Volume 15 Number 2-4 1986 ISSN 0303 – 3902

Interface

Journal of New Music Research

MIDIM/VOSIM Reports

edited by the MIDIM-Group

PUBLISHED BY

SWETS & ZEITLINGER B.V. - LISSE

INTERFACE – Journal of New Music Research – is an international journal devoted to discussion of all questions which fall into the borderline areas between music on the one hand, physical and human sciences or related technologies on the other. New fields of research, as well as new methods of investigation in known fields, receive special emphasis.
In addition, shorter communications concerning the description of new apparatus, compositional devices, short interviews, descriptions of new computer programs, interim reports on current activities in electronic studios, institutions, etc. are eligible for publication.

It is published quarterly, four issues of approx. 64 pages each forming one annual volume.

Authors are invited to send manuscripts and all communications related to the following address:
The Editors of Interface
c/o Swets Publishing Service
347b, Heereweg
2160 AH LISSE, The Netherlands

All correspondence concerning subscriptions and reprints should be sent to the publishers:
Swets Publishing Service
347b, Heereweg
2160 AH LISSE, The Netherlands

Printed in The Netherlands by Offsetdrukkerij Kanters B.V., Alblasserdam
Type-setting Hoi Studio, Rotterdam

Interface, Vol. 15 (1986), pp. 69-70. 0303-3902/86/1502-0069 $3.00

Foreword

The issue which you have before you is consecrated to the VOSIM/MIDIM-system for digital musical composition. The models for VOSIM sound synthesis were published for the first time more than ten years ago in Interface (Vol. 2 and 3). It seemed natural to us therefore that a comprehensive report concerning the developments which VOSIM has led to in the meantime would be entrusted to Interface. These developments began with the design of soft- and hardware generators for the purpose of testing experimentally the VOSIM7-model and to offer composers as well as researchers an effective tool with which to work. Since VOSIM-generators of a sufficient expressive power display a relatively high number of control parameters, this gave rise quite early to the demand for a suitable control system, without which neither a systematic exploration of the model nor its applicability for artistic and scientific aims could be taken into consideration. As a solution the MIDIM-language was developed. Since 1977 at the Institute for Sonology, Utrecht, both VOSIM and MIDIM were joined together to form a unified system, which immediately attracted lively interest from composers and researchers and has led to many-facetted results (and even to the formation of a MIDIM-Group). This shall be spoken about in the following pages.

The six articles presented do not form a MIDIM mini-monograph; they are considered as no more than a collection of reports, in which our work will be reported upon in so factual but many-sided a way as possible. The initial article explicates in short for the reader the origins and principles of VOSIM sound synthesis and makes clear the necessity for an intelligent control system. In the following extensive article (pg. 83) is shown how such a control system can be constructed by means of a formal MIDIM-language; with which not only the syntax and semantics within this formal system shall be presented in detail but also extreme emphasis will be placed upon the interpretability of this formal system in the field of its application. This last, particularly important aspect could give rise to numerous polemical confrontations; we preferred, however, to present in a factual way the problems of duplication and pattern recognition (which are closely related to the above) in two fields of application: Gamelan music (Janssen/Kaegi, pg. 185) and Dutch linguistic interjections (van Berkel, pg. 231). The theme of pattern recognition is taken up once more in the last article (P. Kuipers, pg. 257) and it is shown how the concepts of the abstract and

All manuscripts received April 24, 1986.

the comprehensor as defined in the MIDIM-language can be further developed in the direction of a theory of musical "forms". A contribution by P. Goodman (pg. 163) will finally draw the readers' attention to the fact that our activities are not exhausted by theoretical speculations but are directed to giving fruition in a creative-artistic manner to the insights obtained and to transmit these further to young people.

F-30830 Aubais, September 1986 Werner Kaegi

Interface, Vol. 15 (1986), pp. 71-82. 0303-3902/86/1502-0071 $3.00

Controlling the VOSIM Sound Synthesis System

Werner Kaegi

ABSTRACT

This paper gives first a brief glance at an alternative system for digital sound synthesis which is known under the name VOSIM. The cycle of a VOSIM time function consists of a series of sin^2-pulses. The amplitude of the successive pulses decreases uniformally. After the last pulse there is a delay. The main parameters for describing a VOSIM signal are: pulse width, number of pulses, rate by which the pulses decay, length of delay and overall-amplitude. The model suggests additional parameters which have to do with how the sound changes as it progresses. VOSIM is particularly suitable for producing sounds with formants such as speech and musical sounds. VOSIM oscillators are intelligent ones and as a consequence they are complex to control. For this reason the MIDIM control language has been developed and will be discussed briefly in the last part of this paper. For a more comprehensive presentation of the MIDIM language cf. pp. 83-161 of this issue.

1. INTRODUCTION

Whereas the synthesis-networks in the early days of computer application in music were the work of pioneers, the market is now dominated by industry products. The advantage of the commercial development is certainly the standardization in technical aspects (e.g. the compatibility with the Midi-bus). The disadvantages are the low-level program languages, the uniformity and weakness of abstraction which narrow the field of musical interpretation and, as a consequence, may have an ill effect on the artistic quality of the music and its further development. And the idea frequently comes into play that the ideal tool is the real-time syntheziser and that the more involved and more artistically interesting systems cannot participate before they will achieve the speed of "real-time". Thus everything tends towards real-time performance, which means towards a playable instrument while the aspects of musical composition are in danger of being devalued. We should, however, not forget that musical composition has always had to do with premeditation and can never be confused with improvisation. By this I by no means doubt the value of intelligent interactive systems. Already in the 70's I was aware that VOSIM is particularly suitable for real-time synthesis by means of a *syntheziser* (cf. Buxton 1977, p. 68). However, for artistic as well as scientific reasons I felt more attracted to setting up a *compositional language.*

2. THE VOSIM SOUND SYNTHESIS

As early as 1966, when working at Geneva Radio, I had thought of the application of tone burst signals for musical means. I suspected that it might become an elegant and powerful alternative to the conventional methods for sound synthesis used at that time. This was the case even though I only knew the graph of the amplitude-envelope in frequency-domain for the $\sin^2$ tone burst signal with $N=1$ (cf. Kaegi 1967, p. 62/63) and experimentation was not possible due to the lack of a tone burst generator. In 1970 I submitted a project to the Fonds National Suisse and a few months later was given the necessary funds in order to embark on my project* at Sonology in Utrecht, where I had the possibility of working with two analog tone burst generators. Very soon I was familiar enough with the $\sin^2$ tone burst as to produce speech- and instrumental sounds as well as singing which I presented for the first time at the Saarländische Rundfunk Saarbrücken on December 12, 1971 and later on the 19th of October 1972 at the Goethe Institute in Amsterdam. In 1973, in collaboration with Stan Tempelaars, I opened a second research project the purpose of which was to gain an insight into the parameters of the signal function (by means of tests applying Kruskal multidimensional scaling) as well as into the representation in the frequency domain. The model which later was given the name VOSIM (VOice SIMulation) was published for the first time in the same year (Kaegi 1973) and an extended version (Kaegi 1974) was presented in 1976 by Kaegi and Tempelaars at the AES conference in Zürich. In 1978 followed the publication in the AES journal (Kaegi/Tempelaars 1978). According to the terminology applied by Digital Music Systems inc. (Wallraff 1983) and by Robert Moog (Moog 1984) the VOSIM belongs to the alternative systems.

The VOSIM time function

The time function stated by the VOSIM model is a tone-burst signal which is built from out a series of $\sin^2$ pulses each of the same time duration and with a decreasing amplitude (with a staircase envelope) (Fig. 1).

The refined model (Kaegi 1974) recognizes the following variable vector of the 12 control parameters:

$$(T, \Delta T, M, \Delta M, A, \Delta A, C, N, D, S, MF, Np) \qquad (1)$$

where:

T (duration of the $\sin^2$ pulses)
ΔT (in- or decrement of T in Np periods)
M (delay)

* Fonds National Suisse project 1.559-0.71, final report of November 1st, 1975.

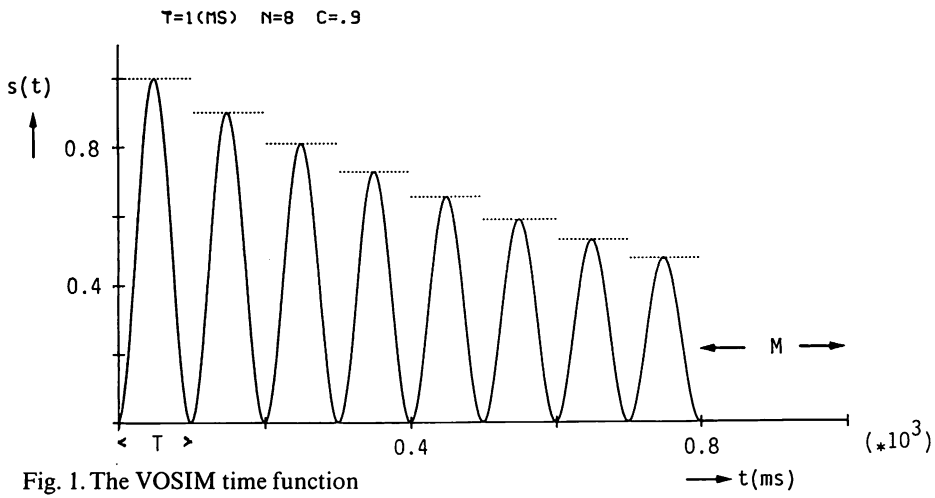


Fig. 1. The VOSIM time function

ΔM	(in- or decrement of M in Np periods)	
A	(amplitude of the first pulse)	
ΔA	(in- or decrement of A in Np periods)	
C	(decay factor = proportion of amplitudes of two subsequent pulses)	
N	(number of pulses in tone-burst)	
D	(deviation = max. change of M)	M-modulation parameters
S	(type of modulation (sin/P-random))	
MF	(speed of modulation)	
Np	(number of periods)	

In what follows the VOSIM model provided with the above input parameters will be called the VOSIM7 model. I tested still other versions of the VOSIM model, e.g. one recognizing amplitude- as well as T-modulation. However, in what follows we will not enter into discussion but concentrate on the VOSIM7 model.

The Spectrum of the VOSIM Signal

The VOSIM model presents a very characteristic property: the time function only recognizes one constant waveform: the $\sin^2$ (Fig. 1). Other properties of the signal may be found in the spectrum (Fig. 2).

The Spectrum of the tone-burst is continuous ($Np=1$). When the signal is repeated ($Np>1$), a discrete spectrum occurs but it exhibits the envelope of the continuous spectrum. The envelope is thus *not* dependent upon the repetition frequency of the signal. An analytic formula is to be found in the pertaining literature (Tempelaars 1977, Kaegi/Tempelaars 1978). I will limit here the discussion only to some of the most characteristic properties of the VOSIM magnitude spectrum.

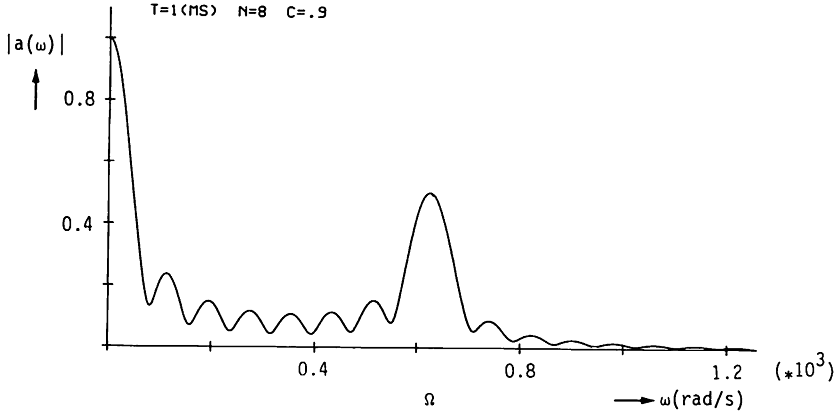


Fig. 2. The Magnitude Spectrum of the VOSIM signal shown in Fig. 1.

(1) In the envelope of the spectrum there is a peak at about Ω, where:

$$\Omega\ (\cdot 10^3 \text{rad/s}) = 2\ \pi/T\ \ (T \text{ in ms}) \tag{2}$$

Frequencies in the neighbourhood of Ω are thus accentuated. In other words, there is a *formant-area* which depends upon T (Fig. 2).

(2) The amplitude value at 2Ω is always 0, thus:

$$|a(2\Omega)| = 0 \tag{3}$$

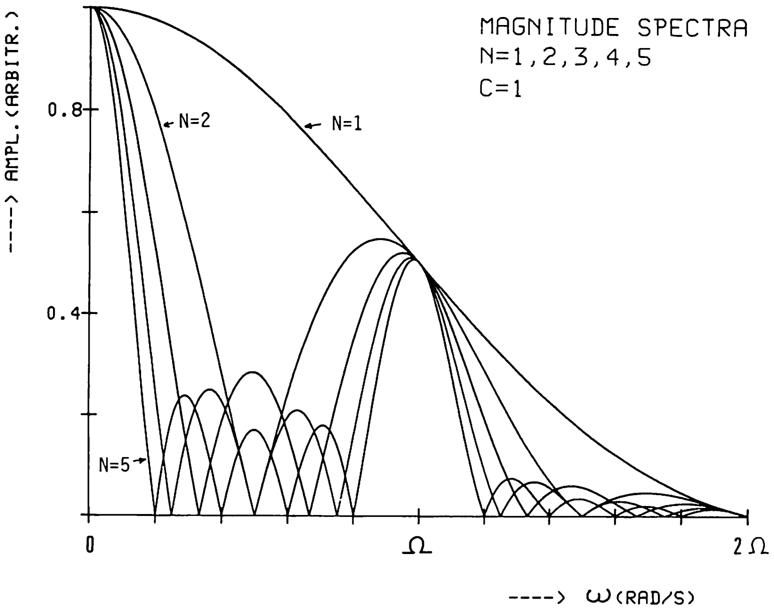
If we assume $|a(0)| = 1$, then the amplitude value at Ω will always be 1/2, thus:

$$|a(\Omega)| = |a(0)|/2 \tag{4}$$

(3) Within the ω-interval $[0.2\Omega]$ there may occur $2(N\text{-}1)$ minima at $i\cdot\Omega/N$, i=1, 2,... and $i \neq N$ (Fig. 3). Whether they will occur and which will be their amplitude values depends upon the decay C. For C=1 (=100%) the value of all minima is 0 (Fig. 4). Above 2Ω the spectrum may be ignored.

(4) For the discrete spectrum it may be interesting to express Ω (where about the peak occurs) as well as $i\cdot\Omega/N$ (where the minima may occur) by means of a function of the repetition frequency:

$$f_1(\cdot 10^6 \text{Hz}) = 1/T' = 1/(N\cdot T+M) \quad (T, M, T' \text{ in } \mu\text{s}) \tag{5}$$



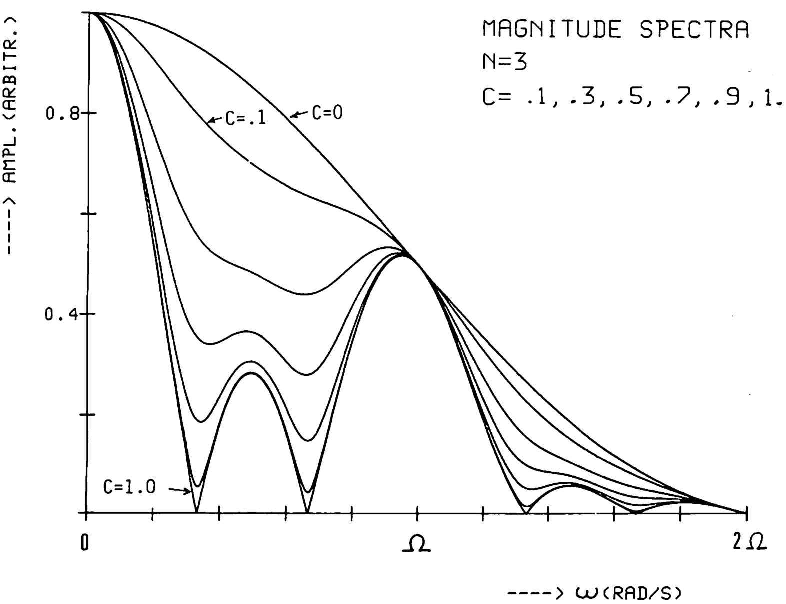
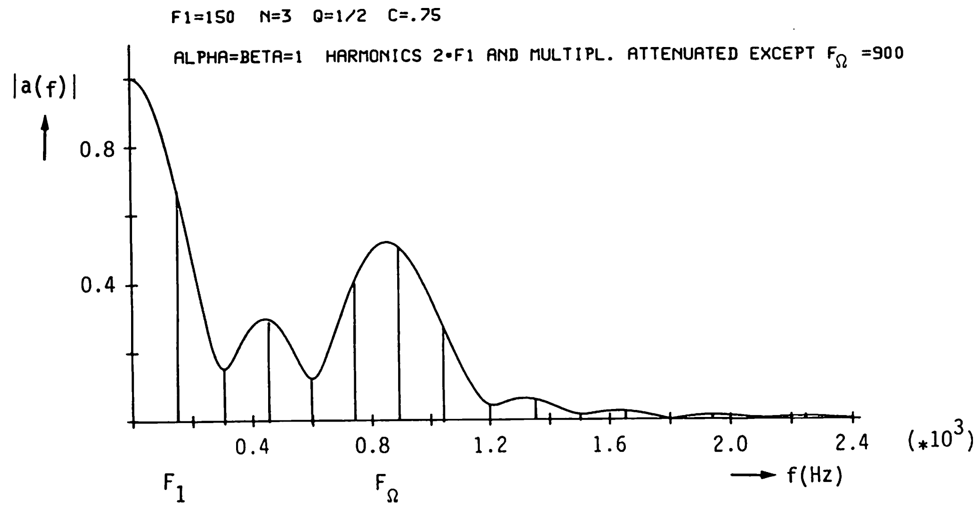


Fig. 3 and 4. Continuous VOSIM magnitude spectra dependent upon *N* (top) and upon *C* (bottom).



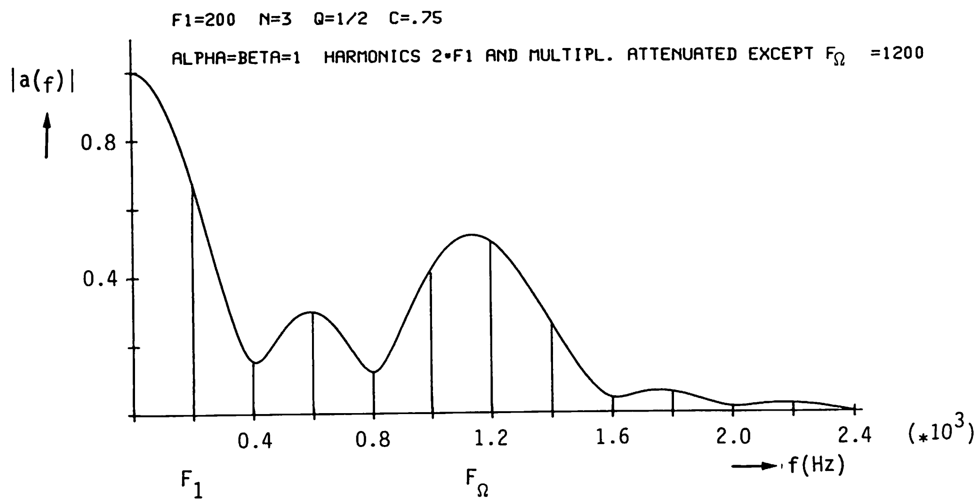


Fig. 5 and 6. Discrete VOSIM magnitude spectra with attenuated harmonics (F_2 and its multiples except F_Ω).

Let us therefore express the ratio $M/(N \cdot T)$ by the smallest possible integers α and β:

$$\frac{\beta}{\alpha} = \frac{M}{N \cdot T} \tag{6}$$

and then introduce q as follows:

$$q = \frac{\alpha}{\alpha + \beta}\left(= \frac{N \cdot T}{N \cdot T + M}\right) \quad 0 < q < 1 \tag{7}$$

We now obtain the following functions for the envelope:

Peak at:

$$F_{\Omega}\,(\mathrm{Hz}) = N \cdot \mathrm{f}_1 \cdot \left(1 + \frac{\beta}{\alpha}\right) = N \cdot \mathrm{f}_1 \cdot \mathrm{q}^{-1} \quad \text{with } N > 1 \tag{8}$$

Possible minima at:

$$F_{\mathrm{m_i}}(\mathrm{Hz}) = \mathrm{i} \cdot F_{\Omega}/N = \mathrm{i} \cdot \mathrm{f}_1 \cdot q^{-1} \quad \mathrm{i} = 1, 2, \ldots \text{ and } \mathrm{i} \neq N \tag{9}$$

(5) Let us still have a look at the harmonics:

$$\mathrm{f_n}\,(\mathrm{Hz}) = \mathrm{n} \cdot \mathrm{f}_1 \quad \mathrm{n} > 1 \tag{10}$$

Some of the harmonics may coincide with possible minima. For which ones this will be the case depends, of course, upon α and β as follows:

$$\mathrm{n} = \mathrm{i} \cdot (\alpha + \beta) \quad \mathrm{i} = 1, 2, \ldots \text{ and } \mathrm{i} \neq N \tag{11}$$

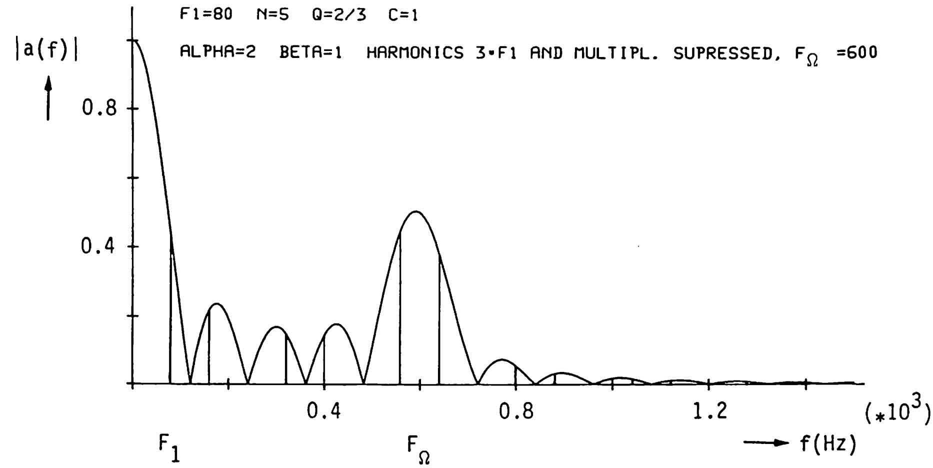
The coinciding harmonics are thus $\mathrm{f}_{(\alpha+\beta)}$ and its multiples. According to the value of C these harmonics will be attenuated or entirely suppressed ("missing harmonics") except for $\mathrm{i} = N$. Cf. Fig. 5-8. Notice that for constant q, C and N the amplitude ratio of the harmonics will be the same for any T'.

With the above VOSIM model it was possible to duplicate vowels, fricatives and plosives as well as musical instruments and singers by synchronizing two VOSIM functions. Sequences in terms of spoken language and musical phrases synthesized by the author proved, already, at this early stage the power of the model. Here is an example which I made in 1975: The following sequence of pairs of V-vectors (= VOSIM7 file) will produce in sound the name "Stan Tempelaars". (Since the example is an excerpt of a file "Guten Tag Professor Stan Tempelaars", the /t/ is aspirated in a german way.)

	T	ΔT	M	ΔM	D	A	ΔA	C	N	S	MF	Np
S	20.	0.	80.	0.	40.	10.	240.	75.	1.	0.	100.	200.
	20.	0.	80.	0.	40.	0.	0.	75.	1.	0.	100.	200.
	20.	0.	80.	0.	40.	250.	-75.	75.	1.	0.	100.	200.
	20.	0.	80.	0.	40.	0.	0.	75.	1.	0.	100.	200.
T	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
	20.	0.	200.	0.	199.	250.	0.	75.	1.	0.	88.	90.
	20.	0.	200.	0.	199.	0.	0.	75.	1.	0.	88.	90.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
H	100.	0.	400.	0.	360.	10.	115.	100.	1.	0.	60.	61.
	100.	0.	400.	0.	360.	0.	0.	100.	1.	0.	60.	61.
	100.	0.	400.	0.	360.	125.	0.	100.	1.	0.	60.	62.
	100.	0.	400.	0.	360.	0.	0.	100.	1.	0.	60.	62.
A	240.	0.	3805.	-600.	160.	511.	0.	75.	4.	1.	31.	32.
	98.	0.	3724.	-600.	160.	50.	0.	100.	10.	1.	31.	32.
N	420.	0.	2005.	0.	0.	250.	0.	75.	3.	0.	24.	25.
	560.	0.	2564.	0.	0.	60.	0.	75.	2.	0.	24.	25.
T	0.	0.	1000.	0.	0.	0.	0.	0.	-1.	0.	0.	50.
	0.	0.	1000.	0.	0.	0.	0.	0.	-1.	0.	0.	50.
	20.	0.	200.	0.	199.	300.	0.	75.	1.	0.	88.	90.
	20.	0.	200.	0.	199.	0.	0.	75.	1.	0.	88.	90.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
H	100.	0.	400.	0.	360.	10.	115.	100.	1.	0.	60.	61.
	100.	0.	400.	0.	360.	0.	0.	100.	1.	0.	60.	61.
	100.	0.	400.	0.	360.	125.	0.	100.	1.	0.	60.	62.
	100.	0.	400.	0.	360.	0.	0.	100.	1.	0.	60.	62.
E	650.	0.	205.	0.	200.	511.	0.	75.	2.	1.	12.	8.
	90.	0.	1485.	0.	200.	50.	0.	100.	14.	1.	12.	8.
M	840.	0.	2045.	-600.	0.	250.	0.	75.	1.	0.	10.	13.
	560.	0.	3165.	-600.	0.	200.	0.	75.	1.	0.	10.	13.
P	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	150.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	150.
	420.	0.	40.	0.	0.	200.	0.	75.	2.	0.	1.	1.
	140.	0.	40.	0.	0.	75.	0.	100.	6.	0.	1.	1.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	5.
Є	490.	0.	2970.	600.	0.	400.	-100.	75.	4.	1.	5.	6.
	65.	0.	3531.	600.	0.	15.	0.	100.	28.	1.	5.	6.
L	550.	0.	2611.	-600.	120.	80.	0.	75.	4.	1.	14.	15.
	90.	0.	3851.	-600.	120.	13.	0.	100.	21.	1.	14.	15.
A	240.	0.	3131.	600.	160.	250.	250.	75.	-8.	1.	12.	13.
	98.	0.	2971.	600.	160.	30.	0.	100.	-20.	1.	12.	13.
R	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
	150.	0.	360.	0.	359.	150.	0.	75.	1.	0.	17.	18.
	150.	0.	4000.	0.	0.	0.	0.	75.	3.	0.	2.	3.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	20.
S	20.	0.	80.	0.	40.	150.	50.	75.	1.	0.	100.	800.
	0.	0.	160.	0.	0.	0.	0.	0.	1.	0.	0.	800.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	2000.
	0.	0.	1000.	0.	0.	0.	0.	0.	1.	0.	0.	2000.

(Domains: ΔT (in $4\mu s$) = [—2047,2047], $T(t)$ (in $4\mu s$) = [20,2047]
ΔM (μs) = [—4095,4095], $M(t)$ (μs) = [1,4095]
ΔA = [—511,511], $A(t)$ = [0,511], 9 bit
D (μs) = [0,4095]
C (%) = [0,100]
MF (in periods of the carrier)
S = [0,1], 0 for random, 1 for sin-modulation)

A vector where $T = 0$ occurs stands for a rest of the duration d (ms) = Np.



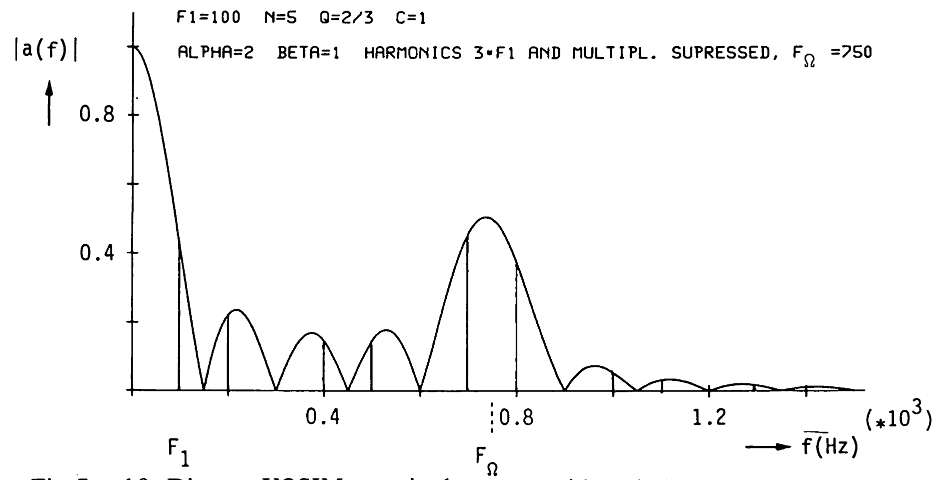


Fig. 7 and 8. Discrete VOSIM magnitude spectra with entirely suppressed harmonics (F_3 and its multiples except).

3. CONTROLLING THE VOSIM SYSTEM

In the beginning I worked with voltage-controlled generators. In order to acquire greater stability and accuracy and to have a protocol at my disposal I stepped over to a solution making use of digital means. At first this consisted of two soft-ware oscillators (programmed by Tempelaars) and very soon of two digital hard-ware oscillators (designed by Kaegi/Scherpenisse/Tempelaars).

The oscillators of this first generation input the values for *T*, *M* and *A* (calculated in turn by the computer and stored in a buffer) and output single $\sin^2$-pulses followed by a delay (equal to zero for *N*-1 pulses and equal to *M* for the *N*th pulse). The oscillator was presented at the First Internat. Conference on

Computer Music Boston 1976; a detailed description is to be found in Kaegi/-Tempelaars 1978. In what follows the soft- and hardware involved will be called the VOSIM7-system. I mention yet the third generation of VOSIM-oscillators (designed by J. Scherpenisse) where the calculation time is sped up by using a microprocessor; oscillators of this type build up the whole sequence of *Np* VOSIM periods indicated by the input vector.

A VOSIM system is an intelligent sound source. In order to have an idea of its intelligence we may look at the input which is an $m \times n$-list of *m* vectors (or control lines) of the dimension $n=12$ (*m* depends, of course, on the domains of the parameters as well as on the sampling rate). If we would print this list with a linefeed of one centimeter (assuming a sampling rate of 25 kHz), the length of the listing would be about 10^{32} kilometers. Think then in terms of light years! We know, however, that the set V^* of all combinations of all vectors of the list *V* will embrace the totality of all sounds which the VOSIM-oscillator can produce. But how can we make a choice? It is impossible to select the vectors from the list "by hand". This task can only be accomplished by means of a directional strategy.

Already in 1973/74 I formulated *rules* for the production of linguistic sounds by VOSIM, and soon I did the same with the sound of wind instruments (clarinet, bassoon, flute, hautbois, english and french horns, recorder etc. by rules). In this way I was able to denote subsets of VOSIM vectors retrieved from the VOSIM input list or matrix *V*. These experiences led me to a more general solution as follows: to build up over the matrix *V* of the VOSIM system a *formal language* by which means I will be able to define *any* subset of vectors I may wish. This language is called MIDIM (MInimum DescrIption of Music).

How does MIDIM operate?

Let us think of the VOSIM input again in terms of the constant $m \times n$-matrix *V* consisting of *m* lines or VOSIM-vectors each of them exhibiting *n* constant coefficients. This matrix may be represented by the above variable vector (1) provided the domains will be defined for each parameter involved. Let us express this by

$$V = \lambda x_1, \ldots, x_{12}. (x_1, \ldots, x_{12}) \tag{12}$$

where the operand $(x_1, \ldots, x_{12})$ stands for the variable VOSIM-vector (1) and where the operator $\lambda x_1, \ldots, x_{12}.$ will define the domain of each variable listed. Our main task is now to define subsets of V^*. We shall do it by operating upon the variable vector (12) as follows (cf. p. 85):

(1) *substituting* functions into the variable coefficients;

e.g. $F_3 = T' + Of - N \cdot T$ into x_3 $(= M)$
($T' + Of$ for the duration of the period (T, Of, T' (us)))
$F_{12} = rof(d \cdot 10^3 / (T' + Of + \frac{1}{2}(\Delta M + N \cdot \Delta T)))$ into x_{12} $(= Np)$
(d for the duration of the signal (ms), rof for rounding off)

Accordingly we will obtain the following vector $\vec{m}_{11}$:

$$\vec{m}_{11} = \lambda Of, \Delta T, \Delta M, C, T, N, d, T', x_5, x_6, x_9, x_{10}, x_{11}. (T, \Delta T, F_3 (T', Of, N, T), \Delta M, x_5, x_6, C, N, x_9, x_{10}, x_{11}, F_{12} (d, T', Of, \Delta M, N, \Delta T)) \quad (13)$$

It obviously denotes the set of all vectors $\in V$ which satisfy the functions F_3 and F_{12}.

(2) *assigning constants* to (some of) the variables listed in the operator (where *a*, *b*, *c*,... are numbers and $(\ )_{11}$ is an abbreviation standing for the operand of the vector $\vec{m}_{11}$), e.g.:

$$\vec{m}_{111} = (\lambda Of, \Delta T, \Delta M, C, T, N, d, T', x_5, x_6, x_9, x_{10}, x_{11} \cdot (\ )_{11}) (c_N, c_T, c_C, c_{\Delta M}, c_{\Delta T}, c_{Of})$$

This new vector $\vec{m}_{111}$ represents the set of vectors $\in V$ which, for any d, T' (and any x_i occurring in the λ-operator), will exhibit a "fixed formant"-area defined by c_T, c_N and c_C. It is easy to see that in this way we may represent vowels, musical instruments etc. Summarizing we say that any M-vector $\vec{m}_j$ stands for a sound-*concept* and denotates a subset of the matrix *V*. This set can be associated with a descriptive meaning (e.g. the property to be a vowel /a/, a fricative /f/, a basson-like sound etc.).

(3) If we now *allocate constants to all the λ-tied variables* occurring in a M-vector *and apply λ-elimination,* then we will arrive (after computation of the functions) at a single VOSIM- or V-vector and thus at a sound *event*. The sounds obtained in this way are, of course, not random sounds but events which exhibit exactly the desired properties wished for; in short they are *instances* of sound concepts formulated in our language.

This was the MIDIM-language in a "nut-shell". Yet in reality one single vector is not enough to describe the behavior of sound in time. It turned out that four vectors at least are necessary for this (prefix, body, suffix, stop). Sounds characterized by two (or more) formants in the frequency domain may be obtained by the sum of the outputs of two (or more) synchronized generators as mentioned above. The MIDIM language represents every sound concept by means of a pair of synchronized *sequences* of four vectors preceded by a λ-operator listing the arguments for pitch, attenuation and time duration (the prosodic controls). These pairs are called a *predicator*. The predicators are assembled in an ordered set called *P-library* which assigns them an index as a name. Forming *formulae* of the MIDIM language means then building up sequences of predicators (viz. of sound concepts) and assigning to each predicator the constants for the prosodic controls. This will be done by means of a list of predicator indices (referring to a P-library) and of the corresponding constants. Such a list is called a *descriptor* since it denotes a sequence of single sound events.

The MIDIM-system M8X for digital composition (and its forerunners M7 and M8) ran since 1977 until the present time on the DEC PDP-15 computers of the Institute for Sonology Utrecht and were used by many composers and researchers from numerous different countries (cf. the composition list on p. 182). This led not only to an artistic harvest but to a testing of the expressive power of the VOSIM model as well as of the structure of the controlling language. In Amsterdam the M8-system has been implemented into a DEC PDP-23 at the Sweelinck conservatory. Actually a new generation of the system, the MIDIM9, is in preparation and will be implemented at the Royal Conservatory in the Hague in 1986.

The M8X-system refers to a design made in the mid 70's for the DEC PDP-15 computer and was set up as a tool to control the first generation of VOSIM digital hardware oscillators. Meanwhile a lot has changed. While storage space on the computers of the early 70's was limited, computers are nowadays supplied with large storage capacities as well as more sophisticated operating systems and peripherals which are micro-processor supported. What is more a 3rd generation of VOSIM generators has been designed where tasks earlier accomplished by means of the host computer are now taken over by microprocessor supported hardware solutions. Given the situation I made the decision not to refine the older systems but to design a new version from the ground up. Once again the central idea was to set up, as in the earlier systems, a MIDIM language upon the input matrix of the VOSIM sound system. In comparison with the MIDIM language which controlled the former systems, the language applied in MIDIM9 is more advanced. It concentrates on a calculus applying two main rules, namely substitution and λ-elimination. Together with the enlarged set of parameters of the new VOSIM generators with their extended domains, the refined language conveys to the MIDIM9 system a considerably improved expressive power. The software is written in PASCAL.

REFERENCES

Buxton, W. A. S. (1977). A Composer's Introduction to Computer Music. Interface 6, 57-72.

Kaegi, W. (1967). Was ist elektronische Musik. Zürich.

Kaegi, W. (1973/1974). A Minimum Description of the Linguistic Sign Repertoire. Interface 2, 141-156 and Interface 3, 137-158.

Kaegi, W./Tempelaars S. (1978) VOSIM — A New sound Synthesis System. Journ. AES 16, 6, 418-425.

Moog, R. (1984). Digital Synthesis, part V: Some alternatives. Keyboard, jan 1984.

Tempelaars S./Scherpenisse J. (1976). Exponential Decay. Interface 5, 207-224.

Tempelaars S. (1977). The VOSIM Signal Spectrum. Interface 6, 81-96.

Wallraff D. et al. (1983). Computer Music Catalog. Digital Music Systems, Inc, Boston Mass.

Interface, Vol. 15 (1986), pp. 83-161. 0303-3902/86/1502-0083 $3.00

The MIDIM Language and Its VOSIM Interpretation

Werner Kaegi

ABSTRACT

In 1974-77 I developed the MIDIM8X system for digital musical composition and research. It is a software package for controlling a pair of VOSIM7 hardware generators. MIDIM8X is conceived in terms of an interpreted formal language. In this paper the language will be presented in an analytical way in order to keep it independent of its material realization. A new MIDIM9 system based on this language is on the way to be implemented.

1. PRELIMINARIES OF THE MIDIM-LANGUAGE

1.1 Alphabet

Individual constants	$c_1, \ldots, c_p$
functional symbols	$F_1, \ldots, F_n, \ldots, G_n, \phi_1, \ldots, \phi_m, \chi_1, \ldots, \chi_m, f, g, h$

For formalization we will still introduce the following:

variables	$x_1, \ldots, x_n, \gamma_1, \ldots, \gamma_q, \alpha, \beta, \gamma \ldots$
lambda operator symbol	λ
brackets, comma, dot	( ) , .
relational symbols	$\leftrightarrow$ $=$

1.2 Terms

A constant is a term,
a variable is a term,
if x is a variable and F is a functional symbol then $F(x)$ is a term.

We shall adopt the representation of functions to be found in λ-conversion. (Cf. Church 1941, Hindley 1972). Functions will thus be represented as follows, where a, b, c are numbers:

$$F = \lambda x . x - y$$
$$F(a) = (\lambda x . x - y)(a) \leftrightarrow a - y \quad \lambda\text{-elimination}$$
$$F' = \lambda y . x - y$$
$$F'(b) = (\lambda y . x - y)(b) \leftrightarrow x - b \quad \lambda\text{-elimination}$$

Two- (or many-)placed functions are represented by functions of one variable, whose value is not a number but a function.

$G \quad = \lambda xy \, . \, x - y$ will accordingly be represented by
$G' \quad = \lambda x \, . \, (\lambda y \, . \, x - y)$
$G'(a) \quad = (\lambda x \, . \, (\lambda y \, . \, x - y))(a) \leftrightarrow \lambda y \, . \, a - y$
$(G'(a))(b) = (\lambda y \, . \, a - y)(b) \leftrightarrow a - b$

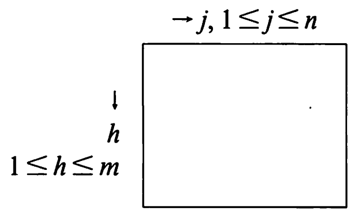
Abbreviations for convenience:
$\lambda x_1, x_2, \ldots, x_n \, . \, Y := (\lambda x_1 \, . \, (\lambda x_2 \, . \, (\ldots (\lambda x_n \, . \, Y) \ldots)))$
$(\lambda x_1, x_2, \ldots, x_n \, . \, Y)(c, \ldots, b, a) := (\lambda x_1 \, . \, (\lambda x_2 \, . \, (\ldots (\lambda x_n \, . \, Y)(c) \ldots))(b))(a)$

1.3 Vectors

If $t_1, \ldots, t_n$ are terms then $(t_1, \ldots, t_n)$ is a *vector* of dimension n.

1.3.1 Vectors with constant coefficients (V-vectors)

Let us assume there to be the constant $m \times n$-*matrix* V as follows:



where it holds:

c_{hj} stands for the *constant coefficient* at the intersection of line h and column j.

In an explicit way the matrix V may be defined by a λ-expression as follows:

$$V = \lambda h, j \, . \, c_{hj} \quad \text{where } h = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, n \tag{1.1}$$

We now define:

A *V-vector* $\vec{v}$ is any vector where all terms occurring are constants.

Hence:

$$\vec{v}_h = \lambda j \, . \, c_{hj} = (c_{h1}, \ldots, c_{hn}) \quad \text{is a line- or } V\text{-vector of } V. \tag{1.2}$$

1.3.2 Vectors with variable coefficients (M-vectors)

Up to this point we dealt with constant coefficients $c_{hj} \in V$. Now we shall proceed to an interpretation of V by the following assignment introducing the M-vector $\vec{m}_0$:

$$\vec{m}_0 = \lambda x_1, \ldots, x_n \, . \, (x_1, \ldots, x_n) \quad \text{for} \quad V = \lambda h, j \, . \, c_{hj} \tag{1.3}$$

where $x_j, j = 1, 2, \ldots, n$ stands for a *variable* coefficient of the vector $(x_1, \ldots, x_n)$, the ordered integer-domain D_j of x_j coinciding with the column j of V.

Accordingly:

$$x_j \in D_j, D_j = \{c_{1j}, c_{2j}, \ldots, c_{mj}\} \tag{1.4}$$

We now define: An *M-vector* $\vec{m}$ is any vector where one term at least is an λ-tied variable x_k or a function F_j of x_k.

Accordingly the following are *M*-vectors:

$$\vec{m}_0$$
$$\lambda x_i \,.\, (c_1, c_2, \ldots, x_i, \ldots, c_n)$$
$$\lambda x_k \,.\, (c_1, c_2, \ldots, F_i(x_k), \ldots, c_n)$$
$$\lambda x_1, \ldots, x_n \,.\, (x_1, \ldots, x_n)$$
$$[\lambda x_1, \ldots, x_n \,.\, (x_1, \ldots, x_n)](c_n, \ldots, c_1)$$

1.3.3 Substitution

We introduce the *substitution*-operator $(\lambda F_j)\, x_k \ldots$. Applied to any arbitrary *M*-vector $\vec{m}$ it replaces any occurrence of the variable x_j by the function $F_j(x_k, \ldots)$. The operator operates, of course, upon λ-expressions and is defined as follows:

$$S_1: x_j \rightarrow F_j(x_k, \ldots) \text{ for the operand of the innermost } \lambda\text{-expression,} \tag{1.5}$$

$$S_2: \lambda\alpha, \beta, x_j \,. \rightarrow \lambda\alpha, \beta, x_k, \ldots \quad \text{for all } \lambda\text{-operators occurring} \tag{1.6}$$

$$S_3: \lambda x, x = \lambda x \quad \text{in the expression.} \tag{1.7}$$

Here are two examples of substitution:

Let there be the functions

$$F_h(x_i, x_k, \ldots, \gamma)$$
$$F_i(\alpha, \beta, \ldots, \gamma) \quad \text{with } h \neq i \text{ and } h, i, k \in [1, n]$$

$$\begin{aligned}\vec{m}_p &= (\lambda F_h)\, x_i, x_k, \ldots, \gamma \,.\, \vec{m}_0 \\ &= (\lambda F_h)\, x_i, x_k, \ldots, \gamma \,.\, \lambda x_1, \ldots, x_h, \ldots, x_n \,.\, (x_1, \ldots, x_h, \ldots, x_n) \\ &= \lambda \ldots, x_i, x_k, \ldots, \gamma \,.\, (\ldots, F_h(x_i, x_k, \ldots, \gamma), \ldots, x_i, \ldots)\end{aligned}$$

$$\begin{aligned}\vec{m}_q &= (\lambda F_i)\, \alpha, \beta, \ldots, \gamma \,.\, \vec{m}_p \\ &= \lambda\alpha, \beta, \ldots, \gamma, x_k, \ldots (\ldots, F_h(F_i, x_k, \ldots, \gamma, \ldots), \ldots, F_i, \ldots)\end{aligned}$$

For convenience we still introduce the *product of substitution operators:*

$$\vec{m}_q = (\lambda F_i F_h)\, x_i, x_k, \alpha, \beta, \ldots, \gamma \,.\, \vec{m}_0$$
$$\text{for } (\lambda F_i)\, \alpha, \beta, \ldots, \gamma \,.\, (\lambda F_h)\, x_i, x_k, \ldots, \gamma \,.\, \vec{m}_0 \tag{1.8}$$

Admitting

$$\vec{m}_0 = (\lambda F_j)\, x_j \,.\, \vec{m}_0 \tag{1.9}$$

it follows that every *M*-vector may be represented by an expression where $\vec{m}_0$ is preceded by the appropriate product of substitution operators.

1.3.4 λ-Elimination

$$\vec{m}_p(c) = (\lambda x_i, x_k, \ldots, \gamma, \delta \,.\, (\ldots, F_h(x_i, x_k, \ldots, \gamma, \delta, \ldots), \ldots, x_i, \ldots))(c)$$
$$\longleftrightarrow \lambda x_k, \ldots, \gamma, \delta \,.\, (\ldots, F_h(c, x_k, \ldots, \gamma, \delta, \ldots), \ldots, c, \ldots) \qquad (1.10)$$

In any M-vector the λ-tied variables to which constants are allocated may be λ-eliminated.

Every M-vector $\vec{m}$ where constants are allocated to all λ-tied variables occurring may be represented by a V-vector $\vec{v}$.

Accordingly we obtain e.g. for the line vector $\vec{v}_h$ stated above:

$$\vec{v}_h = (\lambda x_1, \ldots, x_n \,.\, (x_1, \ldots, x_n))(c_{hn}, \ldots, c_{h1}) \longleftrightarrow (c_{h1}, \ldots, c_{hn}) \quad \text{cf. (1.2)}$$

1.3.5 Interpretation

Every M-vector stands for a subset $S \subset V$ of V-vectors.

(1) Referring to (1.1) we may thus represent any arbitrary S as follows:

$$S = \lambda h, j \,.\, c_{hj} \quad \text{with } h \in H \subset [1, m] \text{ and } j = 1, 2, \ldots, n \qquad (1.11)$$

(2) We may, however, also write:

$$S = \vec{m} = (\lambda F_i \ldots)\, x_k, \ldots, \gamma \,.\, \vec{m}_0 \quad (i, k \text{ not necessarily different}) \quad (1.12)$$

(3) The domain of every variable listed in the λ-operator of vector $\vec{m}$ being defined, we may then allocate constants to all these variables and apply λ-elimination, which results in a single V-vector belonging to S as follows:

$$\vec{v}_a = (c_{a1}, \ldots, c_{an}) \in S \subset V \quad \text{with } a \in H \qquad (1.13)$$

Here is a very simple example:

$$S = \vec{m}_r = (\lambda F_n) \ldots, \gamma \,.\, \vec{m}_0$$
$$= \lambda x_1, \ldots, x_{n-1}, \ldots, \gamma \,.\, (x_1, \ldots, x_{n-1}, F_n(\ldots, \gamma))$$
$$F_n = \lambda x_1, x_3, x_8, \gamma \,.\, \gamma/(x_1 \cdot x_8 + x_3)$$

Accordingly the following V-vector $\vec{v}_a$ is an instance of S:

$$(\vec{m}_r)(\ldots, 2, \ldots, 200, 0, 150, 10^3)$$
$$= [\lambda x_1, x_2, x_3, \ldots, x_8, \ldots, \gamma \,.\, (x_1, x_2, x_3, \ldots, x_9, \ldots, \gamma/(x_1 \cdot x_8 + x_3))]$$
$$(10^3, \ldots, 2, \ldots, 200, 0, 150)$$
$$\longleftrightarrow \vec{v}_a = (150, 0, 200, \ldots, 2, \ldots, 2) \in S,\ a \in H$$

By varying all constants over the corresponding domains defined by the λ-operator of the M-vector under consideration, we eventually may enumerate all V-vectors belonging to S. The list of all V-vectors will be called *the extension* of S, and the representation by means of an M-vector shown in (1.12) will be called the *intension* of S.*

In de MIDIM-language the intensional representation of subsets is adopted. This involves a strategy which focusses upon answering the two questions as to:

a) which are the functions substituted into $\vec{m}_0$, and

b) which are the domains of the variables listed in the λ-operator?

1.4 Functions

For formalization a distinction is made between

V-variables $x_1, x_2, \ldots, x_n$
M-variables $\gamma_1, \gamma_2, \gamma_3, \ldots$

And now we define:

A *function of first order* F_j is any function the value of which is a V-variable x_j, and
a *function of second order* ϕ_j is any function the value of which is an M-variable γ_j.

The functions are listed in the corresponding function tables 1 and 2 and may thus be called by indices (though for convenience labels will also be used in what follows). The MIDIM-language recognizes still more types of functions the discussion of which is, however, postponed. Cf. par. 1.7, 5.3 and par. 9.2.

1.5 Domains

According to the substitution introduced in par. 1.3.3 every variable coefficient $x_j \in D_j$ may be replaced by a function F_j of x_h (or of γ_h). Let us assume A, B, D, E to be non-empty sets of numbers. Accordingly

$(\lambda F_j)\ \gamma_h\ .\ \vec{m}_k$ stands for substituting
$F_j\colon A \rightarrow B$ into $x_j \in D_j$ occurring in the M-vector $\vec{m}_k$ such that:

$$D_j \subset B$$
$$E = \{\gamma_h \in \mathrm{A} \mid F_j(\gamma_h) \in D_j\} \qquad (1.14)$$

The domain E of the function $F_j(\gamma_h)$ depends thus upon D_j of x_j.

Here is a very common example:

Let there be the function f defined by $\lambda\gamma\ .\ \gamma^2$ substituted into the coefficient $x \in D = [1, 9]$. Accordingly it holds:

$\gamma \in E$ and $\gamma^2 \in D$, where $E = [1, 3]$. (Fig. 1.1.)

---

* Carnap (1947)

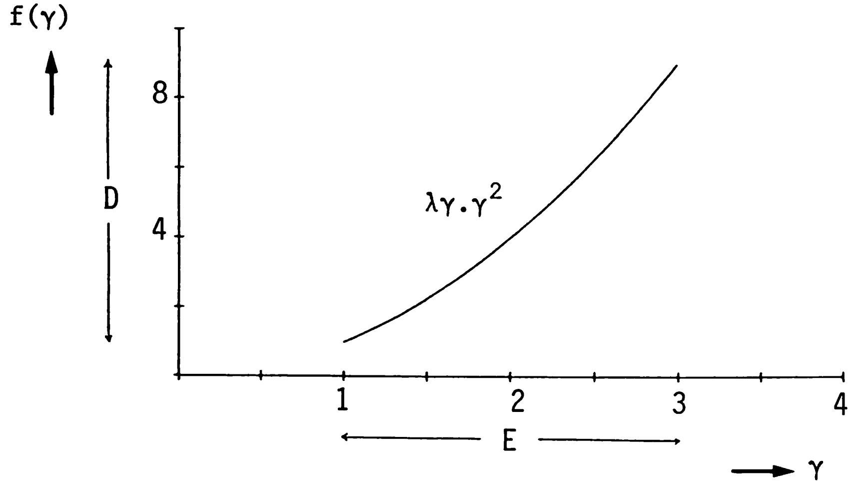


Figure 1.1 Dependency of domain E upon the range D due to substitution.

For many-placed functions (and for compositions of first- and second-order functions) the product of functions:

$$(\prod_{i=1}^{n} f_i)(\gamma_n) \quad \text{standing for} \quad f_1 \cdot f_2 \cdot \ldots \cdot f_n(\gamma_n) \tag{1.15}$$

is introduced into (1.14) as follows (for simplicity we shall omit the index j, e.g. $f_i = (f_j)_i$, $\gamma_n = (\gamma_j)_n$, $D = D_j$):

$$E_n = \{\gamma_n \in A_n | (\prod_{i=1}^{n} f_i)(\gamma_n) \in D\} \tag{1.16}$$

There are, however, cases where the range D_j of the function $F_j(x_k)$ substituted into x_j is not a constant interval but depends upon the argument x_k. This will e.g. be the case for functions mapping the VOSIM-parameter x_i into Δx_i (applying for T, M and A). Since it holds that:

$$\Delta x_i \in [x_{i\min} - x_i, x_{i\max} - x_i] \wedge x_i \in D_i = [x_{i\min}, x_{i\max}] \tag{1.17}$$

we obtain:

$$E_{F_j} = \{x_i \in D_i | F_j(x_i) \in [x_{i\min} - x_i, x_{i\max} - x_i]\} \tag{1.18}$$

Here is the example of the function G_2 (cf. Fig. 1.2 and pg. 94):

$$E_{G_2} = \{T \in D_1 | G_2(T, c_{FS}) \in [T\min - T, T\max - T]\}$$

In order to give the reader an idea what the instances will be, the following list and a visualization for $s = 4$ are presented:

$$C(0) = [\lambda v, d_i, DUR, \ldots, X, Y \,.\, S_1 \ldots S_4](0) = [\lambda v, \ldots, X, Y \,.\, SSSS](0)$$
$$C(1) = [\lambda v, d_2, d_3, d_4, DUR, \ldots, X, Y \,.\, ASSS](1)$$
$$C(2) = [\lambda v, d_1, d_3, d_4, DUR, \ldots, X, Y \,.\, SASS](2)$$
$$C(3) = [\lambda v, d_1, d_2, d_4, DUR, \ldots, X, Y \,.\, SSAS](3)$$
$$C(4) = [\lambda v, d_1, d_2, d_3, DUR, \ldots, X, Y \,.\, SSSA](4)$$

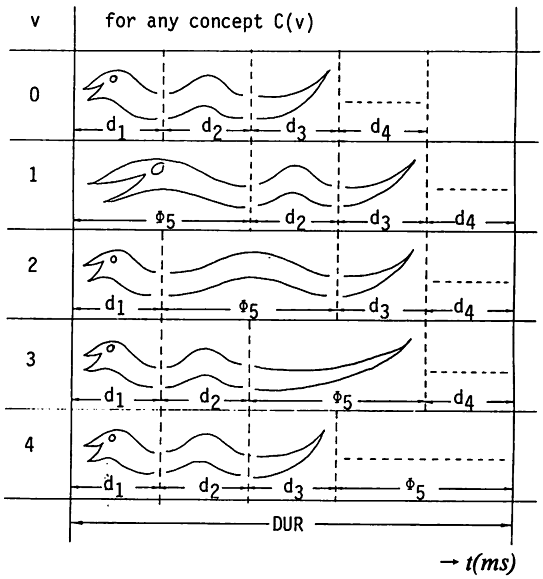


Figure 1.3. The concept *C(v)*. Visualization.

1.7 Pairs of linked *M*-Concepts and the Predicator

Let us assume there to be the pair of *M*-concepts $C_l(v)$ and $C_L(v)$, where $v_L = v_l$ and $2 \leq L \leq n$. We now introduce the L(ink)-functions by means of which a link is laid between the two concepts as follows:

An *L-function* L_j is any function the value of which is an *M*-variable $(\gamma_j)_L$ occurring in the concept $C_L(v)$.

The *L*-functions are listed in the open function table 3 (cf. par. 5.3) which will at least contain the following two functions:

$$L_1 = \lambda(\gamma_1)_1 \,.\, (\gamma_1)_1 = \lambda DUR_1 \,.\, DUR_1 \quad \text{(cf. par. 1.6.2 and 2.3)} \qquad (1.29)$$

$$L_5 = \lambda(\gamma_5)_1 \,.\, (\gamma_5)_1 = \lambda(d_i)_1 \,.\, (d_i) \qquad (1.30)$$

Common controls $(\gamma_j)_L = (\gamma_j)_1$ occurring in $C_1(v)$ and $C_L(v)$ respectively are thus obtained by the application of substitution as follows:

$$(\lambda L_j)(\gamma_j)_1 \,.\, (\gamma_j)_L \tag{1.31}$$

Assuming that there is a common absolute start point in time for both concepts involved, *synchronization* is then obtained by:

$$(\lambda L_5 L_1)(d_i)_1, DUR_1 \,.\, C_L(v) \tag{1.32}$$

We now define:

The *predicator* $P(c_v)$ is a n-uple of synchronized concepts as follows (where DUR, d_i stand for $(DUR, d_i)_1$):

$$P(c_v) = [(\lambda \ldots L_5 L_1)\ldots, d_i, DUR \,.\, (C_1(v), \ldots, C_n(v))](c_v) \tag{1.33}$$

Interpretation: Each of the synchronized concepts occurring in the predicator will denotate a subset $\in V^*$ from which the input for the n corresponding sound generating units will be selected. Since the n concepts will now have in common the start point, the constant c_v as well as the variable d_i and DUR, they will operate over the same absolute time interval for each segment S_i.

1.8 The application of the Terminal Rule

For the time being we shall limit ourselves to a sketch of the main lines of the MIDIM-syntaxis and take up a more extensive presentation at a later stage (par. 6). The reason for this is that further developments of the syntaxis will hardly be graspable for the reader before he has gained an insight into the interpretation of the language, because it is exclusively by means of the interpretation aimed at that it can be explained why this and no other way has been taken when setting up the language. This, however, does not exclude that a strict syntactical representation of the MIDIM-language could have been given here.

What I would like to discuss now is the role played by λ-elimination within the frame of our language. In the previous pages we familiarized the reader with how, by means of substituting functions into $\vec{m}_0$, any sort of M-vectors can be formed and then assembled into concepts which themselves are further assembled into predicators. The extensional interpretation which was given to this was the denotation of subsets of V^*. Within these subsets further subsets can be selected by allocating constants to the λ-tied variables until a single n-uple of synchronized sequences of V-vectors is defined. This is the case when constants are applied to all variables occurring. Although the expression obtained denotes V-vectors, it still contains all substitution operators applied and thus all information concerning its derivation from $\vec{m}_0$. λ-elimination states that every λ-expression where all tied variables are assigned constants can be converted into an expression without λ-tied variables. This means in our case that, once λ-elimination is applied and the functions are computed, the expression consists merely of *V*-vectors. In our language, application of λ-elimination and the computation of the functions will thus be the *terminal rule* by means of which we shall leave the intensional MIDIM-representation and step over to the extensional *V*-representation of sound events. As a consequence we lose the predicator with its networks of functions, but we obtain a terminal expression which can be converted into sound. It is obvious that, if we will proceed to extend our syntaxis by forming *sequences of predicators*, subject these sequences to grammars, and allocate constants to all λ-tied variables occurring, then an application of the above terminal rule will result in an extensional *V*-representation of sequences of sound events wished for.

2. THE VOSIM-INTERPRETATION OF THE MIDIM-LANGUAGE

2.1 Generalities

Up to this point we have given a physical interpretation of the MIDIM-language exclusively for the parameter *time-duration.* We shall now extend the physical interpretation. It is obvious that this will depend upon the input-parameters recognized by the particular sound synthesis system we wish to control. For instance a system in terms of additive synthesis would involve ex- or implicit amplitude- and phase-controls for the sin-functions involved, whilst any *FM*-system would recognize speed- and width-controls for the carrier as well as for its modulation etc. The MIDIM-language as it has been designed hitherto could be constructed upon the input-list of both systems mentioned as well as upon any other type of system (wave shaping, linear prediction etc.). In what follows we shall, however, concentrate upon the interpretation of the MIDIM-language in terms of the VOSIM sound synthesis. Accordingly we shall proceed to the formation of a MIDIM/VOSIM-language. (The MIDIM-systems based upon this language are the M8 and the M8X (implemented on the DEC PDP-15 computers at Sonology Utrecht) and the new MIDIM-9 (implemented on the DEC PDP-11/23 at Sonology The Hague.))

2.2 The interpreted matrix $\vec{m}_0$

Assigning the *parameters* is realized as follows:

$$\begin{aligned}\vec{m}_0 &= \lambda x_1, \ldots, x_{12} \,.\, (x_1, \ldots, x_{12})\\ &= \lambda T, \Delta T, M, \Delta M, A, \Delta A, C, N, D, S, MF, Np \,.\\ &\quad (T, \Delta T, M, \Delta M, A, \Delta A, C, N, D, S, MF, Np)\end{aligned} \tag{2.1}$$

The *domains* D_j of the variable coefficient x_j are defined in accordance with the particular VOSIM-generator(s) intended to be controlled:

$$D_j = [c_{1j}, c_{zj}] \quad \text{where } z \leq m \ (m \text{ is the number of line-vectors of } V) \tag{2.2}$$

The following table of domains applies to the M8X-system. The numerical examples submitted to the reader in the following pages will refer to this table.

parameter		integer-domain		parameter	integer-domain	
$T(t)$	(us)	[80, 8190]	hardware	D (us)	[0, 4095]	software
$M(t)$	(us)	[1, 4095]	hardware	MF (number of T')	[1,]	software
$A(t)$	(arb.)	[0, 511]	hardware	S	[0, 1]	software
C	(%)	[1, 100]	software	Np	[1, 4096]	software
N		[1, 1024]	software			

2.3 The interpreted *M*-functions

The functions are defined by the following *function tables;* cf. also par. 5 (function table 3).

2.3.1 *The function table 1 of first-order-functions*

It lists from left to right in

column 1: the vector $(x_1, \ldots, x_{12})$ of *V*-variables,

column 2: the variable VOSIM vector $(T, \ldots, Np)$, i.e. the physical interpretation of the above vector $(x_1, \ldots, x_{12})$,

column 3: the vectors $(F_1, \ldots, F_{12})$ and $(G_1, \ldots, G_{12})$ and ... of first order functions.

coeff. x_j	parameter $\in \mathbb{R}$	function F_j, G_j $\quad (R = 2^{(1200^{-1})} = \text{const.} = 1.00057779)$	
x_1	T(us)	$F_1 = \lambda N, q, Of, T' \,.\, q(T' + Of)/N$	
x_2	ΔT(us)	$F_2 = \lambda N, q, \Delta T' \,.\, q\Delta T'/N$	
x_3	ΔT	$G_2 = \lambda T, FS \,.\, (R^{FS} - 1)T$	
x_3	M(us)	$F_3 = \lambda N, T, Of, T' \,.\, T' + Of - NT$	
x_4	ΔM(us)	$F_4 = \lambda N, \Delta T, \Delta T' \,.\, \Delta T' - N\Delta T$	
x_5	A	$F_5 = \lambda y_2, e, At, Am \,.\, ((y_2 - 1)At^e + 1)Am \cdot At$	see par. 2.3.3
x_6	ΔA	$F_6 = \lambda y_2, e, At, \Delta Am \,.\, ((y_2 - 1)At^e + 1)\,\Delta Am \cdot At$	
x_7	$C \in \mathbb{Q}$	F_7 = open	
x_8	$N \in \mathbb{N}$	$F_8 = \lambda q, Of, T\text{max}, T' \,.\, \text{int}_1(q(T' + Of)/T\text{max})$	see par. 2.3.1
x_8	N	$G_8 = \text{int}_2(g_8(\Delta T, D, Ef, N\text{max}, \Delta T', Of, T, T')$	see par. 2.3.2
x_9	D	$F_9 = \lambda W, Of, T' \,.\, (T' + Of)(R^W - 1)$	
x_9	D	$G_9 = \lambda y, M \,.\, y \cdot M$	
x_{10}	S	F_{10} = open	
x_{11}	$MF \in \mathbb{N}$	$F_{11} = \lambda Sp, Of, T' \,.\, rof(10^6/(Sp(T' + Of)))$	
x_{12}	$Np \in \mathbb{N}$	$F_{12} = \lambda N, T, M, \Delta T, \Delta M, d \,.\, \text{rof}\left(\dfrac{d \cdot 10^3}{N \cdot T + M + \frac{1}{2}(\Delta M + N \cdot \Delta T)}\right)$	

$$\text{rof}(x) = \begin{cases} \text{int}(x + 0.5) \text{ if } x \geq 1. \\ 1 \text{ otherwise} \end{cases} \quad \text{(rounding off)}$$

2.3.2 *The function table 2 of second-order-functions*

It lists from left to right in

column 1: the vector $(\gamma_1, \gamma_2, \ldots)$ of M-variables,

column 2: the vector of variable MIDIM-parameters, i.e. the physical interpretation of the above vector $(\gamma_1, \gamma_2, \ldots)$,

column 3: the vectors $(\Phi_1, \Phi_2, \ldots)$, $(\chi_1, \chi_2, \ldots)$, ... of second order functions.

coeff. γ	parameter $\in \mathbb{R}$	function Φ $(c = 16.35159784\ (\text{Hz}))$	meaning
γ_1	DUR(ms)	$\Phi_1 = \lambda \text{met}, du\ .\ du \cdot 6 \cdot 10^4/\text{met}$	duration of word
γ_2	$At \in [0, 1]$	$\Phi_2 = \lambda at\ .\ 2^{-at}$	attenuation
γ_3	T'(us)	$\Phi_3 = \lambda pi, oc, \text{sub}\ .\ 10^6(c \cdot 2^{oc+(pi/\text{sub})})^{-1}$	durat. of V-period
γ_3	T'(us)	$\chi_3 = \lambda N, T, M, Of\ .\ N \cdot T + M - Of$	
γ_4	Of(us)	$\Phi_4 = \lambda P, T'\ .\ T'(1/R^P - 1)\ (= -\Delta T')$	offset, T'-shift
γ_4	Of(us)	$\chi_4 = \begin{cases} i > h \rightarrow \lambda h, i, \Delta T'_j\ .\ \sum_{j=h}^{i-1} \Delta T'_j \\ \text{else} \rightarrow 0 \end{cases}$	
γ_5	d(ms)	$\Phi_5 = \lambda v, DUR, d_i\ .\ DUR - \sum_{i=1, i \neq v}^{s} d_i$	durat. of segment
γ_6	$\Delta T'$(us)	$\Phi_6 = \lambda T', Of, P\ .\ (T' + Of)(R^P - 1)$	increment rate of T'
γ_6	$(\Delta T'_i)_k$	$\chi_6 = \lambda h, i, k, T'_k, T'_{k+1}, (Of_i)_k, (Of_h)_{k+1}\ .$ $(T' + Of_h)_{k+1} - (T' + Of_i)_k$	increment rate per segm. S_i and word k
γ_7	$q\ 0 < q < 1$	$\Phi_7 =$(open)	ratio NT/T'
γ_8	P(cent)	$\Phi_8 =$(open)	portamento rate
γ_9	W(cent)	$\Phi_9 =$(open)	width of M-modulation
γ_{10}	SP(Hz)	$\Phi_{10} =$(open)	speed of M-modulation
γ_{11}	Am_i	$\Phi_{11} = \lambda i, Am, \Delta Am\ .\ (Am + \Delta Am)_{i-1}$	relat. amplitude per segm. S_i
γ_{12}	ΔAm	$\Phi_{12} = \lambda \alpha, d\ .\ \alpha \cdot d$	increm. rate for Am
γ_{12}	ΔAm	$\chi_{12} = \lambda Am\ .\ -Am$	
γ_{13}	Tmax(us)	$\Phi_{13} =$(open)	max. value of T
γ_{14}	$N\text{max} \in \mathbb{N}$	$\Phi_{14} =$(open)	max. value of N
γ_{15}	y	$\Phi_{15} = \lambda y_2, y_1, x_2, x_1, e, x\ .\ \frac{y_2 - y_1}{(x_2 - x_1)^e}(x - x_1)^e + y_1$	multipl. factor
γ_{16}	Ef(us)	$\Phi_{16} =$(open)	min. value of M
γ_{17}	EF(us)	$\Phi_{17} = \lambda Ef, D\ .\ \max(Ef, D)$ see 2.3.2	
γ_{18}	f_1(Hz)	$\Phi_{18} = \lambda T'\ .\ 10^6/T'$	fundam.freq. "f_0"
γ_{19}	FS(Cent)	$\Phi_{19} =$(open)	Ω-shift rate
.			
.			
.			

For obvious reasons we shall frequently apply the following products of functions:

x_2	ΔT	$\Phi_6 F_2 = \lambda N, q, P, Of, T' \,.\, q(T' + Of)(R^P - 1)/N$	(2.3)
x_4	ΔM	$\Phi_6 F_4 = \lambda N, \Delta T, Of, P, T' \,.\, (T' + Of)(R^P - 1) - N\Delta T$	(2.4)
x_{12}	Np	$F_{120} = \Phi_6 F_4 F_3 F_{12} = \lambda P, Of, d, T' \,.\, \mathrm{rof}\left(\dfrac{d \cdot 10^3}{(T' + Of)(1 + \frac{R^P - 1}{2})}\right)$	(2.5)

2.3.3 Some functions

We shall present here a few functions which need more elaboration than could be given within the above tables.

2.3.3.1 The function F_8

$$N = F_8 = \mathrm{int}_1(EN) = \mathrm{int}_1(f_8(T\mathrm{max}, Of, q, T')) \quad \text{with } EN \in \mathbb{R}, N \in \mathbb{N}$$

$$EN = \lambda T\mathrm{max}, Of, q, T' \,.\, q(T' + Of)/T\mathrm{max} \tag{2.6}$$

$$N = \mathrm{int}_1(EN) \equiv \begin{pmatrix} \text{if } EN \leq 1 \text{ then } N = 1 \\ \text{otherwise } N \geq EN \text{ with } N \text{ minimal integer} \end{pmatrix} \tag{2.7}$$

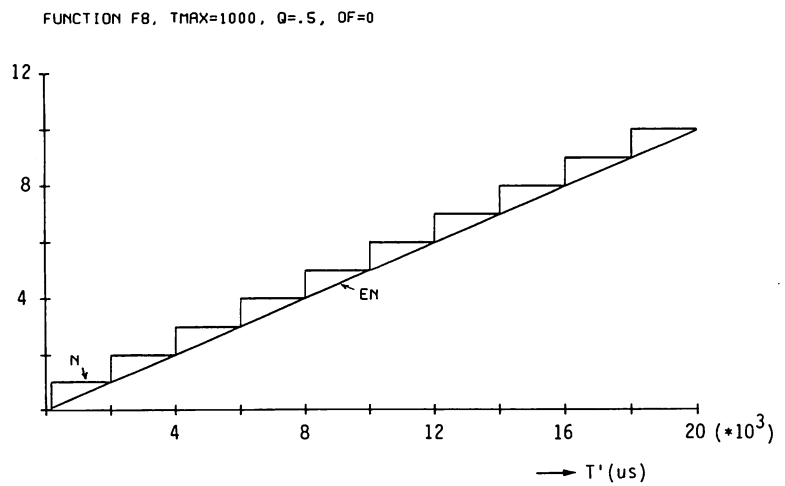


Figure 2.1. Function F_8. The graph shows the mapping of T' into the real EN and its mapping into the integer N, cf. par. 3.7

2.3.3.2 The function G_8
$N = G_8 = \text{int}_2(EN) = \text{int}_2(g_8(\Delta T, D, Ef, N\text{max}, \Delta T', Of, T, T'))$ with $EN \in \mathbb{R}$ (for simplicity we assume: $\Delta T' = 0$).

Assumptions:

$$E = \{n \in N \mid n/2 \in \mathbb{N}\}$$

$$O = \{n \in N \mid (n+1)/2 \in \mathbb{N}\}$$

$$EF = \Phi_{17} = \max(Ef, D) \text{ such that } \begin{cases} Ef < D \rightarrow EF = D \\ Ef \geq D \rightarrow EF = Ef \end{cases}$$

$$Ef \leq M \leq (mT + EF) \quad \text{with } m \in \mathbb{N} \tag{2.8}$$

where

$$m = 2 \quad \text{for} \quad \frac{M\text{max} - EF}{A} \geq 2.$$

$$m = 1 \quad \text{for} \quad \frac{M\text{max} - EF}{A} < 2.$$

$$A = T + \Delta T \quad \text{for} \quad \Delta T \geq 0.$$

$$A = T \qquad \text{for} \quad \Delta T < 0.$$

$$EN = \frac{T' + Of - M}{A} \quad \text{where} \quad M = mA + EF \tag{2.9}$$

$$N = \text{int}_2(EN) \quad \text{with} \tag{2.10}$$

for $m = 2$: a) either $N\text{max}, N \in E \wedge 2 \leq N \leq N\text{max}$
or else $N\text{max}, N \in O \wedge 1 \leq N \leq N\text{max}$
b) $M \leq 2T + EF$
hence: $N \geq EN = \frac{T' + Of - (2A + EF)}{A} = \frac{T' + Of - Ef}{A} - 2$
c) N minimal

for $m = 1$: a) $N\text{max}, N \in \mathbb{N} \wedge 1 \leq N \leq N\text{max}$
b) $M \leq T + EF$
hence: $N \geq EN = \frac{T' + Of - (A + EF)}{A} = \frac{T' + Of - Ef}{A} - 1$
c) N minimal

For the application cf. par. 3.8.

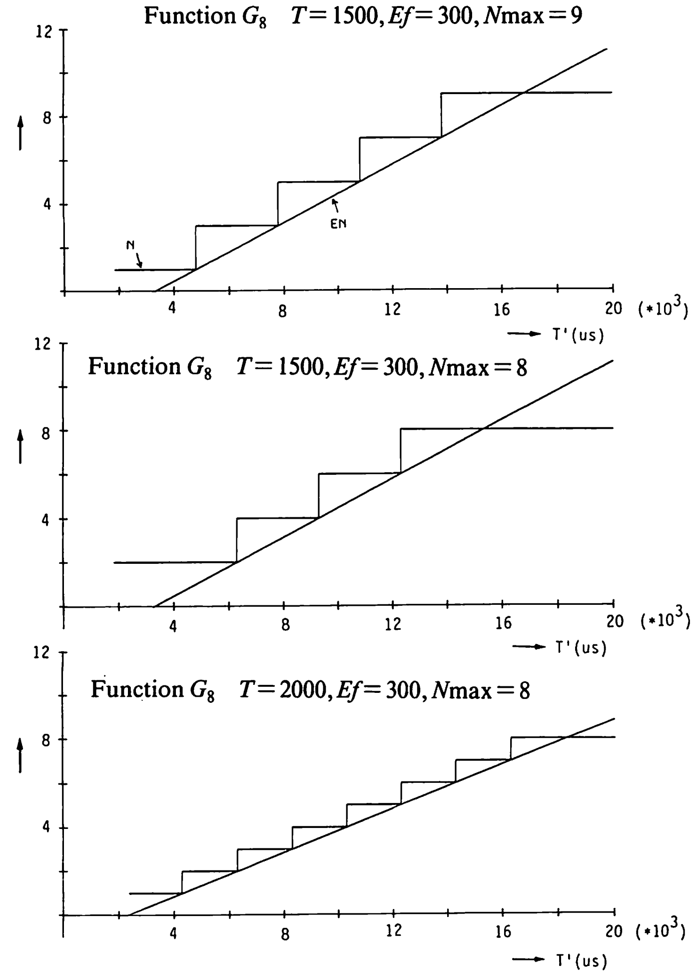


Figure 2.2-2.4. Function G_8 for Nmax odd and $m = 2$ (top), for Nmax even and $m = 2$ (middle) and for $m = 1$ (bottom).

2.3.3.3 The functions F_5 and F_6

We assume there to be the interpolation through the points (x_1, y_1) and (x_2, y_2)

$$\Phi_{15} = \frac{y_2 - y_1}{(x_2 - x_1)^e}(x - x_1)^e + y_1 \quad \text{cf. par. 2.3.2}$$

For the points (0, 1) and $(1, y_2)$ we accordingly obtain:

$$\Phi'_{15} = (\lambda x_1, x_2, y_1, y_2, e, At \,.\, \frac{y_2 - y_1}{(x_2 - x_1)^e}(At - x_1)^e + y_1)(1, 1, 0)$$

$$\longleftrightarrow \lambda y_2, e, At \,.\, (y_2 - 1)At^e + 1 \quad \text{cf. Fig. 2.5.} \qquad (2.11)$$

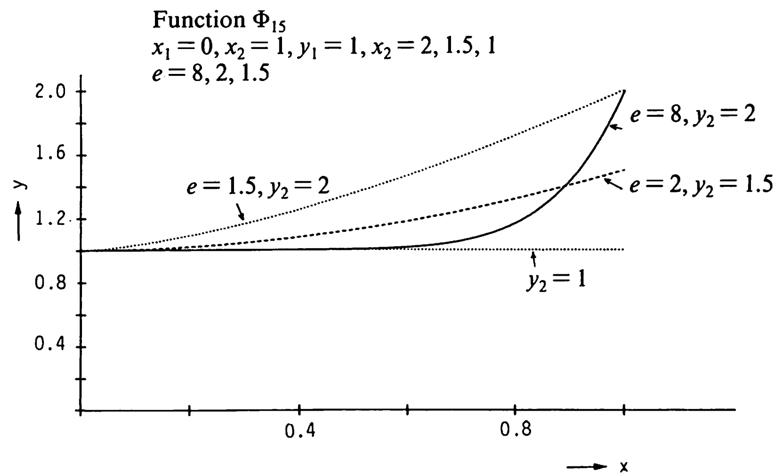


Figure 2.5. The interpolation function Φ_{15}.

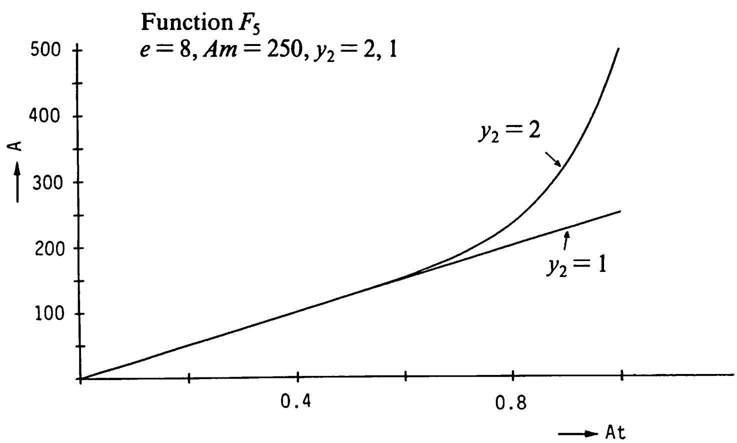


Figure 2.6. Function F_5. Application of interpolation Φ_{15} to the mapping of At into A.

Let us assume there to be in addition the functions:

$$H_1 = \lambda At, Am, \gamma_{15} \,.\, At \cdot Am \cdot \gamma_{15}$$
$$H_2 = \lambda At, \Delta Am, \gamma_{15} \,.\, At \cdot \Delta Am \cdot \gamma_{15}$$

We now form:

$$F_5 = (\lambda \Phi'_{15}) y_2, e, At \,.\, H_1 = \lambda y_2, e, At, Am \,.\, ((y_2 - 1) At^e + 1)) Am \cdot At$$
cf. Fig. 2.6.
$$F_6 = (\lambda \Phi'_{15}) y_2, e, At \,.\, H_2 = \lambda y_2, e, At, \Delta Am \,.\, ((y_2 - 1) At^e + 1)) \Delta Am \cdot At$$

For $y_2 = 1$ the function Φ'_{15} takes the constant value 1 for any At and any e:

$$F'_5 = (\lambda y_2, e, At, Am \,.\, ((y_2 - 1) At^e + 1) Am \cdot At)(1)$$
$$\longleftrightarrow \lambda At, Am \,.\, Am \cdot At$$

and likewise for F_6, which will be suitable for most cases. For the application of F_5 and F_6 with $y_2 > 1$ see par. 5.2, p. 125. For the At-domain we obtain:

$$At \in [0, At\max] \text{ with: } At\max = \frac{A\max}{y_2(Am + B)} \quad \text{where } B = \begin{cases} \Delta Am \text{ for } \Delta Am > 0 \\ 0 \text{ otherwise } y_2 \geq 1 \end{cases}$$

2.3.3.4 The function $F_{120} = \Phi_6 F_4 F_3 F_{12}$ cf. par. 2.3.2.

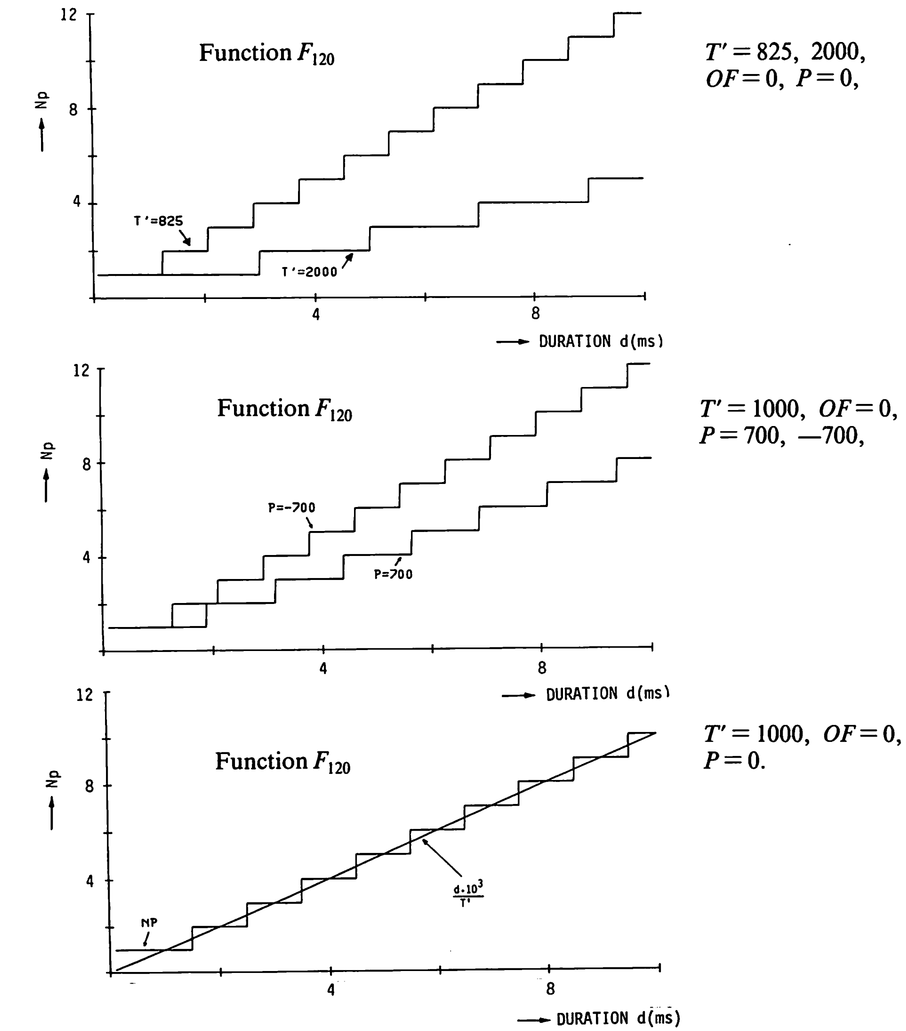


Figure 2.7-2.9. Function F_{120}. Dependency of Np upon d and T' (top), upon d and P (middle) and rounding-off therein.

2.4 Selecting the functions

The function tables 1 and 2 presented above are open lists which may at any moment be enlarged by introducing new functions. This leads to the question as to how the functions presented hitherto were selected and by which criteria new functions should by chosen?

There are two main criteria:
(1) the *model* for sound synthesis applied *and its physical realization* (in our case the VOSIM7-system); $\vec{m}_0$ stands for the VOSIM7 input matrix (in what follows *V*), and accordingly every function should exhibit as its value a VOSIM-interpreted *V*- or *M*-variable; (2) the *field of application* aimed at by our language (in the present case speech- and muscial sound synthesis covering the main concepts of the indo-european tradition*); the *M*-functions should thus be such that every product of substitution operators applied to $\vec{m}_0$ will denotate a subset of *V* which can be given an interpretation in the field of application.

Since we deal now with an interpreted language, a semantical interpretation will be conveyed to the *M*-vectors. Accordingly (a) the subsets of *V*-vectors $\in V$ designated by the *M*-vectors are their *M-extensions* and the corresponding product of substitution-operators applied to $\vec{m}_0$ are their *M-intensions.* (b) The latter will be associated with *descriptive* intensions or *properties* established in the field of application aimed at by our language. (Notice that the set of these properties will embrace any possible sort of musical and linguistic concepts whatsoever and is not limited either in its scope or its representation.**) The properties may be seen as the intensions of *predicates* established in the musical and linguistic culture under consideration, and, of course, descriptive extensions may be allocated to them (e.g. the sound-output of musical instruments (including synthezisers, music-computers), the extension of the sound system of standard English RP etc.).*** Seen in this light *duplicating* sound concepts external to the MIDIM system will then be the attempt to associate elements of descriptive extensions with *M*-intensions. Cf. Janssen/Kaegi on p. 185.

---

* There are reasons to believe that the field of application for the MIDIM/VOSIM-language is much larger.

** E.g. properties like "andante" (introduced into Italian 18th century music and still in use today, denotating ca. 84 human paces/min = ca. 0.7 (s) per pace = ca. 1.4 (Hz); "vowel" (standardized set of linguistic sounds since time immemorial, representable by the formant theory since Helmholtz); "roaring like lions" (which depicted the sound of the new Winchester organ for the inhabitants of that town in 980 by means of a sound concept established by common sense rather than by experience, a common sense at that time based upon the bible and in our own age based upon the logo of 20th century fox); "like the trumpets of judgement day" (representing a supranatural voice by means of a precise musical concept); but also "pizzicato", "con legno", "flageoletto", "muted", "short", "pink noise", "saw-tooth", "wet", "warm" etc.

*** This presupposes that in each case the domain of discourse as well as the set of predicates are defined. In speech this fact is rarely contested whilst the conciseness of musical concepts is often underestimated. In the case where it is difficult or even impossible to establish two-valued predicates, one can still refer to the evaluation-techniques suggested by fluid logics.

3. THE INTERPRETED *M*-VECTORS

We open the discussion with the interpretation of V given in (1.3). V now stands for the VOSIM7 input matrix.

$$\vec{m}_0 = \lambda x_1, \ldots, x_{12} \,.\, (x_1, \ldots, x_{12}) \quad \text{for } V \quad \text{cf. (1.3)}$$

Every arbitrary M-vector stands for a subset of V. If $\vec{m}_0$ is assigned in turn one through twelve constants, it will result in the family of the vectors where merely V-variables and/or constants occur:

$$(\lambda x_1, x_2, \ldots, x_{12} \,.\, (x_1, x_2, \ldots, x_{12}))\,(c_1)$$
$$(\lambda x_1, x_2, \ldots, x_{12} \,.\, (x_1, x_2, \ldots, x_{12}))\,(c_2, c_1)$$
$$\vdots$$
$$(\lambda x_1, x_2, \ldots, x_{12} \,.\, (x_1, x_2, \ldots, x_{12}))\,(c_{12}, \ldots, c_2, c_1)$$

However no vector of the above family will apply to the MIDIM-concept $C(v)$ since nowhere does the function F_{12} occur. Cf. par. 1.6.2. In what follows we shall present a selection of V-interpreted M-vectors, each of them being associated with descriptive intensions like "time duration", "pitch" etc.

3.1 Time duration

(For easy reading we shall omit in the λ-operators the V-variables not occurring in the functions substituted into the vectors.) The most simple vector satisfying $C(v)$ is as follows:

$$\begin{aligned}\vec{m}_1 &= (\lambda F_{12})\, d, T, \Delta T, M, \Delta M, N \,.\, \vec{m}_0 \qquad (3.1)\\ &= \lambda d, T, \Delta T, M, \Delta M, N \,.\, (T, \Delta T, M, \Delta M, \ldots, N, \ldots,\\ &\quad F_{12}(T, \Delta T, M, \Delta M, N, d))\\ &= \lambda d, T, \Delta T, M, \Delta M, \mathrm{N} \,.\, (T, \Delta T, M, \Delta M, \ldots, N, \ldots,\\ &\quad \mathrm{rof}\left(\frac{d \cdot 10^3}{N \cdot T + M + \tfrac{1}{2}(\Delta M + N \cdot \Delta T)}\right))\end{aligned}$$

This vector introduces among its λ-tied variables the M-variable d(ms) called in what follows the *time duration control*. The vector obviously stands for the subset of vectors $\in V$ which satisfy the function F_{12}. Assigning constants to $\vec{m}_1$ will result in a large family of subsets of vectors satisfying F_{12}.

$$\begin{aligned}\vec{m}_{11} &= (\lambda F_3) T, N, Of, T' \,.\, \vec{m}_1 = (\lambda F_3 F_{12}) T', Of, d, T, \Delta T, M, \Delta M, N \,.\, \vec{m}_0 \qquad (3.2)\\ &= \lambda T, \Delta M, \Delta T, N, Of, T', d \,.\, (T, \Delta T, F_3(T, N, Of, T'), \Delta M, \ldots, N, \ldots,\end{aligned}$$

$$F_{12}(T, F_3(T, N, Of, T'), \Delta M, \Delta T, N, d))$$
$$= \lambda T, \Delta M, \Delta T, N, Of, T', d\,.\,(T, \Delta T, T' + Of - N \cdot T, \Delta M, \ldots, N, \ldots,$$
$$\mathrm{rof}\left(\frac{d \cdot 10^3}{T' + Of + \frac{1}{2}(\Delta M + N \cdot \Delta T)}\right))$$

introduces still the M-variables T'(us) and Of(us) controlling the time duration of the first period (the initial state) of the VOSIM signal within the time interval d. We shall not yet talk about "pitch control" since the time dependent period-duration $N \cdot T(t) + M(t)$ will still depend upon ΔM and ΔT. Once again, assigning stepwise constants to $\vec{m}_{11}$ will result in a family of subsets of V. The same will apply to all vectors we shall present in what follows. In each case *the family will be characterized by an M-intension,* that is by *the corresponding product of substitution operators applied to* $\vec{m}_0$.

3.2 Pitch

$$\vec{m}_2 = (\lambda F_4)\Delta T, N, \Delta T'\,.\,\vec{m}_{11} \tag{3.3}$$
$$= (\lambda F_4 F_3 F_{12})\,\Delta T', T', Of, d, T, \Delta T, M, \Delta M, M\,.\,\vec{m}_0$$
$$= \lambda T, \Delta T, \Delta T', N, Of, T', d\,.\,(T, \Delta T, F_3(T, N, Of, T'),$$
$$F_4(\Delta T, N, \Delta T'), \ldots, F_{12}(T, F_3(T, N, Of, T'), F_4(\Delta T, N, \Delta T'),$$
$$\Delta T, N, d))$$
$$= \lambda T, \Delta T, \Delta T', N, Of, T', d\,.\,(T, \Delta T, T' + Of - N \cdot T, \Delta T' -$$
$$N\,.\,\Delta T, \ldots, N, \ldots, \mathrm{rof}\left(\frac{d \cdot 10^3}{T' + Of + \frac{1}{2}\Delta T'}\right))$$

This vector introduces the M-variable $\Delta T'$ called the T'-incrementing control. Since ΔT will be compensated by the term $-N \cdot \Delta T$ occurring in F_4 the triple T', Of, $\Delta T'$ is the *pitch control.* The values of F_4 will not depend upon T'; accordingly raising and lowering the pitch over the time interval d may be defined by $\Delta T'$(us) but a $(T' + Of)$-dependent portamento rate is not yet available.

Interpretation: In the frame of indo-european culture a minimum requirement is satisfied by all its musical sound concepts; it consists of the independent controllability of at least the time duration and the pitch. That is precisely what vector $\vec{m}_2$ offers.*

---

* The way Janssen/Kaegi p. 215 applied the vector $\vec{m}_1$ for musical purposes might thus seem to be in contradiction with the above requirement. However, they assumed that for $\vec{m}_1$ it holds: $\Delta T = \Delta M = 0$. For $\Delta T \neq 0$ e.g. the vector $\vec{m}_1$ would any more be suitable (since not satisfying the independent control of pitch) and it would have to be replaced by at least the vector $\vec{m}_2$ (or an even more refined one).

3.3 Portamento

$$\vec{m}_3 = (\lambda\Phi_6)T', Of, P\ .\ \vec{m}_2 = (\Phi_6 F_4 F_3 F_{12})P, \ldots \vec{m}_0 \tag{3.4}$$
$$= \lambda T, \Delta T, P, N, Of, T', d\ .\ (T, \Delta T, F_3(T, N, Of, T'), F_4(\Delta T, N, \Phi_6(T', Of, P)), \ldots, N, \ldots, F_{12}(T, F_3(T, N, Of, T'), F_4(\Delta T, N, \Phi_6(T', Of, P)), \Delta T, N, d))$$
$$= \lambda T, \Delta T, P, N, Of, T', d\ .\ (T, \Delta T, T' + Of - N \cdot T, T'(R^P - 1) - N \cdot \Delta T, \ldots, N, \ldots, \mathrm{rof}\left(\frac{d \cdot 10^3}{(T' + Of)(1 + \frac{R^P - 1}{2})}\right))$$

The M-variable P(cent) introduced provides a $(T' + Of)$-dependent *Portamento control* over the time interval d (for any pitch). Fig. 3.2. We still mention a subset of $\vec{m}_3$ which will later be applied (par. 4.2.1):

$$\vec{m}_{31} = (\lambda\Phi_4)P, T'\ .\ \vec{m}_3 = (\lambda\Phi_4\Phi_6 F_4 F_3 F_{12})\ldots\vec{m}_0 \tag{3.5}$$
$$= \lambda N, T, \Delta T, P, T', d\ .\ (T, \Delta T, F_3(N, T, \Phi_4(P, T')), F_4(N, \Delta T, \Phi_6(T', \Phi_4(P, T'), P)), \ldots, N, \ldots, F_{12}(\Phi_6(T', \Phi_4(P, T'), P), \Phi_4(P, T'), d, T'))$$

Since Of now equals $-\Delta T'$ the time-function over d starts at the fundamental frequency $f_{\text{init}}(\text{Hz}) = 10^6/(T' - \Delta T')$ and ends at $f_{\text{fin}} = 10^6/T'$.

Interpretation: The intonations occurring in speech and vocal music, but also in performing any kind of unfretted string instruments etc.

3.4 Ω- or Formant-Shifting

$$\vec{m}_4 = (\lambda G_2)T, FS\ .\ \vec{m}_3 = (\lambda G_2\Phi_6 F_4 F_3 F_{12})\ldots\vec{m}_0 \tag{3.6}$$
$$= \lambda T, FS, P, N, Of, T', d\ .\ (T, G_2(T, FS), F_3(T, N, Of, T'), F_4(G_2(T, FS), N, \Phi_6(T', Of, P), \ldots, N, F_{12}(T, F_3(T, N, Of, T'), F_4(G_2(T, FS), N, \Phi_6(T', Of, P)), G_2(T, FS), N, d))$$
$$= \lambda T, FS, P, N, Of, T', d\ .\ (T, T(R^{FS} - 1), T' + Of - N \cdot T, T'(R^P - 1) - N \cdot T(R^{FS} - 1), \ldots, N, \ldots, \mathrm{rof}\left(\frac{d \cdot 10^3}{(T' + Of)(1 + \frac{R^P - 1}{2})}\right))$$

Since the values of G_2 will depend upon T and FS the M-variable FS(cent) introduced by this vector will be called the *Ω- or Formant-Shift control* over the time interval d. Fig. 3.1. Obviously the function G_2 may also be substituted into the vector $\vec{m}_{31}$.

Interpretation: The λ-operator of $\vec{m}_4$ shows that pitch, formant, formant-shifting and portamento may be controlled independently of each other within the stated domains. Fig. 3.3. It is obvious that in the indo-european culture all sound concepts applied in speech as well as in vocal and instrumental music satisfy this independency as well. In order to illustrate this we may for instance think of the concepts applied in highly stylized spoken stage languages encountered in the dramatic arts, of the pitch-independent formant control in some Tibetan vocal music (where the pitch is always tied by the same constant, cf. 5.1), of the rigorous formant stability required in the style of Belcanto (despite the virtuosity of the pitch patterns applied), of the extremely precise timbral definitions of all highly stylized instrumental performances etc. We may thus conclude that with the vector $\vec{m}_4$ we took an important step towards the interpretation aimed at by the MIDIM language.

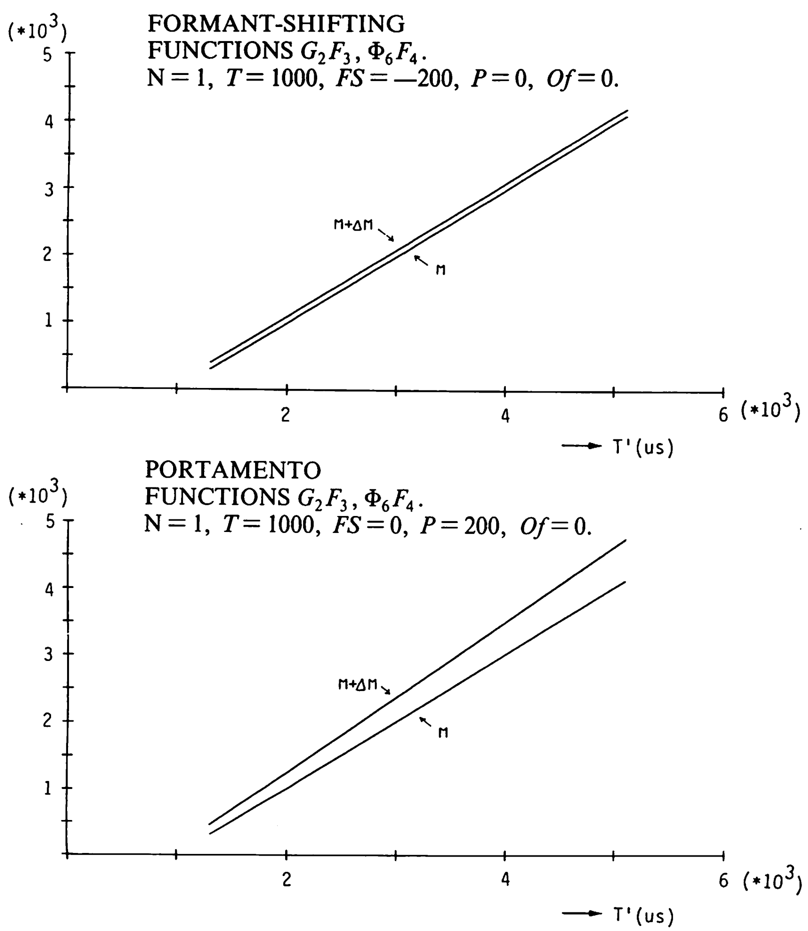


Figure 3.1. and 3.2. Functions G_2F_3 and Φ_6F_4. $P = 0$ and $FS \neq 0$ results in formant-shifting (top). The graph shows how ΔT (expressed by FS via G_2) is compensated by ΔM in order to keep the pitch constant. $P \neq 0$ and $FS = 0$ results in portamento (bottom). Notice that the values for M are the same for both graphs.

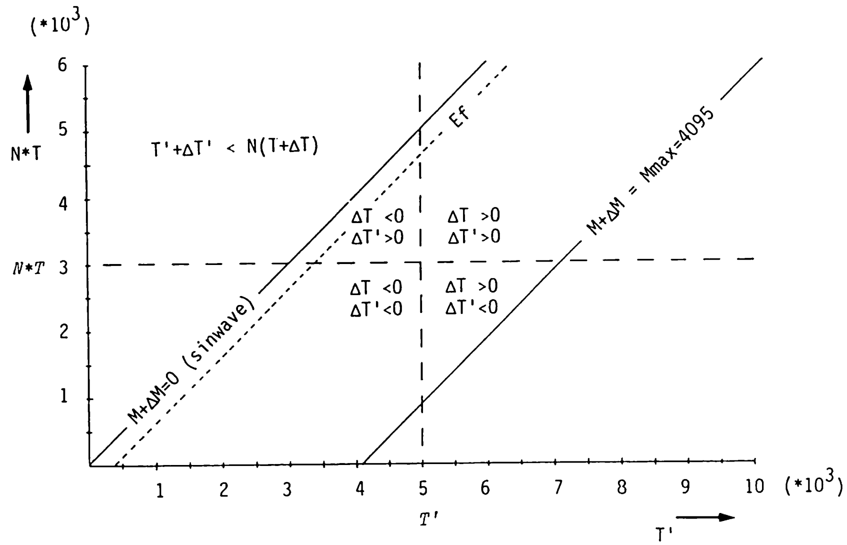


Figure 3.3. The range for ΔT and $\Delta T'$. For a given initial state $(T', N * T)$ indicated by the intersection of the broken horizontal and vertical lines, the final state $(T' + \Delta T', N(T + \Delta T))$ has to ly between the boundary lines Ef and $M + \Delta M = M\text{max}$. The domain $M = [0, 4095]$ adopted for the VOSIM-oscillators of the first generation has been extended to 13 bit for the oscillators of the third generation used for MIDIM9).

3.5 *M*-Modulation

For simplicity we shall first operate upon $\vec{m}_{11}$.

$$\begin{aligned}\vec{m}_{501} &= (\lambda F_9)W, Of, T' \;.\; \vec{m}_{11} = (\lambda F_9 F_3 F_{12})\ldots\vec{m}_0 \qquad (3.7)\\ &= \lambda S, MF, W, T, \Delta M, \Delta T, N, Of, T', d \cdot (T, \Delta T, F_3(T, N, Of, T'),\\ &\quad \Delta M, \ldots, N, F_9(Of, W, T'), S, MF, F_{12}(T, F_3(T, N, Of, T'),\\ &\quad \Delta M, \Delta T, N, d))\\ &= \lambda S, MF, W, T, \Delta M, \Delta T, N, Of, T', d \cdot (T, \Delta T, T' + Of - N \cdot T,\\ &\quad \Delta M, \ldots, N, (T' + Of)(R^W - 1), S, MF,\\ &\quad \text{rof}\left(\frac{d \cdot 10^3}{T' + Of + \tfrac{1}{2}(\Delta M + N \cdot \Delta T)}\right))\end{aligned}$$

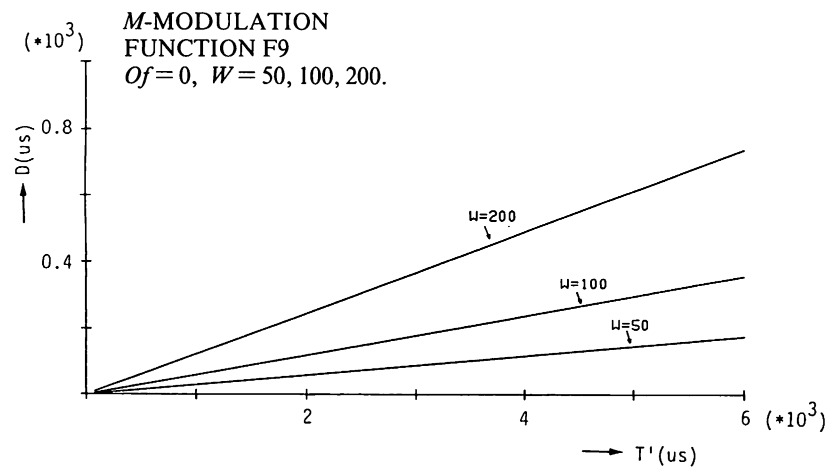
The *M*-variable *W*(cent) is the *Modulation-width control.* Since *MF* is not replaced by a function, the modulation frequency will be equal to f_1/MF (the *MF*th subharmonic) over the time interval *d*. Fig. 3.4.

$$\vec{m}_{502} = (\lambda G_9)M, y \;.\; \vec{m}_{11} = (\lambda G_9 F_3 F_{12})\ldots\vec{m}_0 \qquad (3.8)$$

is like $\vec{m}_{501}$ but now the modulation width W is replaced by $G_9 = F_3 \cdot y, 0 \le y \le 1$.

$$\begin{aligned}\vec{m}_{503} &= (\lambda F_{11})\, Sp,\, Of,\, T' \,.\, \vec{m}_{501} = (\lambda F_{11} F_9 F_3 F_{12}) \ldots \vec{m}_0 \\ &= \lambda S, Sp, W, T, \Delta M, \Delta T, N, Of, T', d \cdot (T, \Delta T, T' + Of - N \cdot T, \\ &\quad \Delta M, \ldots, N, (T' + Of)(R^W - 1), S, \mathrm{rof}(10^6/(Sp\,(T' + Of)), \\ &\quad \mathrm{rof}\left(\frac{d \cdot 10^3}{T' + Of, \tfrac{1}{2}(\Delta M + N \cdot \Delta T)}\right)\end{aligned} \tag{3.9}$$

The *M*-variable Sp (Hz) introduced by substituting F_{11} into *MF* is the *Modulation-speed control.* $\vec{m}_{503}$ provides an independent control of the *M*-modulation by means of the two *M*-variables *W* and *Sp* and the *V*-variable *S* (shape of the modulation). Obviously for $S = 0$ (P-random) the control *FM* is meaningless. Fig. 3.5.



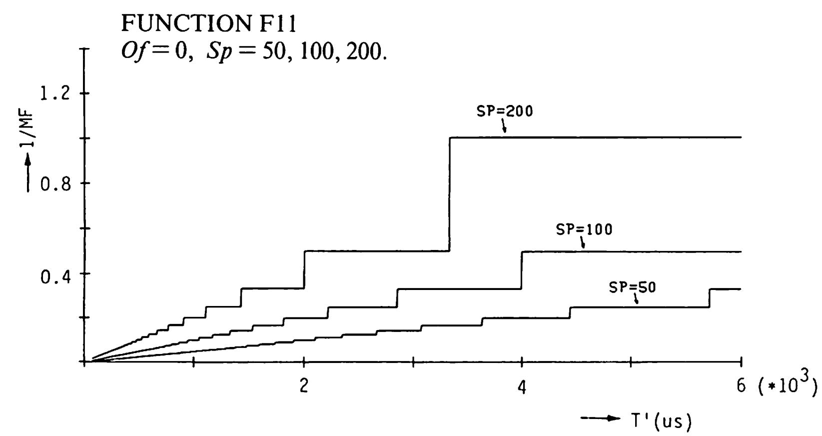


Figure 3.4. and 3.5. Functions F_9 (top) and F_{11} (bottom). *M*-modulation.

In an analogous way we may also form:

$$\vec{m}_{504} = (\lambda F_{11})\, Sp, Of, T' \,.\, \vec{m}_{502} \qquad = (\lambda F_{11}\, G_9 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.10)$$

Eventually we proceed to:

$$\vec{m}_{51} = (\lambda F_9)\, W, Of, T' \,.\, \vec{m}_4 \qquad = (\lambda F_9\, G_2 \Phi_6 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.11)$$

$$\vec{m}_{52} = (\lambda F_3 G_9)\, T, N, Of, T', M, \underline{y} \,.\, \vec{m}_4 = (\lambda G_2 \Phi_6 F_4 F_3 F_9 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.12)$$

$$\vec{m}_{53} = (\lambda F_{11} F_9)\, W, Sp, Of, T' \,.\, \vec{m}_4 \qquad = (\lambda F_{11} F_9\, G_2 \Phi_6 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.13)$$

$$\vec{m}_{54} = (\lambda F_{11} F_3 G_9)\, T, N, Of, T', M, y, Sp \,.\, \vec{m}_4 = (\lambda F_{11}\, G_2 \Phi_6 F_4 F_3\, G_9 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.14)$$

which enlarges the power of $\vec{m}_4$ by the M-modulation facilities.

Interpretation: For $W > 0$ a distinction is made between the vectors provided with a periodical (triangular) modulation ($S = 1$) and the vectors provided with a pseudo-random modulation ($S = 0$).

1a) $S = 1$ and $MF =$ const. will convey to the sound the MFth subharmonic (as already mentioned above), the amplitude ratio of the latter and the first harmonic still dependent upon W.

1b) $S = 1$ and $\Phi_{11}(Sp, Of, T')$: here the domain of Sp may be divided into subdomains which can be associated with sound properties established in music and speech. (However, the following does not pretend to be more than a rough sketch.)

$10^3/Sp > d$(ms) will convey to the segment less than one modulation cycle and may be suitable for the simulation of non-linear intonation patterns over one segment (also in combination with $\Delta T'$).

$Sp \in [3, 10]$ will represent the vibrato provided an appropriate value for W is supplied. ($Sp \in [3, 5]$ for wood winds, $Sp \in [5, 7]$ for femal singers (with an increasing W for higher pitches which for a dramatic Wagner singer may attain more than 300 Cent at ca. 880 (Hz)!), $Sp \in [7, 10]$ for string instruments).*

$Sp \in [20, 40]$ represents roughness, e.g. occurring in the attack of brass instruments.

$Sp > 40$ will entail FM-effects, although we bear in mind that FM is different from M-modulation. The idea of introducing a modulation index came, of course, into my mind, but it has not (yet) been realized because of the complexity of an analytical representation of the corresponding VOSIM signal spectrum.

2) $S = 0$ will open the way to the representation of any kind of noise bands, the width dependent upon W. $T'(t)$ is now the center frequency. Cf. the linguistic fricatives presented on p. 120 and the attacks of the gender sound in Janssen/Kaegi on p. 206 ff.

---

* In the new MIDIM9-system the phase angle of the periodical modulation can be controlled (rendering possible e.g. the duplication of different types of vibrato in violin playing (Russian, Belgian styles etc.)).

3.6 "Missing Harmonics" and a *(T' + Of)*-dependent (moving) formant

$$\begin{aligned}\vec{m}_6 &= (\lambda F_1)\,N, q, Of, T' \,.\, \vec{m}_2 = (\lambda F_1 F_4 F_3 F_{12})\ldots, d \,.\, \vec{m}_0 \qquad (3.15)\\ &= \lambda q, C, \Delta T, \Delta T', N, Of, T', d \,.\, (F_1(N, q, Of, T'), \Delta T,\\ &\quad F_3(F_1(N, q, Of, T'), N, Of, T'), F_4(\Delta T, N, \Delta T'), \ldots, N, \ldots,\\ &\quad (F_{12}(F_1(N, q, Of, T'), F_3(F_1(N, q, Of, T'), F_4(\Delta T, N, \Delta T'), N, d))\\ &= \lambda q, C, \Delta T, \Delta T', N, Of, T', d \,.\, (q(T' + Of)/N, \Delta T, (T + Of)(1 - q),\\ &\quad \Delta T' - N \cdot \Delta T, \ldots, C, N, \ldots,\\ &\quad \mathrm{rof}\left(\frac{d \cdot 10^3}{T' + Of + \frac{1}{2}\Delta T'}\right))\end{aligned}$$

With the substitution of F_1 into T the M-variable q ($0 < q < 1$) is introduced. It stands for the ratio $N \cdot T/T'$ ($= \alpha/(\alpha+\beta)$, α and β minimal, cf. actually p. 77). The harmonics of order $(\alpha+\beta)$ and its multiples will coincide with possible "minima" occurring in the envelope of the spectrum at the frequencies $n \cdot \Omega/N$, $n = 1, 2, 3, \ldots$, ($n \neq N$), the amplitudes depending upon C. However, since ΔT and $\Delta T'$ are obviously independent of q, the ratio designated by q will merely apply to the initial state of the signal within the time interval d, unless $\Delta T = \Delta T' = 0$.

$$\begin{aligned}\vec{m}_{61} &= (\lambda F_2)\,N, q, \Delta T' \,.\, \vec{m}_6 = (\lambda F_2 F_1 F_4 F_3 F_{12})\ldots, d \,.\, \vec{m}_0 \qquad (3.16)\\ &= \lambda q, C, \Delta T', N, Of, T', d \,.\, (F_1(N, q, Of, T'), F_2(N, q, \Delta T'),\\ &\quad F_3(F_1(N, q, Of, T'), N, Of, T'), F_4(F_2(N, q, \Delta T'), N, \Delta T'), \ldots,\\ &\quad C, N, \ldots, F_{12}(F_1(N, q, Of, T'), F_4(F_2(N, q, \Delta T'), N, \Delta T'), N, d))\\ &= \lambda q, C, \Delta T', N, Of, T', d \,.\, (q(T' + Of)/N, q \cdot \Delta T'/N,\\ &\quad (T + Of)(1 - q), \Delta T'(1 - q), \ldots, C, N, \ldots,\\ &\quad \mathrm{rof}\left(\frac{d \cdot 10^3}{T' + Of + \frac{1}{2}\Delta T'}\right))\end{aligned}$$

Replacing ΔT by F_2 this vector provides the *"missing-harmonic"-control* q for every $\Delta T'$, since it substitutes the q, $\Delta T'$, N-dependent product $F_2 F_4$ into ΔM. Eventually the Portamento may be introduced as follows (Fig. 3.6.):

$$\begin{aligned}\vec{m}_{62} &= (\lambda \Phi_6)T', Of, P \,.\, \vec{m}_{61} = (\lambda \Phi_6 F_2 F_1 F_4 F_3 F_{12})\ldots, d \,.\, \vec{m}_0 \qquad (3.17)\\ &= \lambda q, C, P, N, Of, T', d \,.\, (q(T' + Of)/N,\\ &\quad q(T' + Of)(R^P - 1)/N, (T + Of)(1 - q), (T' + Of)(R^P - 1)(1 - q),\\ &\quad \ldots, C, N, \ldots, \mathrm{rof}\left(\frac{d \cdot 10^3}{(T' + Of)(1 + \frac{R^P - 1}{2})}\right))\end{aligned}$$

M-modulation may, of course, be applied to $\vec{m}_{62}$ by means of the corresponding substitutions (where $()_i$ stands for the operand of the vector $\vec{m}_i$):

$$\vec{m}_{63} = (\lambda F_9)W, Of, T' \; . \; \vec{m}_{62} \tag{3.18}$$
$$\vec{m}_{64} = (\lambda F_1\, G_9)N, q, Of, T', M, y \; . \; \vec{m}_{62} \tag{3.19}$$
$$\vec{m}_{65} = (\lambda F_{11} F_9)W, Sp, Of, T' \; . \; \vec{m}_{62} \tag{3.20}$$
$$= (\lambda F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12}) \ldots, d \; . \; \vec{m}_0$$
$$= \lambda W, Sp, S, q, C, P, N, Of, T', d \; . \; ()_{\underline{65}}$$
$$\vec{m}_{66} = (\lambda F_{11} F_1\, G_9)N, q, Of, T', Sp, M, y \; . \; \vec{m}_{62} \tag{3.21}$$

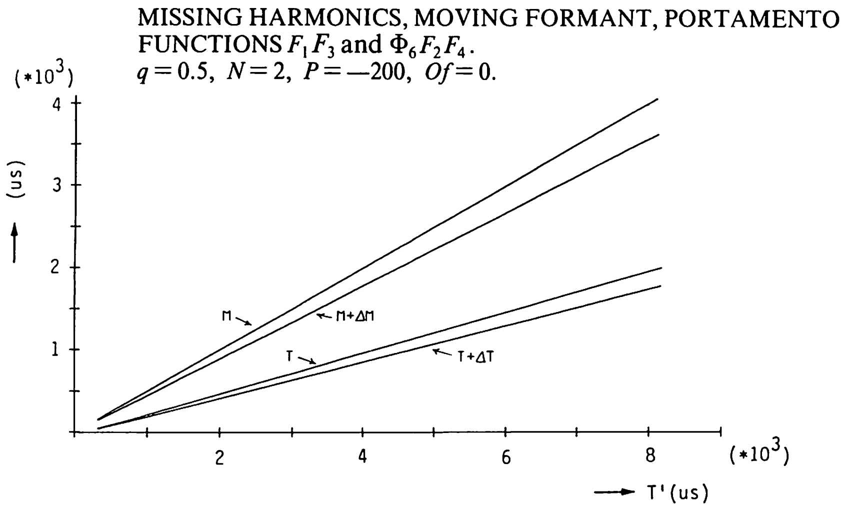


Figure 3.6. Functions $F_1 F_3$ and $\Phi_6 F_2 F_4$. Missing harmonics, moving formant and portamento. ΔM is compensated by $\Delta T (= F_2)$ in order to satisfy the constant q.

Here are some interpretations:

1) Representation of metallophones, e.g.:

javanese gender sounds

$$\vec{m}_{651} = (\lambda F_6 F_5) y_2, e, Am, \Delta Am, At \; . \; \vec{m}_{65} \tag{3.22}$$
$$= (\lambda F_6 F_5 F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12}) \ldots, d \; . \; \vec{m}_0$$
$$= \lambda N, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \; . \; ()_{651}$$

Cf. Janssen/Kaegi, p. 215.

bell- and gonglike sounds e.g.

$$\vec{m}_{641} = (\lambda F_6 F_5)\ldots\vec{m}_{64} \quad \text{cf. par. 3.9 and 5.3, the predicator } P(3)_{\text{Bel}} \qquad (3.23)$$

2) Representation of wind instruments with a moving formant, e.g.:

$$\begin{aligned}\vec{m}_{652} &= (\vec{m}_{651})(1, 1, 1, 0, 0, 1001/1003, c_N) \qquad (3.24)\\ &= (\lambda N, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \, . \, ()_{651})\\ &\quad (1, 1, 1, 0, 0, 1001/1003, c_N)\end{aligned}$$

Since q here is almost equal to 1 the harmonic $f_N = N \cdot f_1$ will be close to $F_\Omega = N \cdot f_1 \cdot q^{-1}$, i.e. the Nth harmonic will be stressed. Cf. pag. 126, the predicator $P(2)_{\text{Clar}}$, track 2.

3) Representation of linguistic fricatives, cf. par. 5.1 p. 120.

3.7 "Missing Harmonics" and a fixed Formant-Area

$$\vec{m}_7 = (\lambda F_8)Of, T\text{max}, q, T' \, . \, \vec{m}_6 = (\lambda F_8 F_1 F_4 F_3 F_{12})\ldots, d \, . \, \vec{m}_0 \qquad (3.25)$$
$$\vec{m}_{71} = (\lambda F_8)Of, T\text{max}, q, T' \, . \, \vec{m}_{61} = (\lambda F_8 F_2 F_1 F_4 F_3 F_{12})\ldots, d \, . \, \vec{m}_0 \qquad (3.26)$$
$$\vec{m}_{72} = (\lambda F_8)Of, T\text{max}, q, T' \, . \, \vec{m}_{62} = (\lambda F_8 \Phi_6 F_2 F_1 F_4 F_3 F_{12})\ldots, d \, . \, \vec{m}_0 \qquad (3.27)$$

and likewise for $\vec{m}_{63}$ through $\vec{m}_{66}$.

By substituting the function F_8 into the V-variable N occurring in the vectors presented in par. 3.6, we may convey to them the property of satisfying the λ-tied variable q ("missing harmonics" control) as well as a formant-area defined by the M-variable Tmax. Introduced by means of F_8, the λ-tied variable Tmax is called the *formant-area control.* Fig. 3.7.-3.9.

An interpretation: the representation of a b-flat clarinet (cf. par. 3.6).

$$\begin{aligned}\vec{m}_{751} &= (\lambda F_8)Of, T\text{max}, q, T' \, . \, \vec{m}_{651} \qquad (3.28)\\ &= (\lambda F_8 F_6 F_5 F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12})\ldots, d \, . \, \vec{m}_0\\ &= \lambda T\text{max}, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \, . \, ()_{751}\\ \vec{m}_{752} &= (\vec{m}_{751})(1, 1, 1, 0, 0, \tfrac{1}{2}, 460) \qquad (3.29)\\ &= (\lambda T\text{max}, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \, . \, ()_{751})\\ &\quad (1, 1, 1, 0, 0, \tfrac{1}{2}, 460)\end{aligned}$$

Since $q = \frac{1}{2}$ and $C = .9$ the even harmonics are almost entirely suppressed except for $N \cdot f_1$. N being a function entails $T \in [T\text{max}, T\text{min}] = [460, 230]$, where $10^6/T\text{max} = 2173.91$ (Hz). Cf. par. 5.3, the predicator $P(2)_{\text{Clar}}$, track 1.

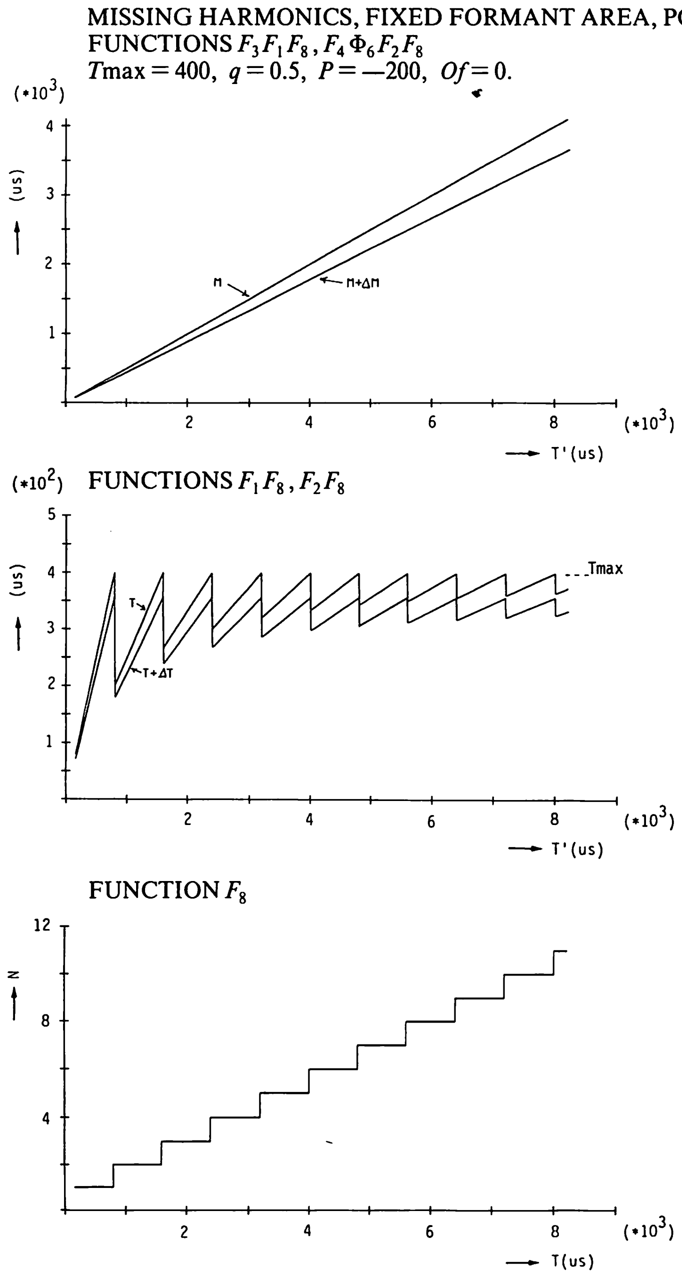


Figure 3.7.-3.9. Functions F_8 (bottom), F_1F_8 and F_2F_8 (middle) and $F_3F_1F_8$ and $F_4\Phi_6F_2F_8$ (top). Missing harmonics, a fixed formant area and portamento. The graph shows that ΔM (top) is compensated by $\Delta T (= F_2)$ in order to satisfy the constant q.

3.8 "Fixed Formant"

$$\vec{m}_8 = (\lambda G_8)\Delta T, D, Ef, N\max, \Delta T', Of, T, T', d \,.\, \vec{m}_2$$
$$= (\lambda G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (3.30)$$

The vector introduces the *M*-variables *Ef* and *N*max. Since it holds that:

$$(T' + Of) \in [T \cdot N\min + EF, T \cdot N\max + M\max]$$

the variables *Ef* and *N*max are called the *(T' + Of)- or pitch-domain controls*. Cf. par. 8.2. In the spectrum Ω will depend upon $T(\Omega(\cdot\ 10^6\text{rad}/s) = 2\pi/T)$, and for $N > 1$ the amplitude maximum of the envelope in the neighbourhood of Ω will depend upon *N*, whereas the envelope itself will still be dependent upon *C*. Accordingly the *V*-variables *T*, *N* and *C* are now called the *"fixed formant"-controls*. Fig. 3.10. and 3.11.

In an analogous way we may form:

$$\vec{m}_{81} = (\lambda \Phi_6)T', Of, P \,.\, \vec{m}_8 = (\lambda \Phi_6\, G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \ \ (\text{cf. } \vec{m}_3) \qquad (3.31)$$

introducing the portamento control *P*, and:

$$\vec{m}_{82} = (\lambda G_2)T, FS \,.\, \vec{m}_{81} = (\lambda G_2\, \Phi_6\, G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \ \ (\text{cf. } \vec{m}_4) \qquad (3.32)$$

introducing the Ω- or formant-shift-control *FS*, and eventually we may still introduce *M*-modulation facilities by setting up the following:

$$\vec{m}_{83} = (\lambda F_9)W, Of, T' \,.\, \vec{m}_{82} \qquad (3.33)$$
$$= (\lambda F_9\, G_2\, \Phi_6\, G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (\text{cf. } \vec{m}_{51})$$
$$\vec{m}_{84} = (\lambda F_3\, G_9)T, N, Of, T', M, y \,.\, \vec{m}_{82} \qquad (3.34)$$
$$= (\lambda G_2\, \Phi_6\, G_8 F_4 F_3\, G_9 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (\text{cf. } \vec{m}_{52})$$
$$\vec{m}_{85} = (\lambda F_{11} F_9)W, Sp, Of, T' \,.\, \vec{m}_{82} \qquad (3.35)$$
$$= (\lambda F_{11} F_9\, G_2\, \Phi_6\, G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (\text{cf. } \vec{m}_{53})$$
$$\vec{m}_{86} = (\lambda F_{11} F_3\, G_9)T, N, Of, T', M, y \,.\, \vec{m}_{82} \qquad (3.36)$$
$$= (\lambda F_{11}\, G_2\, \Phi_6\, G_8 F_4 F_3\, G_9 F_{12}) \ldots, d \,.\, \vec{m}_0 \qquad (\text{cf. } \vec{m}_{54})$$

Here are some interpretations. The representations of:
1) vowels cf. par. 5.2 p. 123, and v. Berkel, p. 249 of this issue.
2) a harpsichord

$$\vec{m}_{800} = (\lambda F_6 F_5)y_2, e, \Delta Am, Am, At \cdot \vec{m}_8 \qquad (3.37)$$
$$= \lambda D, Of, \Delta T', Ef, y_2, e, N\max, \Delta Am, Am, \Delta T, T, At, T', d \,.\, ()_{800}$$
$$\vec{m}_{801} = (\vec{m}_{800})(49, 1, 1, 10, 0, 0, 0) \qquad (3.38)$$
$$= (\lambda D, Of, \Delta T', Ef, y_2, e, N\max, \Delta Am, Am, \Delta T, T, At, T', d \,.\, ()_{800})$$
$$(49, 1, 1, 10, 0, 0, 0)$$

cf. par. 5.3 the predicator $P(2)_{\text{Cem}}$, (5.24)

3) a basson

$$\vec{m}_{850} = (\lambda F_6 F_5) \ldots \vec{m}_{85} \qquad (3.39)$$

cf. par. 3.9 and 5.3, (5.23)

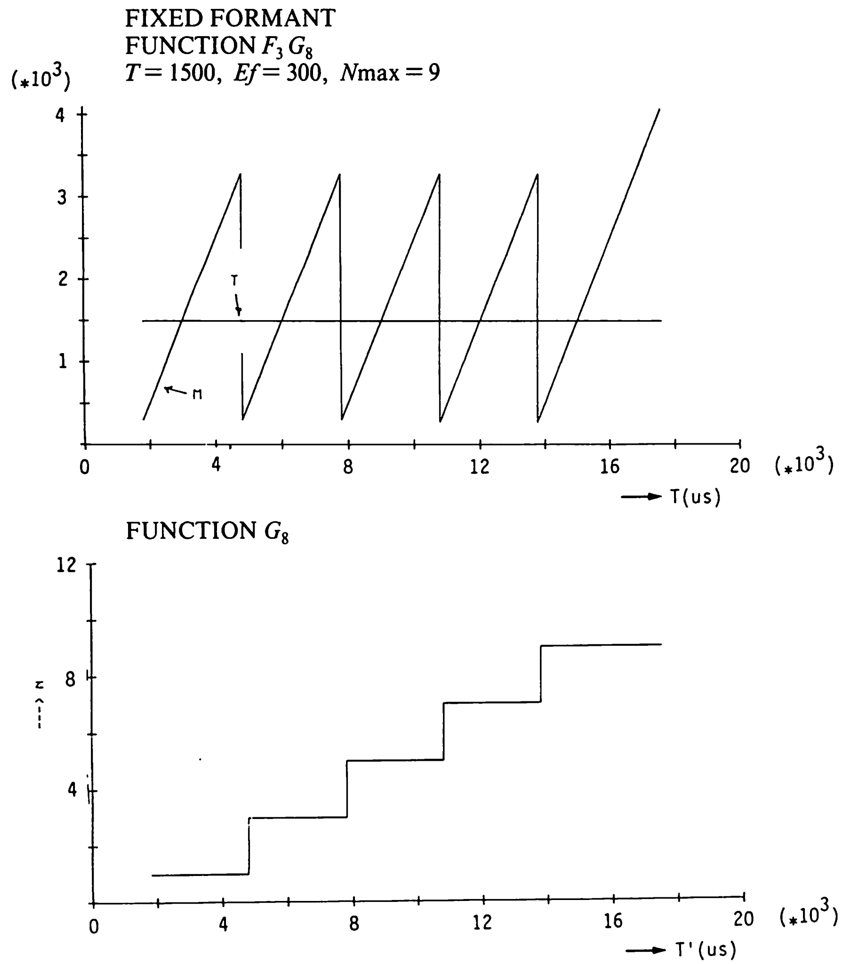


Figure 3.10. and 3.11. Function G_8 (bottom) and $G_8 F_3$ (top).

3.9 Amplitude Envelope and Attenuation

$$
\begin{aligned}
\vec{m}_9 &= (\lambda F_6 F_5) y_2, e, At, Am, \Delta Am\ .\ \vec{m}_1 = (\lambda F_6 F_5 F_{12}) \ldots, d\ .\ \vec{m}_0 \qquad (3.40)\\
&= \lambda Am, \Delta Am, T, \Delta T, M, \Delta M, N, y_2, e, At, d\ .\\
&\quad (T, \Delta T, M, \Delta M, ((y_2 - 1) At^e + 1) Am \cdot At,\\
&\quad ((y_2 - 1) At^e + 1) Am \cdot At, \ldots, N, \ldots,\\
&\quad \mathrm{rof}\left(\frac{d \cdot 10^3}{N \cdot T + M + \tfrac{1}{2}(\Delta M + N \cdot \Delta T)}\right))
\end{aligned}
$$

Among the M-variables introduced by this vector the variables Am, ΔAm are called the *relative amplitude-envelope controls*, and y_2, e, At are called the *attenuation controls.* It has already been mentioned (par. 2.3.3) that for any e it holds:

$$\vec{m}_{91} = (\lambda y_2, e, At, Am, \Delta Am, T, \Delta T, M, \Delta M, N, d\ .\ ()_9)(1) \qquad (3.41)$$

is equivalent with:

$$\lambda At, Am, \Delta Am, T, \Delta T, M, \Delta M, N, d\ .\ (T, \Delta T, M, \Delta M, At \cdot Am, At \cdot \Delta Am, \ldots, N, \ldots, F_{12}(N, T, M, \Delta M, \Delta T, \mathrm{d}))$$

It goes, of course, without saying that by the application of $(\lambda F_6 F_5)$ any arbitrary M-vector which satisfies the M-concept $C(v)$ may be provided with the above facilities for controlling the amplitude envelope.

By eventually substituting the function Φ_{12} into ΔAm we introduce the *control* α *of the amplitude in(de-)crement*, the latter depending also now upon the time duration d. This substitution obviously makes sense when d is replaced by the function Φ_5, thus in any segement $A = S_v$. Here is an example:

$$S_v = \vec{m}_{90} = (\lambda \Phi_5 \Phi_{12}) \ldots \vec{m}_9 \qquad (3.42)$$

which entails:

$$\Delta Am_v = \alpha \cdot (DUR - \sum_{\substack{i=1 \\ i \neq v}}^{4} d_i) \quad \text{(cf. par. 4.1)}$$

An interpretation: e.g. the decay of any plucked or struck sound.

Here are still more formal examples:

$$\begin{aligned} \vec{m}_{41} &= (\lambda F_6 F_5) y_2, e, At, Am, \Delta Am\ .\ \vec{m}_4 \qquad (3.43) \\ &= (\lambda F_6 F_5 G_2 \Phi_6 F_4 F_3 F_{12}) \ldots, d\ .\ \vec{m}_0 \\ &= \lambda Am, \Delta Am, T, N, FS, P, Of, y_2, e, T', d\ .\ (T, G_2(T, FS), \\ &\quad F_3(T, N, Of, T'), F_4(G_2(T, FS), N, \Phi_6(T', Of, P)), \\ &\quad F_5(y_2, e, At, Am), F_6(y_2, e, At, \Delta Am), \ldots, N, \ldots, F_{120}(P, Of, T', d)) \end{aligned}$$

for F_{120} see par. 2.3, cf. also par. 4.1.

$$\begin{aligned} \vec{m}_{641} &= (\lambda F_6 F_5) \ldots \vec{m}_{64} = (\lambda F_6 F_5 F_1 G_9) \ldots \vec{m}_{62} \qquad (3.44) \\ &= (\lambda F_6 F_5 F_1 G_9 \Phi_6) \ldots \vec{m}_{61} = (\lambda F_6 F_5 F_1 G_9 \Phi_6 F_2) \ldots \vec{m}_6 \\ &= (\lambda F_6 F_5 F_1 G_9 \Phi_6 F_4 F_3 F_{12}) \ldots, d\ .\ \vec{m}_0 \end{aligned}$$

$$\begin{aligned} \vec{m}_{651} &= (\lambda F_6 F_5) y_2, e, Am, \Delta Am, At\ .\ \vec{m}_{65} \qquad (3.45) \\ &= (\lambda F_6 F_5 F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12}) \ldots, d\ .\ \vec{m}_0 \\ &= \lambda N, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T'\ .\ ()_{651} \end{aligned}$$

$$\vec{m}_{850} = (\lambda F_6 F_5) \ldots \vec{m}_{85} = (\lambda F_6 F_5 F_{11} F_9 G_2 \Phi_6 G_8 F_4 F_3 F_{12}) \ldots, d\ .\ \vec{m}_0$$

(cf. 3.39) and likewise for

$$\vec{m}_{860} = (\lambda F_6 F_5) \ldots \vec{m}_{86} \qquad (3.46)$$

Cf. also par. 5.3, the predicators $P(3)_{\mathrm{Bel}}$ and $P(2)_{\mathrm{Cem}}$ and par. 7.

4. THE INTERPRETED *M*-CONCEPTS

By means of substitution and assigning constants (numbers) we may in a systematic way form a set of *M*-vectors which denote *relevant* subsets of *V* in the way shown above.

Relevant subsets are extensions of *M*-vectors which can be associated with descriptive intensions (properties) established in the field of application aimed at by the MIDIM-language. In this way we built up *V*-interpreted *M*-vectors and allocated to their extensions descriptive properties such as time-duration (emission time), pitch, portamento, formant-shifting, moving formant, attenuation etc. We know, however, that in most cases single *M*-vectors are not powerful enough for the representation of the temporal comportment of sound. For this reason the *M*-concept $C(v)$ was introduced which is a sequence of s segments $S_1 \ldots S_s$, where S_v is replaced by A_v (cf. par. 1.6.2). Each *M*-concept denotes a sequence of s subsets of V, that is a subset of V^*.

We now define:

A *V-interpreted concept* $C'(v)$ is any concept $C(v)$ where:

(a) exclusively *V*-interpreted *M*-vectors occur,

(b) constants are allocated to all λ-tied variables occurring except v, d_i, DUR and $\ldots, X, Y$.

For our present language the following assumptions are made:

$X = T'$
$Y = At$
$v \in [0, s]$ and $v \neq 1$
$s = 4$

Although allocating constants to v is not admitted before the concepts will be assembled in the predicator, I would right now like to give the reader an initial idea by means of the following schemes:

$$C'(0) = [\lambda v, d_i, DUR, T', At \,.\, S_1 \ldots S_4](0) = [\lambda v, T', At \,.\, SSSS](0)$$
$$C'(2) = [\lambda v, d_1, d_3, d_4, DUR, T', At \,.\, SASS](2)$$
$$C'(3) = [\lambda v, d_1, d_2, d_4, DUR, T', At \,.\, SSAS](3)$$
$$C'(4) = [\lambda v, d_1, d_2, d_3, DUR, T', At \,.\, SSSA](4)$$

The segments occurring in a concept $C'(v)$ may then be assigned *V*-interpreted *M*-vectors as follows (where x, y, z, w are not necessarily different):

$$(C'(v))(\vec{m}_x, \vec{m}_y, \vec{m}_z, \vec{m}_w) \tag{4.1}$$

Interpretation: $v \neq 1$ and $s = 4$ assumed by the present language entails the following:

(a) there is no concept where the time duration d_1 of the first segement S_1 is a function. Reason: the recognizability of the concept. Cf. Fig. 1.3 on p. 91.

(b) all the concepts are of the same degree $s = 4$. Reason: MIDIM aimes at a minimum description adopting the basic model: prefix, body, suffix, stop.

The parameters T' and At are introduced into the outermost λ-operator in view of the field of application chosen for our language. Cf. par. 5.1.

In analogy with the M-vectors presented in par. 3.1 through 3.9 we may now form numerous V-interpreted concepts $\subset V^*$. Once again we shall associate them with properties established in the field of application. Here is a rough sketch:

4.1 An amplitude-envelope generator

$$(C'(v))(\vec{m}_9)_i = (C'(v))(\vec{m}_9, \vec{m}_9, \vec{m}_9, \vec{m}_9) \quad 1 \leq i \leq 4 \tag{4.2}$$

is (for any v) an amplitude-envelope "generator" operating over the time duration DUR. And here is another example:

$$(C'(3))(\vec{m}_{41})\,(c_{\Delta Am}, c_{Am})_i \;\; \begin{array}{rr|l} 500 & 0 & 1 \\ -100 & 500 & 2 \\ -380 & 400 & 3 \\ -20 & 20 & 4 \end{array} \;=\; \begin{array}{l} \qquad\qquad\quad c_{\Delta Am}, \quad c_{Am} \\ \hline (C'(3))((\vec{m}_{41})(\;500,\;\; 0), \\ \qquad\quad (\vec{m}_{41})(-100, 500), \\ \qquad\quad (\vec{m}_{41})(-380, 400), \\ \qquad\quad (\vec{m}_{41})(\;-20,\;\; 20)) \end{array} \tag{4.3}$$

is an amplitude-envelope *pattern* over DUR still dependent upon $d_1, d_3, d_4, At, T', T, N, FS, P, Of, y_2$ and e. (For $\vec{m}_{41}$ see par. 3.9).

The above "generator"can still be improved by the following substitutions applied to the segments S_v and S_{v+1} respectively:

$$(\lambda \Phi_5 \Phi_{12}) \ldots (\vec{m}_9)_v \tag{4.4}$$

$$(\lambda \Phi_{11}) \ldots (\vec{m}_9)_{v+1} \quad \text{with } v < 4 \text{ or even } (\lambda \chi_{12} \Phi_{11}) \ldots (\vec{m}_9)_{v+1} \tag{4.5}$$

Accordingsly we obtain:

$$\Delta Am_v \;\; = \Phi_{12}(\alpha, \Phi_5(v, d_i, DUR)) \tag{4.6}$$

$$Am_{v+1} \;\; = \Phi_{11}(Am_v, \Delta Am_v) = (Am + \Delta Am)_v \tag{4.7}$$

$$\Delta Am_{v+1} = \chi_{12}(Am_{v+1}) = -Am_{v+1} \tag{4.8}$$

Needless to say the amplitude envelope patterns are of prime importance for characterizing the sound concepts. Cf. Janssen/Kaegi p. 212 and v. Berkel p. 238 ff.

4.2 Some more interpretations

In a similar way we may form the following *M*-concepts $C'(v)$ by assigning to each segment e.g. one of the vectors:

portamento contours (intonation)	$\vec{m}_3, \vec{m}_{31}, \vec{m}_4, \vec{m}_{62}$, through $\vec{m}_{66}$
formant tracks	$\vec{m}_4, \vec{m}_{41}$
M-modulation	$\vec{m}_{501}$ through $\vec{m}_{504}$
miss. harm., moving formant	$\vec{m}_6, \vec{m}_{61}$ or $\vec{m}_{62}$
miss. harm., fixed formant-area	$\vec{m}_7, \vec{m}_{71}$ or $\vec{m}_{72}$
fixed formant	$\vec{m}_8, \vec{m}_{81}$ through $\vec{m}_{86}$

By assigning constants in various ways we may then form any type of corresponding patterns over the time interval *DUR*.

Here is an example:

4.2.1. An elabored intonation generator

$$(C'(v))(\vec{m}_{31}, \vec{m}_{31}, \vec{m}_{31}, \vec{m}_{31}) = (C'(v))(\vec{m}_{31})_i \tag{4.9}$$

is an intonation "generator".

We now introduce:

$$\begin{aligned} \vec{m}_{32} &= (\lambda \Phi_6 \chi_4)T', Of, P, h, i, \Delta T'_j \,.\, \vec{m}_{31} \\ &= (\lambda h, i, \Delta T'_j, N, T, \Delta T, P, T', d \,.\, (T, \Delta T, F_3(N, T, \chi_4, T'), \\ &\quad F_4(N, \Delta T, \Phi_6(T', \chi_4, P)), \ldots, N, \ldots, F_{12}(\Phi_6(T', \chi_4, P), \chi_4, d, T')))(2) \end{aligned} \tag{4.10}$$

and form the concept

$$(C'(v))(\vec{m}_{31}, \vec{m}_{32}, \vec{m}_{32}, \vec{m}_{32}) \quad \text{(for } \vec{m}_{31} \text{ cf. par. 3.3)} \tag{4.11}$$

where it holds:

$$Of_1 = \Phi_4 = \lambda P_1, T' \,.\, T'(1/R^{P_1} - 1) = -\Delta T'_1 \quad \text{cf. par. 2.3 func-tab. 2} \tag{4.12}$$

and

$$Of_i = \Phi_6 \chi_4 = \begin{cases} i > 2 \rightarrow \lambda i, P_j, Of_j, T' \cdot \sum_{j=2}^{i-1} ((T' + Of_j)(R^{P_j} - 1)) \\ \text{otherwise } 0 \end{cases} \tag{4.13}$$

Accordingly we obtain:

$$Of_2 = 0 \tag{4.14}$$

$$Of_3 = \lambda T' \,.\, (T' + Of_2)(R^{P_2} - 1) = T'(R^{P_2} - 1) \tag{4.15}$$

$$Of_4 = \lambda T' \,.\, T'(R^{P_2} - 1) + (T' + Of_3)(R^{P_3} - 1) \tag{4.16}$$

The concept stands for an intonation contour (= a set of intonation patterns) which has proven to be *musically highly suitable*, since for the listener it allocates the "right pitch" to the sound irrespective of the constants allocated to it. Fig. 4.1.

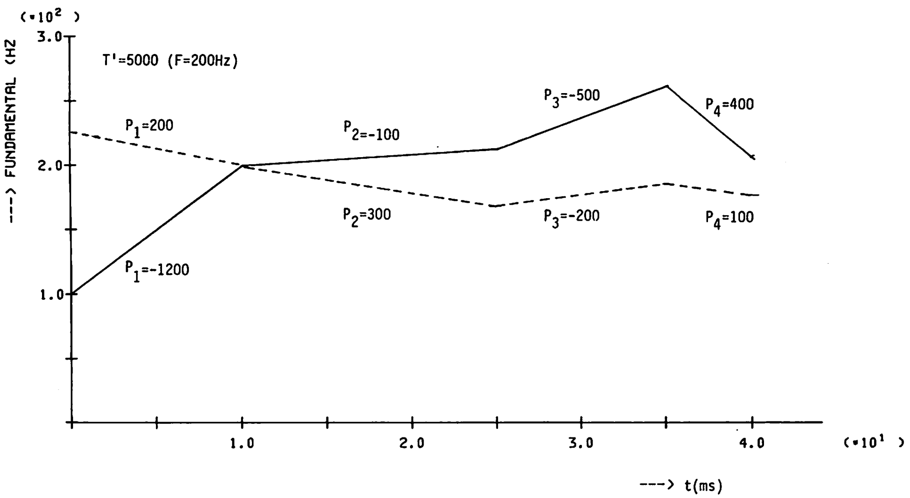


Figure 4.1. Intonation contour over the four segments S_1 to S_4. For every pattern satisfying this contour the function value at the beginning of segment S_2 is equal to T'.

For completeness we still mention the following possible substitution:

$$(\Delta T'_4)_k = \chi_6 = (\lambda h, i, T'_k, T'_{k+1}, (Of_i)_k, (Of_h)_{k+1} \cdot (T' + Of_h)_{k+1} - (T' + Of_i)_k)(4, 1) \qquad (4.17)$$

This function provides a smooth portamento between the segment i of the word k and the segment h of the word $k+1$, $k = 1$, $h = 1$, $i = 4$. Notice that it will also be possible to replace h and i by functions of the masks applied to the two words involved. Cf. par. 6.

5. THE PREDICATOR

5.1 Generalities

In the previous pages we presented a selection of M-concepts $C'(v)$ in an intensional way (that is by means of a product of substitution-operators applied to $\vec{m}_0$; to each concept corresponds, of course, uniquely an extension or subset of 4-uples $\in V^*$ of V-vectors). Furthermore we made an attempt to associate with each M-concept certain properties such as a "moving formant", "M-modulation" etc. In this way we tried to lay a link between the V-interpreted MIDIM-language and the descriptive concepts established in the field of application. E.g. for constructing the linguistic concept "unvoiced fricative" we shall refer to the properties "moving formant" and "M-modulation" by forming the following M-vector (cf. par. 3.6):

$$\begin{aligned}\vec{m}_{64} &= (\lambda F_1 G_9)N, q, Of, T', M, y \;.\; \vec{m}_{62} = (\lambda \Phi_6 F_2 F_1 G_9 F_4 F_3 F_{12})\ldots, d \;.\; m_0 \quad (5.1)\\ &= \lambda y, q, P, N, Of, S, C, T', d \;.\; (q(T'+Of)/N, q(T'+Of)(R^P-1)/N,\\ &\qquad (T'+Of)(1-q), (T'+Of)(R^P-1)(1-q),\\ &\qquad \ldots, C, N, y(T'+Of)(1-q), S, \ldots,\\ &\qquad F_{120}(P, T', Of, d))\end{aligned}$$

By assigning constants we then form:

$$\vec{m}_{641} = (\lambda y, q, P, N, Of, S, C, T', d \cdot ()_{64})(1, 0, 0, 1, 0, \tfrac{1}{2}, \tfrac{1}{2}) \qquad (5.2)$$

This new M-vector denotes a subset of V within which we shall find representations of the following unvoiced fricatives:

$$\vec{m}_{642} = (\vec{m}_{641})(c_{T'}) \qquad (5.3)$$

$c_{T'}$
160 for /s/
280 for /ç/
400 for /ʃ/

A glance at the corresponding magnitude-spectrum may be obtained by applying the function Φ_{18} as well as the formulae (8) through (11) mentioned on p. 77. Cf. also Kaegi 1974, 143-147. In order to run sound-tests one may then still allocate constants to the M-variable d (and to the λ-tied V-variables omitted in the operator by convenience) and proceed via λ-elimination to V-vectors representing sound-instances of the above fricatives.* But let us still add another M-vector:

$$\vec{m}_{643} = (\lambda y, q, P, N, Of, S, C, T', d \;.\; ()_{64})(1, 0, 0, 1, 0, 4/7, 1) \qquad (5.4)$$

* Testing instances of M-vectors in this way is not in contradiction with the definition given in par. 4.

It will embrace a large family of "wet" sounds including:

$$\vec{m}_{644} = (\vec{m}_{643})(140) \quad \text{for the fricative } /f/ \tag{5.5}$$

Up until now we have not paid much attention to the amplitude (admitting the whole domain of $Am(t)$ via the "silent" occurrence of the variables Am, ΔAm in the λ-operator). We may now refine the above vectors by conveying to them the attenuation control as well as specific constants for Am and ΔAm. Therefore we proceed as follows:

$$\begin{aligned}\vec{m}_{67} &= (\lambda F_6 F_5)y_2, e, At, Am, \Delta Am \,.\, \vec{m}_{64}(5.6) \\ &= (\lambda F_6 F_5 \Phi_6 F_2 F_1\, G_9 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0 \\ &= \lambda y, q, P, N, Of, S, C, y_2, e, T', Am, \Delta Am, At, d \,.\, ()_{67}\end{aligned} \tag{5.6}$$

$$\vec{m}_{671} = (\vec{m}_{67})(1, 1, 1, 0, 0, 1, 0, \tfrac{1}{2}, \tfrac{1}{2})(5.7) \tag{5.7}$$

In the subset denotated by $\vec{m}_{671}$ we shall, of course, find once more the sounds /s, ç, ʃ/ (by assigning to T' the corresponding constants presented above), but now they are controllable by Am, ΔAm, At and d.

Eventually we may proceed to the corresponding *M*-concepts and convey to them specific amplitude-envelope contours over *DUR* (cf. 4.1). In the same manner concepts may be formed combining different vectors in any possible way from degree 1 up until 4 (e.g. */fs/*, */sf/*, */sç/*, . . . , */fʃs/*, . . . , */çsfʃ/*, . . .).

Notice that finding the means of building *M*-concepts up like the ones presented above, is due to our knowledge of both the *V*-interpretation of the MIDIM-language as well as the mapping of the properties (e.g. to be a */s/* or a */fs/*. . .) into physical parameters. The latter information is, however, of a *descriptive* nature and thus *external* to the frame of our *V*-interpreted formal language. (A thorough discussion of how to incorporate descriptive information into the MIDIM is found in Janssen/Kaegi 1986 on p. 185 of this volume.) The important point is, however, that *once introduced into the formal language any descriptive information will become a part of the semantics of the language.* Accordingly the latter will develop by the duplication of any arbitrary sound concept still external to it.

The question as to how sound-concepts are formed in different cultural traditions merits extreme attention. A thorough discussion is, alas!, out of the scope of this paper. Only a few remarks may be allowed. In accordance with the indo-european musical tradition, the scheme of the interpreted *M*-concept $C'(v)$ presented in par. 4 exhibits (at most) the three controls T', At and *DUR*. This is obtained by allocating constants to the remainder occurring in the operators of the *M*-vectors involved. Here is an example (where $a, b, \ldots, s$ are numbers, but not d, e and q):

$$C'(2) = (\lambda d_1, d_3, d_4, T', At, DUR\ . \qquad (5.8)^*$$
$$(\lambda T, FS, \ldots (\ldots, T, FS, d_1, At, T', \ldots))(\ldots, b, a),$$
$$(\lambda T, P, \ldots (\ldots, T, P, \Phi_5(d_1, d_3, d_4, DUR), T', \ldots))(\ldots, f, c),$$
$$(\lambda q, N, \ldots (\ldots, q, N, d_3, At, T', \ldots))(\ldots, h, g)$$
$$(\lambda T, M, N, \ldots (\ldots, T, M, N, d_4, At, T', \ldots))(\ldots, n, m, l))(s, r, p)$$
$$\longleftrightarrow \lambda T', At, DUR\ .\ (\ldots, a, b, p, At, T', \ldots),$$
$$(\ldots, c, f, \Phi_5(p, r, s, DUR), T', \ldots),$$
$$(\ldots, g, h, r, At, T', \ldots),$$
$$(\ldots, l, m, n, s, At, T', \ldots)$$

However, in the case of Tibetan formant-singing we would probably rather form a concept along the line:**

$$(\lambda T', DUR, At, \ldots, T, \mathrm{FS}\ .\ (\ldots, T', DUR, At, \ldots))(\ldots, c, b, a) \qquad (5.9)^*$$
$$\longleftrightarrow \lambda T, FS\ .\ (\ldots, a, b, c, \ldots)$$

And aiming at speech synthesis the following scheme would make sense for singing with text:

$$\vec{m}_{\mathrm{ling}} = (\lambda T, FS, \ldots, T', P, At, DUR\ . \qquad (5.10)^*$$
$$(\ldots, T, \mathrm{FS}, \ldots, T', P, At, DUR \ldots))(\ldots, b, a)$$
$$\longleftrightarrow \lambda T', P, At, DUR\ .\ (\ldots, a, b, \ldots, T', P, At, DUR \ldots)$$

whereas for spoken language the following substitutions would be suitable:

$$(\lambda LI_1 LI_2 LI_3 LI_4) \ldots \vec{m}_{\mathrm{ling}} \qquad (5.11)$$

The functions LI_1 through LI_4 might then stand for linguistic functions providing the corresponding stress- and intonation-patterns. For completeness it should be mentioned that I did many tests with a MIDIL system for linguistic application. The results obtained can be found among others in my composition "Vers d'autres jeux" from CONSOLATIONS (1981/82). Cf. p. 182.

MIDIM stands, however, for "Minimum description of Music" and for this reason we shall in what follows apply the *M*-concept $C'(v)$ in the way defined in par. 4.

5.2 Pairs of synchronized *M*-vectors

In order to represent the temporal comportment of sound we took the step from the single *M*-vector to the *M*-concept embracing a sequence of 4 segments viz. *M*-vectors. When discussing the VOSIM we mentioned on p. 77 that vowels can be duplicated by means of two synchronized *V*-generators.

---

* This presentation has been chosen for reader's convenience, although it is not valid in our language before the concepts will be assembled in the predicator. Cf. par. 4.

** The same would apply for the click-languages of South-African Bushmen.

In the same way M-representations of *vowels* may be formed by pairs of M-vectors assembled in a M-concept exhibiting two synchronized sequences of four segments each. Let us first have a look at the vectors. We shall once more refer to some descriptive properties being supposed to characterize spoken vowels. Among them we mention: emission time, pitch, fixed formant, formant-shifting, portamento (intonation), amplitude-envelope contour, attenuation. These properties are realized by the vector $\vec{m}_{850}$ (cf. par. 3.9). We assume thus that the subset denoted by this vector will contain the extensions of the M-vectors which, assembled in pairs, represent vowels, preferably even the totality of them. We shall give some examples which may contribute to corroborate the above assumption.
First we tie by constants the controls S, Ef, Of (for obvious reasons) and C as a result of tests:

$$\vec{m}_{851} = (\lambda S, Ef, Of, C, FS, W, Sp, P, e, y_2, \Delta Am, Am, N\max, T, T', At, d \cdot ()_{850})(0.75, 0, 1, 1) \tag{5.12}$$

Restricting our representation of vowels to the minimum amount of distinctional features (and the prosodic controls) we may now allocate constants to all variables listed in the λ-operator except $Am, N\max, T$ (and the prosodic ones) as follows:

$$\vec{m}_{852} = (\vec{m}_{851})(0, 1, 1, 0, 0, 0, 0) \tag{5.13}$$

We now form pairs of synchronized M-vectors (marking all the information shared in common only once) and obtain:

$$(\vec{m}_{852})\begin{matrix}(c_T, c_{N\max}, c_{Am})\\(c_T, c_{N\max}, c_{Am})\end{matrix} = (\begin{matrix}\lambda Am, N\max, T\cdot\\ \lambda Am, N\max, T\cdot\end{matrix}(\lambda T', At, d \cdot ()_{852}))\begin{matrix}(c_T, c_{N\max}, c_{Am})\\(c_T, c_{N\max}, c_{Am})\end{matrix} \tag{5.14}$$

c_T	$c_{N\max}$	c_{Am}	
960	4	511	for /*a*/
394	10	50	
2560	2	511	for /*o*/
960	5	50	
4320	1	511	for /*u*/
1320	3	35	
2400	2	511	for /*e*/
394	10	50	
4480	1	511	for /*i*/
240	18	35	

$$T' \in ([N\min \cdot T + Ef, N\max \cdot T + M\max]_1 \cup [N\min \cdot T + Ef, N\max \cdot T + M\max]_2)$$

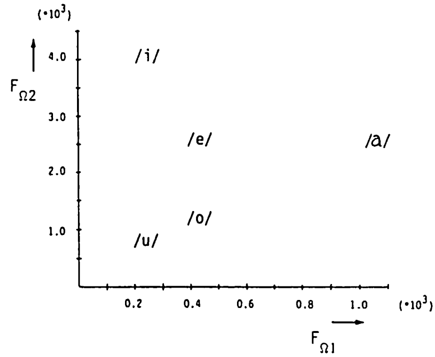


Figure 5.1. The vowel-concepts given on page 123 are represented in the space F_{Ω_2} versus F_{Ω_1}.

By means of these pairs of vectors we may proceed to the formation of the corresponding (rather schematic) vowel-concept. If, on the contrary, we wish to convey to our vowels a maximum of flexibility viz. controllability then we simply go back to $\vec{m}_{851}$ and represent it as follows:

$$\vec{m}_{851} = (\lambda S, Ef, Of, C, N\text{max}, T, FS, W, Sp, e, y_2, d, P, Am, \Delta Am, T', At\ .\ ()_{851})(0.75, 0, 1, 1) \tag{5.15}$$

and we form:

$$\vec{m}_{853} = (\vec{m}_{851})(c_T, c_{N\text{max}}) \tag{5.16}$$

For completeness we add yet a standard vector for spoken vowels:

$$\vec{m}_{854} = (\vec{m}_{851})(c_{\Delta Am}, c_{Am}, c_P, c_d, 1, 1, 3, 50, 0, c_T, c_{N\text{max}}) \tag{5.17}$$

and the corresponding constants per segment for the formation of concepts (for c_T, $c_{N\text{max}}$ and c_{Am} see the previous table):

segment i	$c_{\Delta Am}$	c_{Am}	c_P	c_d
1	0	511	—200	200
	0	c_{Am}	0	200
2	0	511	0	Φ_5
	0	c_{Am}	0	Φ_5
3	—511	511	0	10
	$-c_{Am}$	c_{Am}	0	10
4	0	0	0	100
	0	0	0	100

In the above vector $\vec{m}_{854}$ it holds: $e = y_2 = 1$ which entails: $\Phi'_{15} = 1$ (cf. 2.3.3.3).

If now we assume e.g. $y_2 = 2$, $e = 8$ for the segments of the "second track", then the amplitude ratios Am_1/Am_2 and $\Delta Am_1/\Delta Am_2$ will depend upon At ($Am_1/Am_2 = 511/50$ for $At = 0.4$ but $511/100$ for $At = 1$). The spectrum will thus be a function of At which is known to apply for speech as well as instrumental sounds.

Since the concepts may assemble different sorts of vectors we may form diphtongs, transitory sounds (e.g. */u→a/*) as well as any compounds of fricatives and vowels (e.g. */fa/*, */suo/*) up to degree 4.* The two tracks available render possible the formation of any voiced fricatives and plosives (*b*, *d*, *g*). The close relationship existing between the signs of spoken language and musical concepts will favour the setting-up of *M*-concepts which can be associated with any sort of descriptive musical sound-concepts, and eventually the composer may let soar his creative imagination seeking for his own universe of sounds.

5.3 The standard predicator

The formation of concepts embracing pairs (*n*-uples) of synchronized sequences of vectors needs, however, more precision. Since two (*n*) generators are now controlled they will be indexed. The concepts embracing *n*-uples of synchronized vectors are then obtained by forming *n*-uples of interpreted concepts and by applying to them the so-called "link"-functions. Here is the function table 3, which is an open list of functions applying to pairs of concepts.

Function table 3 (cf. Function table 2 on p. 95)**

coeff.	parameter	function	
$(\gamma_1)_2$	DUR_2	$L_1 = \lambda DUR_1 \,.\, DUR_1$	
$(\gamma_2)_2$	At_2	$L_2 = \lambda At_1 \,.\, At_1$	
$(\gamma_3)_2$	T'_2	$L_{31} = \lambda T'_1 \,.\, T'_1$	
$(\gamma_3)_2$	T'_2	$L_{32} = \lambda y, T'_1 \,.\, y \cdot T'_1$	
$(\gamma_3)_2$	T'_2	$L_{33} = \lambda \mathrm{bea}, T'_1 \,.\, 10^6/\lvert 10^6/T'_1 - \mathrm{bea}\rvert$	(bea (Hz) is the
$(\gamma_5)_2$	$(d_i)_2$	$L_5 = \lambda (d_i)_1 \,.\, (d_i)_1$	beat-freq.)

Synchronization is obtained by:

$$(\lambda L_5 L_1)(d_i)_1, DUR_1 \,.\, C_2'(v) \quad \text{cf. par. 1.7} \tag{5.18}$$

* By the application of compound predicators degrees higher than four may be obtained. Cf. par. 10.

** The M8X-system offers among other facilities the so-called "stereo"-option. It allows to define the position of the sound between two sound sources by a function either of T' and Of or of T. We refrain from discussing this item.

and common T'- and At-control is realized by:

$$(\lambda L_2)At_1 \,.\, C_2'(v) \qquad (5.19)$$

$$(\lambda L_{3j})\ldots T_1' \,.\, C_2'(v) \qquad (5.20)$$

We now define:

The *standard predicator* $P'(c_v)$ is a pair of synchronized interpreted concepts as follows:

$$P'(c_v) = [(\lambda L_1 L_2 L_{3j} L_5)\ldots d_i, DUR, T', At \,.\, (C_1'(v), C_2'(v))](c_{d_i}, \ldots, c_v) \qquad (5.21)$$

where d_i, DUR, T', At stand for $(d_i, DUR, T', At)_1$. For convenience the members of the pair of concepts will be called *track 1* and *2*. (Note that the M8X-system recognizes also single track predicators, that is interpreted concepts.)

Let us still have a glance at the application of the functions L_{31} through L_{33}. The substitution into $(\gamma_3)_2$ of

L_{31} needs no comment.

L_{32} with $y \in \mathbb{Q}$ will result again in harmonic spectra. For $y < 1$ and $T_2 = \text{func}(T')$ the relevant frequency domain of the spectrum will be enlarged from $2F_{\Omega 2}$ to $y^{-1} \cdot 2F_{\Omega 2}$
with $y \in (\mathbb{R} \setminus \mathbb{Q})$ the function L_{32} provides *non-harmonic* spectra, which may still be accentuated by the application of subharmonics.

L_{33} allocates a constant beat rate to the predicator (optimal for $Am_2 = Am_1$ and $\Delta Am_2 = \Delta Am_2$). An interpretation: metallophones.

Here are some standard predicators (where the labels given in the topline indicate descriptive properties which may be associated with the predicators.*) For simplicity we shall write $P(c)$ for $P'(c)$.

A b-flat clarinet (cf. 3.6 and 3.7)

$$P(2)_{\text{Cla}} = [(\lambda L_1 L_2 L_{31} L_5)v, d_1, d_3, d_4, DUR, T', At \cdot (\text{Cla}_1, \text{Cla}_2)] \qquad (5.22)$$

$$\frac{(c_{d_4},\ c_{d_3}, c_{d_1}, c_v)}{100, 30, 40, 2}$$

$\text{Cla}_1 = (C'(2))((\vec{m}_{752})(c_{Am},\ c_{\Delta Am},\ c_C,\ c_{Sp},\ c_W))_i$

c_{Am}	$c_{\Delta Am}$	c_C	c_{Sp}	c_W	i
0	300	0.9	4.5	0	1
300	0	0.9	4.5	7	2
300	—300	0.9	4.5	0	3
0	0	0.9	4.5	0	4

* The labels could just as well have been given in the bottom line as Debussy did in the preludes for piano.

$\text{Cla}_1 = (C'(2))((\vec{m}_{652})(c_{Am}, c_{\Delta Am}, c_C, c_{Sp}, c_W))_i$

c_{Am}	$c_{\Delta Am}$	c_C	c_{Sp}	c_W	i
0	10	1.	4.5	0	1
10	0	1.	4.5	7	2
10	—10	1.	4.5	0	3
0	0	1.	4.5	0	4

$$\vec{m}_{651} = (\lambda F_6 F_5) y_2, e, Am, \Delta Am, At \,.\, \vec{m}_{65}$$
$$= (\lambda F_6 F_5 F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0$$
$$= \lambda N, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \,.\, ()_{651}$$

$$\vec{m}_{652} = (\vec{m}_{651})(1, 1, 1, 0, 0, 1001/1003, 3)$$

$$\vec{m}_{751} = (\lambda F_8) Of, T\text{max}, q, T' \,.\, \vec{m}_{751}$$
$$= (\lambda F_8 F_6 F_5 F_{11} F_9 \Phi_6 F_2 F_1 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0$$
$$= \lambda T\text{max}, q, P, Of, e, y_2, S, W, Sp, C, \Delta Am, Am, d, At, T' \,.\, ()_{751}$$

$$\vec{m}_{752} = (\vec{m}_{751})(1, 1, 1, 0, 0, \tfrac{1}{2}, 460)$$

The vibrato may be varied in any suitable way. While the above data stand for the duplication of a Boehm clarinet (played in the french style) the German Oehler clarinet may be obtained by $C = 1$ with no vibrato. Decreasing the C-rate simulates in general a lessening of the reed pressure.

A bassoon (cf. par. 3.8 and 8.2)

$$P(2)_{\text{Bon}} = [(\lambda L_1 L_2 L_{31} L_5) v, d_1, d_3, d_4, DUR, T', At \cdot (\text{Bon}_1, \text{Bon}_2)] \qquad (5.23)$$

(c_{d_4},	c_{d_2},	c_{d_1},	c_v)
10,	10,	50,	2

$\text{Bon}_1 = (C'(2))((\vec{m}_{855})(c_{\Delta Am}, c_{Am}, c_T, c_{N\text{max}}, c_{Ef}, c_{Sp}, c_W, c_S, c_P))_i$

$c_{\Delta Am}$	c_{Am}	c_T	$c_{N\text{max}}$	c_{Ef}	c_{Sp}	c_W	c_S	c_P	i
500	0	478	4	300	5	5	1	0	1
0	500	478	4	300	5	15	1	0	2
—500	500	478	4	300	5	15	1	0	3
0	0	478	4	300	5	15	1	0	4

$\text{Bon}_2 = (C'(2))((\vec{m}_{855})(c_{\Delta Am}, c_{Am}, c_T, c_{N\text{max}}, c_{Ef}, c_{Sp}, c_W, c_S, c_P))_i$

$c_{\Delta Am}$	c_{Am}	c_T	$c_{N\text{max}}$	c_{Ef}	c_{Sp}	c_W	c_S	c_P	i
50	0	330	6	300	5	5	1	0	1
0	50	330	6	300	5	15	1	0	2
—50	50	330	6	300	5	15	1	0	3
0	0	330	6	300	5	15	1	0	4

$$\vec{m}_{850} = (\lambda F_6 F_5) \ldots \vec{m}_{85} = (\lambda F_6 F_5 F_{11} F_9 G_2 \Phi_6 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0$$
$$= \lambda e, y_2, Of, FS, C, P, S, W, Sp, Ef, N\text{max}, T, Am, \Delta Am, T', At, d \,.\, ()_{850}$$

$$\vec{m}_{855} = (\vec{m}_{850})(0.75, 0, 0, 1, 1)$$

Possible variation: $T_2 \epsilon[330, 357]$. In the signals of natural bassoon sounds we observed frequently a strong instability of the pitch essentially for the attack. For the MIDIM-duplication the intonation pattern discussed in par. 4.2.1 will be suitable. Applying the spreading function will eventually render the vibrato more living. Cf. p. 130.

A harpsichord-like sound (cf. par. 3.8)

$$P(2)_{\text{Cem}} = [(\lambda L_1 L_2 L_{32} L_5\, v, y, At, T, DUR, d_1, d_3, d_4 \,.\, (\text{Cem}_1, \text{Cem}_1)] \tag{5.24}$$

$$\frac{(c_{d_4}, c_{d_3}, c_{d_1}, c_y, c_v)}{10, 20, 5, \tfrac{1}{2}, 2}$$

$\text{Cem}_1 = (C'(2))((\vec{m}_{801})(c_{\Delta Am}, c_{Am}, c_{\Delta T}, c_T))_i$

$c_{\Delta Am}$	c_{Am}	$c_{\Delta T}$	c_T	i
500	0	17	53	1
—200	300	0	70	2
—100	100	0	70	3
0	0	0	70	4

$\vec{m}_{800} = (\lambda F_6 F_5) y_2, e, \Delta Am, Am, At \,.\, \vec{m}_8 = (\lambda F_6 F_5 G_8 F_4 F_3 F_{12}) \ldots, d \,.\, \vec{m}_0$
$= \lambda D, Of, \Delta T', Ef, y_2, e, N\text{max}, T, \Delta T, Am, \Delta Am, At, T', d \,.\, ()_{800}$

$\vec{m}_{801} = (\vec{m}_{800})(49, 1, 1, 10, 0, 0, 0)$

I applied this predicator in my composition "Consolations", in the opening sequence of the first part "In Memoriam".

A low gong-like sound (cf. par. 3.6)

$$P(3)_{\text{Bel}} = [(\lambda L_1 L_2 L_{32} L_5) v, y, At, T', DUR, d_1, d_2, d_4 \,.\, (\text{Bel}_1, \text{Bel}_2)] \tag{5.25}$$

$$\frac{(c_{d_4}, c_{d_2}, c_{d_1}, c_y, c_v)}{10, 120, 10, 111/117, 3}$$

$\text{Bel}_1 = (C'(3))((\vec{m}_{642})\ (c_{\Delta Am}, c_{Am}, c_{MF}, c_S, c_N, c_y))_i$

$c_{\Delta Am}$	c_{Am}	c_{MF}	c_S	c_N	c_y	i
500	0	/	0	2	.9	1
150	100	/	0	2	.9	2
—250	250	9	1	2	.9	3
0	0	/	0	2	.9	4

(/ for dummy)

$\text{Bel}_2 = (C'(3))((\vec{m}_{643})\ (c_{\Delta Am}, c_{Am}, c_{MF}, c_S, c_N, c_y))_i$

$c_{\Delta Am}$	c_{Am}	c_{MF}	c_S	c_N	c_y	i
300	0	/	0	3	.5	1
—220	300	/	0	3	.5	2
—80	80	25	1	3	.5	3
0	0	/	0	3	.5	4

$$\vec{m}_{641} = (\lambda F_6 F_5)e, y_2, At \,.\, \vec{m}_{64} = (\lambda F_6 F_5 F_1\, G_9\, \Phi_6 F_2 F_4 F_3 F_{12})\ldots, d \,.\, \vec{m}_0$$
$$= \lambda q, C, P, Of, Ef, e, y_2, y, N, S, MF, Am, \Delta Am, At, T', d \,.\, ()_{641}$$

$$\vec{m}_{642} = (\vec{m}_{641})(1, 1, 10, 0, 0, 1, 24/37)$$

$$\vec{m}_{643} = (\vec{m}_{641})(1, 1, 10, 0, 0, 1, 3/5)$$

Varying $N_1 \in [2, 11]$ and $N_2 \in [3, 17]$ results in many sorts of metallic sounds. Predicators of this type were used in my compositions "Chants magnétiques" (1985) and "Ritournelles I" for soprano and computer (1984/85).

5.4 Substituting functions into the controls

In the standard predicator constants are allocated to all λ-tied variables except *DUR*, *At* and *T′* occurring in the outermost λ-operator.

(1) It goes without saying that these variables will be *ignored* by any vector where they are already tied by a constant. E.g. $\lambda T'$. will act upon the "vowel"-vector $(\vec{m}_{852})$ $(c_T, c_{N\max}, c_{Am})$ but not upon the "fricative"-vector $(\vec{m}_{641})$ $(c_{T'})$; λDUR . will be ignored by any vector occurring in a segment S_i, $i \neq v$ and accordingly by all vectors occurring in a predicator $P(0)$ etc. In a general way the following schemes may be set up for the controls operating in the predicator $P(v)$:

$$\lambda DUR, At, T' \,.\, (\quad)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots T'\ldots)(a)\ldots)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots T'\ldots At\ldots)(b, a)\ldots)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots T'\ldots At\ldots DUR\ldots)(c, b, a)\ldots) = P(0)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots T'\ldots DUR\ldots)(c, a)\ldots) = P(0)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots At\ldots)(b)\ldots))$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots At\ldots DUR\ldots)(c, b)\ldots) = P(0)$$
$$\lambda DUR, At, T' \,.\, (\ldots(\ldots DUR\ldots)(c)\ldots) = P(0)$$

(2) *Functions* may be substituted into the variables *DUR*, *At* and *T′*.

(2a) The M8X-predicator substitutes the *prosodic functions* Φ_1 through Φ_3 into the prosodic variables *T′*, *At* and *DUR*. These functions will bridge the gap separating the physical and musical representation of the prosodic controls by taking their arguments close to musical tradition as follows (cf. par. 2.3 functable 2):

$$(\lambda \bar{\Phi}_1\, \Phi_2\, \Phi_3) met, du, at, sub, pi, oc \,.\, P(v)$$

where

$DUR(\text{ms}) = \Phi_1(met, du)$

du for relative duration (1., .5, .25 for ♩, ♪, 𝅘𝅥𝅯 etc.)

met for metronome (♩$(s) = 60s/met$)

$T'(\text{us}) \quad = \Phi_3(sub, pi, oc)$

pi	for pitch-index (0, 1, 2, ..., sub, ...)
oc	for octave-index (0, 1, 2, ...)
sub	for subdivision ($sub = 12$ entails $pi = 0, 1, 2, \ldots$ for $c, c\#, d, \ldots$)
At	$= \Phi_2(at)$
at	for relative attenuation (from 0 raising for increasing att.)

Accordingly the M8X-predicator reads as follows:+

$$P8X(c_v) = (\lambda \Phi_3 \Phi_2 \Phi_1) sub, met, du, pi, oc, at \;.\; P(c_v)$$

(2b) We mention still the *spreading function*. It may be introduced in order to supply a greater degree of suppleness to the predicator as follows:

$$Spr(\gamma, z) = \lambda\gamma, z \;.\; alea(\gamma - \gamma z, \gamma + \gamma z) \quad 0 \leq z \leq 1 \tag{5.26}$$

The value is selected randomly from the domain indicated and ranges from 0 through 2γ. The product $\bar{F}_j(z, \gamma_h) = (Spr(F_j(\gamma_h), z)$ may be substituted into x_j as follows:

$$(\lambda \bar{F}_j) z, \gamma_h \;.\; \vec{m} \quad \text{for} \quad (\lambda SprF)_j z, \gamma, \gamma_h \;.\; \vec{m}) \tag{5.27}$$

Accordingly $\bar{F}_j$ designates a function F_j provided with spreading. The M8X-predicator recognizes spreading applied to the parameters *DUR*, *D*, *MF*. (Standard: $z_{DUR} = .1$, $z_D = .8$, $z_{MF} = .1$).

5.5 Conclusion

The set of all standard predicators represents a powerful classification and interpretation of the sound-output corresponding to V^*. This power could, of course, still be considerably increased by introducing other types of predicators (1) applying controls different from the present ones (cf. p. 122), (2) recognizing more than four segments as well as (3) controlling more than two sounds tracks. However, the standard predicator presented in this paper has been adopted along the line of the following assumptions:

a) In any culture there exists a close physical and conceptual relationship between the musical and the linguistical expressions, such that

b) if a sound generating system is powerful enough to embrace speech synthesis it can be extended into a system which is powerful enough for musical application as well,

---

+ In an analogous way we could proceed to any sort of predicator; e.g. by substituting into *DUR* of the standard predicator a function of the number of frames and of the projection-speed in making film music, or by replacing *At* by a function of the luminosity-control of a TV-screen etc. There is even no (theoretical) limitation for iterated substitution of any sort.

c) given the strong conceptual nature of music and spoken language (both systems operating upon sign-repertoires) any synthesis applied to them will not intend the duplication of signals but of signs,

d) the minimum amount of information required for the production of any sound event will render exactly possible the decision as to whether or not the sound event belongs to a stated concept viz. sign. (MIDIM stands for MInimum DescrIption of Music.) Since two-formant-structures are sufficient for the construction of linguistic sign-repertoires (within the field of application chosen) the predicator merely controls two synchronized systems. (The more or less constant formants 3 and 4 may easily be added and controlled by suitable functions, increasing the bandwidth and contributing to a better fusion of the two sound outputs summed up.)

6. THE ARTICULATED PREDICATORS OR WORDS

Let there be the empty vector

$$\vec{e} = (0, 0, \ldots, 0) \in \{\vec{e}\} = E, V \cap E = \varnothing$$

and the ordered set of functions MAS_{art} with $art = 1, 2, 3, \ldots$ mapping any arbitrary predicator $P(v) \in V$ into the set $\{V \cup E\}^*$. We first assume $s = 4$, $v = 0$ and define the mapping rule by the following table:

MAS_{art}: art	MAS_{art} $(P(0) \in V^*) \in \{V \cup E\}^*$
1	$S\ S\ S\ S$
2	$S\ S\ S\ \vec{e}$
3	$S\ S\ \vec{e}\ \vec{e}$
4	$\vec{e}\ S\ S\ S$
5	$\vec{e}\ \vec{e}\ S\ S$
6	$\vec{e}\ S\ \vec{e}\ \vec{e}$
7	$\vec{e}\ \vec{e}\ \vec{e}\ S$
8	$S\ \vec{e}\ \vec{e}\ \vec{e}$
9	$\vec{e}\ \vec{e}\ S\ \vec{e}$
10	$\vec{e}\ S\ S\ \vec{e}$
11	$S\ \vec{e}\ \vec{e}\ S$
12	$\vec{e}\ S\ \vec{e}\ S$
13	$S\ \vec{e}\ S\ \vec{e}$
14	$S\ \vec{e}\ S\ S$
15	$S\ S\ \vec{e}\ S$

The mapping operates upon the segments of the two (or *n*) tracks of the predicator in terms of a *mask*. Let us now introduce the variable $v \in \{0, 2, 3, 4\}$. From par. 1.4 we know the definition for $C(v)$; here is a reminder:

v	$C(v)$
0	*S S S S*
2	*S A S S*
3	*S S A S*
4	*S S S A*

The generalized mapping rule will be subject to the following restriction: *Replacing an occurrence of the segment A by* $\vec{e}$ *is not valid* (if not mentioned

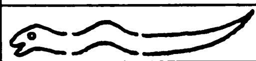
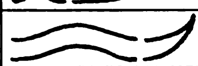
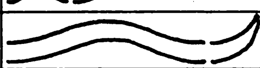
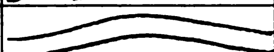
art	*v*	$MAS_{art}(P(v)_p)$	s_1	s_2	s_3	s_4
1	2		d_1	Φ_5	d_3	d_4
1	3		d_1	d_2	Φ_5	d_4
1	4		d_1	d_2	d_3	Φ_5
2	2		d_1	Φ_5	d_3	0
2	3		d_1	d_2	Φ_5	0
3	2		d_1	Φ_5	0	0
4	2		0	Φ_5	d_3	d_4
4	3		0	d_2	Φ_5	d_4
4	4		0	d_2	d_3	Φ_5
5	2		0	Φ_5	d_3	0
5	3		0	d_2	Φ_5	0
6	2		0	Φ_5	0	0
7	4	....	0	0	0	Φ_5
		← *DUR* →	$\Phi_5 = \lambda v, d_i, DUR \cdot DUR - \sum_{i=1, i \neq v}^{5} d_i$			

Figure 6.1. Visualization of the articulated predicators for *art* $\in [1, 7]$.

otherwise). We bear in mind that in any A the time duration d is replaced by the function Φ_5 of DUR (cf. par. 1.6.2); accordingly replacing A by $\vec{e}$ will anihilate the DUR-control.+ And now we define:

> *A M-word* is any articulated predicator $MAS_{c_{art}}(P(c_v))$

The extension of the function MAS_{art} may then be represented by a list of M-words. We present it to the reader by means of a visualization (Fig. 6.1., $v=0$ omitted). For simplicity we limit the presentation to $art \in [1, 7]$. Notice that the variable d occurring in an empty vector $\vec{e}$ will take the value zero since it holds: $N_P = 0$ by definition.

7. SEQUENCES OF WORDS AND FORMULAE

7.1 Formulae and their interpretations

The totality of all predicators is assembled in the ordered set LIB of indexed predicators $P(c_v)_p$. LIB is called the *predicator- or P-library* and p is called the *predicator-index.*

Any arbitrary articulated predicator or word may now be represented in an abridged way by the pair (p, art) of the corresponding indices p and art:

$$(p, art) \quad \text{for} \quad MAS_{art}(P(c_v)_p) \tag{7.1}$$

In an analogous way *any arbitrary sequence of words* $W_1 \ldots W_w$ may be represented by a sequence of indexed pairs $(p, art)_k$, $1 \leq k \leq w$:

$$(p, art)_1 (p, art)_2 \ldots (p, art)_w \in \{V \cup E\}^* \tag{7.2}$$

where k is called the *word-index.*

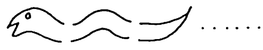
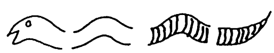
We now define:

> *The set M_g of formulae* is a subset of the set of all possible word-sequences $\in \{V \cup E\}^*$; it is selected by means of a "grammar" g (cf. bottom of this paragraph).

The MIDIM language recognizes different sets of formulae. We shall, however, limit our discussion to a simplified presentation of two of them: the sets M_{g_1} and M_{g_2}. In order to give the reader an initial idea what the formulae of these two sets will be, we present a visualization of the lists of atoms for both sets (Fig. 7.1.). The corresponding formulae could then be recursively defined in the conventional manner by stating:

+ In musical tradition every word (represented by the "notes" of conventional notation) has to be DUR-controllable (except the grace notes). The same holds for the syllables of speech (though DUR will often be replaced by a function; one may e.g. think of metrics in poetry).

Every atom is a formula,
if Z_1, Z_2 are formulae then $Z_1 Z_2$ is a formula.

ATOMS	VISUALIZATION
$(p,1)$	
$(p,2)$	
$(p,7)$	......
$(p_1,3)(p_2,4)$	
$(p_1,3)(p_2,5)$	
$(p_1,3)(p_2,6)$ ... $(p_3,4)$	
$(p_1,3)(p_2,6)$ ... $(p_3,5)$	

$p_1, p_2, p_3, \ldots$ not necessarily different for M_g

$p_1 = p_2 = p_3, \ldots$ for $M_{g'}$

Figure 7.1. List and visualization of the atoms for the sets of formulae (languages) M_{g1} and M_{g2}.

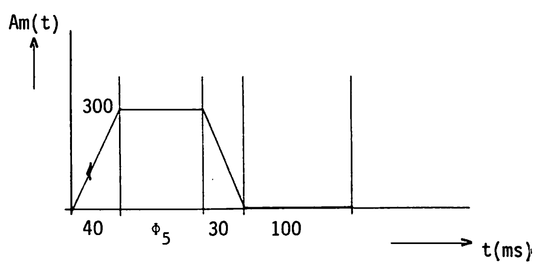
Here is an example:

Let there be a *P*-library $LIB_1 = \{P_1, P_2, \ldots, P_n\}$, which contains (among others) the above "clarinet"- and "bassoon"-predicators ordered as follows:

$$P_1 = P(2)_{\text{cla}}$$

$$P_2 = P(2)_{\text{bon}}$$

The relative amplitude envelope stated by P_1 over its four segments is thus:

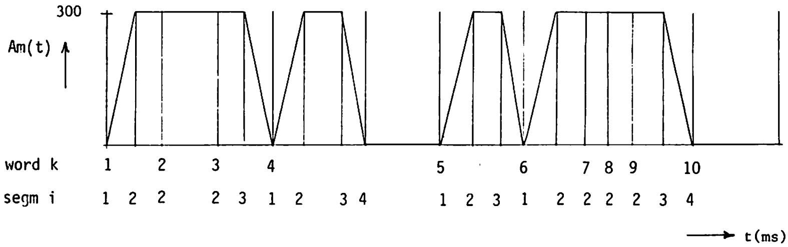


Let there still be the two formulae as follows ($p, q \in \{1, 2, \ldots, n\}$):

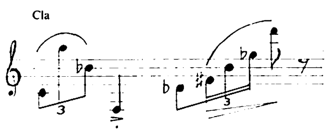
$$Z_1 = (p,3)(p,6)(p,5)(p,1)(p,2)(p,3)(p,6)(p,6)(p,5)(p,7) \in M_{g_1}$$

$$Z_2 = (p,3)(p,5)(q,2)(q,3)(q,4)(p,3)(p,6)(p,5)(q,2)(p,7) \in M_{g_2}$$

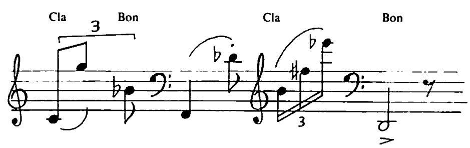
The relative amplitude envelope stated by Z_1 with $p = 1$ is thus:



A possible instance of Z_1 is for $p = 1$ e.g.:



And a possible instance of Z_2 is for $p = 1$, $q = 2$ e.g.:



In an analogous manner we could discuss the melodic, dynamic, temporal and timbral patterns etc.

However, for the user's convenience, *the decision as to whether or not any arbitrary sequence of words is a formula,* is not sufficient. *A rewriting algorithm* is needed *which maps any arbitrary sequence of words into a formula.* The MIDIM-language provides an ordered set $G = \{g_1, g_2, \ldots\}$ of mapping or *"grammars"* which will accomplish the tasks of deciding as well as of rewriting. We shall present here the grammar g_1 which maps any arbitrary sequence of words into the set M_{g_1} of formulae.

7.2 The mapping g_1

Let there be any arbitrary sequence of words:

$$(p, art)_1 (p, art)_2 \ldots (p, art)_w \in \{V \cup E\}^* \quad (7.3)$$

where p for predicator-index
art for articulation index
k for word-index
w for number of words occurring.

g_1 is a recursive function of art, k as follows; for any p_k it holds:

$$g_1 : (art_k, art_{k+1}) \rightarrow f_2(f_1((art_k, art_{k+1}))) \text{ and} \quad (7.4)$$

$$(\lambda x\, .\, (art_0 = art_{w+1} = x)(7) \leftrightarrow art_0 = art_{w+1} = 7$$

(two dummy words are thus introduced for computational means, the words indexed by 0 and $w+1$)

where the functions

$$f_1 : (art_k, art_{k+1}) \rightarrow f_1((art_k, art_{k+1})) = Cof_k$$

$$f_2 : f_1((art_k, art_{k+1})) \rightarrow f_2(f_1((art_k, art_{k+1}))) = f_2(Cof_k)$$

are defined as follows:

$\rightarrow art_k$

f_1	1	2	3	4	5	6	7
1	1	1	-2	1	1	-5	1
2	2	2	-2	2	2	-5	2
3	3	3	-2	3	3	-5	3
4	1	1	4	1	1	4	1
5	2	2	5	2	2	5	2
6	3	3	6	3	3	6	3
↓ 7	7	7	-2	7	7	-5	7
art_{k+1}							

Cof_k is the intersection of line art_k and column art_{k+1} of the table on the left. (7.5)

$$f_2(Cof_k) = \begin{cases} (|Cof_k|, art_{k+1}) \text{ if } Cof_k \leq 0 \\ (art_k, Cof_k) \text{ otherwise} \end{cases} \quad (7.6)$$

Since $g_1 = f_2 f_1$ we obtain:

$$g_1((p, art)_1 (p, art)_2 \ldots (p, art)_w) \in M_{g_1} \subset \{V \cup E\}^* \quad (7.7)$$

Accordingly g_1 maps any arbitrary sequence of articulated predicators into a formula $\in M_{g_1}$.

Here is an example of computation:

Let there be the arbitrary sequence of articulated predicators

$$(p, 4)(p, 3)(p, 6)\ (p, 7)(p, 5)(p, 4)(p, 3)(p, 6)$$

The mapping g_1 is now stepwise as follows; for any p_k:

									→	k		
g_l:	0	1	2	3	4	5	6	7	8	$w+1$		
art_k		4	3	6	7	5	4	3	6			
	7	4	3	6	7	5	4	3	6	7		
	7	4									f_1:	$(7,4) \to 1$
	7	1									f_2:	$1 \to (7,1)$
		1	3								f_1:	$(1,3) \to 3$
		1	3								f_2:	$3 \to (1,3)$
			3	6							etc.	$(3,6) \to 6$
			3	6								$6 \to (3,6)$
				6	7							$(6,7) \to -5$
				5	7							$-5 \to (5,7)$
					7	5						$(7,5) \to 2$
					7	2						$2 \to (7,2)$
						2	4					$(2,4) \to 1$
						2	1					$1 \to (2,1)$
							1	3				$(1,3) \to 3$
							1	3				$3 \to (1,3)$
								3	6			$(3,6) \to 6$
								3	6			$6 \to (3,6)$
									6	7		$(6,7) \to -5$
									5	7		$-5 \to (5,7)$
$f_2(f_1(art_k))$		1	3	5	7	2	1	3	5			

Accordingly for the above example it holds for any p_k:

$$g_1((p, 4)(p, 3)(p, 6)(p, 7)(p, 5)(p, 4)(p, 3)(p, 6))$$

$$= (p, 1)(p, 3)(p, 5)(p, 7)(p, 2)(p, 1)(p, 3)(p, 5) \in M_{g_1}$$

7.3 The Descriptor and the Terminal Formulae

Summarizing a distinction will be made between the *non-terminal* formulae $\in M_g (g \in G = \{g_1, g_2, \ldots\})$ and the terminal formulae $\in V^*$.

(1) A *non-terminal formula* is any formula $\in M_g \subset \{V \cup E\}^*$ where λ-tied variables occur.

(1a) If no constants $c_{DUR}, c_{At}, c_{T'}$ are assigned to the pairs $(p, art)_k$ occurring then the non-terminal formula is called an *abstract*, since it denotes a sequence of words, that is of articulated predicators or sound concepts:

$$(p, art)_1 (p, art)_2 \ldots (p, art)_w \in M_g \tag{7.8}$$

where every predicator p_k, $1 \leq k \leq w$ will define the domains for the corresponding λ-tied variables DUR, At, T'.

(1b) If constants satisfying the stated domains are assigned to *all* pairs $(p, art)_k$ occurring then the non-terminal formula is called a *descriptor*, since it denotes a sequence of individual instances or sound events (where a, b, c, for $c_{T'}, c_{At}, c_{DUR}$ are numbers):

$$((p, art)(c, b, a))_1 ((p, art)(c, b, a))_2 \ldots ((p, art)(c, b, a))_w \in M_g \tag{7.9}$$

(2) A *terminal formula* is any pair (n-uple) of synchronized sequences $\in V^*$ of V-vectors obtained from a descriptor by application of the terminal rule.

The *terminal rule* consists of λ-elimination and computation of the functions occurring. When it is applied to the descriptor then any M-vector occurring in an articulated predicator is replaced by a V-vector, whilst any empty vector $\vec{e}$ occurring disappears since it does not belong to V.

Here is a schematic example:

$$\begin{aligned}
D_1 &= \ldots((3,6)(c, b, a))_k \ldots && \in M_{g_1} \text{ (descriptor)} \\
&= \ldots((MAS_6(P(2)_3))(c, b, a))_k \ldots && \in M_{g_1} \\
&= \ldots((\lambda DUR, T', At \,.\, \frac{\vec{e}\,\vec{m}_f\ \vec{e}\,\vec{e}}{\vec{e}\,\vec{m}_h\ \vec{e}\,\vec{e}})(c, b, a))_k \ldots && \in M_{g_1} \\
&\longleftrightarrow \ldots \frac{\vec{v}_p}{\vec{v}_q} \ldots && \in V^* \text{ (sequence of } V\text{-vectors)} \\
& && p, q \in [1, m], \text{ cf. par. 1.3.1}
\end{aligned}$$

Obviously we leave the intensional representation of the MIDIM-language when proceeding to terminal formulae and step over to the extensional

VOSIM-representation. As a consequence we lose the predicators involved with their networks of M-functions, but we obtain an expression which fits into the input of our sound generating system. Accordingly we will be given the sound events wished for.

7.4 The M8X-Descriptor

We mentioned in par. 5.4 that the predicator of the M8X-system substitutes the functions Φ_1, Φ_2 and Φ_3 into the controls DUR, T' and At. As a consequence every word occurring in a M8X-descriptor will be assigned the 4-uple of constants (c_{du}, c_{pi}, c_{oc}, c_{at}), whereas the pair of constants (c_{met}, c_{sub}) is considered to apply for all words listed and will therefore be stored in the *header* or word $k=0$ preceding the expression. The M8X-descriptor is thus a $(w+1)\times 6$-list. Cf. par. 9.2.

Here is an example: $P_3 = (\lambda\Phi_3\Phi_2\Phi_1)\ldots P(3)_{\mathrm{Bel}}$ (cf. par. 5.3)

k						
0	$sub = 1200$, $met = 60$					
	p	art	du	pi	oc	at
1	3	2	3500	1050	3	0
2	3	1	2100	1150	3	1
3	3	2	4500	950	3	0.5
.						
.						
.						
w						

8. THE DOMAIN-PROTECTION

In our language the constants allocated to the λ-tied variables DUR, At, T' of the predicator are always supposed to satisfy the corresponding E-domains (cf. par. 1.5). This presupposes, of course, *that all predicator-indices occurring in any descriptor are defined* by a corresponding P-library. Obviously any descriptor where such would not be the case would be *meaningless* since it would designate vectors of V as well as vectors beyond V.

Here is an example:

For every predicator where $\Delta T' = 0$ we get according to the formula (1.14) for the E-domain of the λ-tied variable DUR:

$$E_{DUR} = \{\gamma_1 \in D_{\gamma_1} | F_{12}(\Phi_5(\gamma_1)) \in D_{x_{12}}\} = D_{\gamma_1} \cap [\alpha + T', \alpha + T' \cdot 4096] \quad (8.1)$$

$$\text{with } F_{12}(\Phi_5(\gamma_1)) = \mathrm{rof}\left(\frac{DUR - \alpha}{T'}\right) \in [1{,}4096] \quad \alpha = \sum_{\substack{i=1 \\ i \neq v}}^{4} d_i$$

8.1 *E*-Domain Checking and Rewriting

In order to assure domain-protection any predicator will (1) decide whether or not the constants assigned to its λ-tied variables *DUR*, *At*, *T'* by the descriptor will satisfy the corresponding *E*-domains. This task is accomplished by the mapping *CHE* (checking). (2) For users convenience the mapping *CHE* is combined with a rewriting-algorithm *REW* mapping any argument which is not valid into a valid one. We thus obtain for any $\gamma_h \in D_{\gamma_h}$:

$$CHE(\gamma_h) = \begin{cases} \gamma_h \notin E \rightarrow REW(\gamma_h) \in E \\ \gamma_h \text{ otherwise} \end{cases} \tag{8.2}$$

There are two different rules applied to *REW*; for $\gamma_h \in D_{\gamma_h}$ and $E_{\gamma_h} = [\min, \max]$:

$$REW_1(\gamma_h) = \begin{cases} (\gamma_h > \max \rightarrow \gamma_h = \max) \wedge (\gamma_h < \min \rightarrow \gamma_h = \min) \\ \text{otherwise } \gamma_h \end{cases} \tag{8.3}$$

$$REW_2(\gamma_h) = \begin{cases} (\gamma_h > \max \rightarrow \gamma_h/2) \wedge (\gamma_h < \min \rightarrow 2 \cdot \gamma_h) \\ \text{otherwise } \gamma_h \end{cases} \tag{8.4}$$

Interpretation: REW_1 will be applied to *At* and *DUR*. REW_2 is designed for application to *T'* in the case that the interval $E_{T'}$ will be equal to or larger than an octave; otherwise REW_1 is applied to *T'*. These interpretations are in tune with the musical tradition.

The *E*-domains are defined by the *M*-vectors $\vec{m}$ occurring in the predicator; for this reason the arguments will be checked and, if necessary, rewritten per word for any segement which is not an empty vector $\vec{e}$.

8.2 Enlarging the expressive Power of the Predicator

If you ask a soprano to produce the vowel */u/* at a low pitch and then to raise the pitch gradually, the singer will (possibly) follow the instructions but stop at a certain pitch telling you that the upper pitch-limit of the */u/*-sound has been reached; or she will continue by moving the */u/*-sound gradually into an */o/*-, an /ɔ/- and an */a/*-sound the more she raises the pitch. Enlargment of the pitch-domain is thus obtained by *changing* the sound-concept. Whether this change is admissible or not depends exclusively upon conceptual criteria. In any *speech-communication altering an /u/* into an */a/* will be inacceptable since it affects the meaning; but in singing it is easily admitted if not common use since priority is given to the melodic development of pitch within the pitch-domain defined by the singer (soprano, mezzo, counter-alto). Notice that on the contrary the substitution of a */u/*- into an */a/*-sound would be refuted by an accomplished singer (cf. Fig. 8.1., 8.2). Unfortunately a thorough discussion of the quarrels concerning the relationship of music and text is out of the scope of these pages. We still mention, however, an important compromise made for almost every pitch-controllable musical instrument: the *registers*. Each register of an instrument is a concept defining its own *E*-domains; accordingly the

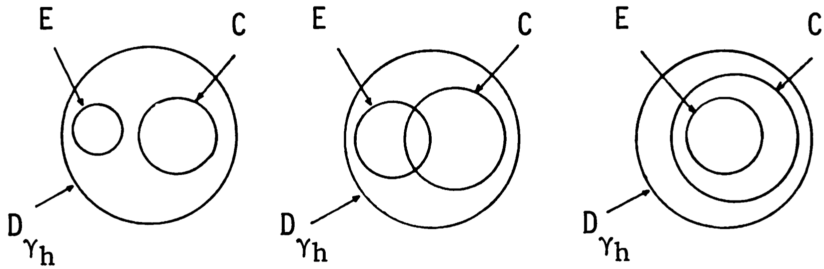
Figure 8.1. Cl. Debussy (ariettes oubliées 1888, Spleen, bars 22-34). The famous text of Verlaine: "...et de tout, fors de vous, hélas!" will sound in Debussy's song like: "et de */ta/*, fors de vous, hélas!"

instrument is an assembly of (compatible) concepts. One may for instance think of the bassoon which alters the characteristic */o/*-sound of the medium register into an */u/*-sound for the low keys and into an */ä/*-sound for the high keys. Cf. Fig. 8.6 as well as par. 3.8 p. 113 and par. 5.3 p. 127.

Figure 8.2. W. A. Mozart (Nozze di Figaro 1785/86, No. 27, Susanna, "Rosen-Aria", bars 11/15). In the text of Da Ponte: "...not-tur-na fa-ce,..."' the */na/* and */fa/* may by no means turn into */no/*, */fo/* or */nu/*, */fu/* in spite of the extremely low pitch.

In order to form in the above spirit an enlarged predicator which will assemble more than one single register we assume there to be

(1) the non-empty domain C with $C \not\subset E$ as follows:



(2) and the mapping:

$$COR: F_j(\gamma_h) \rightarrow F_j'(\gamma_h) \tag{8.5}$$

which changes the function F_j primitively defined by the predicator into the function F_j' such that it holds:

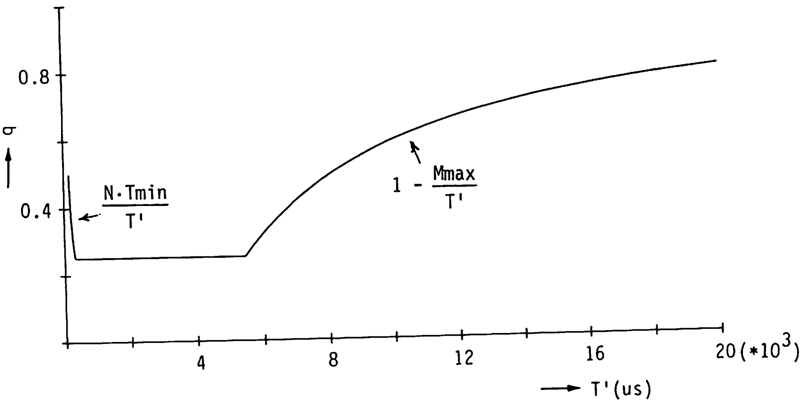
$$\{C \backslash E\}_{\gamma_h} = \{\gamma_h \in D_{\gamma_h} | F_j'(\gamma_h) \in D_{x_j}\} \tag{8.6}$$

The following scheme may give an idea of the domain-protection assured by the enlarged predicator:

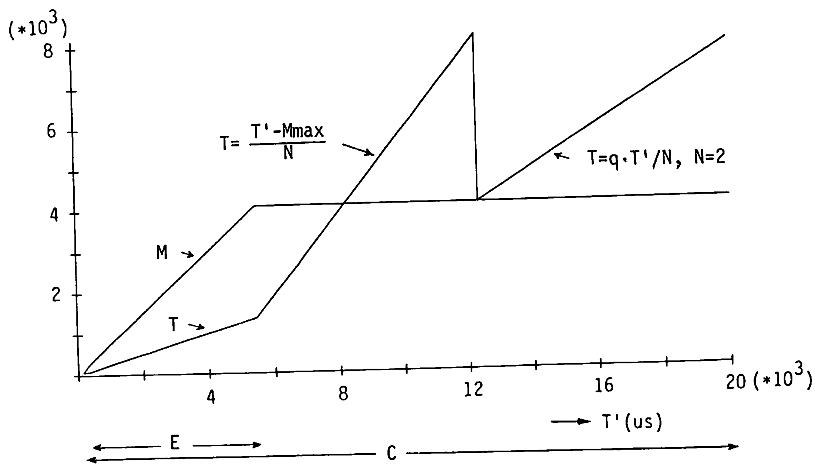
$$CHE(\gamma_h) = \begin{cases} \gamma_h \notin E \begin{pmatrix} \gamma_h \notin C \rightarrow REW(\gamma_h) \in C,\ CHE(REW(\gamma_h)) = \ldots \text{etc.} \\ \gamma_h \in C \rightarrow \gamma_h \in \{C \backslash E\} \rightarrow COR: F_j(\gamma_h) \rightarrow F_j'(\gamma_h) \end{pmatrix} \\ \text{otherwise } \gamma_h \end{cases} \tag{8.7}$$

Every image of the mapping *COR* (correction) will, of course, be incorporated into the (open) function-tables 1 or 2, cf. par. 2.3. In the following pages we submit to the reader some examples. An elaborared theory of MIDIM-substitution with domain-definition is the subject of a special study in preparation.

MISSING HARMONICS, MOVING FORMANT



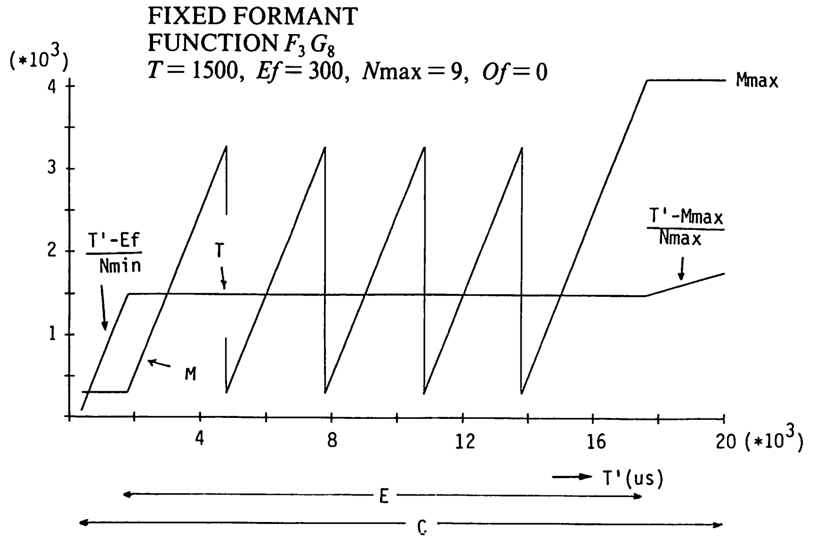
FUNCTION $F_1 F_3$ $q = 0.25$, $N = 1$, $Of = 0$.



$$E = \left[\frac{N \cdot Tmin}{q} - Of, \frac{Mmax}{(1-q)} - Of \right] = [320, 5460]$$

$$C = [40, 20 \cdot 10^3]$$

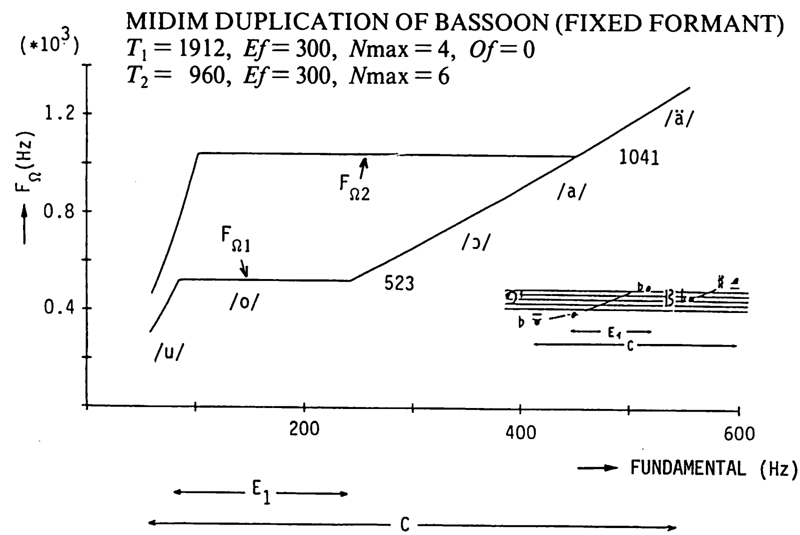
Figure 8.3. and 8.4. Missing harmonics and a moving formant. Beyond the T' domain E the mapping *COR* replaces the primitively defines functions $F_1 F_3$ by new functions for the domain $C \backslash E$. (Beyond C rewriting is applied). Cf. Fig. 3.6.



$E = [N\min \cdot T + Ef - Of, N\max \cdot T + M\max - Of] = [1800, 17595]$

$C = [380, 2 \cdot 10^4]$

Figure 8.5. Fixed formant. Beyond the T' domain E the primitively defined functions $G_8 F_3$ are replaced by new functions. Cf. Fig. 3.10., 3.11.



$E_1 = [85.16, 242.48]$ (Hz)

$C = [58.27, 554.30]$ (Hz)

Figure 8.6. Fixed formant, M-duplication of a bassoon represented by F_{Ω_1} and F_{Ω_2} in the frequency domain. In order to simulate the registers of the bassoon the

boundaries of the domain E are defined by Ef and Nmax for the functions $G_8 F_3$. Beyond the functions shown in Fig. 8.5 are applied. Corresponding vowel sounds are associated with the denoted sound.

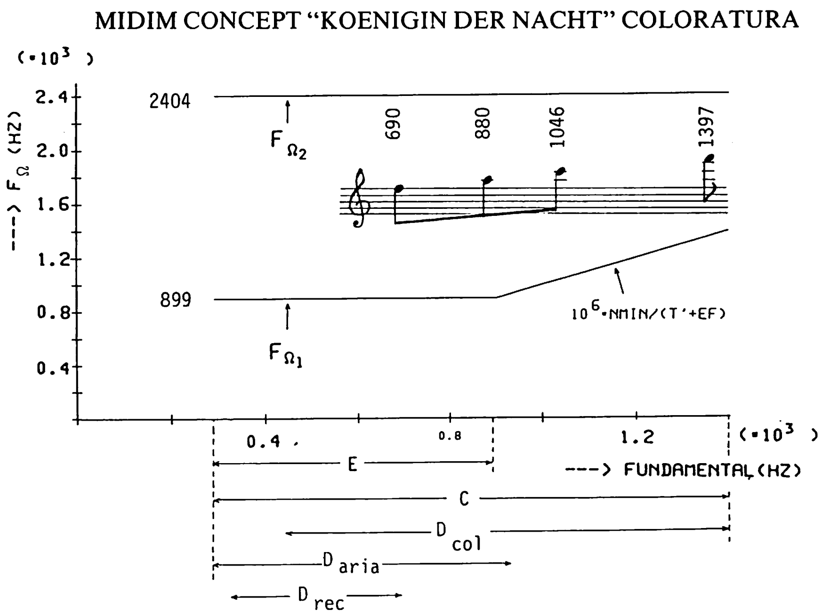


Figure 8.7. Fixed formant. M-duplication of the Königin der Nacht from Mozarts Magic Flute (excerpts of various coluratura sections) represented by F_{Ω_1} and F_{Ω_2} in the frequency domain. The domains E and C are related to the pitch domains applied by Mozart for the recitativo, the aria (with text) and the coloratura. Note the coincidence of the domains E and D_{aria}. The formant-areas within E may be associated with the sound of an open o-vowel.

9. OPERATIONS APPLIED TO THE FORMULAE

9.1 Binary operation

Any arbitrary pair of formulae may be concatenated in order to form a corresponding new formula.

9.2 Unary operations

Any non-terminal formula may be mapped into a corresponding new formula. *Deriving* new formulae from old ones is done by the application of the product of mappings as follows:

$$D: M_g \rightarrow \{V \cup E\}^* \qquad (9.1)$$

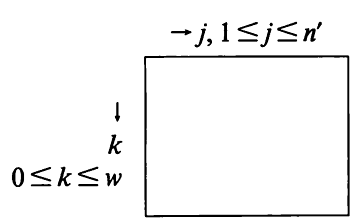
$$g: \{V \cup E\}^* \rightarrow M_g$$

Any non-terminal formula may thus be mapped into a sequence of words which is in turn mapped back into a new formula. Accordingly it holds for any non-terminal formula:

$$g(D(\text{formula})) \in M_g \qquad (9.2)$$

I would like to remind the reader at this point that what follows is in part quite familiar to him despite of the formalized representation. It is true that the *M*-words occurring in the *M*-abstracts as well as the abstracts themselves represent sound concepts and taged sequences of concepts for which it is not easy (though not impossible) to find counterparts in conventional music theory (e.g. phrasing, articulation applying to musical instruments and combinations of them). The *M*-words occurring in the descriptor, however, denote the same thing as the *interpreted* "notes" in a conventional score, namely instances of sound concepts, and the descriptors denote phrases and super-phrases etc. built upon them. Accordingly the derivational rules presented in what follows may be understood in terms of interpreted "composing" rules.

In what follows the *operand* is supposed to be a descriptor. For the representation of the descriptor we shall choose a $(w+1)\times n'$-list:



Accordingly we shall write

$\lambda k, j, x_{kj} \,.\, x_k$ for the above variable matrix (variable descriptor) (9.3)

$$\vec{d}_k = \lambda j, x_{kj} \,.\, x_{kj} \text{ for a variable line-vector} \begin{cases} k=0 & \text{header} \\ 1 \leq k \leq w & \text{word} \end{cases} \qquad (9.4)$$

$$\vec{d}_j = \lambda k, x_{kj} \,.\, x_{kj} \text{ for a variable column-vector} \qquad (9.5)$$

$$\lambda x_{kj} \,.\, x_{kj} \text{ for a variable coefficient with fixed position} \qquad (9.6)$$

The *operator D* may, of course, be assigned many kind of derivational rules, linear, non-linear, conditional ones, alea etc. (However, for each of them there should also be a sensible interpretation.)

A distinction is made between:

(1) *the functions of* k $(1 \leq k \leq w)$:

a) operating upon the *position* of a line-vector viz. word $\vec{d}_k$:

$$D\colon \vec{d}_k \to \vec{d}_{k'} \, (\equiv D\colon \lambda x_{kj}, j \,.\, x_{kj} \to \lambda x_{k'j}, j \,.\, x_{k'j}) \qquad (9.7)$$

b) operating upon the position of the variable coefficient x_{kj}:

$$D'\colon x_{kj} \to x_{k'j} \qquad (9.8)$$

In this latter case the *content* of the word $\vec{d}_k$ will be affected as well.

(2) *the functions of* x_{kj}:

operating upon the content of the variable coefficient x_{kj}:

$$D''\colon x_{kj} \to x'_{kj} \qquad (9.9)$$

affecting thus always the content of the word $\vec{d}_k$.

(3) *the functions of* $\chi(x_{kj})$:

mapping the truth value assigned to the variable coefficient x_{kj} into the content of a selected coefficient $x_{k'j'}$:

$$D'''\colon \chi(x_{kj}) \to x'_{k'j'} \qquad (9.10)$$

affecting thus always the content of the word $\vec{d}_{k'}$.

In the following pages we shall present a selection of D-operations retrieved from a (once more open) D-function table. For more information the reader may refer to Kaegi 1984.*

* For the sake of completeness we will add that any MIDIM-system is a package embracing a set of programs as follows: the predicator- or P-program where predicators are constructed, derived from one another and P-libraries are built up; the descriptor- or D-program where descriptors are formed and derived; the handler or M-program where terminal λ-elimination and computation takes place; and the command constructor- or C-program by means of which it is possible to store complex sound output encoded in command strings which can be read into the handler. The alphabet of the command strings consists of the sets of library-file names, of D-file names and of instructions.

9.2.1 Functions of k

Retrograde:

$D_{ret} = \lambda w, k \,.\, w + 1 - k$ word-index $k = 1, 2, \ldots, w$ (9.11)

number of words w

$(\lambda w, k \,.\, w + 1 - k)(4)$

D_{ret}: k	k'
1	4
2	3
3	2
4	1

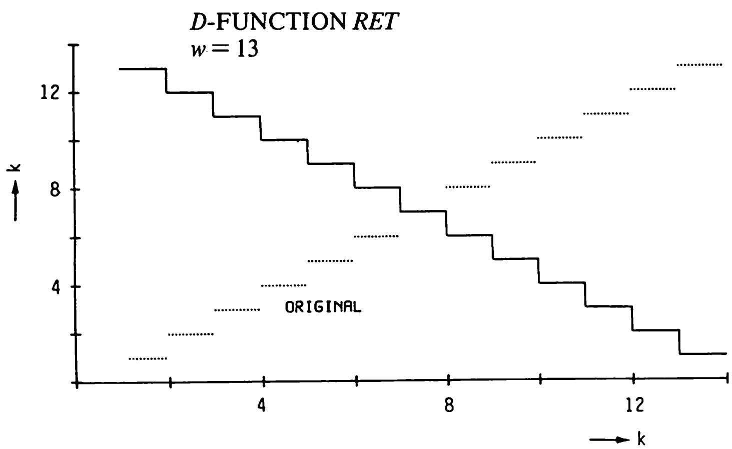


Figure 9.1. *D*-operation "retrograde".

Mirror ($s = 2$) and Mirror + 1 ($s = 1$):

$D_{mir} = \text{formula}MIR(\text{formula})$ (9.12)

$MIR = (\lambda s, w, k \,.\, 2(w + 1) - (k + s))$ (9.13)

$(\lambda s, w, k \,.\, 2(w + 1) - (k + s))(4, 2)$

MIR: k	k'
1	7
2	6
3	5
4	4

$(\lambda s, w, k \,.\, 2(w + 1) - (k + s))(4, 1)$

MIR: k	k'
1	8
2	7
3	6
4	5

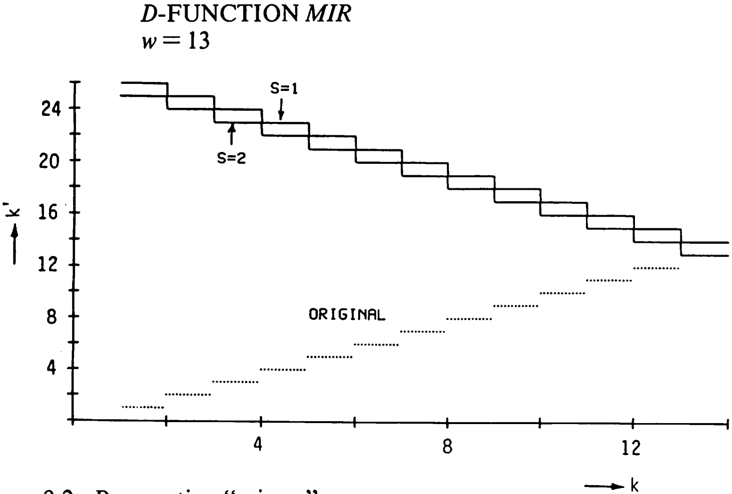


Figure 9.2. D-operation "mirror".

Here are some examples showing how to operate upon the position of words. Let there be the formula $MID \in M_g$; the following are formulae of M_g as well:

$$DIM = g(D_{ret}(MID))$$

$$MIDIM = g(D_{mir}(MID)) \qquad s = 2$$

$$MIDDIM = g(D_{mir}(MID)) \qquad s = 1$$

9.2.2 *Functions of* x_{kj}

Interpolation:
(For easy reading we omit in the formula the column-index j for $\vec{d}_j$. It will be reintroduced below for the interpretation.)

$$D_{ipo} = \lambda ex, w, \alpha, \beta, x_k, k \,.\, \beta(x_k + x_1(\alpha - 1)\left(\frac{k-1}{w}\right)^{ex}) \quad k = 1, 2, \ldots, w \qquad (9.14)$$

$$\alpha = \frac{x_1 + \Delta x}{x_1} \qquad \alpha, \beta \in \mathbb{R}$$

Obviously k may run over any subdomain of $[1, w]$ as well.

Interpretation:

$\vec{d}_j$ for
- DUR accelerando/ritardando, Fig. 9.3, 9.4 } rubato
- At crescendo/diminuendo, Fig. 9.5 } rubato
- T' pitch-contour between words, Fig. 9.6

It might interest the reader to know that (according to my own observations) patterns of accel/rit always satisfy an exponential interpolation in Western music ($ex \neq 0$) but in Javanese music they are always linear. In Western music the non-linearity applies as well to dynamic patterns of cresc/dim.

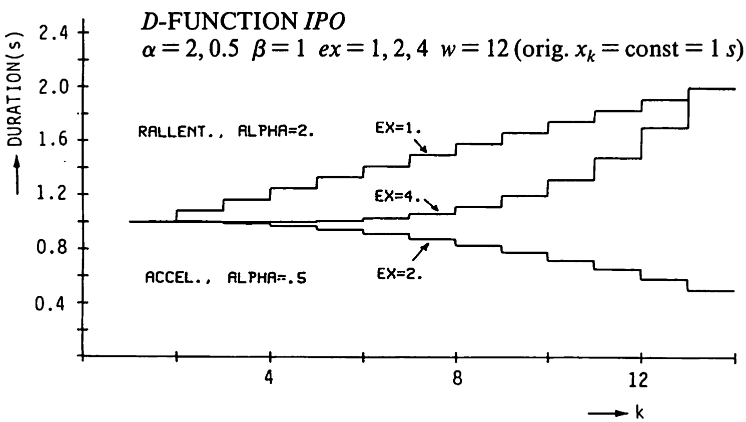


Figure 9.3. D-operation "interpolation"applied to the parameter DUR_k. The original function $x_k = \text{const.} = 1(s)$ is mapped into three different rubato patterns.

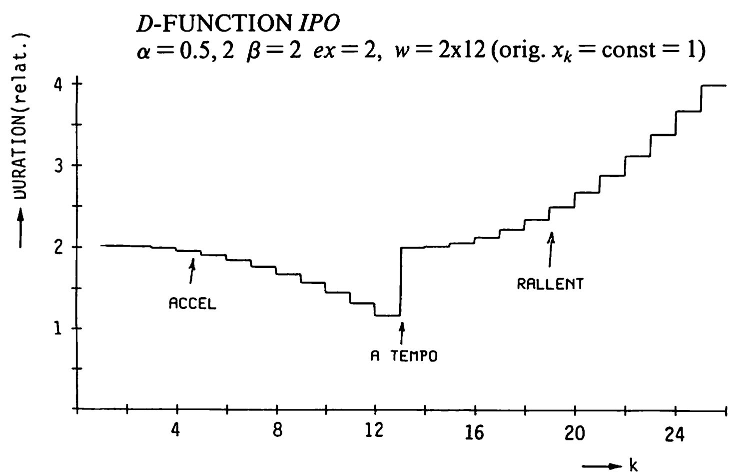


Figure 9.4. D-operation "interpolation"applied to the parameter DUR_k. The original function $x_k = \text{const.} = 1$ is mapped into an "a tempo" preceeded by "accelerando" and followed by "rallentando".

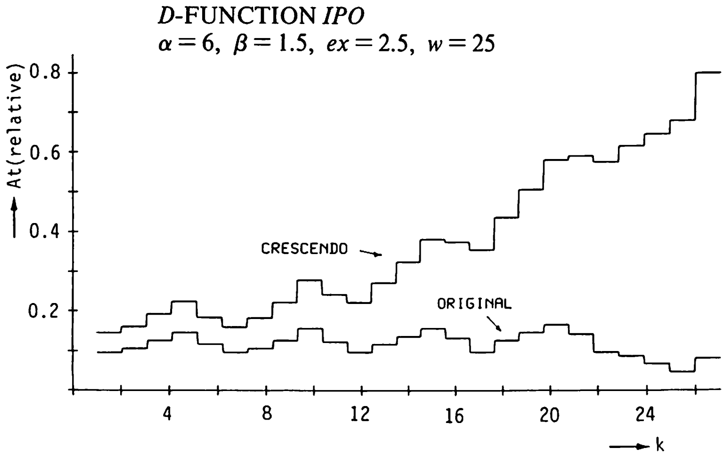


Figure 9.5. D-operation "interpolation"applied to the parameter At_k. A crescendo pattern is superimposed onto the original.

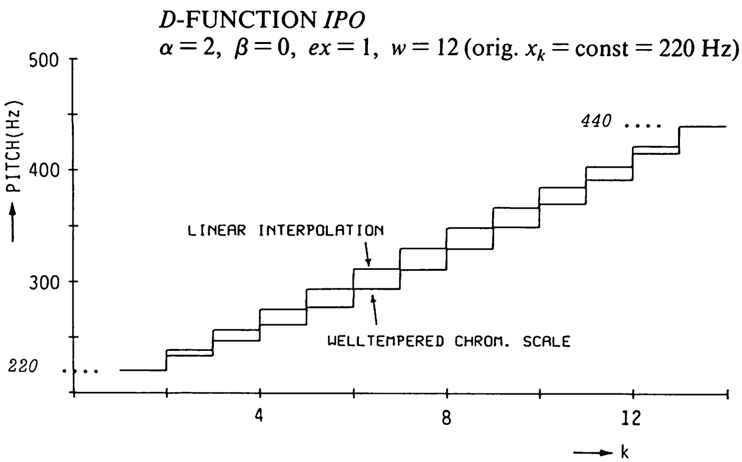


Figure 9.6. D-operation "interpolation"applied to the parameter $10^6/T'$. Linear interpolation over an octave is shown in comparison with the well-tempered chromatic scale.

The real frequency domain is, however, not the means by which the musicians are accustomed to represent the pitch space. It will be easy for a musician to admit that among two time durations the second is twice or half the first, or even that from two sounds the second is four times softer than the first one. But a musician would hardly agree that half or twice the pitch will result in the upper or lower octave. It is true that since ancient times theorists represented the elements of the discrete pitch-space by means of rational numbers, but musicians tended always to represent them merely by *indices*. Almost all the representations known in indo-european music apply the octave in constructing the pitch space. There were many approaches in order to organize the octave-space. However, in all the tunings I know a repertoire of *n* pitches is formed and the *n* elements are assembled in an ordered set in terms of a (falling or rising) scale dividing the octave into *n* fractions. That means, the elements of the scale are numbered or indexed (from 0 through $n - 1$) and any manipulation of pitch operates then merely upon these *pitch-indices*. With the function Φ_3 we introduced a set of well-tempered scales which, in the view of the musician, may be presented as follows (c(Hz) = const., f_1 (Hz) = fundamental, cf. par. 2.3.2):

$$\frac{ln(f_1) - ln(c)}{ln(2)} = {}^2\log\left(\frac{f_1}{c}\right) \sim \lambda oc, pi, sub\ .\ oc + pi/sub \qquad (9.15)$$

Operations in the pitch-space may then be understood in terms of simple linear integer-functions applied to octave- and pitch-indices. Since *sub* is a variable as well, a large amount of well-tempered scales will be available. With $sub = 1200$ the unit Cent is introduced. In what follows we present some operations where the parameter *pi* will be assigned to x_j.

Transposition (familiar to any musician):

$$D_{tra} = \lambda sub, s, z, k\ .\ oc_k + z + (pi_k + s)/sub \quad k = 1, 2, \ldots, w \quad s \text{ for shift} \qquad (9.16)$$

Inversion:

$$D_{inv} = \lambda ocmi, sub, s, k\ .\ s - (pi_k + (oc_k - ocmi)sub) \qquad (9.17)$$

Notice that the pitch-indices pi_k of the original are rewritten in relation with the lowest octave-index occurring *ocmi*, then the *s*-complement is built. $s/2$ is thus the axis applied. Fig. 9.7, 9.8.

An example: $ocmi = 2$, $sub = 12$, $s = 17$

k	$(pi, oc)_k$		D_{inv}	
1	9	2	8	2
2	10	2	7	2
3	11	2	6	2
4	12	2	5	2
5	1	3	4	2

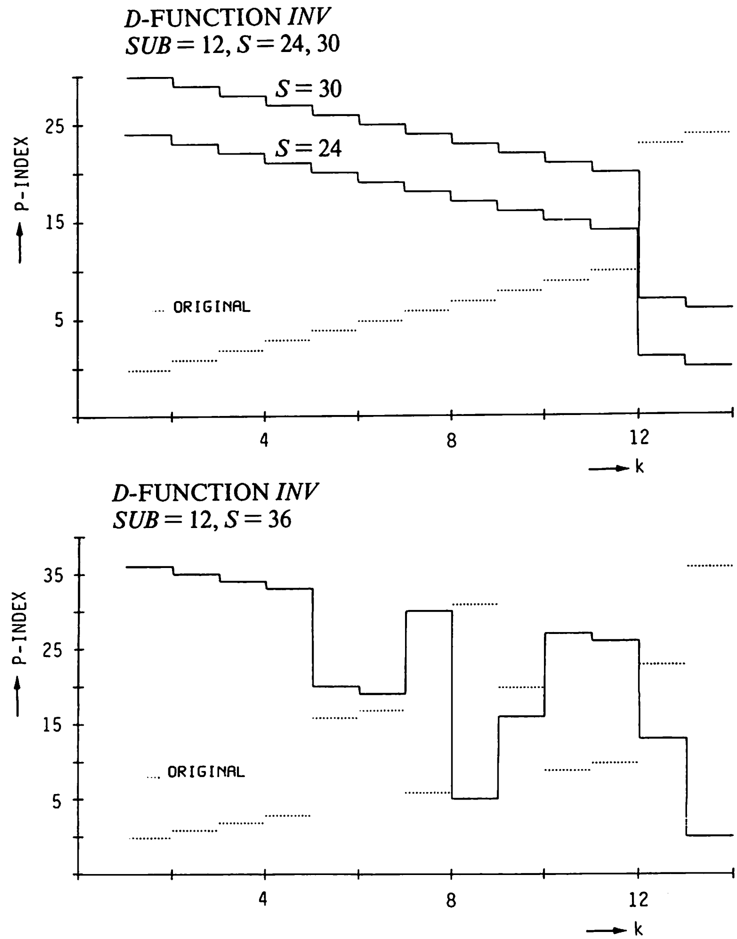


Figure 9.7. and 9.8. *D*-operation "inversion" applied to the pair of parameters (*pi*, *oc*) (pitch- and octave-index, *sub* = 12). *Top:* the original function is mapped into two inversion-patterns with the shifts $s = 24$ and $s = 30$ respectively. The axes between the original and the results are found at $s/2$. *Bottom:* the original is mapped into the inversion-pattern with $s = 36$ (equal to 3 octaves).

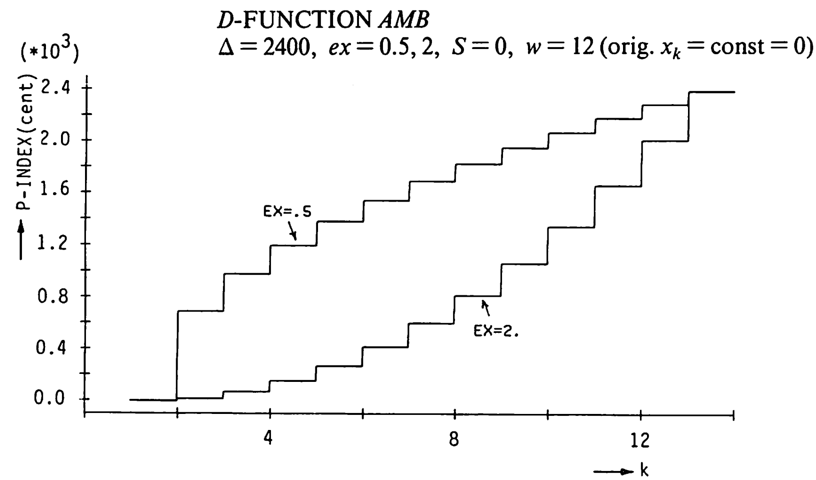
The index-representation of pitch applied by musicians is, of course, not limited to linear functions. Non-linear ones may apply as well provided there is a sufficient pitch-resolution per octave available (e.g. 1200). Here is a small selection.

Ambit (pitch-interpolation over words):

$$D_{amb} = \lambda k, ex, \Delta, sub, s, w, pi_k \,.\, \mathrm{rof}(pi_k \cdot 1200/sub + s + \Delta\left(\frac{k-1}{w}\right)^{ex}) \quad (9.18)$$

shift s (Cent) $k = 1, 2, \ldots, w$
Δ (Cent)
$ex \in \mathbb{R}$

Notice that *sub* is the value stored in the header of the original; it will be replaced by 1200 valid for the result. Fig. 9.9-9.11.



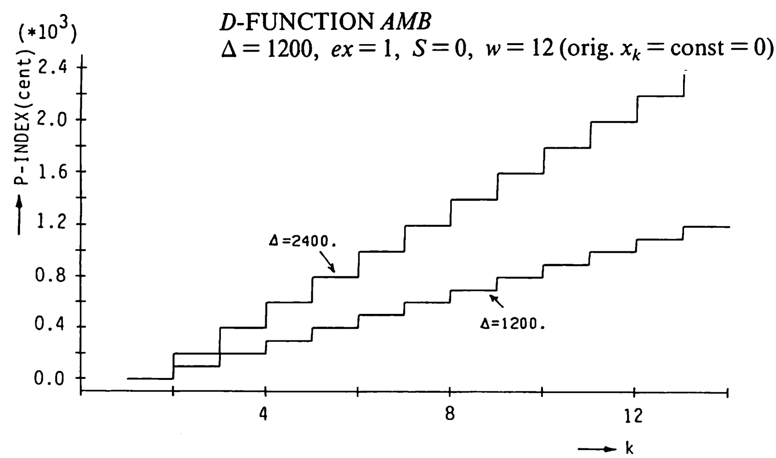


Figure 9.9. and 9.10. *D*-operation "ambit" applied to the parameter *pi* (pitch-index, $sub = 1200$). *Top:* the original function of $k = \text{const.} = 0$ is mapped

into two exponential scales ($ex = 0.5$ and $ex = 2.$) ranging over two octaves ($\Delta = 2400$(cent)). *Bottom:* the same original function of k is mapped into two linear scales ($ex = 0$) ranging over one and two octaves respectively.

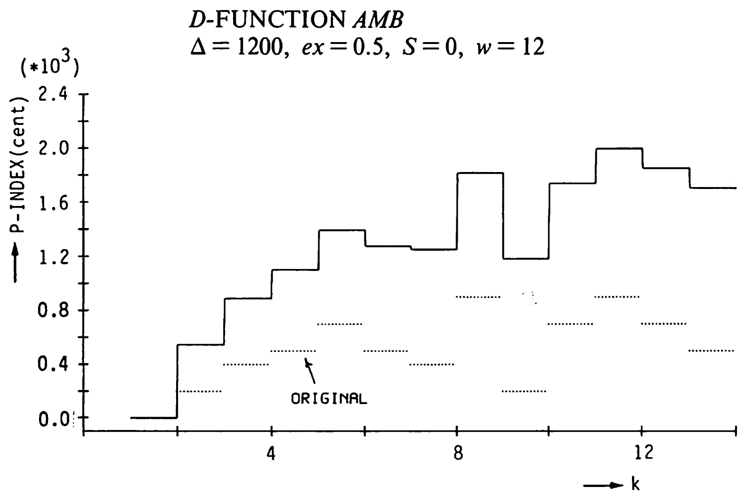


Figure 9.11. The original (theme of the fuga 1 from J.S. Bach's Welltempered Piano I) is mapped into a new pitch-pattern which does any more fit into the well-tempered chromatical scale.

We add still an operation which allows the formation of *non-standard* pitch-scales within a selected interval of pitch-indices [pimin, pimax]. For sufficient resolution the indices are again mapped into $sub = 1200$.

For any $pi_k \in [pi\text{min}, pi\text{max}]$:

$$D_{nst} = \lambda ex, pi\text{min}, pi\text{max}, sub, pi_k, k \;.\; (pi\text{max} - pi\text{min})$$

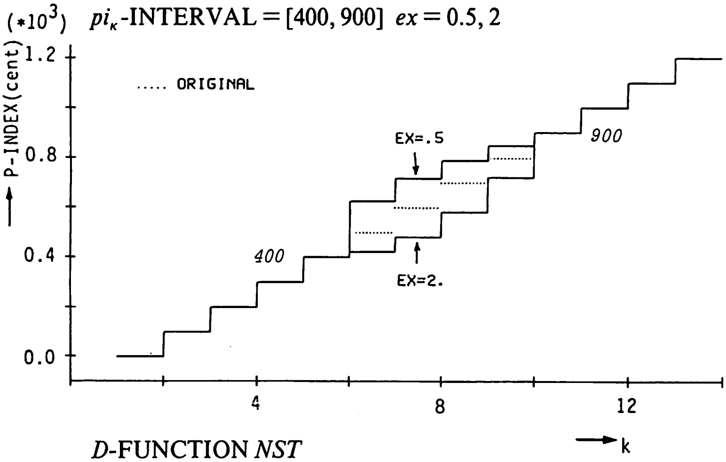
$$\left(\frac{pi_k \cdot 1200/sub - pi\text{min}}{pi\text{max} - pi\text{min}}\right)^{ex} + pi\text{min} \tag{9.19}$$

Among the interpretations are e.g. the formation of the chromatic and enharmonic scales of the ancient Greeks, the "Leitton"-tendencies occurring in vocal music and in playing unfretted string instruments, areas of blue notes etc. Here is an example:

$sub = 12$ $\quad pi\text{min} = 400$
$ex = 2$ $\quad pi\text{min} = 900$

k	pi_k	D_{nst}
1	0	0
2	1	100
3	2	200
4	3	300
5	4	400
6	5	625
7	6	715
8	7	785
9	8	845
10	9	900
11	10	1000
12	11	1100
(13	12	1200)

D-FUNCTION *NST*
pi_κ-INTERVAL = [400, 900] $ex = 0.5, 2$



D-FUNCTION *NST*
pi_κ-INTERVAL = [0, 200], [1000, 1200], $ex = 0.5, 2$

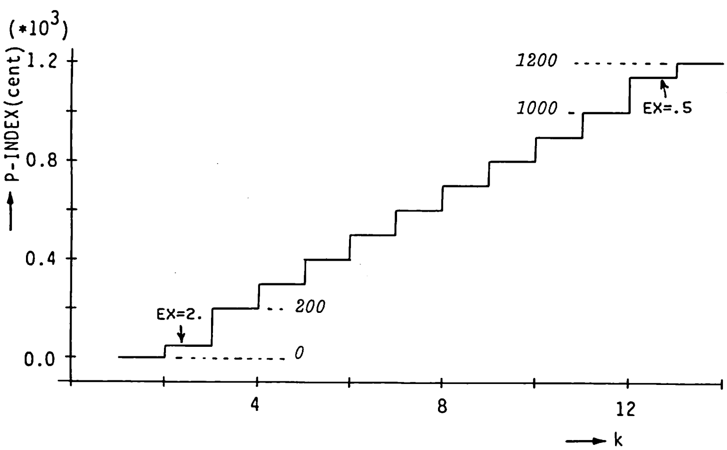


Figure 9.12. and 9.13. (pg. 156) *D*-operation "non-standard" applied to the parameter *pi* (pitch-index, *sub* = 1200). *Top:* the original chromatic scale is mapped into two scales with non-linear pitch-instances within the interval [400, 900] (cent). *Bottom:* the original chromatic scale is mapped into a scale with stressed "Leitton"-tendency within the intervals [0, 200] and [1000, 1200] (cent).

Summarizing we may say: In manipulating pitch the musician operates upon the indices for octave and pitch and accordingly he deals with integers as mentioned above*. By this he keeps his calculations close to his aural experience. But for the musician there is still another advantage of the index representation applied to pitch since time immemorial. Since the pitches are not represented by their contents but by the indices allocated to them the mapping of one tuning into another is rendered extremely simple. All the musician has to do is to chose the appropriate subdivision and to assign the new contents to the indices (in fact, to change the instrument and apply the old indices!). This is the way the MIDIM language introduces non standard pitch systems like the pythagorean and various Gamelan tunings. Cf. Kaegi 1984 and Janssen/Kaegi 1986 on p. 201 of this issue.

9.2.3 Functions of $\chi(x_{kj})$ or Conditionals

We assume there to be a domain $A = [a, b] \subset D_j$, a constant $c \in D_j$, and a characteristic function of x_{kj} as follows:

$$\chi(x_{kj}) = \begin{cases} \text{if } x_{kj} \in A \text{ then } 1 \\ \text{otherwise } 0 \end{cases} \tag{9.20}$$

The mapping D_{con} is now as follows:

$$D_{con}: \chi(x_{kj}) = 1 \text{ implies } x_{k'j'} = c \tag{9.21}$$

The mapping selects within the column j of the descriptor all x_{kj} where $\chi(x_{kj}) = 1$ and rewrites $x_{k'j'}$ into c. Notice that $k_{j'} = k_j + s$, where s is a line-shift.

The function may be dependent upon more than one argument as follows:

$$D_{con}: (\chi(x_{kj}) \wedge \ldots \wedge \chi(x_{kj''})) = 1 \quad \text{implies } x_{k'j'} = c \tag{9.22}$$

For $j' = j$ and $s = 0$ the mapping is a CONVERT-operation.

Interpretation: x_j and $x_{j'}$ may be assigned any descriptor-column.

* Serial pitch-organization is entirely based on this representation.

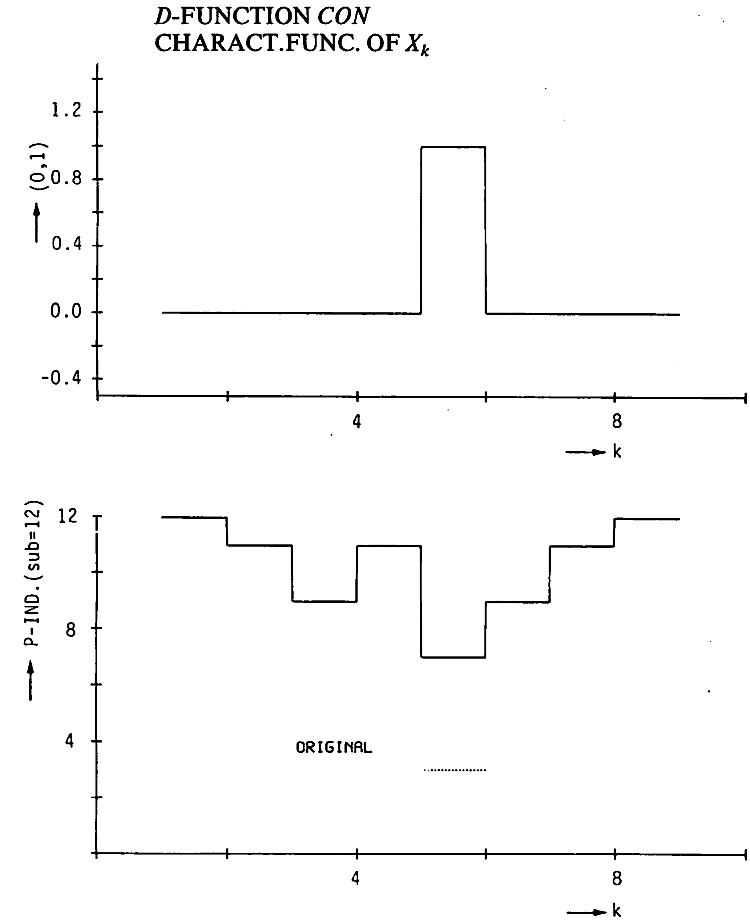


Figure 9.14. *D*-operation "conditional". The original (pitch-)function of *k* is mapped into the correct melody of Martin Luther's song "Vom Himmel hoch da kom ich her" by means of the convert-option.

10. COMPOUND PREDICATORS AND *D*-COMPREHENSORS

10.1 Representation of compound sound concepts

We know that any predicator occurring in a formula stands for a single sound-concept. There are, however, also *compound* sound concepts. On may e.g. think of the sound of a cuckoo which is impossible to represent by merely one single sound "koo"; a pair of these is needed, in which the prosodic controls are tied by constants representing the pattern of a falling third where pitch, tempo and loudness are fixed. Another example would be "laughing" where a sequence of at least two single concepts "ha" is needed in which the prosodic variables are once more tied by a pattern which, however, now varies in many ways.* Such compounds may, of course, also consist of different sound concepts. Almost all the words of spoken language where more than one single syllable occur are of this type ("hello", "MIDIM" etc). In this case the prosodic variables are replaced by "linguistic" functions which depend upon the intonation- and stress-rules of the linguistic sound system concerned. Cf. par. 5.1.

In the MIDIM-language any kind of compound sound concept may be represented by compound predicators or *P-compounds*. The latter are constructed as follows:

(1) We assume there to be the (non-empty) *P*-library *LIB* and the mapping $SUC(p)$ of the variable *P*-index p (where $a, b, c, \ldots$ are integer constants). *SUC* allocates to the indices of selected predicators $\in LIB$ the index of the successor according to a user defined table as follows, e.g.:

SUC: p	$suc(p)$	*P*-compounds where $(\ldots)$ is a cycle:
a	b	
b	c	$= (abcd)$
c	d	
d	a	
e	f	$= (ef)$
f	e	
g	h	$= gh(fe)$
h	f	
i	i	$= (i)$

(2) We assume there to be any arbitrary formula where the *P*-indices p_k occurring are defined recursively as follows:

$$\lambda x(p_1 = x) \tag{10.1}$$

$$p_{k+1} = SUC(p_k)$$

Accordingly the formula will exhibit for p_k, $k = 1, 2, \ldots, w$ the following instances:

$$(x)_1 (suc(x))_2 (suc(suc(x)))_3 \ldots \tag{10.2}$$

Such an expression will be called in what follows a *D-comprehensor*.

---

* I remember one case where a person expressed "laughing" by means of a single "ha". Otherwise it conveys meanings like astonishment or hate according to the amplitude-envelope, intonation-contour and timbre.

(3) By assigning a constant to x the *D*-comprehensor will be linked with the above *P*-library via the assignment table. E.g. for any *art*-index *art* (where $g \in G =$ set of grammars):

$$\lambda x(p_1 = x)(a) \leftrightarrow p_1 = a$$

entails the formula

$$g((a, art)_1 (b, art)_2 (c, art)_3 (d, art)_4 (a, art)_5 \ldots) \in M_g \qquad (10.3)$$

and likewise for

$p_1 = e, p_1 = g$ and $p_1 = i$ respectively:

$$g((e, art)_1 (f, art)_2 (e, art)_3 (f, art)_4 \ldots) \in M_g \qquad (10.4)$$

$$g((g, art)_1 (h, art)_2 (f, art)_3 (e, art)_4 (f, art)_5 (e, art)_6 \ldots) \in M_g \qquad (10.5)$$

$$g((i, art)_1 (i, art)_2 (i, art)_3 \ldots) \in M_g \qquad (10.6)$$

10.2 Context sensitivity

We still mention the possibility of defining various types of *context sensitivity* for the *D*-comprehensors. The instructions concerned are stored in the header. When two formulae are concatenated where the second is a *D*-comprehensor then the preceding formula is considered to be the context of the comprehensor upon which the latter will react. This renders it possible to form prosodic patterns which may be assigned to *P*-compounds, and in this way a field of powerful compositional strategies is opened. For more information concerning context-sensitivity the reader may refer to Kaegi 1984. Cf. also Kuipers on p. 257 and v. Berkel on p. 231 of this issue.

ACKNOWLEDGEMENTS

I wish to thank the numerous users of the MIDIM-system who provided me with precious suggestions; the technical staff of the Institute for Sonology who looked after the computers and peripherals; the members of the MIDIM Group and among them especially my son Heinerich for the constructive criticism concerning the formalization of the language, and my assistant Paul Goodman for his help in the english presentation of this paper.

REFERENCES

Banerij, R.B. (1969). Theory of Problem Solving An Approach to Artificial Intelligence. New York, London, Amsterdam.

Carnap, R. (1947). Meaning and Necessity. Enlarged ed. Chicago 1956.

Church, A. (1941). The Calculi of Lambda-Conversion. London.

Hindley, J. R. et al. (1972). Introduction to Combinatorial Logic. Cambridge.

Janssen, J./Kaegi, H. (1986). MIDIM-Duplication of a Central-Javanese Sound Concept. Interface Vol. 15, 185-229.

Kaegi, W. (1973/1974). A Minimum Description of the Linguistic Sign Repertoire. Interface 2, 141-156 and Interface 3, 137-158.

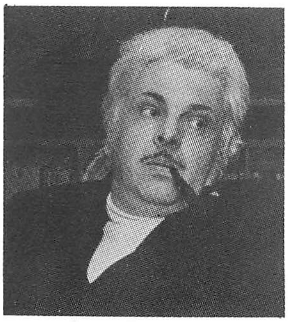
Kaegi, W. (1984). The MIDIM8X System, DESC users Manual. Utrecht.

Kaegi, W. (1985). The Design of Musical Discovery. Utrecht.

Kaegi, W./Tempelaars, S. (1978). VOSIM - A New Sound Synthesis System. Journ. AES 16, 6, 418-425.

Kuipers, P. (1986). CANON, Interface, Vol. 15, 257-269.

Tempelaars, S. (1977). The VOSIM Signal Spectrum. Interface 6, 81-96.



WERNER KAEGI, born 1926 in Switzerland. Studies in musical composition, piano, clarinet, musicology and mathematics in Zürich, Basel, Heidelberg and Paris. Ph.D. 1952. Student of Paul Hindemith, Arthur Honegger and Louis Aubert. Comes into contact with electro-acoustic music for the first time in 1951 at the Paris radio. Numerous compositions of instrumental and electronic music (among others electronic music at the Swiss pavillion at the Osaka World Fair, Japan 1970). From 1963 to 1970 he worked for the Centre de Recherches Sonores at the Radio Suisse Romande in Geneva, and since 1971 he has been staff member at the Institute for Sonology where he developed his sound synthesis system VOSIM and the MIDIM programs.

Interface, Vol. 15 (1986), pp. 163-184. 0303-3902/86/1502-0163 $3.00

The User and The MIDIM System

Paul Goodman

The MIDIM Group would like to dedicate this article to the memory of Peter Groenland (1956-1985).

ABSTRACT

A description of a practical workshop for users of the MIDIM8X system is given followed by a report on some concert activities over the past four years and commentary on a few compositions composed using the MIDIM system. The descriptions are both short and informal and where necessary the reader is referred to sources for more detailed information.

THE PRACTICUM

You might imagine after reading through the theory concerning the MIDIM system that it is a horror to teach and that anyone making the attempt to come to terms with it will remain stranded in theoretical questions and never reach the stage where they may begin to work practically. This has been shown by experience not to be the case.

The programs which comprise the system (see Fig. 1), of which we will concern ourselves here with the Descriptor and the Predicator programs, are set up in such a way that the user encounters a series of high level questions (actually a questionnaire) to fill in as they work their way through the programs. The above mentioned programs are constructive programs which give the user the possibility of creating sounds. Of the two the Predicator program is the most unfamiliar and therefore the most difficult. The terminology is sometimes less accessible at first but once the vocabulary has been learnt one can handle the questions with relative ease. It is also the case that the results are always well-formed and the domains within which one must work are clearly defined, which allows the user to work without undue attention needing to be devoted to technical problems.

In order for users to reach this point a practicum course was given by myself and Jos Janssen for two years at the Institute for Sonology of the University of Utrecht. The course was set up in such a way that the morning classes of Dr. Kaegi dealt with some item of the theory and then in the afternoon class Jos and I would work through the system with the students in order to give practical examples of the items discussed in the morning (see Fig. 2).

It is important to point out that none of the composer-students had any acquaintance with the system prior to attending the classes and yet by the end of the year they were able to open up individual projects resulting in many of the

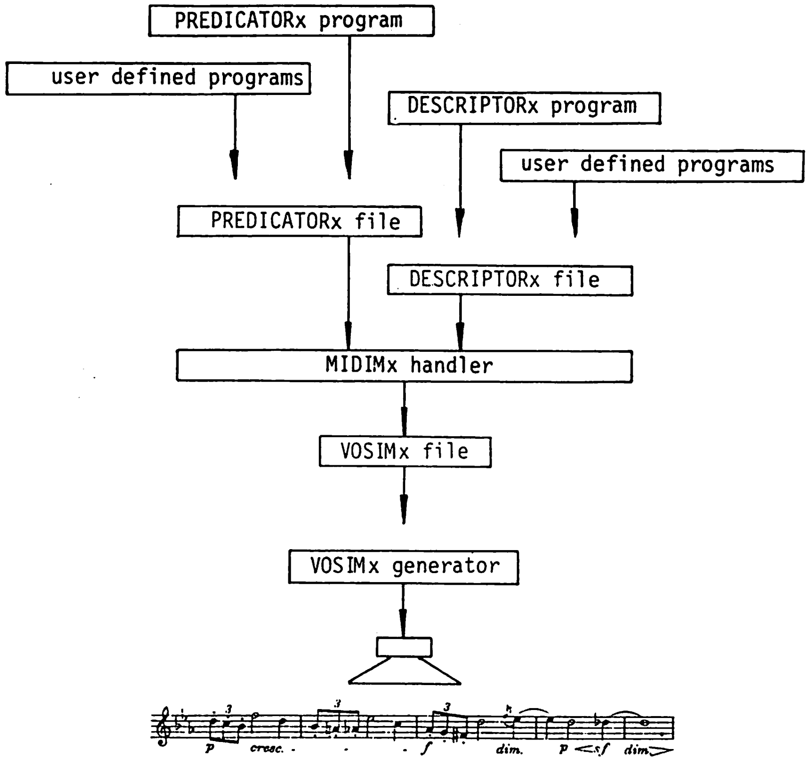


Fig. 1. Survey of the MIDIM system

cases in completed compositions which were of high enough quality to be given in concert.

In general one nearly always begins with the Predicator program which is to say with a concrete sound and then commences to work on the Descriptor level (macro-level) but by experience we found it more convenient to introduce the system by means of the Descriptor program and then to work our way towards the construction of the sounds themselves in the Predicator program.

As was stated above, the terminology of the Descriptor program is kept wherever possible to that used in musical practice and is therefore familiar to students coming from conservatories or studying musicology at the university. In the Descriptor program one builds up a sort of score where first values for a Subdivision and a Metronome must be given as reference points. The Subdivision may take any value between 1 and 1200 which allows for any audible segmentation of the octave. The Metronome corresponds exactly to musical practice; so for example if one should give a value of 60 to the Metronome then a quarter-note will equal one second. These are the first two elements of the questionnaire and afterwards one proceeds to fill in Arrays for each of the Parameters.

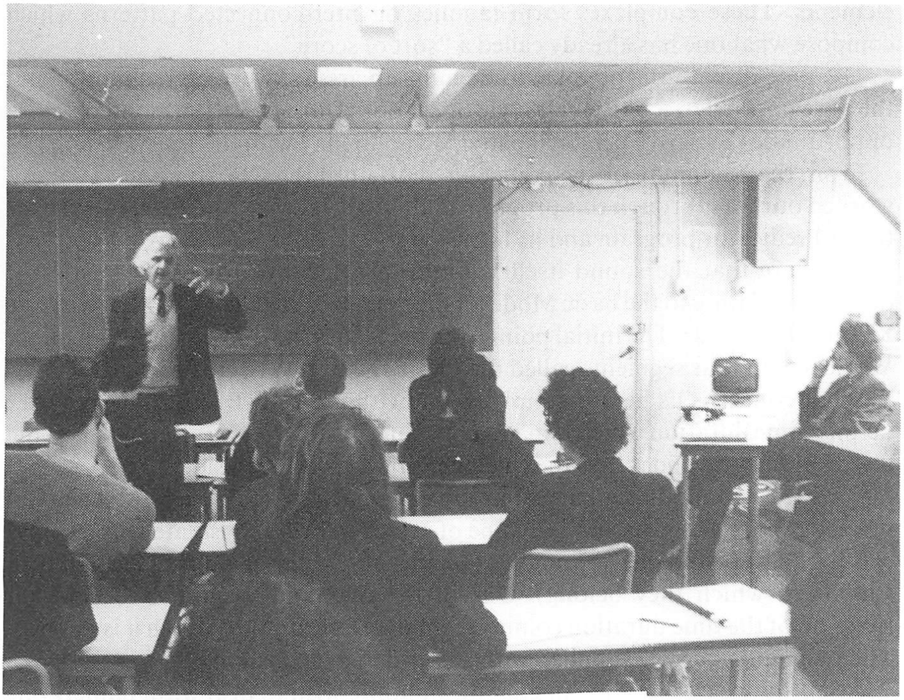


Fig. 2. Morning class for the theory.

These are filled in by answering two questions for each parameter, namely "How many?" and "Which ones?". The parameters being in the following order: Pitch, Octave, Duration, Intensity, Articulation and Predicator Index. Once these Arrays have been filled in for two separate tracks, then they are combined by means of Ordering Operations to form a completed Descriptor File. The operations range from totally ordered to totally random orderings. It is possible to give Subarrays of the arrays and to operate on them apart as well as to "accumulate" operations in order to produce complex orderings of the basic elements stored in the Arrays.

Once this stage has been reached then it is offered one the possibility to go a step further and to operate on the Descriptor file(s) in order to produce families of Descriptors each derived from the primitive file. The operations offered at this level extend from a simple printing out of the file to editing it, transposing it, overlapping the D-tracks onto each other (which is necessary for example if one is using a piano duplication), creating functional networks from the elements of the Descriptor file, accelerando, crescendo, retrograde, inversion, mirroring, tuning operations and so on (see Kaegi, Descriptor users Manual, 1984 and Interface, 1986, MIDIM, par. 9).

In short one forms in a constructive way complexes built up upon basic elements. These complexes form families of inter-connected patterns which compose what one has already called a "sort of score".

The type of patterns possible to be built up are of a very high quantity and they are not tied to a particular style or school. This gives the user much more opportunities to derive personal patterns and to play with them.

From the beginning of the course in Sept. until the Christmas vacation we worked our way through this program and then after the holidays stepped over to the Predicator program and its intricacies.

It is here that the sound itself is built up again by answering a series of questions relating to the basic Model of the system called the VOSIM model (see Kaegi, 1986, p. 72). The initial point to be decided is the time basis of the sound. We deal with four segments called the PREFIX, BODY, SUFFIX and STOP.

What concerns us here is simply to know that the PREFIX is given a user defined time duration which remains constant. The PREFIX is equivalent to the attack of a sound and is thus the chief identifier for fixing the identity of a sound (eg. a struck sound has a particular type of attack (see Janssen/Kaegi, 1986) while a French Horn has its own type of attack in order to isolate it as a sound class or concept and they are not inter-changeable without altering the family of sounds to which they belong). One of the remaining segments becomes a function of the time duration coming from the Descriptor to which it is applied. The remaining two are supplied with a user-defined constant value.

After this we choose the category of sound we wish to deal with, namely one exhibiting a Moving Formant (function 1 in the theory) or a Fixed Formant. We always know in advance which we would like to use. Once we have stated which class we deal with, we are offered a number of features which our sound may display such as Portamento, M-modulation (which gives an intonation contour or vibrato to the sound), a standard linguistic characteristic fixed to one of the segments (eg. S, T, H, Sch, K, ...in the PREFIX), etc.

An amplitude envelope contour must be stated for the four segments which gives them their deportment in time.

As in the Descriptor program we have the possibility of using two independent tracks or voices but in some cases it is necessary to combine both predicator tracks into a single voice in order to produce more complex spectra. So for example we have three possible ways to LINK the tracks (a Linked predicator merely reads the first D-track, while an unlinked predicator reads both D-tracks). The first states that the fundamental coming from the Descriptor will be equivalent for both tracks so $F1 = F2$. The second states that a ratio may be set (eg. at a perfect 5th, or a minor 3rd or any interval one may wish) between the two tracks. The third and last possibility is to define a constant Beat frequency for the two voices.

If we should choose not to LINK the tracks, then information for the second track must be filled in either by running once more through the questionnaire or by simply telling the program to copy the same information from the first track onto the second which is then automatically done.

The resulting predicator file is given a name and then is stored as a 12 × 10 matrix. As in the case with a Descriptor file this primitive file may now be operated upon to produce a family of sound concepts derived from the original. This is accomplished by means of a sophisticated screen editor as well as P-operations. In this way one can step slowly away from the original sound to a derived one (s).

The original predicator and its derivations can be stored in a P-library under a given name in which they are indexed according to their rank order and can then be retrieved through assigning to the Descriptor locations the various Predicator Indexes stored in the Library. So for example we call for the fifth Predicator in our Library with the Index number 5 and so on for the remaining Predicators.

After Easter we started on the last phase of the practicum which consisted of putting into practise the skills learnt in the previous months. There are many ways of dealing with the system, for instance in a highly reflective way or in a quick, spontaneous-intuitive manner. As an experiment (one of many which we made) the process known as 'Ecriture Automatique' (Automatic Writing) was attempted. This process was devised by the french surrealist writers André Breton and Philippe Soupault and used for the first time in their collection entitled "Champs Magnétiques" (see Breton, A./Soupault, P., 1919). It proceeds as so in practise. We set a time duration (limit) within which one may work with one's material (elements of the language) and then set a tempo which is fast enough that we cannot reflect or rationalize what we are doing. This has the purpose of blocking out the conscious-logical area of the mind in order to reach unconscious sub-strata and to give free expression to this. The technique presupposes a language with a well-defined grammar and knowledge of how to use the language. In this case the language is the MIDIM language whereas in the case of Breton and Soupault it was literary french.

After a practise session to become acquainted with the working procedure we immediately began. Each of us taking a turn at the terminal producing a Descriptor and a Predicator(s) within a time duration of 7 min. Once the material had been produced we collected it and then Dr. Kaegi produced in an "automatic" way an ordering and an overlapping of the material by means of a command string. The resulting VOSIM file was then directly recorded on tape. The last step was the mixing in the studio and the decision as to an end point. We disposed the material in quadrophonic space and added some reverberation and in this way the compositions 'Impromptu 1, 2 and 3' were produced. All three were given a number of times in concert, combined with texts produced in real time by means of a computer program for écriture automatique.

Once the practicum was finished students concentrated on personal projects. So in this way the practicum extended over into the area of independent artistic activity.

The group consisted of:

Composition Group:	(Impromptu 1, 2 and 3)
	Annechien Cosijn (Holland) 3
	Ramon Gonzalez-Arroyo (Spain) 3
	Paul Goodman (Canada) 1, 2, 3
	Peter Groenland (Holland) 1, 2
	Ben Guttman (Israel) 3
	Jos Janssen (Holland) 1, 2, 3
	Pieter Kuipers (Holland) 1, 2
	Jack MacDermot (Canada) 3
	Tim Plowman (United States) 1
	Anna Rubin (United States) 1
Automatic Texts:	Paul Goodman (Canada)
	Heinerich Kaegi (Switzerland)
	Han van der Vegt (Holland)
Automatic Painting:	Pieter Entrop (Holland)
	Wim Gregoor (Holland)
	Antoine America (Holland)
Automatic Sculpture:	Gerda van der Krans (Holland)

REPORT ON SOME CONCERTS GIVEN BY THE MIDIM GROUP

For the last four years there has been given annually at the Music Centre Vredenburg (Utrecht) and at the Geertekerk (Utrecht), concerts devoted to the most recent compositions produced with the MIDIM system. Each concert is an attempt to try something new for ourselves and for the audience. They are considered actually as experiments.

The first MIDIM concert took place on Oct. 11, 1982 at the Geertekerk. As a theme for the event the title "Associative Music" was chosen which was defined in the program notes by Werner Kaegi as follows:

What is associative music?

The MIDIM system opens new ways of conceiving music. Among the most interesting possibilities I will mention the following:

1. Every sound output of the MIDIM system is *associative but not commutative.* MIDIM is a language, therefore the MIDIM system contains a grammar which forms musical sound sequences according to the rules of a formal language. Every sound sequence is an element of a sound universe defined by the composer within the MIDIM language and follows its rules. Therefore it is associative but not commutative. Thus (ab) c=a (bc), but not abc = acb.

2. Every sound output of the MIDIM system is *associative but not sequential.* The MIDIM system ignores "scores" going to be, so to speak, orchestrated note by note. Working with the MIDIM system the composer defines his own sound objects and their behaviour. Each sound object which is called for determines then its own behaviour, either totally or partially by calling in its turn defined behavioural patterns and activating them through fields E.g. an "oboe" would realize its behaviour in a different manner than would a percussion sound or an

f-like noise band. The objects called for may also define their own followers which leads to the construction of complexes. So the process of composing accordingly is not a sequential progression within a list of notes but a network of calls on different levels. The calls operate by random access.

3. The MIDIM-system allows the composer to form his music *associatively in a colloquial sense.* Composing becomes a process of mental associations in the form of sound complexes which can in many ways metamorphize one into the other. By this the composing process comes close to the way music is understood by the listener in my opinion, not in terms of a progressive summation of sounded notes along a time axis, but in terms of an associative formation of mutual relations between sound complexes which make part of a defined universe.

The concert consisted of the following six compositions:

"Suggestions from Limbo"	by Paul Goodman
"Voyager II"	by Claude Fatus
"Circe"	by Cecilie Ore
"Consolations":	by Werner Kaegi

a) In Memoriam
b) Automne
c) Vers d'autres Jeux

Each composer in their own particular way made use of the MIDIM system to compose a sound world distinctly their own. All the sounds heard were built up step by step by the composers (except certain material in "Suggestions from Limbo", see commentary on piece for details) to fit the ideas they wished to make concrete. This is what gives the MIDIM system such rich possibilities for the composer-user since they are free to create very personal sounds (if they have the imagination and skill to do so) which express their feelings and not to be caught in the situation where standard sounds must be used to express things deeply personal which belong solely to the artists themselves.

In this case the compositions were all purely tape compositions (in other words no soloists were involved in the concert) and played back in quadrophonic space. From the computer itself it is possible to hear directly back one or two voices at a time, so for multi-voiced pieces one records the material on tape and overlaps and mixes it in an analog studio by means of multi-track machines. (This process was dependant upon the format of the studios at Sonology and not determined by any limitation in the program itself). The disposition of the material in space was the last stage in the production of the master tape.

This concert was later given at the Stedelijk Museum in Amsterdam on May 4th, 1983.

The following year the main focus of interest was the participation of the Indonesion musician Supanggah Rahayu and a Gamelan orchestra. The composition "Dialogue" by Werner Kaegi was composed for Supanggah Rahayu who performed it on a Kendhang, which is an Indonesian percussion instrument. The player reacted to the tape which was composed of sounds derived from his instrument. Since the sounds presented to the player were recognizable to him he could react quite spontaneously just as he would if playing with another instrumentalist in place of the tape. Thus an interaction between performer and tape arises of a very intense kind.

The MIDIM system is capable of duplicating the sounds of spoken language (e.g. Kaegi, Consolations, Part 3, synthesis of a poem of Charles Cros) and for this reason it is possible for a composition such as "Dialogue" to be produced. Since the drummer's musical vocabulary is heavily linguistic, the cues on the tape must also be linguistic in nature and of a high enough quality that he has no trouble in recognizing them as belonging to or derived from the standard musical repertoire he accepts as belonging to his instrument (see Janssen/Kaegi, p. 192).

A similar situation arose for Jos Janssen and myself when we worked on a composition for Jazz drummer and tape (in collaboration with the jazz musician Pierre Courbois). We submitted a tape of the composition to the Jazz drummer Martin van Duinhoven who after listening to it said in approximately these words, "Yeah, I'll do it, its my vocabulary".

Also involving the use of a linguistic repertoire is the synthesized spoken language of the sound-poem "Tjaktjakai" by Pierre van Berkel. This is an abstract language composition in two parts, the one for computer tape and the other for live recitation by a male voice. The performer for all the concerts involving this piece has been Sjabbe van Selfhout. The tape is composed completely of computer synthesized speech sounds created by means of the MIDIM system (see van Berkel, 1986).

The sound-poem follows in the tradition of Hugo Ball, Kurt Schwitters and Theo van Doesburg in creating from abstracted linguistic elements, patterns and particles of speech and ordering them.

Included in the concert as well was an experiment conducted by Jos Janssen with a Gamelan orchestra. One of the instruments of the orchestra, namely the Gender (see Janssen/Kaegi, 1986) was replaced by a computer-synthesized duplication of the instrument and the part it had to play in the composition performed by the ensemble. The composition was a traditional piece of music well known to all Indonesian musicians and therefore made judgement concerning the results possible. It is not within the scope of this article to discuss the outcome or the reactions which took place but simply to say that the composition was capable of being performed in this way.

The composition "Wounded" by Paul Goodman (see commentary in next section) and an improvisation upon a Gender by Supanggah Rahayu were also presented that evening.

The concert devoted to Ecriture Automatique was given thrice last year in three varied forms but remaining basically the same concert. The concerts took place in Utrecht at the Music Centre Vredenburg, at the Festival for Modern Music in Arnhem and in Rotterdam at a conference for Visual Arts and Computer.

In order to give an idea of the atmosphere accompanying each of the concerts (each has its own particular character) there now follows a description, in my best journalistic style, of a concert given in 1985 by the MIDIM group.

We can imagine the concert occurring in this way. The audience is called via the tone to the concert hall. They seat themselves in a lit room and wait. Two painters enter from stage right and go towards a table filled with paint supplies and prepare themselves. Kaegi ascends the podium and gives an explanation of what is about to happen (the rules of the game) and then gives the artists 7 min to complete their work. They start and the strained sheet of plastic stretched across the centre-back of the stage slaps back and forth as they strike and stroke it on both sides as if they are fighting with each other. The time is up and they stop reluctantly and a dual sided composition in colour and line is presented to the audience who have watched its production. The lights go down and two spots throw down a column of light on each side of the stage. Two young men with typewriters (in the Rotterdam version the typewriters were replaced by Apple computers and an associative text program) sit stage left and right in the spots with the coloured plastic displayed behind. A tape of Impromptu I begins to fill the hall with strange patterns and they begin to type furiously and trance-like in response to the music. This proceeds for 5 min and then there is silence with the plastic still displayed behind. Then Impromptu II begins and the writers recite their texts into microphones as the tape accompanies them. The texts are brooding, desolate word patterns and the tape is composed of weird and moody sounds which are strange to the ear (see list p. 168).

> "......no one sees them but myself and the rose it is no more a garden and it is now so dark I cannot breath suffocating no sense around me but this silly singular plant that entwines like the moonglow around my feet moon plant and leaves that burst with song. You are frightened of them because they bite and smoke and chew the earth to pieces that midnight is this world that is all....."

Next coloured spots appear in the darkness and Kaegi's piece "Chants Magnétiques" dedicated to Breton and Soupault begins. The sounds of the piece with their full-modulated spectra are aggressive and angry and the sound level is up high to increase the atmosphere of violence.

There follows the piece "Episodes" for tape and recitation. The soloist Dieuwke Aalbers stands centre-stage facing the plastic sheet as if contemplating it while quasi-keyboard clouds surround the audience. As the tape reaches the end of the keyboard-sounds she turns to face the audience and recites a text composed as if in a dream and deep BOOMS are heard from the loudspeakers. Then a wild cry reaches an hysterical pitch and she begins with abstracted and dreamy phrases until everything fades to nothingness and organ-like sounds

STADSSCHOUWBURG ARNHEM
12 JUNI 20.30
KLEINE ZAAL

écriture automatique

COMPUTERMUZIEK
SCULPTUUR
POËZIE

MIDIM-groep voor computermuziek (o.l.v. Werner Kaegi) / Dieuwke Aalbers, sopraan / Gerda van der Krans, sculptuur / Paul Goodman en Heinrich Kaegi, poëzie

Festival KIJK . . .MUZIEK, toegang: f 12,50, cjp: f 7,50

Fig. 3a. Poster announcing concert in Arnhem, June 12, 1985, given by Midim-group.

Fig. 3b. Poster announcing concert in Utrecht, given Jan. 14, 1986 by Midim-group.

collect and then fade leaving the BOOMS alone repeating and her small, silent figure, head bent forward, waiting.

After the pause the composition "La Belle et La Bête" commences. It is a pure tape piece mixing tunings from different Indonesian instruments and traditional western pitch systems. It is pained and sad just as the fairy tale and animal-like sounds come and go inter-married with soft, gentle sounds of a feminine nature. It ends with the Beast and the Belle blended in a duet.

Then the coloured spots change once more and Dieuwke Aalbers is once more on the stage. The composition is entitled "Ritournelles" for Soprano and computer by Werner Kaegi. Her beautiful voice interplays with the tape in a highly dramatic way. It is as if she must extend herself, reach further in order to stay with the tape which runs on without effort. She sings only linguistic sounds in melismatic lines and must constantly adapt herself to the tunings used for the tape. The piece is surely a high point in computer music where one no longer speaks about this or that technique or computer program but can simply say it is music which has gone beyond questions of technology and has reached the realm of aesthetics.

The audience responds with a standing ovation and the concert is finished (see "Appendix I" for a list of concerts 1982-86).

A FEW SHORT COMPOSITION DESCRIPTIONS

The reason for the little descriptions which follow is to give some idea of actually working with the MIDIM system. It is not a bloodless reasoning process devoid of intuition and emotional content but as with most creative work the intuition should be free to test and play when building the material and ordering it. Kaegi wrote his "automatic" piece "Chants Magnétiques" on Jan. 12th/85 in merely two hours in an atmosphere he described as "thoughts come tumbling in, unordered, and often without any connection with the composition on which you are working." One's intuition can lead one to ideas and feelings which are then formulated (expressed) in terms of the system one is using but it takes practise to learn how to work with a system and have it under ones fingers. To wake up in the morning after a night of dreaming and then proceed to express this in a certain medium demands of one technique and knowledge of using this technique to an artistic end. What follows are the composers attempts at expressing their world as opposed to yours or someone elses. Although they will have correspondances and relationships between them they are anything but standardized and as Renoir (the painter) stated "if something contains the personality of its maker, whether it is a spoon or a painting it has value but otherwise it is pointless and of no interest."

"Suggestions from Limbo"

"Limbo" was the first composition I produced making use of the MIDIM system and dates from 1981. It was an attempt to capture within the MIDIM language a dream atmosphere. The composition opens with the sounds of blackbirds deep in conversation. This is my own voice transformed reciting a poem. A new section then starts in which by means of a large M-modulation I was able to alter one of my basic predicators into a dog barking while superimposed upon this is an accompanying material composed of small D-compre-

Fig. 4. Repetitions for Ritournelles. Werner Kaegi and Dieuwke Aalbers.

hensors. This material was then subjected to a dephase operation which allowed me to create a cohesive material not openly repetitive. In the Bass register is a timbre which leans in the direction of a Contrabass. It is divided into two voices the one playing melismatic lines and the other occurring so low as to acquire a rasp and gives the bottom limit of the sound world presented to the listener. All of a sudden a new section appears which displays predicators suggesting eastern instruments. Again use was made of D-comprehensors. The reason for this is that it gives one the possibility of working in a highly associative manner by letting one assign freely the D-comprehensor patterns to whatever predicator desired. The composition ends as it began with the blackbirds.

The atmosphere of the piece results from the strange combination of timbres which occur together. Suggestions of instruments well known to the listener and animals as well as the superimposed contrasts of the material come together to try to create a strong associative environment. My own late-Stockhausean alchemical playground if you like.

"Wounded"

As one can ascertain from the title of this composition it deals with pain and makes the attempt to give this expression and form within the MIDIM language.

Material I: The predicator used for this material applies a moving formant as well as a DUR dependent BODY. The ratio "q" is 100/127. This value was found by testing various possibilities and seeing if they gave the proper audible result. The working process as is always the case with myself being primarily auditory, that is testing, reacting and then accepting or rejecting the material. Since the formant and timbre are dependent upon the pitch the subdivision applied will also determine the possibilities one has for various timbres. I chose a subdivision of 12 for this material but for one of the later materials I decided on a subdivision of 24 in order to arrive at a particular timbral effect. This has much to do with determining as well the atmosphere of the section. The descriptors for this predicator are derived from an initial descriptor to which operations are applied as follows:

(ret(inv(ret(Desc))))

This gives the stockpile of operations from out which I chose those to be used. The derived descriptors were then concatenated and computed. The result being then recorded directly onto tape from the computer.

Material II: The idea was to create a pedal note which slowly over the course of the composition ascends discretely (each time by an octave) over three full octaves and then abruptly falls at the end to its original position. This gives the piece a certain feeling of monotony corresponding to the monotonous feelings pain arouses when it never seems to alter or reside. The predicator used displays two fixed formants. Its most characteristic feature is a portamento which gives it a distinct quality of despair. In Kaegi, 1986 on p. 119 of this publication in par. 4.2, Fig. 4.1 one can find the portamento option which enabled me to express the emotion I intended.

Material III: This last group of predicators makes use of extremely long time durations for its segments and applies a portamento rate which gives a sort of howling effect which one critic found reminded him of an "animal" licking its wounds. This portamento occurs in the BODY of the predicator which being variable and the time duration being from .5 to 4 seconds in length give a sound composed like a wail.

The original composition was very dense consisting of 6 tracks. Not satisfied I reduced until I was left with a sort of skeleton, the basic structure without any superfluous items at all and then remade the piece on the basis of this to produce the final composition. One might compare the final structure to a "minimal description".

"La belle et la bête"

"La Belle et la Bête" is a collaboration between myself and Jos Janssen. It began as a commission from an Art Gallery in the Netherlands and became as we worked on it an experiment in tuning systems with Indonesian and Western elements interplaying throughout the composition (see Janssen/Kaegi). It opens with a melody taken by a "Woodwind"-like Predicator displaying as in the composition "Wounded" a Moving Formant. The melody is an Arabesque which revolves around a set of figures. We can classify this Predicator and Descriptor pattern into "La Belle" so to speak as it is quite delicate and refined in character. What follows is an eruption of "animal"-like sounds inundating the melody until it disappears in a cloud of rough, aggressive lines out of which emerges the melodic predicator but altered in its patterns to become syncopated and pained. The "animal" patterns display a healthy amount of noise or random modulation values and this is the first predicator transformed in this way. This "war" continues until the melody reasserts itself and emerges once more to end the first section.

A second section then begins with the Wind predicator and a Gender duplication playing in a duet as if improvising together. The florid Wind predicator and the high and bright Gender predicator continue happily until the Beast asserts itself once more and swallows them up. A further transformation of the "Wind-Animal" predicator occurs when it becomes a sort of "Zoom" sound which plays in quadrophonic space by itself. The Gender returns in short fragmented patterns and then the Beast is once more present as the layers multiply and the Gender repeats its patterns over and over until a rising scale on a second Gender ends the section with its final note disappearing away revealing the coda which is a short duet between a predicator in the direction of a female voice (sung) and the "animal" subdued and sad and so ends the composition.

The piece makes use of the facilities the MIDIM system offers for using different tunings (easily defined within the descriptor program) and applied the following:

1) Well-tempered tuning (with a subdivision of 12)
2) Pythagorean tuning
3) Indonesian tuning Slendro
4) Indonesian tuning Pelog

along with:

5) transformation of Predicators
6) contrasts of Indonesian and Western elements to create an atmosphere and mood.

Opposing elements are interplayed, contrasted and disturbed by one another to create the composition.

"Dorian"

The composition Dorian was composed by Ramon Gonzalez-Arroyo in 1985 and premiered on Jan. 14/86 at the muziekcentrum in Utrecht, the Netherlands. A description in the composers own words is the best way to present the piece to the reader.

"In Dorian life is the action of opposites. A world of sounds which would sound like a unit: the building up of an architectonic structure which when mapped into sound perception would be heard as one. What of the elements forming this world? Units in themselves, self-sufficient, individualistic and differentiated: the sculpting into the parameters of sound to obtain colourful elements with very distinct comportments. A flow. The clear perception of movement in time: the system going through the several states due to different circumstances caused by a number of force fields and at the same time the forcing of intersections: obliging the sounds to escape their own sonorous quality through interaction with the others. Shading their actual comportment and sound characteristics: hiding from perception "the real movement of the system".

A meticulous planning of every little point: the cross-relationships, the patterns, the structural games .... and the ruling in the actual production of "other" criteria which could alter that which by law was already fixed. Intermingled with all this was being occupied with a strong formal language – the MIDIM – which could apparently force you to act in a certain direction. Only by a deep study of it, by getting to know it and assimilating its characteristics was it possible to work in a dignified way with it. Again, the interaction between my own abstract ideas and images with the no less abstract laws and characteristics of the language in itself proved to be a very rich collision. One november night Dorian had cunningly gathered all his "out-of-control" force and showed me a side of himself very different from that which I had dreamt to be his. I was amazed, surprised, even frightened. What could I have done? I had to kill him so that it could be born and yet so many times he made me feel that once he had decided to show himself he knew he would remain free whether I desired it or not."

To conclude this section some compositions of Werner Kaegi shall be dealt with and as in the previous descriptions the composer himself can best speak about his work and attitude by means of the corresponding program notes.

"Composing in my view by means of a computer should marry the strength of the formal language with the spontaneity of my musical imagination. Exploring the expressive power of the formal language by means of my musical imagination will not be possible until the tool and composer operate as one and problem solving becomes a spontaneous activity."

"Consolations"

"Often it will be happen to me that I will note in my "diary" musical ideas which occur to me, a habit taught to me about thirty years ago by Arthur Honegger. Then I threw notes on paper, now I store them in digital MIDIM data files: predicators, whole libraries, comprehensors and so forth which can quickly be converted into sound and listened to. 'CONSOLATION' draws its material from data I stored over the period 1978 to 1981. The formation of the material took place in 81/82. The piece consists of three parts:

'In Memoriam' is written in memory of my wife Regine Brandt. It is a game of associations obtained by various applications of two comprehensors to a small library. The latter including among others predicators acquired from my investigations into string sounds.

'Automne' displays a leading melodic phrase which is repeated during the whole piece as in a chaconne. Around this a game of melodic lines of different lengths develop which sound together in constantly new phase relationships. The piece is an expression of a static image of an emotional state which might be expressed by the following Japanese poem (Kakuren, 1202):

The cries of the insects
are buried at the roots of
the sparse pampas grass -
the end of autumn is in
the colour of the last leaves.

'Vers D'Autres Jeux' (Toward Other Games) is based upon a poem by the French writer and inventor of the phonograph Charles Cros. The speech and singing displayed are entirely synthetic. The first part exhibits an exponential development of speed and pitch, resulting in a kind of continuously modulating harmony. The epilogue of this very sensual piece quotes among others the composition JEUX of Debussy."

Sensation de Haschisch (Charles Cros)

Tiède et blanc était le sein.
Toute blanche était la chatte.
Le sein soulevait la chatte.
La chatte griffait le sein.

Les oreilles de la chatte
Faisaient ombre sur le sein,
Rose était le bout du sein,
Comme le nez de la chatte.

Un signe noir sur le sein
Intrigua longtemps, la chatte
Puis, vers d'autres jeux, la chatte
Courut, laissant nu le sein.

"Ritournelles I and II for soprano and computer"

"The human voice has always excited me. In my 16th year under the huge impression of the great hungarian singer Ilona Durigo I began to write songs which in the course of time were followed by further compositions for song (1959/61: "Miracles" for soprano and 8 instruments, 1968: "Voce magna and Dominum clamo" for mezzo-soprano and large orchestra which was premiered by Bruno Maderna, etc.). In 1967 I made the attempt to connect the singing voice with electronic sounds; my music "Les Vêtements de la demoiselle" for soprano and tape attained the Prix Suisse. In the same year I proposed the idea of basing the organization of the electronic sound world upon the characteristics of the human voice (Was ist elektronische Musik, Zurich, 1967), an idea that I later in my VOSIM-model for sound synthesis (VOice SIMulation) and the MIDIM language was able to make concrete.

The two RITOURNELLES consist of strophes from where the title arises. The eleven strophes of the first piece are developed from a single basic pattern in a micro-tonal manner. This could absolutely not have been realized without the aid of a computer. All the predicators of the P-library applied are strongly characterized by periodical M-modulation; the low and high sounds are in addition shaped by intonation patterns. The second piece consists of three strophes and explores thoroughly sound material first applied in "Chants Magnétiques". The two pieces will be followed by a third piece which I am on the way to compose for the International Computer Music Conference in the Hague. The final version of Ritournelles will be a piece for soprano, mixed Choir and Computer in which the three soprano solos will sparkle".

"Chants Magnétiques" (hommage à A. Breton et Ph. Soupault)

"Thoughts come tumbling in, unordered and often without any connection with the composition on which you are working. This was also the case during the time I was composing my Ritournelles for Soprano and computer. A flood of material and studies came into existence which could not be fit into my Ritournelles. From out this material resulted my piece "Chants Magnétiques" on January 20th, 85 in two hours by writing a command string at a high tempo and reading it into the MIDIM system. The piece was immediately recorded on 4-track tape and no corrections were later applied. It had its premiere on January 22nd, 85 in Utrecht" (see Appendix II for list of MIDIM compositions).

CONCLUSIONS

The MIDIM system has up to now been used by many different composers from various lands to create compositions and interest is growing in the system as well as in the activities revolving around it. A group has been formed which puts on concerts and gives lectures and demonstrations about the system. The name of the group is called the "MIDIM Group" and is situated in Utrecht, the Netherlands.

ACKNOWLEDGEMENTS

The author wishes to thank Werner Kaegi, Heinerich Kaegi, Jos Janssen and everyone connected with the activities described in this article, in particular the students of the practicum with whom it was a great pleasure to work.

APPENDIX I

SOME CONCERTS 1982-86 (the numbers below the city/country column indicate the index of the compositions to be found in Appendix II of this article.)

Espace 300	Annecy, France (1,12b,5)	06-03-82
Geertekerk*	Utrecht, the Netherlands (8,14,13,12a,b,c)	11-10-82
Stedelijk Museum	Amsterdam, the Netherlands (8,14,13,12a,b,c)	14-05-83
I.G.N.M*	Zurich, Switzerland (8,14,3,4,12a,b,c)	30-05-83
Festival Bourges	Bourges, France (12a,b,c)	12-06-83
Dansk Elektronmusik Selskab	Copenhagen, Denmark (12a,b,c)	14-12-83
Muziekcentrum Vredenburg*	Utrecht, the Netherlands (12a,17,18,19,21,20)	24-01-84
Stedelijk Museum	Amsterdam, the Netherlands (12b,17,18,19,23)	03-03-84
Era Studio Gabriele de Agostini	Geneva, Switzerland (12a,b,c,8,22)	13-04-84
Festival Bourges	Bourges, France (22)	12-06-84
Festival 84'	Geneva, Switzerland (23)	12-09-84
Radio Suisse Romande	Geneva, Switzerland (23)	12-09-84
Gele Rijder*	Arnhem, the Netherlands (25)	18-10-84
WDR, Musik Der Zeit I, 30 Jahre Elektronische Musik	Cologne, Germany (12a)	26-10-84
Glockenhaus	Luneberg, Germany (19,12a,22,16)	02-12-84
Muziekcentrum Vredenburg*	Utrecht, the Netherlands (27,28,30,24,25,26)	22-01-85
Festival Bourges	Bourges, France (30)	10-06-85

Stadsschouwburg*	Arnhem, the Netherlands (29,28,30,24,25,26)	12-06-85
Swiss Computer Music Centre	Zurich, Switzerland (26,30)	13-06-85
Television	Zurich, Switzerland (26)	13-06-85
Musica Electroacustica 85'	Buenos Aires, Argentinia (30)	11-09-85
Lantaren, Venster*	Rotterdam, the Netherlands (17,18,29,28,33,25,26)	17-11-85
A.S.K.I.-Academy*	Solo, Java (12c,24,25,23)	27-12-85
Muziekcentrum Vredenberg*	Utrecht, the Netherlands (17,12c,31,35,34,26,32)	14-01-86
Paradiso	Amsterdam, the Netherlands (31)	17-01-86
Levis Art Center	Laval, Canada (8,34,12a)	86
Gele Rijder*	Arnhem, the Netherlands (31)	04-03-86
Berlin	Berlin, Germany (30)	86
Festival Bourges	Bourges, France (26,32)	09-06-86

*Concerts organized by the MIDIM group

APPENDIX II

- MIDIM COMPOSITIONS -

1)	1979	K. Samkopf	Etude Nr.1	15'09"	Stereo
2)	1979	F. J. Sacci	Passage	9'04"	Stereo
3)	1980	F. J. Sacci	Dance 1	4'39"	Stereo
4)	1980	F. J. Sacci	Dance 2	5'10"	Stereo
5)	1980	C. Fatus	Voyager	8'45"	4-track
6)	1980	R. Temmingh	Blomsit 2 (organ/tape)	23'00"	Mono
7)	1980	R. Temmingh	Blomsit (vers 2) (organ/tape)	15'00"	Mono
8)	1981	P. Goodman	Suggestions From Limbo	7'00"	4-track
9)	1981	C. Fatus	Pieces pour Guitar (guitar/tape)	16'42"	Stereo
10)	1982	S. Dydo	Stelling 1	5'15"	Stereo
11)	1982	S. Dydo	Stelling 2	8'45"	Stereo
12)	1982	W. Kaegi	Consolations:		4-track
	a)		- In memoriam	12'00"	
	b)		- Automne	17'00"	
	c)		- Vers d'autres Jeux	16'00"	
13)	1982	C. Ore	Circe	16'00"	4-track

14)	1982	C. Fatus	Voyager 2	10'00"	4-track
15)	1983	C. Fatus	Recht is...	60'00"	4-track
16)	1983	C. Fatus	Turbulences	11'20"	4-track
17)	1983	P. v. Berkel	Tjaktjakai (voice/tape)	5'00"	4-track
18)	1983	P. v. Berkel	Heegothee	7'00"	4-track
19)	1983	P. Goodman	Wounded	8'55"	4-track
20)	1983	W. Kaegi	Dialogue for Kendhang solo computer and gamelan	12'00"	4-track
21)	1983	Ki Nartosabdo/ Jos Janssen	Subakastawa (gamelan/tape)	8'00"	4-track
22)	1984	W. Kaegi	Dialogue for Kendhang playback and computer	12'00"	4-track
23)	1984	W. Kaegi	Dialogue for Kendhang solo and computer	12'00"	4-track
24	1984	P. Goodman	Episodes (recitation/tape)	13'00"	4-track
25)	1984	P. Goodman/ J. Janssen	La Belle et La Bête	9'49"	4-track
26)	1984	W. Kaegi	Ritournelles I for Soprano and computer	9'30"	4-track
27)	1984	Group	Impromptu 1	5'00"	4-track
28)	1984	Group	Impromptu 2	8'45"	4-track
29)	1984	Group	Impromptu 3	7'45"	4-track
30)	1985	W. Kaegi	Chants magnétiques (Hommage à A. Breton et P. Soupault)	7'00"	4-track
31)	1985	Courbois/Goodman/ Janssen	Independence (Jazzdrummer/and tape)	6'00"	4-track
32)	1985	W. Kaegi	Ritournelles II for Soprano and computer	3'00"	4-track
33)	1985	A. Cosijn	Qui Vive	5'00"	Stereo
34)	1985	R. Gonzalez-Arroyo	Dorian	10'25"	4-track
35)	1985	B. Guttman	Different Attitudes	10'00"	4-track

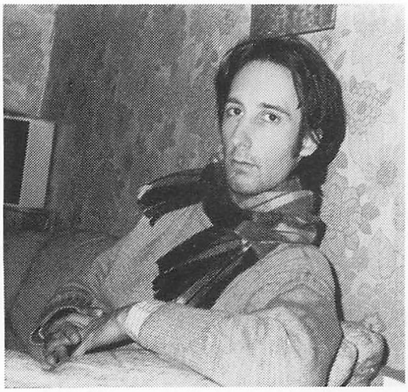
- BIBLIOGRAPHY -

1. Berkel, P. van – Notes by a Film-maker. Interface, Vol. 15, pp. 231-257.
2. Breton, A./Soupault, P. – Champs Magnétiques, Gallimard, 1919
3. Janssen, J./Kaegi, H. – Midim-duplication of a Central-Javanese Sound Concept. Interface, Vol. 15, pp. 185-229.
4. Kaegi, W. – MIDIM8X System Descriptor Users Manual. Last Edit: May, 1984. Institute for Sonology, Utrecht University.
5. Kaegi, W. – The Design of Musical Discovery. A Selection, revised version, Nov. 1985. Institute for Sonology, Utrecht University.
6. Kaegi, W. – Concert Program, Utrecht, Geertekerk. Oct. 11, 1982.
7. Kaegi, W. – Controlling the Vosim Sound Synthesis System. Interface, Vol. 15, pp. 71-82.
8. Kaegi, W. – The Midim Language. Interface, Vol. 15, pp. 83-162.

Recommended

1. Behar, H./Carasson, M. – Le Surréalisme, texte et débats. Le Livre de Poche, 1984.

2. Breton, A. – Entretiens. Idées/Gallimard, 1973.
3. Breton, A. – Manifestes du Surrealisme. Idées/Gallimards, 1981.
4. Carnap, R. – Der Logische Aufbau der Welt. Im Weltkreis-Verlag, Berlin-Schlachtensee, 1928.
5. Carnap, R. – Meaning and Necessity. University of Chicago Press, 1970.
6. Kaegi, W. – Was ist Elektronische Musik. Orell Fussli Verlag, Zurich, 1967.



Paul Goodman: was born in Vancouver, Canada on March 13/1955. After studying music privately he took his degree in theory from the Royal Conservatory of Toronto and left to pursue his education in Holland at the Institute for Sonology at the University of Utrecht where he worked as student (1977/79), composer (project worker, 1981/86) and substitute lecturer. As assistant to Dr. Werner Kaegi (1983/85) he gave with Jos Janssen a practical course in the use of the MIDIM-system for two years at the Institute for Sonology and is now a member of the MIDIM-Group for Associative Computer Music which is busy lecturing and giving concerts throughout the Netherlands. He has had many concerts in and out of Holland, the most notable being at the Stedelijk Museum in Amsterdam; Geneva and Zurich (Festival for Contemporary Music); Luneburg, Germany; Solo, Java (at the first computer music concert ever given in the country).

Interface, Vol. 15 (1986), pp. 185-229. 0303-3902/86/1502-0185 $3.00

MIDIM-Duplication of a Central-Javanese Sound Concept

Jos Janssen and Heinerich Kaegi

ABSTRACT

One of the many practical applications of the MIDIM-system in the last few years has been *the duplication of existing sound families*, in particular those of musical and speech sounds. In the first part of this article we will define the concept of a sound duplication and show a method by which means duplications may be found. This method is particularly suitable for the MIDIM/VOSIM-system. In the second part an application of this method is shown for the instrumental concept *Gender* drawn from the *Central-javanese Karawitan.* We shall also concern ourselves with a number of other concepts from this music culture.

PART I - MIDIM SOUND DUPLICATIONS

Apart from continual investigations into western instruments a few members of the MIDIM-group have devoted themselves to sound duplications of non-western instruments[0]), in particular Javanese instruments. In our article we shall discuss this work. In part I what exactly sound duplication means is formally described and a method shall be shown by which MIDIM-duplications may be found.

1 WHAT IS A SOUND DUPLICATION?

A sound duplication is a bundle of rules defining a set of synthetically generatable sound events, which are associated with an existing sound concept.[1]) Each sound event of the duplication corresponds ideally to a sound event which belongs to this existing concept. Duplications are formulated by means of a

[0]) Already in 1978 K. Samkopf did investigations on Chinese tempelblocks which he used in his composition Etude nr. 1 (1979); in 1983/84 Dr. Kaegi analyzed the alphabet of the Central-javanese Kendhang Ciblon (Jav. drum) which was the starting point for his composition Dialogue for Kendhang and computer (1984); in 1983/84 Jos Janssen did an experiment using his Gender duplications in the context of a classical Javanese composition (Subakastawa Pl.6 for Gamelan Gadhon and computer). See also composition list pag. 182. (All documentation concerning the duplications are available in the archives of the MIDIM-group).

[1]) The duplication of sounds does not have to be limited to natural sounds. It would be quite possible to make a MIDIM-duplication of the coloratura of the KÖNIGIN DER NACHT, synthesized by means of the sound synthesis system CHANT of IRCAM in Paris (see Rodet, 1984).

formal language (to which soft- and hard-ware designs belong), in particular formal languages which are suitable for sound-synthesis, such as in our case the MIDIM-language.

A sound duplication in our sense describes thus *collections of sounds* (W. Kaegi calls this concept-duplication) *and should not be confused with the duplication of unrelated sound events* (event-duplication).[2])

2 WHY MAKE MIDIM SOUND DUPLICATIONS?

Before giving a formal definition of what a MIDIM-duplication is, we shall mention a few important and meaningful scientific and artistic applications of it:

1. Duplications show the expressive power of the MIDIM-system, because the truth or falsity of our duplicated sounds is easily tested by comparing them with the originals. (Kaegi, 1986, p. 101) On the other side via duplications new functions (which have an interpretation in our music culture) may be found in order to expand the existing function table.

2. Every MIDIM-duplication means a formalization of the desired concept (see part II) thus an increase in our knowledge of this concept.

3. The formalization of existing sound concepts (which we have already mentioned) makes it possible to compare these concepts in a formal way. (Think about the comparison of musical and speech sound concepts stemming from one and the same culture. See Kaegi, 1986, p. 130 and the composition list: DIALOGUE, Goodman, 1986, p. 169).

4. A musician can use duplications as a starting point in order to derive sound families. (In MIDIM-compositions for computer and soloist this procedure is often applied, for example in CONSOLATIONS and RITOURNELLES of W. Kaegi. See composition list, Goodman, 1986, p. 182).

[2]) That from the beginning we concerned ourselves with sound concepts (signs) and not with instances (in other words with signals), was already apparent in "Das Problem der mentalen..., Kaegi, 1975" from which we quote the following:
"Es wurde vielmehr von Beginn an unterstellt, dass nicht von Signalen bezw. Signalfunktionen, sondern von Zeichen auszugehen und das Problem in umgekehrter Richtung anzugehen sei. Nämlich durch den Versuch, von musikalischen Zeichen auszugehen und zu einer metrischen Beschreibung der Zeichen in einem n-dimensionalen physikalischen Raum vorzustossen". (See also Kaegi, MIDIM, 1986, p. 131).

5. Users of the system learn via the construction of duplications to listen to and to abstract elements from the sounds they hear and then to translate these elements into the MIDIM-language (See Goodman, 1986). This opens the way to the formalization of imaginary sounds (e.g. in artistic work).

6. Duplications make possible the development and testing of pattern recognition algorithms, which fit with the MIDIM-functions.

In what follows below we shall show by which methods duplications may be found.

3 DEFINITIONS AND ASSUMPTIONS

We introduce first the set MC which is *the Music Culture under consideration* and which contains on the one hand *the set* U (universe of discourse) of *all musical sound events of* MC and on the other hand all *expressions, concepts, definitions* etc. ("the context"), that are related to the sounds. K is a particular sound concept within the music culture MC. We call this concept: *the descriptive concept* (Kaegi, 86, p. 101).

The *extension* of K, which we call K_e is *the set of all sound events or instances* i_K, which belong to the descriptive concept K. K_e is a subset of U.
Thus $i_K \in K_e \quad K_e \subset U$.

In analogy to this, in the MIDIM-language the *predicator P is a MIDIM sound concept,* which we call the M-concept. The *extension* of P, which we call P_e is *the set of MIDIM sound events or instances* $i_P \in P_e$, selected from out the VOSIM-matrix V*. The instances are thus (sequences of) VOSIM-vectors.
Thus $i_p \in P_e \quad P_e \subset V^*$.[3])

The *intension* of P, which we call P_i is *a bundle of MIDIM-functions* (taken from out the function tables and substituted into a MIDIM-vector $\vec{m}_0$, see Kaegi, p. 93), whose *course of values is the extension* P_e (Carnap, 1947).

Descriptive concept K (predicate)	⟨	INTENSION: properties
		EXTENSION: K_e, which is the set of all sound events i_K, which satisfy the properties.

[3]) We tacitly assume here that V* contains all sound concepts from U, in other words the music culture MC falls within the "field of application" of the MIDIM-language. This implies at the same time that the music culture MC belongs to the Indo-european music culture. See Kaegi, 1986. p. 101.

MIDIM sound concept P (predicator)

- INTENSION: P_i, which is a bundle of MIDIM-functions (substituted into a $\vec{m}_0$-vector).
- EXTENSION: P_e, which is the set of all MIDIM sound events i_p (sequences of VOSIM-vectors), satisfying these functions.

The above scheme is now given for a Böhm clarinet (b-flat; played in the french manner):

Descriptive Concept: Böhm clarinet

- INTENSION: property to be a Böhm clarinet sound.
- EXTENSION: the set of all sounds produced when performing on a Böhm clarinet.

MIDIM-concept: Böhm clarinet predicator

- INTENSION: The bundle of MIDIM-functions belonging to this predicator.
- EXTENSION: The set of all VOSIM-sounds supposed to be accepted by listeners as Böhm clarinet sounds.

The MIDIM-formalization of this descriptive concept is shown by Kaegi (1986) p. 126/7.

We now call P_i *a duplication of* K *if and only if within all pairs* (i_p, i_K) *of corresponding instances,* i_P *and* i_K *cannot be distinguished in sound.*
A formal (e.g. MIDIM-)concept is thus accepted as being a duplication *by means of experimentation.* The design of the appropriate experiments will be discussed in par. 6.7.

4 FROM AN INTENSION TO AN EXTENSION AND VICE VERSA

In the MIDIM-language a predicator is an expression where at most the prosodic parameters T′, DUR, and At are λ-tied variables. The transition from an intension P_i to a particular sound event, or instance i_p in the MIDIM-language takes place in the formulization of a descriptor and λ-elimination of the prosodic parameters T′, DUR, and At. P_e can now be calculated by a repetition of this procedure, whereby the prosodic parameters become varied over their domains. One can thus say that the MIDIM-language allocates to the intension P_i the extension P_e.

MIDIM-language: $P_i \rightarrow P_e$

As is known from logic (Carnap, 1947) *every intension uniquely determines an extension. It is impossible to move in the opposite direction namely from an extension to its primitive intension without meta-information.* If we know for example only a set of VOSIM-vectors, *without any further information* about the

MIDIM-language, then it is impossible to find the bundle of MIDIM-functions which generated these VOSIM-vectors. Information about the MIDIM-language and its functions is required in order to rediscover the intension.[4])

$$P_e + \text{meta-information} \rightarrow P_i$$

This is one of the most important problems of pattern recognition. The discovery of a pattern from a signal always calls for knowledge (meta-information) concerning the pattern sought. (Banerij, 1969).

5 SEEKING A DUPLICATION THEORETICALLY

Even though the instances i_p of a duplication sound like the corresponding instances i_K of the descriptive concept K, they are actually *not* in most of the cases *physically identical,* i.e. they have not the same signal function.[5])

Thus $i_K \neq i_p$ (physically)

Thus seeking a duplicate necessitates not only the step from the extension to the intension of the M-concept $P_e \rightarrow P_i$ but also the steps from the descriptive extension to the extension of the M-concept $K_e \rightarrow P_e$. The signal functions of the descriptive extension K_e could first be reduced to VOSIM-vectors (with VOSIM-signal functions as an intermediate stage) from which we should have

[4]) Experiments have been performed in order to calculate from an arbitrary sequence of VOSIM-vectors the corresponding MIDIM-vectors. In most of the cases this appeaeed to be impossible. We note here that the set of all possible MIDIM-functions was known in the form of meta-information.

When mapping MIDIM-vectors onto VOSIM-vectors one actually changes languages; so we may call the meta-information which is necessary in order to proceed backwards from the VOSIM to the MIDIM-language "extra-linguistic information within the VOSIM-language".

[5]) For simplicity the differences existing between the concepts "VOSIM-vector", "signal" of a sound, "perception" and "cognitive mapping of a sound" have been ignored. That these differences are in principle quite important is shown by the fact that for auditors instances may sound identical, while not having the same signal functions. A natural signal contains much "non-relevant information" or redundancy, which, when eliminated, does not effect the auditory experience. Kaegi has shown this in "A new approach to ..., 1972". The idea of a minimum description is based upon this. (MIDIM = MInimal DescrIption of Music).

to determine (with the help of the necessary meta-information) the appropriate bundle of M-functions P_i.[6])

The step $K_e \rightarrow P_e$ may in physics be summarized as the mapping of empirical data onto the most appropriate function. This is called *curve fitting* (see par. 6.6). Meta-information is once more needed for this transition. (Imagine for example statistical algorithms of abstractors, selecting the data which is needed for the approximations.)

Curve fitting: K_e + meta-information $\rightarrow P_e$

In practise one does not in most cases know *all* elements of K_e but simply a selection. For this reason it is necessary to interpolate between the instances in order to fill in those missing. Also the choice of the type of interpolation to be used depends upon the meta-information. (For example: if we know simply two points it is possible to connect them with many types of different functions. Which function we should choose finally depends on the meta-information.)

By preference the step from the descriptive extension to the intension of the M-concept $K_e \rightarrow P_i$ is not performed via P_e but by constructing *a spectral space* R_s *within which the descriptive instances* i_K *are fit to the MIDIM-functions.* We obtain then the following analysis-chain:

Analysis: $K_e \rightarrow R_s \rightarrow P_i$

In the following paragraphs these analysis-chains are dealt with in detail. Part II of this article shows an example of a duplication of the descriptive concept *GENDER* (metallophone from out the *Gamelan)*[7]) which belongs to the music culture (MC) of *Central-Java.* In order to prepare for this we shall have a brief glance at this concept in par. 6.

6 SEEKING A MIDIM-DUPLICATION IN PRACTISE

The above logical foundation of the concept-duplication could be formulated in *any* powerful sound-synthesis language. We will limit ourselves to MIDIM-terminology, because on the one hand the language has proven to be extremely powerful and on the other hand because a large number of duplications have been realized with the MIDIM-system. Moreover certain parts of the analysis-chains were formalized and effectivized while others were as yet still performed by hand. How we proceeded in practise is shown in Fig. 1.

[6]) There have been made attempts to map an arbitrary signal into a VOSIM-signal but simple algorithms for this could not be found.

[7]) A 'GAMELAN' is a collection of bronze percussion instruments and diverse aerophones, cordophones and membraphones.

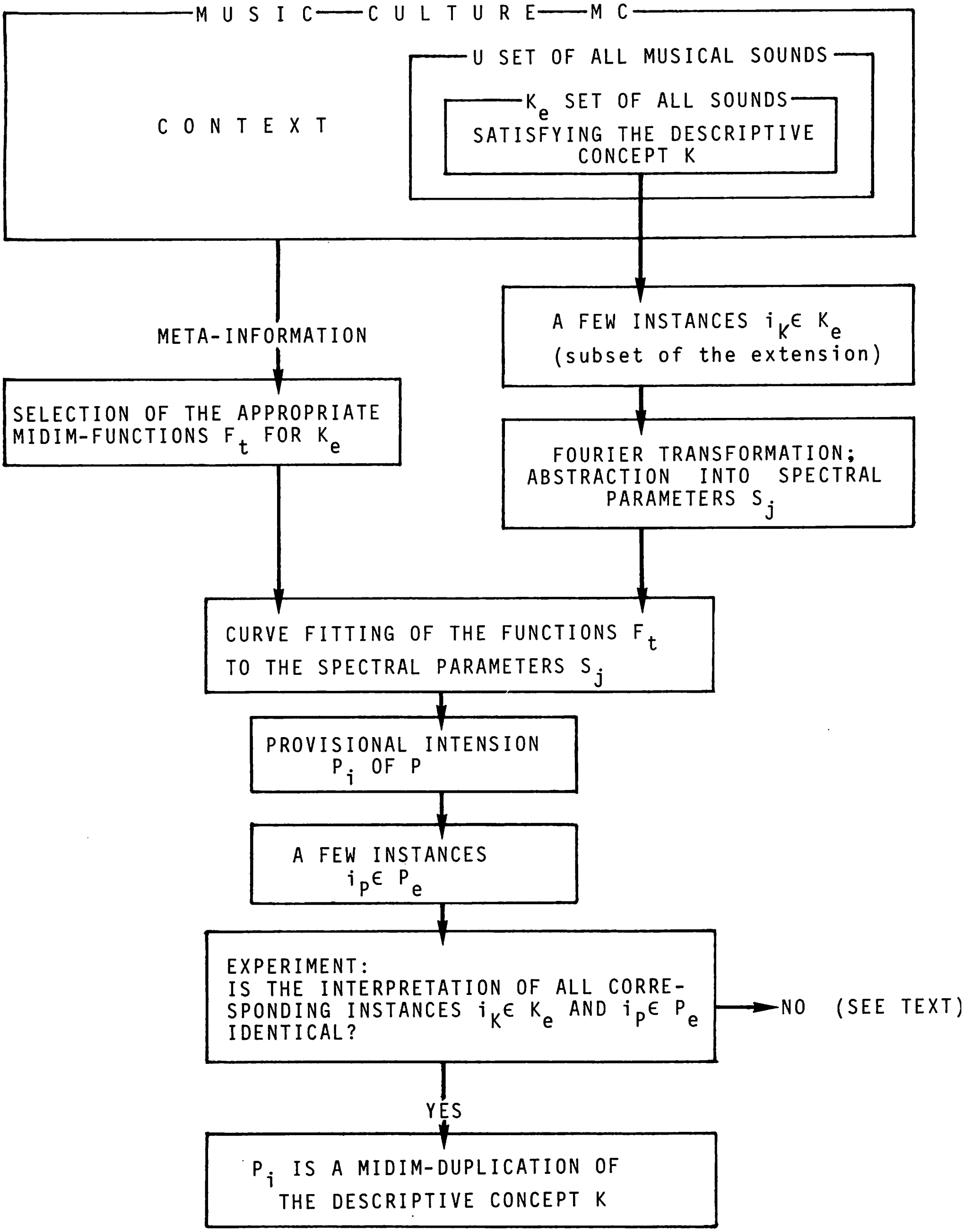


Figure 1. Scheme for determination of a MIDIM-duplication.

6.1 Concept definition and acquiring knowledge

The first step towards a duplication is the definition of the desired descriptive concept K. For this *knowledge concerning concept K* and *the context* is needed; knowledge in the sense that it is possible *to judge whether or not any arbitrary element of U belongs to* K_e. Our concept must thus have a *clear* and *testable* meaning within the music culture (MC) to which it belongs. Let us take as an example of this the concept G: "Gender sound" which belongs to the Central-javanese Karawitan. We are *not* free to decide arbitrarily which sounds (instances) belong to this concept, because it has a very specific meaning deeply anchored in the Central-javanese music culture. We must ask Javanese musicians which sounds are to be considered Gender sounds. The clearest answer are the sounds themselves: the masters of Gender performance can show us by playing their instrument what a Gender sound is and then we become aware that there is an *unbreakable marriage between instrument, performer and the music culture* which has given birth to them (something which is often forgotten. This is also the reason we use the word *Karawitan* in this article in place of the term *Gamelan* which is more familiar in the west. Gamelan indicates only the instruments, while Karawitan designates the whole music culture.). It appears as well that musical concepts are sharply determined, although it is not always possible to describe them in colloquial language. *The musicians are able to show precisely what does and what does not belong to a concept.*[8]) See appendix I.

The choice of sound concepts for duplication which are readily testable is of the essence. On the one hand the duplication is made clearly testable by this (this does not for example hold for imaginary sounds), while on the other hand in this way *the functions used for the description are given a meaning within the music culture and are thus not anymore foundated on a purely physical basis.* Only when these conditions have been satisfied for a large number of very different families of duplications may we assert that our sound-synthesis language can give a formal description of the sound world in a music culture.

6.2 Registration of instances

Proceeding from a clearly testable descriptive concept K, we collect a number of instances $i_K \in K_e$, which together form the subset K_G of K_e. These *registrated instances* are stored on tape (or directly in the computer). For a registration it is essential that one after the other of the prosodic parameters T′, DUR, or At vary, while the other two remain constant. The instances can then be divided over the three sub-sets $K_G(T')$, $K_G(DUR)$, $K_G(At)$.[9]) (Holding for the MIDIM8X standard predicator; see Kaegi, 1986, MIDIM, p. 126).

[8]) The most sharply testable sound concepts are to be found in natural languages, because they are based entirely upon common sense (within a culture). Whats more, everyone can speak but not everyone can make music.

[9]) We assume here that the prosodic variables are not correlated. In practise one chooses configurations which have the most influence upon the sound.

6.3 Fourier transform

First fourier transform is applied to the registrations by means of effective algorithms performed by a computer and *careful listening.* (Kaegi calls this his "personal fourier analyzer")[10]) Fourier transform *does not change the signal information but only the representation of this information.* Application of fourier transform is based upon the assumption that the ear performs fourier transformation and that fourier representation is thus better able to reproduce a sound experience. In *fast fourier transform* (FFT) the signal is divided into small time segments and the spectrum is calculated per segment. Next the calculated data is mapped into a 3-dimensional space spent up by the axes time, frequency and amplitude of the coefficients, which is called *the running spectrum.*[11]) A clear representation of this as well as the interpretation which follows obliges us to apply *data reduction by abstracting certain tendencies by means of abstractors.* The so-called *peak-tracking abstractor* which we often applied connects the amplitude maxima (of the fourier coefficients) in the 3-dimensional space described in slices of either *constant time,* or *constant frequency* or eventually *a constant frequency band.* The peak tracks which result can be clearly plotted in a pseudo-3-dimensional representation (which is a projection of the 3-dimensional spectral space on an appropriate plane). The plots are abstracts. For convenience we call them "the spectra".

6.4 Division into MIDIM-segments

The MIDIM-language proceeds from *four segments* S_i in time. We seek thus in the spectral information an appropriate time segmentation. Therefore each peak-track abstracted has to be divided over the four desired segments. For this reason within a particular time segment a "constant deportment"[12]) in the peak-tracks is sought as well as in the amplitude envelope of the whole signal. From all segments the final MIDIM-segmentation is chosen and the time-durations of the segments S_i determined.[13]) A good example of a possible segmentation method is described by van Berkel, p. 241.

[10]) One can scan the spectrum by listening to sounds sent through filters, whereby the transmitted amplitude is read upon a dB-meter. W. Kaegi calls this "the poor persons fourier analysis". It should be noted that scientists tend to turn their noses up at this method as they consider it unscientific. They have the tendancy to prefer to represent their results pictorially.

[11]) The sampling rate for a digital signal representation determines the highest analyzable frequency, the "window width", represented in time the lowest frequency.

[12]) With constant deportment per segment we mean that the signal can be described (approximately) by a MIDIM-function of which the coefficients are not yet determined. For example: the constant behavior could be that the amplitude progresses linearly, so it can be described by a linear equation. The gradient is actually not yet known and is not determined until a statistical approximation takes place.

[13]) There are various methods for segmenting the signal. Strong changements in the envelope contour of the whole signal or of the peak-tracks (and eventually in the frequency changements of the peak-tracks) can be discovered by determining the derivative. The maxima in the derivative then give possible break-points.

Comparing the instances taken from the subset $K_G(DUR)$ (in which DUR varies while At and T′ remain constant) shows us which segment S is dependent upon DUR (one calls this segment A) and which value v in the M-concept C′(v) is thus applicable (Kaegi 1986, p. 116). From this point per segment S_i the whole analysis is performed.

6.5 Representation of the instances in the spectral space R_S

For the following steps we introduce the so-called spectral parameters s_j with $j = 1$ to n_j. We assume that n_j parameters (over one segment) are capable of characterizing a peak-track. For the description of the peak-tracks (given by the fourier transform and abstraction) by means of these parameters data-reduction (averaging) is applied which is dependent upon the type of spectral parameters which we have at our disposal. Although one can, if desired, introduce other spectral parameters, we make use of s_1 for the frequency f, s_2 for the amplitude a, s_3 for the amplitude changement Δa and assume a linear interpolation between a and $a + \Delta a$. (If the peak-track would shift in frequency as well over the segment then one would need to introduce a parameter Δf). In Fig. 2 a theoretical peak-track is shown. The original peak-track, from which the track shown is derived is left out for clarity.

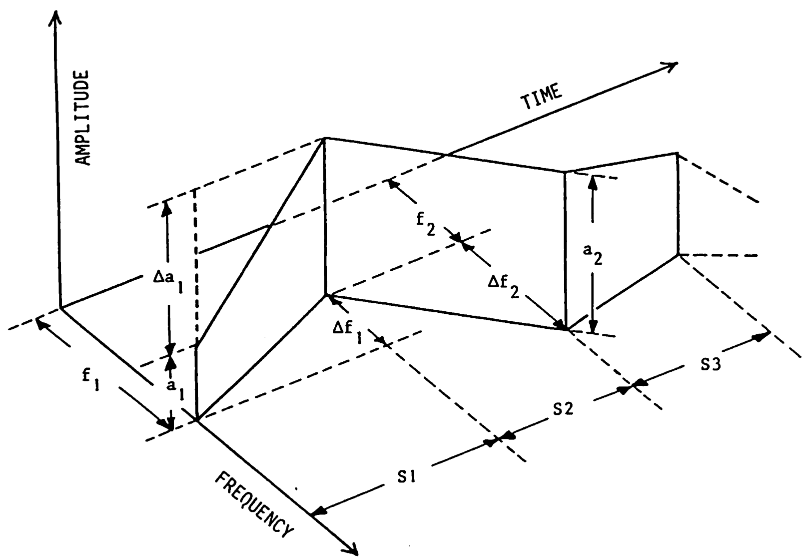


Figure 2. A theoretical peak-track (over three segments) described by means of the spectral parameters f, Δf, a, Δa in a 3-dimensional spectral representation.

Since by application of fourier transform and abstraction for each instance $(i_K)_m \in K_G$ (we index the instances with m) a series of peak-tracks (per segment) is produced, we can now assign to each instance various series of spectral parameters, namely:

$$(s_1, \ldots, s_{n_j})_h$$

The index h shows to which peak-track the parameters belong and runs from h=1 until n_h. (There are thus exactly n_h peak-tracks abstracted). Each instance $(i_K)_m$ can, as is known, be characterized by means of three constants $(c_{T'}, c_{DUR}, c_{At})_m$ which stand for the prosodic parameters T', DUR and At.[14]) We now construct a linear space R_s ("spectral space"), by means of the prosodic parameters and the spectral parameters, namely, T', DUR, At, $s_1,\ldots,s_{n_j}$ (dimension $n_j + 3$). With the spectral parameters which we have described and used f, a, Δa, our linear space R_s contains the coordinates (T', DUR, At, f, a, Δa). To each instance within this space there can be assigned n_h points. All instances from out K_G together thus form in R_s *a point-cloud.*

If we look now not from the standpoint of the registered instance K_G but from the standpoint of the MIDIM-language into the space R_s, than we see that to each instance $i_p \in P_e$ there *also correspond points* (again per segment). *Each MIDIM-function or product of functions (intension) from the function tables* (Kaegi, 1986, p. 94) *describes a set of VOSIM-vectors (extension), which one can conceive as a point-cloud in* R_s. *Finding a duplication thus comes close to finding the intension of a MIDIM point-cloud which fits best with the point-cloud of the instance* $i_K \in K_G$. This procedure has been called earlier by us *curve fitting.*

[14]) This assertion assumes that the prosodic parameters are already known in an extensional form. How do we know actually what pitch corresponds to an arbitrary instance? There are techniques for determining the fundamental via spectral analysis (F0-tracking) but in most cases this is not necessary because we have information taken from the context. If we record a piano, for example, then the pianist can tell us exactly which note was played, imagine the chamber tone a', then we know according to tuning standards that the fundamental equals 440 Hz, from which $T'=1/F_1$ can be directly calculated. The investigation of non-western sounds is more problematic. The tunings are often not a priori (extensionally) known or are even determined within an area, instrument or orchestra. In anticipation of this problem Jos Janssen began by measuring first (at the start of his investigations into Javanese sounds) the frequency values corresponding to the various Javanese tunings and instruments, which served as the reference standards. This shall be shown in Part II, par. 2. From the contextual givens, namely the pitch name within Javanese culture (and the name of the tuning and instrument which are used as references) T' can be directly determined. The functions ϕ_3 and ϕ_1 in the MIDIM-system, resolve this problem for western music because they connect respectively the pitch/octave names and subdivision with the pitch-frequency and the pitch-duration (in fractions) and metronome with the time-duration. The parameter At gives few problems because it is easily measurable (dB-meter) and is not of importance until we wish to study the relative amplitude differences between various pitches.

6.6 Curve fitting

Fitting the MIDIM point-cloud to the point-cloud of the instances i_K must result in the following concrete information concerning:

a) The indexes t of the MIDIM-functions F_t (taken from the function tables) which are appropriate for the descriptive concept K.
Through the indexes we find the λ-tied variables and their actual domains.[15])

b) The numerical values of all parameters (within the above functions) which are as yet not eliminated (with the exception of the prosodic parameters).

In order to clarify this we will give an example: the point-cloud which belongs to our Karawitan concept "Gender sound" appears to be best approximated by means of the function F_1: $T = q(T' + Of)/N$ for the reason that T is linearly dependent upon T'. In step a) our analysis system gives us thus the function index $t = 1$. Still unknown are the particular constants for q, Of and N (and all VOSIM-variables tied by functions not discussed here). In step b) the system must give us the values for q, Of and N which fit best. (We shall return to this in Part II).

The information desired cannot be abstracted from the signal information without meta-information (as shown in paragraph 5). Even if we know the basic set of functions (the function tables of the MIDIM-language) in the form of meta-information discovering the specific functions (point a) is in practise extremely involved. In contrast to this if the function rules of suitable fitting functions are known by means of meta-information then the procedure becomes drastically simplified and one can make use of the conventional approximation techniques.[16]) In this case as well a rather large number of instances is necessary (as an illustration: for a fit to a polynomial of degree n one needs at least n+1 points). One prefers to have a small number of instances because working one's way from a registration to the spectral parameters is rather cumbersome. *For these reasons gathering meta-information is of essential importance.* (See the branching to the left in Fig. 1). If we should desire that the analysis system itself would interpret the meta-information (often intensional), then the system must contain a very refined library of sound classes formulated within the MIDIM-

[15]) For simplicity the role played by the domains in seeking a duplication has not been dealt with in this article. Just as in the general MIDIM-theory (Kaegi, 1986) the problem of the domains (with their musical interpretation) is very involved.

[16]) Concerning this widely known problem there exist many publications. We indicate here Ludwig 1969. The best known algorithm is based upon the chi-square-criterium.

language. In the case that our library would contain the concept "Metallophone" then from the meta-expression "a Gender is a Metallophone" the system will "know" which basic MIDIM-functions will describe the Gender sound and only the missing coefficients would have to be abstracted from the instances. (Kaegi 1986, p. 121). There are many differing types of meta-expressions comparable to the one above from which it is possible to derive the necessary MIDIM-functions for the duplication of an instrument. (One does not necessarily need to know the sound in order to apply this procedure. One can ascertain from *the image* of a key-board for example that the pitches will be fixed or from that of large and small instruments (such as an alp-horn or a piccolo) that they shall resp. produce low and high sounds. Naturally previous knowledge is needed).

6.7 The testing experiment

Following our scheme (Fig. 1) we arrived at *a provisional intension* P_i. In order to test our analysis-chain we generate a number of *representative instances* i_p and perform an experiment *in order to investigate whether or not our provisional predicator can be called a duplication of our descriptive concept within the MIDIM-language.* (The condition that the concept must be clearly testable has been earlier stated by us). The experiment can be performed in many different ways, of which we mention three:

1) We let *musicians from the music culture* (MC) listen to the instances and ask if it holds for all i_p that: i_p belongs to K.

2) We let *an arbitrary test person* listen to the pair (i_p,i_K) and ask if the instances are similar in sound (similarity test, see Kruskal, 1964,[17]).

3) One performs spectral analysis and abstractions upon i_K and i_p and compares the corresponding spectra. (These tests are necessary if we should wish to use automatic trial-and-error techniques within an extended and effectivized analysis system.)

If our experiment is negative, then there are numerous methods in order to optimalize P_i. Indications for this are given by test persons and analysis data.

The above description shows how in the last few years many MIDIM-formalizations were arrived at. The quality of these duplications (which are considered as minimum descriptions) have proven on a large scale the enormous power of the MIDIM/VOSIM-system.

[17]) The Kruskal method was used often by Kaegi and Tempelaars in the past (Kaegi, A new approach..., 1972; Tempelaars, Testing elements..., 1973).

PART II - MIDIM-DUPLICATIONS OF THE CENTRAL-JAVANESE INSTRUMENT GENDER

In the second part of this article the results of the analysis method described will be shown for the *Central-javanese* instrument *Gender.* In order to give the reader a general overview of this non-western instrument there follows first a short description (a few important points of "the context"). That the extensional representation of pitch in frequencies is the first step towards a MIDIM-formulization is discussed in paragraph 2 together with an explication of a number of Javanese tuning systems. Afterwards we give a description of various Gender families. The description of the registered instances $i_K \in K_G$ and the concept K will then be the next stage in working towards a systematic presentation of the MIDIM-duplications of the Gender.

1 THE GENDER: AN INITIAL DESCRIPTION

The *Gender* is an *Indonesian metallophone,* which appears in the *Balinese* and *Javanese Gamelan.*[1]) The *Central-javanese* Gender, to which we shall limit ourselves in this article, contains generally 13 or 14 bronze keys which hang horizontally beside each other in a wooden frame by means of a cord. The height of the frame is ca. 40 to 45 cm and the length ca. 1 m, while the width runs from 12 to 16 cm. These sizes hold for the average Gender (usually *Gender Barung).* The keys can vary widely in length, width and thickness, depending on the type of instrument. The keys of an instrument beginning at the highest pitch increase gradually in length and width, while the thickness *decreases.* (In general for a new Gender the measurements of the keys are as follows: length from 19 to 27 cm, width from 5 to 9 cm, thickness from 6 to 1 mm)[2]). See Figure 1.

Beneath every key there is a resonation tube which is stopped on the bottom and which is tuned slightly lower than the corresponding key.[3]) Nowadays the tubes are made from zinc and are all approximately 35 cm in length. The diameter and the effective length (which is determined by a zinc disk soldered on the inside of the tube, the Javanese term of which is: "tumbengan") vary with each key and determine the tuning of the tubes.[4]) Older resonators are made from bamboo which are sawed off in such a way that the joint (the so-called nodium) occurs at exactly the proper height.

1) The Gender was developed in approximately the 11th century (Supanggah, 1985).
2) There exist keys made from brass and iron (a less expensive version) while even aluminum has been experimented with. The keys are not entirely right-angled.
3) The difference in frequency is, according to Mr. Supanggah (1985), approximately 6 to 10 Hz. He has determined the tuning of the new Gamelan of the A.S.K.I.-academy.
4) With the lower keys (2nd and 3rd octave) one would expect tubes longer than 35 cm. This is the case with very old Genders. In younger instruments the opening at the top has been partially covered in order to produce a lower resonance frequency.

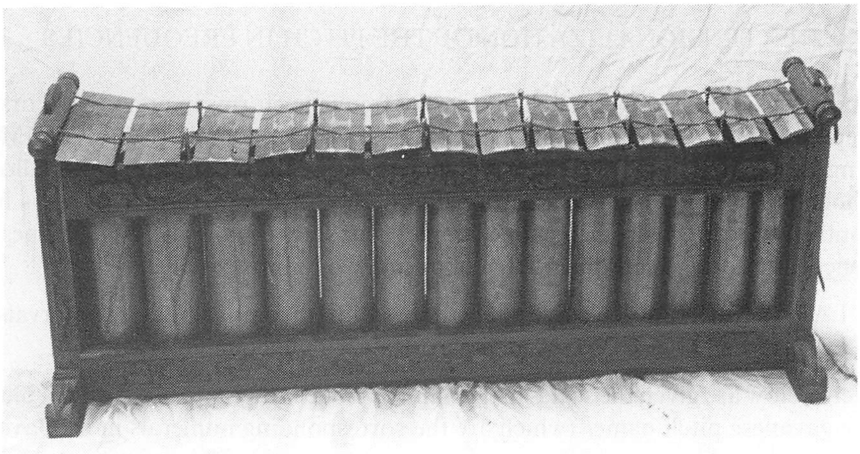


Figure 1. A Central-javanese Gender Barung.

The keys are tuned by filing them either in the middle (to lower the pitch) or on the ends (in order to raise the pitch).

The Gender is played by means of two mallets (in Javanese: "tabuh") with disk-shaped wooden heads wrapped around by cords, which gives the Gender its typically clear sound. See Figure 2. (This contrasts with the hard mallets used by Balinese Gender performers).

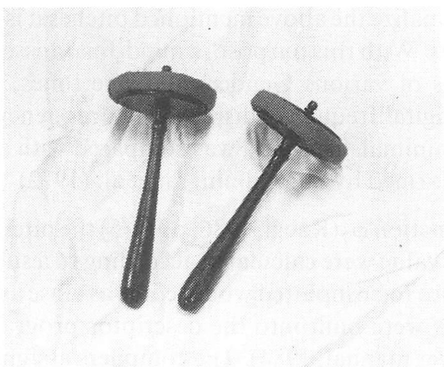


Figure 2. Gender mallets.

The Gender is produced in small factories where each craftsman gives to each instrument he makes a part of his own personality (tuning, timbre, decoration etc.).

2 EXTENSIONALIZATION OF THE PITCH IN FREQUENCIES

The extensionalization of pitch in terms of frequency forms in practice the first step towards formulizing the concept of a Gender-sound in the MIDIM-language (partI, note 14). This was obvious because the set pitches of a metallophone are a property of the instrument and not of the performer. (This is in contrast to string instruments for example). Our starting point was the Javanese concept of tuning according to the following two tuning systems.

1. LARAS SLENDRO – a five pitched scale with (theoretically) equal intervals;
2. LARAS PELOG – a seven pitched scale with unequal intervals.

Here below are given the scales in the number-notation system used in Java plus the Javanese pitch names (which are the corresponding numerals in the Javanese language):

LARAS SLENDRO	1	2	3	5	6	1		
Jav. pitch name	siji	loro	telu	lima	enem	siji		
abb.	ji	ro	lu	ma	nem	ji		
LARAS PELOG	1	2	3	4	5	6	7	1
Jav. pitch name	siji	loro	telu	papat	lima	enem	pitu	siji
abb.	ji	ro	lu	pat	ma	nem	pi	ji

In practice large differences in the tunings can occur. There is actually no standard pitch, so it is not possible to use an instrument in any Gamelan whatever. Therefore the tuning used at the radio (R.R.I.-Surakarta) is accepted more and more as reference.[5])

In order to extensionalize the above mentioned pitches it is necessary *to measure the pitch frequencies.* With this purpose in mind Jos Janssen made comparisons between recordings of various Genders and sine tones, produced by a sine-generator with a digital frequency display. The frequency was regulated until the beats became minimal. This data was compared with and supplemented by measurements performed by Suryodiningrat et al. (1972).

Via the MIDIM-function ϕ_3 (Kaegi, 1986, p. 94/5) the pitch and octave numbers for each frequency value were calculated according to a subdivision n=1200. In order that the pitches to be inputted would come as close as possible to Javanese notation, *compilers* were built into the descriptor program. (Goodman, 1986 and Kaegi, Desc user manual, 1984). The compilers assign automatically to the Javanese pitch numbers, (which are inputted via a subdivision of n=7 or n=6) the corresponding pitch numbers within a subdivision of n=1200.

[5]) A few years ago a new Gamelan was set up at the R.R.I.-Surakarta in which this reference was changed into the so-called new R.R.I.-tuning. The new Gamelan appeared not to satisfy all the expectations which one had for it, so the old tuning was once more implemented.

2.1 The slendro compiler

The Slendro system divides the octave into six segments whereby the fourth note is not used. (This representation is in accordance with Javanese pitch notation). The compilation is only applied if the subdivision is n=6. If a fourth note is made use of, then the compiler steps over to the Pelog system. (This is actually a modulation, which is very normal in Javanese music). See Table 1.

Table 1. Pitch measurements of two Slendro gamelans and their MIDIM-representations.

The following data are given: Javanese notation, MIDIM-input representation n = 6, MIDIM-output representation n = 1200 (after translation), the octave and frequencies in Hz (measurement accuracy ± 0.5 Hz).

Slendro R.R.I. (radio-Surakarta)[7]

Jav. not. [6]	MIDIM input pitch	MIDIM output pitch	oct.	Freq(Hz)
6	0	991	3	231.8
1	1	35	4	266.9
2	2	276	4	306.8
3	3	498	4	348.8
5	5	765	4	406.9
6	6	995	4	464.8
1	1+6	48	5	537.9
2	2+6	307	5	624.7
3	3+6	540	5	714.7
5	5+6	771	5	816.8
6	6+6	1036	5	951.9
1	1+12	106	6	1112.5
2	2+12	337	6	1271.3
3	3+12	566	6	1451.1
5	5+12	816	6	1676.6
6	6+12	1061	6	1931.5
1	1+18	129	7	2254.9

Gangsa Kyai Hardja Winangun (Kraton-Surakarta)

Jav. not.	MIDIM input pitch	MIDIM output pitch	oct.	Freq(Hz)
1	1	14	3	131.8
2	2	255	3	151.5
3	3	524	3	177.0
5	5	752	3	201.9
6	6	1014	3	234.9
1	1+6	42	4	268.0
2	2+6	289	4	309.1
3	3+6	528	4	354.9
5	5+6	769	4	407.9
6	6+6	1017	4	470.7
1	1+12	70	5	544.8
2	2+12	316	5	628.0
3	3+12	570	5	727.2
5	5+12	846	5	852.9
6	6+12	1091	5	982.6
1	1+18	145	6	1137.9
2	2+18	367	6	1293.6
3	3+18	623	6	1499.7
5	5+18	863	6	1722.7
6	6+18	1108	6	1984.6
1	1+24	128	7	2253.6

[6] Raising and lowering a pitch by one or more octaves is designated within Javanese notation by one or more points above or below the given pitch number.

[7] The names of the various Gamelans imply already for insiders whether the pitch scale Pelog or Slendro is concerned.

2.2 The Pelog Compiler

The Pelog system divides the octave into seven segments, so that the compilation is only applied to a descriptor track with a subdivision n=7. If this is the case, then the subdivision is automatically rewritten into n=1200 and the compilation is applied. Table 2 shows two Pelog tunings which can easily be extended.

Table 2. Pitch measurements of two Pelog gamelans and their MIDIM-representation.

The following data are given: Javanese notation, MIDIM-input representation n = 7, MIDIM-output representation n = 1200 (after translation), the octave and frequencies in Hz (measurement accuracy ± 0.5 Hz).

Pelog R.R.I. (radio-Surakarta)

Jav. not. [6]	MIDIM input pitch	MIDIM output pitch	oct.	Freq(Hz)
7̣	0	1065	3	241.9
1	1	201	4	293.8
2	2	282	4	307.9
3	3	443	4	337.9
4	4	765	4	406.9
5	5	876	4	433.9
6	6	1003	4	466.9
7	7	1138	4	504.8
1̇	1+7	228	5	596.9
2̇	2+7	354	5	641.9
3̇	3+7	508	5	701.6
4̇	4+7	777	5	819.6
5̇	5+7	919	5	889.7
6̇	6+7	1032	5	949.7
7̇	7+7	23	6	1060.4
1̈	1+14	232	6	1196.5
2̈	2+14	364	6	1291.3
3̈	3+14	506	6	1401.7
4̈	4+14	816	6	1676.6
5̈	5+14	948	6	1809.4

Gangsa Kyai Mangun Hardja (Kraton-Surakarta)

Jav. not.	MIDIM input pitch	MIDIM output pitch	oct.	Freq(Hz)
7̤	0	976	2	114.9
1̣	1	81	3	137.0
2̣	2	188	3	145.8
3̣	3	304	3	155.9
4̣	4	668	3	192.4
5̣	5	752	3	201.9
6̣	6	858	3	214.7
7̣	7	1021	3	235.9
1	1+7	54	4	269.9
2	2+7	201	4	293.8
3	3+7	348	4	319.8
4	5+7	668	4	384.8
5	3+7	765	4	406.9
6	5+7	872	4	432.9
7	6+7	1039	4	476.7
1̇	1+14	83	5	548.9
2̇	2+14	236	5	599.6
3̇	3+14	388	5	654.7
4̇	4+14	688	5	778.5
5̇	5+14	807	5	833.9
6̇	6+14	925	5	892.8
7̇	7+14	1082	5	977.5
1̈	1+21	164	6	1150.4
2̈	2+21	285	6	1233.7
3̈	3+21	442	6	1350.8
4̈	4+21	713	6	1579.8
5̈	5+21	868	6	1727.7
6̈	6+21	976	6	1838.9
7̈	7+21	1118	6	1996.1

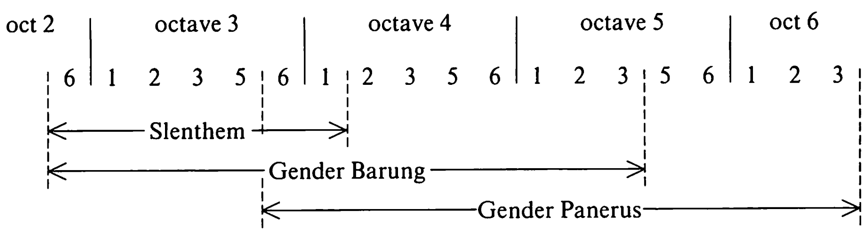
When using a predicator library (Kaegi, 1986, p. 133) every tone is directly assigned the proper predicator at the time of compilation. (In this way it was possible to build up standard Gender libraries. This marries well with the original instance duplications. See also par. 5)

3 VARIOUS GENDER FAMILIES: AN INITIAL DESCRIPTION

Now that the reader has a general impression of the Gender and the common tuning systems, we shall in what follows give a short description of a number of *Gender families.*

The most important three Genders are:
1. Gender Panerus
2. Gender Barung
3. Gender Panembung (usually called Slenthem)

Their pitch ranges are (for simplicity we limit ourselves to the tuning system Slendro; reference: chamber tone a′=440 Hz which is located in the 4th octave.):



The keys, resonators and mallets vary in size according to instrument. The Gender Panerus is the smallest, followed by the Gender Barung and then the Slenthem, which is the largest. The thickness of the keys varies in the *opposite* direction. This holds even in the case of common notes occurring between the different Gender families (see for example pitch 6, octave 3, Fig. 9). The timbral properties also vary amongst the three Gender families. We shall come back to this in detail in par. 5.4

4 THE INSTANCES i_K AND THE GENDER CONCEPT G

A physical analysis demands, as was stated in part I, on the one hand registrations of numerous instances i_K, which belong to a descriptive sound concept K, and on the other hand knowledge of the music culture MC and the context.

In 1980, 1981 and 1985 Jos Janssen undertook study trips to Central-Java in order to become acquainted with both the theory and practise of Gender performance. Via lessons and various interviews with prominent Gamelan musicians (among which Mr. Martopangrawit)[8]) he has attempted to gain an insight into the existing concepts of the Gender.

In the Kraton (the court of Surakarta) Janssen made recordings of Genders taken from four different Gamelans in 1980. These recordings are unique for various reasons, namely, the antique instruments situated there are agreed to be among the best in quality, and it was only possible to move them to one of the Pendåpå's (an open hall) where they sound to perfection, thanks to the permission of Prince Praboewidjojo. (In 1984 a large portion of the Kraton was destroyed, including the Pendåpå where the recordings took place). Further recordings took place in the museum at Bronbeek in Arnhem, the Netherlands, where six Genders (all originating from the Kraton) are in the collection.

All notes were recorded on the one hand isolated (the whole pitch range of the various instruments from near by), and on the other hand applied in a musical context (in modal improvisations, the so-called Pathetan).

An important condition for a duplication is (as we showed) a sensible definition of the concept K, so that K has a clear interpretation in the music culture MC. *Our main concept* G *is: 'The Gender sound of Genders taken from the three Gender families of the Central-Javanese Gamelan'.* It is necessary that the instrument is of a high quality (for a Javanese musician this is a clear concept, see 5.5). Naturally the instrument must be played by a reputed Gender performer. *The recordings mentioned satisfy in every way these conditions.*

Supplementary information concerning the instruments registrated can be found in Appendix II. In the following the appendix shall be referred to by means of a simple code (PAN, BAR, etc.)

5 SYSTEMATIC DESCRIPTION OF THE CONCEPT G

Now that the reader has been introduced to the duplication methods which we will follow and to a basic knowledge of the Javanese instrument Gender, we shall return to our purpose: To find a MIDIM-duplication for the Gender sound according to concept G.

[8]) Mr. Martopangrawit is one of the most important Javanese musicians. He has published a large number of books over music theory and his contributions to the Central-Javanese Karawitan are of extreme importance. (see also App. I).

A start was made in this direction with the extensionalization of pitch (par. 2). Below we proceed further by investigating the time-segmentation, the amplitude envelope and the spectra of the recorded signal.

For clarity we show each time a comparison between a registered instance and the corresponding instance derived from a MIDIM-duplication. This is in accordance with the path originally followed: in the beginning there is sought an exact MIDIM-duplication for each separate Gender note. In a later stage MIDIM-concepts are derived from out these descriptions.

5.1 Time segmentation and amplitude envelope

A time segmentation of a signal registration is the first step in the direction of a systematic analysis of a particular sound. In Figure 3 is shown the start of the signal functions and the amplitude envelopes of the 1st pitch of the 5th octave. Top left and middle show the originals struck on a Gender Panerus (PAN11), top right and bottom show its MIDIM-duplications. (The amplitude scale is linear).

The attack (top left), lasting a few milliseconds, is clearly seen, which is followed by a fairly regular signal in which the fundamental frequency is dominant. In the MIDIM-description a PREFIX of 1 to 3 ms worked wonderfully. The whole amplitude envelope of the same registration (middle) furnishes us with information concerning the other segments. The attack cannot be seen here due to the insufficient time resolution. The segmentations for the BODY, SUFFIX and STOP applied in the duplication are indicated. (bottom).

The amplitude resolution on the Y-axis is chosen in accordance with the VOSIM-variables. (see Kaegi, 1986, VOSIM) To summarize we find: (The data of the prefix can not be abstracted from the figures.)

	d (ms)	A	ΔA
PREFIX	1 à 3	(45)	(0)
BODY	54	511	-136
SUFFIX	1180	375	-175
STOP	250	200	-195
total	ca. 1490		

The amplitude decay in the SUFFIX can be interpreted as a *physical damping of the key,* while the STOP describes a *damping by means of the performers hand.* The time point of the last damping determines the *note duration.* The SUFFIX has thus a variable duration, dependent of the prosodic variable DUR. Thus: v=3 in the MIDIM-concept C'(3) (See Kaegi, 86, p. 116).

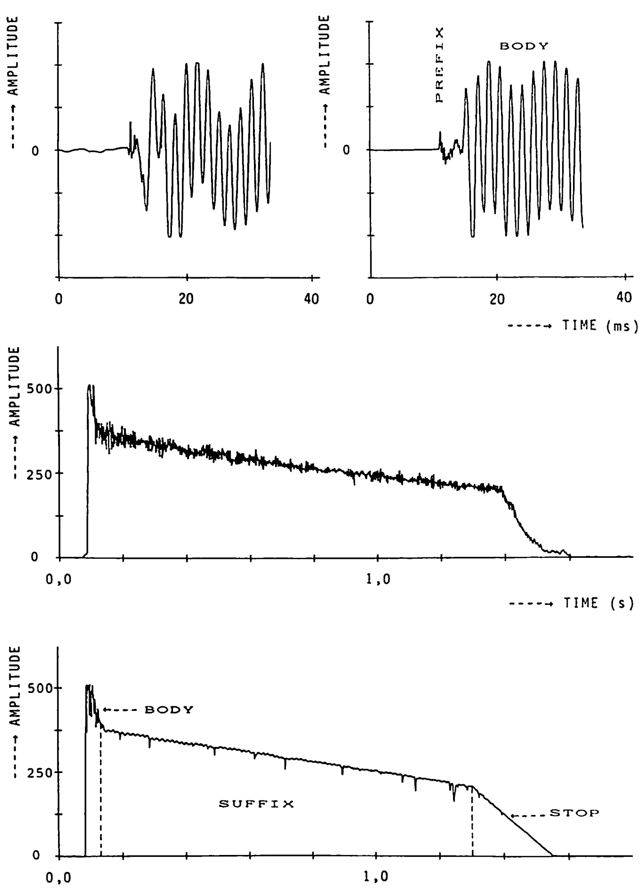


Figure 3. Signal functions and amplitude envelopes of a Gender sound, pitch 1 octave 5, $F_1 = 580$ Hz.
Compare the original signal of a struck Gender Panerus (PAN11) (top left and middle and its MIDIM-duplication (top right and bottom). The different segments are clearly seen.

5.2 A Gender Spectrum

To both the above mentioned signal regïstration and the corresponding MIDIM-duplication (see Figure 3), fast fourier transform is applied. (Program SIGPAC, developed by S. Tempelaars of the Institute for Sonology; for more information see Tempelaars, SIGPAC manual, 1982). Thanks to the abstraction described (see part I), the so-called peak-tracking of the data, the spectrum is clearly reproduced: see Figure 4. The frequency range runs from 125 Hz to 8 kHz and the time resolution is 16 ms.

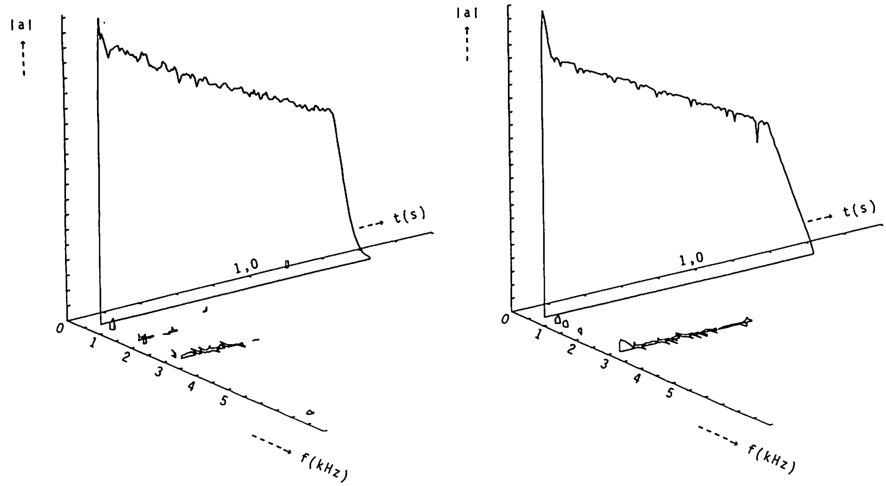


Figure 4. Spectra of a Gender sound, pitch 1 octave 5, $F_1 = 580$ Hz.

Compare the spectrum of the original signal of a struck Gender Panerus (PAN11) (left) and the spectrum of its MIDIM-duplication (right).

The most notable peak-track is found at the position of the fundamental frequency F_1, which we know from the measurements described earlier (par.2). In this case $F_1 = 580$ Hz. (Note 1 of the 5th octave in Pelog tuning (PAN11)). The amplitude as a function of time of this F_1 track agrees well with amplitude envelope of the whole signal (figure 3 middle). The second clear frequency component is found at 3 kHz although it is very weak. We shall see later that this is the 5th harmonic ($F_5 = 2.9$kHz).

Because the time resolution in Figure 4 is too insufficient the prefix is not reproduced, so as an illustration the first 20 ms (of BAR33) were analyzed and plotted; see Figure 5.
The formants and amplitudes of the plosive-like attack were determined by carefully listening to and comparing the registration and duplication.

Before we go further into the relations between spectral components and the prosodic variable pitch, we wish to draw the reader's attention to the high quality of the MIDIM-duplications. They are still *minimum* descriptions.

5.3 Dependence on the prosodic parameters
Seeking for a MIDIM-duplication demands a knowledge of the relation existing between spectral parameters and the prosodic variables T', DUR and At. Since we are dealing only with isolated notes, and in registrations the volume can be

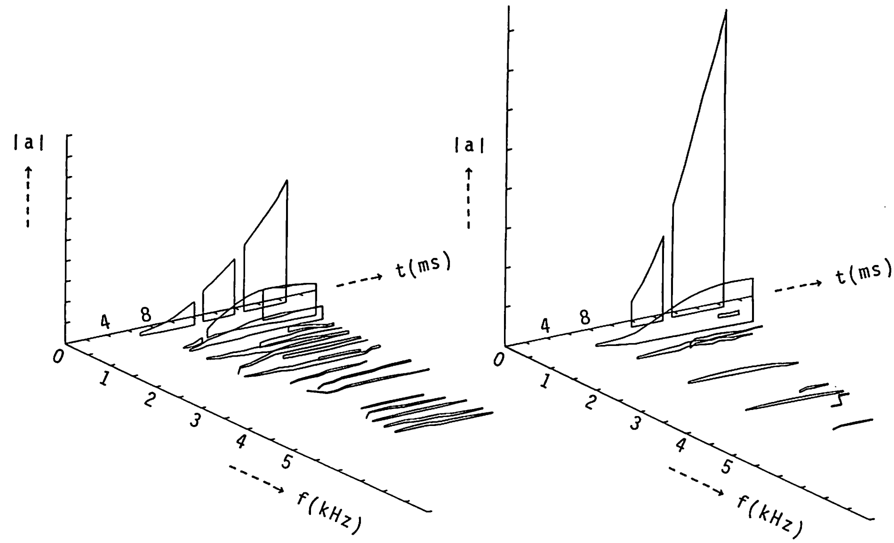


Figure 5. Spectra of the attack of the original (BAR33) (left) and the PREFIX of its MIDIM-duplication (right).

varied arbitrarily, the loudness At can be considered as a rather constant quantity. The note duration DUR, which is determined by the damping of the key by the performer, has (provided it is not too short) an influence only upon the breaking off of the amplitude envelope. The frequency of the spectral components is pretty much constant in time.

The pitch $T' = 1/F_l$ influences, on the contrary, the position of the peak-tracks in the spectra. It is sufficient to construct the spectral space R_s described earlier from the spectral parameters f,a and Δa and the prosodic parameter T′. Figure 6 shows a 2-dimensional subspace of R_s. One can see here the f-dependence of F_l for the whole range of the Gender Barung (Mangun Hardja, Pelog).

As stated in part I, the points in this subspace stand for instances. Our purpose is to connect these with a proper MIDIM-function, which is transformed into the spectral representation. In our case the *function* F_1 is sufficient (see Kaegi, 86, p. 109). As is known T (VOSIM-formant) determines the (second) maximum in the envelope of the VOSIM-spectrum (Kaegi, 86, VOSIM, p. 74). We assume that a harmonic F_n is found at this position, then it holds that (with Of=0).

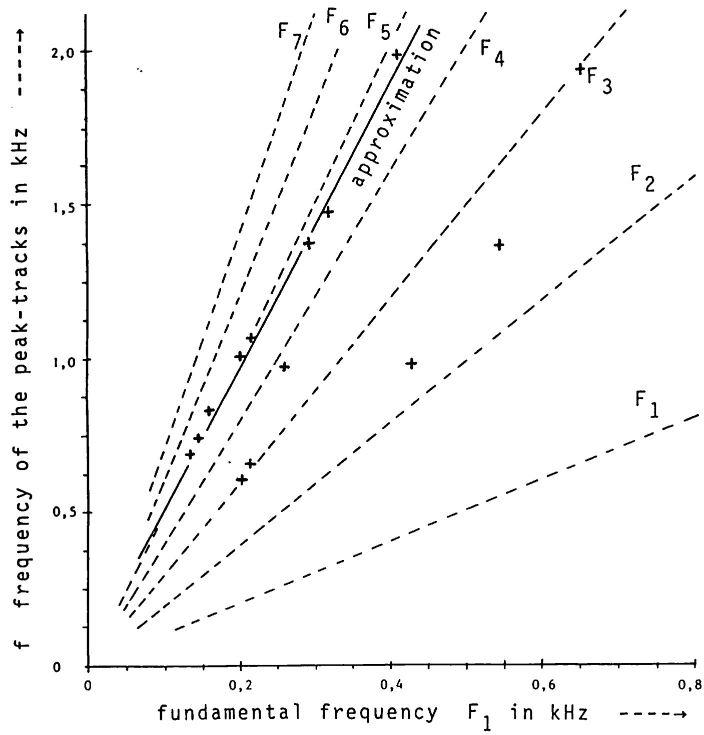


Figure 6. Subspace of R_5. Plotted are the co-ordinates (F_1,f) corresponding to the instances of a Gender (+) (Gender Mangun Hardja) and the approximation of these in the neighbourhood of the fifth harmonic F_5. The theoretical harmonics are indicated by dotted lines.

$$\left.\begin{array}{lll} F_n = n \,.\, F_1 = 1/T & & \text{(requirement)} \\ q = N \,.\, T/T' & & \text{(function } F_1\text{)} \end{array}\right\} F_n = N/(q \,.\, T') = N \,.\, F_n/q \qquad [1]$$

The supplementary condition that all harmonics except F_1 and F_2 must be suppressed, sets q at about 1 (assuming that c = 100%).
Formula [1] then becomes

$$F_n = N \,.\, F_1 \qquad [2]$$

N then appears to correspond with the number n of the dominant harmonic, which we have demanded. In Figure 6 are drawn the theoretical lines for the harmonics F_1 until F_7. The points in the neighbourhood of F_5 are fitted by a straight line (chi-square-criterium) of which the gradient amounts to 4.6.

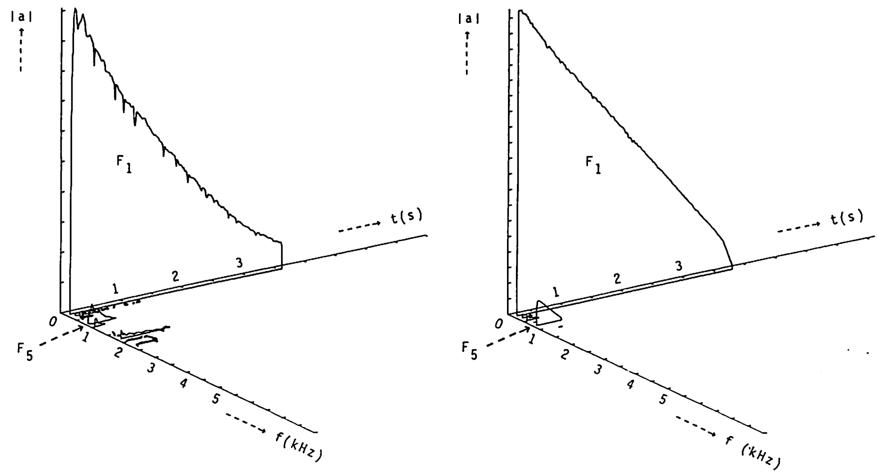


Figure 7. Spectra of a Gender sound, pitch 3 octave 3, $F_1 = 156$ Hz.
The fifth harmonic F_5 is indicated.

Compare the spectrum of the original signal of a struck Gender Barung (BAR33) (left) and the spectrum of its MIDIM-duplication (right).

We round this off to 5 and use in the described formula [2] with N=5:

$$F_5 = 5 \cdot F_1 \qquad [3]$$

We shall speak of *the 5th harmonic in the Gender spectrum.* This is an example of how the adaptation of instances by means of MIDIM-functions can be realized within the space R_s.

Naturally there are more changes in the spectrum as a function of pitch than has been shown above. For example the amplitude of the F_5-track is pitch dependent, as is shown in Figure 7 (spectra of pitch 3, octave 3 (BAR33)).

Comparing this Figure with Figure 4 (pitch 1, octave 5) shows that the harmonic F_5 in the low pitch range is more dominant at the beginning but rapidly decreases. Two octaves higher (pitch 3, octave 5) F_5 is weak although it is played on the same instrument. On the other hand, we find a strong component with a frequency lower than that of the fundamental. See Figure 8. Also the timbre as well as the intensity of the attack vary with the pitch.

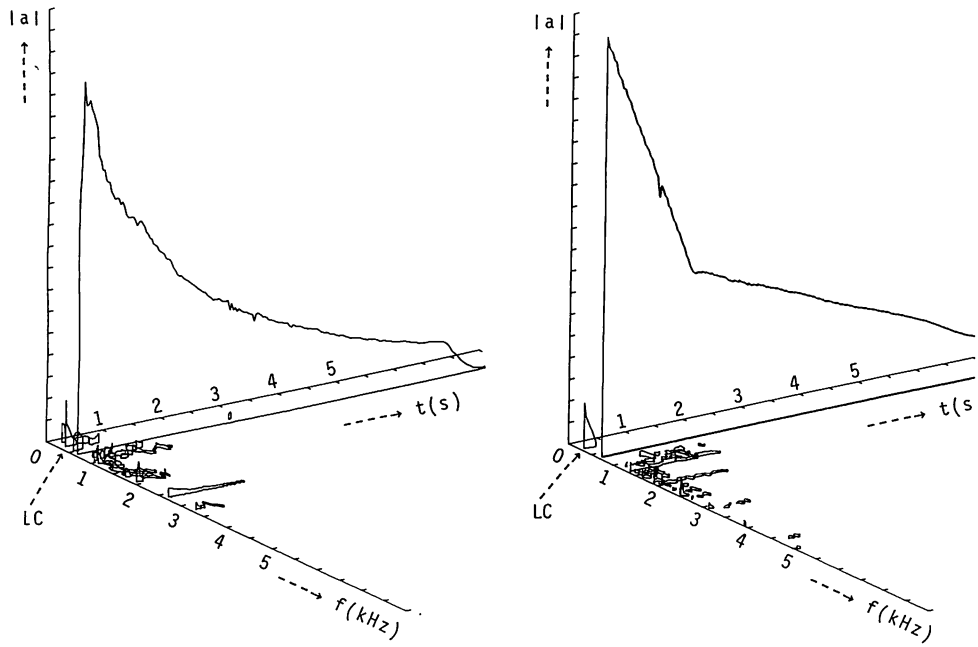


Figure 8. Spectra of a Gender sound, pitch 3 octave 5, $F_l = 655$ Hz.
The low component (LC) can be clearly seen.

Compare the spectrum of the original signal of a struck Gender Barung (BAR03) (left) and the spectrum of its MIDIM-duplication (right).

5.4 The three Gender families

In paragraph 3 we gave the first description of the three current Central-javanese Gender families, which together embrace practically the entire pitch range. The physical difference in the signals of the three Genders takes place in the amplitude envelope. This clearly audible difference is illustrated in Figure 9, which reproduces pitch 6 in the 3rd octave struck on resp. the Gender Panerus, the Gender Barung and the Slenthem. While the keys of the Panerus damp *logarithmically*, those of the Gender Barung approximate a *linear* decay. The amplitude envelope of the Slenthem is in the first 3/4 seconds nearly constant after which a slight logarithmic decay takes place. A correct tuning of the keys and resonator is here of great influence. In this way the amplitude envelope differs considerably per key.

The spectra as well are not entirely identical in the various instruments. Although some show correspondences in the three families,the corresponding keys are not equally large. The large keys bring about the fifth harmonic while the small keys on the other hand produce the low component, analogous to the difference between low and high notes. Also by means of these characteristics the three Gender families are thus characterized.

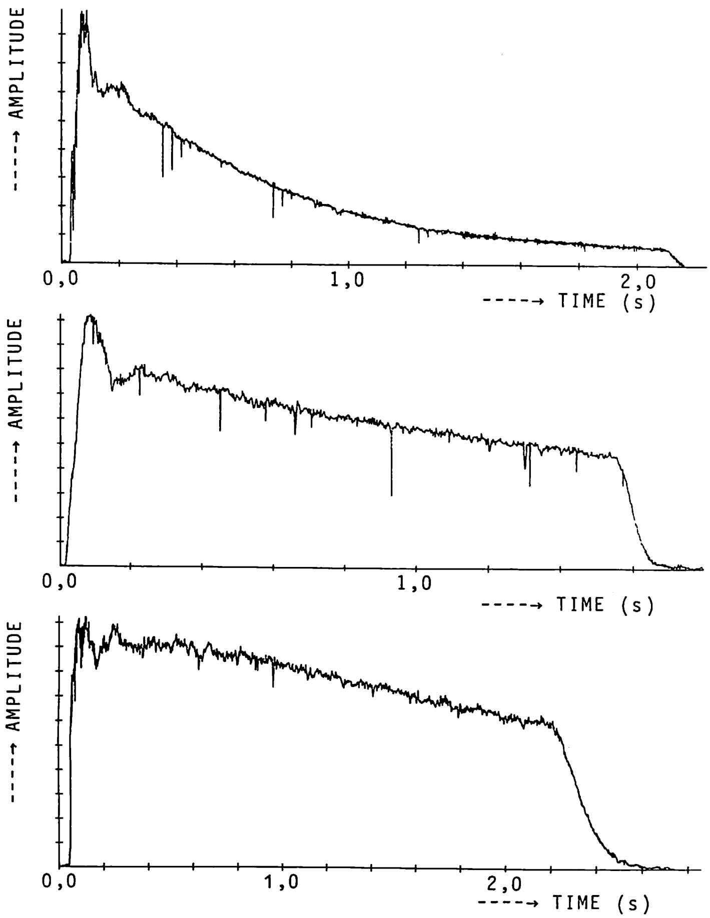


Figure 9. Amplitude envelopes of instruments of the three different Gender families.

At the top the Gender Panerus (PANCO), in the middle the Gender Barung (BARCO) and at the bottom the Slenthem (SLECO).

5.5 Quality differences

The Genders which we have spoken of up to this point are of particularly good quality. They approximate the sound ideal of Javanese musicians for what concerns a Gender:[9])

1) The fundamental must be the strongest component in the spectrum. By this the clear strident sound is created.
2) The instrument must sound the same from a distance as well as up close (This depends strongly upon point 1).

Generally old instruments satisfy these points very well, because they have been brought into balance in contrast with the new Genders. We can see the difference very well in the spectra (Figure 10). Represented are time (on the horizontal axis) and frequency and amplitude (both on the vertical axis; the reduced scale at the beginning of the axis indicates the amplitude, peak-tracks are drawn as closed figures. See van Berkel, 86, p. 239).

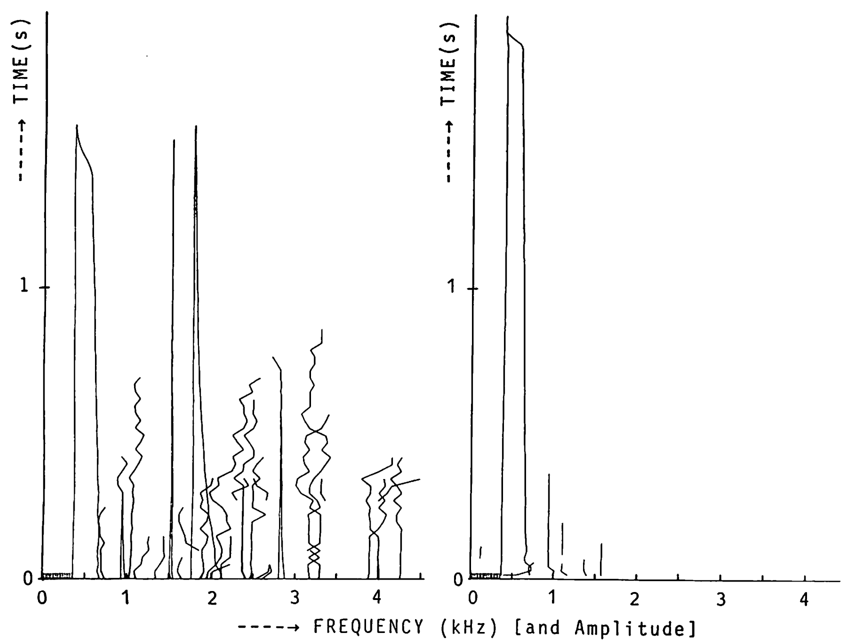


Figure 10.Spectra of a 3 year old Gender (BARNEW) (left) and a ca. 80 year old Gender (BAROLD) (right). (pitch 3 octave 4).

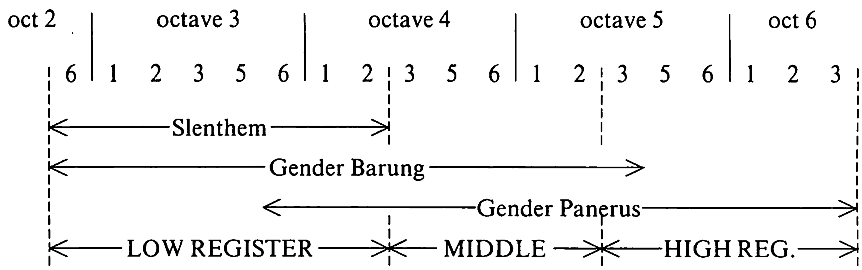
[9]) This sound ideal was notated by Janssen in interviews with S.D. Humardani (†83, previous-director of the A.S.K.I.-academy of Surakarta) and Mr. Martopangrawit (head-master at the A.S.K.I.-academy).

6. A GENERAL MIDIM-DESCRIPTION OF THE CONCEPT G

When proceeding from the MIDIM-duplication we discussed in detail the physical properties of the Gender sound (in light of the VOSIM-spectrum). How one shall realize these within the MIDIM-language has not been dealt with as yet.

6.1 Segmentation of the pitch range into registers

In the frame of the MIDIM8X-system it appeared not to be possible to describe the general concept G for the Gender sound with one single predicator which could be applied over the whole pitch range (par.3). A division into the registers "high", "middle" and "low" was necessary for this purpose. The sound of each register is described by a specific predicator[10]).



In frequencies in Hz:
Low reg.: $F_l = [117,309]$, mid.reg.: $F_l = [309,628]$, high reg.: $F_l = [628,1500]$.

The MIDIM-vectors needed will be shown per segment for the three registers.

6.2 The MIDIM-description applicable in the middle register

Here we shall discuss the MIDIM-vectors used, the concepts C'(v) and the most optimal values, applicable for the *middle* register.

We refer generally to the theory introduced by W. Kaegi in his articles concerning the MIDIM-language and the VOSIM-system (this issue).

[10]) The existing register segmentation could be avoided by expanding the function tables so that the amplitude contour of specific spectral components can be made dependent upon T'.

6.2.1 The following MIDIM-vectors are used:

$\vec{m}_I = (\lambda F_{12}).\vec{m}_0 = \lambda T', \Delta T', Sp, Of.(T, \Delta T, M, \Delta M, A, \Delta A, C, N, D, S, MF, F_{12})$

$\vec{m}g'_0 = (\vec{m}_I\ )(c_S, c_{MF}, c_D, c_C\ , c_{\Delta A}, c_{\Delta M}, c_{\Delta T}) =$
$= (\vec{m}_I\ )(0\ , /\ \ , 15, 100, 0\ \ , 0\ \ , 0\ \ )$

$\vec{m}g_0 = (mg'_0)(c_d, c_N\ , c_A, c_{\Delta A}\ , c_M\ , c_T)$

Within vector $\vec{m}g_0$ none of the prosodic parameters is λ-tied so that $\vec{m}g_0$ becomes suitable for the PREFIX of our Gender predicator.

The following vector $\vec{m}g$ is derived by us from the vector $\vec{m}_{651}$, which is presented by W. Kaegi in paragraph 3.6 concerning metallophone-like sounds (p. 110):

$\vec{m}_{651} = (\lambda\ F_6\ F_5\ F_{11}\ F_9\ \phi_6\ F_2\ F_1\ F_4\ F_3\ F_{12})\ldots, d.\vec{m}_0$

$\vec{m}g' = (\vec{m}_{651})(c_N, c_q\ , c_P, c_{Of}, c_e, c_{y2}, c_C\ , c_S, c_{Sp}, c_W) =$
$= (\vec{m}_{651})(c_N, \frac{3000}{3006}, 0\ , 15, 1\ , 1\ \ , 100, 0\ , 0\ \ , 0\ \ )$

$\vec{m}g = (\vec{m}g'\ )(c_{Am}, c_{\Delta Am})$. This vector is implicitly described in par. 5.3. It produces the *fundamental* F_I and the *harmonic* F_N. The other harmonics are suppressed.

By simplicity we add to these vectors an index, which indicates the value of c_N. Thus $\vec{m}g_I$ produces only the fundamental, because $c_N = 1$.

6.2.2 The concepts

By means of these vectors the following concepts are built up, wherein the third segment, the SUFFIX, has a variable time duration $d_3 = \phi_5\ (d_i, DUR)$. Thus $v = 3$. In the PREFIX is $c_{\Delta A} = 0$.

$Gm = (C'(3))(\vec{m}g_0, \vec{m}g_I, \vec{m}g_I, \vec{m}g_I)$ applicable in *both* tracks.

In the first track:

$Gm1 = (C'(3))\ \ \{(\vec{m}g'_0)(c_d, c_N, c_A, c_M, c_T)$

	1 1 55 15 360,			PREFIX
$(\vec{m}g'_I)$	$(c_{Am},$	$c_{\Delta Am})$	i	
	500,	—250	2	BODY
	250,	—175	3	SUFFIX
	75,	— 75	4 }	STOP

In the second track:

$$Gm2 = (C'(v)) \quad \{ (\vec{m}g_0')\, (c_d, c_N, c_A, c_M\ , c_T)$$

$(\vec{m}g_0')$	$(c_d, c_N, c_A, c_M, c_T)$		
	1 1 40 150 200,		PREFIX
$(\vec{m}g_i')$	$(c_{Am},$ | $c_{\Delta Am})$	i	

c_{Am}	$c_{\Delta Am}$	i	
65 ,	—35	2	BODY
30 ,	—20	3	SUFFIX
10 ,	—10	4 }	STOP

The 'm' in the names 'Gm' indicates the middle register.

6.2.3 The predicator used is:

$$PGm = (\lambda L_1\ L_2\ L_{33}\ L_5)\, v, d_2, d_4, bea, DUR, T', At.(Gm1, Gm2)(7, 160, 60, 3)$$

The link function L_{33} gives us a constant *beat frequency* of (in our case) 7 Hz, and simulates thus the difference between the tuning of the key and the corresponding resonator (See par. 1, note 3).

6.2.4 The domains of the parameters used

The values shown in the above tables are applicable for an optimal case. Within the general Gender concept presented these parameters may vary over particular domains of which we shall now mention a few:

a) The most important parameter in the PREFIX is the time-duration d_1. It holds approximately: $d_1 \in [1,5]$.
The other parameters may vary extremely and should be chosen experimentally (Kaegi, 1967).
b) The amplitude contour can be varied by which means *the differences existing between the various Gender families* may be formalized.
(These values can be abstracted from fig. 9, par. 5.4)[11]).
c) As has been shown in par. 5.5 the amplitude A_{F_5} and the duration of the harmonic F_5 determine the differences between one Gender and another.

[11]) As we have seen the Gender families differ mainly in their amplitude contours. The linear amplitude changements can be described by means of the present MIDIM/VOSIM-model very well (parameters A and ΔA). The logarithmic envelope of the Gender Panerus offers more problems. In order to solve this problem one can extend the number of segments, so that a step approximation of the envelope becomes possible. Another solution lies in the hardware of the VOSIM-generators which up until now could only give a linear interpolation over a segment.

In general the following extreme values are valid:

New instruments: – $A_{F_5} \simeq A_{F_1}$ *in the beginning of the BODY, but the amplitude of* F_5 *decreases rapidly.*
– *The duration of* $F_5 \simeq$ *duration of* F_1.
Old instruments: – $A_{F_5} \simeq 0$.

See Figure 10.

6.3 The MIDIM-description applicable in the other registers

Starting from the MIDIM-description of our Gender concept G for the *middle* register, we can extend this duplication to the *low* and the *high* register *by replacing some vectors in the second track.* In the PREFIX only *some characteristic values* are changed.

6.3.1 The concepts

In par. 5.3. was shown that the spectrum of low pitches contains the fifth harmonic F_5 with a small amplitude in the first 60 ms of the signal. To simulate this we replace in the BODY and the SUFFIX of concept GM2 the vector $\vec{mg}_1$ by the vector $\vec{mg}_5$ (thus $c_N = 5$) and adapt the amplitude within this track.

For the high pitches there appeared, as was shown, a low component in the spectrum. We introduce this in our concept GM2 by replacing vector $\vec{mg}_1$ by $\vec{mg}_0$ in the BODY ($\vec{mg}_0$ is pitch-independant).

	PRE	BOD	SUF	STO	
GH2 = (C′(3))	($\vec{mg}_0$,	$\vec{mg}_0$,	$\vec{mg}_1$,	$\vec{mg}_1$)	high register
GM2 = (C′(3))	($\vec{mg}_0$,	$\vec{mg}_1$,	$\vec{mg}_1$,	$\vec{mg}_1$)	middle register
GL2 = (C′(3))	($\vec{mg}_0$,	$\vec{mg}_5$,	$\vec{mg}_5$,	$\vec{mg}_1$)	low register

For comparison the middle register is also indicated.
Compare the used concepts with the figures 4, 7, 8.

6.3.2 The values used:

In the first track: (assumed is $c_{\Delta A} = 0$)

GH1 = (C′(3)) {($\vec{mg}'_0$) (c_d, c_N, c_A, c_M, c_T), PREFIX high reg.
1 1 63 15 100

($\vec{mg}'_i$) see middle register for i = [2, 4]}

GL1 = (C′(3)) {($\vec{mg}'_0$) (1, 1, 45, 15, 90), PREFIX low reg.

($\vec{mg}'_i$) see middle register for i = [2, 4]}

In the second track:

$GH2 = (C'(3))\ \{(\vec{mg}'_0)$	$(c_d,$	$c_N,$	$c_A,$	$(c_{\Delta A}),$	$c_M,$	c_T	$)\ i$	high register
	1	1	45	0	15	60	1,	PREFIX
	60	1	125	−125	10	2000	2,	BODY

$(\vec{mg}'_1)\ (c_{Am}, c_{\Delta Am}) = (\vec{mg}'_1)\,(0{,}0)$ for $i = [\,3{,}4\,]\}$

$GL2 = (C'(3))\ \{(\vec{mg}'_0)$	$(c_d,$	$c_N,$	$c_A,$	$c_M,$	$c_T,$	$)$	low register
	2	2	36	152	56		PREFIX

$(\vec{mg}'_5)$	$(c_{Am},$	$c_{\Delta Am})$	i
	45	40	2, BODY
	5	5	3, SUFFIX

$(\vec{mg}'_1)\ (c_{Am}, c_{\Delta Am}) = (\vec{mg}'_1)\,(0{,}0)$ for $i = 4\}$

6.3.3 The predicators PG1 (low) and PGh (high) are analogous to PGm (middle).

7. TESTING THE DUPLICATIONS

It was made clear in part I that *an experiment* was necessary in order to decide if the instances belonging to a predicator P sound like the corresponding instances, derived from the concept K, in our case G (see scheme part I, Fig. 1).
For the testing of the various Gender predicators (of which we have shown here the most general form) all three of the methods described (part I par. 6.7) were applied:

1) During his second study trip to Java in 1981, Jos Janssen played to Javanese musicians Gender duplications made with the MIDIM-system which had been recorded on tape. Further, the well known Javanese musician *Supanggah Rahayu,* while on a tour to the Netherlands (he was soloist in the composition Dialogue, see composition list: Goodman, 1986, p. 182), listened to duplications which he then could criticize. On the basis of his comments the original description in three segments was extended to four. By this means it was possible to bring into play the typical *damping of the performer* within the concept (STOP in the described segmentation)[12]. Also the compilers (discussed in par. 2) were

[12]) In the case that the STOP should be missing, then the amplitude must decrease to zero during the SUFFIX. Mr. Supanggah commented that, when this was so, a Gambang (Javanese Xylophone) was approximated rather then a Gender. After his remarks we naturally tried to make a duplication of the Gambang. A few instances were duplicated derived from the Gender predicator with roughly the following alterations: the amplitude contour (a very strong decrease in the body) the Stop is the variable segment in duration.

tested in this way. Mr. Supanggah himself recognized the tunings he had applied to the original instruments.

2) In order to compare the instances *test tapes* were set up in which an original sound of a struck key was always followed by its duplication. Many test persons could not differentiate the orignal instances from the duplicated instances. (At the dutch AES-convention, 1985, Dr. W. Kaegi showed among others some MIDIM-duplications of the Gender. The critical audience kept silent when they were asked to distinguish the original from the duplicate).

3) The instances were compared *spectrally* many times. We do not need to go any deeper into this because the reader has already met a number of comparison examples in the text.
These tests proved sufficiently that the duplications satisfied the concepts to which G belongs.

8. MUSICAL APPLICATIONS OF GENDER CONCEPTS

Apart from the general or family related concepts of the Gender, attempts were made to duplicate *individual* instruments so accurately as possible. For this reason *predicator libraries* were built up with a large number of registers. For artistic applications it appears that these descriptions are highly satisfactory because of their variety which makes them living. (La Belle et La Bête by Goodman/Janssen and Dialogue by W. Kaegi, see composition list page 182).

In order to illustrate this we will show here an example, once more presented in the spectral space. See figure 11. (The signals are filtered in such a way that only the fundamental can be seen). It is the beginning of a so-called *Modal improvisation Laras Pelog Pathet 7*, performed by Doyopangrawit in the Kraton of Surakarta upon the Gender Mangun Hardja. The melody in Javanese notation looks as follows:

$7 \; . \; . \; . \; . \; \dot{2} \; 6 \; 7$ right hand
$. \; \underset{\cdot}{7} \; \underset{\cdot}{5} \; \underset{\cdot}{6} \; \underset{\cdot}{7} \; . \; . \; .$ left hand

A point above a number raises a note by an octave. The points under lower the note by an octave (see pag. 201, note 6). It is important to state that a performer damps the keys whenever the following note is struck so that the notes overlap one another. The overlapments can be clearly seen in the spectra. (These overlapments in the duplication are calculated by a special overlap facility, implemented in the descriptor program).(See Kaegi, 1984.)

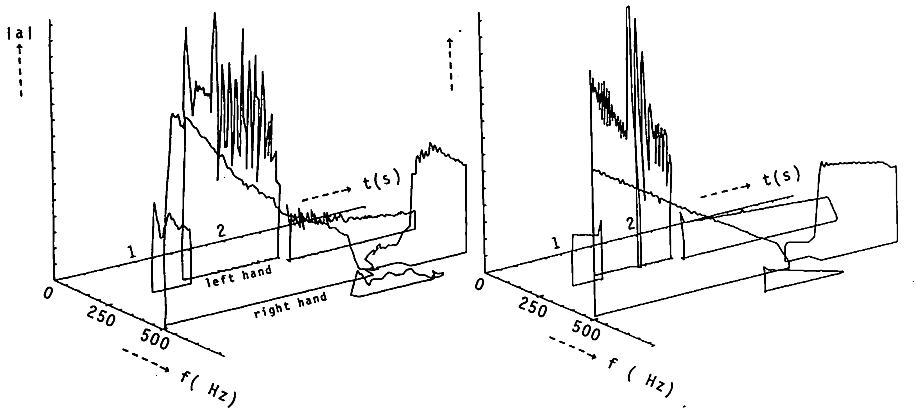


Figure 11.Spectra indicating the fundamental frequencies of the following melody:

7 . . . . . $\dot{2}$ 6 7	(right hand)	(above 400 Hz)
. $\underset{\cdot}{7}$ $\underset{\cdot}{5}$ $\underset{\cdot}{6}$ $\underset{\cdot}{7}$ . . .	(left hand)	(below 400 Hz)

Compare the original played on a Gender Barung (BAROV)(left) and its MIDIM-duplication (right). Note that the pitches 6 and 7 played by the right hand are represented by one peak-track; the same holds for the pitches 5 and 6 played by the left hand.

9. FUTURE DEVELOPMENTS

It is intended that, with the help of the new MIDIM9-system and compatible pattern recognition systems, further experiments shall take place with duplications. On the one hand investigations into the non-western sound world will receive attention, through which *ethnomusicological researches* will be possible on a large scale. In connection with this one can for example think about *formal comparisons of instruments and language within the same culture.* By means of this the questions as to *what is a musical sound* shall acquire a formal foundation. On the other hand in the future there shall, by means of duplications and through sound libraries (which will be developed on the basis of these duplications), arise *intensional descriptions of the sounds* which the composer wishes to have at his disposal. The expression "I would like something which sounds like a Gender 'and' the song of a blackbird" can then be *directly* extensionalized in the MIDIM-language, after which the desired sound – if it is physically realizable – will be supplied by the system (Kaegi/Janssen/Goodman, 1986).

APPENDIX I

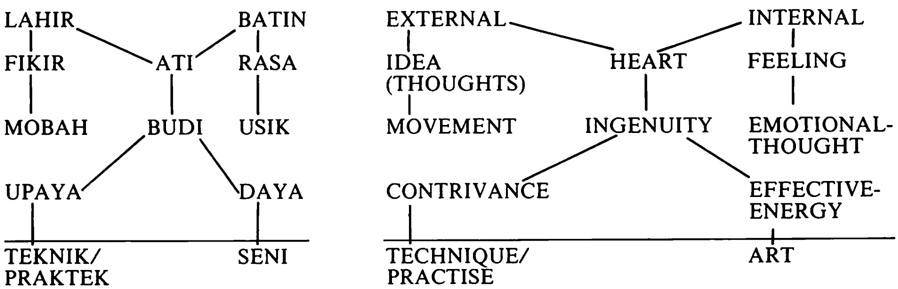
In this appendix we will deal with some supplementary basic concepts which are valid within *the Music Culture (MC) of Java:* the musical concepts 'GARAP', 'GARAP PATHET' as well as the concept 'MODULATION' and the sound concept 'GONG'.

We shall base our discussion upon the doctoral dissertation 'INTRODUCTION AUX STYLES ET INTERPRETATION DANS LA MUSIQUE JAVANAISE' written by Dr. Rahayu Supanggah (1985), as well as the book 'PENGETAHUAN KARAWITAN' written by R. L. Martopangrawit (1975). Moreover use shall be made of personal interviews conducted by Jos Janssen with Mr. Martopangrawit in Solo (1981 and 1985).

1. INTRODUCTION

In one of the interviews Mr. Martopangrawit (nov./dec '85) gave a scheme concerning the arts which according to him must hold before the term 'art' may be applied. We will now show the scheme but without going into too many details. (There appear to be certain correspondances between the following scheme and various aspects of the philosophy of Aristotle.)

'SCHEMA KEBUDAYAAN' (literally: scheme for the arts) constructed by Martopangrawit



In the scheme above there are two main processes 1. the external,
2. the internal

Within the two processes we constantly find complementary levels which must be harmonized, such as fikir-rasa (idea/feeling), mobah-usik (movement/emotional-thought), etc. Imagine that you suddenly want to go to the cinema (emotional-thought), then of course you have to get to the cinema somehow (movement) and to have sense enough to buy your ticket and stay out of jail (ingenuity). According to Mr. Martopangrawit every artist in a state of development passes through these various stages and for each stage there are intensionally and extensionally interpreted concepts. Although these concepts were in the past hardly notated and thus may seem somewhat unclear to Western musicologists, *they are quite clearly defined what concerns the local musicians and moreover are easily demonstrable on their instruments.*
(In the last few decades many publications have appeared with the purpose of setting these concepts down in writing. Most especially at the A.S.K.I.-academy, Surakarta, Indonesia.)

In what follows a number of these concepts shall be dealt with in detail.

2. GARAP

The Javanese concept *Garap*, which may be translated as *"interpretation"* when first making contact with the term, plays a central role in the Karawitan. Garap is defined by Dr. Supanggah in the following way: *"Garap is concerned with the domains of creation, interpretation, and even inspiration and imagination. (...) The Javanese artist possesses a great deal of freedom within the Garap."* One should pay attention that the notion of freedom must not be confused with the concept of improvisation.

"The Javanese shadow-play, the Wayang Kulit, could be helpful in clarifying the term Garap and the freedom the artist enjoys: for the same drama (Lakon) the effect, the style and the content may vary according to the puppeteer (Dhalang) and the circumstances, which means the moments when the play comes to life or the scenery (ambiance). Thus for a single drama every puppeteer may create a different "Garap" or "Sanggit" (Sanggit is the creativity of the Dhalang).

The Dhalang may vary the following among others:
1. the order of the scenes
2. the accent laid upon the personalities of certain characters
3. the plot in the progress of the drama
4. the choice of the movements of the puppets (Sabet)
5. the music accompanyment
6. the rhythm of the performance
etc."

"The Lakon is nothing more than a framework within which the puppeteer may build or arrange the drame entirely in accordance with his talent and imagination (...)."

Within the general concept of Garap one may mention among others the following sub-concepts:

1. GARAP LARAS (tuning system): there are two tuning systems namely Pelog and Slendro[1]). It is possible to perform very many melodies in both systems but it obliges one to reinterpret the individual parts/melodies. The tuning systems are based upon the Javanese concept *Embat Alam* (natural interval). When tuning the instruments this fact is taken into account. During a performance, the musicians give their own interpretation to the tuning system particularly the *Pesindhen* (female singer), the *Rebab* (two-stringed fiddle) and the *Siter*. One could speak of *dynamic tuning systems*[2]).

2. GARAP IRAMA (rhythm, tempo): every rhythm necessitates another interpretation (Garap). Rhythm in the sense of "(...) *the expanding and contracting of structural units such as the* Gastra *(musical unit of four basic melody-notes) and the degree or level at which the Gatra is subdivided (or filled in)."* (Becker, 1984).

[1]) In antique Javanese poetry there is sometimes references to 3 tuning systems, namely Slendro, Pelog (Pelog 6) and Barang. (Information derived from B. Arps, Javanist in Leiden, the Netherlands.)

[2]) Mr. Martopangrawit related in one of the interviews that once he was asked to play a siter. At the beginning of the concert he tuned the instrument to the Gender. During the first composition it appeared that the tuning did not harmonize with the rest of the orchestra (Gamelan) thus for the next composition he tuned the instrument to the Slenthem. Afterwards it still seemed to be the case that the tuning was not right and he decided to tune to the sound of the entire orchestra which produced the desired results.

3. GARAP PATHET (mode, tonality): there are three modes per tuning system, which obliges a particular musical interpretation (as far as ones knowledge extends). See also under the header *Garap pathet.*

4. GARAP of the DYNAMICS: a less developed concept within the total Garap. Only in the last ten years has there been much work done to develop this concept by means of new compositions with strong contrasts in the dynamic level.

5. GARAP of the ORCHESTRATION: traditionally the orchestration within a composition is fixed. A term which would be equivalent to *'solo-instrument'* does not exist. Even the singers follow the same rules as those pertaining for the instrumentalists and only through recording techniques are the male or female singers allowed to dominate the rest of the orchestra. Since 10 or 15 years experiments have been conducted in the field of orchestration such as using a drum set for a dance performance or using various instruments from different parts of Indonesia, but before this time it was not common practise.

One of the dangers arising from an admixture of various instruments and styles is according to Mr. Martopangrawit (interview, nov. 1985)
"that music and dance etc. is becoming less and less abstract; tending to refer too much to daily events such as the grinding of rice or the influence of radio and television (...)". One of the most important foundations of Javanese culture is just this abstraction in form within music, dance, the Wayang kulit, the Batik, etc.

3. GARAP PATHET

Pathet can according to Dr. Supanggah be defined in the following way: *"It is a Javanese musical system which classifies the compositional repertoire according to its range (or register), the proper moments when they may be played by someone (the artist), or the order in which someone may play them, the feeling the composition evokes, the vocabulary of the execution and the tonal hierarchy".* (Supanggah, 1985).

There exist disagreements concerning the concept Pathet but we will not go any further into this at the moment. A thesis on this subject is being written by Mr. Sri Hastanto at Durham University, England. Mr. Supanggah points out that the interpretation of Pathet (Garap Pathet) is also dependent upon the artistic meaning each player gives to it and depends on the choice of the melodic patterns (*Cengkok*), their variations (*Wiled*) with respect to the other instrumental and vocal parts. A Javanese musician is able to show his talent and richdom qua *Garap* by means of different Wiled, which differ with each performance.

A composition in Javanese music only achieves its final realization with its interpretation (Garap) during a performance. For this reason the musicians dislike to have their parts written out as they feel it limits their personal interpretation, although a concise guide for a performance is given by means of a basic notated melody (called in Javanese *Balungan*).

To leave scope for Musical freedom Javanese music is severely structured, which Mr. Martopangrawit termed during one of the interviews *"freedom within knowledge"* which every artist must attain (...) *within the Karawitan one is very free but not so free as to take a train without buying a ticket (...), only someone who understands the concept of Garap is able to explain the composition* (Gending), *the mode* (Pathet) *and the tuning system* (Laras)".

Within each tuning system the following Pathet terms and corresponding pitch scales are stated:

LARAS SLENDRO PATHET	SANGA	(=9)	5 6 1 2 3
	NEM	(=6)	2 3 5 6 1
	MANYURA	(=?)	6 1 2 3 5
LARAS PELOG PATHET	LIMA	(=5)	5 6 1 2 3
	NEM	(=6)	2 3 5 6 1
	BARANG	(=7)	6 7 2 3 5

The modal term *Manyura* means "peacock" in Sanskrit while the other terms designate a number in Javanese. (In India the pitches were designated by names for animals, continents and heavenly bodies etc. (te Neyenhuis, 1970). Concerning the pitch scales of *Laras Pelog* there are many differences in opinion as to which notes are the tonic, dominant, sub-dominant etc. It is possible that the dissertation mentioned above of Sri Hastanto in Durham will clear up this matter. (The scales mentioned above were taken over from Martopangrawit (1975) discussing the item.)

The *Gender* (*Gender Barung*) is the most important instrument for the accompanyment of the puppeteer (*Dhalang*). The Gender player must keep the puppeteer in the correct Pathet for the whole 9 hour duration of the performance (three Pathets each lasting 3 hours). Each Pathet expresses a particular state of mind and a good performance demands of the musician an extensive knowledge of the basic concepts Garap and Pathet mentioned above.

The wife of a puppeteer often plays the Gender during a Wayang performance, thus there exists a *female style of Gender performance:* certain special variations (Wiled) are only used by female Gender players. However it is nowadays difficult to point out exactly what the difference is between the Gender-performance of male or female.

Among Javanese musicians there exists a special term for the quality of an instrumental sound: *Nyopak* (*"This Gender is Nyopak"* (is good)). The tuning of a Gender does not seem to be terribly difficult at first sight (when one tunes the tubes and keys in accordance with each other, then the Gender seems to be Nyopak), but it is possible to tune a Gender in such a way that a particular Pathet becomes more "precisely" tuned than the other two Pathet. This demands knowledge of the concepts Garap, Pathet and naturally Laras. Nowadays the Pathet Manyura is often used for the "precise" tuning in Laras Slendro, while not so long ago Pathet Sanga was also used. For modulations the tuning is of importance as well.

4. MODULATION

There are two modulations
1. modulation of Pathet
2. modulation of tuning system

In order to make it possible to modulate from one tuning system to another it is necessary that one of the notes must be common (*the Tumbuk-note*). A melodic phrase or sentence is always closed by striking a *Gong*; with a modulation the Gong-note is identical to the

Tumbuk-note. Usually this is note 5 (Tumbuk 5) or 6 (Tumbuk 6): note 5 for *Laras Slendro Pathet Sanga* and *Laras Pelog Pathet Bem* (Pathet 5 and 6) and note 6 for *Laras Slendro Pathet Manyura and Laras Pelog Pathet Barang* (and the seldom used fourth pathet Pelog Manyura with the scale 6 1 2 3 5). When constructing the compilers we took into consideration the Tumbuk-note. (The gamelans from out the Kraton Mangun Hardja (Pelog) and Hardja Winangun (Slendro) use Tumbuk 5, the R.R.I. tuning Tumbuk 6, see page 201/202).

In the following scheme which has been taken over in part from Mr. Martopangrawit (1975) the theoretical intervals within the systems Slendro and Pelog are shown.

Slendro:	pitch	1	2	3	5	6	1		
	cent		240	240	240	240	240		
Pelog:	pitch	1	2	3	4	5	6	7	1
	cent		150	150	225	150	150	150	225

The Pelog scale could possibly give an explanation why exactly the three Pathet which we have named are used: the theoretical intervals in the three Pathet made use of are identical, namely:
150,375,150,375,150 cent. This might be why these are used so often while Pelog Pathet Manyura (375, 150, 150, 375, 150 cent) is seldom applied.

The fact that there is *no standard pitch* may also play a role in the Garap of a composition: certain compositions sound better when played upon certain Gamelans. In the court of Surakarta (Kraton) one will find various different Gamelans in residence. One of the Gamelans of the A.S.K.I.-academy which was tuned by Dr. Supanggah (who apart from his work as head-teacher at the academy performs sometimes as a puppeteer) is particularly suitable for the *Wayang Kulit.*

5. THE SOUND CONCEPT "GONG"

One finds in the Music Culture of Java *duplications of instrumental sounds made with the help of other instruments. Gong sounds* for example can be made by the following differing instruments:

1. The GONG AGENG (large gong): consists in its usual form from out a round bronze plate with a large bulge in the middle, where he is struck with the fist or a cudgel.

2. The GONG KEMODHONG: is composed from out two bronze keys with earthenware resonators mounted in a wooden cabinet. The keys and resonators are tuned slightly apart (ca. 6 Hz). The Gong Kemodhong gives an extremely good duplication of the sound concept Gong (this instrument is very suitable for small ensembles, the so-called *Gamelan Gadhon*).

3. The GONG BUMBUNG: is composed from out a thin bamboo tube and a thick stopped resonator. The resonator is brought into vibration by blowing into the thin tube. A good performer is capable of producing a surprising duplication of a large Gong; only through the decay does it become clear that it is a blown instrument (sudden breath pause).

4. The Human Voice: The word Gong is onomatopoeic. By calling out the word one duplicates the concept via speech and in this way it is even possible to imitate complete Gamelans.

NOTE: If mistakes and incorrect interpretations concerning Javanese theory are found in this article the fault lies with the authors.

APPENDIX II

Information concerning the Genders made use of:

BAR33,BAR03 en BAROV
Gender Barung Laras Pelog taken from out the Gamelan KYAI MANGUN HARDJA which is situated at the KRATON SURAKARTA, Indonesia: figures 5-8 and 11

PAN11,BAROLD,PANCO,BARCO en SLECO
Gender Panerus, Gender Barung, Slenthem Laras Slendro and Pelog taken from out the Gamelan KYAI NUGROHO PURADININGRAT MAESO which makes a part of the collection of the Bronbeek museum in Arnhem, Holland. This Gamelan comes originally from the KRATON of SURAKARTA: figures 3,4,9 en 10

BARNEW
Gender Barung Laras Slendro, RRI-tuning, manufacture date 1982: figure 10

APPENDIX III

Information concerning the spectral analyses:

The spectral analyses are performed by means of the signal processing programs SIGPAC, developed by S. Tempelaars.

figure	code	sampling rate	window widt/shift	type
3,4	PAN11	16 kHz	256/128 samples	H
3,4	PAN11	16 kHz	256/128 samples	H
5	BAR33	16 kHz	256/16 samples	H
7	BAR33	16 kHz	512/256 samples	H
8	BAR03	16 kHz	512/512 samples	O
9	PANCO	16 kHz		
9	PANCO	16 kHz		
9	SLECO	16 kHz		
10	BARNEW/BAROLD	16 kHz	256/128 samples	H
11	BAROV	16 kHz	512/256 samples	H

(type of window: O=open window, H=Hanning window).

BIBLIOGRAPHY

REFERENCES

Banerij, R.B. (1969). Theory of problem solving. An approach to artificial intelligence, Modern analytic and computational methods in science and mathematics, number 17, Elsevier.
Becker, J., see Martopangrawit 1975.
Berkel, P. van, Notes by a Film-Maker, Interface, 1986.
Carnap, R., Meaning and Necessity, University of Chicago Press, 1947 seventh edition 1975.

Goodman, P. The user and the Midim system, Interface, 1986.
Kaegi, W., 1986. A new approach to a theory of sound classification, Interface, 1 (1972) pp. 93-109.
— A minimum description of the linguistic sign repertoire, Interface, part one: 2(1973) pp. 141-156, part two: 3(1974) pp. 137-158.
— Design of musical discovery, A selection, Revised edition, Inst. for Sonology, University of Utrecht, 1985.
— The MIDIM 8 X-System, Desc. user manual, Institute for Sonology, University of Utrecht, 1984.
— Das Problem der mentalen Kodierung von musikalischer Information, Final Report, Schweizerischer Nationalfonds, Gesuch-Nr. 1. 559-0.71, 1975.
— Controlling the Vosim Sound Synthesis System. Interface, Vol. 15, pp. 71-82.
— The Midim Language. Interface, Vol. 15, pp. 83-162.
Kaegi, H., Janssen, J., Goodman, P. (1986). Proceedings of the ICMC, Den Haag.
Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrica 29, 319 (1964).
Ludwig, R., Methoden der Fehler und Ausgleichsrechnung, Uni-text, VIEWEG, 1969.
Martopangrawit, R. L., Pengetahuan karawitan I and II, Akademi Seni Karawitan Indonesia (ASKI), Surakarta, 1975.
English translation: Becker, J. and Feinstein, A. H., Karawitan, source readings in Javanese gamelan and vocal music, vol. 1, Center for South and Southeast Asian studies, Univ. of Michigan, Ann Arbor, USA, 1984.
Nijenhuis, E. te, Dattilam, a compendium of ancient indian music, E. J. Brill, Leiden, 1970.
Rodet, X. et al., The chant project, Computer music journal vol. 8 Nr. 3, pag. 15-31, MIT-press, 1984.
Simon, J.-C. Informatique et reconnaissance des formes, Extrait de la Revue de Defense Nationale, oct. 72.
Supanggah, R., Introduction aux styles et interpretation dans la musique Javanaise, Université de Paris VII, UER et ethnologie, 1985.
Suryodiningrat, W. et al, Tone measurements of outstanding Javanese gamelans in Yogjakarta and Surakarta, Gajah Mada University press, Yogjakarta, 1972.
Tempelaars, S. Testing elements from the Kaegi minimum description of signs from the linguistic repertoire using the Kruskal method, Interface, 2 (1973), pp. 157-158.
— The VOSIM signal spectrum, Interface 6 (1977), pp. 57-72.
— SIGPAC manual, Institute for Sonology, 1972.

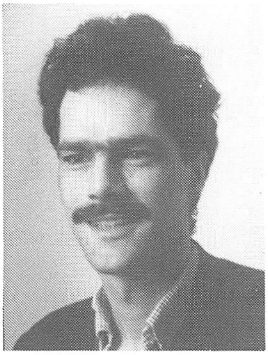
RECOMMENDED

Bevington, P. R. Data reduction and error analysis for the physical sciences, McGraw-Hill book Company, 1969.
Hardjito, P. D. Analisa getaran pada pelarasan gamelan, departemen mesin fakultas teknologi industri, Institut teknologi, Bandung, Indonesia, 1979.
Kaegi, W. Musicians pitch-frequency-cent tables, Institute for Sonology, University of Utrecht, 1984.
Kunst, J. De toonkunst van Java, Martinus Nijhoff, The Hague, 1934.
Martopangrawit, R. L. Titilaras cengkok-cengkok gender dengan wiletannya, Akademi Seni Karawitan Indonesia (ASKI), Surakarta, 1973.
Risdianto, B., Pengukuran spektrum dari nada gamelan, Departemen fisika teknik, Institut teknologi Bandung, Indonesia, 1977.
Tempelaars, S., Signaalverwerking in spraak en muziek, Institute for Sonololy, University of Utrecht, 1985.

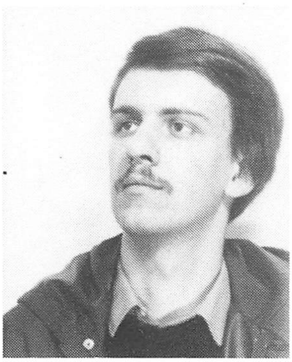
ACKNOWLEDGEMENTS

The authors wish to thank Werner Kaegi for his enlightening discussions by which means the first part of this article could come into being; Paul Goodman for his translation; Marcel Bottema, Frank Wolters and Ben Arps for their critical comments; M. W. Vos for the photos; and Mr. van de Velden for kindly allowing us to make use of his instruments in the museum Bronbeek, Arnhem; Prins K.G.P.H. Praboewidjojo, for his untiring support to realize the numerous sound recordings in the Kraton of Surakarta; and last but not least the masters of Javanese music: Raden Lurah Martopangrawit, Supanggah Rahayu and the ASKI-academy.

While writing this article we received the sad news that R. L. Martopangrawit had died april 13, 1986 (photo taken December 31, 1985 at his home in Surakarta).



JOS JANSSEN, gamelan musician and recording engineer, was born in Arnhem, the Netherlands, 18-5-1953 (The birthday of Russell and Carnap...). After studying at the music school in Arnhem and working as a recording engineer, he became interested in electronic (in particular computer music) and javanese gamelan music. In 1976 he started his studies at the Institute for Sonology and at the embassy of the Republic of Indonesia in the Hague. He has undertaken three study trips to Indonesia, the first in 1980, the second in 1981 and the third in 1985. He has had private lessons at the A.S.K.I.-Academy in Solo with numerous gamelan teachers and most notably with R. L. Martopangrawit with whom he has held many interviews. In 1980 he applied measurements to various genders in the Kraton of Surakarta. Since 1979 he has been involved in a gamelan project under the mentorship of Dr. Werner Kaegi. Together they have set up a special gamelan compiler for the MIDIM-system, the outcome of which has been shown via compositions, concerts and publications. Most notably the MIDIM concert of associative computer music, Utrecht, Jan. 1984 and the first concert of computer music at the A.S.K.I.-Academy in Solo, Java, Dec. 1985. He has been an assistant of Dr. Werner Kaegi since 1983 and together with Paul Goodman gave a MIDIM workshop at the Institute for Sonology for two years. He is a member of the MIDIM-Group.



HEINERICH KAEGI was born in 1961 (an extremely good vintage year) in Switzerland where he spent his earliest years. In 1971 he emigrated along with his parents to Holland. At the moment he is studying physics at the University of Utrecht and is specialised in the foundations of this science. His interests extend to the relations existing between the arts (in its widest sense), technology and science. For this reason aside from his academic studies he is intensively busy in the fields of theatre (acting, directing), literature, language and music (in particular sonology). He has a special love for nature whose fecundity inspires and forms a reference point for all artistic activity. As a member of the MIDIM-group he has taken an active part in the artistic experiments as well as the scientific research of the group.

Interface, Vol. 15 (1986), pp. 231-256. 0303-3902/86/1501-0231 $3.00

Notes by a Film-Maker

Pierre van Berkel

ABSTRACT

For a number of years the author has been occupied with the investigation of movements of the mouth and the accompanying speech sounds of Dutch interjections with the purpose of synthesizing with a computer the sounds and images for a video film of "speaking heads". In this article some aspects of speech synthesis are touched upon which can be realized with the help of the MIDIM/VOSIM-system. The phenomena of the 'auditive pause', the so-called 'joint' for bridging quick and strong formant springs and the analysis and synthesis (via D-comprehensors) of vowels in the interjections are also shown.

INTRODUCTION

My motive for working with the MIDIM-system is the desire to be able to recreate speech sounds and to place them in another context in the same way a painter recreates elements from his environment by means of drawings and then takes them as a basis for a painting. I was able to realize this with the help of the MIDIM-system in my sound poems "Tjaktjakai" (1983) and "Heegothe" (1984). At the moment I am working on a video-film which is in its entirety computer synthesized and in which 'speakers' visible on the screen will recite in short episodes sound poems composed from interjections stemming from the Dutch language. (In English for example "wow!".) The realization of this intention demands on the one hand a thorough knowledge of speech sounds and on the other hand of the movements of the mouth corresponding to the particular articulations. For this reason I recorded three male and three female voices on film and video. I studied on a montage table the articulation movements as well as the corresponding speech sounds after application of fourier transform. From the first a pseudo-3-dimensional synthesis of heads resulted and from the second synthetic MIDIM speech sounds were derived.

In order to give the reader some idea of the synthesis of images which will not be discussed further there is shown, in Figure 1, one of my "speaking heads". On the contrary, in what follows in this article, various aspects of speech synthesis shall be dealt with.

In part one I will briefly show the striking phenomenon that there can exist an apparent "asynchronicity" between speech sounds and the movements of the mouth. Part two describes how, with the MIDIM-system, I realized extremely quick phoneme transitions preceding from out fourier analyses of the sound

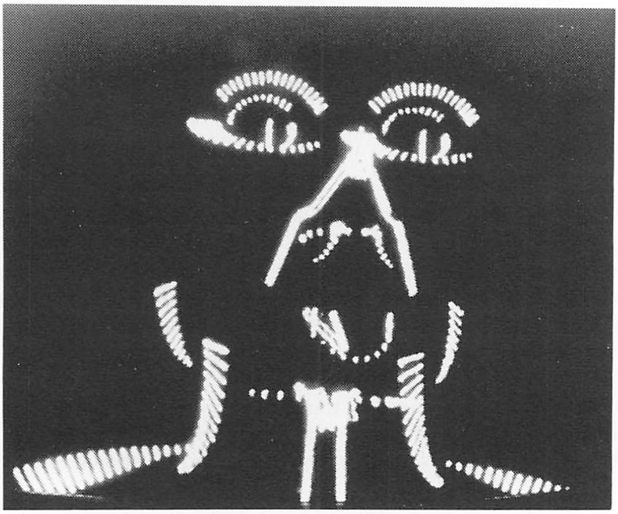


Figure 1. One of the "speaking heads", from the videofilm.

registrations. In part three is presented the analysis and synthesis of vowels and an artistic application of these. An appendix containing a list of phonetic symbols is supplied for the reason that repeated use is made of these symbols in the course of the article; thus readers who are not familiar with them are directed towards Appendix I.

1. AUDITIVE PAUSE

As was sketched in the introduction I did not only investigate speech sounds and the movements of the mouth but also the relationships existing between the two. Repeatedly I observed that in speech a large changement in the mouth could take place while there was no sound produced. Just such an auditive pause (with a duration of 120 ms) is found in /Xɔt/ during the transition from /ɔ/ to /t/. The corresponding image and sound registration is shown in Figure 2.

2. QUICK TRANSITIONS IN SPEECH SOUNDS

Kaegi has demonstrated that it is possible to synthesize speech signals with the sum of two VOSIM signals. The maximum of the spectrum which corresponds with the period duration *T* can be interpreted as a speech formant. (See Kaegi, 1986, p. 77). Although it is designed for musical application, the MIDIM-system controlling the VOSIM-generator lends itself extremely well for speech synthesis. (A number of functions have been introduced especially for this aim.)

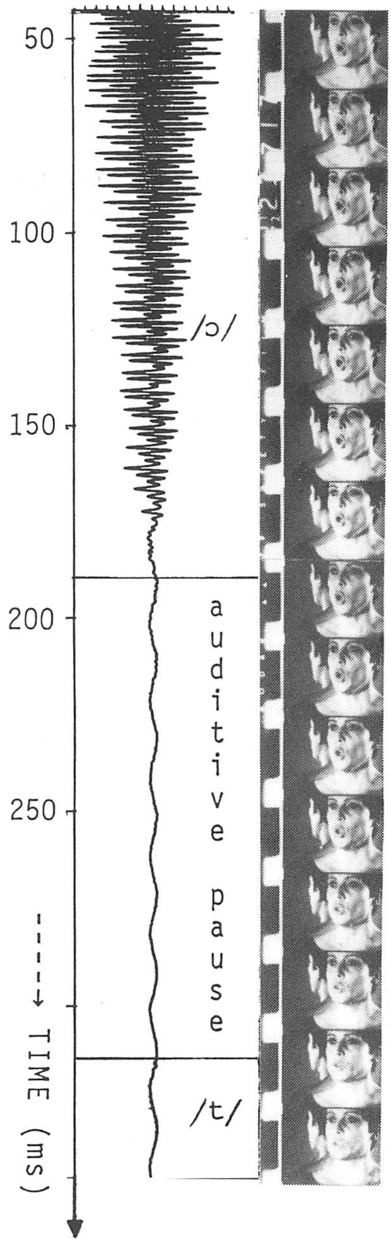


Figure 2. Left: Fragment of the signal function of the compound /Xɔt/.
Right: The corresponding movements of the face of the speaker. (The speed of the camera during shooting was 50 frames per second, twice the standard speed of 25 frames per second.) The changement in the position of the mouth is clearly seen during the auditive pause.

2.1 Limitations in the application of formant-shifting and portamento
Speech synthesis demands high standards from a sound synthesis system because (among other reasons) very strong changements are necessary within often very short time intervals. This may involve the problems discussed in what follows.

The VOSIM-signal is built up from a $\sin^2$ pulse which is repeated N times with a pulse width T followed by a delay M. One period of the signal has the length:

$$T'(t) = N(t) \cdot T(t) + M(t) \qquad [1]$$

The whole period is repeated Np times. Within this sequence T and M can be interpolated linearly from T to $T+DT$ resp. from M to $M+DM$. If we shift $M(t)$ in [1] to the left of the equation and call the time point where the segment begins t_i and the time point where it ends t_e than it holds that:

$$\text{Begin} \quad M(t_i) = T' - N \cdot T \qquad [2]$$

$$\text{End} \quad M(t_e) = (T' + DT') - N \cdot (T + DT) \qquad [3]$$

in which $T' = T'(t_i)$, $T = T(t_i)$, $T' + DT' = T'(t_e)$ and $T + DT = T(t_e)$. There is thus a gradual changement of T and T', while N remains constant:

$$N(t) = \text{constant for } t \in [t_i, t_e] \qquad [4]$$

At the end of the segment N can spring to a different value. This stepwise changement of N is naturally connected with the definition of the VOSIM-signal.

The reason that one chooses various N values in the various successive segments is on the one hand because the VOSIM-spectrum changes strongly with variations of N (above all for low values of n and with the transition from N = even to N = odd) and on the other hand because $M(t)$ is connected with a particular domain: (See Kaegi, 1986, p. 106)

a) $M(t) > 0$ as the fundamental limitation of the VOSIM-model. (The value zero is excluded in order to prevent the VOSIM-signal from becoming a sine).

b) $M(t) < M\text{max} = 4095$ μs as technical limitation of the VOSIM7-system. (4095 corresponds with 12 bits. In the new MIDIM9-system 13 bits shall be available).

The variations T and T' corresponding with a shifting of the VOSIM-formant and the harmonics within one segment are via formula [1] connected to these limitations. Certain interpolations are thus impossible.

The following two cases of interpolation can cause limitation problems:

1) $\Delta T > 0$ and $\Delta T' < 0$, which means a decline of the formant and an increase of the fundamental of the frequency, so that $M(t) < 0$.

2) $\Delta T < 0$ and $\Delta T' > 0$, which means an increase of the formant and a decrease of the fundamental in frequency, so that $M(t) > 4095\ \mu s$.

The limitations which have just been sketched are reached above all in speech synthesis because one is here strongly tied to particular formant values and values for the fundamental.

2.2 A Numerical Example and the MIDIM-function G8

In order to avoid that $M(t)$ should overreach the described limits the function G8 was introduced into the MIDIM-system, which for certain values of T', $\Delta T'$, T and ΔT investigates if $M(t)$ overreaches the limits and if necessary adapts the value of N. (See Kaegi, 1986, p. 97). We will illustrate the problem by means of Figure 3 in which the following transitions are indicated with the help of arrows: (We here present the principle but not the exact way in which G8 applies corrections)

	T'	T	$\Delta T'$	ΔT	N	$M(t_i)$	$M(t_e)$
1.	4,0	1,0	2,0	0	1	3,0	5,0> *M*max
2.	4,0	1,0	2,0	0	3	1,0	3,0
3.	4,0	1,0	1,1	0	1	3,0	4,1
4.	5,1	1,0	1,0	1,0	1	4,0	4,1= *M*max
5.	4,0	1,0	—2,0	0	3	1,0	—1,0< *M*max
6.	4,0	1,0	—2,0	0	1	3,0	1,0

We start with the initial values (ms) in case one. The increment is not possible, because $M(t_e) > M$max. We can now change N from 1 to 3 (one may not step over to 2, because this is an even number, see Kaegi, 1986, p. 97). With $N = 3$ increments of T' and T are permissable as we can see in the diagram (case 2). Another solution is to split the segment in two and as a first step to interpolate until the boundary of the limit (case 3). Afterwards we may continue the increment along the value $M(t) = M$max (case 4). Also the other limit where $M(t) = 0$ can be overreached, see case 5 and compare it with case 6.

When it would be possible to change N gradually, as is the case for T' and T, and there would exist a ΔN than we could realize cases 7 and 8: (see Fig. 3)

	T'	T	$\Delta T'$	ΔT	N	ΔN	$M(t_i)$	$M(t_e)$
7.	4,0	1,0	2,0	0	1	2	3,0	3,0
8.	4,0	1,0	—2,0	0	3	—2	1,0	1,0

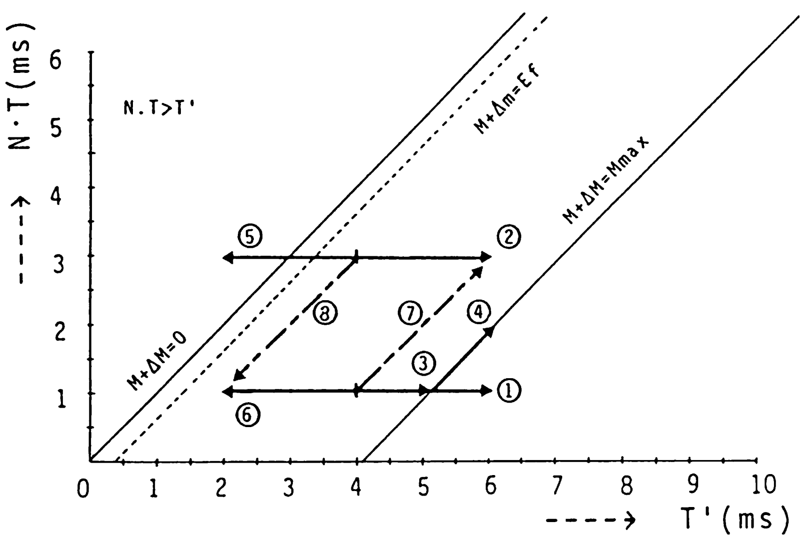


Figure 3. The range for ΔT and $\Delta T'$ and some examples of interpolations indicated by the numbered arrows. For a given initial state (T', $N \cdot i$) the final state ($T' + \Delta T', N \cdot T + N \cdot \Delta T$) has to lie between the boundary lines $M + \Delta M = 0$ and $M + \Delta M = M\text{max}$.

2.3 Simulation of an interpolation of *N* within the segment

In order to simulate a gradual interpolation of *N*, I took the following course. In place of an interpolation between a start and end position within a particular segment I divided the start and end position over two tracks so that I could overlap the two positions within one segment. By allowing the amplitude of the start position on a first track to decrease until zero and the amplitude of the end position to increase to the demanded value the transition could be simulated excellently. In analogy with the terminology used in film I have termed the transition segment a 'dissolve' and the decaying and increasing components resp. a 'fade-out' and 'fade-in'. A disadvantage of the dissolve can be that one has two tracks necessary for a transition. In practise this limitation gives few problems because the segments are very short in duration (e.g. 20 ms).

2.4 The Joint

Experiments have shown that in speech synthesis it is often sensible to make use of not only one dissolve but to join a number of dissolves into a series. This series of dissolves I term a 'joint'. As we shall show in the example which follows, the joint can be rather easily derived from the representation of the spectral information which I used and received from the analysis of the registered sounds. As I applied the joint it forms usually the transition from a consonant to a vowel.

2.5 Rules for Forming the Joint within the MIDIM-Language

Application of the dissolve demands use of two synchronized VOSIM-signals which are in the MIDIM-predicator given the labels track 1 and 2. We assume for simplicity that every VOSIM-signal produces a formant which is determined by the pulse width T for the two tracks and are called resp. $T1$ and $T2$. (In reality the VOSIM-formant is the second maximum in the one-shot spectrum and thus forms the contour of the harmonic which are set by T'; see Kaegi, 1986, p. 74 ff). It is now possible to construct a joint, namely a series of dissolves D_i with $i = 1$ until n (with particular values for $(T1)_i$ and $(T2)_i$ and the corresponding amplitudes and amplitude changements $(A1)_i$, $(DA1)_i$, $(A2)_i$, $(DA2)_i$) with the observance of the following rules:

1) In each dissolve one track describes the fade-in and the other the fade-out.
2) In the first dissolve D_1 of the series the amplitude and the formant of the fade-out connect onto the track in the previous segment which ended on an amplitude inequal to zero. The other track must end with an amplitude of zero.
3) A fade-in in D_i is followed in dissolve D_{i+1} by a fade-out, such that the amplitude and the formant connect. The fade-out in D_i is followed in disolve D_{i+1} by a fade-in. These do not need to connect. (Of course it does not make any sense to connect the formants in this case.)
4) The segment after the last dissolve D_n of the series must connect in one track with the fade-in of D_n. The other track has an amplitude zero.

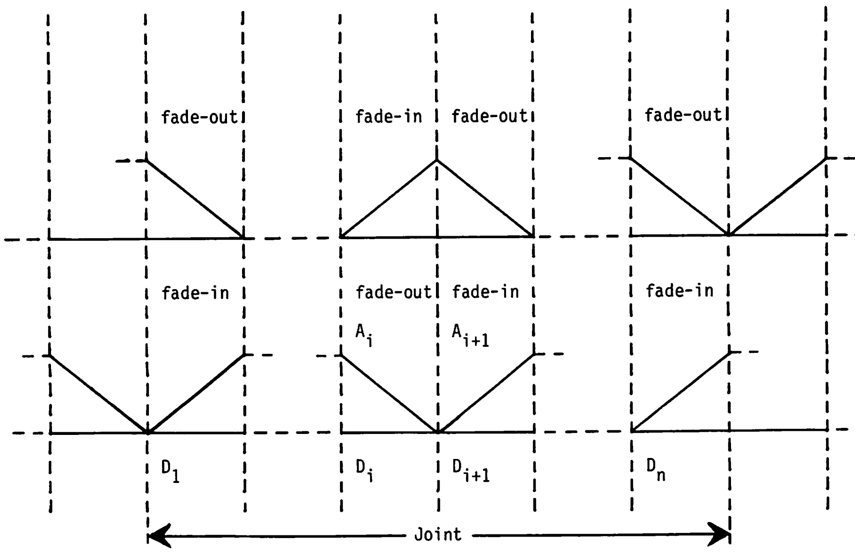


Figure 4. Schematic diagram of the amplitude contour within a joint. (For simplicity we have chosen equal amplitudes in all segments, but they may vary.)

Before we show an application of the joint, fourier analysis, -abstractions and -representations which I employed shall be discussed.

2.6 Fourier Analysis and Abstractions

Investigations into sound signals take place by means of (1) signal functions and (2) spectra. In order to be able to make a selection of relevant information from the flood of information one receives, abstractions are needed. The most important of these abstractions being for us the amplitude envelope as an abstractor of the signal function and the peak-track as an abstractor within the pseudo-3-dimensional fourier representation. Both abstractors are based upon the connecting of maxima in the amplitude. They form the first step in data reduction.

In Figure 5, a signal function (top) and various abstractions and transformations are shown. The amplitude envelope (middle) are determined by connecting

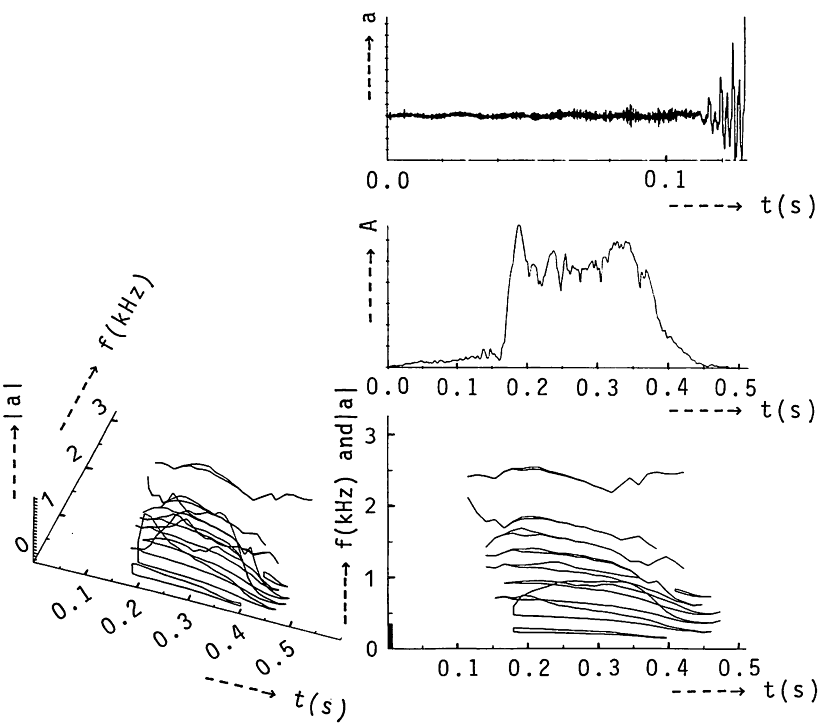


Figure 5. Some different representations of sound signals.
Top: A fragment of the signal function of the compound /so/;
Middle: Amplitude envelope of the whole signal;
Bottom
left: Pseudo 3-dimensional representation of the running spectrum of the same signal;
right: The same spectral abstraction in a 2-dimensional representation.

certain maxima in the signal with one another. The signal function and its envelope give us a general impression of the changements the signal undergoes in the course of time.

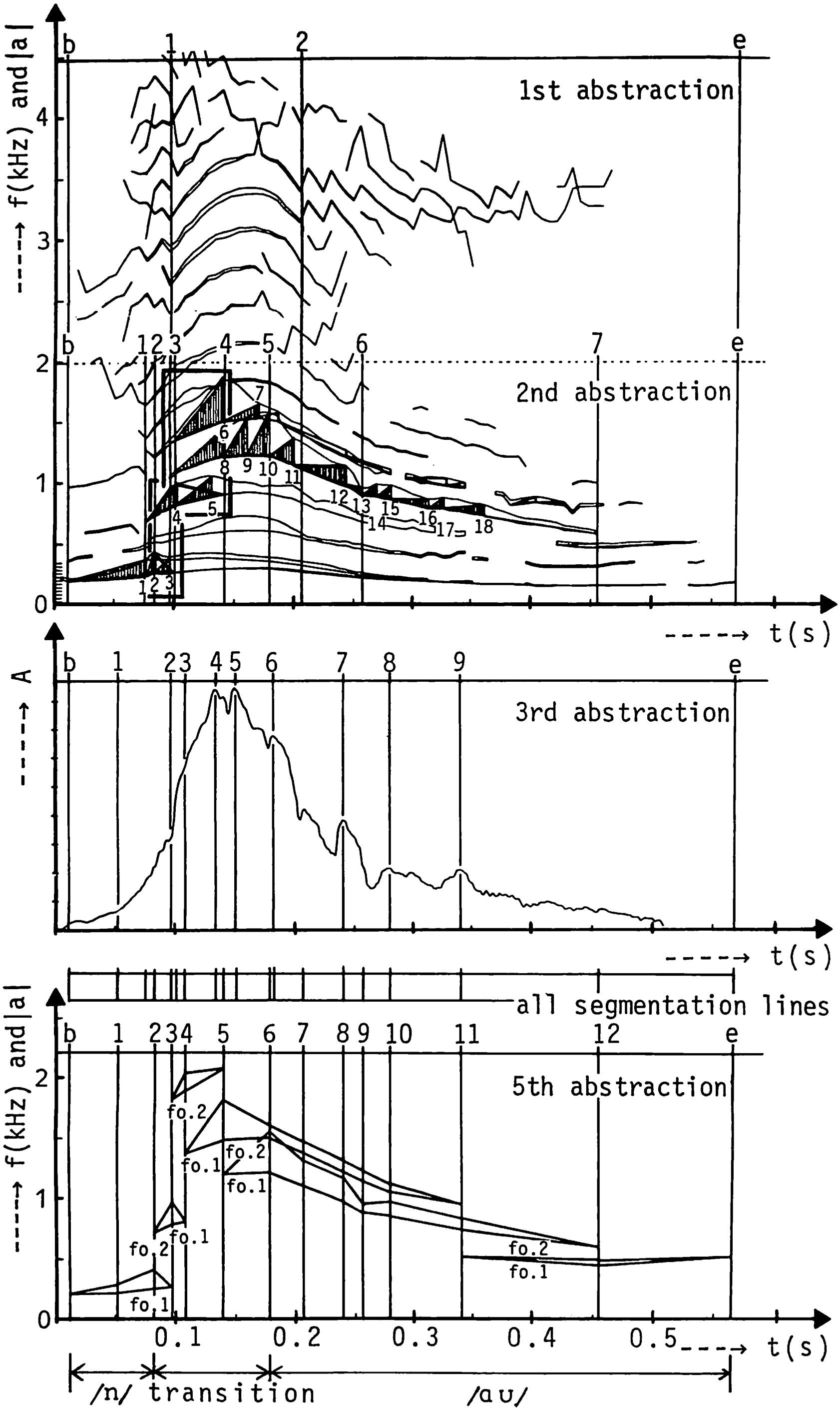
Information concerning the spectral characteristics must be abstracted from the fourier transform of the signal in time. The so called FFT-algorithm (fast fourier transform) transforms the signal per small time segments and gives a running spectrum with the dimensions frequency, amplitude (of the coefficients) and time. The spectral information is usually plotted in a pseudo-3-dimensional representation. In this case as well it is preferable to abstract the information directly by means of a relevant abstractor. Peak-tracking is a much applied abstractor which connects the amplitude maxima in the running spectrum after which the tracks can be rather clearly plotted. For ease a graphic representation of these tracks will be called 'the spectrum'.

With speech sounds the pseudo-3-dimensional representation was not satisfactory because the spectra contained too many tracks which lay in the neighbourhood of the others and thus overshadowed them (see Fig. 5 bottom, left). What is more, it was difficult to carefully measure the track as a result of the perspective distortion of the axes. I therefore developed a representation wherein the axes frequency and amplitude fall together (see Fig. 5, bottom, right). The tracks are plotted as closed figures of which the basis gives the frequency (the frequency line) and the contour the superposition of the frequency and amplitude (the amplitude line). (Actually this is a projection of the tracks onto the plane containing the time axis and a very small declivity in relation to the frequency axis. Choice of the correct angles in the existing projection algorithms gives us directly this '2-dimensional representation'.) In this way it is possible by means of a suitable choice for the amplitude and frequency scaling to plot clearly rather complicated track formations, which can above all be graphically measured. In what follows use is exclusively made of this representation for the spectra.

The abstractors described are actually not powerful enough to make a translation directly from the analyses into the MIDIM/VOSIM-parameters so additional abstractions are necessary. Seen the fact that I concentrated upon sound synthesis with the help of the MIDIM8X-system I chose the abstractions in such a way that a translation of the abstracted information to my joints could be formulated within this language.

2.7 Analysis of the Speech Sound /nɑu/

As an example of the procedure which I followed and application of the joints I have chosen a registration of the phoneme series /nɑu/ spoken by a high female voice. The signal was read into the computer from tape via a A/D convertor, was fourier transformed, abstracted and finally represented in the 2-dimensional



f(kHz) and |a|
1st abstraction
2nd abstraction
A
3rd abstraction
all segmentation lines
5th abstraction
fo.1
fo.2
t(s)
/n/ transition /aʊ/

manner already described. The spectrum is shown together with the contour of the signal function in Figure 6. By means of this figure the abstractions which I further applied will be explained.

The first important step is the segmentation in time. I applied the following abstractions in order to detect as small a number of segments as possible. The segments are delimited by the so-called segment lines which we number per abstraction. The segments are given by the numbers of the delimited segment lines.

Start and end: The first segment line b gives the start of the signal. The segmentation is closed with the end line e which gives the end of the signal.

Abstractions:

1st Abstraction — one determines in a rough way the most important evident segment lines in the spectrum. This gives three segments: b-1, 1-2 and 2-e. (Compare the graphic pattern within the three chosen segments!).

2nd Abstraction — one detects per track strong changements in the course of values for the frequency and amplitude of the most dominant peak tracks and gives these graphically with a vertical line which we call L_i and which connects the frequency and the amplitude line. Index i gives the number of the line calculated from off $i = 1$ in the direction of increasing time. The intersection of line L_i with the frequency line of the track we call SF_i and the intersection with the amplitude line of the track is called SA_i. The first intersection SF_o gives per definition the start of the track concerned. One con-

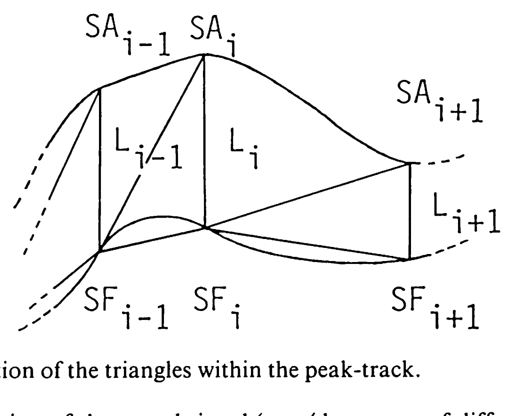


Figure 7. Construction of the triangles within the peak-track.

Figure 6. Segmentation of the sound signal (nau/ by means of different abstractions. ← Abstractions applied to:
Top: the running spectrum (2-dim. representation); Middle: the amplitude envelope. At the bottom the derived result by means of VOSIM-formants is indicated (in the 2-dim. spectral representation). In the transition /n/→/au/ the joint method is applied.

nects now for all i the points SF_{i-1} with SF_i by means of a straight line; one connects for all i the points SF_{i-1} with SA_i. In this manner triangles are constructed which, as we shall show later, can be used in the construction of the joint. The result is seen in Figure 6 (top). The segment lines in Figure 6 give a selection of the most important lines L_i. (See figure 7.)

3rd Abstraction — one detects important changements in the amplitude contour of the signal. See Fig. 6, middle. The changements before the absolute maximum are of more importance than later. After this maximum only relative maxima are chosen for segmentation.

4th Abstraction — This abstraction is passed over in the example. It differentiates noise components in the signal which in certain speech sounds are of high importance. In /nau/ they are hardly present.

Standing clearly shown under Figure 6 are all segment lines which have been found by means of the described abstractions. The last and 5th abstraction selects from these segment lines the most important by making a comparison between the segment lines found in the previous abstractions. Moreover the course of the VOSIM-formants and the amplitudes which are necessary for a MIDIM-duplication is determined.

2.8 Determination of the Necessary VOSIM-Formants and the Joint

Proceeding from the segmentation found and the triangles in the spectrum we seek a series of values for T, ΔT, A and ΔA for both tracks. The following rules therefore hold, whereby we begin with $t = 0$:

1) The segments with only single triangles can be described by one track. The basis of the triangle gives the values for T and ΔT. The rising line gives the values for A and ΔA.

2) Segments with more than one triangle form a dissolve. One chooses the two most important triangles and 'turns' one of these around. According to Figure 7 one does not connect SF_{i-1} with SA_i but SF_i with SA_{i-1}. The reversed triangle describes the fade-out, the other the fade-in. The values for the formants and amplitudes are now easily determined. The dissolve becomes in the described way very clear. See Fig. 8. (This is a portion from Fig. 6, top).

3) In Figure 6, bottom, we find the resulting succession of VOSIM-formants which are shown with *fo*.1 and *fo*.2. The joint runs from segment line 2 until 6 and forms the transition between the /n/ and the /au/. A schematic overview of the four dissolves in the joint is given in Fig. 9.

In the interjections upon which I have directed my work the joint (generally) stands at the transition from the consonant to the vowel.

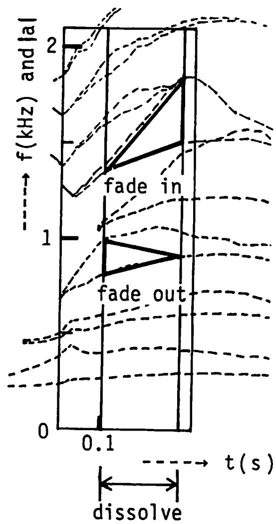


Figure 8. A close-up taken out from Figure 6, top which shows a dissolve within the applied joint.

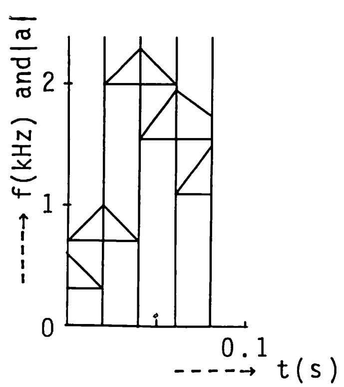


Figure 9. Schematical diagram of the joint applied for the transition /n/ → /ɑu/.

2.9 MIDIM-description of the /nau/

In the example presented we have only examined one sound event. Other analyses led to the following MIDIM-description of /nau/.

We start with the following vectors formulated by Kaegi, 1986, p. 123:

$$\vec{m}_{850} = \lambda S, Ef, C, Of, W, Sp, e, y_2, \Delta Am, Am, Nmax, P, FS, T, T', At, d. ()_{850}$$

From this vector we derive the following vector:

$$\vec{m}' = \vec{m}_{850} \quad \begin{pmatrix} c_S, & c_{Ef}, & c_C, & c_W, & c_{Sp}, & c_e, & c_{y_2} \\ 1 & 20 & 0.75 & 0 & / & 1 & 1 \end{pmatrix}$$

$$\vec{m} = (\lambda\chi_4)\ldots(\vec{m}')_i$$

We define the following concepts:

$Syl_{11} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
3608	—154	—100	2	0	15	1
3300	—243	—250	2	15	335	2

$Syl_{12} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
2456	—153	—100	2	0	100	1
2248	—244	—250	2	100	—50	2

$Syl_{21} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
2864	—149	—150	8	350	—350	1
568	—126	—150	8	0	300	2
528	—80	—150	8	300	—300	3
828	—8	0	8	0	500	4

$Syl_{22} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
1364	—154	—150	8	0	300	1
1248	—109	—150	8	300	—300	2
720	—99	—150	8	0	500	3
680	—62	0	8	500	—500	4

$Syl_{31} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
816	0	—100	6	0	500	1
816	0	0	6	500	0	2
816	848	600	6	500	—350	3
1332	846	300	6	150	—150	4

$Syl_{32} = ((C'))(\vec{m})$

(c_T,	c_{FS},	c_P,	c_{Nmax},	c_{Am},	$c_{\Delta Am}$))	i
652	0	—100	6	500	—300	1
652	0	0	6	200	0	2
652	848	600	6	200	—200	3
1064	843	300	6	0	0	4

We link these concepts in the following predicators.

For the /n/ we form:

$$P(2)_{1a} = ((\lambda L_1 L_2 L_{31} \phi_5)\, v, d_1, d_2, DUR, T', At\, .\, (Syl_{11}, Syl_{12}))\, \underset{40,\ 2}{(c_{d_1}, c_v)}$$

$$P(0)_{1b} = ((\lambda L_2 L_{31})\, v, d_1, d_2, T', At\, .\, (Syl_{11}, Syl_{12}))\, \underset{40,\ 30,\ 0}{(c_{d_2}, c_{d_1}, c_v)}$$

The joint is described by:

$$P(0)_2 = ((\lambda L_2 L_{31})\, v, d_1, d_2, d_3, d_4, T', At\, .\, (Syl_{21}, Syl_{22}))\, \underset{40,\ 30,\ 12,\ 18,\ 0}{(c_{d_4}, c_{d_3}, c_{d_2}, c_{d_1}, c_v)}$$

Finally the /au/ is described by:

$$P(2)_3 = ((\lambda L_1 L_2 L_{31} \phi_5)\, v, d_1, d_2, d_3, d_4, DUR, T', At\, .\, (Syl_{31}, Syl_{32}))$$

$$\underset{220,\ 180,\ 20,\ 2}{(c_{d_4}, c_{d_3}, c_{d_1}, c_v)}$$

The compound may be constructed by a *D*-comprehensor which defines the recursive sequence:

$P(c_v)_n$:

- Start: $n_1 = 1a$ (speaking)
  $n_1 = 1b$ (sound poetry)
- Recursive sequence: $n_{i+1} = n_i + 1$

For the instance-duplication of the analysed /nau/ See par. 2.7) holds:

	T'(us)	DUR(ms)	At
$P(0)_{1b}$	4950	—	0
$P(0)_2$	4292	—	0
$P(2)_3$	3610	380	0

To vary the pitch of the compound /nau/ we apply the transposition rule *Dtra* (see Kaegi, 1986, p. 152), which transposes all T' values within a certain amount of cents.

The transposition rate, if the characteristics of the voice must be kept, equals —400 and +400 cent, and if the phonetic sign must remain recognizable equals —1850 and +400 cents. (By application of predicator $P(2)_{1a}$ the same values can be used.)

The variation of duration DUR has to lie in the following domain:

$$DUR(\text{ms}) \in [0, 1200]$$

In the special case of predicator $P(2)_{1a}$ the following values for DUR are suitable:

$$DUR(\text{ms}) \in [20, 500]$$

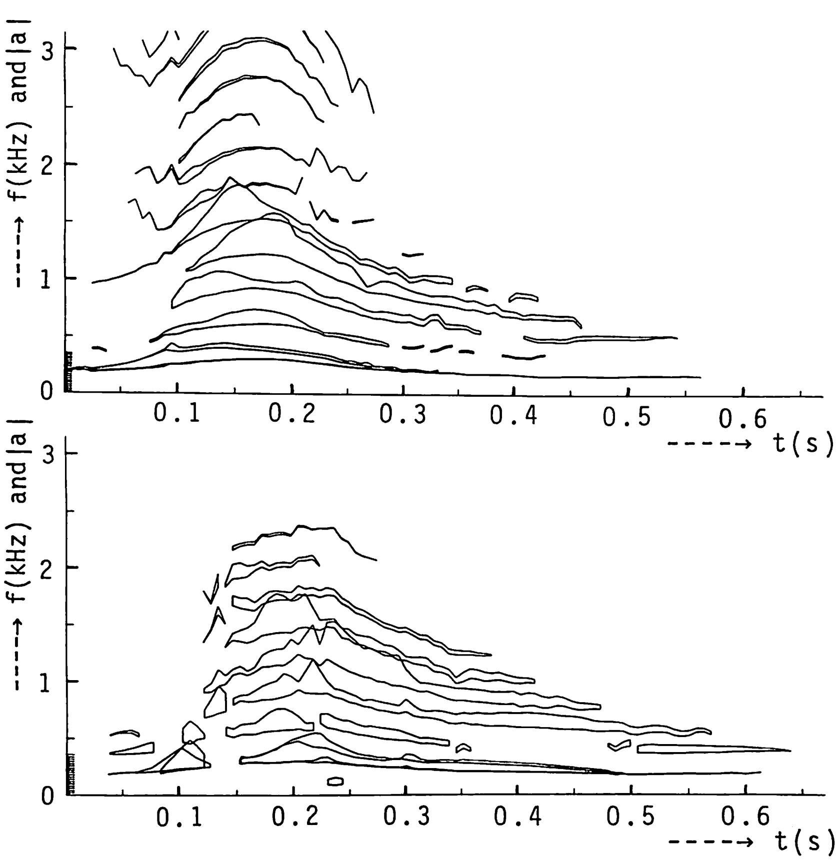


Figure 10. The spectra of the compound /nau/. On top the spectrum of the original registrated signal (spoken by a high female voice) and at the bottom its MIDIM-duplication.

In Fig. 10 the spectra of the original recording and the MIDIM-duplication are shown. We note that after testing the seven primitive segmentations of /au/ within the compound /nau/ are reduced to two. (In the available sound examples one can hear the difference by reducing the number of segments, see references).

3. DERIVATION OF VOWELS FOR AN ARTISTIC APPLICATION

As was stated in the introduction I used as the basis material for my soundtrack sound registrations of Dutch interjections, spoken by six persons, three women and three men. Long before the recordings the speakers were trained to clearly articulate a series of interjections which for completeness is given in Appendix II.

During the recordings the speakers read the interjections (which were notated according to the given standard notation (see Appendix I)) from paper. We call these interjections which served as the starting point the descriptive concepts. The second step was fourier analysis, abstractions and the 2-dimensional representation of the spectra received in the way described earlier (see 2.6). In a third phase characteristic parameters for the sounds were derived from the peak tracks in the spectra. Among others I have determined the frequency and amplitude of the fundamental on the one hand and on the other hand the two dominant formants for the vowels /a/, /ɑ/, /ɔ/, /o/, /ɛ/, /e/ and /i/.

3.1 The Six Voices

The frequencies arrived at for the fundamentals per speaker were averaged and are valid within the segment of the vowel (in most cases directly after the transition of the consonant to the vowel). The results are shown in the table below, expressed in Hz. and in microseconds.

Fundamental		Hz	μs
Female	high	311	3215
	middle	220	4545
	low	185	5405
Male	high	262	3817
	middle	156	6427
	low	131	7635

The accuracy of the measurements was ± 7 Hz = ± 150 ms.These givens confirm a provisional division into high *(h)*, middle *(m)* and low *(l)* female voice *(F)* resp. male voice *(M)*. (We denote them in a further abbreviated form as for example *Fm* = female medium). This division is systematically used in the rest of the article.

3.2 The Measured Vowel Formants

The abstracted vowel formants were averaged per voice and per descriptive vowel concept and are reproduced in Figure 11. Represented horizontally is the first formant fo. 1, vertically the second formant fo. 2 in Hz. The various voices are given with corresponding symbols. The accuracy of the measurements and abstractions amounts to ± 2 Hz and is on account of the clarity of the graph not drawn in. The measurement data which according to the experiment belongs to a descriptive vowel concept are connected with broken lines.

It is striking that the various areas of the different vowels sometimes lie close to the limits of each other, while the various vowels per voice lie so far away from one another that they remain differentiateable to the ear. It would appear that the absolute differentiatibility of the vowels is of less importance in this experiment than the relative differentiatibility per voice. (Compare for a voice the series /a/, /ɑ/, /ɔ/, /o/, /ɛ/, /e/, /i/.)

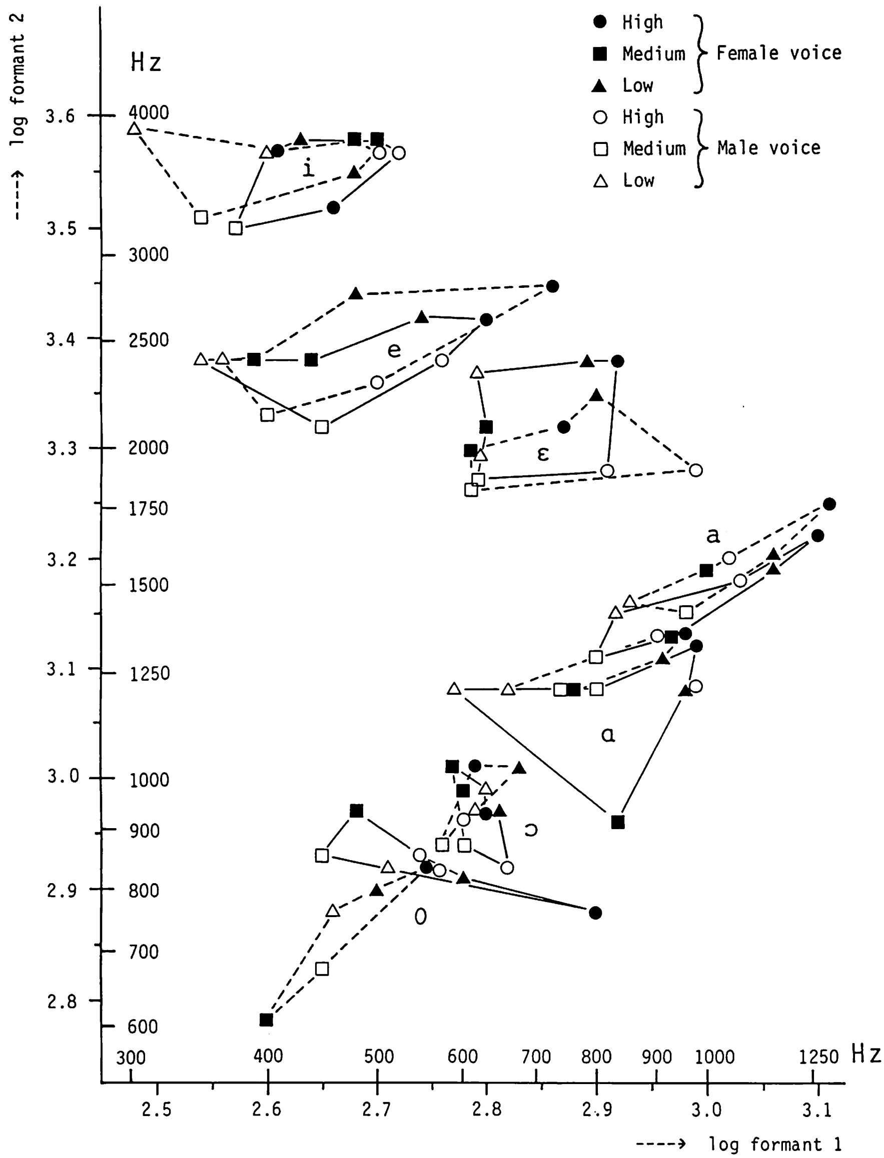


Figure 11. Vowel formants: hortizontal the first formant and vertical the second formant in a ^{10}log representation. The measuring values which were determined from out the spectral analyses are connected with unbroken lines. The values derived from out these for the soundtrack are connected with dotted lines. In the last case the values denote the VOSIM-formants.

If we compare the measurements with the pertaining literature than it appears that there are reasonable correspondences between them. In the case of the /ε/, /e/, and /i/ mine lie somewhat higher for the second formant. This can occur because of the rather artificial stylization of the spoken interjections which I desired for my sound track. (Kaegi, 1986, p. 123, Pols, 77).

3.3 Derived Vowels for an Artistic Application

For the realization of my sound track with the MIDIM-system I proceeded from out the measurements shown. First I translated the measurements into MIDIM-parameters. For this I assumed for simplicity that the necessary VOSIM-formants had the same value as the measured formants fo. 1 and fo. 2. Thus $F_{\Omega 1}$ = fo. 1 and $F_{\Omega 2}$ = fo. 2. By means of a proper choice of the other parameters and application of the correct MIDIM-concepts vowel duplications were arrived at. From these I formulated derivations after listening to the vowel series per voice such that the synthetic vowels fulfilled my wishes for the sound track. The VOSIM-formants arrived at in this way are shown in Figure 11 and connected with dotted lines. As with the measurement data it holds that the various vowel concepts are only sharply differentiateable in the total vowel series of a voice.

3.4 MIDIM Formalization of the Derived Vowels

Our vector for spoken vowels is (according to Kaegi, 1986, p. 123):

$$\vec{m} = (\vec{m}_{853})\,(c_T, c_{Am}, c_{\Delta Am}, c_{Nmax}, c_P, c_{FS}, c_d, \underset{1}{c_e}, \underset{1}{c_{y2}}, \underset{6}{c_{Sp}}, \underset{20}{c_W})$$

The corresponding constants per segment for the formation of concepts are (*i* is the index of the segments):

segment *i*	c_{Am},	$c_{\Delta Am}$,	c_P,	c_{FS},	c_d
1	0	0	0	0	20 or 30
	0	c_a	—100	—100	20 or 30
2	0	c_{Am1}	100	100	20 or 30
	c_a	c_{Am2}—c_a	100	100	20 or 30
3	c_{Am1}	—½c_{Am1}	100	0	ϕ_5
	c_{Am2}	—½c_{Am2}	100	0	ϕ_5
4	½c_{Am1}	—½c_{Am1}	100	0	20 or 30
	½c_{Am2}	—½c_{Am2}	100	0	20 or 30

We apply the duration c_d of 20 ms for *Fh*, *Mh* and *Fm*, and 30 ms for *Fl*, *Ml* and *Mm*.

For the vowels /a, ɑ, ɔ, o, ε/ it holds that: c_a = 500.

For the other vowels the c_a is dependent upon the voice:

	Fh	*Fm*	*Fl*	*Mh*	*Mm*	*Ml*
/e/:	500	400	200	500	500	250
/i/:	240	150	200	200	120	200

The value for the formants, the amplitudes and c_{Nmax} are as follows:

High female voice			Medium female voice			Low female voice			
c_T,	c_{Am},	c_{Nmax}	c_T,	c_{Am},	c_{Nmax}	c_T,	c_{Am},	c_{Nmax}	vowel
768	500	4	1000	500	4	868	500	6	/a/
560	300	6	640	250	8	624	300	10	
1052	500	2	1332	500	4	1108	500	6	/ɑ/
740	300	4	832	250	6	784	400	8	
1612	500	2	1664	500	2	1468	500	4	/ɔ/
980	250	4	1020	250	4	980	100	6	
1784	500	2	2500	500	2	2000	500	2	/o/
1196	160	2	1664	250	2	1248	100	4	
1348	500	2	1612	500	2	1248	500	4	/ε/
476	200	8	500	60	10	444	250	14	
1388	500	2	2380	500	2	2080	500	2	/e/
356	30	10	416	50	12	360	20	18	
2480	500	2	2080	500	2	2080	500	2	/i/
272	25	14	264	20	20	284	30	22	

High male voice			Medium male voice			Low male voice			
c_T,	c_{Am},	c_{Nmax}	c_T,	c_{Am},	c_{Nmax}	c_T,	c_{Am},	c_{Nmax}	vowel
952	500	4	1052	500	6	1188	500	12	/a/
624	400	6	716	250	10	692	300	22	
1108	500	4	1348	500	6	1512	500	10	/ɑ/
744	250	6	832	400	8	832	200	18	
1664	250	2	1724	500	4	1612	500	10	/ɔ/
1108	500	4	1148	250	6	1084	250	14	
1784	500	2	2220	500	2	2172	500	6	/o/
1192	100	4	1468	160	4	1312	300	12	
1220	500	4	1612	500	4	1612	500	10	/ε/
528	250	8	548	250	14	400	50	14	
2000	500	2	2500	500	2	2776	500	4	/e/
436	50	10	464	30	16	432	30	36	
2000	500	2	2856	500	2	3332	500	4	/i/
268	40	16	312	25	24	260	30	62	

The pitch (or T'—) domains in μs are as follows:

		HIGH	MEDIUM	LOW
Female voices	a.	[2551, 4049]	[3610, 5727]	[4292, 6812]
	b.	[2408, 6810]	[3822, 8099]	[3608, 10811]
	(for /i/	[2756, 6810])		
Male voices	a.	[3034. 4808]	[5102, 8097]	[6068, 9634]
	b.	[2551, 8581]	[4816, 16198]	[6067, 20408]

In case a. the characteristics (viewed according to the speech pitch) of the specific voice are preserved with variations of the pitch over the whole domain. In case b. only the characteristics of the phonetic sign remain recognizable. (In my artistic application for many vowels the domain will be larger, e.g. for /a/.)

The pitch domain of the six voices within the MIDIM-system:

$$T' \in ((N\text{min} \cdot T + Ef, N\text{max} \cdot T + M\text{max})_1 \cup (N\text{min} \cdot T + Ef, N\text{max} \cdot T + M\text{max})_2)$$

For domain E holds (see Kaegi, 1986, p. 144):

$$E = (N\text{min} \cdot T + Ef, N\text{max} \cdot T + M\text{max})$$

Dependent upon the maximum value of T' stated in the list Nmax must be changed if $T' \notin E$.

The duration (or DUR—) domains are: $DUR(\text{ms}) \in [20, 800]$.

The attenuation (or At—) values may be chosen arbitrarily.

3.5 An Artistic Application of the Derived Phonemes

The previous descriptions showed a number of facets of my work on the way to creating a soundtrack which I desired to build up from out stylized and with the MIDIM-system synthesized Dutch interjections. As stated this led me to detailed analyses of registered speech signals within which the presented vowels form only a small part. In the same way I also collected material for consonants and studied the transition between consonants and vowels which was touched upon in the sound example /nɑu/ in par. 2.7. The voice characteristics of the six voices which I chose as the starting point were preserved as much as possible with a few exceptions. There thus resulted for each element of my 'game' predicators which I grouped per voice and collected in predicator libraries within which each predicator was assigned an index. By calling for the correct predicators (via the index) from the descriptor it was possible for me to build synthesized series of connected vowels and consonants. It was possible to control the melodic contour and the increased duration of the vowels via the prosodic parameters pitch and duration, the loudness being varied with the attenuation. Seen that the MIDIM8x-system is particularly suitable for musical applications and that I needed a rather large number of segments for the sound description it was necessary to make use of compound predicators and *D*-comprehensors (see Kaegi, 1986, p. 158). Connecting the pitch and amplitudes of the various segments so that no undesired springs took place was accomplished in part by hand and in part calculated automatically with $\phi_6\chi_4$ (see Kaegi. 1986, p. 118). In principle expansions within the function tables (and eventually the grammar) could eliminate the work done by hand. In order to illustrate the sound series which I formed there follows an example. For each voice I built up a predicator library as basic material containing predicators for vowels, consonants and applicable connecting elements which we earlier termed joints (Phonetic notation according to the earlier stated table where it must be noted that the synthetic derivations do not always agree with the sounds in the reference words. The joints used are given for ease with the phonetic symbol for the predecessor and the successor):

Vowels: a, ɑ, ɔ, o, ε, e

For each vowel there exist five versions in which the amplitude changements and the portamento are different. We indicate this by means of two arrows, the first for the pitch and the second for the amplitude. Example: a ↑↓, raising pitch, decreasing amplitude.

Consonants: X, g, h, s, t, j, n, w or v

Matching joints: ha,	hɑ,	hɔ,	ho,	hε,	he
			so		
ja,	jɑ,	jɔ,		jε,	je
na,	nɑ,			nε,	ne
			wo or *v*o		

In order to synthesize the sound series /nεjahe/ I called via the *D*-comprehensor for example the following series of predicators:

n nε ε↓↓ ε↓↓ ε↑↑ j ja a↓↓ a↓↓ a↓↓ a↓↓ a↑↑ a↑↓ h he e↓↓

In this way I determined the course of the pitch (intonation contour), attenuation and duration in such a way that in the terminal expression the course of values for the pitch and amplitude arose, reproduced in Fig. 12. The pitch values are given in cents in relation to the approximated average pitch for vowels of the voice used (Female high). The amplitude of the two tracks is shown by means of connected and dotted lines.

Clearly visible are the joints at the quickly alternating amplitudes of the transition areas between vowels and consonants. (See par. 2.5).

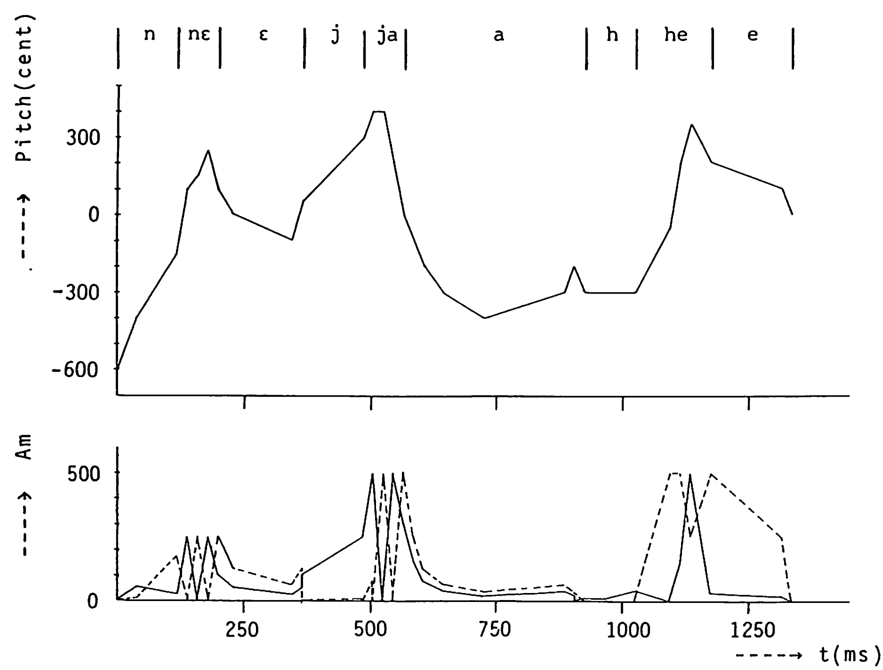


Figure 12. Values for the relative pitch (in relation to the average pitch of the applied vowels, dis′ = 311 Hz) and the amplitude of the synthetic sound series /nεjahe/ of a high female voice. The joints are clearly recognizable.

To summarize we can say that the MIDIM-system is very suitable for work in the area of speech synthesis although the system is directed towards musical applications (MIDIM = MinImal DescrIption of Music; Kaegi, the inventor of the MIDIM/VOSIM-system has done much investigation in the field of speech sounds. The synthesis system MIDIL was especially designed for this but for political reasons was not further practically developed.) More speech directed applications could be made concrete by expansion of the function table (see Kaegi, 1986, p. 94) where successive predicators would influence each other for example. The connection between an Image synthesis system and the MIDIM-system could also in this way be accomplished by the addition of functions which would connect the parameters of the various systems with each other.

APPENDIX I

The phonetic notation

Dutch vowels

short	English	Dutch
/ɑ/	shorter than bath	bad
/ɔ/	pot	pot
/ɛ/	fat	bed

half long or long	English	Dutch
/a/	father	baden
/o/	dote	boot
/e/	face	feest
/i/	free	niet

Dutch diphtongs

/ɑu/	loud	koud

Other symbols who are used:

English short vowel

/ʌ/ but

English diphtong

/ɑi/ line

Dutch consonants

	English	Dutch
/X/	'scotch'	zagen
/γ/	loch, but weaker	zagen
/h/	hitch-hike	hoog
/s/	leisure	stelling
/j/	yes	jong
/ʋ/	like a soft /v/	water
/n/	noon	nee

English consonants

/g/ gag

/w/ wigwam

APPENDIX II

The interjections which were registrated

Spoken by high, medium and low female and male voices. Registrated in 1982 at the Institute for Sonology.

/ha/	/ja/	/na/					
/hɑ/	/jɑ/	/nɑ/	/Xɑ/	/ɑX/	/γɑ/	/ɑγ/	/ɑ/
/hɑha/	majawɑt/*	/gaga/	/nɑu/				
/hɔ/	/jɔ/	/Xɔt/	/Xɔ/	/γɔt/	/γɔ/	/ɔX/	/ɔγ/
/ho/	/wo/	/so/	/o/				
/hɛ/	/jɛ/	/nɛ/					
/he/	/je/	/ne/					
/iʌ/	/ɑi/						

* This word is not an interjection. The person pronounced the Dutch words "maar ja wat" as a whole and because I liked the sound the word was included.

APPENDIX III

Information concerning the spectral analyses:

The spectral analyses are performed by means of the signal processing programs SIGPAC, developed by S. Tempelaars.

FIGURE	SAMPLING RATE	WINDOW WIDTH/SHIFT	RANGE	TYPE
6. top	20 kHz	256/128 samples	0-10 kHz	Hanning
10. top	20 kHz	256/128 —	—	—
10. bot.	20 kHz	256/128 —	—	—
Fundamental tracking:				
	20 kHz	512/512 —	0-0.4 kHz	—
ditto	20 kHz	512/256 —	—	—
	20 kHz	512/256 —	—	—

APPENDIX IV

The six dutch voices:

High female voice: singer, actress, video artist.
Medium female voice: flutist, sound- and concrete poet.
Low female voice: sculptress and stage designer.
High male voice: singer, tenor.
Medium male voice: professionally trained singer, linguist.
Low male voice: film/video producer.

All recordings were made after a number of rehearsals. The reason for these rehearsals was the need for interjections which had to be clearly articulated in short, medium and long time intervals. The average duration of the long interjections was 0.5 s.

ACKNOWLEDGEMENTS

I am highly indebted to Heinerich Kaegi and the other members of the MIDIM-group without whose vigorous assistance this paper would not have come into being. My thanks go also to Paul Goodman for his english translation and to Petra Banning for kind use of the video-tape.

BIBLIOGRAPHY

REFERENCES

Janssen, J./Kaegi, H. MIDIM-duplication of a Central-javanese sound concept, Interface, 15 (1986), pp. 185-229.

Kaegi, W. Controlling the VOSIM sound synthesis system, Interface, 15 (1986) pp. 71-82.

— The MIDIM-language and its VOSIM interpretation, Interface, 15 (1986), pp. 83-161.

Pols, L. Spectral Analysis and Identification of Dutch vowels in Monosyllabic Words. Printed by Academisch Pers B., Amsterdam, 1977.

RECOMMENDED

Collins, B. and Mees, I. The Sounds of English and Dutch, 1984. Leiden University Press, Leiden.

Fudge, E. C. "Phonology" in New Horizons in Linguistics by J. Lyons and D. Crystal. Penguin Books.

Kaegi, W. A new approach to a theory of sound classification, Interface, 1 (1972) pp. 93-109.

— A minimum description of the linguistic sign repertoire, Interface, part one: 2 (1973) pp. 141-156; part two: 3 (1974) pp. 137-158.

— The MIDIM 8X System, Desc. user manual, Institute for Sonology, University of Utrecht, 1984.

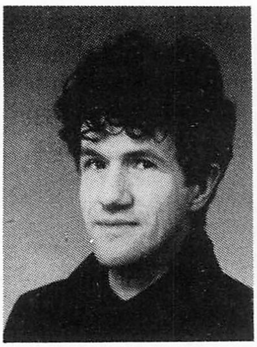
— Musicians pitch-frequency-cent table, Institute for Sonology, University of Utrecht, 1984.

— Design of musical discovery, A selection, Revised edition, Inst. for Sonology, University of Utrecht, 1985.

Tempelaars, S. Testing elements from the Kaegi minimum description of signs from the linguistic repertoire using the Kruksal method, Interface, 2 (1973), pp. 157-158.

— The VOSIM signal spectrum, Interface 6 (1977), pp. 57-72.

— SIGPAC manual, Institute for Sonology, 1982.



Pierre van Berkel was born in 1953 in Waremme (Belgium). He is Dutch by nationality. He studied painting at the academy St. Joost in Breda, animation film at the Rijksacademie van Beeldende Kunsten in Amsterdam and sonology at the Institute for Sonology in Utrecht. He made the animation films "Strepen", 1977, "Gesprek", 1979, and the soundpoems "Tjaktjakai", 1983, and "Heegothe", 1984.

The main focus of his activities is the human voice. The soundpoems were made using the MIDIM-system of Dr. W. Kaegi. At the moment he is working on a video-project called "Abstracte Spraak-film". For the soundtrack of this project the MIDIM-system is used and for the images he has developed software for computer animation.

Interface, Vol. 15 (1986), pp. 257-269. 0303-3902/86/1501-0259 $3.00

CANON:
A System for the Description of Musical Patterns

Pieter Kuipers

ABSTRACT

As an extension to the MIDIM-language formulated by Kaegi, the author has developed a system named CANON for the description of musical patterns. In this article subclasses of the MIDIM-concept PREDICATOR are defined. Such a class is described by a CANON-vector which determines domains for the parameters pitch, duration and intensity instead of the fixed values occurring in a MIDIM-descriptor. These CANON-vectors may be concatenated to a comprehensor with which a great number of one-voiced musical patterns can be described. A system for the description of polyphonic patterns is in development. These systems may be used to research the occurrence of patterns in various musical styles, and at the same time could be useful for composers when forming their ideas.

1. INTRODUCTION

When I arrived at the Institute for Sonology in Utrecht in 1983 I encountered the MIDIM-system of Dr. W. Kaegi. The name of this system: "a Minimum Description of Music" indicates that this description is sufficient but by no means complete. After I had acquired some experience with the system, I felt the need to proceed in the same direction as that taken by Kaegi.

The MIDIM-language is a language based upon an alphabet of VOSIM-vectors. For every vector there exists a one to one relationship with a certain sound event. With the MIDIM-language it is possible to describe classes of sounds, with which a classification is made within the immense repertoire of possible (sequences of) VOSIM-vectors. Such a class of sounds is described by a predicator.

At the Institute for Sonology in Utrecht I developed a system, called CANON, with which a further subdivision of these sound classes was obtained. As in the MIDIM-language this classification was realized with the help of λ-abstraction. The CANON-system may be considered as a formal language, based upon an alphabet of descriptor parameters. By limiting the domains out of which the six variables pitch, octave, duration, intensity, predicator-index and MIDIM-articulation may be chosen, new concepts originate by means of which sound classes of a high musical relevance may be represented.

An example of this is the concept of rhythm, where fixed values are assigned to the duration parameters, while no restrictions are made for the parameters pitch, octave and intensity. For example a tango-rhythm may be described (as is done in Figure 1) by a sequence of five duration values.

du
1: 250
2: 500
3: 250
4: 500
5: 500

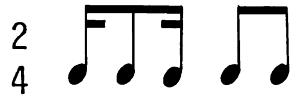


Figure 1. A tango rhythm described by a sequence of five fixed duration values (related to tactus = 1000).

There is a strong relationship between rhetorical and musical concepts, just as there is a strong relation between music and speech.[1] Since antiquity these concepts have received much attention from musicians, orators, actors and theorists.

Around 1600, under the influence of Humanism, rhetoric (which was originally trivial) was introduced on a large scale into the musical culture of the time, which before this had been rather mathematical. The esoteric symbolism which characterizes the music of the Burgundian and Flemish school was replaced by rhetorical concepts by which the "Affekten" of the listener could be controlled in a very accurate way.

Although rhetoric has never since been disputed in such a violent way as was done in the Baroque era, the use of rhetoric and related musical concepts has in no way fallen into the background. An example of this is the use of sounds taken from our daily environment, such as bird-song, the roar of trains or the din of café conversation in electronic compositions. The aim of this is often to stimulate certain associations in the listener in order to bring him into a certain state of mind.

2. DEFINITION OF THE CANON-VECTOR $\vec{c}$

2.1 The MIDIM-descriptor line

The MIDIM-language as formulated by Kaegi is a formal language based upon an alphabet of VOSIM-vectors $(x_1..x_{12})$. After the substitution of functions (λ-operators) within the MIDIM-vector $\vec{m}_0 = (\lambda x_1..x_{12})$. $(x_1..x_{12})$ concepts were defined as sequences of four segments interpreted as sounds. The total duration of these four segments (in ms) is called *DUR*. Finally a predicator P was defined with which it is possible to describe a class of sounds. In a M8X-standard predicator at most the λ-tied parameters *DUR, At* and *T'* occur. These so-called

[1]) In Interface, vol. 12 (1983), S. Dydo published the results of his investigations in this field. By extracting patterns from an input list relations were made between this list and other parameters, especially between pitch and timbre.

prosodic parameters can be eliminated with the help of a MIDIM-descriptor wherein values are assigned to the variables[2] according to:

$$(\lambda\, DUR, At, T' \,.\, \mathrm{P}\,(ind, art))\,(a, b, c) \in \mathrm{P}(ind) \subset V^*$$

In the MIDIM-theory the following functions are substituted for the prosodic variables *DUR, At* and *T'*:

$$(\lambda\, DUR, At, T')\, DUR, At, T' \,.\, \mathrm{P}\,(ind, art) =$$

$$(\lambda\, \phi_1\, \phi_2\, \phi_3)\, met, du, at, sub, pi, oc \,.\, \mathrm{P}\,(ind, art) \qquad (1)$$

The variables occurring in a descriptor line have the following interpretation:

pi and *oc* determine the frequency by:

$$f(\mathrm{Hz}) = 1/T' = 16{,}3516\ \mathrm{Hz} \cdot 2^{(oc + pi/sub)}$$

where *sub* stands for the subdivision of the octave in equal parts. This definition is chosen in such a way that the fourth octave contains the central $a' = 440$ Hz.

du determines the tone duration by means of:

$$DUR\,(\mathrm{ms}) = du \cdot 60\ \mathrm{ms}/met$$

where *met* stands for Mälzel's Metronome number.

at determines the intensity by means of:

$$At = L_0 \cdot 2^{-at}$$

L_0 corresponds to an arbitrary maximum intensity to which *at* is related.

art and *ind* form a MIDIM-word and imply an application of a MIDIM-predicator.[3]

When we assume, for example, that P(1) stands for the class of sounds which are produced by a bassoon in a high register, (in fact such a predicator is formulated in the MIDIM-language cf. Kaegi, 1978), then the first tone of Igor Strawinsky's Sacre du Printemps may be described by means of a descriptor line as is done in Figure 2.

With a MIDIM-descriptor, elements from the concept P can be defined. The aim of the CANON-system is to make possible the description of subsets of P which have more than one element.

A descriptor is in a formal sense a non-terminal: by λ-elimination a terminal $\in V^*$ may be obtained. In an informal sense as well a descriptor may be called a

[2]) All parameters and variables occurring in this article are defined and symbolized in a way corresponding with those used in Kaegi, 1986.

[3]) Cf. Kaegi, 1986, § 6, § 7.

sub = 12 *met* = 50

pi	*oc*	*du*	*at*	*ind*	*art*
0	5	3000	1	1	2

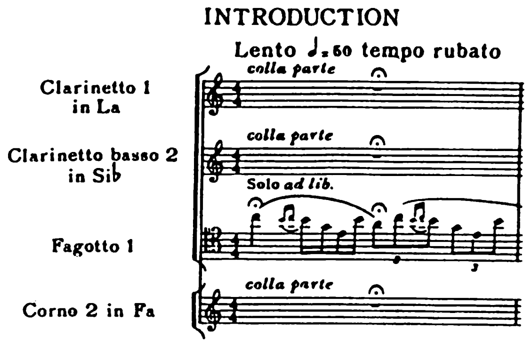


Figure 2. Conventional and MIDIM-description of the first note in Strawinsky's Sacre du Printemps.

non-terminal: by further substitution of functions new non-terminals may be constructed.

2.2 Introducing functions into the MIDIM-descriptor

I introduced the following functions:

f_1 *pi*: $\lambda\, \alpha_1 \,.\, \text{int}\,(c_1^- + \alpha_1 \cdot (c_1^+ - c_1^-))$

f_2 *oc*: $\lambda\, \alpha_2 \,.\, \text{int}\,(c_2^- + \alpha_2 \cdot (c_2^+ - c_2^-))$

f_3 *du*: $\lambda\, \alpha_3 \,.\, \text{int}\,(c_3^- + \alpha_3 \cdot (c_3^+ - c_3^-))$

f_4 *at*: $\lambda\, \alpha_4 \,.\, \text{int}\,(c_4^- + \alpha_4 \cdot (c_4^+ - c_4^-))$

f_5 *ind*: $\lambda\, \alpha_5 \,.\, \text{int}\,(c_5^- + \alpha_5 \cdot (c_5^+ - c_5^-))$

f_6 *art*: $\lambda\, \alpha_6 \,.\, \text{int}\,(c_6^- + \alpha_6 \cdot (c_6^+ - c_6^-))$

f_7 *intv*: $\lambda\, \alpha_7 \,.\, \text{int}\,(c_7^- + \alpha_7 \cdot (c_7^+ - c_7^-))$

f_8 *pitch*: $\lambda\, \alpha_1, \alpha_2 \,.\, f_1 + f_2 \times sub$

Now a canon-vector may be defined by substitution of the λ-operators f_1 to f_6 into formula (1):

$$\vec{c} = (\lambda\, f_1 f_2 f_3 f_4\, \phi_1\, \phi_2\, \phi_3)\, met,\, sub,\, \alpha_1, \alpha_2, \alpha_3, \alpha_4 \,.\, \mathrm{P}\,(\lambda\, f_5, f_6 \,.\, \alpha_5, \alpha_6)$$

The set of all canon-vectors is called C. The domains for the coefficients c_i are derived from those of the corresponding descriptor parameters. Thus:

$pi,\, c_1^-,\, c_1^+ \quad \in [0, sub-1]$

$oc,\, c_2^-,\, c_2^+ \quad \in [0{,}9]$

du, c_3^-, c_3^+	$\in [1,9900]$	
at, c_4^-, c_4^+	$\in [0,..)$	
ind, c_5^-, c_5^+	$\in [1,..)$	The predicator-index *ind* is a name for a sound concept which is only implicitly called for in a MIDIM-descriptor.
art, c_6^-, c_6^+	$\in [1,9]$	The MIDIM-articulation *art* is not an ordinal quantity. For that reason the supplementary restriction below is made: $c_6^- \neq c_6^+ \Rightarrow c_6^- = 1 \wedge c_6^+ = 2$
c_7^-, c_7^+	$\in [-6 \times sub, 6 \times sub]$	
sub	$\in [2,1200]$	
met	$\in [1,999]$	

All coefficients c_i are integers for which hold:

$$\forall_{i \in [3,7]}: c_i^+ \geqslant c_i^- \text{ and}$$

$$c_1^+ + (c_2^+ \times sub) \geqslant c_1^- + (c_2^- \times sub)$$

All variables α_i are real numbers within the interval [0,1].

By this choice for the domains the following holds:[4]

$$\forall_{i \in [2,8]}: f_i \in [c^-, c^+]_i$$

When a musical interpretation is aimed at, and in the MIDIM-system this is the case, an important additional restriction has to be made with respect to the choice of the coefficients c_i^- and c_i^+.

The call for a predicator P(*ind*) requires the satisfaction of a certain domain for the parameters *DUR, At* and *T'* allied with it. This kind of domain restriction may concern the boundaries of a certain pitch range for an instrumental register (e.g. low flute), it may occur in the intensity domain of an instrument (e.g. the moderate sound-level of a clavichord) or restrictions may be made with respect to tone duration. For example low tones of an organ or a double-bass need a certain time to unfold into sound.

In the case of predicators P(*ind*) which do not at all depend on one of the variables *DUR, At* and *T'* (such as is the case for wood-blocks, harpsichord or triangle where *DUR, At* and *T'* respectively are more or less fixed) a choice for c_i will be absolutely irrelevant.

[4]) For convenience here and in what follows the abbreviated notation $[c^-, c^+]_i$ for $[c_i^-, c_i^+]$ is used.

In the MIDIM-system a correction mode is defined in order to preserve the domains of a certain sound concept.[5] When these corrections are ignored, elements from other sound concepts will result. Although this practice is well-known in the history of music (e.g. the grumbling double-basses in Beethoven's pastorale "Gewitter, Sturm", or the barking of a dog in Vivaldi's four seasons (la primavera) where the violas are told to play "sempre molto forte e strappato", i.e. very strong and tearing) we will restrict ourself in this article to the use of the MIDIM E- and C-domain.

2.3 Interpretation of the canon-vector

A canon-vector is a subset of the concept P(*ind*), wherein the descriptor-parameters *pi, oc, du* and *at* are constrained within the domain $[c^-, c^+]_i$. The variables c^- and c^+ represent respectively the lower and upperbound of this domain, while α_i is a pointer indicating an element taken from this domain.

If $c_5^- \neq c_5^+$, then each value of *ind* between c_5^- and c_5^+ implies a subset of P(*ind*) such that the canon-vector may be described by:

$$\vec{c} = \bigcup_{ind \in [c^-, c^+]_5} (\quad) \cdot \mathrm{P}(ind) \tag{2}$$

If $c_i^- = c_i^+$ then

$$\lambda\, \alpha_i \,.\, \mathrm{int}(c_i^- + \alpha_i \cdot (c_i^+ - c_i^-)) = \lambda\, \alpha_i \,.\, c_1^- \Leftrightarrow c_i^-$$

Thus for $c_i^- = c_i^+$ the canon-vector will no longer depend upon α_i.

3. THE CANON-COMPREHENSOR

The set of canon-vectors C may be provided with the operation of concatenation. By a combination of N canon-vectors $\vec{c}(line)$, $(line = 1\,..\,N)$ a sequence of canon-vectors called CANON is produced. This so-called canon-comprehensor is an element of C^*, the product set of all sequences taken from C. In this way to every vector $\vec{c}(line)$ a successor $\mathrm{SUCC}(\vec{c}(line)) = \vec{c}(line + 1)$ may be assigned. $\mathrm{SUCC}(\vec{c}(N)) = \mathscr{E}$, the empty string.

By substituting functions it is possible to make a link between $\vec{c}$ and $\mathrm{SUCC}(\vec{c})$. In this way the content of a vector $\vec{c} \in$ CANON is made context dependent. This is done in the following two ways:

1: With the concatenation of canon-vectors a sequence of MIDIM-words $(ind, art)_1$, $(ind, art)_2, \ldots$ arises. These words are rewritten into well-formed MIDIM-formulae in the same way as is done in the MIDIM-system.[6]

[5]) Cf. Kaegi, 1986, § 8.

[6]) Cf. Kaegi, 1986, § 7.

2: In order to make feasible the representation of the largest possible repertoire of musical patterns, to every canon-vector a supplementary function f_7 is added which designates the domain of a melodic interval. This is done because the melodic interval plays a very important role in the setting up of musical structures in both contemporary and classical music.[7] This redundancy in the description implies the following context-dependency of $\vec{c}\,(line + 1)$:

$$pitch^{-}(line + 1) = \max\,(pitch^{-}(line) + c_7^{-}, pitch^{-}(line + 1))$$

$$pitch^{+}(line + 1) = \min\,(pitch^{+}(line) + c_7^{+}, pitch^{+}(line + 1))$$

Use is made here of the function f_8:

$$pitch^{-} = c_1^{-} + c_2^{-} \text{ x } sub$$

$$pitch^{+} = c_1^{+} + c_2^{+} \text{ x } sub$$

When the domains for c_1 and c_2 are chosen as shown in 2.2, there results (within the context of a fixed subdivision) a one to one relationship between the value for *pitch* and the pair *pi, oc*.

$$pi = pitch \text{ modulo } sub$$

$$oc = \text{int}(pitch/sub)$$

A sequence of canon-vectors satisfying the two above-mentioned conditions is called a CANON-comprehensor. A 14 x N-matrix containing the coefficients $c_i^{-}(line)$ and $c_i^{+}(line)$ (i = 1..7) of a canon-comprehensor is called a CANON-matrix. A canon-matrix stands for a melodic, rhythmic and/or dynamic pattern in which at most the λ-tied variables $\alpha_1 .. \alpha_7$ occur.

Patterns which may be described by means of a CANON-matrix play a crucial role in our musical culture. Apart from music and speech, they are also found in nature. The following relation between $\vec{c}\,(line)$ and $\vec{c}\,(line + 1)$ holds for example for the quacking of a wild duck (♀ only) when in an excited state:

$$at(line + 1) \geqslant at(line)$$

$$pitch(line + 1) \leqslant pitch(line)$$

$$(N \in [4, 24])$$

In the Baroque era many rhetorical figures were explicitly described by theorists such as Christoph Bernhard or Johann Mattheson (1739). These figures were called "Figuren", "Manieren" or "loci topici" and may be abundantly

[7]) An excellent example of this is the theory set up by the Dutch composer Peter Schat. In this theory called "De toonklok" (the tone-clock) he replaced the tone-series as used by the composers of the Viennese school by the concept of interval-series. An english version of his article on this subject has been published in Key Notes 17, 1983/1, by Donemus in Amsterdam.

found in all music between Schütz and Bach. The concept EXCLAMATIO may serve as an example. This rhetorical figure was used in order to characterize an exclamation or merely to indicate the start of a sentence.[8] In the recitative "erwäge doch" from Bach's cantata "Ein feste Burg ist unser Gott" this figure may be observed numerous times.

In Figure 3 the CANON-matrix which gives a description of the EXCLAMATIO can be seen. Characteristic for an EXCLAMATIO is a preparatory rest, and the large upward leap which follows. In the MIDIM-system a rest is represented by an articulation-value 9. The domain for the upward leap is $[c_7^-, c_7^+] = [5, 10]$. The interval is thus described by the function $f_7(\alpha_7) = 5 + \alpha_7 \cdot 5$. The lowest note is to be found in measure 7: $f_1(0) = 9, f_2(0) = 2$, the highest in measure 10 (at "Himmel"!): $f_1(1) = 4, f_2(1) = 4$. In this example the predicator-index 1 stands for the class of sounds produced by a bass.

4. THE CANON-HANDLER

With the definition of a CANON-matrix a classification has been realized within the series of concepts P(*ind*). In order to describe elements taken from these classes it is necessary to fix for every line $\in [1, N]$ values for the variables $\alpha_1 \ldots \alpha_7$. For this purpose I developed the CANON-handler. With the CANON-handler it is possible to combine various CANON-matrices with each other.

The combination is realized by means of an iterative comparison between a dummy CANON-matrix called tabula rasa, and the CANON-matrices to be combined. A tabula rasa is a canon-matrix within which all coefficients are given the extreme values of the domains mentioned in 2.2.

Elimination of $\alpha_1 \ldots \alpha_7$ occurs in three stages.

1: In the FIRST STAGE the mutual configuration of the matrices involved is determined. An example of such a configuration is shown in Figure 4. In this configuration an instance of the concept 'descending scale' is represented. A descending scale may be described by a concatenation of canon-vectors for which $[c^-, c^+]_7 = [-2, -1]$. Instances of this concept may be found everywhere in music, for example in the first measures of Claude Debussy's "Prélude à l'après-midi d'un faune" or in his second prélude for piano (Fig. 5).

In Figure 4 a diagram is shown in which the concept 'descending scale' (described by canon B) is combined with other canon-matrices in order to fix the values for α_1, α_2 and α_3. The construction of such a configuration is at the moment made by the user of the system. The intention is to later construct these configurations by means of functions. Once this is possible, the occurrence of a musical pattern will be context dependent.

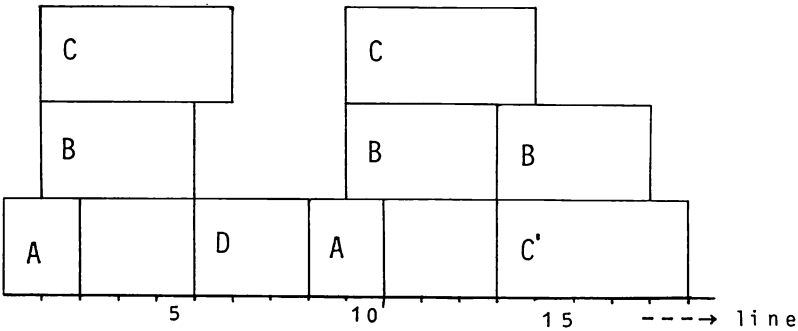
[8]) In Mattheson (1739) Part II, Ch. 9, § 65 to § 68 various kinds of EXCLAMATIO are discussed.

EXCLAMATIO.CAN

sub = 12 *met* = 60

	pi	*oc*	*du*	*at*	*ind*	*art*	*intv*
1:			250..500			9	
2:	9..4	2..4	250..500	0..4	1	1	5..10
3:	2..4	3..4	250..1250	0..4	1	1..2	

Figure 3. J. S. Bach: cantata BWV 80 "Ein feste Burg ist unser Gott" with a CANON description of the rhetoric figure "exclamatio".



A.CAN

	pi	*oc*	*art*
1:			9
2:	4	5	2

B.CAN

	intv
1:	—1
2:	—1
3:	—1
4:	—1

C.CAN

	du
1:	625
2:	125
3:	125
4:	125
5:	250..1500

C'.CAN

	du
1:	1250
2:	250
3:	250
4:	250
5:	500..3000

D.CAN

	pi	*oc*	*du*
1:	0	4	875
2:	0	5	125
3:	5	4	125

Figure 4. Combination of the canon-matrices A, B, C, C' and D (sub = 6, met = 44).

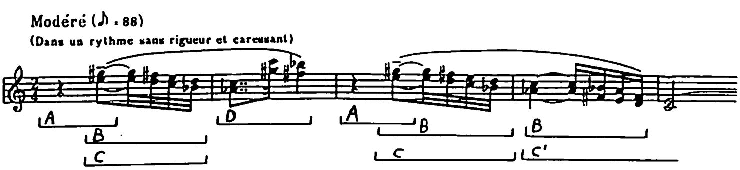


Figure 5. C. Debussy: Prélude no. 2 (. . . Voiles).

2: In the SECOND STAGE several canon-lines are combined with each other. Consider for example the second line from the configuration of Figure 4, where we see the combination of the second line of canon A with the first lines of canon B and canon C.

Let there be a canon-line $\vec{c}$ with coefficients $[c_i^-, c_i^+] = [min, max]$.

Let there be a canon-line $\vec{c}\,'$ with coefficients $[c_i^-, c_i^+] = [p1, p2]$.

When we state: $\left[\begin{array}{l} q1 = \text{maxim}(p1, min) \\ q2 = \text{minim}(p2, max) \end{array}\right.$

then for all $p1$ and $p2$ holds: $\begin{bmatrix} q1 \geqslant min \\ q2 \leqslant max \end{bmatrix}$

The intersection of $[min, max]$ with $[p1, p2]$ then becomes:

for $q1 \leqslant q2$: $[min, max] \cap [p1, p2] = [q1, q2]$

for $q1 > q2$: $[min, max] \cap [p1, p2] = \varnothing$

The intersection of two canon-lines c and c' may now be written as:

$$\vec{c} \cap \vec{c}' = \bigcup_{ind \in [q1,\, q2]_5} : (\lambda f_1 f_2 f_3 f_4 \phi_1 \phi_2 \phi_3)\ met,\ sub\ \alpha_1 \alpha_2 \alpha_3 \alpha_4 . \qquad \mathrm{P}(\lambda f_5, f_6 . \alpha_5 \alpha_6)$$

if $\forall_{i \in [1,7]} : q1_i \leqslant q2_i$.

When $\exists_{i \in [1,7]} : q1_i > q2_i$ then holds $\vec{c} \cap \vec{c}' = \varnothing$.

In this case no combination can take place for those coefficients c_i where $q1_i > q2_i$. In order to make still possible the combination of other coefficients, within the various canon-lines a hierarchy has been introduced. To every pair of coefficients $[c^-, c^+]_i$ a code has been added to which may be given the following interpretation.

1: The domain under consideration is a *LAW*. These values must at all times be respected. (This code may be used for the description of patterns connected with a certain register. Thus for every flageolet on a viola holds $oc \geqslant 4$.)

2: The domain under consideration is a *RULE*. The limit given by the canon-matrix must be respected if possible. Contravention of a rule is only possible after a FIAT given by the user.

3: The domain under consideration is a *WISH*. The domain will be negated when no combination is possible. If combination is indeed possible, the domain will be applied with a probability *P*.

We may now represent the process by which the domains from the canon-lines $\vec{c}$ and $\vec{c}'$ are combined as follows:

On the basis of the results shown above, the content of vector $\vec{c}$ will change for every $i \in [1, 7]$ to $[min', max']_i$ according to:

1: When $[p1, p2]$ is a *LAW* then holds:

a) $q1 \leqslant q2 \quad \Longrightarrow \quad \begin{bmatrix} min' := q1 \\ max' := q2 \end{bmatrix}$

b) $q1 > q2 \quad \Longrightarrow$ Error message, no output will be generated by the canon-handler.

2: When $[p1, p2]$ is a *RULE* then holds:

a) $q1 \leqslant q2 \implies \left[\begin{array}{l} min' := q1 \\ max' := q2 \end{array}\right.$

b) $q1 > q2 \implies$ The user is given a choice between three alternatives:

1) no combination with $[p1, p2]_i$:
$\left[\begin{array}{l} min' := min \\ max' := max \end{array}\right.$

2) approximation: (K is an integer)
$p2 < min \Rightarrow \left[\begin{array}{l} min' := min \\ max' := min + \dfrac{max - min}{K} \end{array}\right.$

$p1 > max \Rightarrow \left[\begin{array}{l} min' := max - \dfrac{max - min}{K} \\ max' := max \end{array}\right.$

3) No output will be generated by the canon-handler.

3: When $[p1, p2]$ is a *WISH* then holds:

a) $q1 > q2 \implies$ no combination with $[p1, p2]_i$:
$\left[\begin{array}{l} min' := min \\ max' := max \end{array}\right.$

b) $q1 \leqslant q2 \implies$ There is a probability P that a combination will be made:
$\left[\begin{array}{l} min' := q1 \\ max' := q2 \end{array}\right.$

There is a probability $1\text{-}P$ that nothing will change:
$\left[\begin{array}{l} min' := min \\ max' := max \end{array}\right.$

After the resulting canon-vector with coefficients $[c^-, c^+]_i = [min', max']_i$ has been controlled in accordance with the context dependent canon-comprehensor grammar (cf. 3), the original vector $\vec{c}$ will be replaced by the resulting vector. Thus a combination of $\vec{c}$ with $\vec{c}'$ has been made.

3: Finally in the THIRD STAGE of the elimination process a value will be assigned to the (eventually) remaining λ-tied variables α_i. This can take place by choosing for every domain a random value ($\alpha_i :=$ alea $[0, 1]$), or by means of a user defined choice.

When in this way for all parameters α_i (*line*) a value is determined, a MIDIM-descriptor may be obtained through λ-elimination by filling these values into the canon-vector (2). When this is combined with a library of predicators by means of the MIDIM-system a subset of the product set of VOSIM-vectors V^* is determined, which corresponds to a series of sounds.

5. CONCLUSION

By means of the CANON-system it has become possible to collect a large variety of musical patterns based upon the parameters pitch, duration and intensity, and to store them in libraries in which they are arranged. In this way a start can be made with the investigation of the occurrence and meaning of musical patterns in diverse styles. Thanks to the MIDIM/VOSIM-system it is possible to apply these patterns in music. The CANON-system can be a tool for composers when organizing their ideas, and structuring their compositions in as far as this is desired. Lastly, a means for the description of musical patterns is a necessary condition for a pattern recognition system which could be used to extensionalize associative processes.

REFERENCES

Bernhard, Chr. Tractatus compositionis augmentatus and Von der Singe-Kunst oder Manier. Both published in an english translation by W. Hilse in The Music Forum III, 1-196.

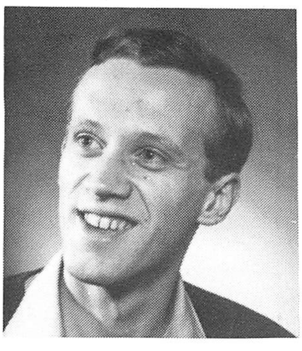
Dydo, S. (1983) Surface Relations between Music and Language as Compositional Aids. Interface 12, 541-556.

Kaegi, W. and Tempelaars, S. (1978) VOSIM - A New Sound Synthesis System. Journal of the Audio Engineering Society 26, n. 6.

Kaegi, W. (1986) The MIDIM-language and its VOSIM-interpretation. Interface, Vol. 15, pp. 83-161.

Mattheson, J. (1739) Der vollkommene Capelmeister. Facsimile published by M. Reimann in Documenta musicologica I, 5, Kassel 1954. English translation by E. Harriss, Michigan 1981.

Schat, P. (1984) De toonklok. Amsterdam 1984.



Pieter Kuipers was born in Amersfoort, Holland, 1959. In 1977 he finished his Gymnasium-B education. Since then he studied experimental physics and informatics at the University in Groningen. In order to obtain a broader insight into our musical tradition, from 1980 until now he combined this with a study in musicology at the University of Utrecht. He is also active as an amateur pianist and viola player in various Dutch student orchestras. Since 1983 he worked as a student under the supervision of Dr. W. Kaegi and is now a member of the MIDIM-group for Associative Computer Music.

MIDIM-Group: Dr. Werner Kaegi
Paul Goodman
Jos Janssen
Heinericht Kaegi
Pieter Kuipers

ADDRESS: MIDIM-Group c/o
Wagenstraat 7
3581 WP Utrecht
The Netherlands

NOTICE TO CONTRIBUTORS

1. Papers should preferably be written in English and not be longer than 5000-8000 words. Articles in French or German will be accepted in exceptional cases. Authors writing in a language not their own are urgently requested to have their manuscripts checked for linguistic correctness before submission.

2. Each article should be preceded by an abstract in English, not exceeding 150 words.

3. Authors should submit two copies of each article. Submission of a paper will be taken to imply that it is unpublished and not under consideration of publication elsewhere.

4. The Editors reserve the right to refuse any manuscript submitted, whether on invitation or otherwise, and to make suggestions as to modifications before publication. Submitted papers should be in a final form ready for publication. All corrections in the proofs other than printing errors are at the author's expense.

5. Bibliographical references should be listed in alphabetical order at the end of the paper.

6. References to periodicals should include the following items: name(s) and initial(s) of author(s); year of publication; title of paper; title of periodical; volume number (Arabic numerals); first and last page number.
   References to books should include: name(s) and initial(s) of author(s); year of publication; exact title; name(s) of publisher; place of publication; page number (where pages are specifically quoted).

7. References should be quoted in the text in parentheses by the name(s) of author(s) followed by the year of publication, e.g. "(Jones, 1961)" except when the author's name is part of the sentence, e.g. "Jones (1961) has shown that...". If there are more than two authors, "et al." may be placed after the first name, e.g. "Jones et al., 1961".

8. Authors will be asked to contribute to the cost of excessive illustrations and elaborate tables. The cost of reproducing coloured plates must be met in full. Estimates may be obtained in advance.

9. Music examples should be suitable for reproduction by offset.

10. Articles accepted by the Editors become the property of Swets & Zeitlinger B.V. No article or any part thereof may be reproduced in any form whatsoever without permission of the Publisher.

11. The Editors and the Publisher of Interface are not responsible for the scientific content and statements of the authors of accepted papers.

12. Offprints may be ordered when proofs are returned.

CONTENTS