Perspectives of New Music
Vol. 22 n. 1/2 (1984)

Symmetry, Regularity, Direction, Velocity

ROBERTO DOATI

My primary interest is the precise temporal organization of sound structures, and for this purpose I have chosen the tool that allows the greatest control over time: the computer. Since the use of psychoacoustics has proven valid for sound perception research, an investigation employing psychology might provide insight into and understanding of the principles that control the perception of musical forms.

More than others the Gestalt psychological school has been devoted to the study of form, not only visual (spatial) perception, but also auditory (temporal) perception of forms. Several musical and, more generally, acoustical examples have been provided by the founders of this school to prove their theories (Köhler, 1929; Koffka, 1935; Wertheimer, 1945).

In 1890 two basic principles were proposed by von Ehrenfels in his article on the psychology of the properties of form: 1. although both spatial and temporal structures can be **decomposed** into separate elements, they cannot be **reduced** to them, and 2. they can be **transposed** with no perceptual change (cfr. Guillaume, 1937).

These principles raise the following question: if a melody is not determined by the separate sounds that form it then why do we perceive it as a whole? To solve this problem the advocates of Gestalt-theory rejected the Ehrenfels sensation concept of perception, and arrived at the formulation of the principles on which perception of form depends through experiments dealing with formal conditions and their laws of transformation (Koffka, 1909, 1922, 1935; Köhler, 1915, 1929; Wertheimer, 1945). These principles have been demonstrated for the visual, auditory, and tactile perceptual systems.

It is worth mentioning here some of these principles using visual and sound examples. The law of proximity (fig. 1): between close elements there is a stronger bond than between distant ones (**sound example 1**). The law of similarity (fig. 2): in a

Figure 1

Figure 2

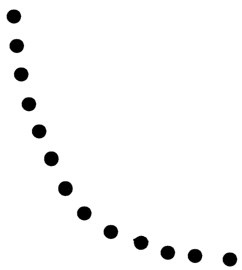


Figure 3

multitude of different elements the similar ones tend to group. (This law is shown in **sound example 2** with timbre, but it could be easily proven with other parameters as well.) The law of good continuation (fig. 3): elements that follow one another in the same direction tend to be unified (**sound example 3**).

A further principle of Gestalt psychology is figure-ground segregation: elements grouped by the above mentioned laws, making up a contour, lead to the formation of a (definite, closer, interesting, significant, smaller) figure standing out from a (amorphous, indefinite, more distant, neutral, bigger) ground. The principles set forth by the Gestaltists have been advantageously used to continue the investigation into sound stimuli arrangement (Guilford, Hilton, 1933; Heise, Miller, 1951; Thurlow, 1957; Bregman, Campbell, 1971; Divenyi, Hirsch, 1978; Rasch, 1978; McAdams, Bregman, 1979; Moore, 1982), and in different fields such as musical analysis (Reichenbach, 1940; Meyer, 1956; Tenney, Polansky, 1980), the creation of auditory images (McAdams, 1982), and individual musical phenomenon perception (melody, rhythm, timbre) (Guilford, Hilton, 1933; Vernon, 1934; Fraisse, Piaget, 1963; Roederer, 1972; Fraisse, 1974, 1975; Deutsch, 1976; McAdams, Bregman, 1979; McAdams, 1982).

Although I recognize music as consisting of many closely related or connected parameters, I started research on a possible compositional application of the above mentioned theories. Some musical examples and two compositions realized by means of a computer at the Centro di Sonologia Computazionale, University of Padua, in 1980-81 and 1981-82 are the results thus far of this research.

The first example to be realized musically is shown in fig. 4. The law of good shape favors the choice of simple, regular, and symmetric shapes, and such is the polyphonic structure I have generated in a two-dimensional space with frequency as the ordinate and time as the abscissa. The construction of this sound shape was based on the Music 5 instrument found in Mathews (1969). This instrument consists of an oscillator with a sine wave whose amplitude is controlled by an envelope which decays exponentially, and permits obtaining several envelope cycles by simply giving the desired number of cycles as score data. The temporal organization is thus produced by the overlapping of n, n-1, n-2, n-3,. . ., n-(n-1) envelope cycles for

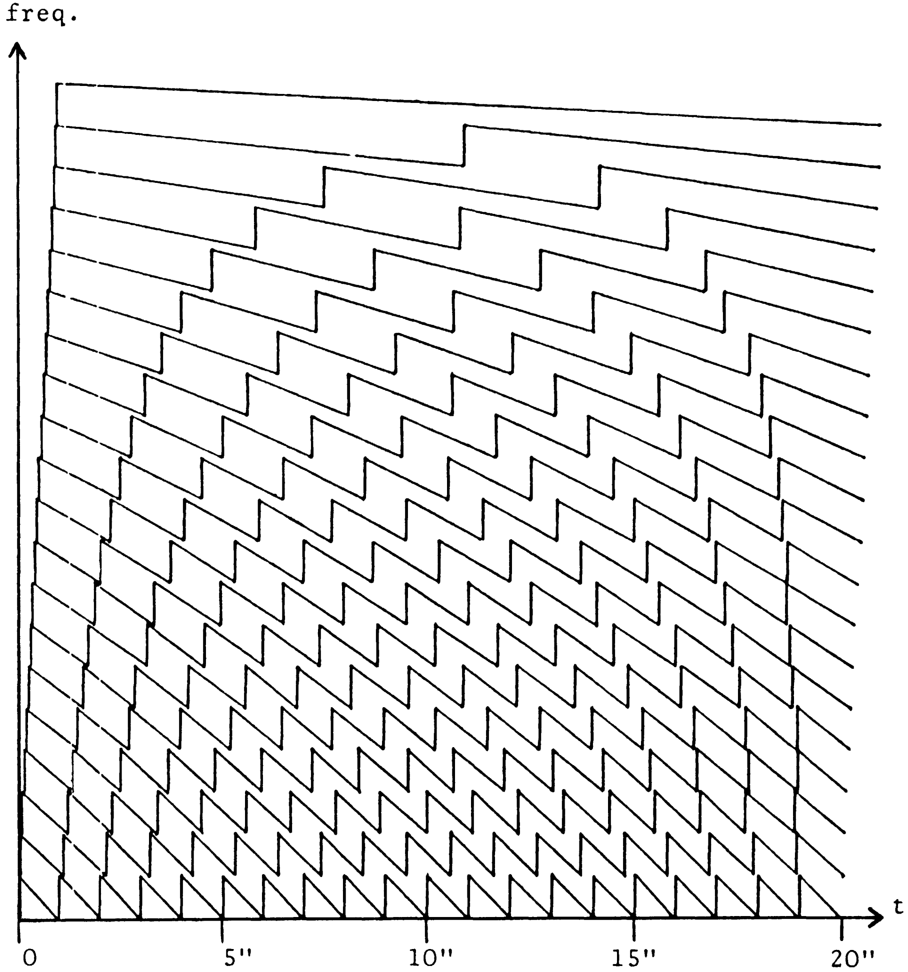


Figure 4

sound number	e. d. in sec.	ratio freq.
1	.95	14.42
2	.90	13.71
3	.85	13
4	.80	12.65
5	.75	11.85
6	.70	11.42
7	.65	10.5
8	.60	9.85
9	.55	9.07
10	.50	8.35
11	.45	7.85
12	.40	6.71
13	.40	6.71
13	.35	5.85
14	.30	4.71
15	.25	3.85
16	.20	3.42
17	.15	2.71
18	.10	2.35
19	.05	1.92
20	0.00	1

different frequencies (with inharmonic ratios between them as can be seen from the y-axis values in fig. 4) whose duration **d** is given by the equation:

$$d_i=T/n_i$$

where i is the voice number
T is the whole event duration in seconds
n_i is the number of attacks of voice i.

The sum of each voice's durations is that of the whole time unit, in this case 20 seconds.

A time delay of .05 second is then added to the entry time between the different voices. In **sound example 4**, a realization of fig. 4, the perception of the structure as a whole is achieved by clearly defining its contours, thus making the shape easy to remember (memory too plays its role in perceptual organization). For this purpose I have used the law of good continuation (cf. sound example 3); the rapid starting succession of tones gives rise to the perception of an ascending motion. The main parameter of this phenomenon is the entry delay time mentioned above; its values should be (for non-overlapping sounds in the frequency range 50 to 3000 Hz) between .05 second and .3 second; values less than .05 second blur the attacks into various degrees of simultaneity, while values greater than .3 second do not have a sufficient cohesive force to avoid the single sounds' being perceived as distinct entities. This is what happens in the middle of the structure when the entry delay, progressively increased, reaches .7 second. The perception of closure (in the definition L. Meyer gives, Meyer, 1956) is then produced by retrograde progression to the starting situation.

In order to enrich the pure sine tone I have modified somewhat the frequency and amplitude parameters; however, the audio signal still remains basically sinusoid because such a signal, more than any other, tends to lose its individual identity in making up the whole as I wished.

Sound example 5 is performed by a Music 5 instrument that, unlike the previous example, has a frequency control function which rises exponentially (fig. 5). This instrument

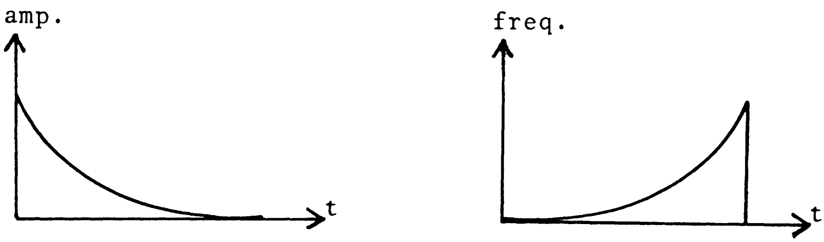


Figure 5

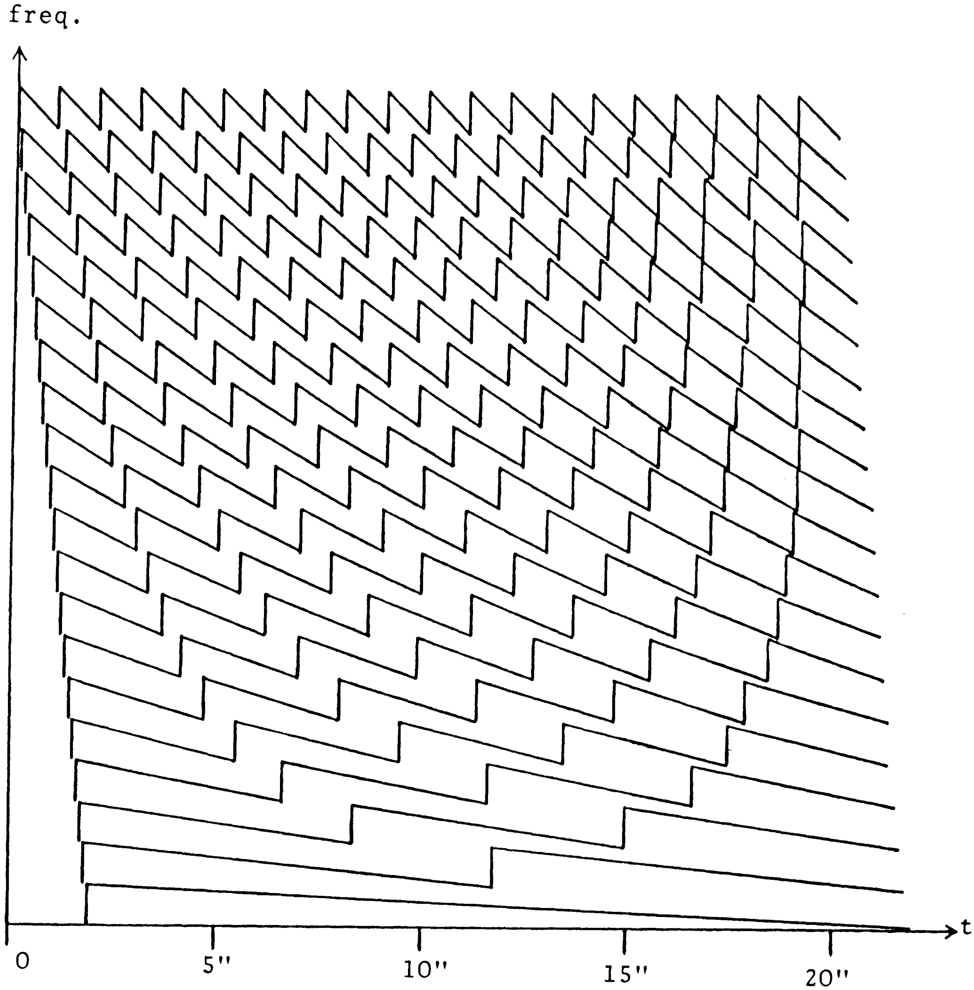


Figure 6

produces the structure in fig. 6: horizontally mirroring fig. 4 but with an entry delay between the voices of .1 second. The frequency deviation is the same for all voices, thus the wideness of the glissandi (of duration equal to that of each note) is inversely proportional to the fundamental frequency of each pitch: a halftone for high frequencies and a minor sixth for low frequencies. Only a few glissandi will be perceived, more exactly those of the tones with longer durations; this does not happen with the other notes because the amplitude is too low by the time the frequency reaches significant deviation values (see function in fig. 5). Glissandi are clearly perceived in the two following examples (both performing fig. 6). In **sound example 6** the glissandi are independent of duration; each voice has only one frequency envelope cycle, the duration of which is that of the entire time unit: 20 seconds.

Fig. 7 shows the function which controls the glissandi in **sound example 7**. It produces a descending-ascending glissando, the range of which (minor third) is the same for all voices. The amplitude controlling function shown in fig. 8b is obtained by multiplying a 6 Hz sine function by a decaying exponential whose duration is equal to that of the note (fig. 8a). The result is used to modify the amplitude of each note in the instrument which performs fig. 4 in **sound example 8**. Although this example is, like the previous ones, monophonic, it gives a sensation of a wide sound source. This happens because of the difficulty the ear has in perceiving the directionality of a sine wave below 4 or 5 kHz.

Sound example 9 presents a structure (fig. 9) that can be seen as the opposite of that in fig. 4; indeed in this case the central 'lines' have been made to stand out by decreasing the starting entry delay in the first half of the structure, and increasing it in the second half. The instrument here used is an oscillator with a sine wave amplitude controlled by an envelope generator with an exponential attack and decay function. A further simple modification was then added to this example: fig 10a shows the frequency spectrum—no longer sinusoidal—of the waveshape for the audio oscillator and the amplitude control function with 'two' attacks: a constant one of .05 second formed by two small line-segments; the second, which will be perceived as the 'real' attack, is obtained with a slope that reaches the amplitude peak at different times (**sound example**

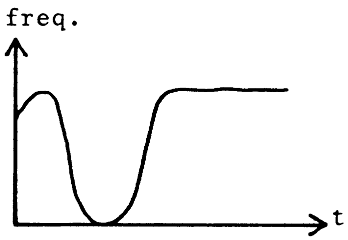


Figure 7

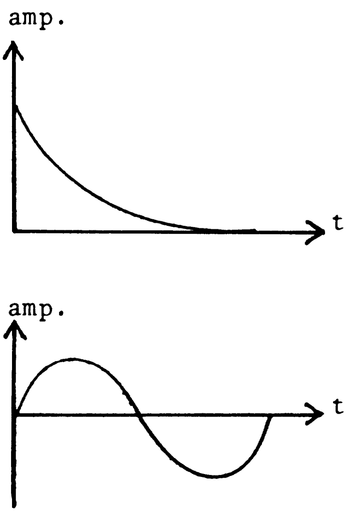


Figure 8a

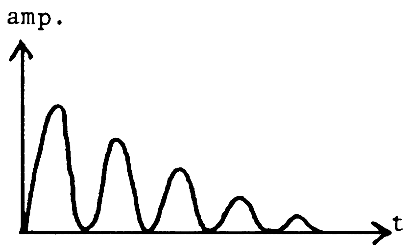


Figure 8b

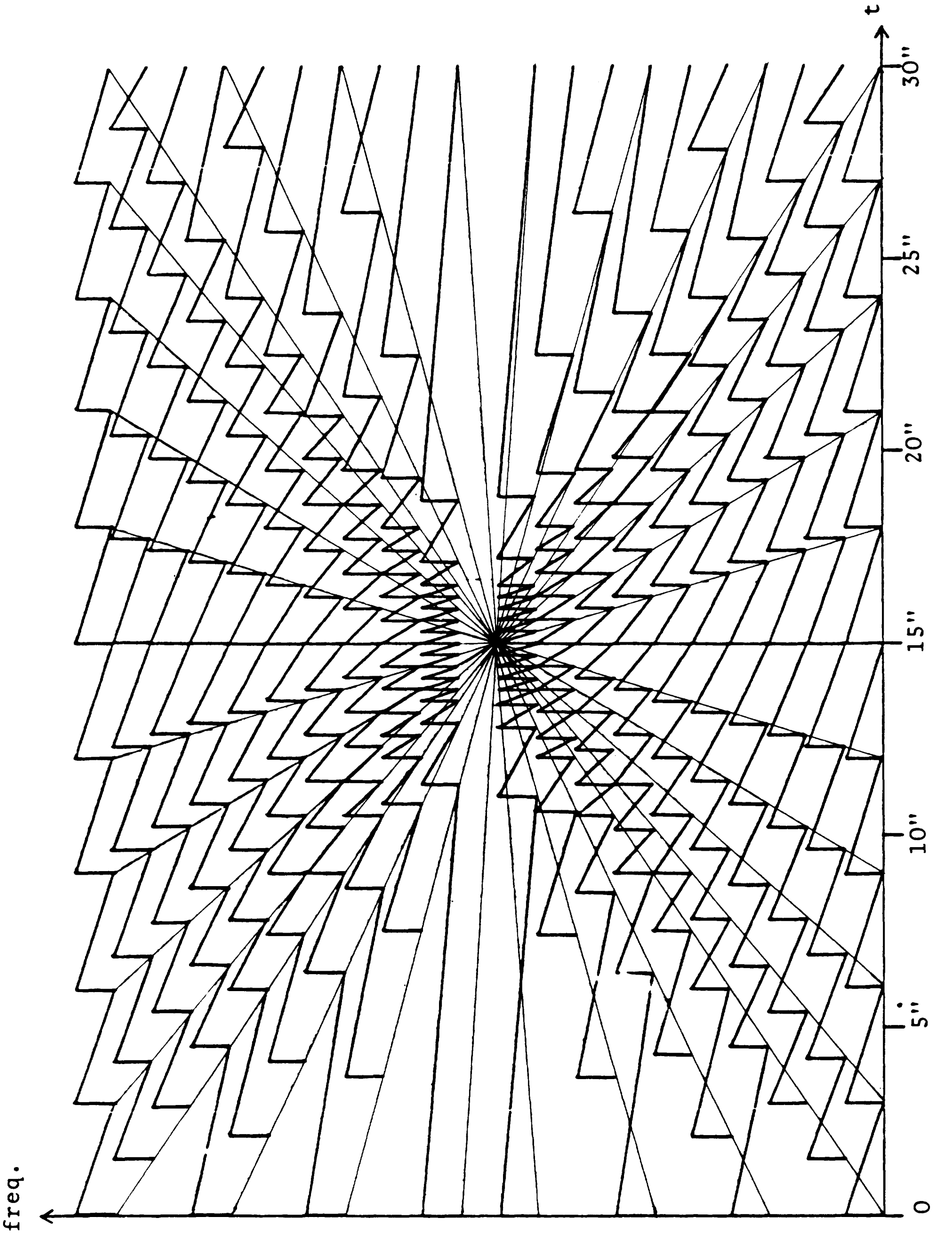


Figure 9

10). Indeed, as shown in fig. 10b, the durations of the second attack and of the decay for each voice do not remain constant. The attack time is increased by .025 second for each single sound while the decay time is reduced until an envelope that is the opposite of the beginning one is obtained: slow attack and sudden decay. This process produces a delay of the perceived onset of the tone compared to the physical one.

The experiences so far described have led to the compositions 'Gioco di velocità' and 'Una pulce da sabbia.'

The macrostructure of 'Gioco di Velocità' (fig. 11) is formed by seven different curves which correspond to the graphic representation, in a two-dimensional space with frequency as the ordinate and time as the abscissa, of the variations of as many circular arcs. This macrostructure determines the formal parameters (repetition series, action time, duration, number of voices, movement in time-frequency space) of seven different kinds of polyphonic rhythmic structures similar to those described above. Some of them are shown in fig. 12. Their graphic generation is performed in two stages: first of all the position of the focus (with the meaning that this word has in visual perspective) is fixed; then the point at which the parallel to the x-axis (each one representing a voice) and the oblique straight lines which converge on the focus meet determines the entry delay (and thus the duration) for each sound.

In the Music 5 instrument used for sound synthesis, the two envelope generators, one controlling the amplitude and the other the frequency of a sine wave, yield two sorts of timbre differentiations between the sine waves so employed. 1. Each structure, besides an amplitude control function (fig. 13a), has its own frequency control function (fig. 13b) which gives a slight deviation of the center frequency (portamento) during the attack time of the individual tones—which is a characteristic of traditional instrument tones. The functions in fig. 14, which I have drawn following the 'Just Noticeable Difference' curve, provide the values for this frequency deviation. 2. The second sort of timbre differentiation is in the dynamic ratio between the durations of the attack transients for the different voices which form each structure. This ratio is reduced both vertically and horizontally. Indeed, in the exposition of the first structure the

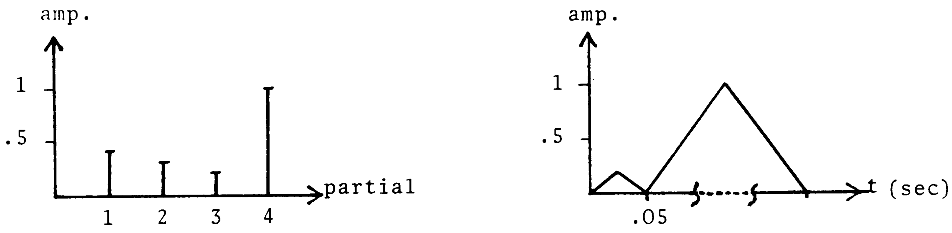


Figure 10a

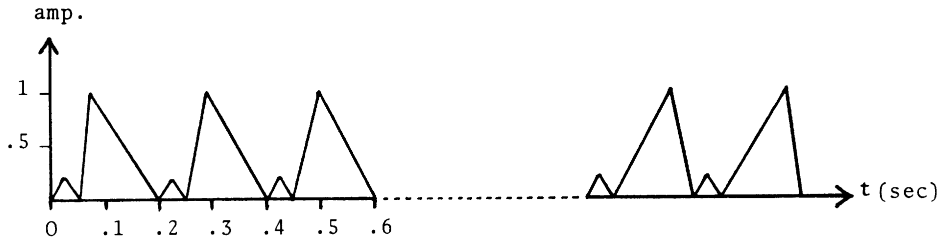


Figure 10b

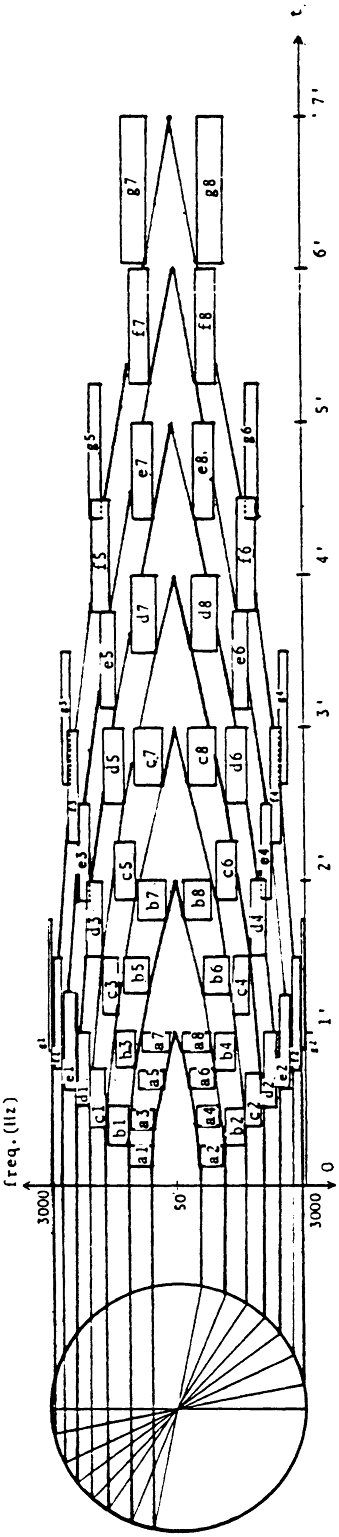

freq. (Hz)
3000
50
3000
0
1'
2'
3'
4'
5'
6'
7'
t
a1
a2
a3
a4
a5
a6
a7
a8
b1
b2
b3
b4
b5
b6
b7
b8
c1
c2
c3
c4
c5
c6
c7
c8
d1
d3
d4
d5
d6
d7
d8
e1
e2
e3
e4
e5
e6
e7
e8
f5
f6
f7
f8
g5
g6
g7
g8

Figure 11

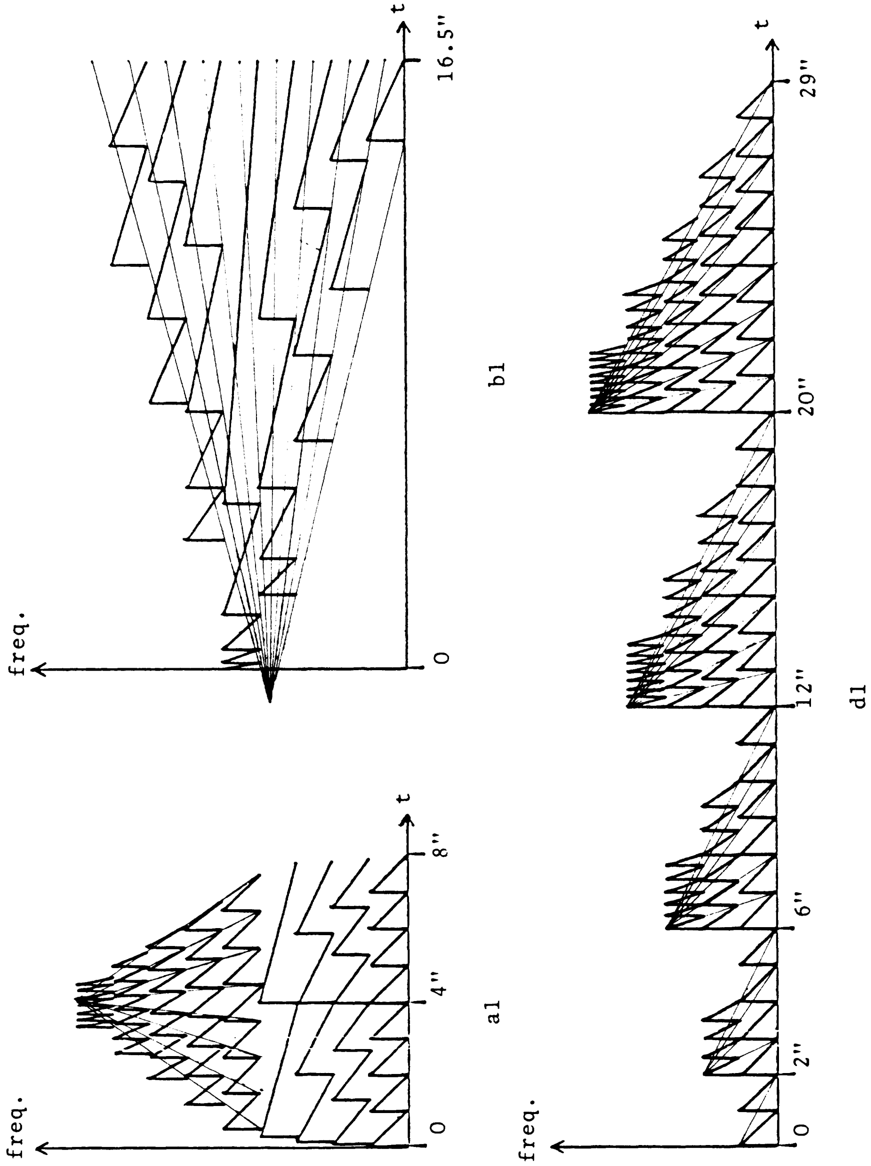


Figure 12

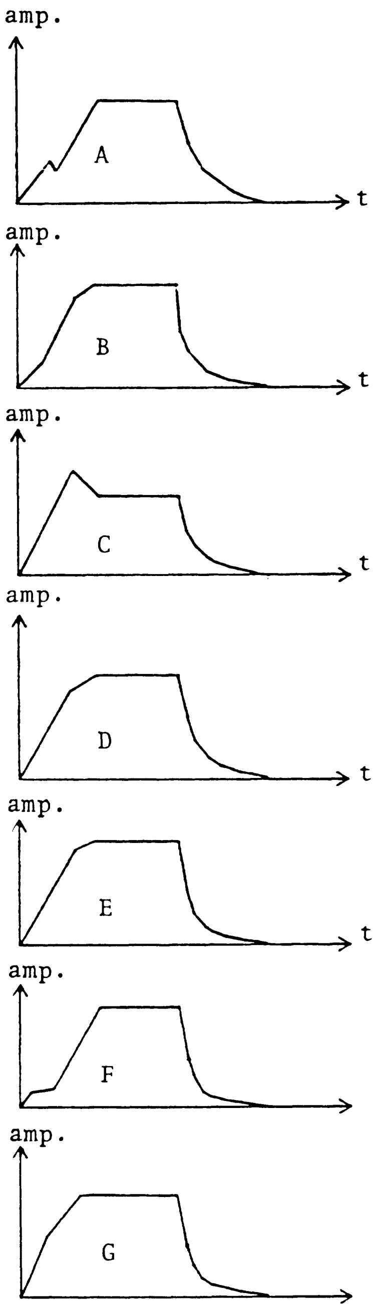


Figure 13a

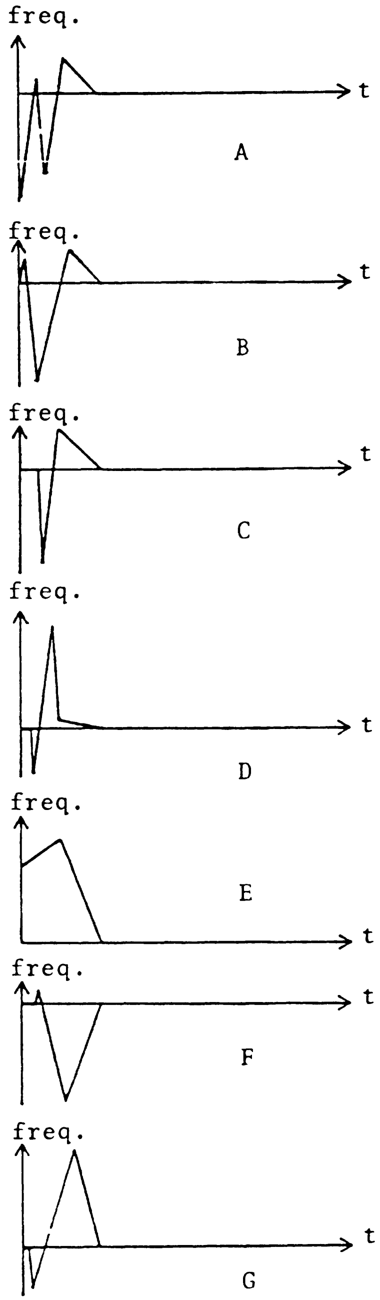
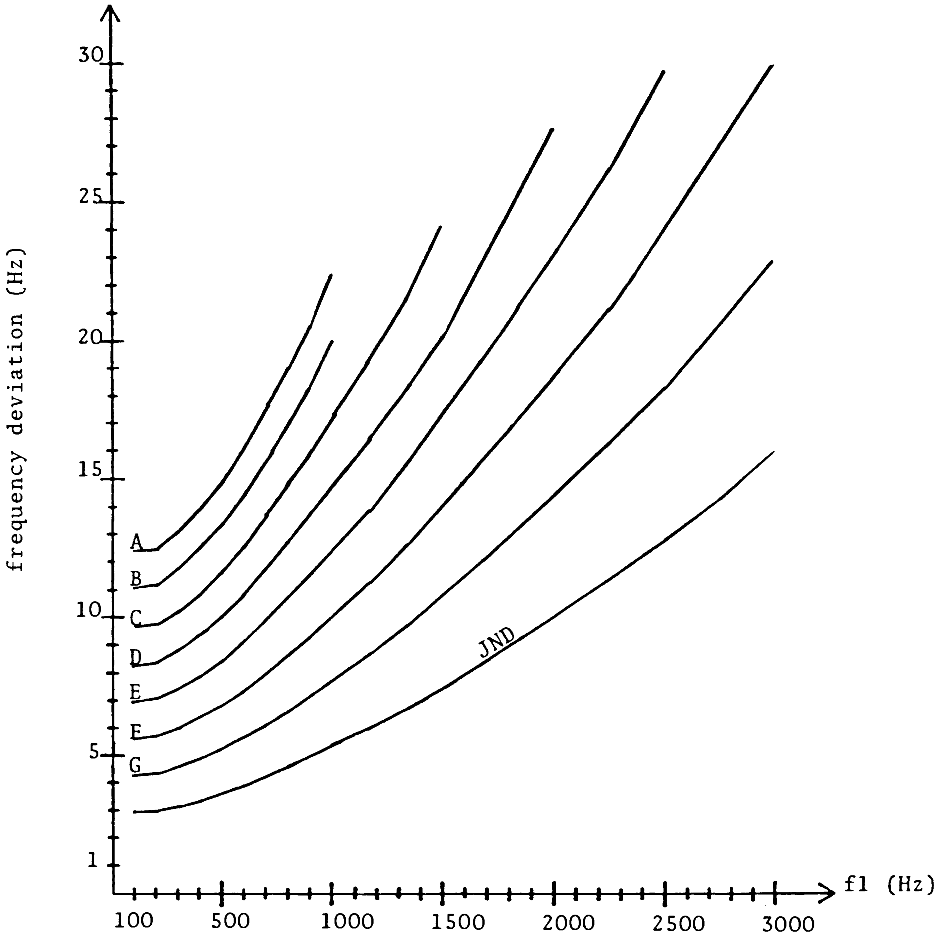


Figure 13b



frequency deviation (Hz)
30
25
20
15
10
5
1
A
B
C
D
E
F
G
JND
100
500
1000
1500
2000
2500
3000
f1 (Hz)

Figure 14

attack time goes from .1 second for the first voice to .01 second for the last, while in the third and last repetition of the same structure for all voices the duration of the attack transient is .01 second.

As regards the distribution of amplitudes for the voices, they have been supplied in Phons (with a Phon-dB conversion)—the structures are considered time-extended complex tones. More precisely, as can be seen in fig. 15, the amplitudes have been fixed according to the formant structure found in the spectra of acoustic instrument tones.

In composing 'Una pulce da sabbia' I have turned my attention even more to timbre. Research on this extremely complex perceptual attribute has recently found in the computer its most powerful investigative tool.

This work is based on the most recent studies of the subject, in particular J. Grey and D. Wessel (Grey, 1975, 1977; Grey, Gordon, 1978; Wessel, 1978, 1979). Experiments in which several subjects were submitted to timbre dissimilarity tests between pairs of synthetic tones (24 or 16 sounds simulating instrument timbres and equalized subjectively for pitch, loudness, and duration) provided the data that have been used by these researchers to build two- and three-dimensional timbre spaces in which each timbre takes its own position along perceptual dimensions related to physical properties of the sounds: I. spectral energy distribution; II. inharmonic energy during the attack time; III. synchronism in the rise time of high order harmonics.

Hence it is possible for the composer to organize at a macrostructural level his own timbre space, a different one for each composition. But the synthesis model used by Grey and Wessel (additive synthesis with amplitude and frequency functions for each partial—about 12—per tone) is not very useful for the great amount of data that the composer has to provide, not to mention the enormous cost involved realizing compositions in this way (unless one has access to Di Giugno 4X!).

What interested me was to see if it was possible, without worrying about orchestral instrument simulation, to maintain the perception of the three dimensions while making a strong data reduction of the synthesis model. For this purpose I used a global synthesis method such as waveshaping (Arfib, 1978; Beauchamp, 1979; Le Brun, 1979) that requires only three

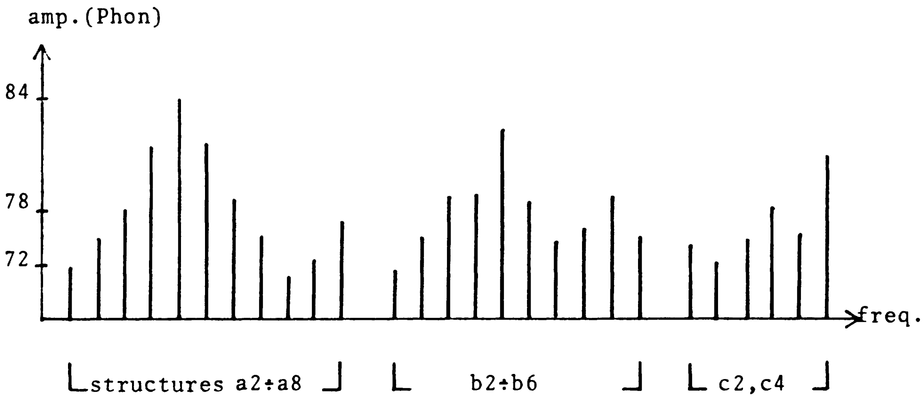


Figure 15

functions for a tone: amplitude envelope; distortion index envelope (which controls proportionally the spectral bandwidth); and transfer function (which determines frequency spectrum).*

A timbre space with fifteen timbres was thus obtained. The first dimension (spectral energy distribution) was still clearly perceivable in spite of the reduced number of harmonics (between 2 and 7) for each spectrum, whereas the second and third dimensions, although recognizable in their extremes, were not perceived to have the timbres as well articulated as had the first dimension. The next step was to build the timbre space for the composition (fig. 16)—the macrostructure of which was already projected—with the three ordered dimensions:

I dimension: spectral energy distribution (concentrated on the lower harmonics at the upper extreme of figure 16, on the higher harmonics at the lower extreme);
II dimension: spectral fluctuation (little at the left extreme, great at the right);
III dimension (coinciding with I): high frequency energy which precedes the full attack of the tone (absent in the upper half, increasing in the lower one).

By means of the synthesis models used (waveshaping and FM) it was possible to control: the first dimension changing the spectral bandwidth and the energy concentration in different spectral regions (in fig. 17, the frequency spectra for the 22 timbres represented in the timbre space); the second dimension with a random fluctuation of the steady state of the distortion index envelope. The third dimension is realized by the superposition of two different attacks as shown in fig. 18: to the sound obtained with waveshaping (a) is added an FM sound (b) with duration equal to the attack time of a, amplitude 25% of a, and with a frequency spectrum (3 or 4 harmonics) centered on higher harmonics of sound a but not always coincident with them because the ratio carrier frequency/modulating frequency is n/.5 (where n=order number of the harmonic of a on which is centered the spectrum of b) (Chowning, 1973).

---

*At the end of my research I was glad to read Charbonneau's article in the **Computer Music Journal** (Charbonneau, 1981) which confirms the need of insight into data reduction perceptual effects. I will be grateful to anyone who could give me information about research on this subject.

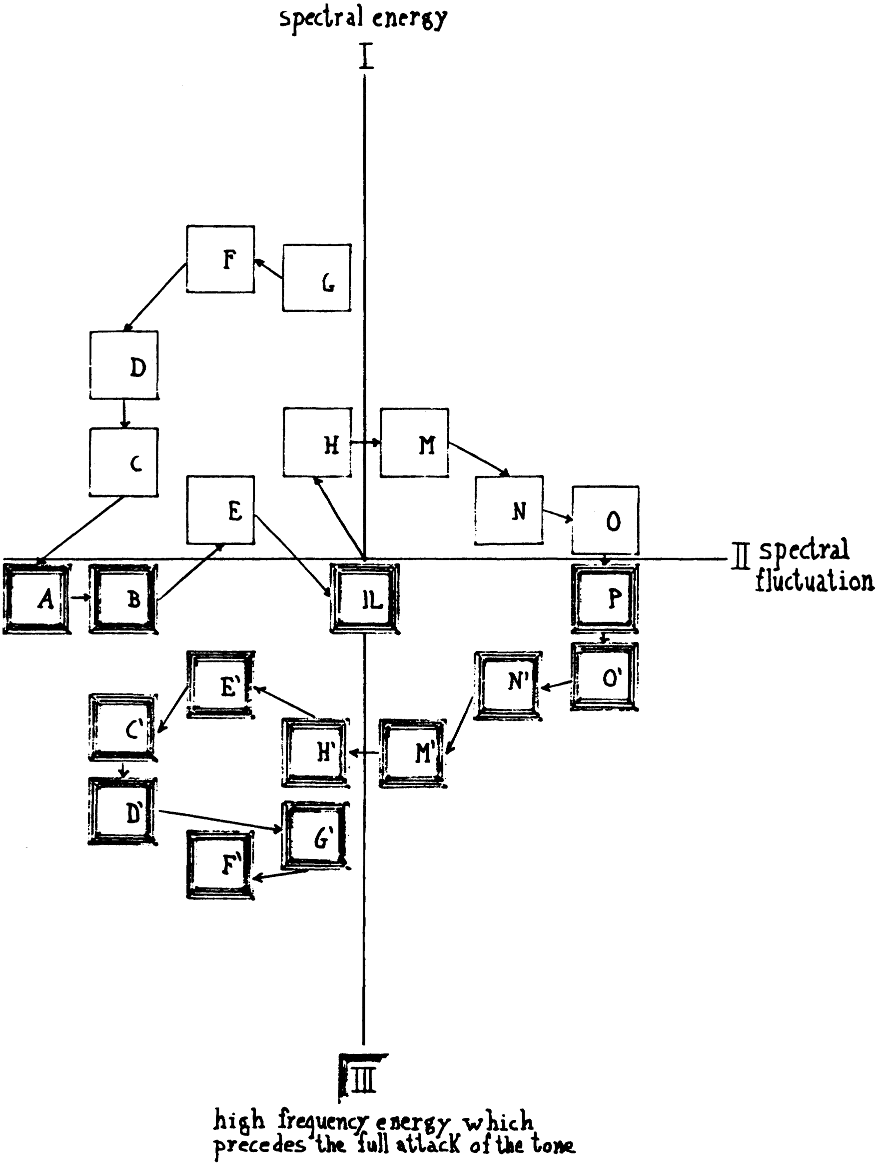

spectral energy
I
F
G
D
H
M
C
E
N
O
II spectral fluctuation
A
B
JL
P
E'
N'
O'
C'
H'
M'
D'
G'
F'
III
high frequency energy which precedes the full attack of the tone

Figure 16

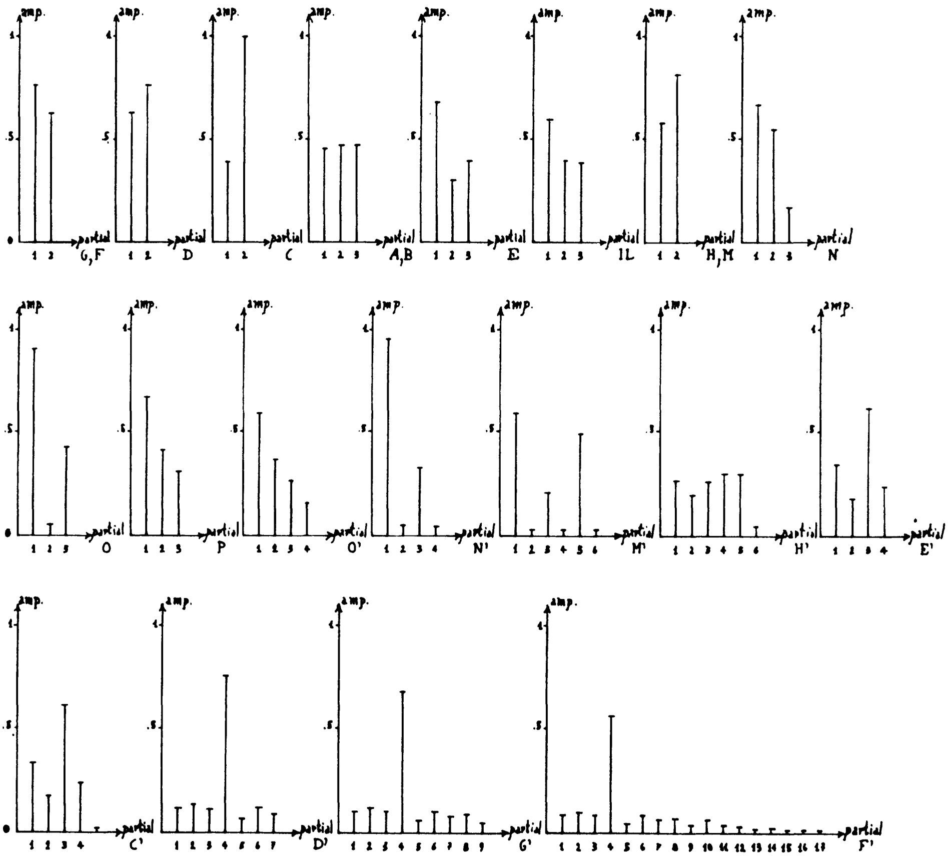


Figure 17

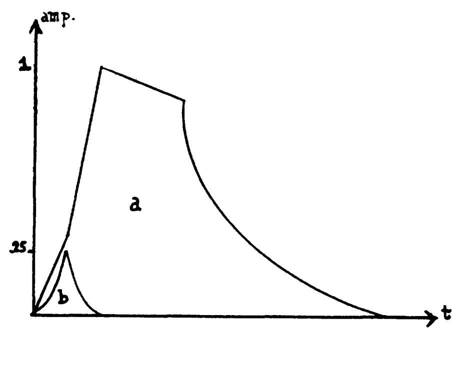


Figure 18

Sound example 11 presents this timbre space following the succession of sounds indicated by the arrows in fig. 16.

Also in this case, as in 'Gioco di velocità', the compositional macrostructure (fig. 19) is given by the projection, in a two-dimensional space with frequency as the ordinate and time as the abscissa, of a figure which is here an architectural structure plan.* By means of the 'slowing perspective' technique I achieved the desired time extension.**

As regards the choice and treatment of the macro-structure, and its internal organization, they depend exclusively on those parameters I have drawn from my previous experiences and which I consider compositional parameters: symmetry, regularity, direction, velocity, 'focus' and 'flight point'. Fig. 20 shows an example of the rhythmic structures which form the macrostructure. Their internal temporal organization is given by the position of the focus, this also determined by the parameters previously mentioned.

Each structure takes the timbre that occupies the corresponding position (indicated by letters) in the timbre space of fig. 16. **Sound example 12** is presented to verify the application of the timbre space to the whole composition. In this example the shape in fig. 20 (an excerpt from the score) is placed on the extremes of the three dimensions G, G', A, P) keeping duration and pitch constant and equalizing loudness.

Conclusion

I have tried to describe how it is possible for the composer to carry out a more precise control of musical forms utilizing knowledge of the sensory and perceptual organization of

---

*More precisely, the plan of the Minerva Medica temple in Rome (260 A.D. and later), but it could have been any other plan. However, one can notice the likeness between this shape and the one used to generate the macrostructure for the preceding composition (a circle). This demonstrates my continuing interest in a shape that gives rise to formal processes which seem to be among the most organic in nature.

**In auditory (musical) perception as well as in visual perception (which is based on temporal organization) the human being listens with a particular kind of 'perspective'; that is, for example, only at the end of a piece of music do we reach a full comprehension of its beginning.

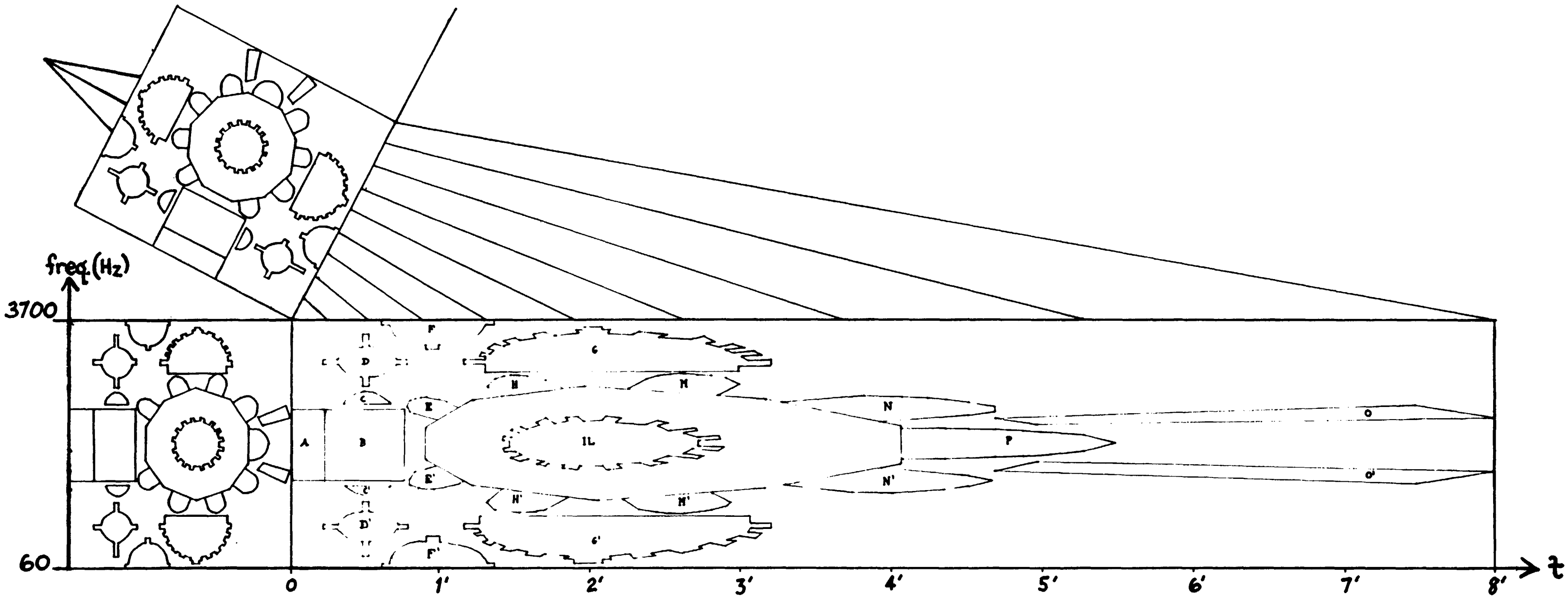


Figure 19

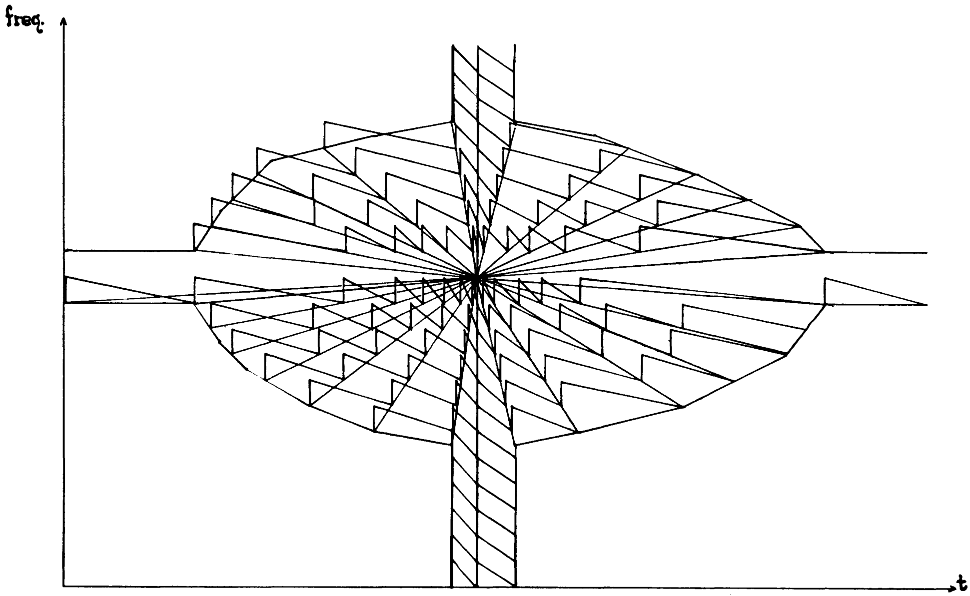


Figure 20

temporal structures. A further development could be a program with a two-dimensional space (not necessarily with time-frequency as coordinates), the position of one or more focuses as input data, and the parametric (thus not only temporal) organization of sounds and structures as output.

Acknowledgements

I would like to thank Mr. Alvise Vidolin for his helpful suggestions during the course of this work, Mr. Graziano Tisato of C.S.C. for his help and support in the technical realization of the compositions and Mr. James Dashow for helping my English.

References

Arfib, D.: Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves. AES preprint n. 1319 (C-2), 1978.

Beauchamp, J.: Brass tone synthesis by spectrum evolution matching with nonlinear functions. **Computer Music Journal**, 1979, 3, n. 2, 35-42.

Bobbitt, J.M.: An experimental study of the phenomenon of closure as a threshold function. **J. Exp. Psychol.**, 1942, 30, 273-294.

Bregman, A.S. and Campbell, J.: Primary auditory stream segregation and perception of order in rapid sequences of tones. **J. Exp. Psychol.**, 1971, 89, 244-249.

Charbonneau, G.R.: Timbre and the perceptual effects of three types of data reduction. **Computer Music Journal**, 1981, 5, n. 2, 10-19.

Chowning, J.M.: The synthesis of complex audio spectra by means of Frequency Modulation. **J. Audio Eng. Soc.**, 1973, 21, 526-534. Reprinted in **Computer Music Journal**, 1977, 1, n. 2, 46-54.

Deutsch, D.: Musical illusions. **Scientific American**, 1975, n. 233, 92-104.

Divenyi, P. L. and Hirsh, I.J.: Some figural properties of auditory patterns. **J. Acoust. Soc. Am.**, 1978, 64, 1369-1385.

Ehrenzweig, A.: **The Psychoanalysis of Artistic Vision and Hearing.** Sheldon Press, London, 1975.

Erickson, R.: **Sound Structure in Music**. University of California Press, Berkeley, 1975.

Fraisse, P.: **Psychologie du rythme.** Presses Univer. de France, Paris, 1974.

______Is rhythm a Gestalt? in **Gestalttheorie in der modernen Psychologie**, 227-232, Ertel-Kemmler-Stadler, 1975.

Fraisse, P. and Piaget, J.: **Traité de psychologie expérimentale**, Presses Univer. de France, Paris, 1963.

Gregg, F.M.: Materializing the ghost of Köhler's Gestalt psychology. **Psychol. Rev.**, 1932, 39, 257-270.

Grey, J.M.: An exploration of musical timbre. Thesis, Stanford University, 1975.

______Multidimensional perceptual scaling of musical timbres. **J. Acoust. Soc. Am.**, 1977, 61, 1270-1277.

Grey, J.M. and Gordon, J.W.: Perception of spectral modifications on orchestral instrument tones. **Computer Music Journal**, 1978, 2, n. 1, 24-31.

Guilford, J.P. and Hilton, R.A.: Some configurational properties of short musical melodies. **J. Exp. Psychol.**, 1933, 16, 32-54.

Guillaume, P.: **La psychologie de la forme.** Flammarion, Paris, 1937.

Heise, G.A. and Miller, G.A.: An experimental study of auditory patterns. **Amer. J. Psychol.**, 1951, 64, 68-77.

von Hornbostel, E.M.: The unity of the senses. **Psyche**, 1927, 7, 83-89.

______Neue Beiträge zur physiologischen Hörtheorie. **Jahresbericht u. d. ges. Physiol.**, 1928, 9, 753-771.

Koffka, K.: Experimental-Untersuchungen zur Lehre v. Rhythmus. **Zeit. f. Psychol.**, 1909, 52, 1-109.

______Perception: an introduction to the Gestalt-Theorie. **Psychol. Bull.**, 1922, 19, 531-581.

______**Principles of Gestalt psychology.** Harcourt, Brace & Co., New York, 1935.

Köhler, W.: Akustische Untersuchungen, III. **Zeit. f. Psychol.**, 1915, 72.

______**Gestalt Psychology.** Liveright Publ. Corp., New York, 1929.

LeBrun, M.: Digital waveshaping synthesis. **J. Audio Eng. Soc.**, 1979, 27, n. 4, 250-266.

Mathews, M.V.: **The Technology of Computer Music**. M.I.T. Press, Cambridge, 1969.

McAdams, S. and Bregman, A.: Hearing musical streams. **Computer Music Journal**, 1979, 3, n. 4, 26-43 (with a cassette of sound examples).

McAdams, S.: Spectral fusion and the creation of auditory images, in M. Clynes: **Music, Mind and Brain: The Neuropsychology of Music**. Plenum, New York, 1982.

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

Moore, B.C.J.: **An Introduction to the Psychology of Hearing**. Academic Press, London, 1982.

Rasch, R.A.: The perception of simultaneous notes such as in polyphonic music. **Acustica**, 1978, 40, 21-33.

Reichenbach, H.: Gestalt psychology and form in music. **J. Musicol.**, 1940, 2, 62-71.

Révész, G.: **Einführung in die Musikpsychologie.** Francke, Bern, 1954.

Risset, J.C.: **Catalog of computer-synthesized sounds.** Bell Labs, Murray Hill, 1969.

Tenney, J. and Polansky, L.: Temporal Gestalt Perception in Music. **Journal of Music Theory**, 1980, 24, 205-241.

Thurlow, W.: An auditory figure-ground effect. **Amer. J. Psychol.**, 1957, 70, 653-654.

Vernon, P.E.: Auditory perception: I. The Gestalt approach; II. The evolutionary approach. **British J. Psychol.**, 1934, 25, 123-139 and 265-283.

Wertheimer, M.: **Productive Thinking.** Harper & Bros., New York, 1945.

Wessel, D.L.: Low dimensional control of musical timbre. IRCAM Report Paris n. 12, 1978 (with a cassette of sound examples).

______Timbre space as a musical control structure. **Computer Music Journal**, 1979, 3, n. 2, 45-52.

Yeston, M.: **The Stratification of Musical Rhythm.** Yale University Press, New Haven, 1976.