Connection Science
Vol. 00, No. 00, January 2008, 1–21

RESEARCH ARTICLE

Analysing Musical Performance through Functional Data Analysis: Rhythmic Structure in Schumann's Träumerei

Josué Almansa[a*] and Pedro Delicado[b]

[a] *Institut Municipal d'Investigació Mèdica, Doctor Aiguader 88, 08003, Barcelona, Spain*; [b] *Departament d'Estadítica i Investigació Operativa, Universitat Politècnica de Catalunya, Jordi Girona 1-3, 08034, Barcelona, Spain*

(*Received 00 Month 200x; final version received 00 Month 200x*)

Functional Data Analysis is a relatively new branch of Statistics devoted to describing and modelling data that are complete functions. Many relevant aspects of musical performance and perception can be understood and quantified as dynamic processes evolving as functions of time. In this paper, we show that Functional Data Analysis is a statistical methodology well-suited for research into the field of quantitative musical performance analysis. To demonstrate this suitability, we consider tempo data for 28 performances of Schumann's Träumerei and analyse them by means of functional principal component analysis (one of the most powerful descriptive tools included in Functional Data Analysis). Specifically, we investigate the commonalities and differences between different performances regarding (expressive) timing, and we cluster similar performances together. We conclude that musical data considered as functional data reveals performance structures that otherwise might go unnoticed.

Keywords: Commonalities; cluster analysis; diversity; local polynomial smoothing; principal component analysis (PCA); tempo; timing.

1. Introduction

Music can be defined as the art of arranging sequences of sounds in time. Dynamic character is an inherent feature of music: performance as well as music perception are time-dependent (or time evolving) abstract processes. Vines et al. (2005) mention fluctuations in tempo or in loudness in musical performance, and the dynamic reactions of human participants to musical stimuli as examples of time dependent processes in music. As these authors point out, *Functional Data Analysis* (FDA) is a particularly well-suited tool-box for analysing the aspects of music that can be represented as continuous measures smoothly evolving in time. FDA (see the monograph Ramsay and Silverman 2005 and other references introduced in Section 2) is the generic name for the statistical techniques recently developed to describe and model situations where a complete function is observed for each individual constituting a random sample. In their work, Vines et al. (2005) focus attention on participants' judgements of tension continuously registered while they were presented with a musical performance.

*Corresponding author. Email: jalmansa@imim.es

ISSN: 0954-0091 print/ISSN 1360-0494 online

DOI: 10.1080/09540090xxxxxxxxxxxx
http://www.informaworld.com

In this paper, we introduce FDA as a statistical framework for research in the field of quantitative musical performance analysis (see Repp 1992 or Gabrielsson 2003 that use the terms *objective* and *empirical* performance analysis, respectively). Our main point is that a FDA approach simplifies interpretation of the results, compared with a multivariate approach, because modelling musical aspects as continuous functions of time (or of score position) automatically takes into account that close moments in a performance are statistically dependent.

There are two basic aspects of music performance analysis: Normative (commonality) and individual (diversity) performance characteristics. The normative characteristics are specified in the score (for example, a *ritardando* or *fermata*) and are common for every performer, but they also include general (possibly unconscious) unwritten rules to which all or most performers adhere. In this sense, Widmer et al. (2003) provide examples showing that inductive rule-learning algorithms (coming from Artificial Intelligence) are able to detect and predict general accepted norms of musical performance. Beyond the normative aspects, the musician can play a score in his or her personal style; even the normative score indications can be performed in slightly different ways, and this is what makes music performance an art.

There are various performance variables that can be measured, such as the rhythmic structure (tempo and timing), intensity (loudness) and articulation (the way in which contiguous notes are connected), among others (Repp 1992, Gabrielsson 2003). The music performance characteristics can be local or global (Repp 1992). Global characteristics affect the performance of the whole piece and can be detected when the piece is listened to all the way through; for example, the global tempo or the final *ritardando* (it is necessary listen to the whole piece to perceive that the notes of the final *ritardando* are played slower than the rest of the piece). Local characteristics affect only a small part of the score, such as the *ritardando* at the end of a phrase, or a more local one: the *rubato* in the *arpeggi* (described below).

The central goal of this paper is to show that FDA is a useful methodology to statistically analysing music performance. We focus on rhythmic structure because of data availability (see Section 4), but other musical dynamical parameters can be studied in a similar way. The main analytical tool we use is *functional principal component analysis* (FPCA, see Section 2.1 for more details). This technique is well-suited for describing diversity in performance characteristics (principal functions identify the strongest and most important modes of variation of individuals around a common mean; see Ramsay and Silverman 2005, p. 149) and this is in fact our aim: to investigate the commonalities and differences between different performances regarding (expressive) timing.

The raw data (Repp 1992) we use as the point of departure for our analysis are time measurements at note-level, which make the study of local performance characteristics possible. At the same time, the FDA framework allows us to represent every complete performance as a function (a single datum in FDA). Therefore our approach is also able to analyse global characteristics.

The paper is organised as follows. Section 2 introduces FDA and FPCA. Then Section 3 presents the Schumann's piece Träumerei, which we use as a case study throughout the paper. Repp's tempo data on Schumann's Träumerei (Repp 1992) are described in Section 4, as well as how they are transformed into functional data by smoothing techniques. The application of FPCA in performance analysis for this piece is detailed in Section 5, which also includes a cluster analysis resulting from FPCA as a by-product. The paper ends with some conclusions summarised in Section 6.

2. Functional Data Analysis (FDA)

Observing and saving complete functions as results of random experiments is nowadays possible by the development of real-time measurement instruments and data storage resources. For instance, continuous-time clinical monitoring is a common practise today. Ramsay and Silverman (2005) express this by saying that random functions are in this cases the *statistical atoms*. Functional Data Analysis (FDA) deals with the statistical description and modelling of samples of random functions. Functional versions for a wide range of statistical tools (ranging from exploratory and descriptive data analysis to linear models and multivariate techniques) have recently been developed. Other techniques are specific to FDA, because they exploit the functional nature of this kind of data: *principal differential analysis* is a type of principal component analysis carried out on the derivatives of the observed functions; *registration* is a pre-process step where a change of variable is carried out on each observed function in order to make them as similar as possible. See Ramsay and Silverman (2005) for a general perspective on FDA, and Ferraty and Vieu (2006) for a non-parametric approach. Ramsay and Silverman (2002) present applications of FDA to a wide range of problems and disciplines. Special issues recently dedicated to this topic by several journals (Davidian et al. 2004, González-Manteiga and Vieu 2007, Valderrama 2007) bear witness to the interest for this topic in the Statistics community.

It is well worthwhile noting that random functions can also be obtained from standard random samples by the application of non-parametric curve estimation methods. For instance, Kneip and Utikal (2001) used non-parametric density estimation methods to obtain annual income densities, which enabled them to study the temporal evolution of income density functions in United Kingdom from 1968 to 1988. The most frequent situation, however, is that of having observations densely sampled over time, space or other continuous parameter spaces. In these situations, interpolation techniques (if the underlying sampled functions are smooth and there is no sampling noise) or smoothing methods (in other cases) allow us to transform the discrete observations into continuous functional objects.

A non-technical introduction to FDA can be found in Levitin et al. (2007), which is illustrated with the musical application presented in Vines et al. (2005), and which as far as we know is the only existing work applying FDA methodology to the analysis of musical data.

2.1. *Functional principal component analysis (FPCA)*

Ferraty and Vieu (2006) define a *functional variable* as a random variable $\boldsymbol{f}$ taking values in an infinite functional space, usually

$$L_2(I) = \{f : I \to \mathbf{R}, \text{ such that } \int_I f(s)^2 ds < \infty\},$$

where $I = [a, b] \subseteq \mathbf{R}$. An observation f of $\boldsymbol{f}$ is called *functional data*. A *functional data set* $f_1, \ldots, f_n$ is the observation of n independent functional variables $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ identically distributed as $\boldsymbol{f}$. In this context, a version of the principal component analysis (PCA, possibly the most popular descriptive statistical technique for multivariate data) has been developed: this is *functional principal component analysis (FPCA)*.

The objective of FPCA can be stated as follows. Given a functional random sample with mean function $\bar{f}(s) = (1/n)\sum_{i=1}^{n} f_i(s)$, for all $s \in I$, we look for functions $g_1, \dots, g_q$ (*principal functions*) in $L_2(I)$ and real numbers ψ_{ij}, $i = 1, \dots, n$, $j = 1, \dots, q$, such that

$$\sum_{i=1}^{n} \int_I \left((f_i(s) - \bar{f}(s)) - \sum_{j=1}^{q} \psi_{ij} g_j(s) \right)^2 ds$$

is minimum. Moreover, the functions $g_1, \dots, g_q$ are required to be orthonormal: $\int_I g_i(s) g_j(s) ds$ is equal to 0 if $i \neq j$ and equal to 1 if $i = j$. In other words, we are looking for a representation of functional data in a q-dimensional space (the space spanned by the functions $g_1(\cdot), \dots, g_q(\cdot)$):

$$f_i(s) \approx \bar{f}(s) + \sum_{j=1}^{q} \psi_{ij} g_j(s), \; s \in I, \; i = 1 \dots n.$$

It can be proved that the principal functions can be estimated, under certain assumptions, as the *eigenfunctions* of the sampling covariance operator:

$$\int_I \Gamma_n(s,t) g_j(t) dt = \lambda_j g_j(s), \text{for all } s \in I, \tag{1}$$

where

$$\Gamma_n(s,t) = \frac{1}{n} \sum_{i=1}^{n} (f_i(s) - \bar{f}(s))(f_i(t) - \bar{f}(t)),$$

and that

$$\psi_{ij} = \int_a^b (f_i(s) - \bar{f}(s)) g_j(s) ds, \; i = 1, \dots, n, \; j = 1, \dots, q.$$

Coefficient ψ_{ij} is the score of the observation i on the j-th principal function. The numbers $\lambda_1, \dots, \lambda_q$, known as *eigenvalues*, are sorted in decreasing order and are equal to a common constant times the proportion of total variability explained by the corresponding principal functions.

A way to interpret the meaning of the principal functions is that they represent the main variation modes of the observed functions around the global mean function. The mean function $\bar{f}$ represents what is common to all the data (commonality, if we are dealing with music performance), the centred functions $(f_i - \bar{f})$ account for individual differences (diversity) and the principal functions summarise what is common in the way individual are diverse.

There are different approaches to solving equation (1) in practise. A solution by expanding sample-paths in terms of B-spline functions was proposed in Aguilera et al. (1996). Ramsay and Silverman (2005) also propose the expression of observed functions as linear combinations of B-splines functions forming an approximate base of $L_2(I)$. In this way, equation (1) can be re-expressed as a matrix equation to be solved by standard methods.

A different solution is suggested by Kneip and Utikal (2001). Once the original functions have been properly smoothed (if required), the centred functions are evaluated in a fine grid of evenly spaced points of I: $s_1 = a, \dots, s_M = b$. Let F be the $n \times M$ the resulting data matrix. It can be proved that for large values of M, the solutions of (1) can be derived from eigenvalues and eigenvectors of $M \cdot M^T$

or $M^T \cdot M$, the last one having the advantage of having dimension $n \times n$, which is very convenient given that usually $n << M$. We follow this approach in Section 5.

3. The piece: Schumann's Träumerei

Schumann's Träumerei op.15/7 is the seventh piece from the album Kinderszenen. Composed by Robert Schumann in 1838. It was dedicated to Clara Wieck who later became his wife. The score is shown in Figure 1. The album, on which Träumerei is the best known piece, is representative of the Romantic period, characterised by its musical expressiveness, subjectivity and psychological nuances of a state of mind. Träumerei consists of three phrases of 8-bar length in a ternary form: A B A′. The first one (A) should be repeated by score indications. The last one (A′) is very similar to the first (A) but the last bars of the phrase are changed to give a conclusive sense to the whole piece. Each phrase consists of two periods of four bars, and all of which have a very similar rhythmic structure. Musicological analyses of the Träumerei can be found in Reti (1951), Brendel (1981) and Traub (1981), for instance.

"Tension and release" is a term frequently used in the analysis of music to describe how a piece retains the interest of the listener (see, for instance, Meyer 1956, Stein 1962 or Huron 2006). In the Romantic style, the clear emphasis of movement from moments of tension and release (and vice-versa) plays an important role in the performance. Träumerei begins in a relaxed state and the moment of maximum tension occurs in phrase B (B-flat note in the upper voice in bar 14); in then relaxes again until the end of the piece is reached. The same relax-tension-relax pattern is reproduced with less emphasis within each period of each phrase. Some other normative aspects in the performance of this piece are:

(1) There is no indication in the score about the repetition of the phrase A, so it is expected to be performed twice with similar agogics.
(2) At the end of each phrase, there is a *ritardando* (intentional slowing of tempo), and the final *ritardando* is the longest one (it involves more notes and greater slowing tempo). The *ritardando* at the end of phrase B is located just after the moment of maximum tension, and moves clearly to a relaxed phrase A′. Thus, one expects to find differences in the performance of *ritardando* in phrases A and B.
(3) In bar number 22, there is a *fermata* indicating that this note has to be played longer.
(4) Apart from the explicit instructions on the score, some other performance techniques are implicit and commonly accepted by musicians, such as the *rubato*: a specific *accelerando-ritardando* rhythmic shape. In this score a *rubato* is generally accepted, involving the six *arpeggiated* notes: five quavers that ascend in increasingly longer pitch steps to a final note. This melodic gesture is located in the last three quavers of the second bar and the first three quavers in the third bar of every period.

Träumerei has a very regular (and simple) musical form, and its main interest lies in its ability to allow performers to exhibit expressiveness. Therefore, it provides a high degree of freedom in individual performance beyond the normative aspect.

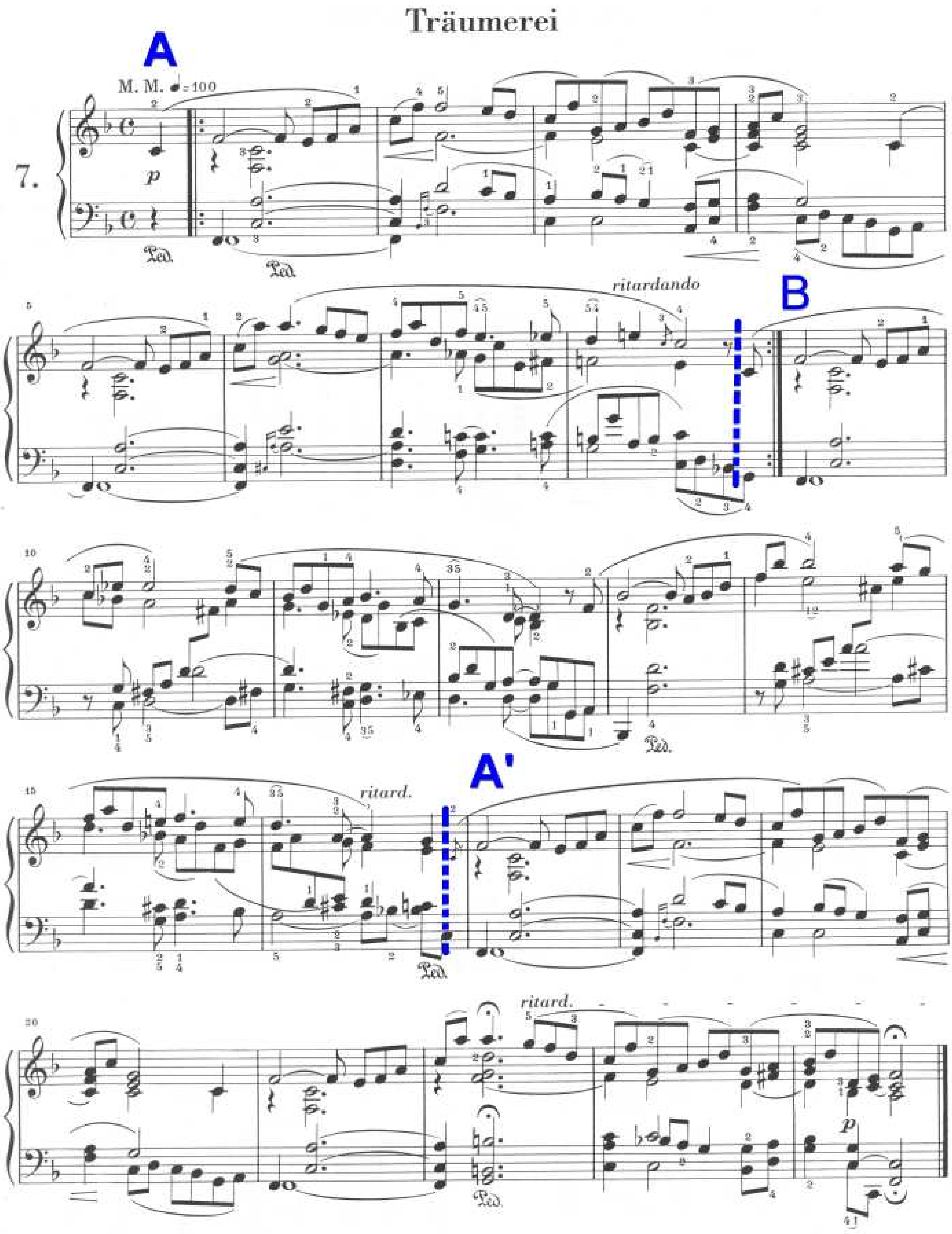

Figure 1. Schumann's Träumerei op. 15/7. Reprinted here with the permission of G. Henle Publishers, http://www.henle.de.

4. Repp's tempo data on Schumann's Träumerei

The analysed data were kindly provided by Bruno H. Repp (see Repp 1992 for a detailed description). They comprised the duration of each note of the melody in milliseconds. Grace notes were omitted. Measurement process consisted in calculating

Table 1. List of performers.

Code	Artist (year of recording)	Code	Artist (year of recording)
ARG	Martha Argerich (<1983)	HO1	Vladimir Horowitz (1947)
ARR	Claudio Arrau (1974)	HO2	Vladimir Horowitz (<1963)
ASH	Vladimir Ashkenazy (1987)	HO3	Vladimir Horowitz (1965)
BRE	Alfred Brendel (<1980)	KAT	Cyprien Katsaris (1980)
BUN	Stanislav Bunin (1988)	KLI	Walter Klien (?)
CAP	Sylvia Capova (<1987)	KRU	André Krust (1960)
CO1	Alfred Cortot (1935)	KUB	Antonin Kubalek (1968)
CO2	Alfred Cortot (1947)	MOI	Benno Moiseiwitsch (1950)
CO3	Alfred Cortot (1953)	NEY	Elly Ney (1935)
CUR	Clifford Curzon (1955)	NOV	Guiomar Novaes (<1954)
DAV	Fanny Davies (1929)	ORT	Cristina Ortiz (<1988)
DEM	Jfirg Demus (1960)	SCH	Artur Schnabel (1947)
ESC	Christoph Eschenbach (<1966)	SHE	Howard Shelley (<1990)
GIA	Reine Gianoli (1974)	ZAK	Yakov Zak (1960)

the time difference between two consecutive notes onset (IOI: interonset intervals) using a waveform editing program. Some measurement error was assumed. Notes were divided into quaver length (eighth-notes in American terminology), so IOIs longer than a nominal quaver in the score were divided into IOIs of equal duration. Given that measures were done through IOIs, the last note had no corresponding measure. A complete performance thus yielded 253 quavers for each performance.

This measurement process resulted in a data set that includes 28 different interpretations of the Schumann's Träumerei as performed by 24 prominent pianists (Table 1). Two of these artists (Cortot and Horowitz) are each represented by three different recordings each. There are two interpretations (F. Davies and A. Krust) in whose recordings phrase A was not repeated. Therefore in the data file, the values for the performance of phrase A are repeated twice.

Data therefore consist of a 28×253 matrix, where each row corresponds to a performance and the n-th column gives the duration of the n-th quaver of the score in millisecond.

Previous studies have analysed this data in different ways. Repp (1992) studied microstructure performances within the piece, and analysed commonality and diversity using PCA and cluster analysis on multivariate data consisting of short fragments of raw data previously described. Repp (1995) added 10 piano student performances (three performances each) and compared student performances with professional ones. Repp (1996) added information about dynamics (pitch intensity), focusing on student performances recorded in 1995.

All Repp's papers analysed each note as one different variable, ignoring the fact that notes that are closer in the score are more statistically related than those that are farther away. So a great deal of effort in standard multivariate techniques is spent on re-discovering this dependence structure. An automatic way of taking into account that close notes are dependent is by considering performances as continuous functions of the position score (as we do in this paper).

Beran and Mazzola (1999a) employed the RUBATO software (Mazzola and Zahorka 1994) for analysing a score according to explicit rules derived from general music theory and practise, and for transforming the results into numerical weights: each note in the score can be assigned a *weight* that measures its metric, harmonic or melodic *importance* respectively. The purpose of the paper was to introduce a statistical approach to the analysis of metric, melodic and harmonic structures of a score and their influence on musical performance. The measures of metric, harmony and melody, as well as the regression coefficients estimated by Beran and Mazzola, are assumed to be vectors of high dimension. In fact they could also be modelled as functions and, in this case, FDA techniques (functional regression with scalar response, for instance) could be used. We believe that the functional approach would make the interpretation of the results easier.

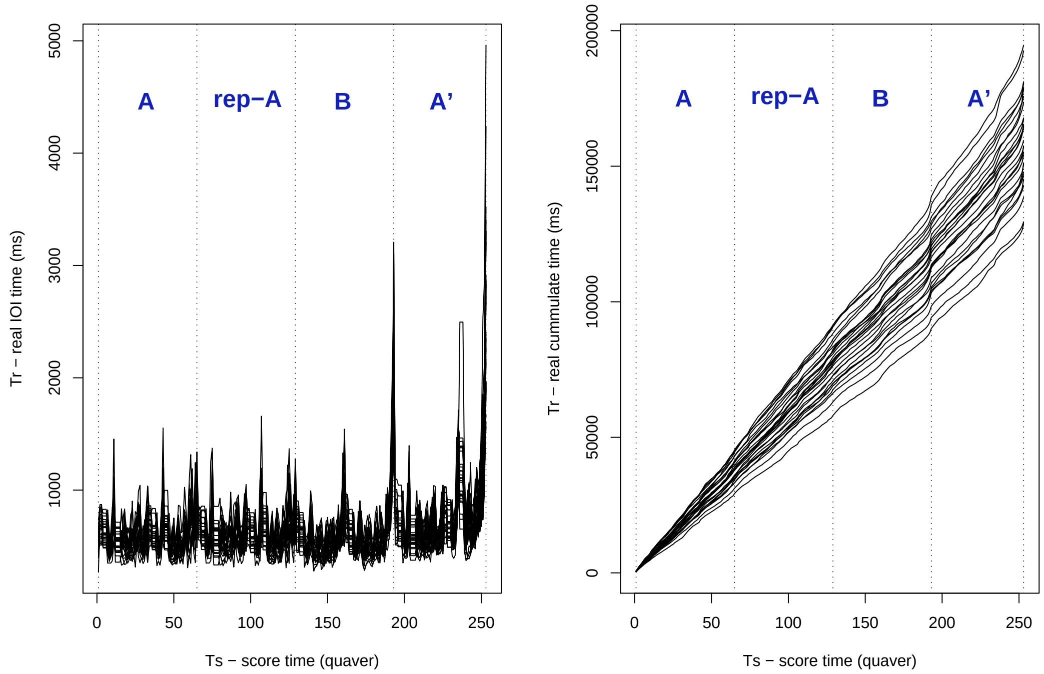

Figure 2. Original and cumulated data for the 28 performances of Träumerei.

Beran and Mazzola (1999b) developed the Hierarchical Smoothing Model, as an attempt to better understanding the relationship between the symbolic structure of a music score and its performance. According to the observation that musical structure typically consists of a hierarchy of global and local structures, in the Hierarchical Smoothing Model data are decomposed in different smoothed functions, each of them with their own bandwidth, so each smoothed function describes different aspects of the original data information. This is in contrast with usual nonparametric smoothness methods, where the optimum bandwidth is looked for. In the final remarks, Beran and Mazzola (1999b) cite the 1997 version of Ramsay and Silverman (2005), as a possible way to complement their approach.

Vines et al. (2005) mention the possibility of analysing music performance with FDA techniques. For music perception analysis, they introduce some concepts (such as velocity and acceleration functions) that are related with other functions used by ourselves (see below for our definition of slowness and deceleration functions). In a final note, they point out the possibility of carrying out their type of analysis with the Träumerei Repp's data. The degree thesis of Almansa (2005) made the first attempt to do this, independently of Vines et al. (2005).

4.1. *Creating functional data*

The original data, forming a 28×253 data matrix, have no functional form. They just give the duration of each quaver of the score for each piano performance. The way we transform these data into functional data is as follows. For each performance we compute a function giving the elapsed time t (in millisecond units) of the performance from the beginning up to any given position s (in quaver units) in the score. The estimation of this function is done simply by cumulating the original performance times (Figure 2). Note that this data now has a continuous sense; for any non-integer $s \in [1, 253]$ it is possible to estimate (by linear interpolation, for instance) the corresponding real time t. We can denote this function by $t(s)$. These functions are very similar throughout all the performances, and any peculiarities among the pianists are hardly detected.

The inverse of these cumulative functions (namely $s(t)$) are *position* functions: given a real time t, they return to what position of the score s the pianist is playing at t. By applying basic notions of physics, we can compute the *velocity* ($v(t)$) and *acceleration* ($a(t)$) functions from the position function (first and second derivatives: $v(t) = s'(t)$, $a(t) = v'(t) = s''(t)$). These new functions (velocity and acceleration) are more meaningful than position, and provide better discrimination of the performances (Vines et al. 2005).

Our option is to work with the cumulative functions $t(s)$ as the raw functional data. The main advantage of working with $t(s)$ is that this way all the observed functions ($t_i(s)$, $i = 1, \ldots, n$) have thereby a common support ($[1, 253]$). This is not the case when using position functions $s(t)$, because different performances have different durations. Having a common support is very convenient for comparing different functions and for relating their values directly to the score. We call the function $t(s)$ the *elapsed-time*; its first derivative $w(s) = t'(s)$ is *slowness* and its second derivative $d(s) = w'(s) = t''(s)$ is *deceleration*. Taking into account that $s(t)$ and $t(s)$ are inverse functions ($s(t(s)) = s$ and that $t(s(t)) = t$), it is easy to prove the following relations,

$$w(t) = \frac{1}{v(s(t))},\ d(t) = -\frac{a(s(t))}{v(s(t))},$$

which can help to interpret results expressed in terms of slowness and deceleration. For instance, higher values in the slowness function around a score position s mean that the note placed in s takes more time to be played and consequently its real velocity is lower. Although slowness and deceleration functions are both estimated, in the analyses presented below slowness functions are mainly used.

4.1.1. Non-parametric adequacy

Smoothed elapsed-time functions and their first and second derivatives are estimated from the elapsed-time raw functions by a non-parametric regression method. The choice of non-parametric estimation methodology has several advantages. First, the performance of a musical piece will not be well fitted by a function expressed in a closed parametric form, because within a single performance there is so much variability and tendency changes that it cannot be explained through a specific parametric expression. Secondly, the adjustment of a non-parametric model for smoothing is preferable to interpolation, because observed values are measured with error. In addition, a musician will never play the same piece twice in the same way; the trend is to do it within the same style. So we have a sample of how a pianist interprets a piece, and it makes sense to make an inference about it rather than taking it as an exact value. Thirdly, what is played in music at any particular time is strongly determined by what is played before as well as what is going to be played later: the execution of a note is determined by its musical context. (Vines et al. 2005 talk about *"the effect of temporal context in music —what has been played before and what is about to be played"*) Therefore, it makes perfect sense to estimate the value of a note by using the values of the notes next to it.

At this point it is worth citing Langner and Goebl (2003), in which a visualising technique for expressive performance is introduced. These authors extract information on tempo and loudness (as functions of time) from MIDI instruments or audio recordings; then they smooth the data by computing local means (which is equivalent to the Nadaraja-Watson non-parametric regression; see Fan and Gijbels 1996), and finally display the smoothed data in a two-dimensional space (loudness against tempo) as animation over time: a red dot moves in synchrony with the music, leaving behind it a trajectory that vanishes over time. These ideas have been implemented in a visualisation program called PERFORMANCE WORM (Dixon et al. 2002). Observe that the smoothed loudness and tempo functions (depending on time) are an example of a two-dimensional functional data, and that FDA techniques could enhance the visualisation capabilities of performance worms for analysing musical performance.

4.1.2. Function estimation

Smoothed elapsed-time, slowness and deceleration functions are estimated through *local polynomial regression* (see, e.g., Fan and Gijbels 1996). The general idea is to fit a polynomial regression of degree p to the data (s_j, t_{ij}), $j = 1, \dots, 253$, (corresponding to the i-th performance) locally around a specific point $s \in [1, 253]$, giving higher weight to nearby s points and lower weight to more distant ones. Based on this estimated polynomial function a response value is predicted for s. By repeating this process for a dense grid of points s in $[1, 253]$, the estimated function $\hat{t}_i(s)$ is obtained. The weights of neighbouring observations s_j are chosen to be proportional to $K((s - s_j)/h)$, where K is a positive symmetric function non-increasing on $[0, \infty)$, known as the *kernel*, and $h > 0$ is a smoothing parameter (larger values of h correspond to smoother estimated functions) called the *bandwidth*. The bandwidth choice is a crucial step in all non-parametric smoothing techniques. Here we choose the bandwidth by following the *rule of thumb* proposed by Fan and Gijbels (1996) in order to optimise the estimation of derivative functions. This rule starts with the asymptotically optimal constant bandwidth expression, which contains some unknown quantities that must be estimated from the data. In particular, a polynomial of degree $p + 3$ is fitted globally to (s_j, t_{ij}) in order to obtain a pilot estimation of the quantities, depending on the unknown function $t_i(s)$ that we are estimating.

We use $p = 2$ for the estimation of the slowness functions and $p = 3$ for deceleration. The kernel is Gaussian. The function `locpoly` of R (R Development Core Team 2008) is used. The bandwidths obtained from the rule of thumbs for the 28 slowness functions go from 3.595 to 6.051 quavers (1st Quartile: 4.224, Median: 4.509, 3rd Quartile: 4.820). For deceleration functions, the bandwidth values go from 2.650 to 4.065 quavers (1st Quartile: 3.289, Median: 3.525, 3rd Quartile: 3.720). As pointed out by Fan and Gijbels (1996), the rule of thumb for choosing bandwidths gives an approximate idea about how large the amount of smoothness should be, but in certain cases this approximate selection might suffice. This is what occurs in our case, because the estimated functions obtained prove extremely useful for analysing the different performances by functional principal components.

As an example, Figure 3 shows the original data and the estimated slowness function for performers DAV and ESC. Figure 4 shows the mean (accounting for performance commonality) and the standard deviation of the 28 estimated functions. At the end of each period there is a slowing in the tempo performance, which is more evident in phrase B and at the end of the piece. The largest variability remains in the final part of the piece. The *fermata* and the end of phrase B also have high variability among the 28 performances.

One disadvantage of the non-parametric regression is the estimation in the boundaries of the interval $I = [1, 253]$ where functions are defined. Because of this effect the end of every performance is overestimated (the functions here are not so smooth as in the rest of interval). As the components of a FPCA analysis are orthogonal, this problem will only be reflected in the overestimation of the

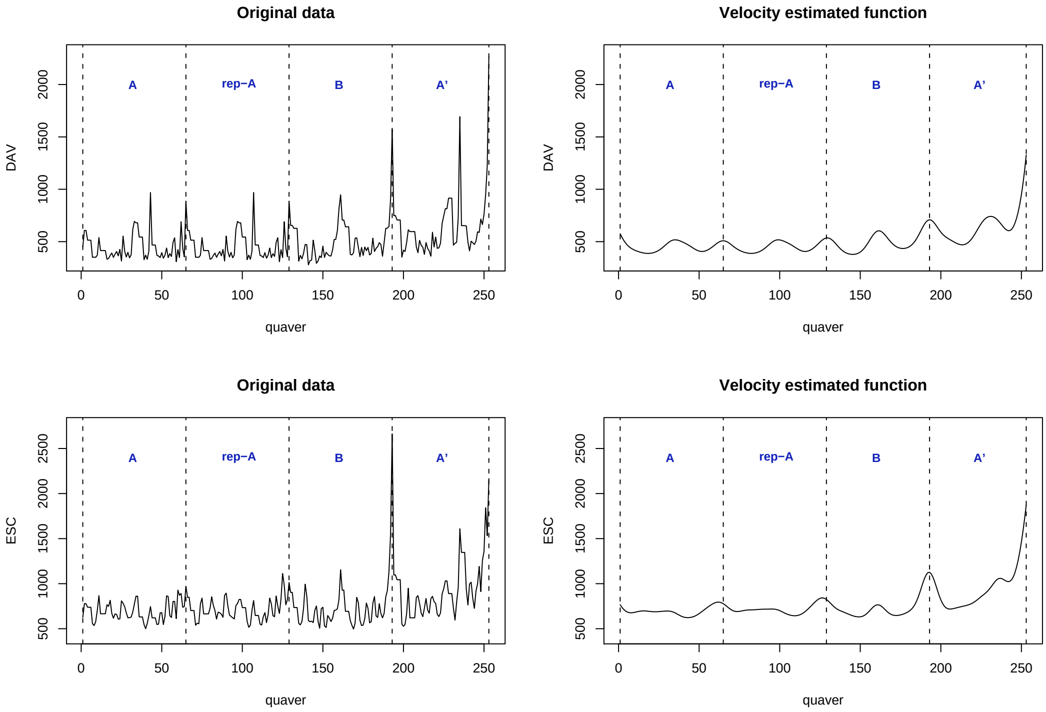

Figure 3. Original and estimated slowness for two performers (DAV and ESC).

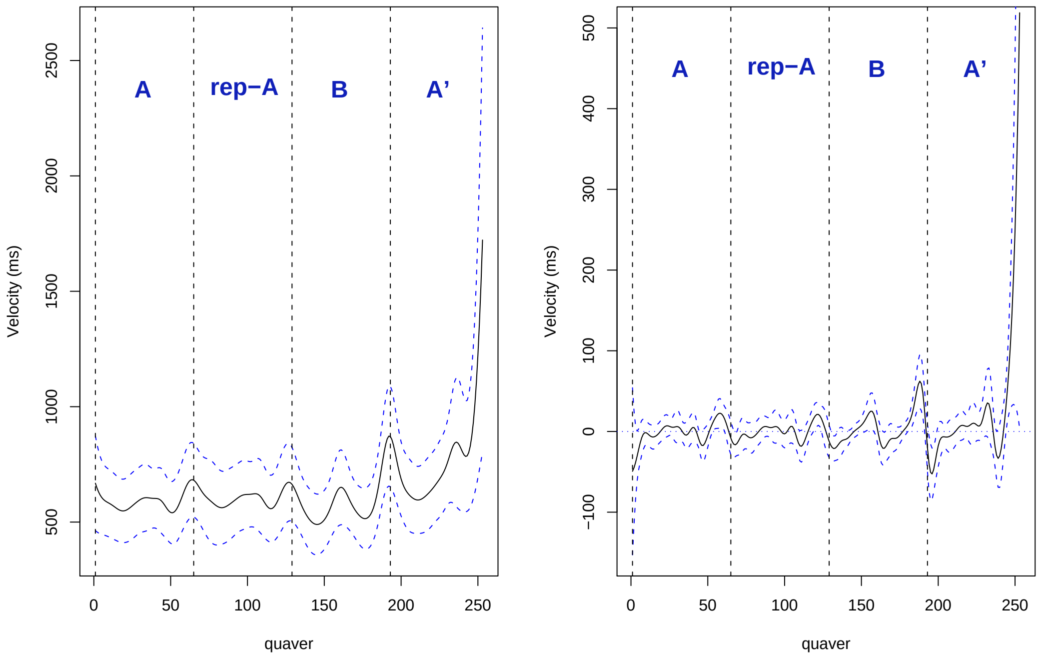

Figure 4. Mean slowness and deceleration functions (continuous lines) with bands (dashed lines) at +/- 2 point-wise standard deviations.

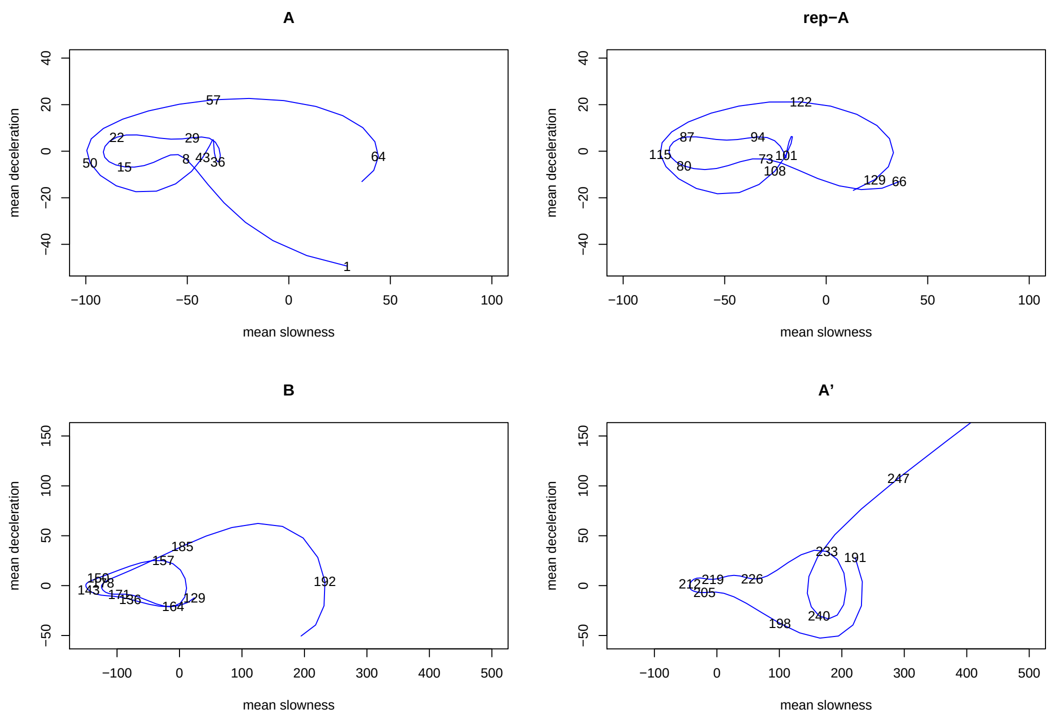

Figure 5. Phase-plane plot. Numbers indicate the quaver order in the score.

percentage of variability explained by one of the factors, without interfering in the interpretation of other components.

The phase-plane plot, as explained in Vines et al. (2005) and in Ramsay and Silverman (2002), shows the second derivative function against the first derivative. Pure oscillatory function yields to a circular phase-plane plot. In Figure 5 the mean deceleration function against the mean slowness function is shown, as a complementary way of describing performance commonality. The mean slowness function is centred in order to obtain zero-mean. The plot is divided by phrases to facilitate the interpretation. Phrases A and rep-A are very similar, only significantly differing at the beginning of the phrase, where phrase A shows considerable changes in the deceleration function, but this is produced by the boundary estimation bias. So we can consider that the global mean performance of phrase A and its repetition are equally performed. The first and second periods of phrase A follow the same flat-shape structure, (an oscillatory movement with little change in deceleration), but in the second period there is a greater variation in velocity. At the end of the first period and in the *arpeggi* of the second period (quavers 32 to 42 for phrase A, and 96 to 106 for rep-A) there is a vertical ellipse shape showing rapid changes in the deceleration function. This is clearly a *rubato*. Phrase B has a different structure compared with A, both periods having a very similar oscillatory shape where the end of both periods have remarkably slower timings. The end of the second period (end of phrase B) is remarkably slower than the end of the first period. The first period of phrase A' has a shape similar to first period of phrase B, but with slower velocity, which releases the tension created in phrase B. Around the *fermata* (bar with notes 234 to 242) there is a rapid change in in the deceleration function, and then the piece ends with an exaggerated deceleration timing.

5. Performance analysis by FPCA

Non-Standardised and standardised FPCA are conducted on the slowness function. Both methods aim to summarise the maximum amount of data variability in a small number of components. The difference between non-standardised and standardised FPCA is that, in the first case, the functions are analysed in their original measurement units (the covariance operator $\Gamma_n(s,t)$, defined in Section 2.1, is used), whereas in the second case all notes are forced to have the same variability across performers (unit variance, because the correlation operator $\rho_n(s,t) = \Gamma_n(s,t)/\sqrt{\Gamma_n(s,s)\Gamma_n(t,t)}$ is used). Therefore, in the non-standardised analysis, the larger variability corresponds to notes that are usually played slowly by all pianists (*fermata*, *ritardandos*, etc.), allowing for larger differences among performers. Typically, this information will be kept by the first component. When data are standardised, slow notes are no longer more important than others. Then the first principal component captures differences among performers that are relegated to second or posterior components in the non-standardised analysis.

The position of an individual in a principal function is the scalar product between its original functional data and the corresponding eigenfunction, so score-positions where an eigenfunction takes large values correspond to notes that have great importance in this principal function. Another graphical way of interpreting the principal functions is to sum and subtract from the grand mean the eigenfunction multiplied by an appropriate constant. This gives us an idea of how the performance pattern differs from the mean for performances having significant positive or negative values in the principal functions (looking at score parts where the shifted mean function is above or below the original mean function and where the maxim and minim values are found). Principal functions show several independent (orthogonal) performance patterns in a decreasing order of importance. When the first principal function has constant sign, it can then be used as a measure of global *size* (whatever that means in the specific context; see Mardia et al. 1979). The following ones (taking positive and negative signs) explain local *shape* characteristics.

5.1. *Non-standardised FPCA*

The first component (Figures 6 and 8) is a *size component* (it is always negative) showing that the main variation characteristic of the piece performance is the global tempo. It explains 60.3% of the whole variability. The performances with greatest negative value in the first principal function (see Figure 8) are ESC, CAT and BUN, which means that these are the slowest performances. On the right hand side of Figure 8 we find the fastest ones, which correspond to CO2 and DAV.

The second component (20.02% of total variability; see Figures 7 and 8) shows the contrast between the global tempo and the final tempo. This means that this component separates relatively fast performances that have a slow conclusion to the piece (the final *ritardando*, after the *fermata*) from others with opposite properties. This makes sense because a fast tempo produces more tension and requires more time to produce a relaxed sensation at the end of the piece. The performances that provide a greater contrast between the global tempo and the tempo at the end of the piece (those having large negative value in the second component; see the the vertical axis of Figure 8) are ORT and BUN, and the most homogeneous ones in the final tempo are ARG, CAP, NEY, KLI and ESC (see the top of Figure 8). So, we can describe BUN as a lower tempo performance with great tempo-contrast at the end of the piece (even lower), because it is placed in the bottom left-hand corner of Figure 8.

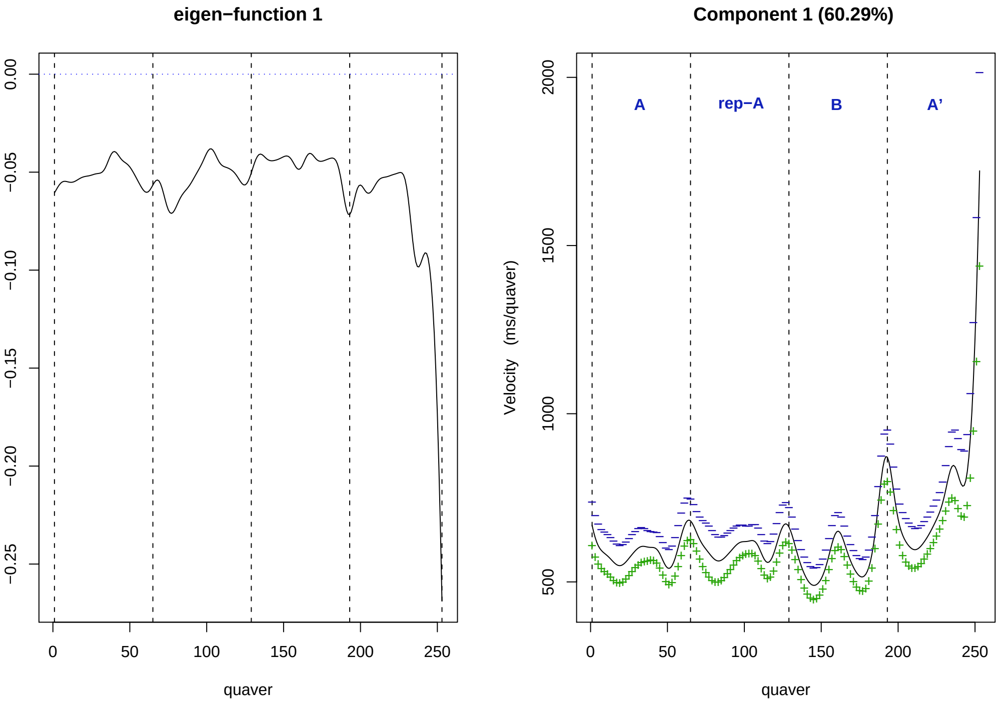

Figure 6. *Left:* First principal function (non-standardised FPCA). *Right:* Global mean function +/- a multiple of the first principal function.

The third component (5.00% of total variability; see Figure 7) shows a pattern where faster performance of phrases A and B contrast with slower tempo at the end of phrase B and the *fermata* (but not in the final *ritardando*). Performers who accentuate this contrast are ASH, KAT and ARG. This assertion is based on the score of each performance at the third component. The corresponding figure (analogous to Figure 8 for first and second components) is not shown here due to space limitations; the same occurs for the fourth and fifth principal components below.

The fourth component (4.53% of total variability; see Figure 7) emphasises the mean rhythmic pattern where there is a slowing tempo in the half of each period, and accelerates until the end of the period. Negative component values (CO1, CUB, CAP) indicate more emphasis within the rhythm of the periods, and positive values (HO3, KLI, BRE) indicate more homogeneous performances in the rhythm of the periods.

The fifth component (2.18% of total variability; see Figure 7) shows a different oscillatory pattern in the periods of the performance of A-phrases. Performances with high positive values in this component (ARG, HO1) are characterised by accelerating the tempo until the end of the *arpeggi*, then decelerating until the half-note in the last bar of the period, and finally accelerating again toward the next period. In contrast, negative values (DEM, ASH, SHE) are characterised by the opposite style: decelerating in the *arpeggi*, accelerating in the next notes until the last bar of the period and then decelerating at the entrance to the next period. The latter also have a slower tempo in the *fermata*, but are faster at the end of phrase B. It can be said that the oscillations in the fifth component are shifted half a bar from the global mean.

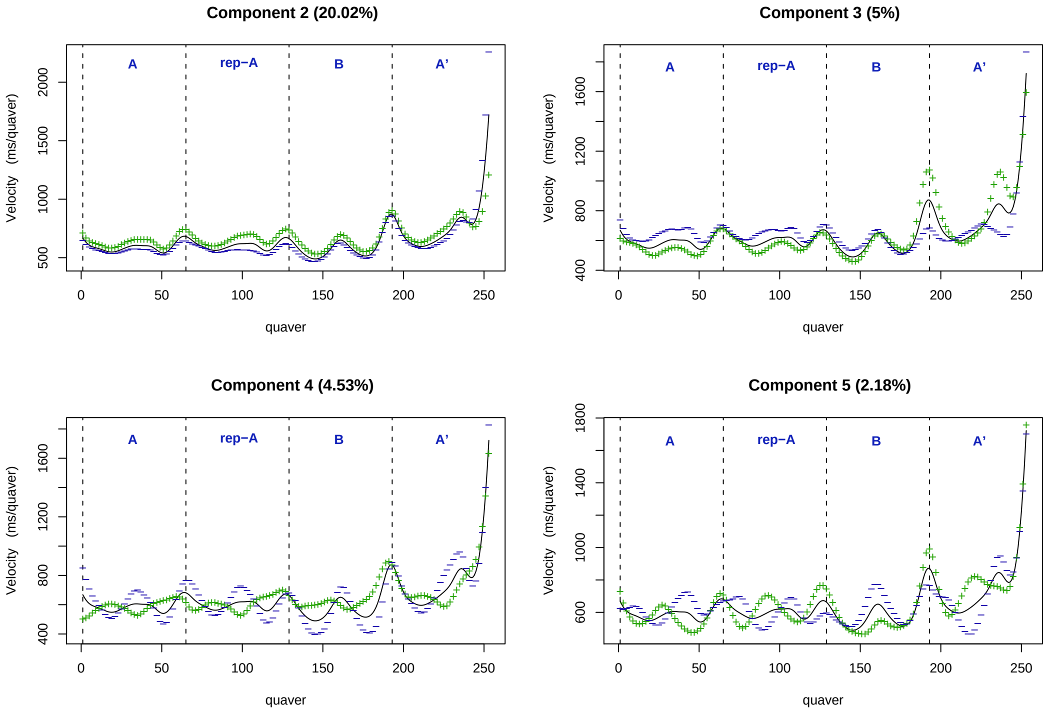

Figure 7. Global mean function +/- a multiple of the principal functions 2 to 5 (non-standardised FPCA).

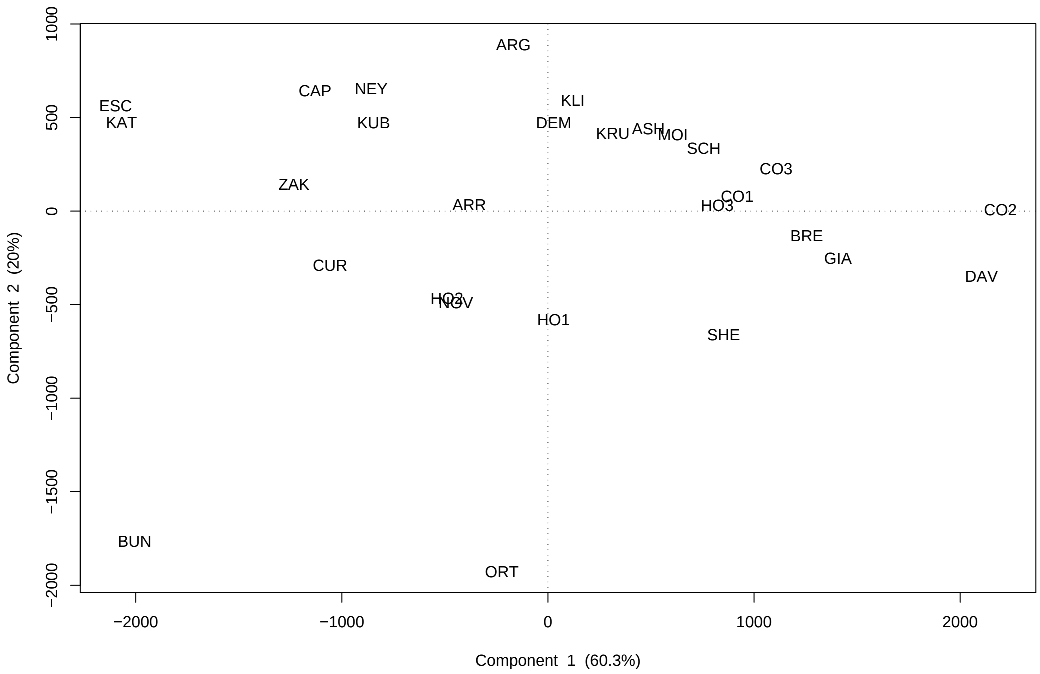

Figure 8. Projection of performers on the plane of first and second principal components (non-standardised FPCA).

5.2. *Standardised FPCA*

The main results are shown in Figures 9 and 10. The first component is also a *size component*, showing that the first characteristic of the piece performance is the global tempo. It explains 70.2% of the total variability. In contrast to the non-standardised FPCA, the final *ritardando* is not so differentiated. In this component BUN is not so well represented as compared with the non-standardised first component, because the notes where BUN tended to play longer are also the notes where all the performers played longer as well.

The second component is very similar to the fourth in the non-standardised analysis (except that now the *fermata* is not a discriminant note). The fifth component is also very similar to the fifth in the non-standardised analysis. So it can be said than both performance patterns are highly characteristic of this piece.

In the third component, those with positive value (BUN, KAT and ORT, at the top of the third graph in Figure 10) follow the mean rhythmic pattern with slightly more emphasis -the longer notes (i.e. *ritardandos*) generally being played longer and the faster ones faster. However, on the negative side (NEY and MOI, at the bottom of the third graph in Figure 10) there is a differentiated performance in the two periods of A phrases: in the first period accelerating in the *arpeggi* and then decelerating, but in the second period decelerating in the *arpeggi* and accelerating afterwards; these pianists also perform the end of phrase B and most of phrase A′ faster.

The fourth factor pattern has two rhythmic change-points: the end of the first period in the A-repetition phrase, and the *fermata*. Performances with a positive component (ARG; the corresponding figure is not included due to space limitations) tend to be faster from the beginning to the end of the first period of the A-repetition, then slower until the *fermata* with the end of the piece being played significantly faster. One may surmise that different pianists anticipate phrase B in different ways, and this is reflected in the repetition of phrase A. As regards the negative component, BUN is the most representative pianist (figure not shown here); he tends to play the second half of the A-repetition and the B phrases faster.

We identify eight meaningful components, although the last ones have a more complex interpretation (see Figure 11; scatterplots for the pianists on the component planes are not included here for these components due to space limitations). For example, pianists on the negative side of the eighth component (ARR,BRE,,ARG) perform the end of phrases A and rep-A more slowly, and are more expressive in the first period of B and A′. On the side of the positive component side (KAT, ZAK) pianists perform the end of the A phrases faster and have more homogeneous timing in phrase B, although slightly slower *fermata*. Inside phrase A and A′, an oscillatory movement similar to the fifth component (but not so marked) appears. The eighth component differs from the fifth component in phrase A′, which does not follow the same structure as A and rep-A.

Comparison between non-standardised and standardised analysis can be summarised as follows: the most important characteristics of the data have been discovered by both methods (components 1, 2, 4 and 5 in the non-standardised analysis, and 1, 2 and 5 in the standardised one), but other are pointed out only by one of them (component 3 in the non-standardised analysis, and 3 to 8 in the standardised one).

5.3. *Cluster analysis*

Once principal functions are computed, a cluster analysis is conducted with the first 8 components from the standardised FPCA. This choice is justified because

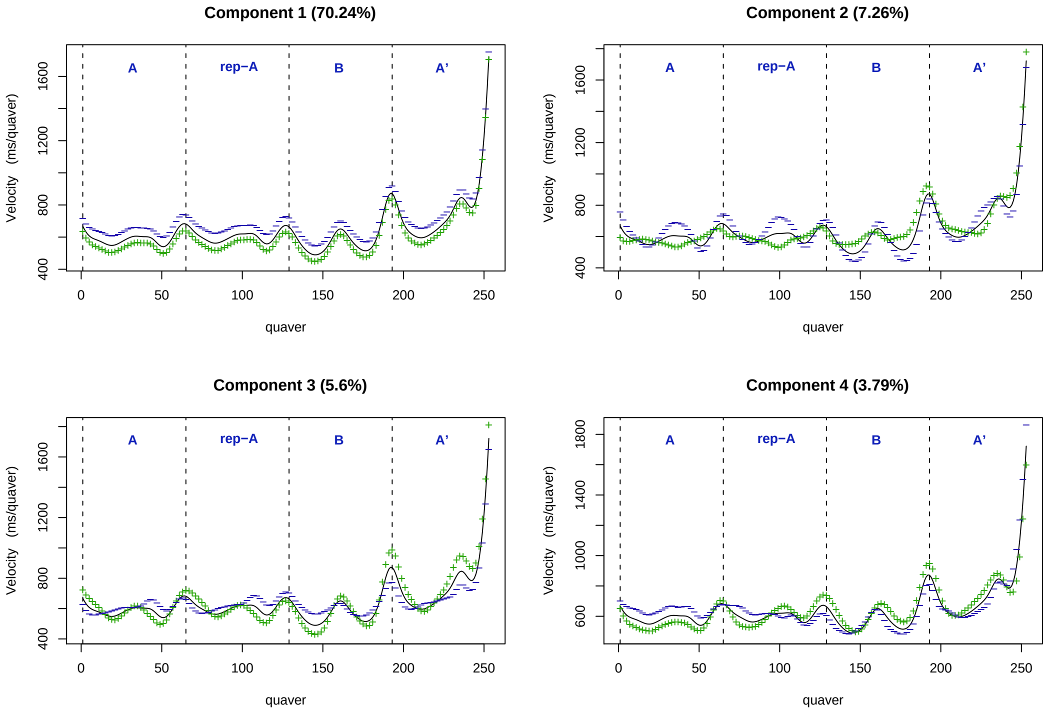

Figure 9. Global mean function +/- a multiple of the principal functions 1 to 4 (standardised FPCA).

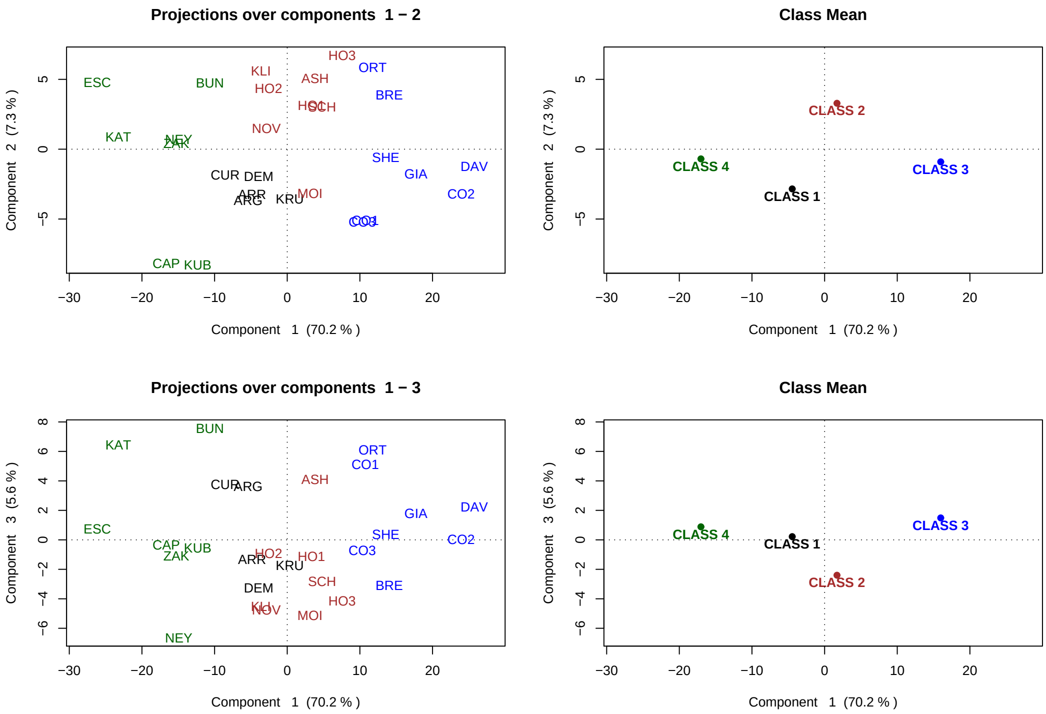

Figure 10. Projection of performers and classes on the planes of principal components (standardised FPCA).

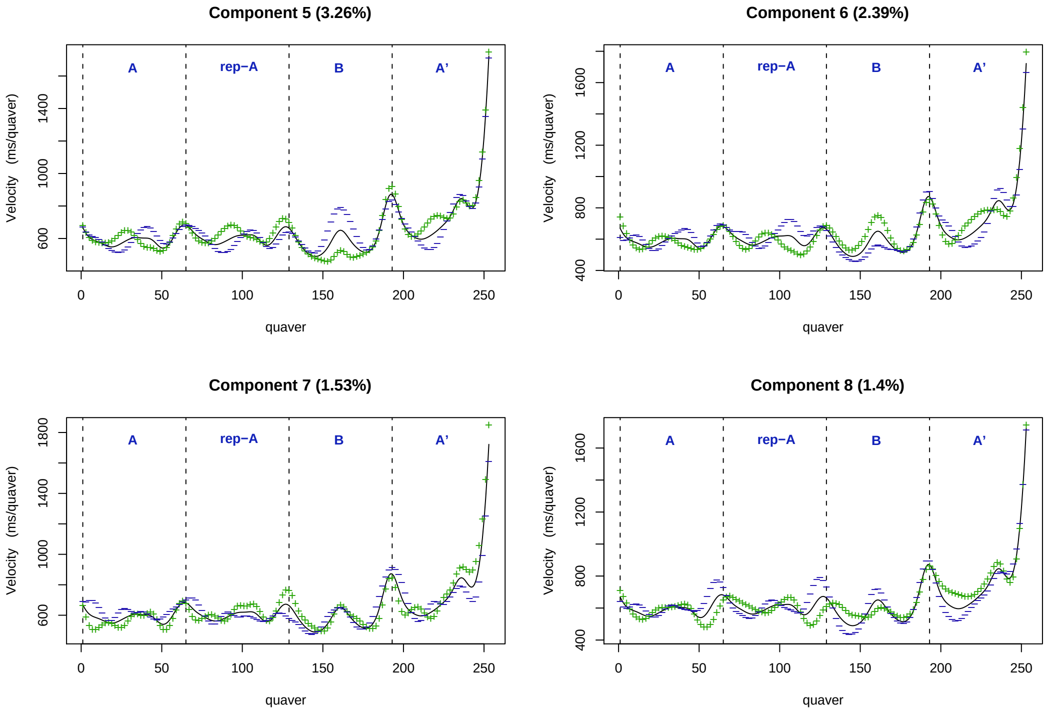

Figure 11. Global mean function +/- a multiple of the principal functions 5 to 8 (standardised FPCA).

standardised analysis show individual characteristics better. The input matrix has dimension 28×8, where the 28 performances have 8 estimated values of their scores in principal components, so any standard cluster algorithm can be used in this context. A hierarchical algorithm is applied with Ward's minimum variance method (the function `agnes` in the R software is used). Cluster analysis defines groups of performances that are similar to other performances in the same group as regards their main characteristics, but are different from elements in other groups. The description of clusters is the only difference between using functional data and using multivariate data.

A standard graphical output of cluster analysis is a *dendrogram*, such as that represented in Figure 12. From the dendrogram we can see that there are three or four clusters. The four cluster partition gives a greater between-class variance relative to the within-class variance, and more meaningful clusters. Class descriptions are shown in Figure 13. Cluster 1 is the closest one to the global mean performance. Cluster 2 is also close to the mean in the global tempo, but with a homogeneous rhythmic (flatter slowness function). Cluster 3 has a faster global tempo and Cluster 4 has a lower tempo. Class membership is as follows (see Figure 10): Class 1: ARG, ARR, CUR, DEM, KRU; Class 2: ASH, HO1, HO2, HO3, KLI, MOI, NOV, SCH; Class 3: BUN, CAP, ESC, KAT, KUB, NEY, ZAK; Class 4: BRE, CO1, CO2, CO3, DAV, GIA, ORT, SHE. The three performances of Horowitz and Cortot are grouped in the same class (Horowitz in class 2, Cortot in Class 4), so although these three performances are different they share the most relevant characteristics.

A complementary class description can be performed by considering the values taken by the principal functions in the elements in each class. Figure 10 (right panels) shows the mean values of principal functions for each class. Cluster 1 is characterised by low values in the second principal function, and significantly it is also located on the negative side of the eighth principal function (figure not

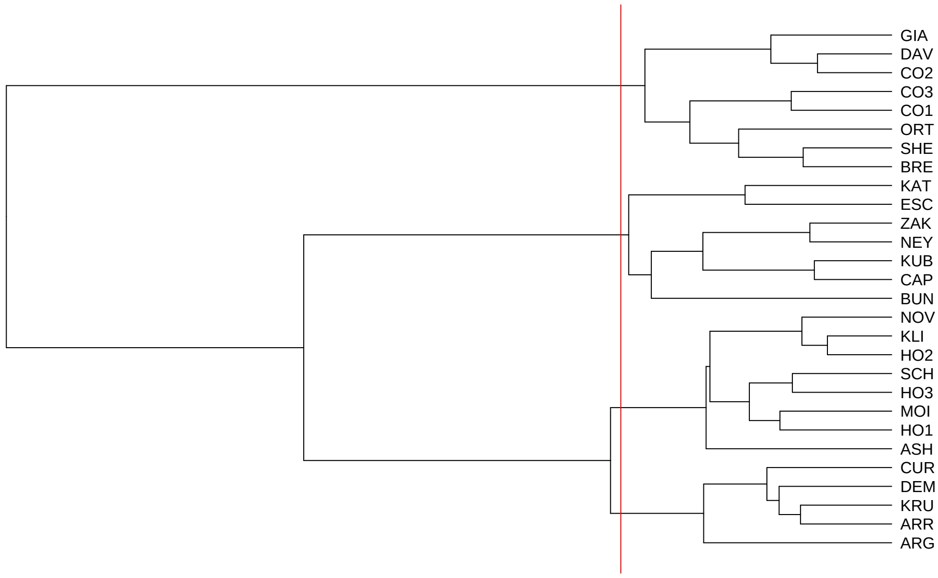

Figure 12. Dendrogram corresponding to cluster analysis from standardised FPCA. The vertical line indicates the tree cut in order to obtain four clusters.

shown here). Cluster 2 is characterised by high second function values and low third function values. Performances in cluster 3 have high first principal function values, whereas the opposite occurs for those in cluster 2.

6. Conclusions

The FPCA analysis enables us to interpret the information on rhythmic diversity in Träumerei performances. Some of the information obtained is obvious (global differences in tempo, for instance), but this type of analysis also looks for significant information that is not so easily perceived. The ability of FPCA to recover *a priori* known information indicates the plausibility of the novel information provided by the method.

All the expected performance characteristics are found in some principal functions (*ritardandos*, *fermata*, phrase structure, etc.). The rhythmic structure of phrase A and its repetition are very similar in all the principal functions, except for the fourth component in the standardised FPCA. The main rhythmic structures in the performances (see, for instance, the fourth principal function in Figure 7) follow the regularity of the musical form of the piece. As an example of non-expected patterns, we mention the fact that the oscillations in the fifth component are shifted half a bar from the global mean.

FDA is shown to be a useful technique for analysing musical performance data. Among its merits are the following: it considers in a natural way musical events as dynamic processes evolving over time; it allows both global and local aspects of musical performances to be combined; the statistical tools involved are no more complex than those employed with multivariate data. A broad variety of musical data (the performance worm of Langner and Goebl 2003, for instance) could benefit

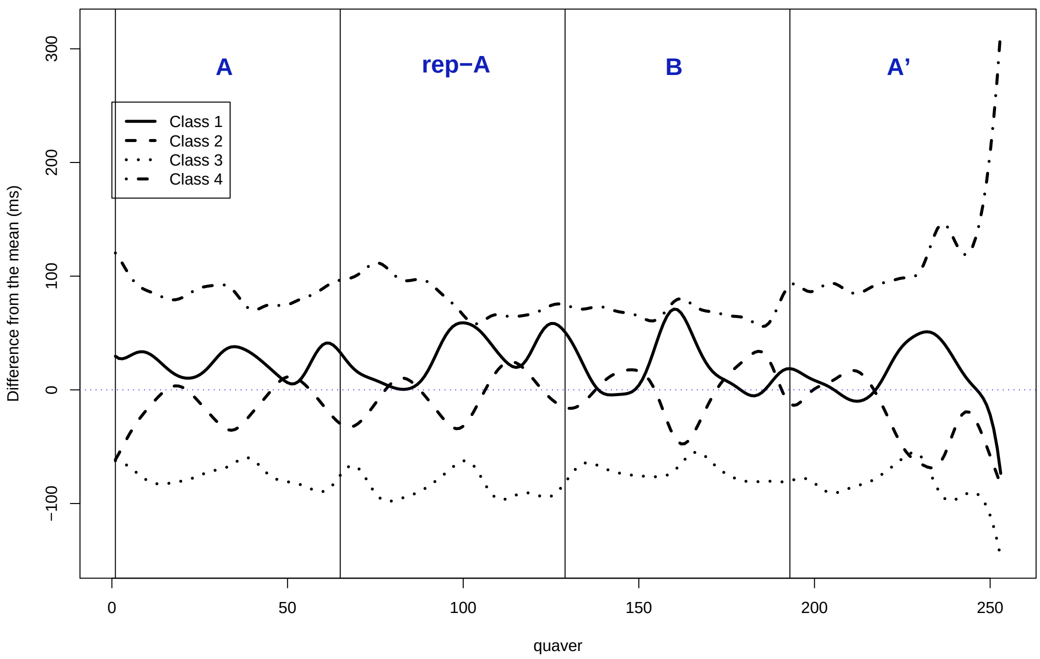

Figure 13. Class descriptions. Group mean functions minus global mean function.

from interaction with FDA methodology.

Acknowledgements

We thank three anonymous referees for their valuable comments, which have enabled us to improve considerably an earlier version of the manuscript. We are very grateful to Bruno H. Repp who provided us with the Träumerei tempo data set. This research is partially supported by the Spanish Ministry of Education and Science and FEDER, MTM2006-09920.

References

Aguilera, A., Gutiérrez, R., and Valderrama, M. (1996), "Approximations of estimators in the PCA of a stochastic process using B-splines," *Communication in Statisitcs, Simulation and Computation*, 25(3), 671–690.

Almansa, J., "Análisis de datos musicales mediante componentes principales funcionales: estudio interpretativo de la obra "Träumerei"," Projecte de fi de carrera, Facultat de Matemàtiques i Estadística. Universitat Politècnica de Catalunya, In Spanish (2005).

Beran, J., and Mazzola, G. (1999a), "Analyzing Musical Structure and Performance–A Statistical Approach," *Statistical Science*, 14(1), 47–79.

Beran, J., and Mazzola, G. (1999b), "Visualizing the Relationship Between Two Time Series by Hierarchical Smoothing Models," *Journal of Computational and Graphical Statistics*, 8(2), 213–238.

Brendel, A. (1981), "Der Interprer muss erwachsen sein. Zu Schumann's 'Kinderszenen'," *Musica*, 6, 429–433.

Davidian, M., Lin, X., and Wang, J. (2004), "Introduction: Emerging Issues in Longitudinal and Functional Data Analysis," *Statistica Sinica*, 14, 613–614.

Dixon, S., Goebl, W., and Widmer, G. (2002), "The PERFORMANCE WORM: Real Time Visualisation of Expression Based on Langner's Tempo-Loudness Animation," in *Proceedings of the International Computer Music Conference (ICMC'2002)*, Göteborg, Sweden.

Fan, J., and Gijbels, I., *Local polynomial modelling and its applications*, London: Chapman & Hall (1996).

Ferraty, F., and Vieu, P., *Non parametric functional data analysis. Theory and practice*, Springer (2006).

Gabrielsson, A. (2003), "Music performance research at the millenium," *Psychology of Music*, 24, 68–91.

González-Manteiga, W., and Vieu, P. (2007), “Editorial: Statistics for Functional Data,” *Comput. Stat. Data Anal.*, 51, 4788–4792.

Huron, D., *Sweet Anticipation: Music and the Psychology of Expectation*, Cambridge, Massachusetts: MIT Press (2006).

Kneip, A., and Utikal, K. (2001), “Inference for Density Families Using Functional Principal Component Analysis,” *Journal of the American Statistical Association*, 96, 519–542.

Langner, J., and Goebl, W. (2003), “Visualizing expressive performance in tempo-loudness space,” *Computer Music Journal*, 27, 69–83.

Levitin, D.J., Nuzzo, R.L., Vines, B.W., and Ramsay, J.O. (2007), “Introduction to Functional Data Analysis,” *Canadian Psychology*, 48(3), 135–155.

Mardia, K.V., Kent, J.T., and Bibby, J.M., *Multivariate Analysis*, Academic Press (1979).

Mazzola, G., and Zahorka, O. (1994), “The RUBATO Performance Workstation on NeXTSTEP,” in *Proceedings of the ICMC 1994*, Aarhus: ICMA.

Meyer, L.B., *Emotion and Meaning in Music*, Chicago: Chicago University Press (1956).

R Development Core Team,, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2008).

Ramsay, J., and Silverman, B.W., *Applied Functional Data Analysis: Methods and Case Studies*, New York: Springer-Verlag (2002).

Ramsay, J., and Silverman, B.W., *Functional Data Analysis*, second ed., New York: Springer (2005).

Repp, B.H. (1992), “Diversity and commonality in music performance: An analysis of timing microstructure in Schumann’s “Träumerei”,” *Journal of the Acoustical Society of America*, 92(5), 2546–2568.

Repp, B.H. (1995), “Expressive timing in Schumanns “Träumerei:” An analysis of performances by graduate student pianists,” *Journal of the Acoustical Society of America*, 98(5), 2413–2427.

Repp, B.H. (1996), “The dynamics of expressive piano performance: Schumanns “Träumerei” revisited,” *Journal of the Acoustical Society of America*, 100(1), 641–650.

Reti, R., *The Thematic Process in Music*, New York: Macmillan (1951).

Stein, E., *Form and Performance*, London: Faber and Faber (1962).

Traub, A. (1981), “Die ’Kinderszenen’ als zyklisehes Werk,” *Musica*, 6, 424–426.

Valderrama, M.J. (2007), “Editorial: An overview to modelling functional data,” *Computational Statistics*, 22, 331–334, Special issue on Modelling Functional Data in Practice.

Vines, B.W., Nuzzo, R.L., and Levitin, D.J. (2005), “Analyzing temporal dynamics in music: Differential calculus, physics, and functional data techniques,” *Music Perception*, 23(2), 137–152.

Widmer, G., Dixon, S., Goebl, W., Pampalk, E., and A., T. (2003), “In Search of the Horowitz Factor,” *AI Magazine*, 24(3), 111–130.