* op.27/II, the *canon cancricans* from Bach's *Musikalisches Opfer* BWV 1079, and Schumann's *Kuriose Geschichte* op.15/2. We have calculated metrical, motivic, and hormonic weights for all of these compositions.

The main task was then to transform this data into a format that was adequate for statistical processing. Since we were focusing on agogics, which had been measured by Bruno Repp for 28 famous performances, our analytical weights were all "boiled down" to functions of onset only. Therefore, we have taken the average values at a given onset for melodic and harmonic weights.

We should add here that Repp's measurements can not be done with much more precision and also regarding parameters other than time. The software Melodyne [105] editor (figure 23.1) is capable of transforming audio data to note data and, after an unavoidable amount of editing, into MIDI data. Therefore, the performance research is open to a huge repertory of historical recordings.

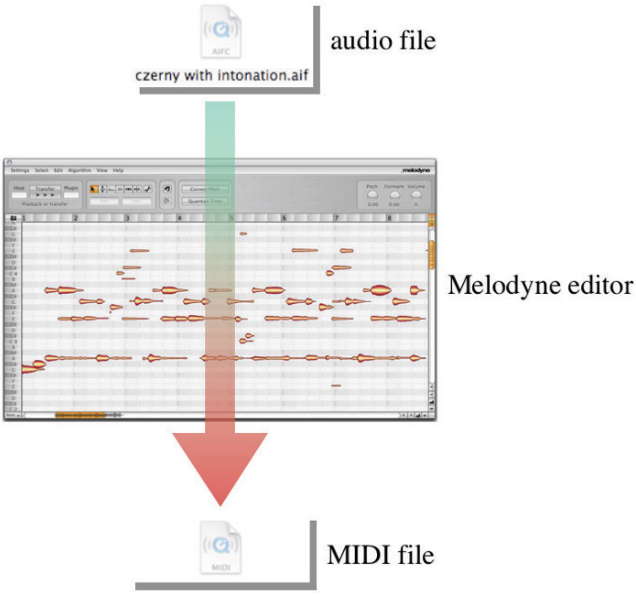

Fig. 23.1. The software Melodyne editor can transform polyphonic audio data into MIDI data and thereby opens research to performance analysis of historical recordings.

23.1 Hierarchical Decomposition of Weights

The statistically relevant decomposition of weights runs as follows: We start from a weight function $w(E)$ being defined for all onsets E, so it is a splined weight function, not just the discrete weight. The weight is not decomposed according to a Fourier procedure, because there is no reason to suppose that periodic weights should play any particular role in this context. Instead, we have chosen a hierarchical averaging procedure. Intuitively, this means that we start with a broad averaging of the weight, then deduct this from the weight and make a slightly less broad averaging, etc., thereby getting more and more local information represented on the finer averaging levels.

More precisely, we take a triangular support function

$$\begin{aligned} K(s) &= 1 - |s| \text{ for } s \in [-1, +1] \\ &= 0 \text{ else.} \end{aligned}$$

Given a sequence $(t_i)_{i=1,\ldots n}$ of times and a non-negative real number b, we next define the Naradaya-Watson kernel function by

$$K_b(t, t_i) = \frac{K(\frac{t-t_i}{b})}{\sum_{j=1}^{n} K(\frac{t-t_i}{b})}.$$

We then suppose given a time series of dimension k

$$(x_s(t_i))_{s=1,\ldots k}, i = 1, \ldots n.$$

The averaging formula then is this:

$$K_b x_s(t) = \sum_{i=1}^{n} K_b(t, t_i) x_s(t_i).$$

For $b = 0$, we have $K_0 x_s(t) = x_s(t)$.

The averaging process now works when we suppose that a decreasing sequence of bandwidths $b_1 > b_2 > \ldots b_m = 0$ is given. We first average according to b_1. This gives the new smoothed functions

$$x_{1,s} = K_{b_1} x_s.$$

We then proceed by induction. Suppose we have constructed smoothed functions $x_{1,s}, x_{2,s}, \ldots j-1$. Then we define the jth smoothed function by

$$x_{j,s} = K_{b_j}(x_s - \sum_{l=1}^{l=j-1} x_{l,s}).$$

In figure 23.2, we show the smoothing curves for a succession of bandwidths $8 > 4 > 2 > 1 > 0.5 > 0.1 > 0$ and *Träumerei.*

For our hierarchical smoothing process, we now start with the triangular support function

$$\begin{aligned} \hat{b}(s) &= 1 - |s|/b \text{ for } s \in [-b, +b] \\ &= 0 \text{ else.} \end{aligned}$$

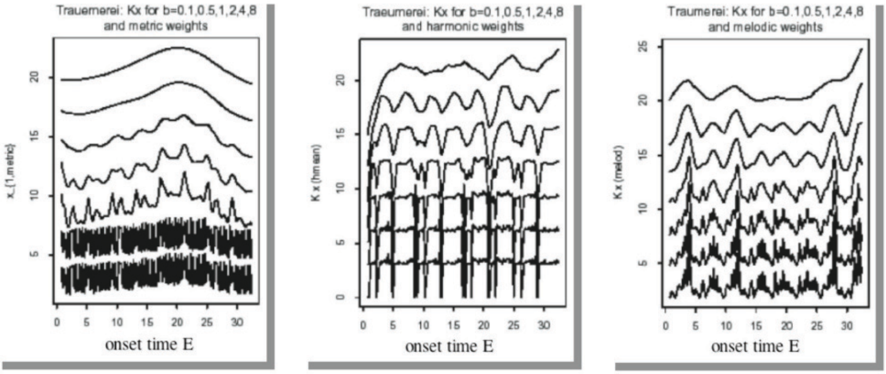

Fig. 23.2. Hierarchical smoothing curves for *Träumerei*—metric, harmonic, and melodic, from left to right—and bandwidths $8 > 4 > 2 > 1 > 0.5 > 0.1 > 0$.

and define the smoothed function for function f by

$$b \diamond f(E) = \int \hat{b}(t - E) \cdot f(t). \tag{23.1}$$

It averages f around E with weighted center E and *bandwidth* b. If this function is a weight, this means that the weight's analysis within the entire bandwidth neighborhood of a given onset is included instead of spiking the analysis to the singular onset. In the following process, this kernel smoothing process has been applied to a hierarchy of bandwidths, starting with $b = 4$ (= eight measures), then $b = 2$, then $b = 1$. The averaging process is taken to define successive remainder functions as follows:

$$f_1 = 4 \diamond f,\ f_2 = 2 \diamond (f - f_1),\ f_3 = 1 \diamond (f - f_1 - f_2),\ f_4 = f - f_1 - f_2 - f_3 \tag{23.2}$$

This means that the decomposition

$$x = x_1 + x_2 + x_3 + x_4 \tag{23.3}$$

for a smooth weight x defines a "spectrum" of that weight with respect to successively refined neighborhoods of its ambit.

Musically speaking, as already observed, this kernel smoothing process is completely natural. In fact, the kernel function alters the original time function $f(E)$ by a weighted integration of f-values in the kernel neighborhood of a given time E. This means that we now include the information about f from the neighboring times to make an analytical judgment. This latter is a well-known and common consideration in musical performance: The interpreter looks up a full neighborhood of a time point to derive what has to be played in that point. Moreover, the repeated application of the kernel smoothing process with increasingly narrowed neighborhoods is understood as a succession of a refinement in local analysis: First, the interpreter makes a coarse analysis over eight measures ($b = 4$), then he/she looks for the remainder $f - f_1$ and goes on with refined actions, if necessary.

This procedure is applied to the metric, melodic, and maximal and mean harmonic weights $x_{metric}, x_{melodic}, x_{hmax}, x_{hmean}$ and to their first and second derivatives $d_E x, d_E^2 x$. This gives the following list of a total of 48 *spectral analytical* functions:

$$\begin{array}{llll}
x_{metric,1} & x_{metric,2} & x_{metric,3} & x_{metric,4} \\
d_E x_{metric,1} & d_E x_{metric,2} & d_E x_{metric,3} & d_E x_{metric,4} \\
d_E^2 x_{metric,1} & d_E^2 x_{metric,2} & d_E^2 x_{metric,3} & d_E^2 x_{metric,4}
\end{array}$$

$$\begin{array}{llll}
x_{melodic,1} & x_{melodic,2} & x_{melodic,3} & x_{melodic,4} \\
d_E x_{melodic,1} & d_E x_{melodic,2} & d_E x_{melodic,3} & d_E x_{melodic,4} \\
d_E^2 x_{melodic,1} & d_E^2 x_{melodic,2} & d_E^2 x_{melodic,3} & d_E^2 x_{melodic,4}
\end{array}$$

$$\begin{array}{llll}
x_{hmax,1} & x_{hmax,2} & x_{hmax,3} & x_{hmax,4} \\
d_E x_{hmax,1} & d_E x_{hmax,2} & d_E x_{hmax,3} & d_E x_{hmax,4} \\
d_E^2 x_{hmax,1} & d_E^2 x_{hmax,2} & d_E^2 x_{hmax,3} & d_E^2 x_{hmax,4}
\end{array}$$

$$\begin{array}{llll}
x_{hmean,1} & x_{hmean,2} & x_{hmean,3} & x_{hmean,4} \\
d_E x_{hmean,1} & d_E x_{hmean,2} & d_E x_{hmean,3} & d_E x_{hmean,4} \\
d_E^2 x_{hmean,1} & d_E^2 x_{hmean,2} & d_E^2 x_{hmean,3} & d_E^2 x_{hmean,4}
\end{array}$$

For which *musical reasons* are these derivatives added to the analytical input data? The first derivatives measure the local change rate of analytical weights. Musically speaking, this is an expression of transitions from important to less important analytical weights (or vice versa), i.e., a transition from analytically meaningful points to less meaningful ones (or vice versa). This is crucial information to the interpreter: It means that he/she should change expressive shaping to communicate the ongoing structural drama. In the same vein, information about second derivatives is musically relevant because it lets the interpreter know that the ongoing structural drama is being inflected. Evidently, one could add higher derivatives, but we argue that an interpreter is already highly skilled if he/she can take care of all these functions and also observe different analytical aspects, from metrics to harmonics, simultaneously.

Besides these analytical input functions, we add three types of 'sight-reading' functions. They regard the following three primavista instances: ritardandi, suspensions, and fermatas. We omit these weights and refer to [84, chapter 44] for details. The entire spectral averaging procedure yields 58 functions of symbolic time E. Their vector, with all functions given a fixed order of coordinates, is denoted by $X(E) \in \mathbb{R}^{58}$.

Next, we look for a connection of this big analytical vector function to the tempo function found from Repp's analysis. We introduce this operator for $\omega \in \mathbb{R}^{58}$:

$$\Omega_\omega^X = (X, \omega),$$

the scalar product of ω and the analytical vector X. This means that for every onset E, we have $\Omega_\omega^X(E) = (X(E), \omega)$. Recapitulating the meaning of the analytical vector X, we are dealing with a second-order differential operator that we call a "Beran operator" since it was introduced by Jan Beran in [6].

On this basis, the central question of the following is whether tempo curves T of the *Träumerei* as they appear in the context measured by Repp in [111] may be approximated via Ω_ω^X by an appropriate choice of the shaping vector ω. The main result of this approach states that there is strong statistical evidence for the equation

$$\ln(T) = \Omega_\omega^X + C \tag{23.4}$$

for the given analytical vector X, a suitable shaping vector ω, and a constant C.

This means that the 58 coefficients of the shaping vector ω are random variables and that we prove a significant statistical correlation—in the mathematical form described by the Beran operator—between a certain subset of the analytical vector X and tempo as it is measured for the 28 performances by Repp.

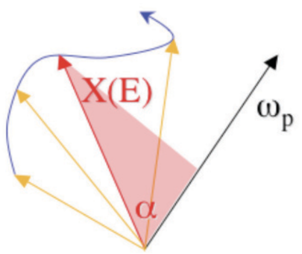

Fig. 23.3. The Beran operator uses the scalar product of the shaping vector ω and the analytical vector X at time E.

Observe that the formula in 23.4 uses the logarithm of tempo, a remarkable fact, which fits in the general fact that logarithms are important for cognitive processes, pitch and loudness being the classical cases. Moreover, taking the logarithm of tempo turns the set of all tempi into a real vector space: $\ln(T_1) + \ln(T_2) = \ln(T_1 T_2)$, and $\lambda \ln(T) = \ln(T^\lambda)$ are reasonable operations of tempo curves!

This being so, the hypothesis to be verified as a statistical statement is that for each p of those famous 28 artists playing *Träumerei*, the measurements of Repp enable a vector $\omega_p \in \mathbb{R}^{58}$ such that

$$\ln(T_p) = \Omega_{\omega_p}^X + C$$

is well approximated (figure 23.3).

This Beran operator formula is strongly supported by the present data set. Moreover, it can be shown that a small number of weights is already significant for the overall effect.

The main statistical conclusions from the analysis can be summarized as follows:

- There is a clear association between metric, melodic, and harmonic weights and the tempo.
- The exact relationship between the analytic weights and an individual tempo curve is very complex. However, a large part of the complexity can be covered by our model.
- Commonalities and diversities among tempo curves may be characterized by a relatively small number of curves. There is in principle no unique way of attributing features of the tempo to exactly one cause (harmonic, metric, or melodic analysis). Which curves need to be used depends partially on which of the three analyses (harmonic, metric, melodic) has 'priority.' However, there seems to be a small number of canonical curves that are essentially independent of the priorities and which determine a large part of the commonality and diversity among tempo curves. Natural clusters can be defined.
- There is a natural way of reducing an individual tempo curve to a series of simplified tempo curves containing an increasing number of features.

The results here are closely related to Repp's work [111]. Repp applied principal component analysis to the 28 tempo curves. One of his main results is that Cortot and Horowitz appear to represent two extreme types of performances. Thus, in a heuristic way, Repp suggested classifying the performances according to their factor loadings into a Cortot and a Horowitz cluster, respectively. Repp's Horowitz and Cortot clusters are confirmed.