Journal of Mathematics and Music
Vol. 3, No. 3, November 2009, 165–173

What pre-whitened music can tell us about multi-instrument compositions

R.E. Dumas[a,b] and A.P. Georgopoulos[a,b,c]*

[a] *Brain Sciences Center, Department of Veterans Affairs, Minneapolis, MN 55417, USA;* [b] *Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA;* [c] *Center for Cognitive Sciences, University of Minnesota, Minneapolis, MN 55455, USA*

(*Received 30 August 2006; final version received 16 March 2009*)

We have discovered remarkable temporal associations between instruments in multi-instrument musical pieces after elimination of melody. This indicates the presence of a compositional 'structural' framework among instruments based on dynamic, temporal interactions. We hypothesize that such a framework is at the heart of the process of musical composition. Since similar melodies can be arranged in different ways, resulting in unique compositions, the instrument framework above may be a fundamental aspect of the musical arrangement in the composer's palette.

Keywords: music; time series; composition; cross-correlation

AMS Subject Classifications: 62H05; 62H20; 62H30; 37M10; 62M10

1. Introduction

Early references to orchestras indicate that they were haphazard arrangements of assorted instruments, devoid of the homogenous sections that make up their modern counterparts. For example, the orchestra of Nebuchadnezzar (c. 630–562 BC) was comprised of sackbuts, psaltery, wooden cornets, recorders, harps, citterns, viols, and other instruments either blown or plucked and was designed chiefly to call the people to worship [1, p. 11]. Their strength was in their numbers: the more musicians that gathered, the louder the sound. Along the same lines, town musicians in the Middle Ages were recruited primarily as watchmen who stood at the gates and blew their shawms (primitive oboes) to sound the alarm for fire or approaching strangers [1, p. 13]. With the exception of brass players and drummers, whose services were reserved exclusively for royal fanfares and celebrations,[1] Renaissance instrumentalists of the mid-1500s performed elaborate polyphonic music at dances and at church services squeezed together into the balcony, often as a substitute for voices [1, p. 15]. Even 100 years later, a painting of the 1656 carnival celebrating Sweden's Queen Christina's arrival in Rome shows the musicians 'packed at random into their balcony. Some are seated, some are standing; some face left, others right; strings, winds,

*Corresponding author. Email: omega@umn.edu

ISSN 1745-9737 print/ISSN 1745-9745 online

DOI: 10.1080/17459730903378273
http://www.informaworld.com

and singers are mixed together in no apparent order' [3]. In 1607, the Italian composer Claudio Monteverdi introduced the exploitation of instrumental tone colour by scoring for three parts: vocal, bass and all others [1, p. 17], paving the way for the eventual arrangement of instruments of the same family (strings, woodwinds, brasses) into orchestral sections. The Baroque Mannheim School, founded by Bohemian musician Johann Stamitz in 1742, produced Europe's first virtuoso orchestra. Known for its precision, broad range of dynamics and exquisite blending of instrumental timbres, the Mannheim Court model inspired Mozart and became the foundation for the classical symphony orchestra [1, p. 62].

In this study, we analysed compositions by Western European classical composers because of the ready availability of these works in Standard MIDI File (SMF) format (see Section 2). Musical forms that help to define compositions are known as archetypes [4]. Archetypes are frameworks upon which composers may 'hang' their musical ideas. A list of the most obvious elements (constraints) of an archetype could include key, tempo, and time signature. More specific properties, such as scale, rhythm, and timbre (instrument sound) might also be identified by the casual listener. In this study, we examined, instead, the subtle relations in time between pairs of instrument parts in orchestral arrangements, stripped of their melody, in order to determine their interactions and evaluate them with respect to the section to which an instrument part belongs. First, we tested the hypothesis that grouping of instruments in sections, as reflected by the seating arrangement of the symphony orchestra, depends on the presence of a melody/harmony within parts played. For that purpose, we removed all melody and harmony (pitch space) through a pre-whitening process [5] and calculated the cross-correlation function within approximately a three-measure window. And second, we used multi-dimensional scaling (MDS) to visualize constellations of instruments in an orchestral setting under various conditions of data processing. We found that suitable instrument grouping could be obtained from the strength of cross-correlation between pre-whitened (i.e. noisy) pitch sequences, in the absence of melody in individual sequences. This finding documents the importance of temporal interactions between instrument parts, among others (e.g. physical characteristics of the instrument), in determining their grouping.

2. Methods

2.1. *Score-to-MIDI-to-data*

The parts of an orchestral score were converted into individual data-streams using several transformations, as follows. SMF were used. The files used were downloaded from an online collection [6]. MIDI events were converted to ASCII text format using a custom-designed MAX/MSP® program running on a personal computer (see Supplement). This program produced one data-stream (note range: 1–127, rest = 0) for every instrument part in a given piece of a composer. The data-streams were then concatenated into a single, multi-column text file where columns represented instruments and rows represented time in 16th-note units. We analysed four pieces from J. S. Bach, L.v. Beethoven, W.A. Mozart, and A. Vivaldi for a total of 16 pieces. Details are given in Table 1.

We also produced data-streams containing information regarding the temporal position of attacks (note onsets) for every instrumental part for use in our analysis of rhythmic (as opposed to melodic) components of the MIDI sequences. These data-streams were binary in nature (onset of note $\rightarrow$ 1, the balance of note duration and any rest $\rightarrow$ 0).

2.2. *Instrument sections*

Based on common compositional practice, we assigned each instrument to the following sections: violin, viola, cello, contrabass, keyboard (organ, harpsichord, harpsichord basso continuo),

Table 1. Orchestral works used.

Piece	Composer	Title	Year	Instrument parts
1	J.S. Bach	Brandenburg Concerto No. 3 in G, (BWV1048)	1713	11
2	J.S. Bach	Werde munter, mein Gemüthe (Jesu, Joy of Man's Desiring), (BWV147)	1716	11
3	J.S. Bach	Orchestral Suite No. 3, (BWV1068) Overture	1723	11
4	J.S. Bach	Passacaglia and Fugue in C minor, (BWV582)	1708–1712	22
5	L.v. Beethoven	Symphony No. 7 in A major, Opus 92, Allegretto	1725	21
6	L.v. Beethoven	Symphony No. 3 E^b major, Opus 55. 'Eroica', Allegro con brio	1807–1808	22
7	L.v. Beethoven	Symphony No. 5 in C minor Opus 67, Allegro con brio	1807–1808	22
8	L.v. Beethoven	Symphony No. 6 in F major 'Pastorale', Opus 68, Allegretto	1808	23
9	W.A. Mozart	Cosi fan tutte, (K. 588) Overture	1790	22
10	W.A. Mozart	Don Giovanni, (K. 527) Overture	1787	21
11	W.A. Mozart	The Marriage of Figaro, (K. 492) Overture	1786	21
12	W.A. Mozart	The Magic Flute, (K. 620) Overture	1791	21
13	A. Vivaldi	'The Seasons', La Primavera (Spring) in E major, (RV.269) Allegro	c. 1725	13
14	A. Vivaldi	'The Seasons', L'Estate (Summer) in G minor, (RV.315) Allegro non molto	c. 1725	10
15	A. Vivaldi	'The Seasons', L'Autunno (Autumn) in F major, (RV.293)	c. 1725	11
16	A. Vivaldi	'The Seasons', L'Inverno (Winter) in F minor, (RV.297)	c. 1725	11

woodwind (flute, clarinet, bassoon, oboe), brass (trumpet, trombone, French horn, tuba), and timpani.

2.3. *General statistical analyses*

Standard statistical analyses were used to analyse data [7], including analysis of variance (ANOVA).

2.4. *Time series analyses*

The main objective of the present study was to assess the interactions between time series (of pitch) in pairs of instruments. For that purpose, individual time series of pitched notes need to be stationary, i.e. 'pre-whitened' [4]; otherwise, non-stationarities in the series themselves can lead to erroneous associations [8,9]. Therefore, the first step in our analyses was to model the time series and derive stationary (or quasi-stationary) residuals from which to compute pairwise association measures, such as cross-correlations. For that purpose, a Box–Jenkins ARIMA modelling analysis [4] was performed to identify the temporal structure of the data time series, using 25 lags, corresponding to 25 16th-notes. We carried out these analyses on the entire time series (from 1279 to 18,676 time points, depending on the piece). After extensive ARIMA modelling and diagnostic checking, including computation and evaluation of the autocorrelation function (ACF) and partial ACF (PACF) of the residuals, we derived practically stationary residual series for each instrument part. The exact model used differed for different pieces, and frequently included periodic components. The ARIMA modelling was carried out using the BMPD®/Dynamic statistical package (Los Angeles, 1992). All possible cross-correlations between pairs of stationary instrument parts in a piece were calculated using the DCCF routine of the IMSL statistical library (Compaq Visual Fortran® Professional edition version 6.6B). The statistical significance of specific cross-correlations was assessed using standard errors calculated by Bartlett's formula [10], based on a general asymptotic expression for the variance of the sample cross-correlation coefficient of two jointly stationary time series with independent, identically distributed normal errors. For each cross-correlogram, we identified the maximum absolute cross-correlation r_{ij} between instruments i and j, and noted its sign and lag. For statistical analyses, r_{ij} was transformed to z_{ij} using Fisher's z-transformation to normalize its distribution: $z_{ij} = 0.5[\ln(1 + r_{ij}) - \ln(1 - r_{ij})]$.

2.5. *MDS*

Four MDS analyses were performed on each one of the 16 pieces, using (a) the original MIDI values, (b) the residuals after pre-whitening, (c) the attack data, and (d) the z_{ij} from the cross-correlogram. In all cases, the procedure ALSCAL of the SPSS® statistical package (SPSS for Windows, version 10.1.0, SPSS Inc., Chicago, IL, 2000) was employed for a two-dimensional solution using the data at the ordinal level. For data (a) and (b) above, the Euclidean distance was used as the proximity measure. For the binary data (c), the variance measure was used as a proximity distance. This measure ranges from 0 to 1 and was computed from a fourfold table as $(b + c)/4n$, where b and c represent the diagonal cells corresponding to cases present on one item but absent on the other and n is the total number of observations. Finally, for data (d), the signed z_{ij} was subtracted from a positive constant ($= 2$) to obtain a proximity matrix consisting of positive distances only. The MDS analyses yielded values for the stress and R^2, and provided a two-dimensional configuration of the instruments for each piece.

3. Results

Figure 1 shows the dramatic effect of pre-whitening on the autocorrelation and partial autocorrelation structure of a single instrument part from a score. Figure 2 shows examples of cross-correlograms before and after pre-whitening. We focused on the maximum (absolute peak value) cross-correlation in each cross-correlogram. The large majority (87.2%) of these peaks were positive and occurred mostly at zero lag, indicating synchrony (Figure 3). (The few negative peak cross-correlations were distributed across lags without a peak at zero lag.) A crucial factor for the magnitude of the peak cross-correlation was whether the two instruments in a pair belonged to the same or different section (see Section 2), such that cross-correlations were much higher between instruments belonging in the same section. Specifically, the average ($\pm$SEM) z_{ij} was 0.853 $\pm$0.024 ($n = 473$ pairs) for intra-section vs. 0.228 $\pm$0.012 ($n = 1942$) for intersection instrument pairs. These means differed significantly overall ($p < 10^{-14}$, F-test in ANOVA) and for each composer individually (at least $p < 10^{-5}$).

Figures 4–7 illustrate the results of MDS analyses for Beethoven's 'Symphony No. 7, Movement No. 2' (22 instruments). (a) Figure 4 shows the MDS solution for the original MIDI data. It can be seen that the arrangement of the instruments in the two-dimensional MDS space provides for the clustering of instruments belonging to the same family (e.g. strings, woodwind, brass, percussion). This closely resembles the segregation of such instrument families in a typical seating chart of a symphony orchestra. (b) This orderly arrangement essentially disappeared for the pre-whitened data (Figure 6), that is when the melody was stripped and the resulting MIDI residuals were used for the MDS analysis. This result is important because it demonstrates the effectiveness of the pre-whitening process to destroy the orchestral arrangement. (c) The MDS analysis of the attack data provided a degree of instrument clustering (Figure 7) but the placement of the clusters was peculiar, as, for example, violins were placed next to the woodwinds and brass, but apart from the other strings. By contrast, when (d) the temporal association (z_{ij}) between instrument parts was used, the MDS solution yielded an orderly clustering of instrument families (Figure 5) which,

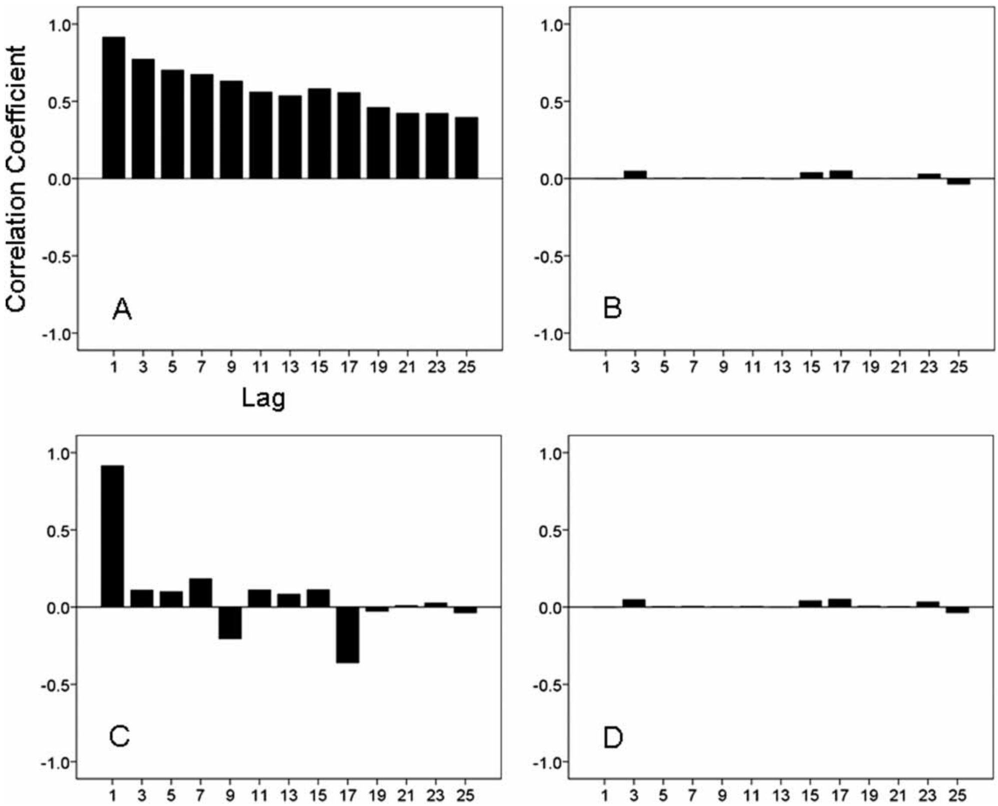

Figure 1. Autocorrelograms before (A) and after pre-whitening (B) for instrument part flute 1 Mozart's overture to the opera 'Don Giovanni'. PACFs before (C) and after pre-whitening (D) are also shown for the same data. Lag is in 16th-notes. Notice the high autocorrelations in the raw data and their dramatic reduction after pre-whitening. The ARIMA model used comprised seven autoregressive orders (1, 5, 8, 9, 13, 16, 24) and 2 moving average orders (1 and 16).

in fact, seemed more coherent than that derived from the original data (Figure 4; compare, for example, the differential clustering of oboes, bassoons, flutes, clarinets, and violas). This finding underscores the point that it is the *relations* between the pre-whitened data that retain the orchestral arrangement information, and not the pre-whitened series themselves, which yielded a disordered arrangement (Figure 4). Finally, another example is illustrated in Figure 8 for Vivaldi's 'The Seasons – Autumn' and is in keeping with the finding above that the strength of the maximum cross-correlation was substantially higher between instruments belonging to the same rather than different sections. Overall, the MDS solutions for z_{ij} were excellent (median stress $= 0.154$, range 0.02–0.239; median $R^2 = 0.903$, range 0.749–0.998).

4. Discussion

These results revealed significant temporal associations between instruments in the absence of melody. Moreover, these associations yielded a remarkable clustering of instruments in two-dimensional solutions of MDS analysis, at least as good as that yielded by the original MIDI data and superior to that yielded by rhythm. This indicates a compositional archetype [5] based on dynamic, temporal interactions among instruments. We hypothesize that such a framework is at the origin of the process of musical composition, and that melodies follow. Of course, the

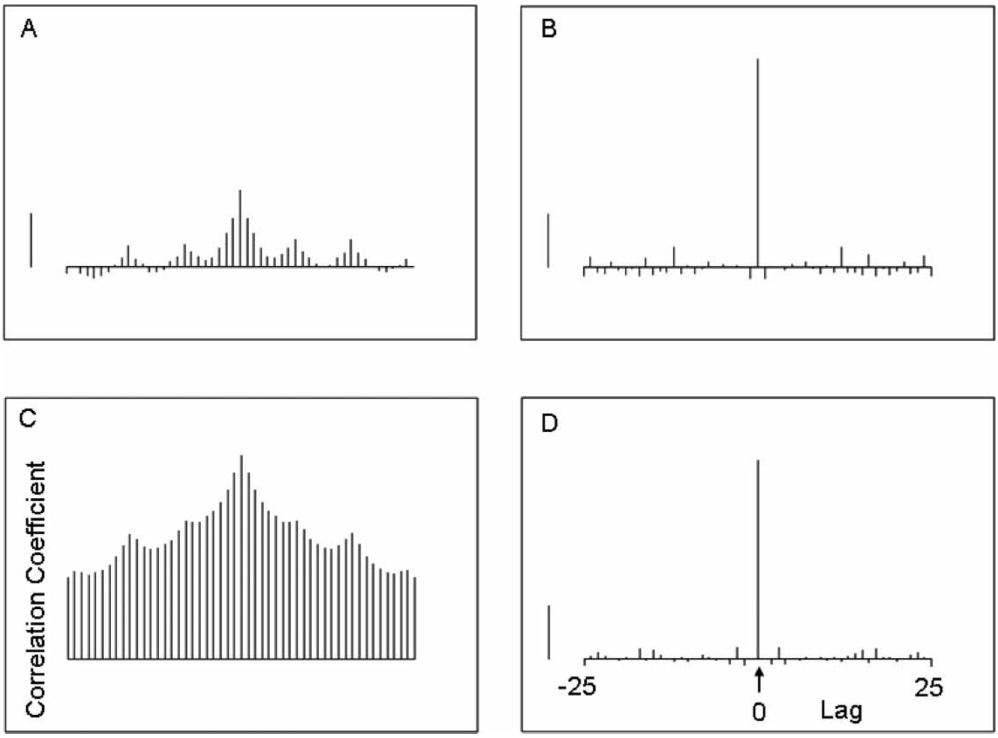

Figure 2. Examples of cross-correlograms before (A, C) and after pre-whitening (B, D). A and B are cross-correlograms between instrumental parts (violin 1 and violin 2) in Mozart's overture to the opera 'The Marriage of Figaro'. C and D are cross-correlograms between instrumental parts (flute 1 and flute 2) in Mozart's overture to the opera 'Don Giovanni'. Arrow indicates zero lag (range: -25 to 25 16th-notes); vertical scale bar indicates a cross-correlation of 0.25. Notice the exquisite temporal focusing of the temporal interaction between the instruments after pre-whitening. The cross-correlograms between non-pre-whitened data reflect spurious correlations due, most probably, to dependencies within the series themselves.

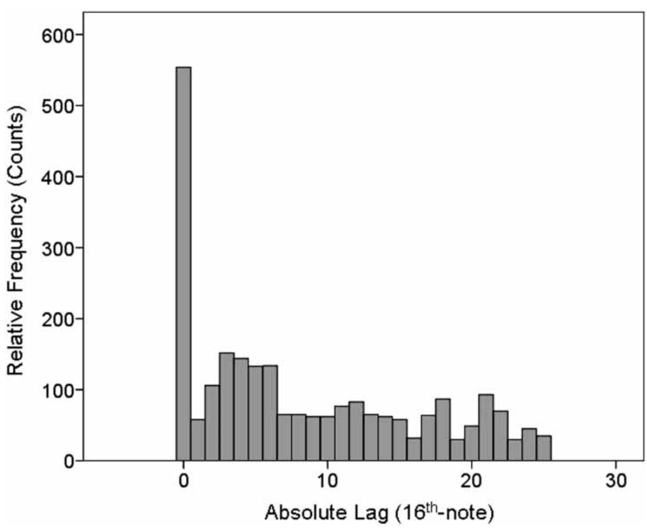

Figure 3. Histogram of absolute lags of all cross-correlogram peaks for all pieces ($n = 2415$ values).

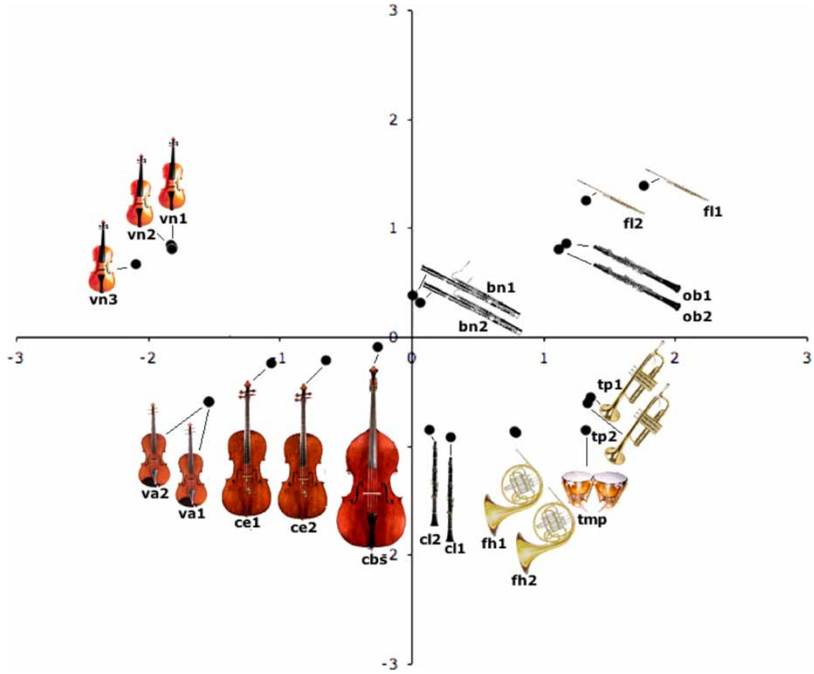

Figure 4. Two-dimensional MDS solution for the original MIDI data of L.v. Beethoven's 'Symphony No. 7, Movement No. 2' (stress $= 0.124$, $R^2 = 0.924$). **Vn1, vn2, vn3**: violin 1, 2, 3; **va1, va2**: viola 1, 2; **ce1, ce2**: cello 1, 2; **cbs**: contrabass; **fl1, fl2**: flute 1, 2; **cl1, cl2**: clarinet 1, 2; **ob1, ob2**: oboe 1, 2; **bn1, bn2**: bassoon 1, 2; **tp1, tp2**: trumpet 1, 2; **fh1, fh2**: French horn 1, 2; **tmp**: tympani. For this piece, the derived dimensions relate to tone colour (Dimension 1, abscissa) and, to a good extent, pitch range (Dimension 2, ordinate).

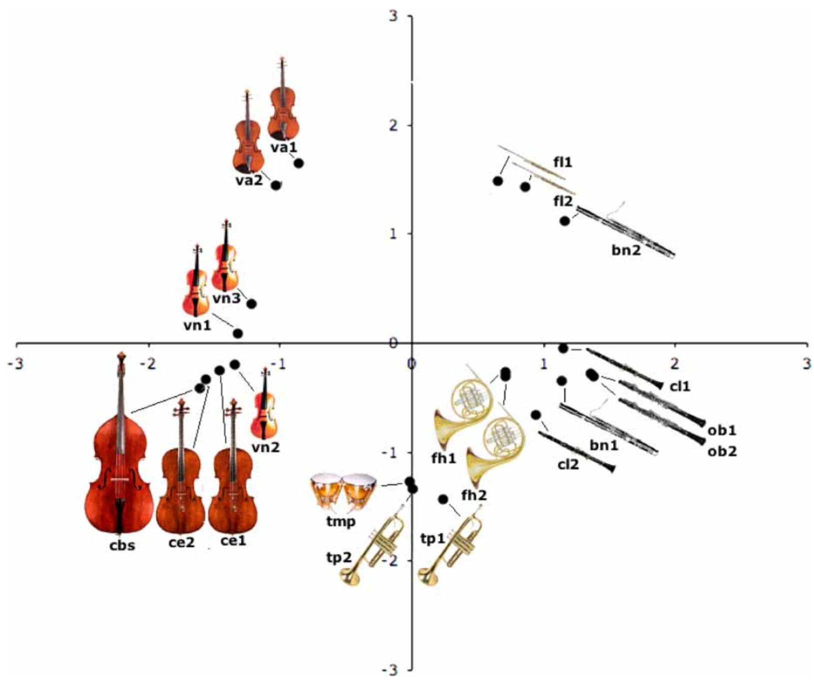

Figure 5. Two-dimensional MDS solution for the z_{ij} between pre-whitened residuals of L.v. Beethoven's 'Symphony No. 7, Movement No. 2' (stress $= 0.172$, $R^2 = 0.848$).

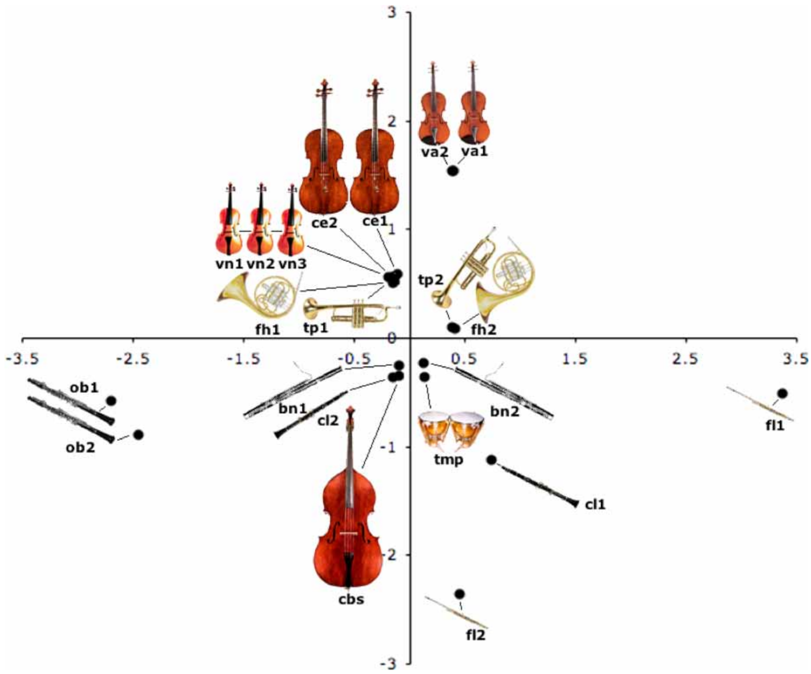

Figure 6. Two-dimensional MDS solution for pre-whitened MIDI residuals of L.v. Beethoven's 'Symphony No. 7, Movement No. 2' (stress $= 0.278$, $R^2 = 0.793$).

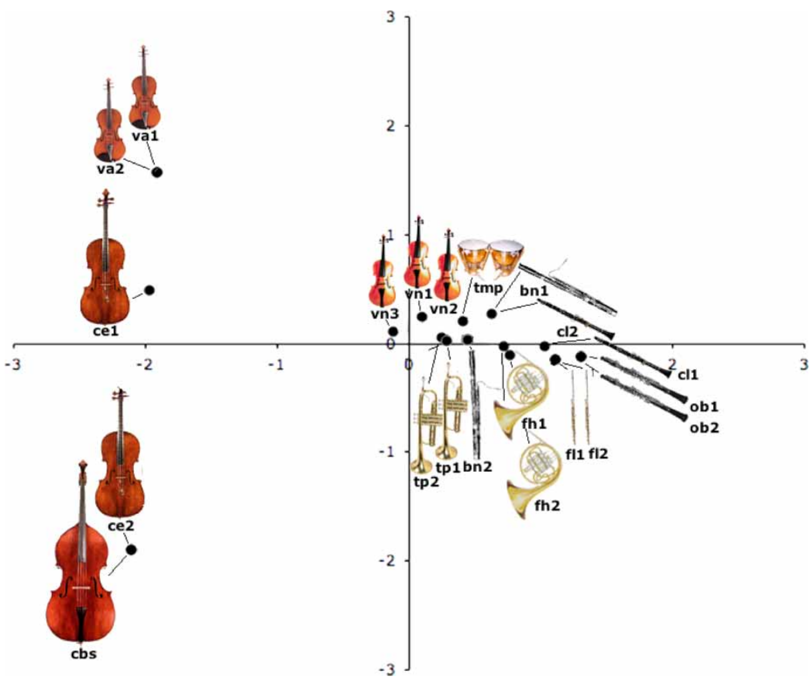

Figure 7. Two-dimensional MDS solution for the attack data of L.v. Beethoven's 'Symphony No. 7, Movement No. 2' (stress $= 0.0714$, $R^2 = 0.988$).

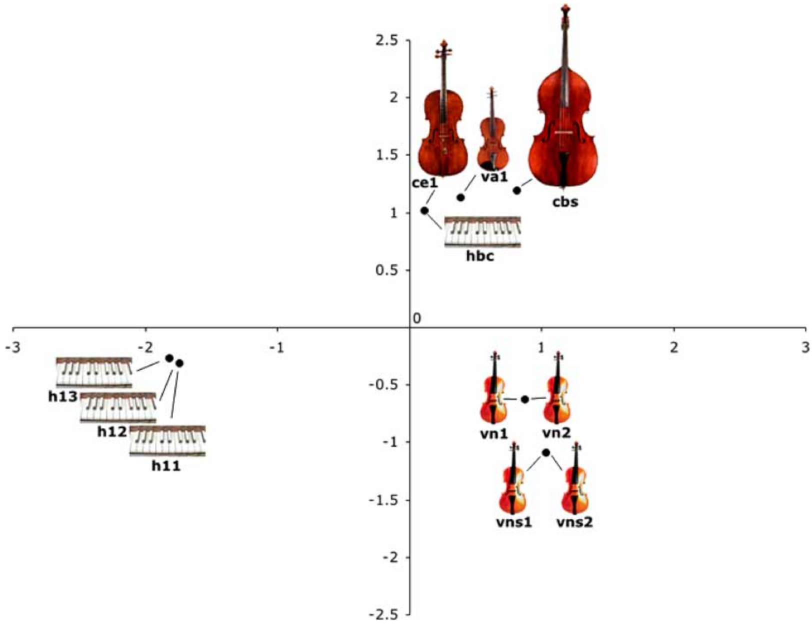

Figure 8. Two-dimensional MDS solution for the z_{ij} between prewhitened residuals of A. Vivaldi's 'The Seasons – Autumn, RV293'. The fit was excellent (stress $= 0.02$, $R^2 = 0.998$). The coordinates for the instruments illustrated are indicated by the filled circles. **Vn1, vn2**: violin 1, 2; **vns1, vns2**: solo violins 1, 2; **va1**: viola; **ce1**: cello; **cbs**: contrabass; **h11, h12, h13**: harpsichord voices 1, 2, 3; **hbc**: harpsichord basso continuo. For this piece, the derived dimensions apparently refer to tone colour (Dimension 1) and pitch range (Dimension 2).

inter-instrument framework and the melodies assigned to specific instruments should be somehow congruent, such that the final outcome intended by the composer is successfully achieved. However, similar melodies can be arranged in different ways, resulting in unique compositions. Indeed, we propose that this arrangement evolves from the instrument framework above.

Note

1. Queen Elizabeth I (1533–1603) occasionally engaged two dozen trumpeters and drummers to serenade her dinner guests [2].

References

[1] H. Raynor, *Orchestra, A History*, Charles Scribner's Sons, New York, 1978, pp. 13–22.
[2] P.M. Young, *Phoenix Music Guides: 2 Symphony*, Crescendo, Boston, p. 34.
[3] J. Spitzer and N. Zaslaw, *The Birth of the Orchestra: History of an Institution, 1650–1815*, Oxford University Press, New York, 2004.
[4] B.S. Rosner and L.B. Meyer, *Melodic processes and the perception of music*, in *The Psychology of Music*, D. Deutsch, ed., Academic Press, New York, 1982.
[5] G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco, 1970.
[6] P. Schwob, *The Classical Music Archives*, 2005. Available at www.classicalarchives.com (accessed 03 February 2003).
[7] G.W. Snedecor and W.G. Cochran, *Statistical Methods*, 8th ed., Iowa State University Press, Ames, IA, 1989.
[8] G.M. Jenkins and D.G. Watts, *Spectral Analysis and its Applications*, Holden-Day, San Francisco, 1968.
[9] M.B. Priestley, *Spectral Analysis and Time Series*, Academic Press, London, 1981.
[10] M.S. Bartlett, *Stochastic Processes*, 3rd ed., Cambridge University Press, Cambridge, 1978.