AUDIBLE DESIGN

TREVOR WISHART

APPENDIX 2

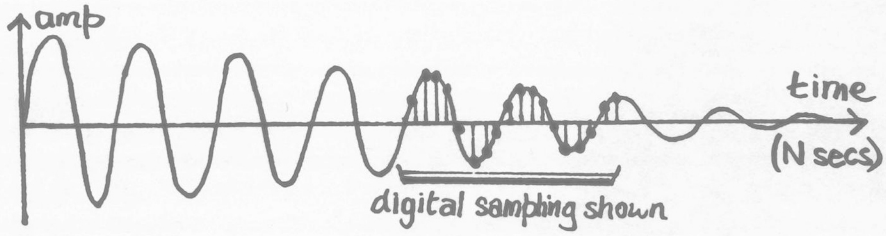
A diagrammatic guide
to sound compositional processes.

Published by Orpheus the Pantomime Ltd : 1994

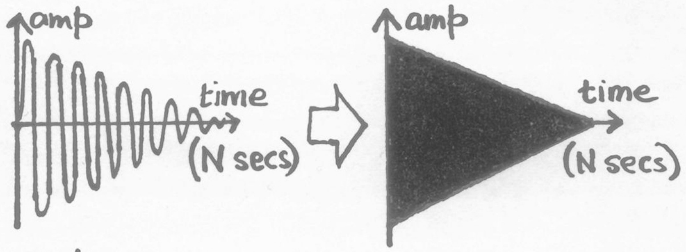
KEY

WAVEFORM :

sound represented in the time domain.

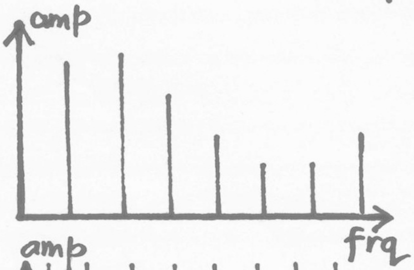


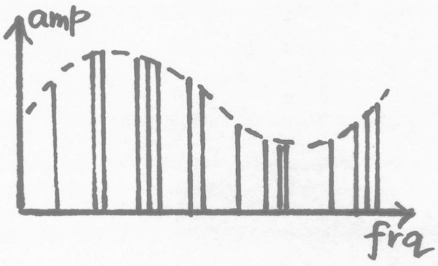
When looking at much longer blocks of time, waveform would be very compressed in display, so a different format is used.



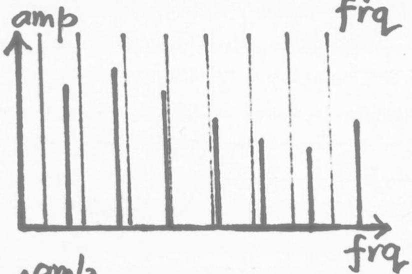
SPECTRUM :

sound represented in the frequency domain. N.B. Frequency axis evenly spaced with respect to frequency, & NOT to pitch.

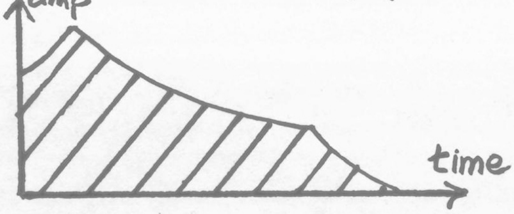




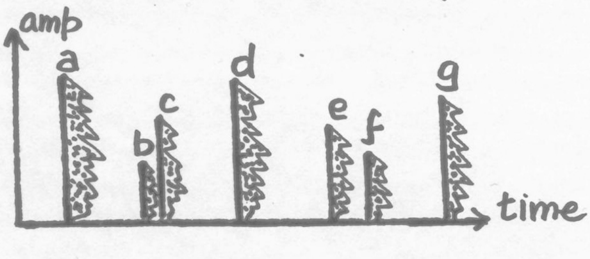
Spectrum, displaying analysis channels.



LOUDNESS TRAJECTORY : (loudness envelope) of a sound.



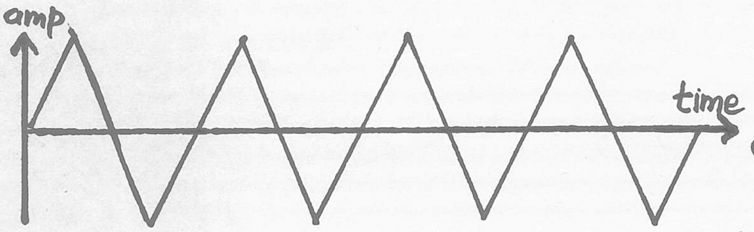
MIX PLACEMENT : represents the times at which sounds in a mix begin. Does NOT display the whole of each sound.

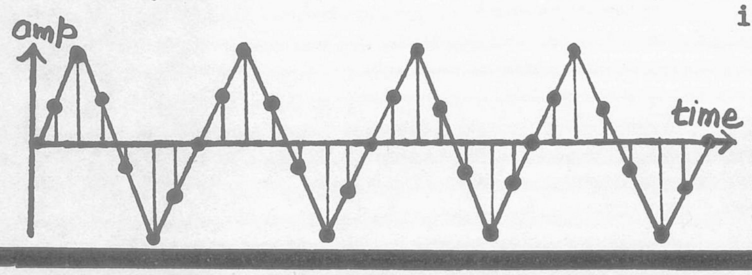


N.B. ALL REPRESENTATIONS ARE SCHEMATIC.

We do not attempt to represent realistic waveforms, spectra, loudness trajectories or mix placements, but only to illustrate the principles involved in the various compositional processes described. Waveform, spectral etc shapes have been chosen to make their transformations as clear as possible in diagrammatic form.

SAMPLING

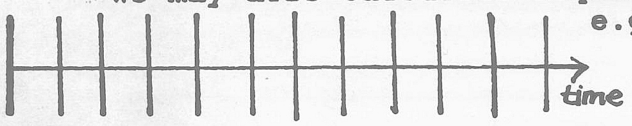




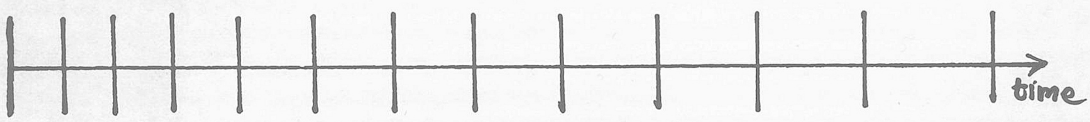
On the computer, and in other digital media, the waveform is stored as a sequence of individual values, or samples, derived from the waveform as shown.

SEQUENCE GENERATION

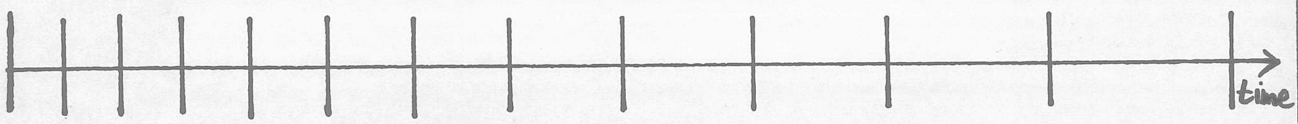
It is often useful to generate sequences of timing information for musical composition. A MIDI instrument does this by sending data when a key is pressed or released.
We may also calculate sequences of times automatically e.g.



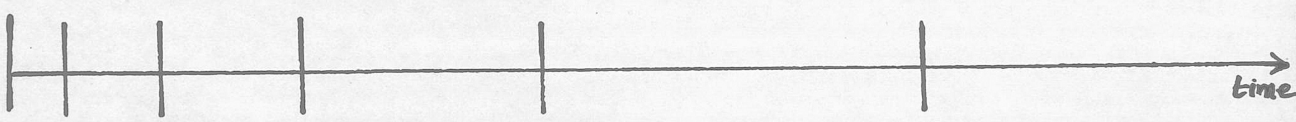
Regular times.



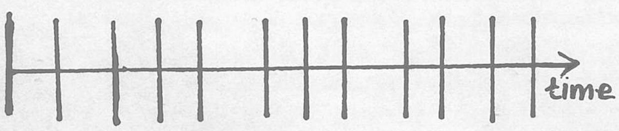
Times growing by addition of a time-unit.



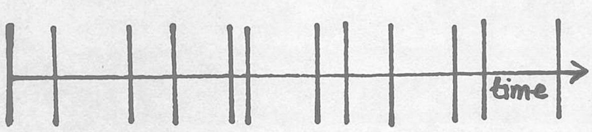
Times growing exponentially.



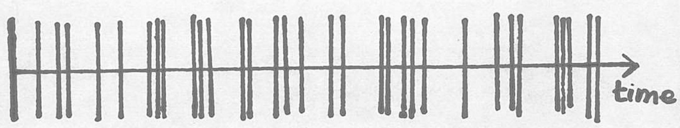
Times following Fibonacci series.



Regular times slightly randomised.

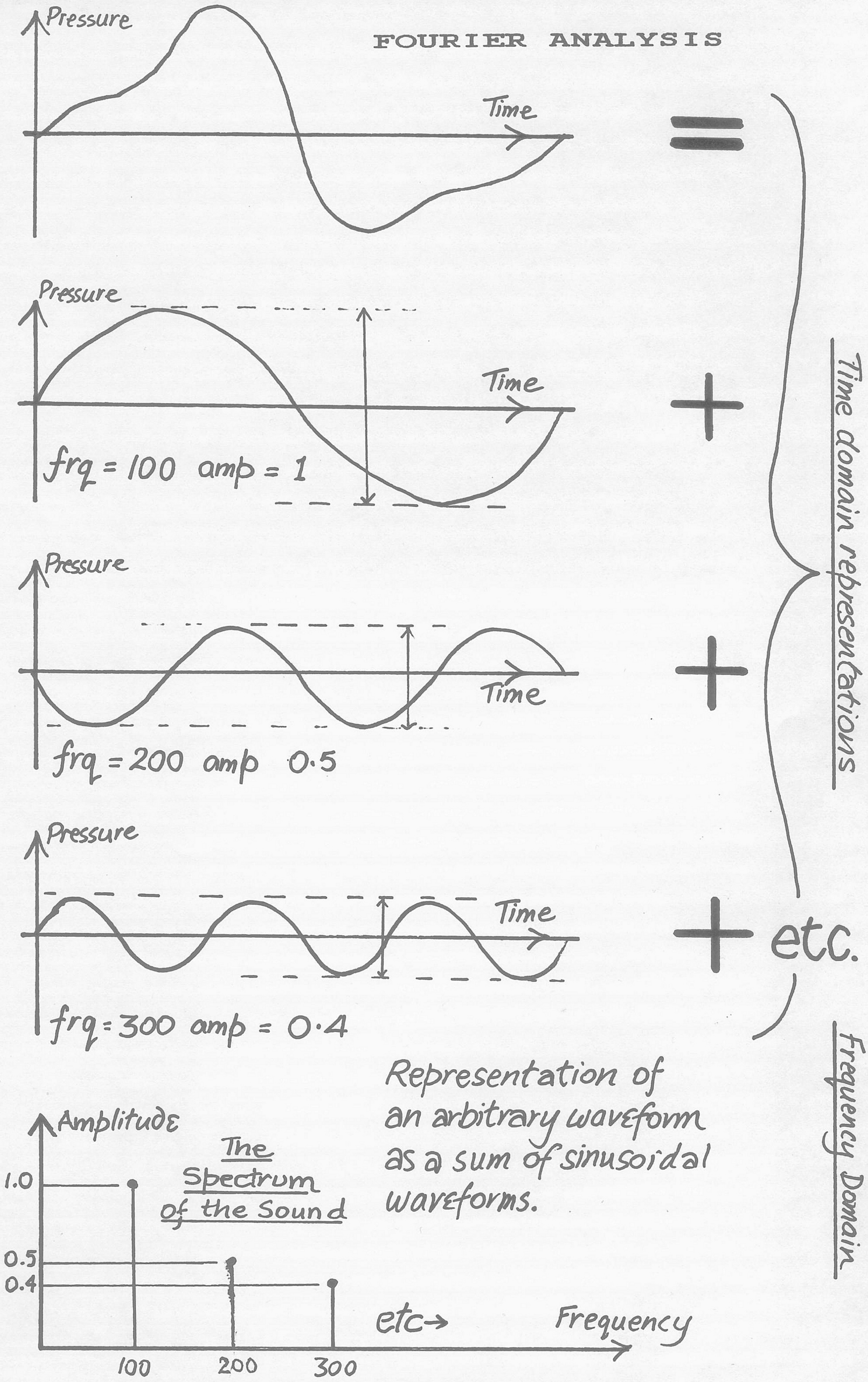


Times completely random.



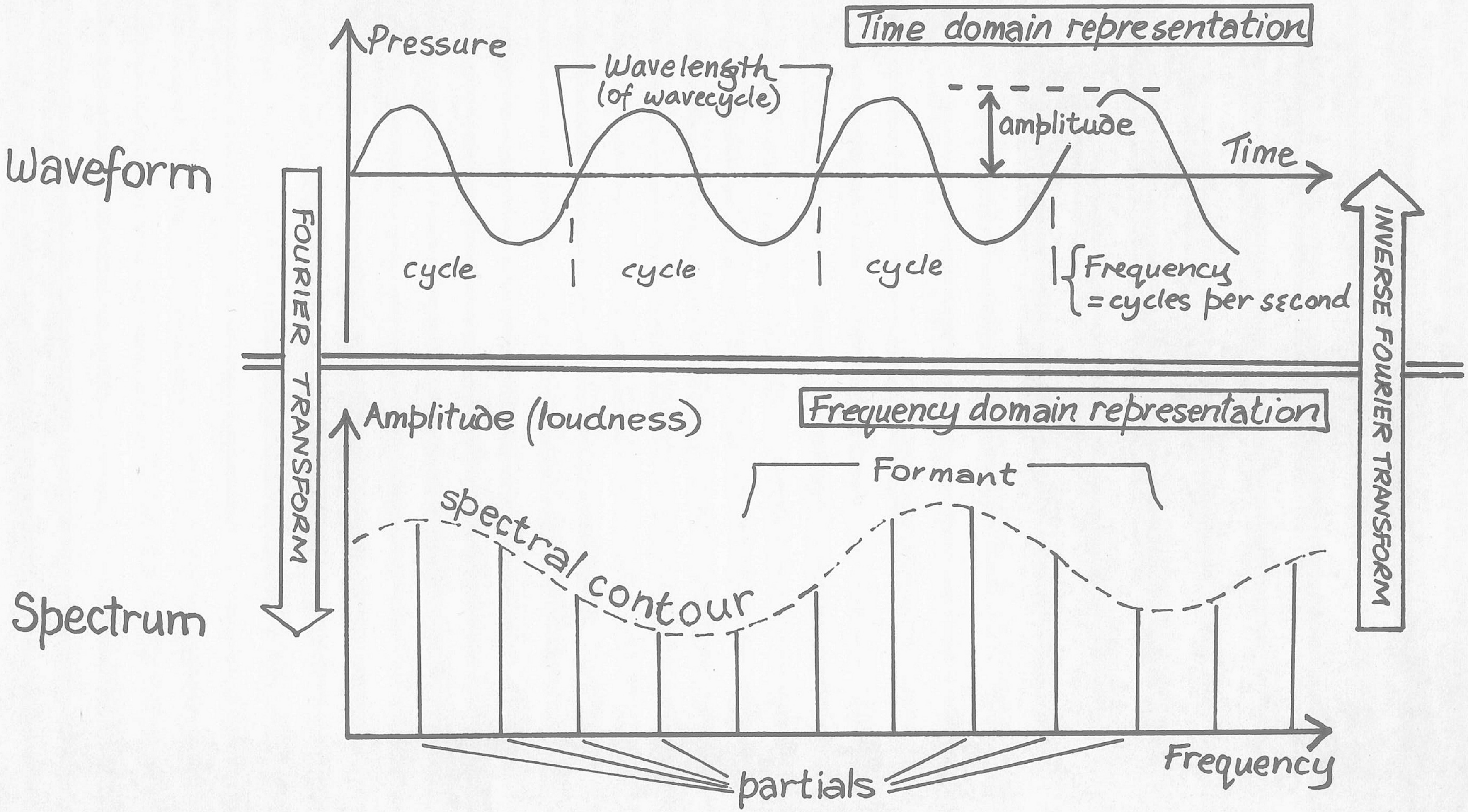
Times completely random, but denser.

FOURIER ANALYSIS



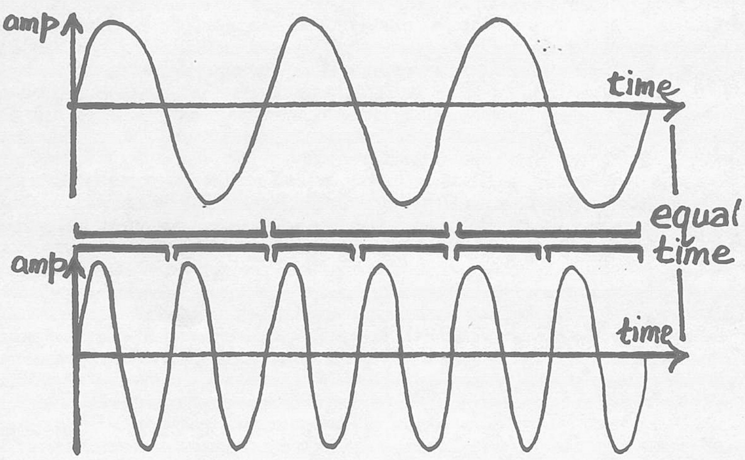
Representation of an arbitrary waveform as a sum of sinusoidal waveforms.

TIME DOMAIN & FREQUENCY DOMAIN

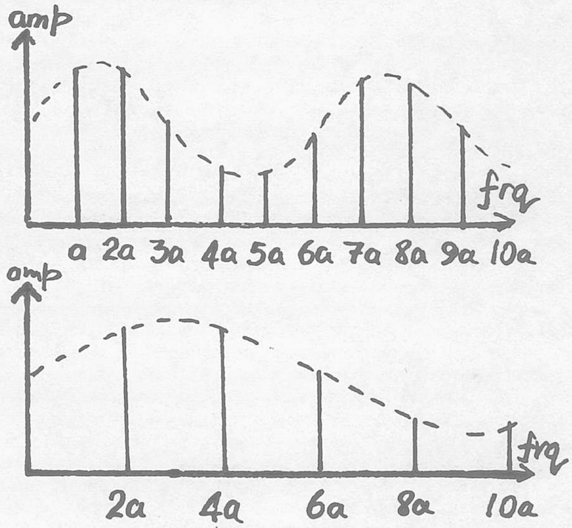


WAVELENGTH, FREQUENCY, PITCH

Time domain



Frequency Domain



When wavelength becomes 1/2 as long,
there are twice as many wavecycles
in the same time.
Frequency = wavecycles-per-second
So
Halve wavelength to double frequency.
Double wavelength to halve frequency.

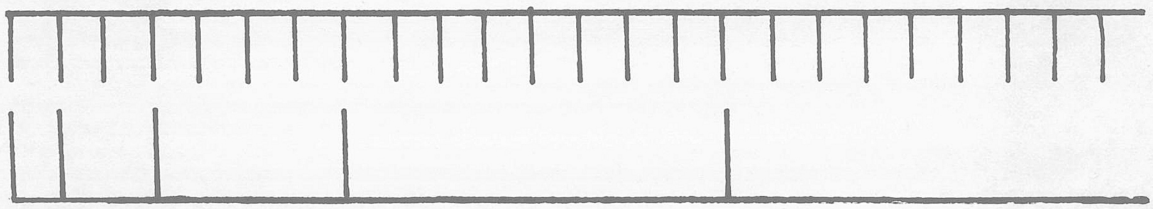
A harmonic spectrum has partials
which are exact multiples of the
1st partial (or fundamental).

Multiplying all partial frequencies
by 2 (or any number)
preserves this relationship.

**WHEN THE FREQUENCY DOUBLES
THE PITCH GOES UP AN OCTAVE.**

Pitch is therefore distributed differently to frequency.

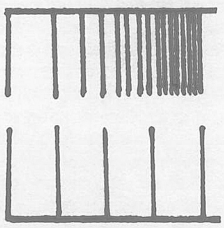
Equal frequency steps.



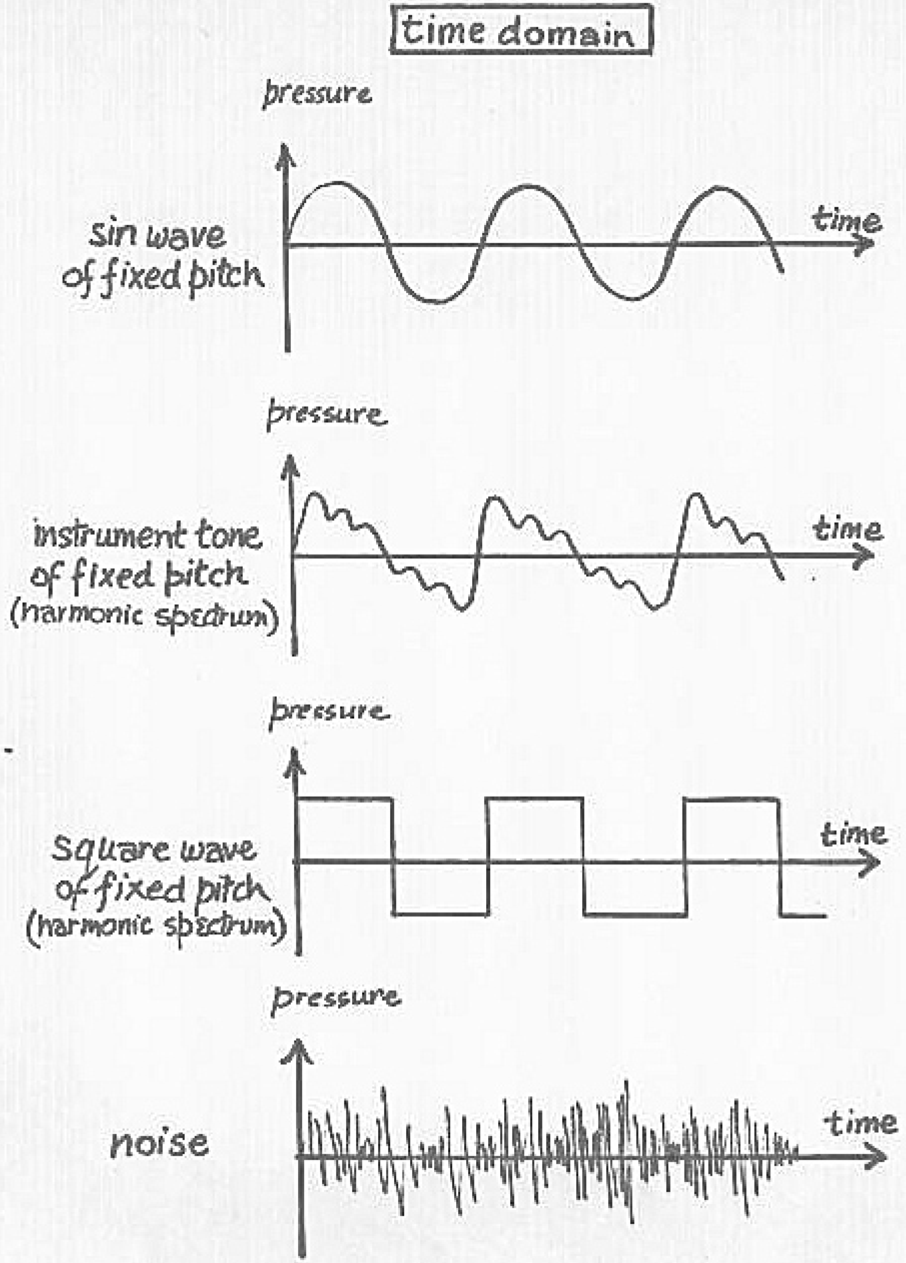
Equal pitch steps (e.g. octaves).

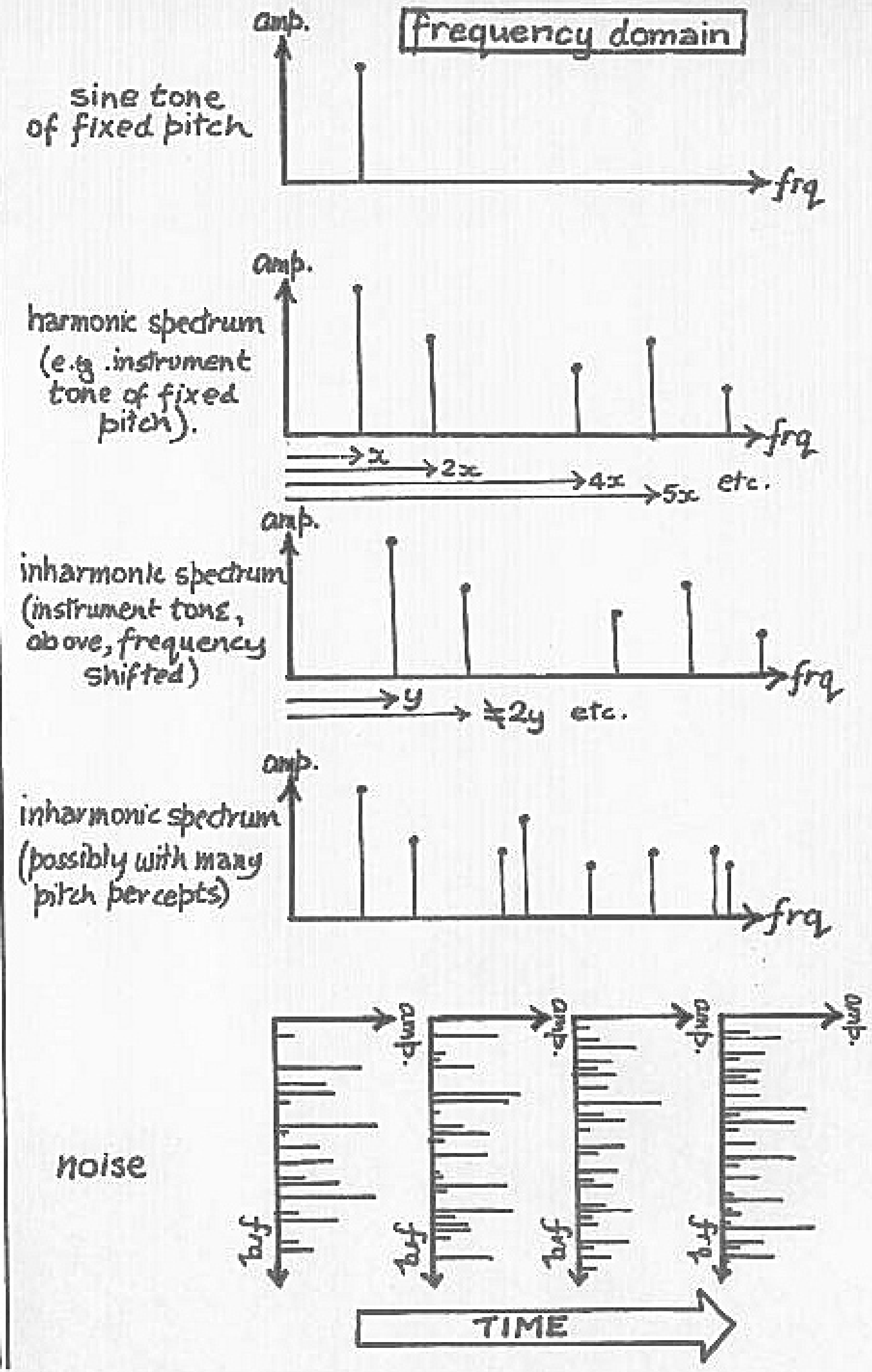
Or, viewed on a scale which makes equal <u>pitch</u> steps look the same...

Equal frequency steps.



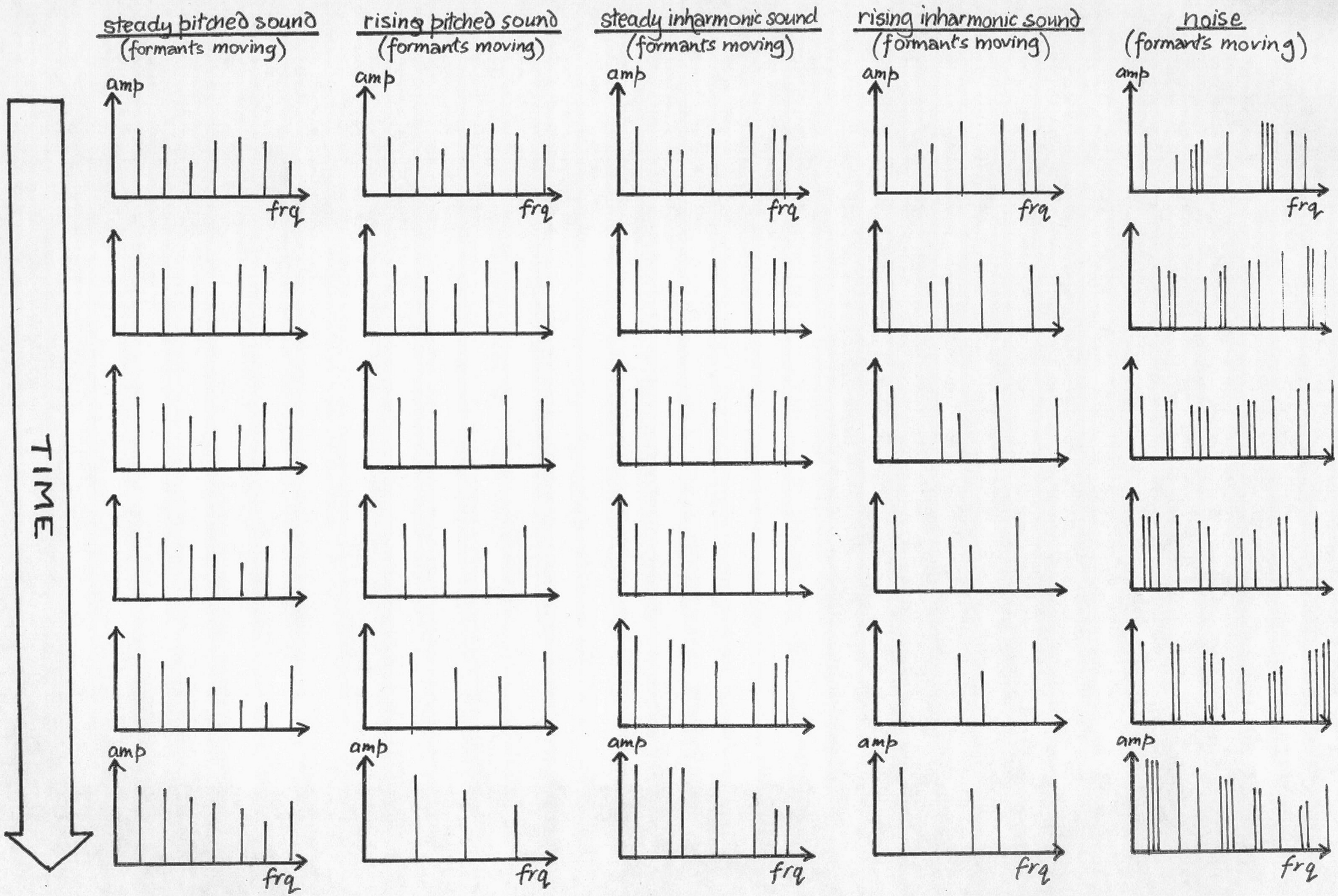
Equal pitch steps.



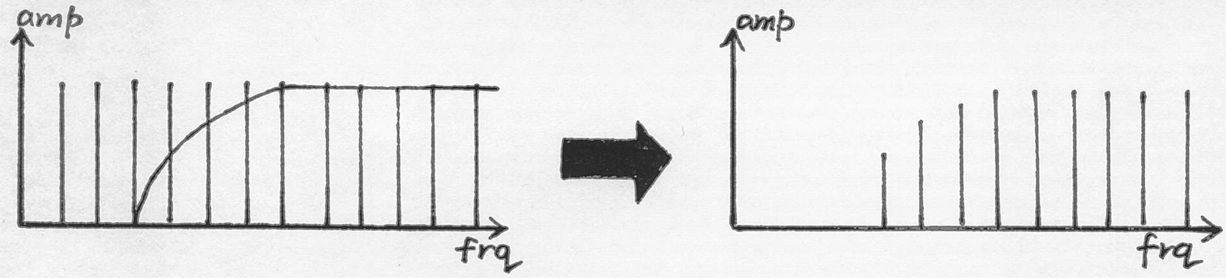


SINE WAVE : SQUARE WAVE

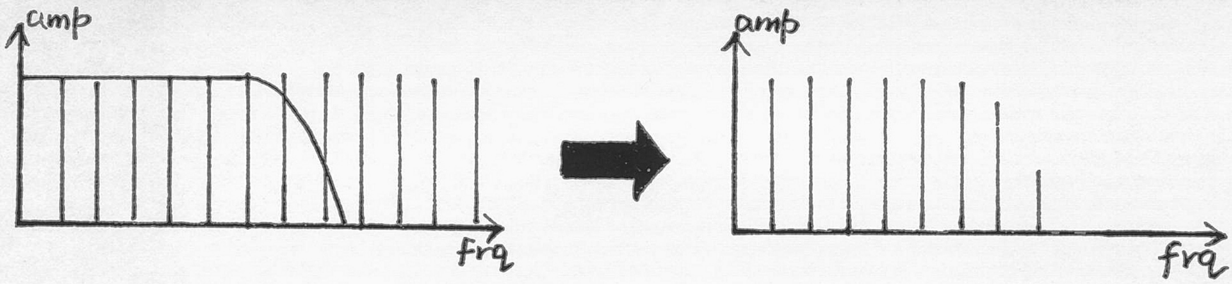
HARMONIC / INHARMONIC / NOISE



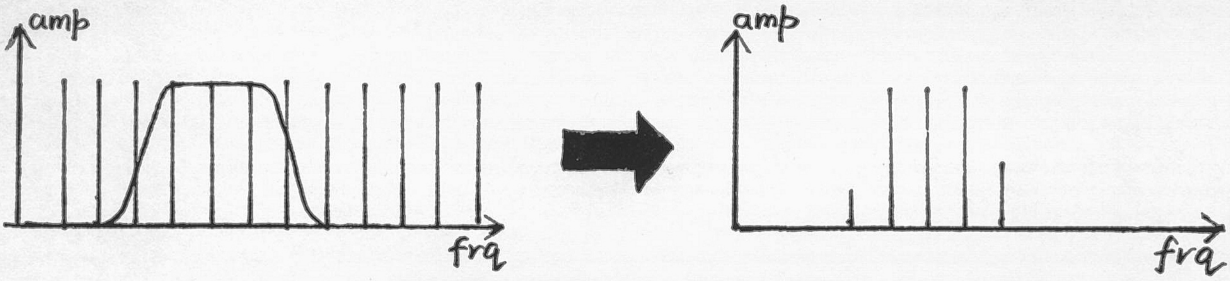
TIME-EVOLUTION OF HARMONIC \ INHARMONIC \ NOISE SPECTRA



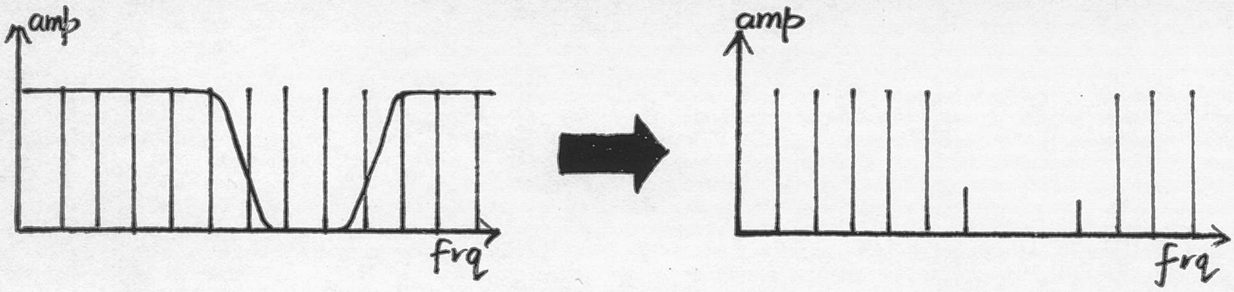
HIGH PASS FILTER



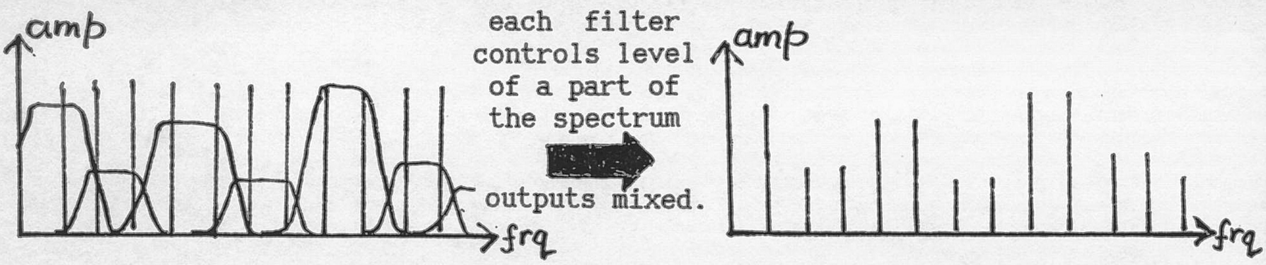
LOW PASS FILTER



BAND PASS FILTER



BAND REJECT or NOTCH FILTER

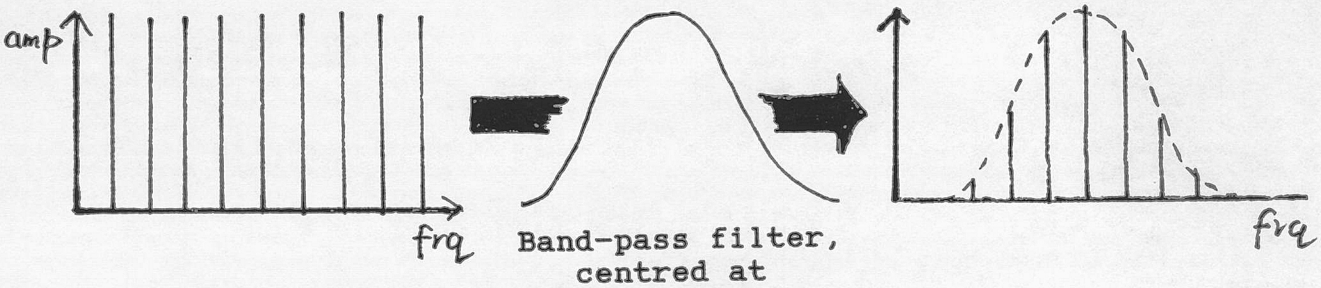


FILTER BANK

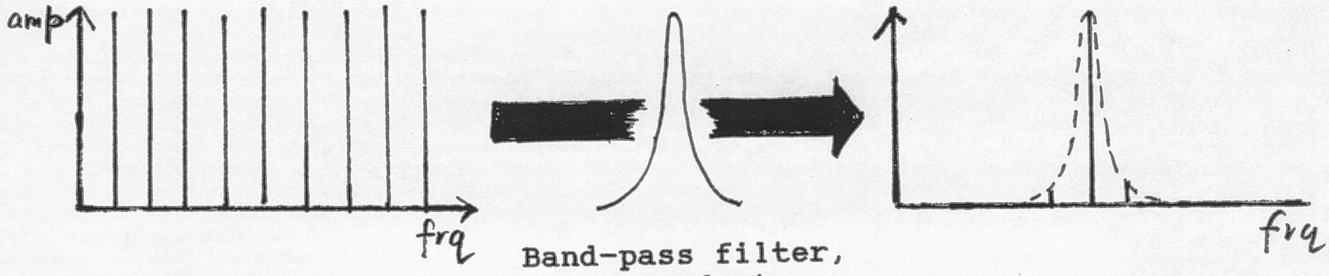
FILTERS continued

Q

The Q of a filter determines steepness of cut-off (see diagrams).

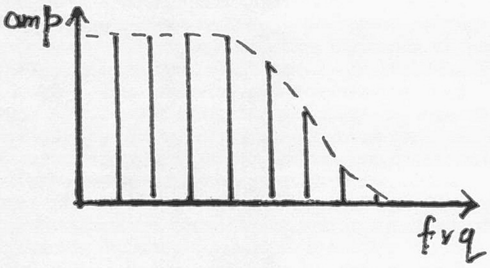


Band-pass filter,
centred at
some frequency,
with low Q.

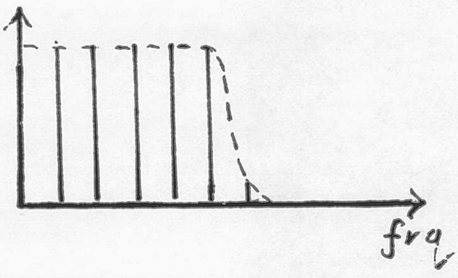


Band-pass filter,
centred at
same frequency,
with high Q.

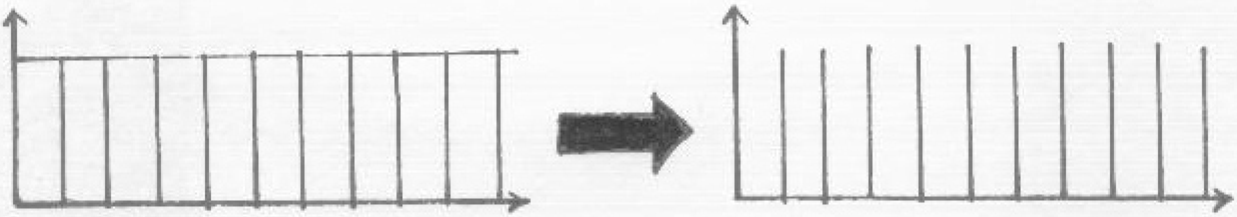
Low pass filter with low Q.



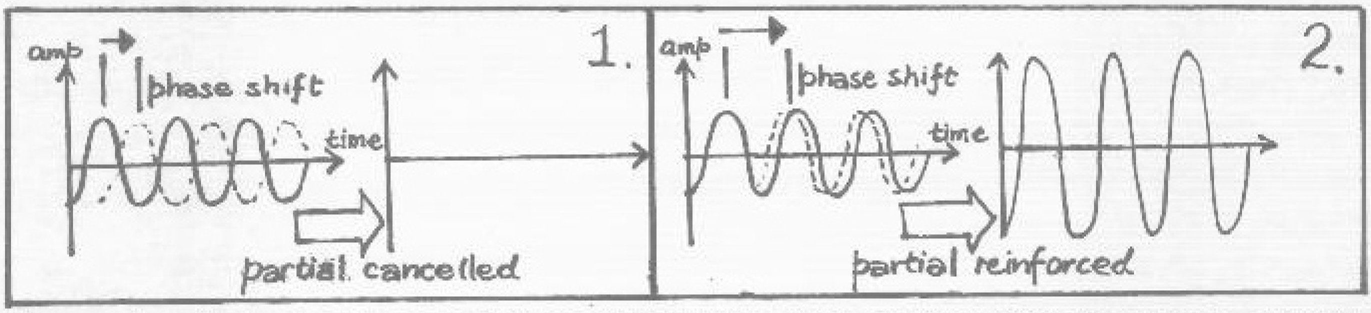
Low pass filter with high Q.
(cut off point at same frequency)



ALL PASS FILTER : PHASING



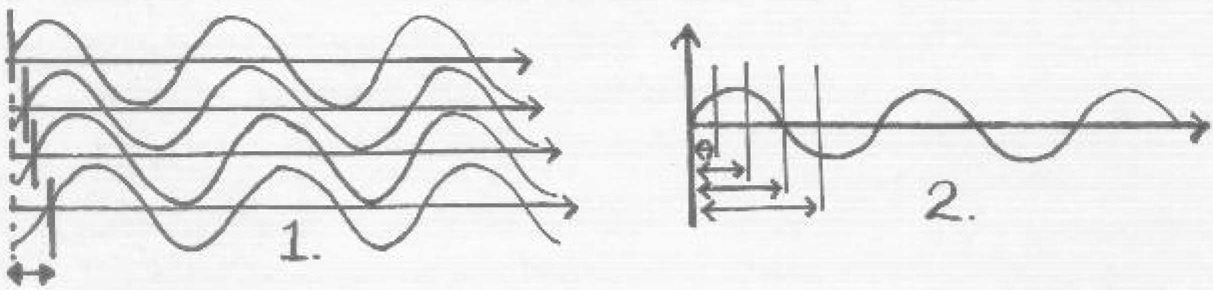
An all pass filter does not alter the frequency content of the spectrum. However, all filters change the phase of the filtered sound (see below), and different frequency ranges in the sound change their phase differently from one another. Hence, if we superimpose an all-pass-filtered sound on the original, we will find that phase shifts in some ranges will cause partials to be cancelled out (see diagram), whilst in other ranges partials will be reinfoced (see diagram).



The way in which an all-pass filter changes the phase of different frequency ranges in a sound is an aspect of its design. And it can be made to vary with time. If a sound is all-pass-filtered in a time-varying manner, and the result mixed with the original sound, we will hear frequency bands sweeping up or down across the sound, as these bands are reinforced or cancelled. This 'effect' is often used in popular music, and is known as PHASING.

PHASE

Two sounds may have the same waveform, but the waveforms may begin at different points in the wavecycle (see diagram). The point in the cycle at which the waveform begins is referred to as the phase.

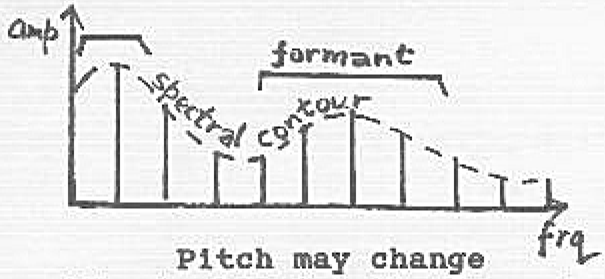


When the waveform begins at the start of the sound, the phase is zero. If it begins later the phase is greater, and has a maximum when the sound begins just before the end of the waveform.

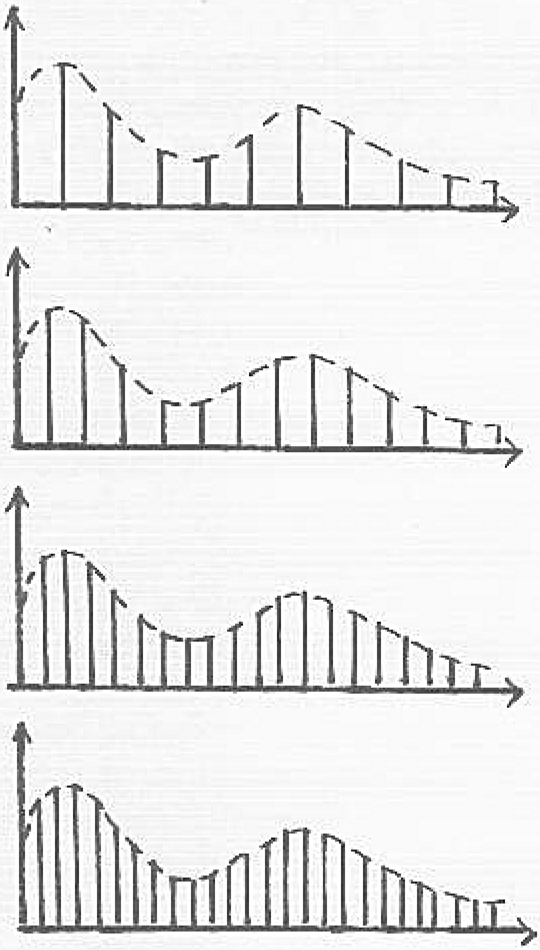
Alternatively, moving along a given waveform, the phase increases (see diagram 2). The rate at which the phase increases is equal to the frequency of the wave. This fact enables the Phase Vocoder to track the frequencies of partials in a sound.

FORMANTS

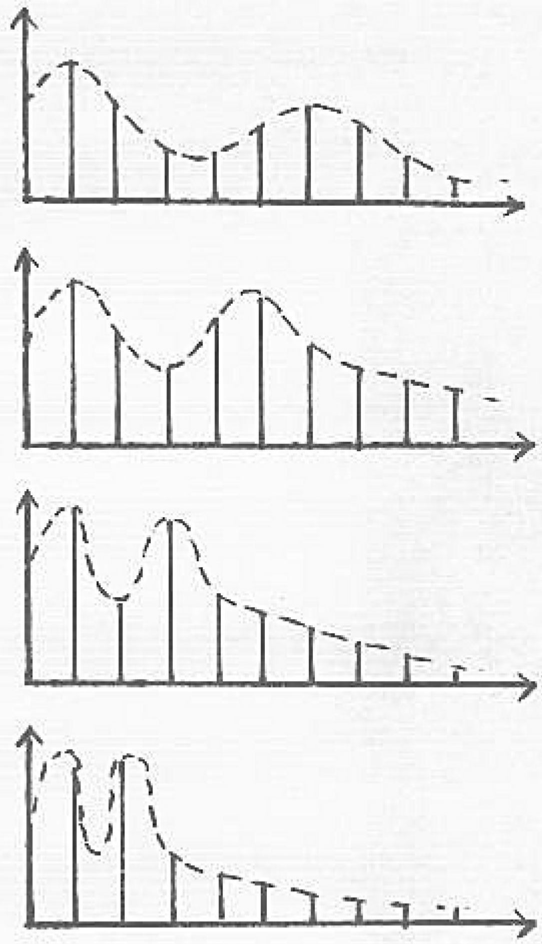
Formants are peaks in the spectral contour.



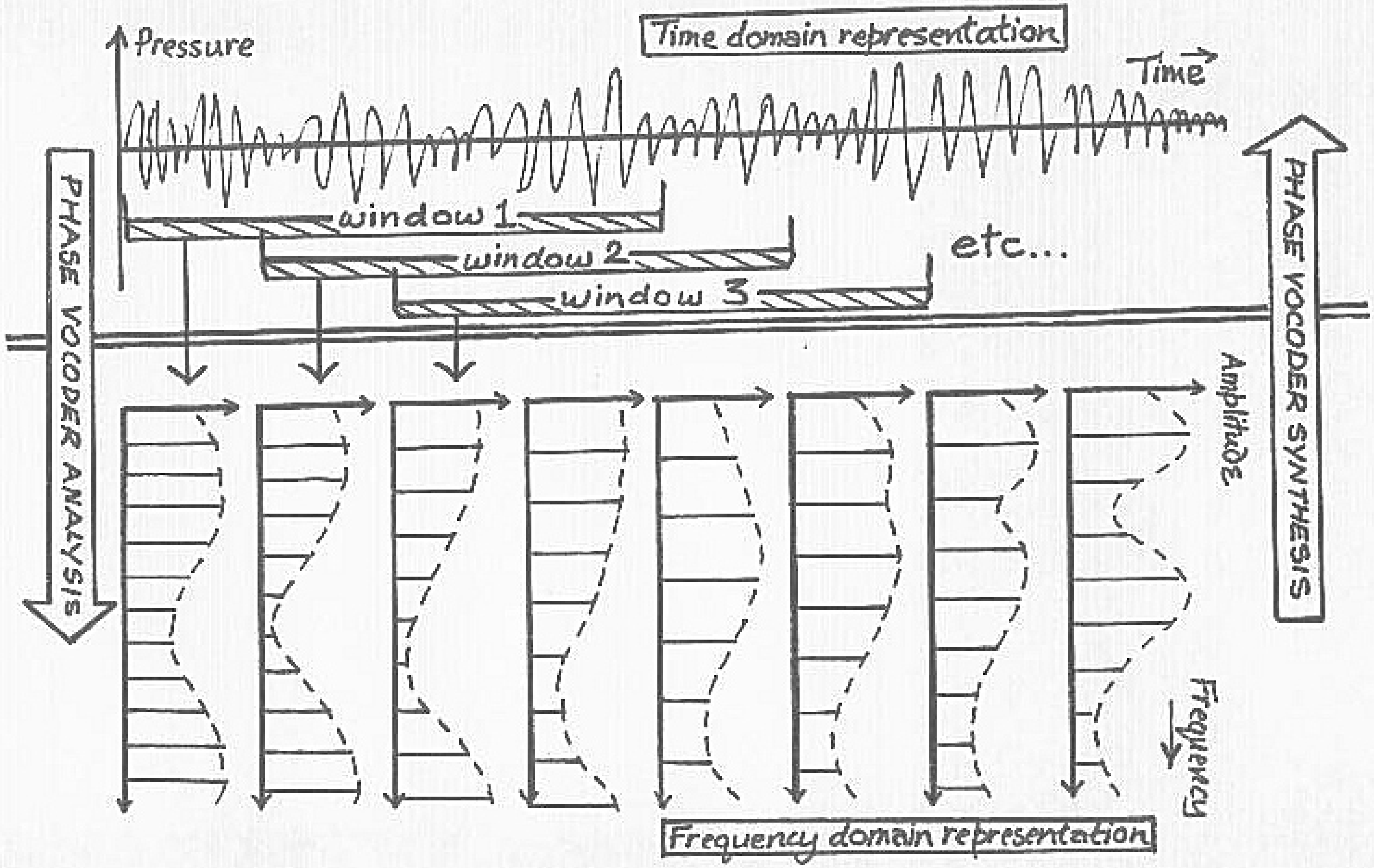
Pitch may change
without changing the formants.
e.g. sing a downward sliding
pitch on a fixed vowel 'a'.



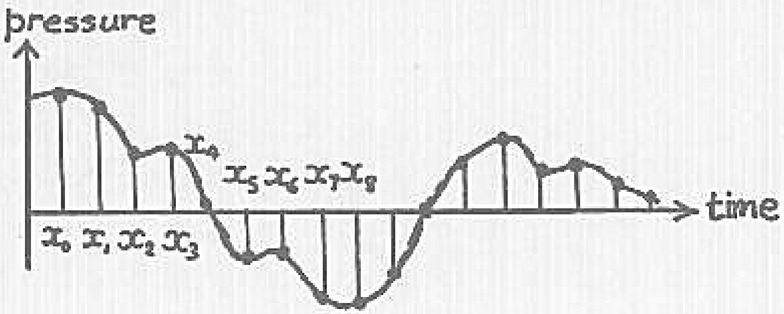
Formants may change
without changing the pitch.
e.g. sing 'a' -> 'u'
on a fixed pitch.



PHASE VOCODER



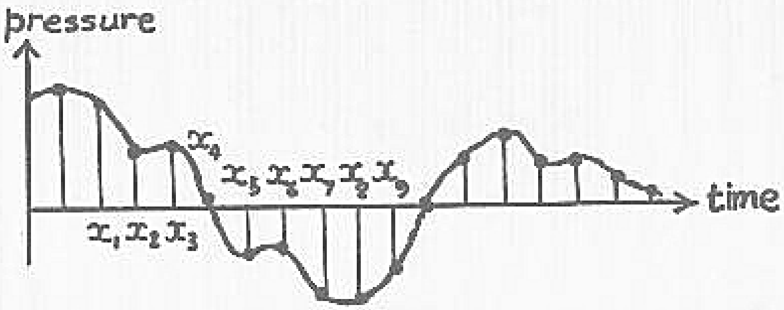
LINEAR PREDICTIVE CODING (LPC)



The waveform above has been sampled where the vertical lines occur, to give the values $x_0, x_1, x_2 \ldots . x_8$.
Can we <u>predict</u> the value of x_8 from the previous 8 values. If we can, we could write down an equation...

$$(1)\quad x_8 = Ax_0 + Bx_1 + Cx_2 + \ldots\ldots + Hx_7$$

Clearly, by chosing appropriate values for A-H (the coefficients) we can always make this equation true. In fact there are many, many ways to chose a set of A-H which will all make the equation true.



Can we predict the value of x_9 from the previous 8 values, <u>using the same set of coeefficients A-H</u>. If we can, we could write down an equation...

$$(2)\quad x_9 = Ax_1 + Bx_2 + Cx_3 + \ldots\ldots + Hx_8$$

This is a much harder problem. We cannot choose just any set A-H which satisfies equation 1, and expect them also to satisfy equation 2. In fact, in general, there will be no set A-H which satisfies both these equations.
We are therefore looking for a <u>best fit</u> so that there is the minimum of error in the predicted values of both x_8 and x_9 using our chosen coefficients.

Let us now define a time-window of (say) 64 samples. Can we choose coefficients A-H such that we can predict every sample in the window from the previous 8, using these <u>same</u> coefficients?

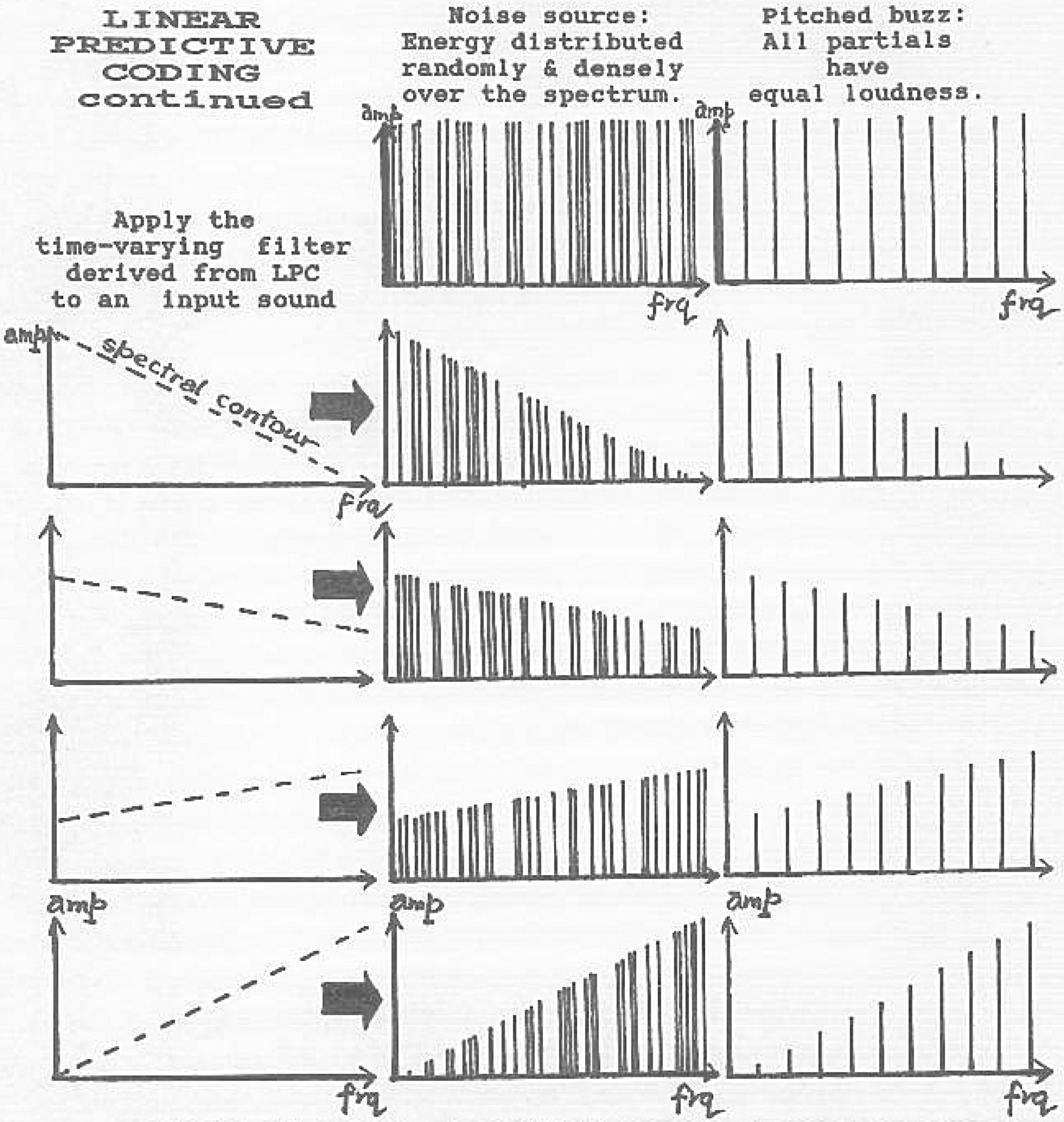
If we can do this we can use them to make the signal <u>predict itself</u>.

We now proceed to the next time-window and look for a <u>new</u> set of coefficients (I-P) to predict the samples there, in the same fashion.

The coefficients turn out to be the numbers we need to define a filter - and all the sets of coefficients together, a time-varying filter. These filters define the contour of the <u>spectrum</u> of our input sound, from moment to moment.

In practice we need about 40 coefficients (predicting each sample in our window from the previous 40 samples) to achieve a reasonable resolution of the spectrum.

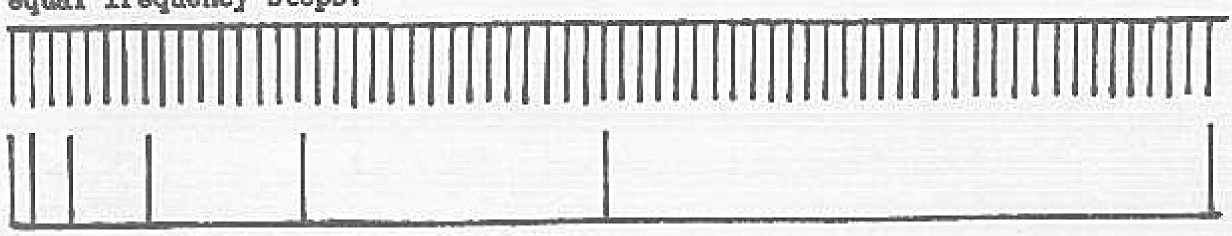
LINEAR PREDICTIVE CODING continued



e.g. analysis of speech -> unvoiced speech -> monotone speech (but the input buzz could have time-varying pitch)

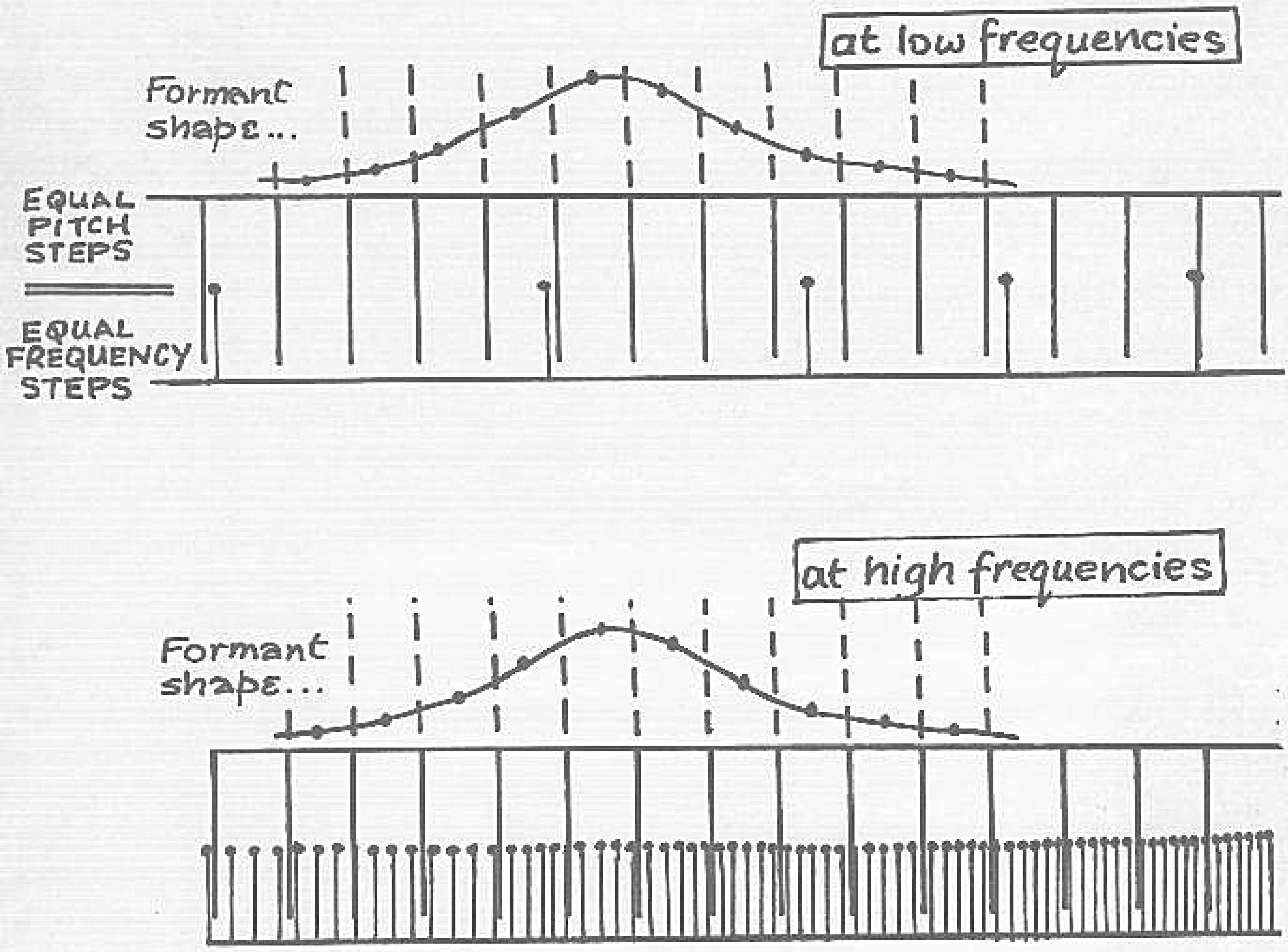
LPC can be set up to generate a filter bank which is distributed evenly with respect to pitch. In contrast, the Phase Vocoder channels are distributed evenly with respect to frequency.

equal frequency steps.



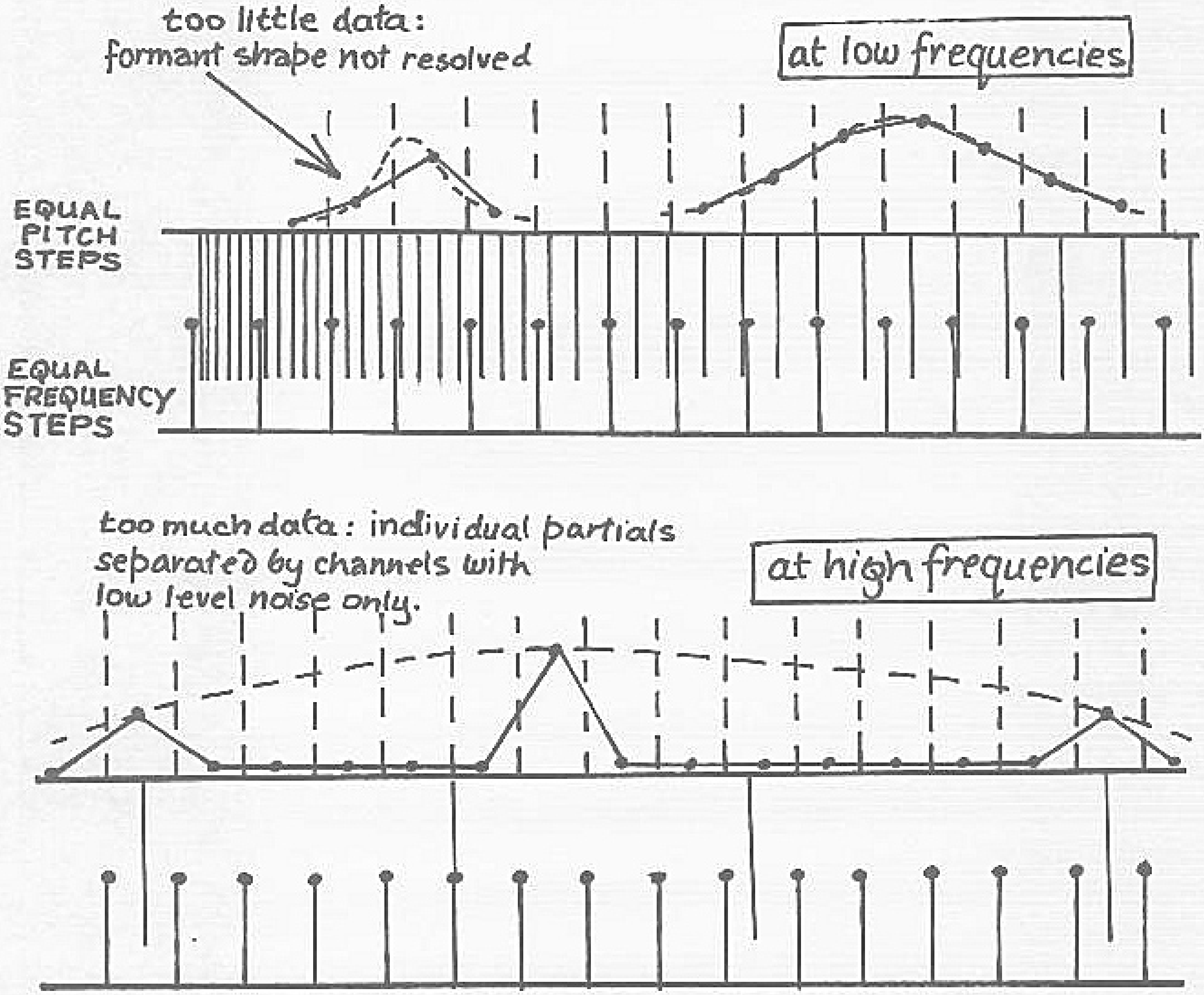
equal pitch steps.

Distribution of Formants relative to Pitch



Formants have similar pitch-width. If analysis channels are equally spaced in pitch (as they may be, in Linear Predictive Coding) then formant envelopes are resolved equally well in all ranges.

Distribution of Formants relative to Frequency

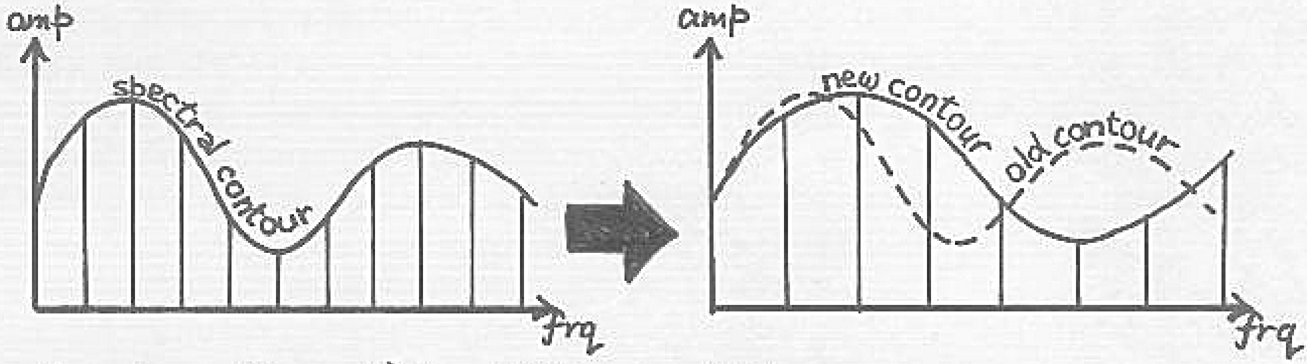


Formants have similar pitch-width. Hence their frequency-width increases as we go up the spectrum. The phase vocoder channels are equally spaced in frequency so do not resolve the formant envelopes well in all ranges.

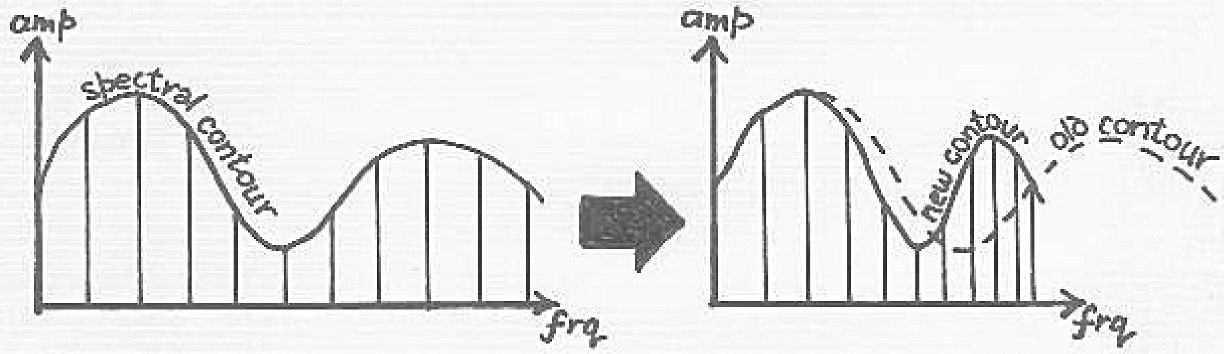
PHASE VOCODER AND FORMANTS

CHANGING THE SPECTRUM

USUALLY CHANGES SPECTRAL CONTOUR



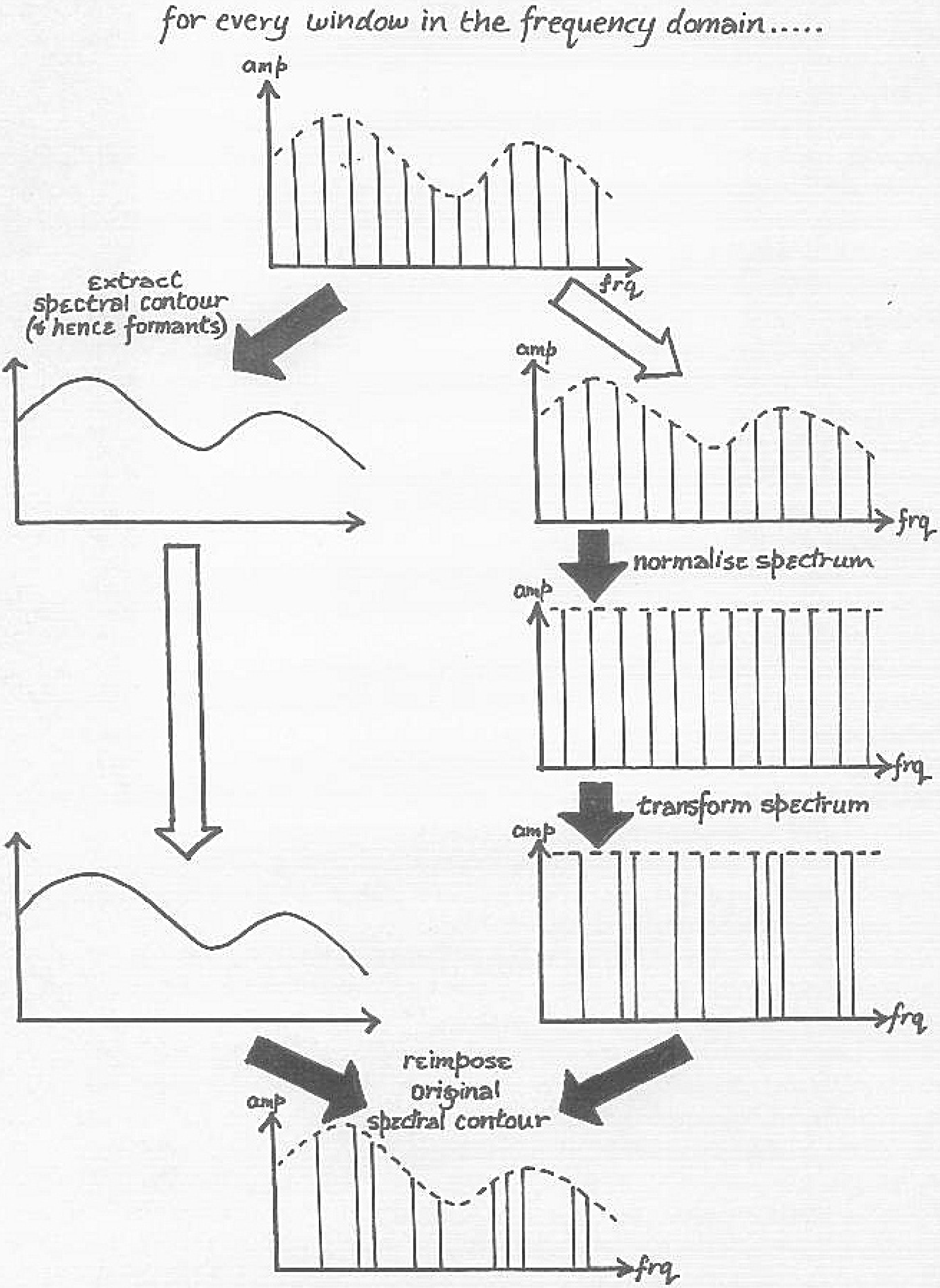
Transpose (upwards) by shifting partials.



Transform (spectral stretch by less than 1 = shrink) by shifting partials.

Pitch transposition & spectral transformation do not normally preserve the spectral contour.

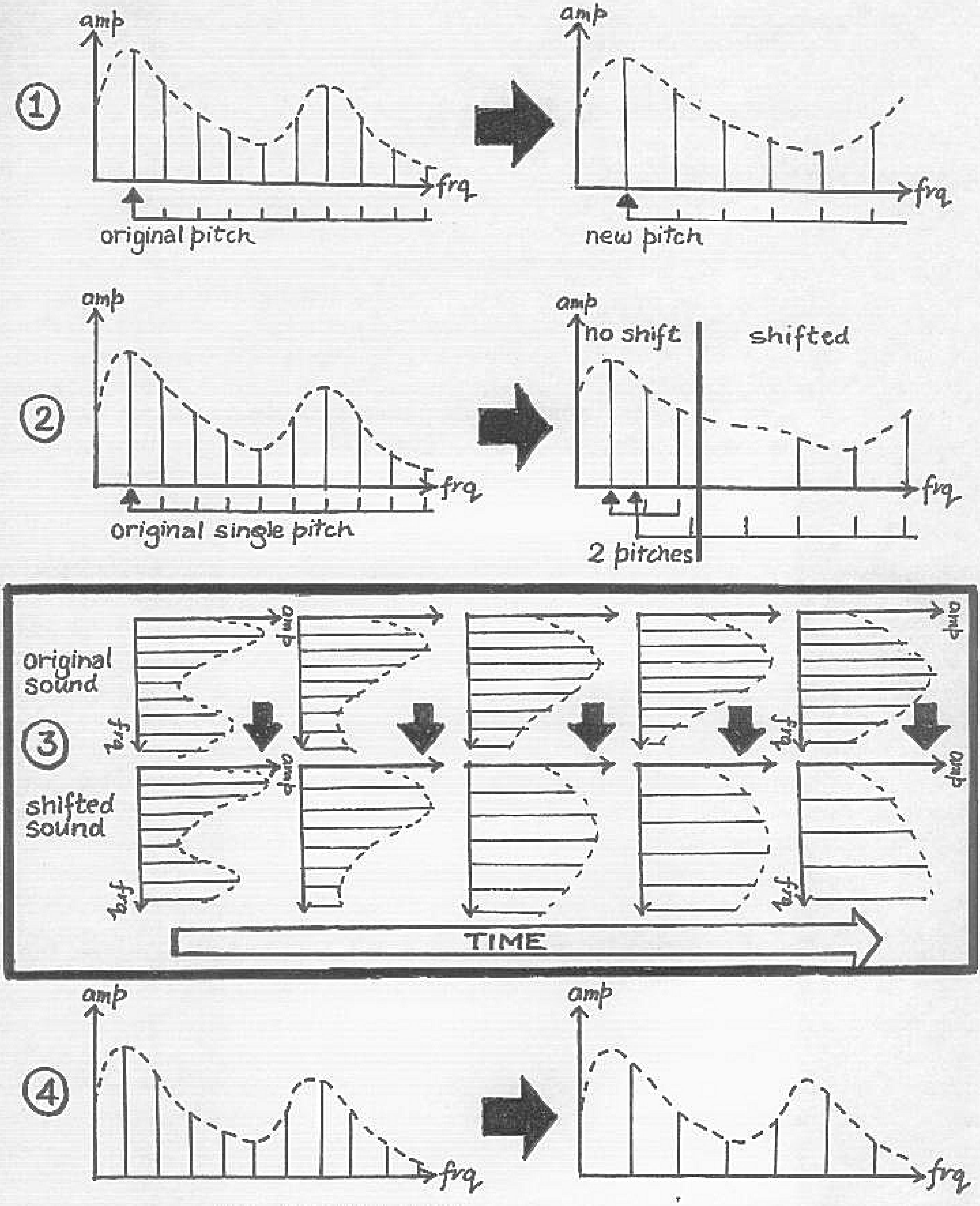
FORMANT PRESERVING SPECTRAL MANIPULATION



FORMANT PRESERVING PITCH SHIFTING

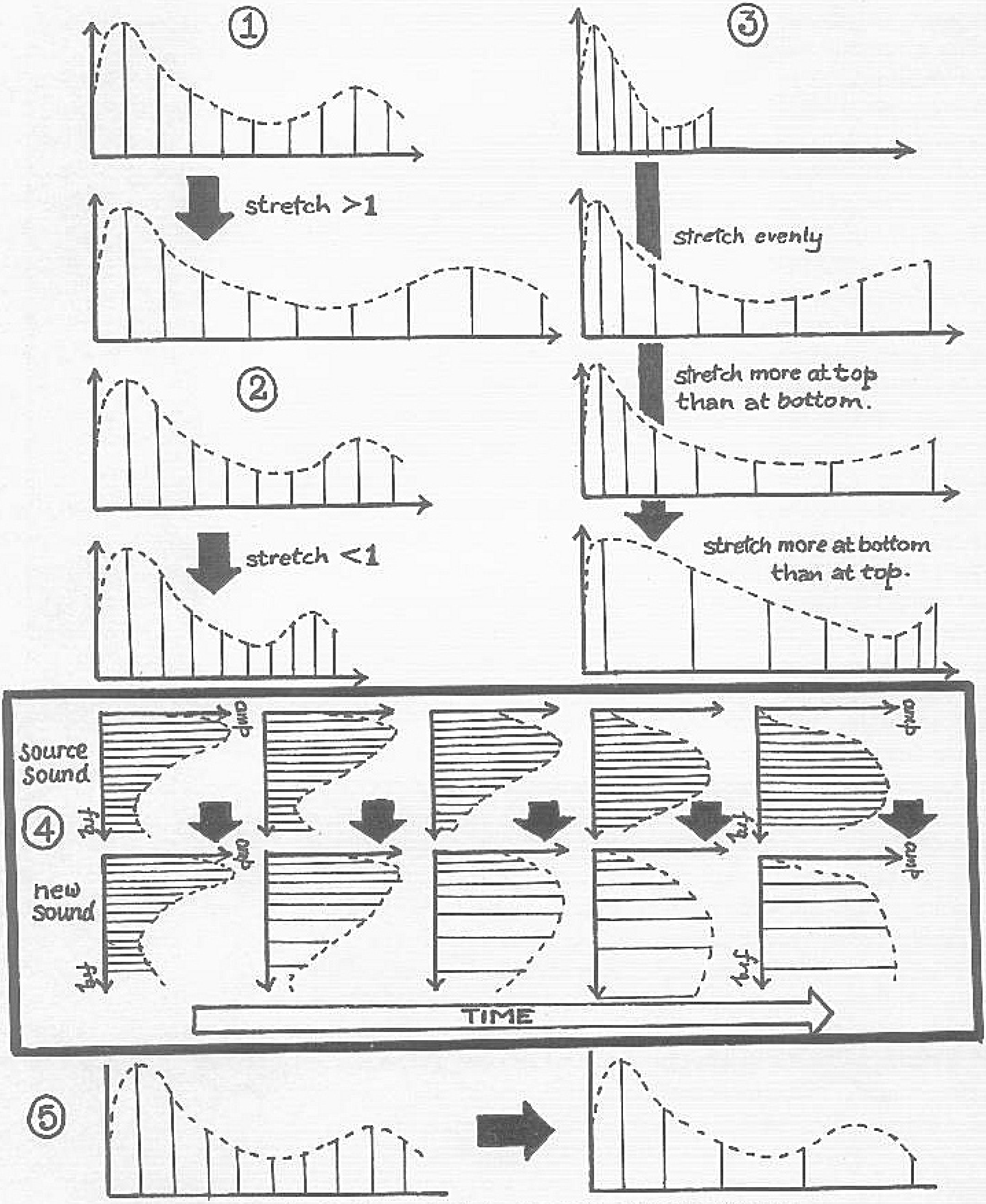
works similarly

SPECTRAL SHIFTING



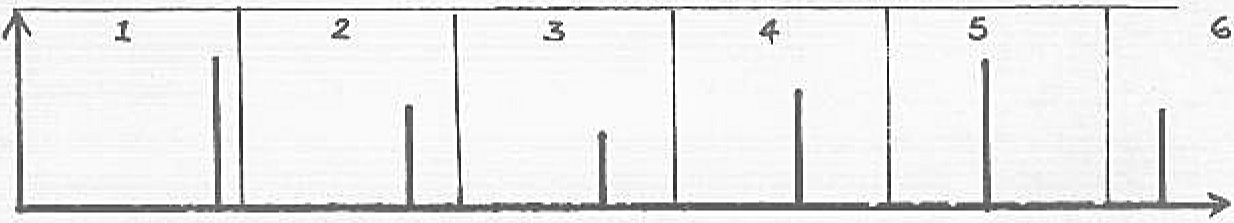
(1) Spectral shift.
(2) Spectral shift of part of spectrum only.
(3) Time-varying (increasing) spectral shift.
(4) Formant-preserving spectral shift.

SPECTRAL STRETCHING

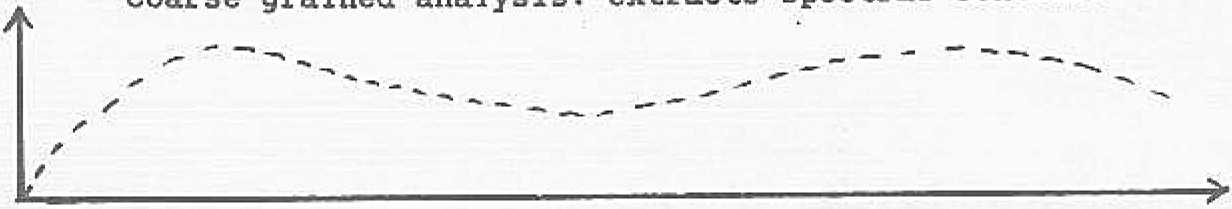


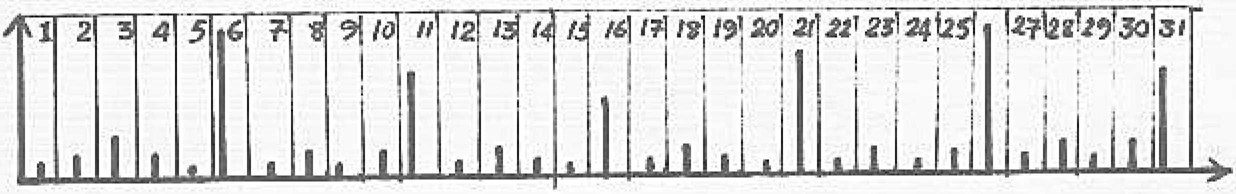
(1) Spectral stretching: total stretch greater than 1.
(2) Total stretch less than 1 (spectral shrinking).
(3) Different types of stretching.
(4) Time-variable (total stretch increasing) spectral stretching.
(5) Formant preserving spectral stretching.

SPECTRAL FOCUSING

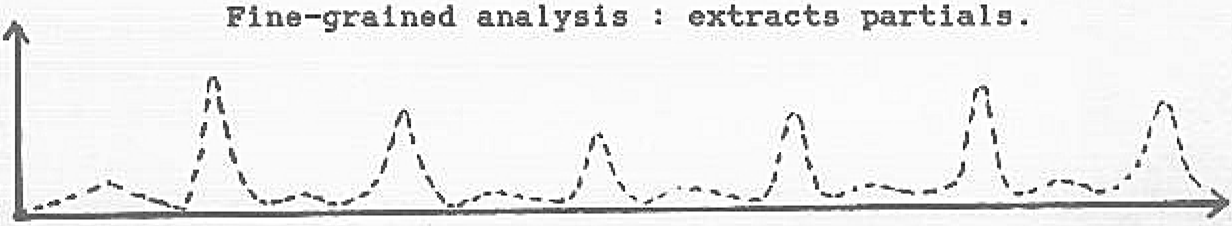


Coarse grained analysis: extracts spectral contour.

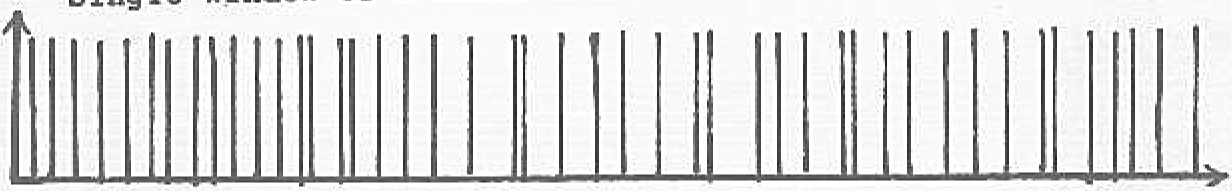




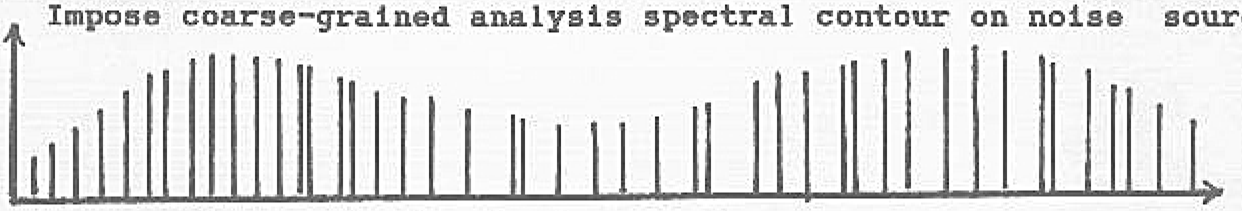
Fine-grained analysis : extracts partials.



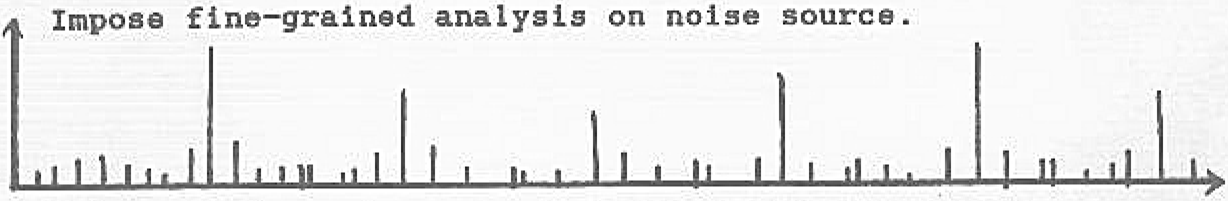
Single window of a noise source.



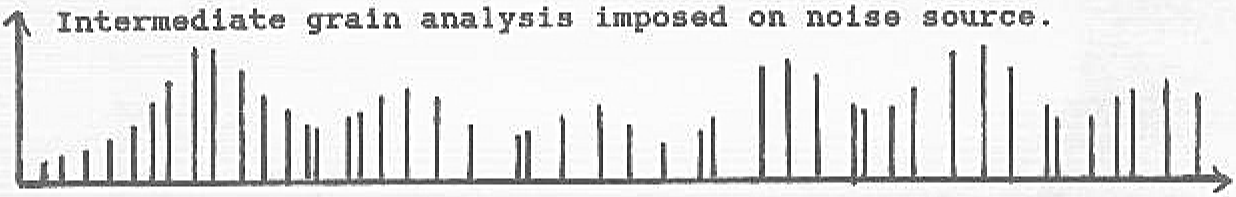
Impose coarse-grained analysis spectral contour on noise source



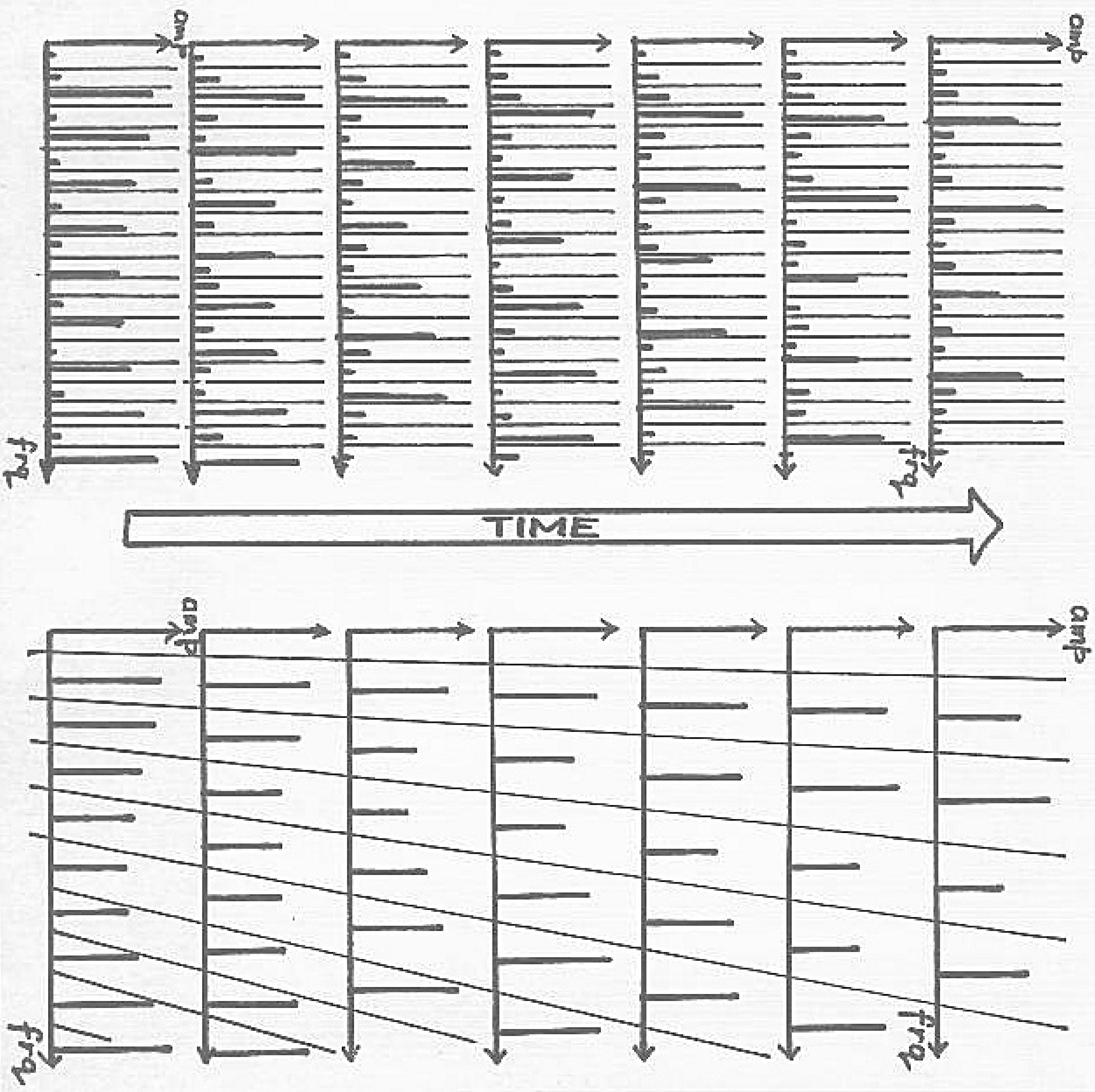
Impose fine-grained analysis on noise source.



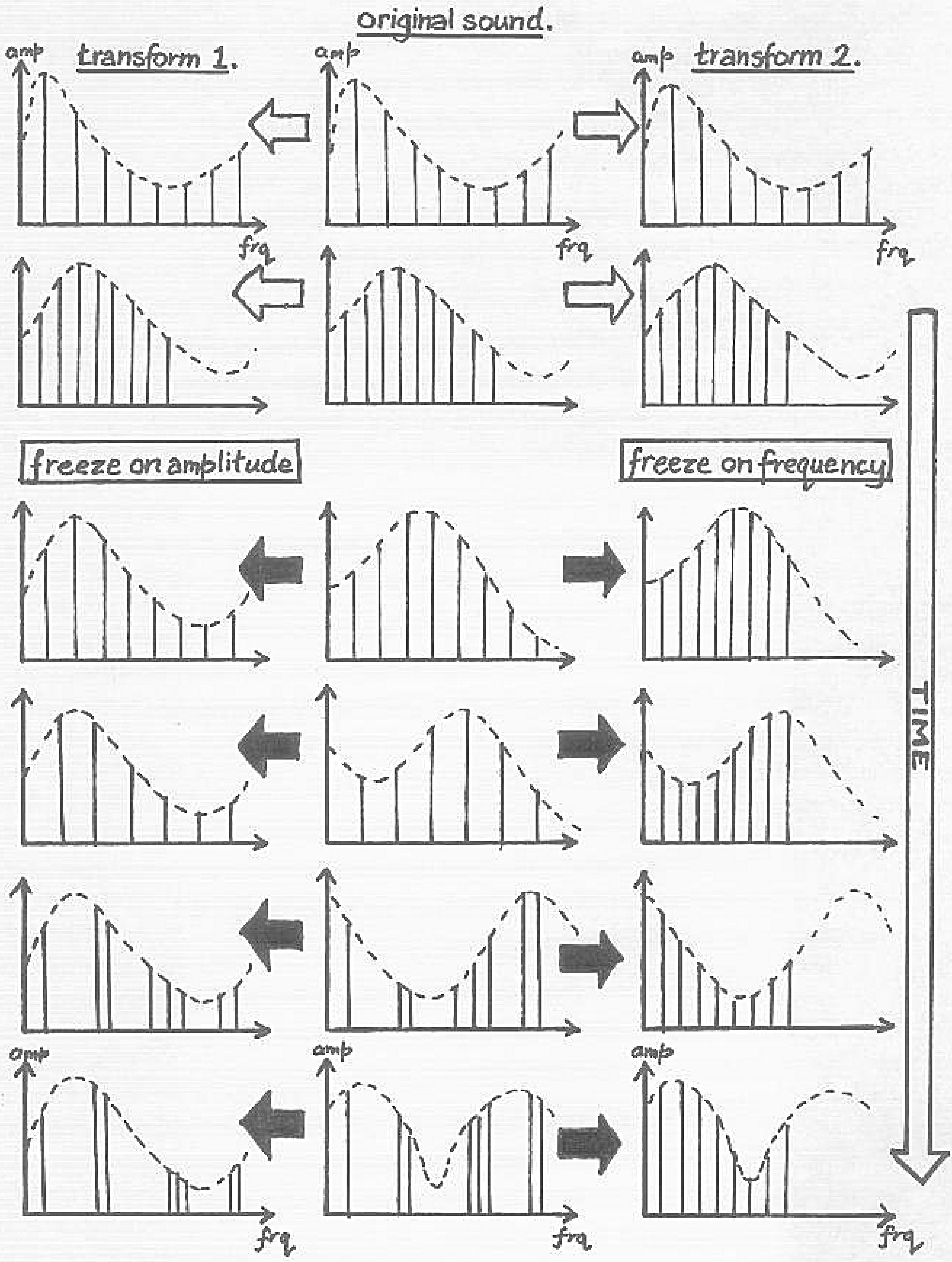
Intermediate grain analysis imposed on noise source.



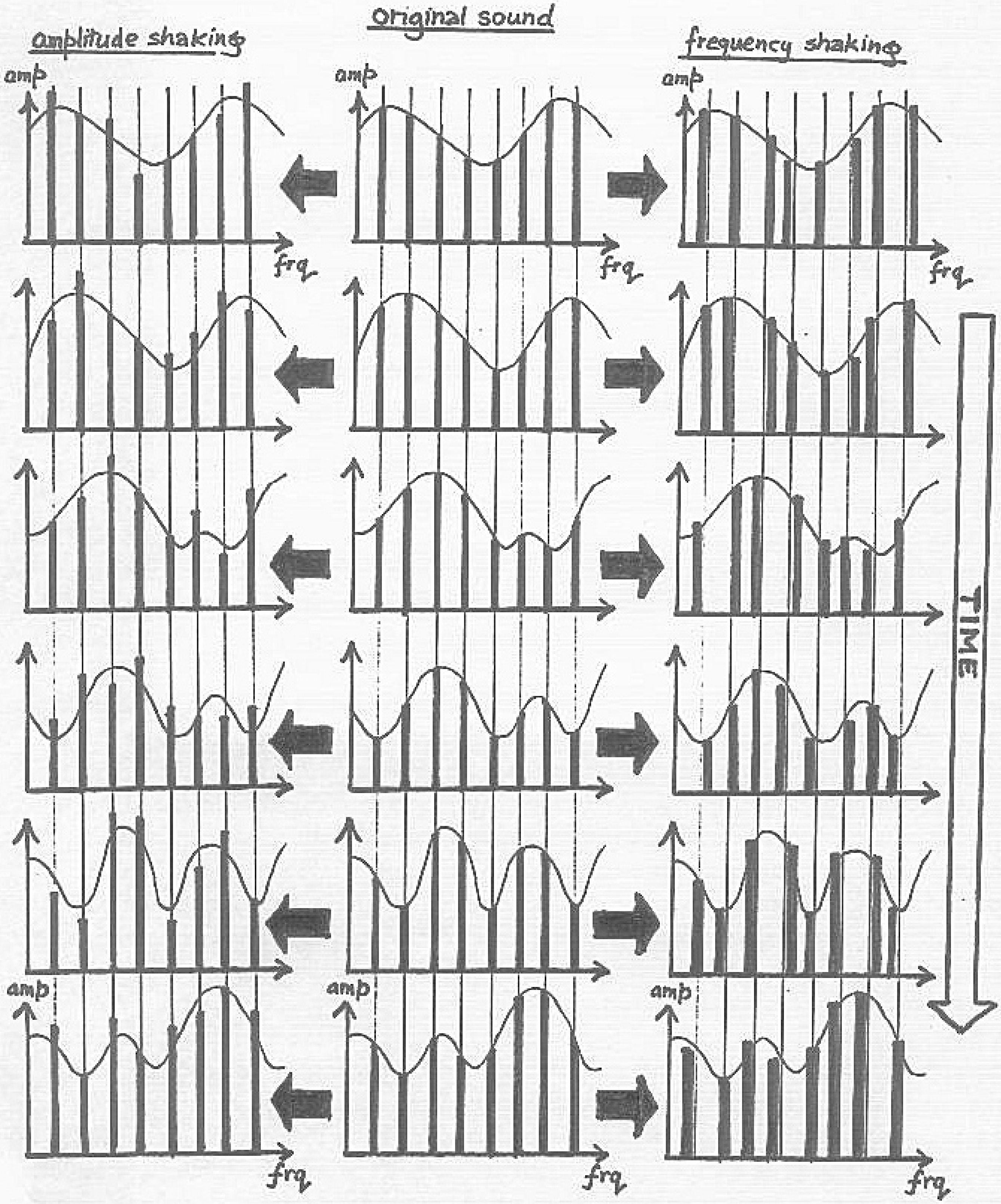
PARTIAL TRACKING



Track each partial, regardless of which analysis channel it falls in.

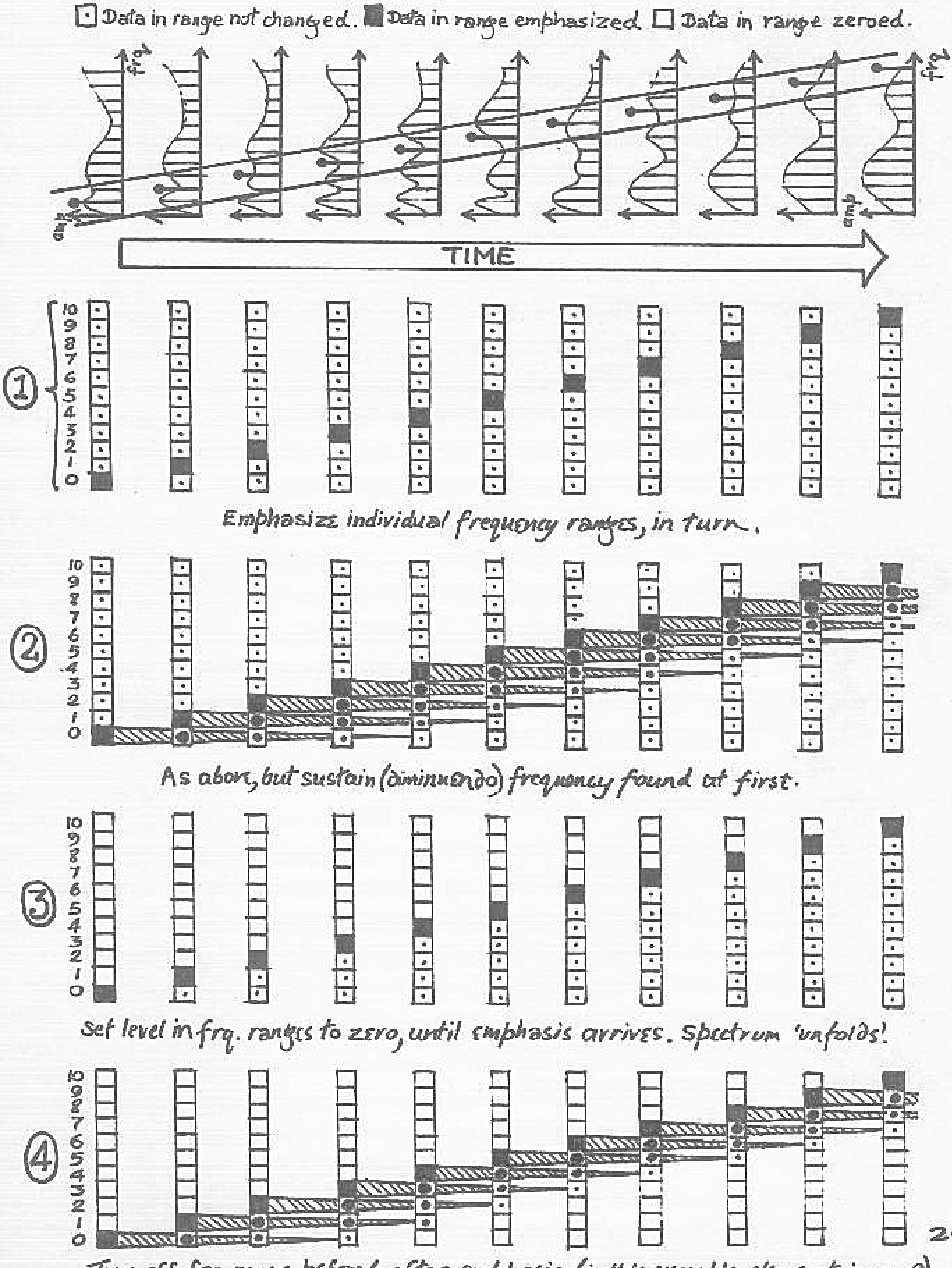

original sound.
transform 1.
transform 2.
amp
frq
freeze on amplitude
freeze on frequency
TIME

SPECTRAL SHAKING



SPECTRAL ARPEGGIATION

Emphasize different elements of spectrum in successive time-windows. Emphasis-by-loudness may sweep over phase vocoder channels or composer-defined frequency ranges. The sweeping motion may be in any composer-specified shape (e.g. sinewave, ramp etc.) and at any (possibly time-varying) speed.



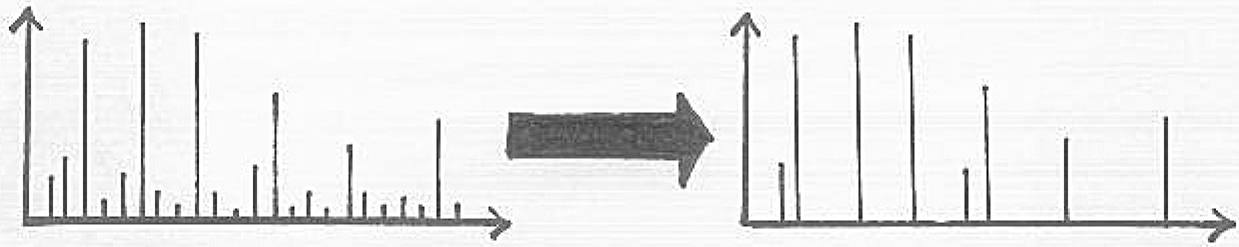
Emphasize individual frequency ranges, in turn.

As above, but sustain (diminuendo) frequency found at first.

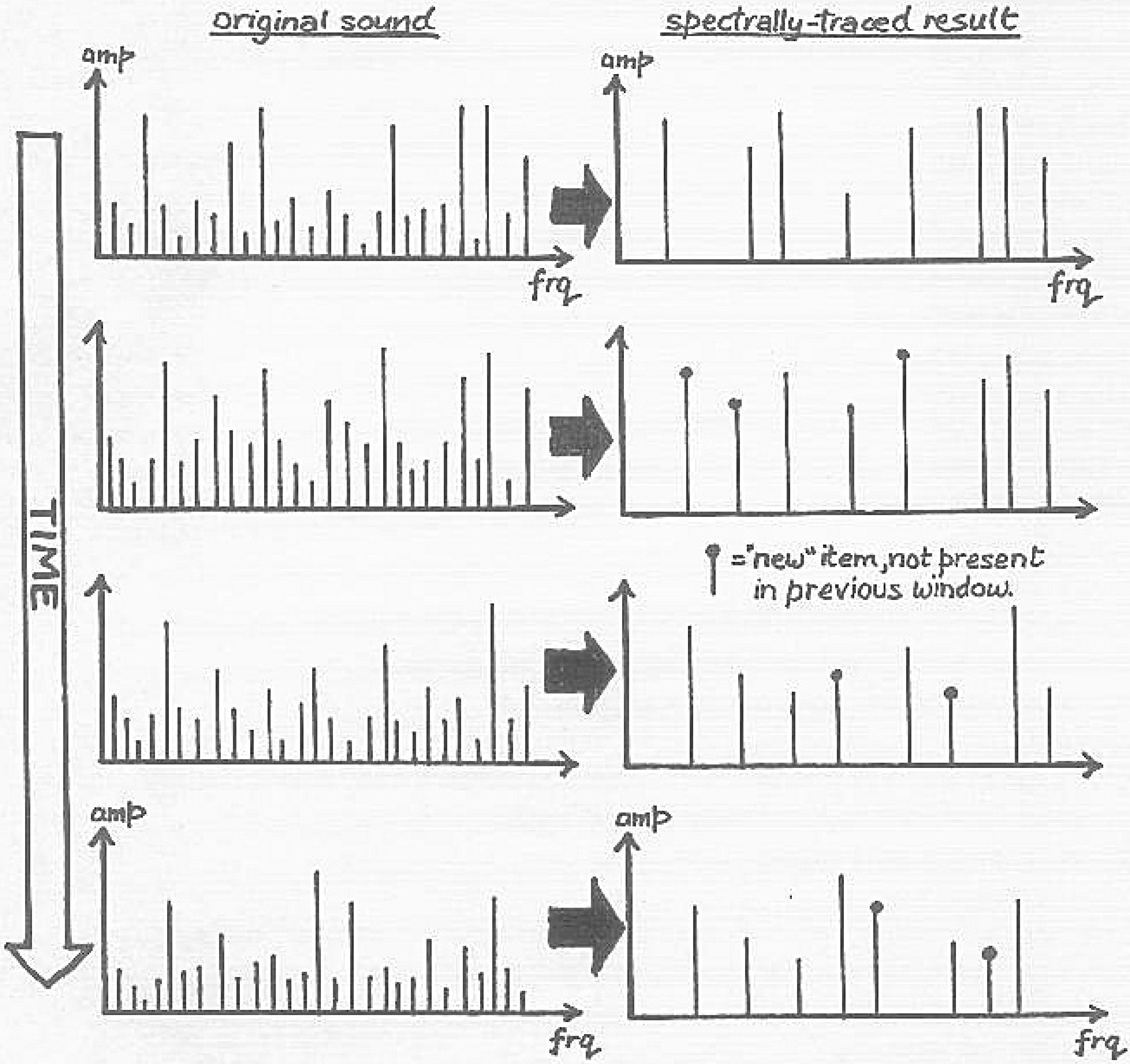
Set level in frq. ranges to zero, until emphasis arrives. Spectrum 'unfolds'.

Turn off frq. range before & after emphasis (in this example, also sustain, as 2).

SPECTRAL TRACING



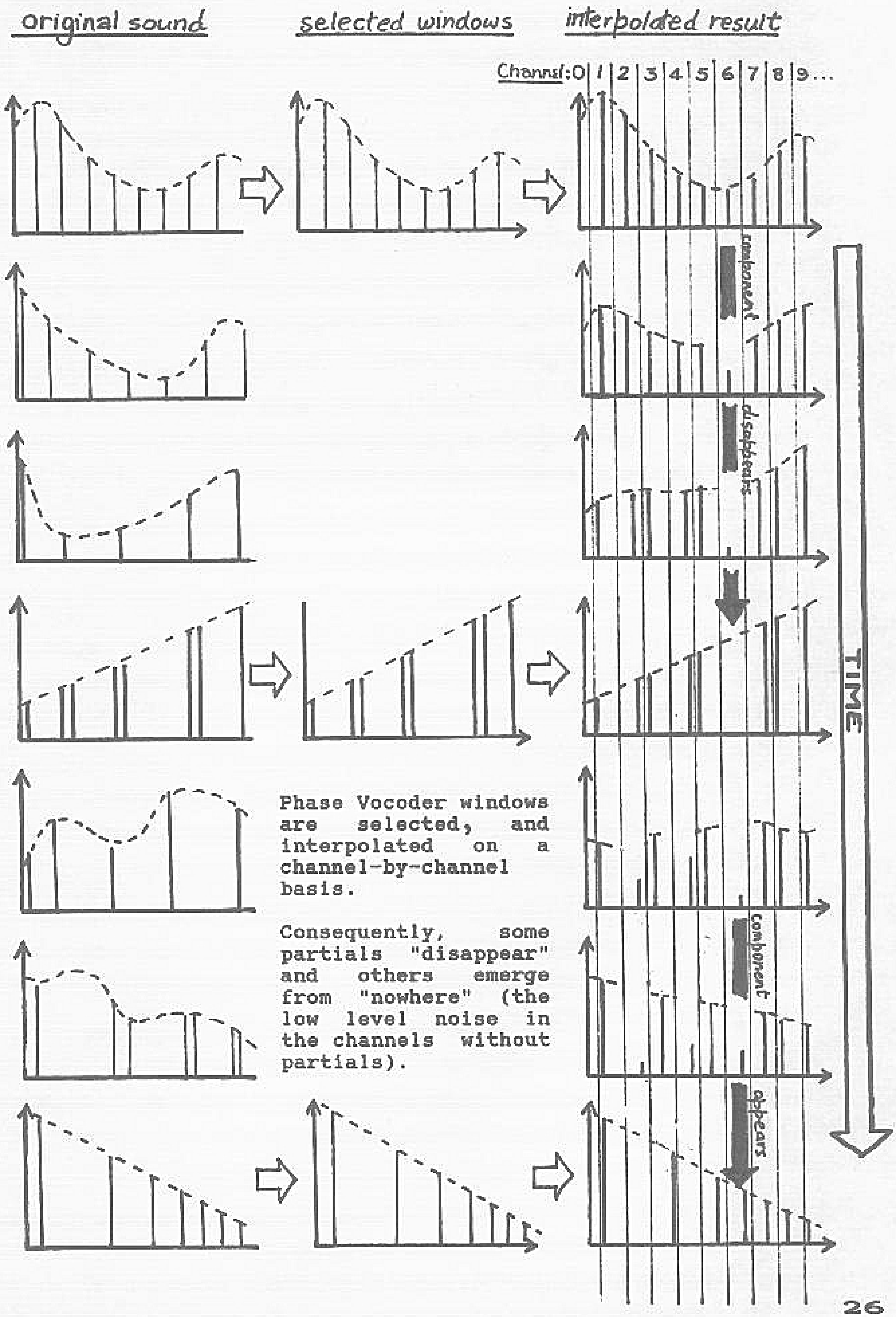
Spectral tracing for noise-reduction.



Spectral tracing creating revealed melodies as new partials enter the "most-prominent" set.

In the above examples we retain only the 8 most prominent data items. In a realistic situation, we would be working with many more channels (e.g. 512), retaining perhaps 128 of these for noise-reduction, and between 32 & 8 to create revealed melodies.

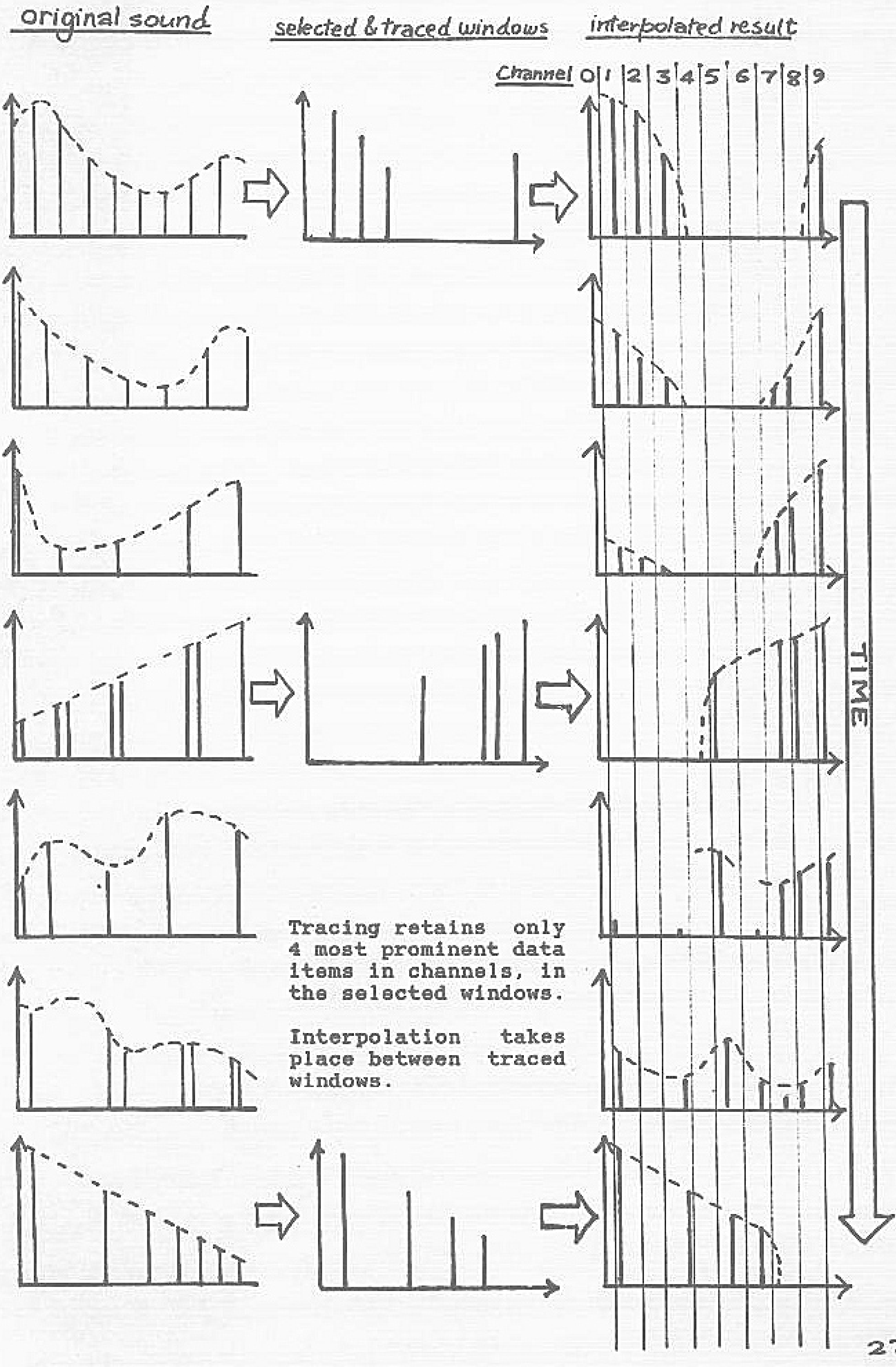
SPECTRAL BLURRING



Phase Vocoder windows are selected, and interpolated on a channel-by-channel basis.

Consequently, some partials "disappear" and others emerge from "nowhere" (the low level noise in the channels without partials).

SPECTRAL TRACE AND BLUR

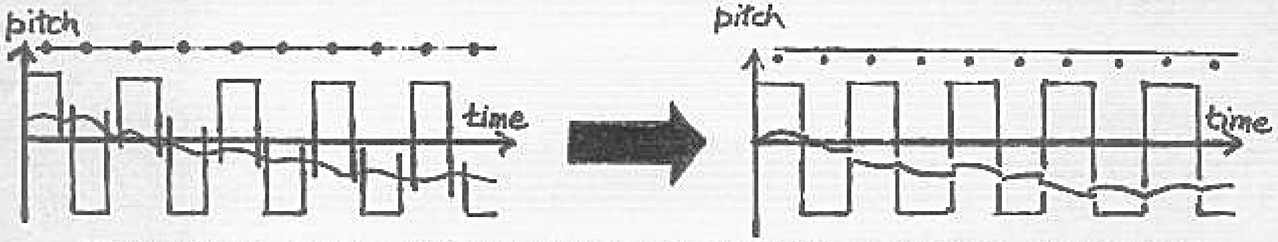


Tracing retains only 4 most prominent data items in channels, in the selected windows.

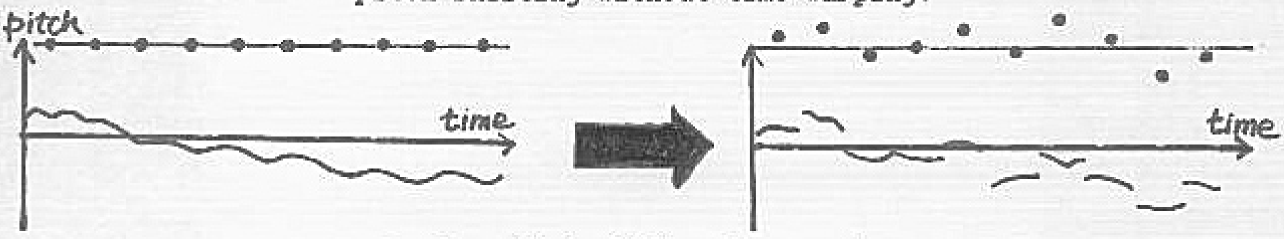
Interpolation takes place between traced windows.

BRASSAGE (continued)

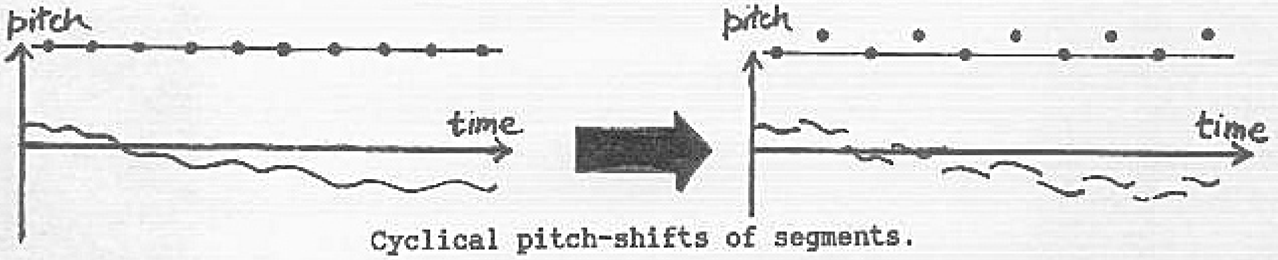
PITCH-SHIFTING OPTIONS



'Harmoniser' (only works with segments of grain dimension) or montage pitch-shifting without time-warping.

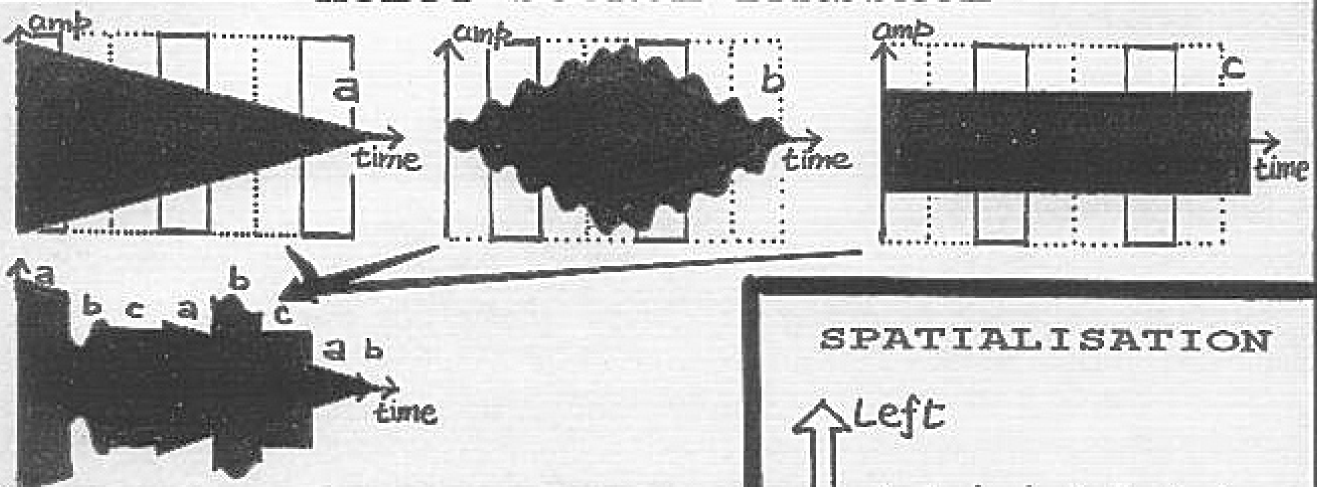


Random pitch-shifts of segments.

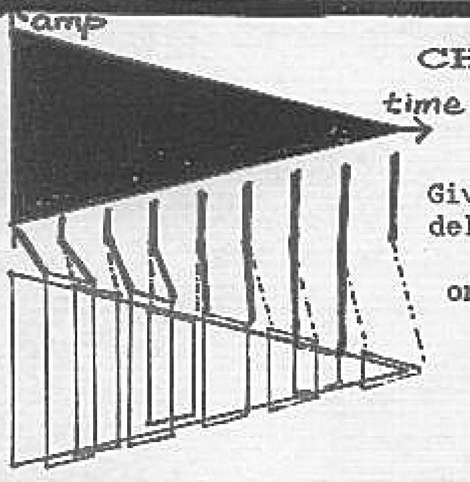


Cyclical pitch-shifts of segments.

MULTI-SOURCE BRASSAGE

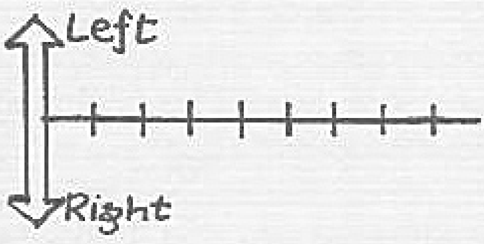


CHORUSING

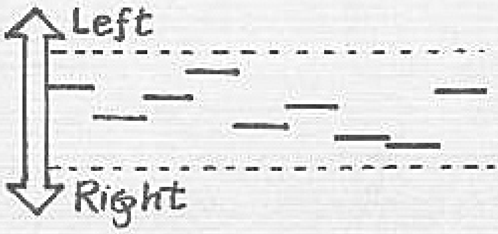


Give small random delays to segments and mix with original source.

SPATIALISATION

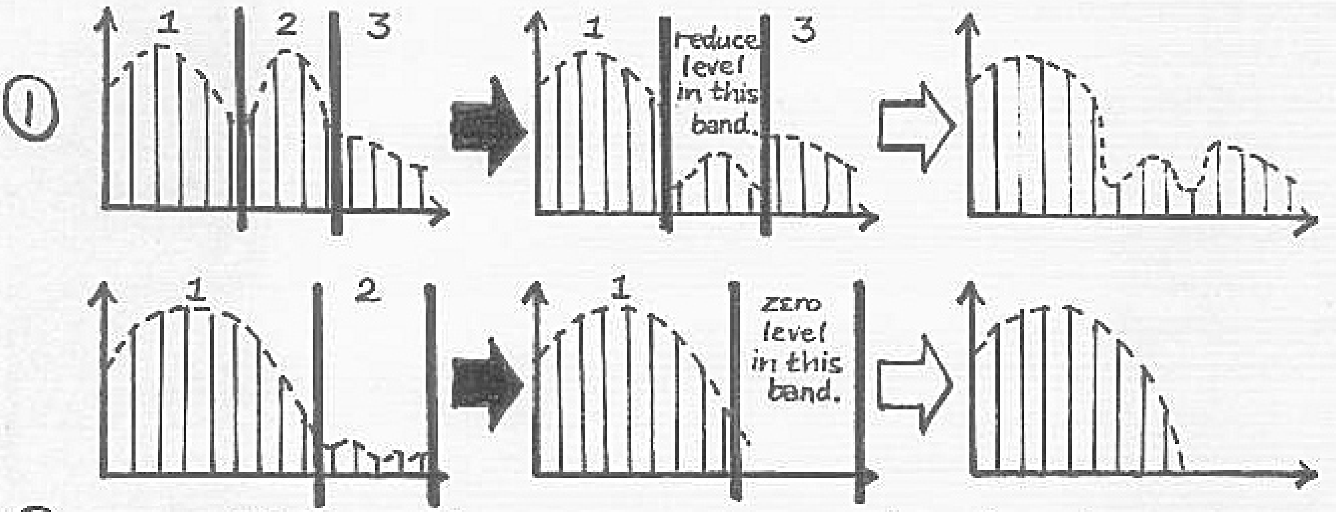


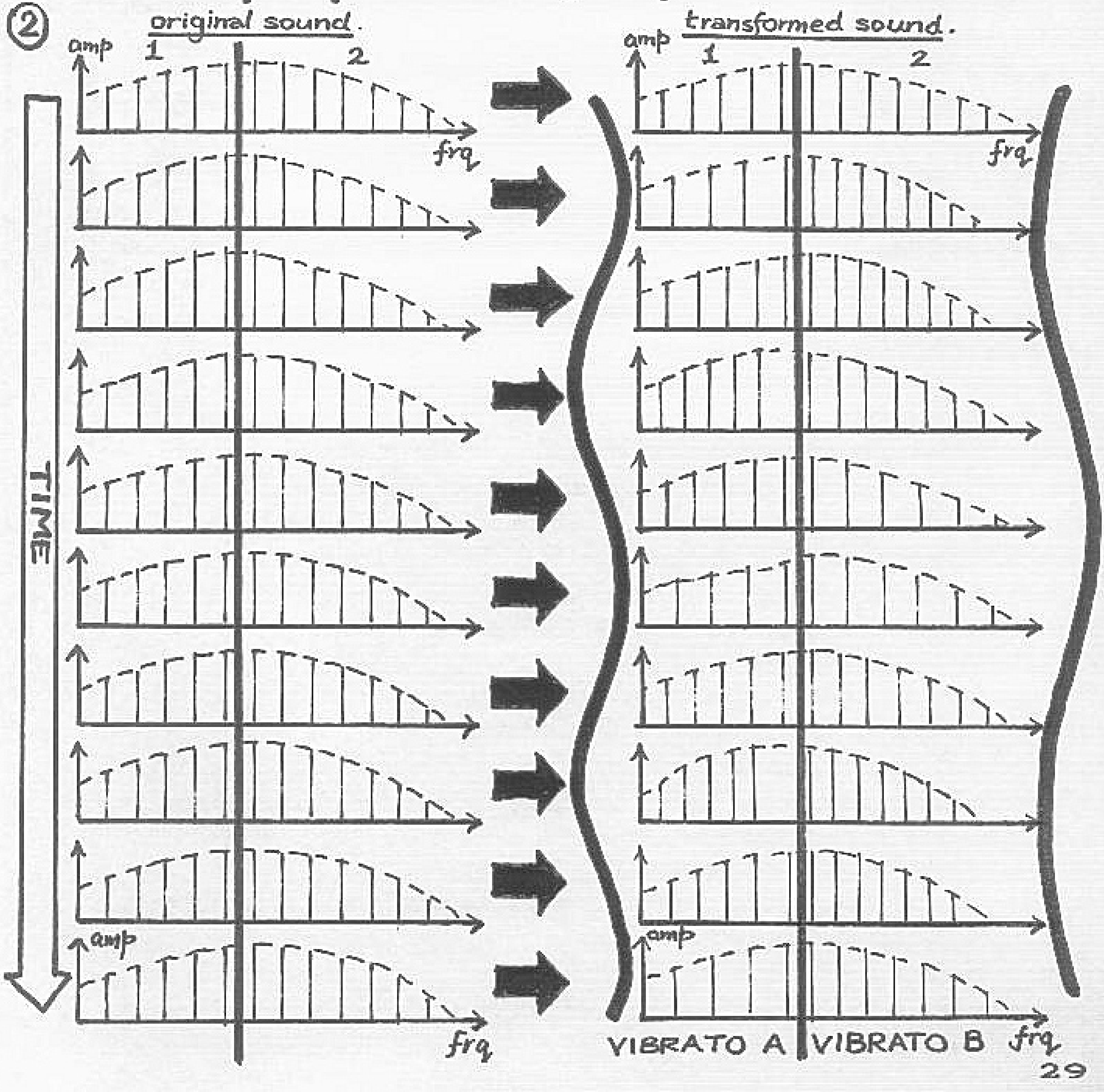
e.g. random spatialisation within a range.



SPECTRAL SPLITTING

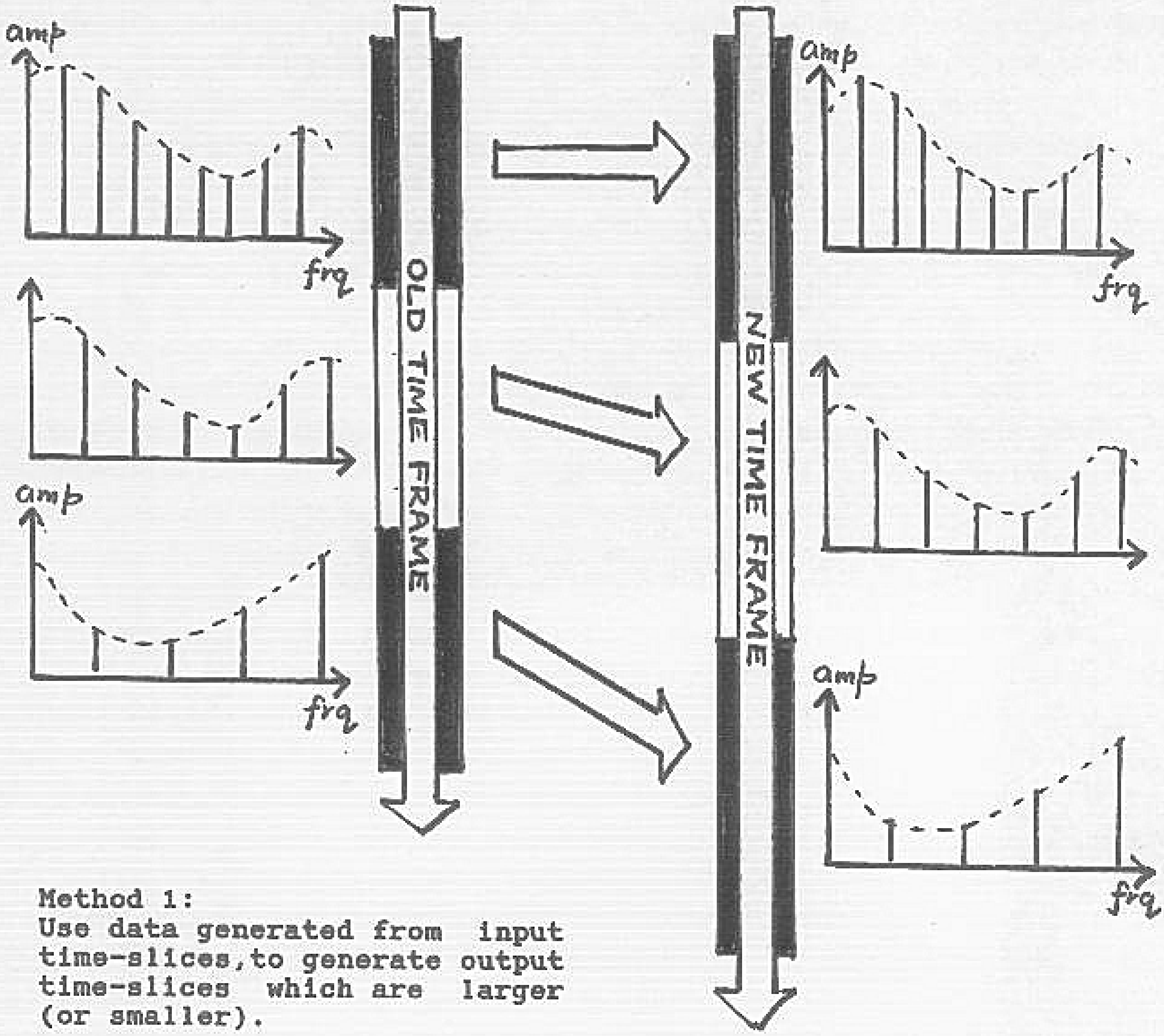
Divide spectrum into bands for e.g.
(1) Filtering, or noise removal.
(2) Splitting the aural image (here we add different vibrato to 2 bands).





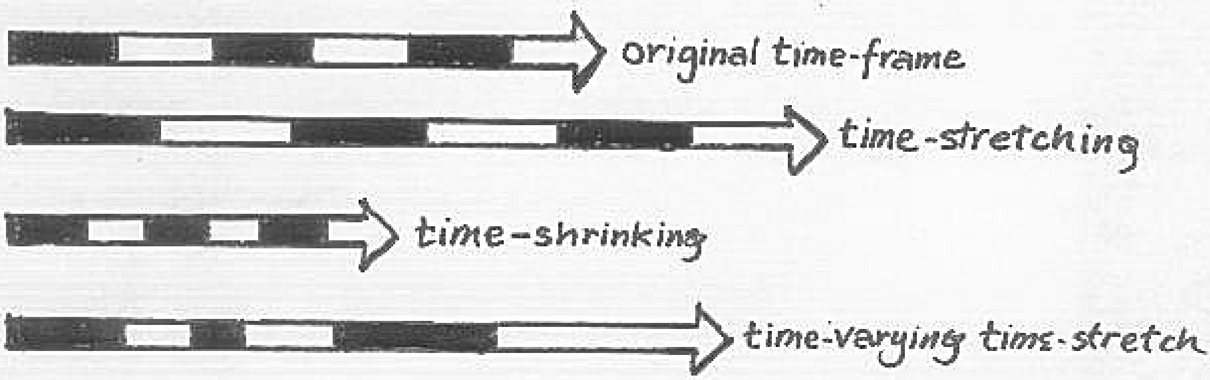
SPECTRAL TIME-STRETCHING

Method 1



Method 1:
Use data generated from input time-slices, to generate output time-slices which are larger (or smaller).

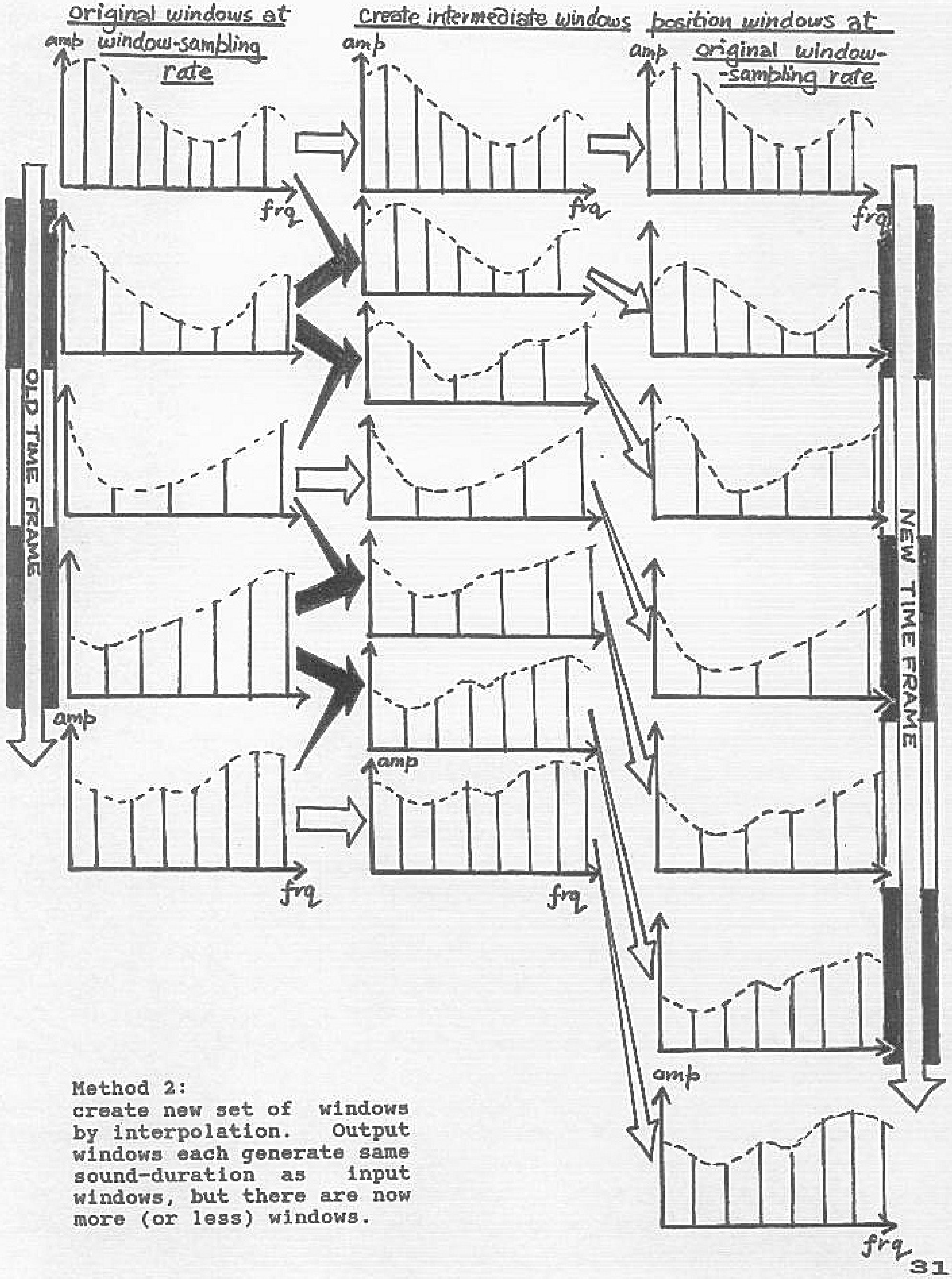
TIME-WARPING



Time-varying time-stretching (time-warping) is achieved ...
<u>Using method 1</u> : by varying durations of output time-slices generated from input windows.
<u>Using method 2</u> : by varying interpolation process which generates new windows.

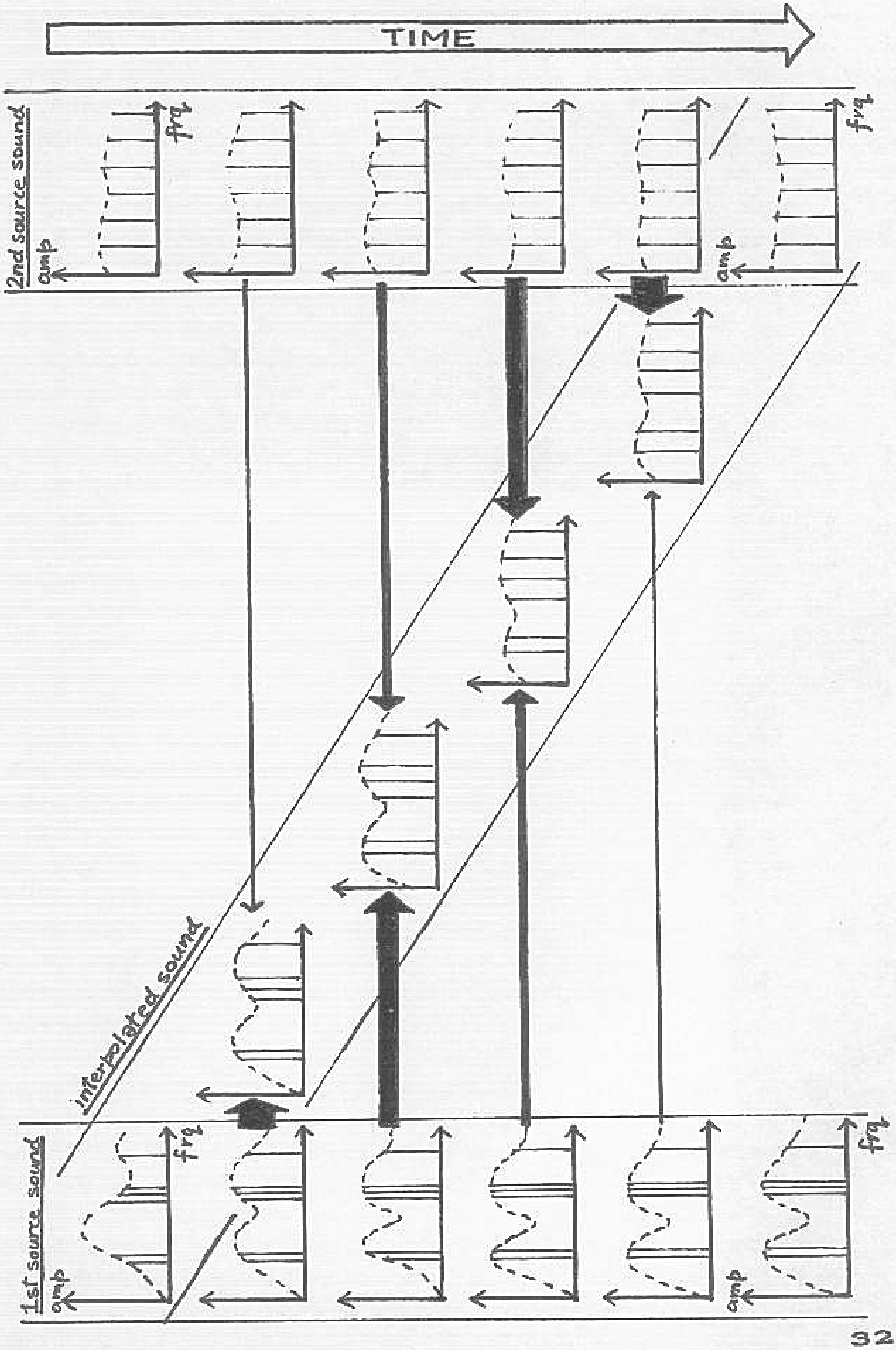
SPECTRAL TIME-STRETCHING

Method 2



Method 2:
create new set of windows by interpolation. Output windows each generate same sound-duration as input windows, but there are now more (or less) windows.

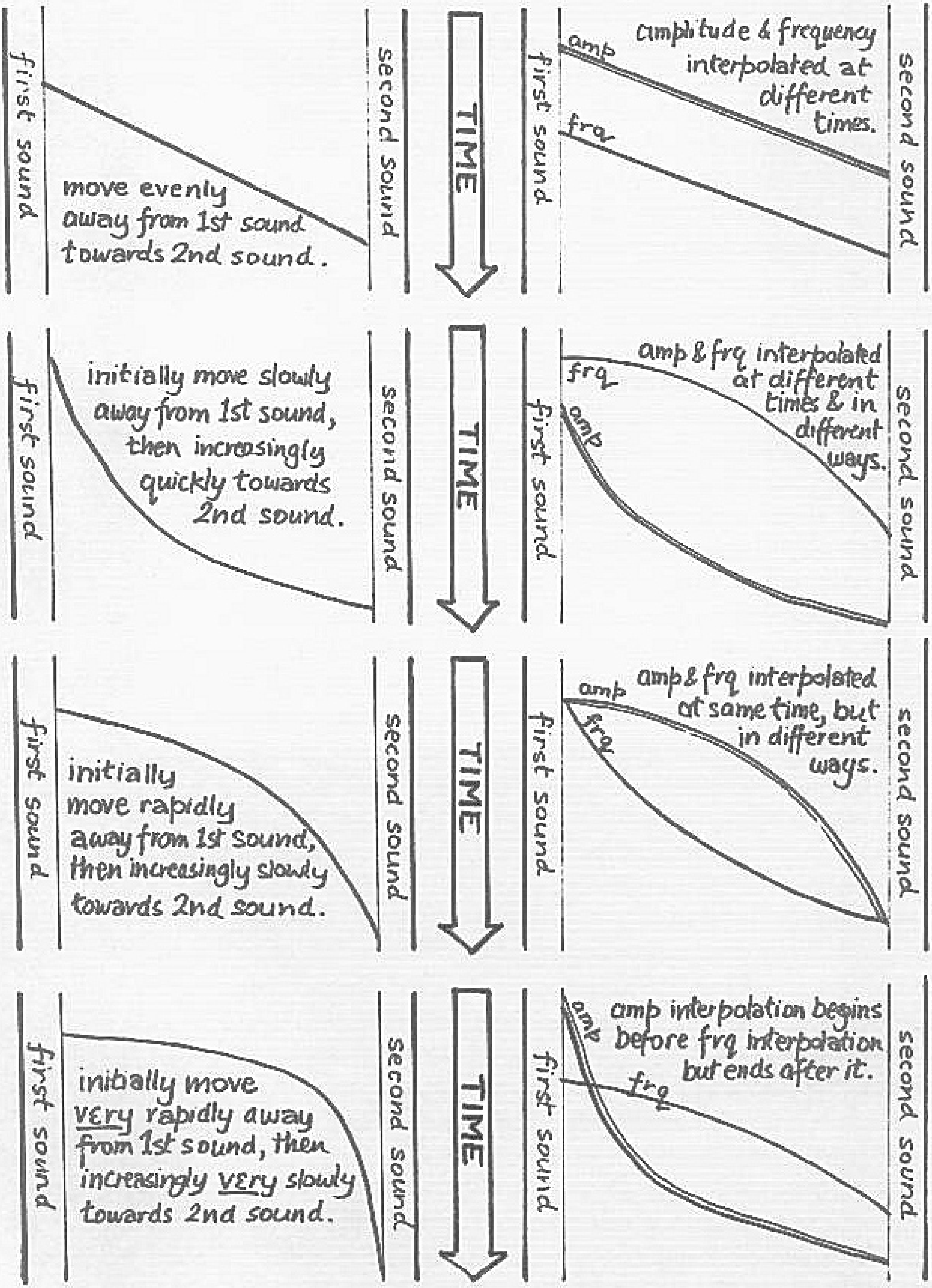
SPECTRAL INTERPOLATION



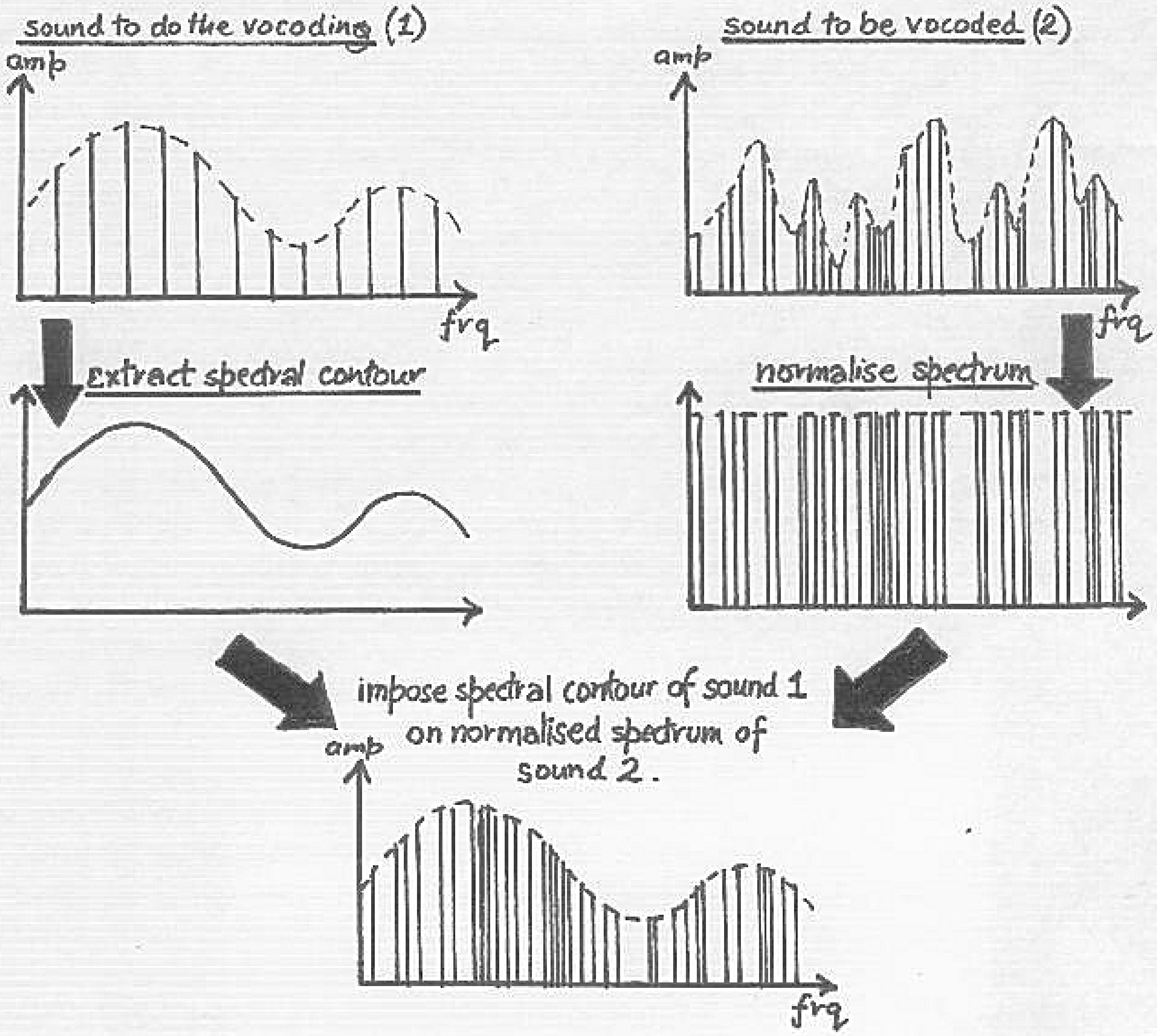
TYPES OF SPECTRAL INTERPOLATION

Different types of interpolation.

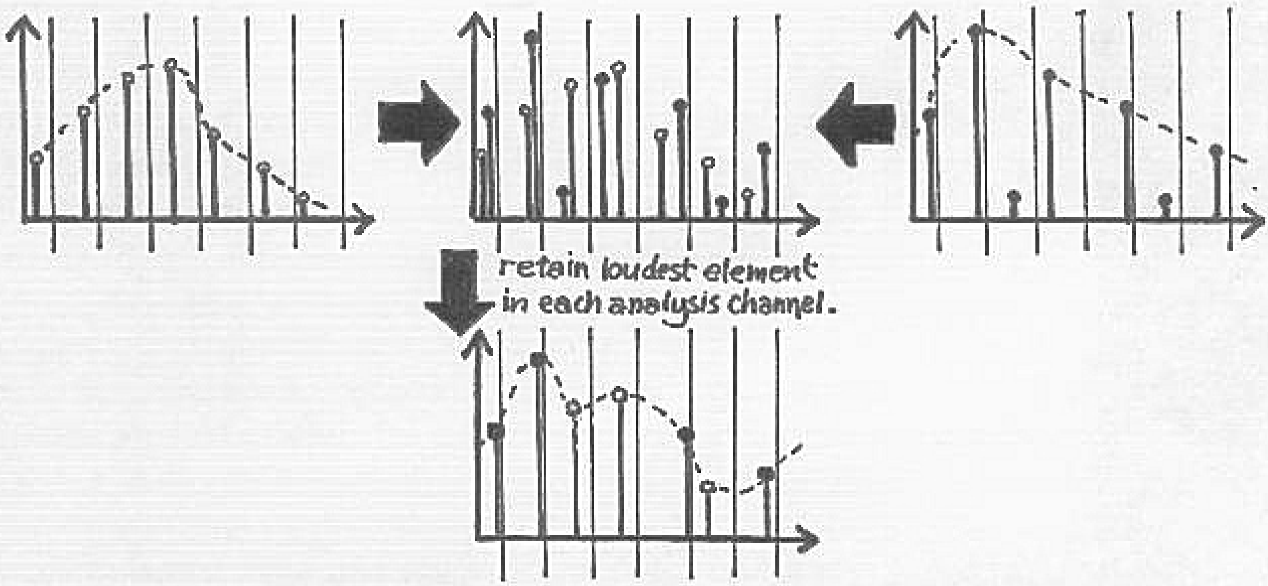
Independent interpolation of amplitude and frequency.



VOCODING



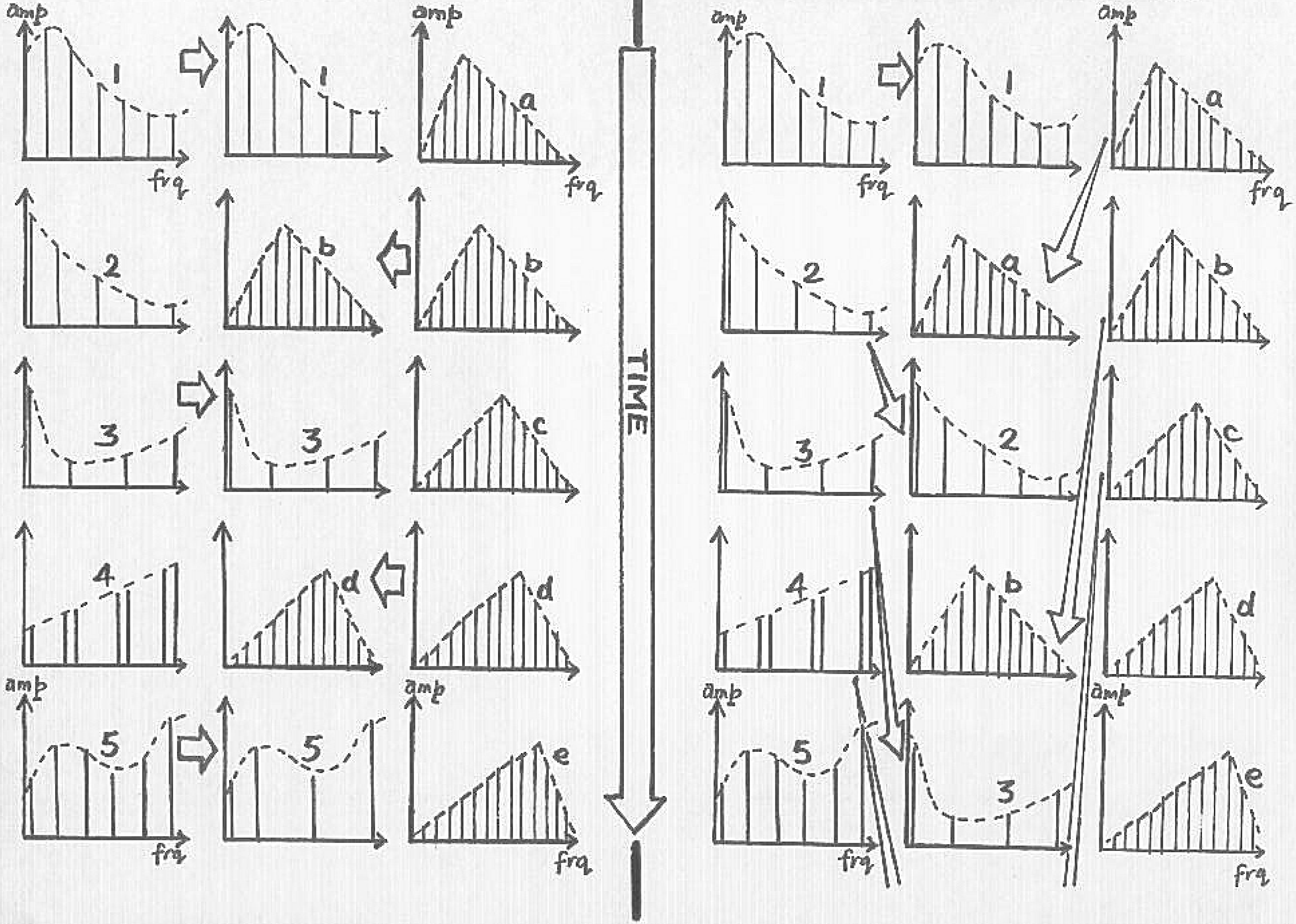
SPECTRAL MASKING



SPECTRAL INTERLEAVING

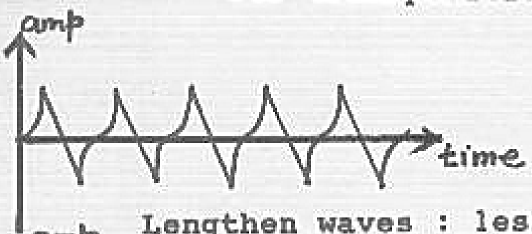
(1) In original time-frame.

(2) With time-frame extended X 2.

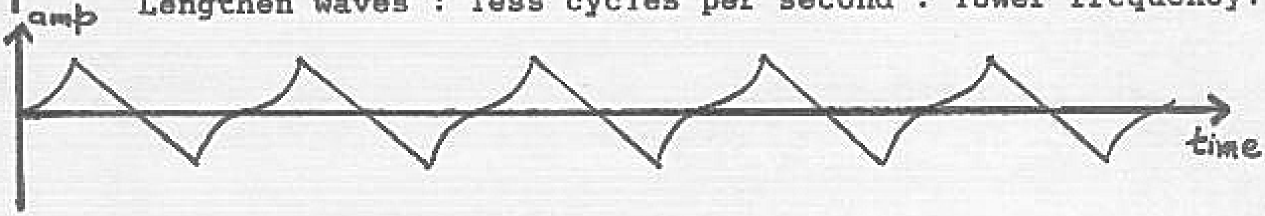


TAPE-SPEED VARIATION

For transposition and time-stretching.

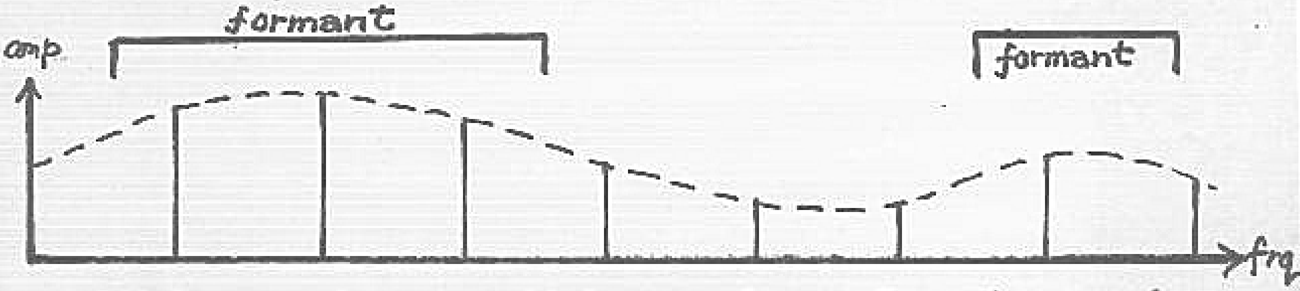


Lengthen waves : less cycles per second : lower frequency.

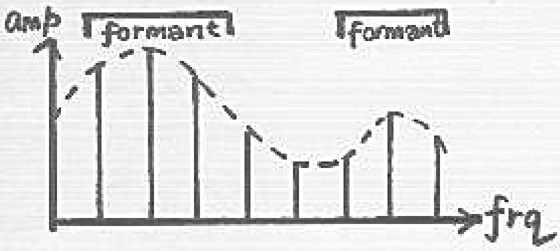


but also increases duration...

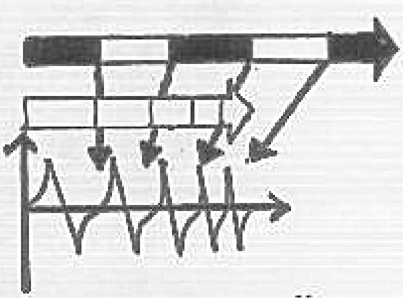
Original spectrum and spectral contour.



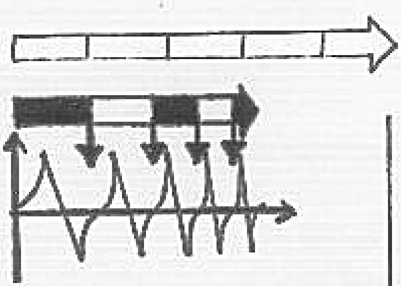
Pitch moves down : partials move down : formants move down.



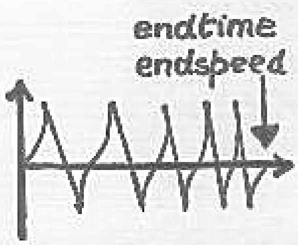
TAPE ACCELERATION



Specify goal speeds relative to source time.

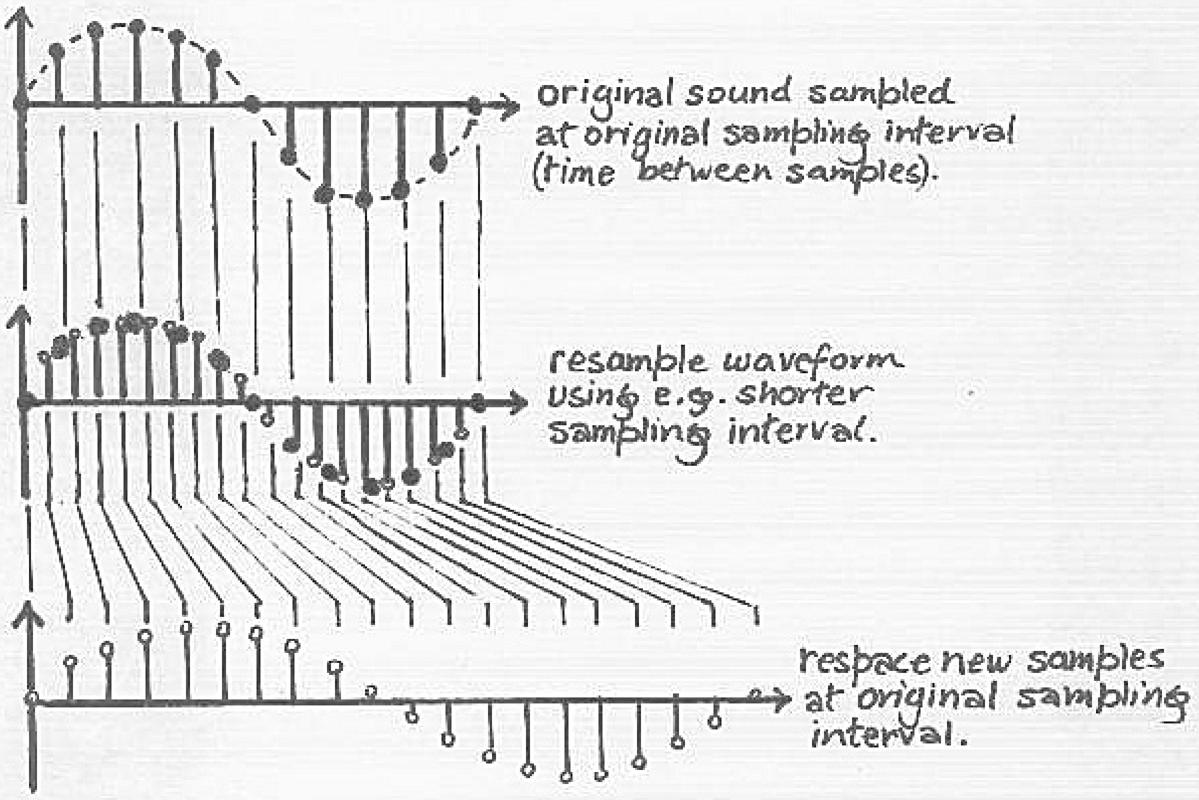


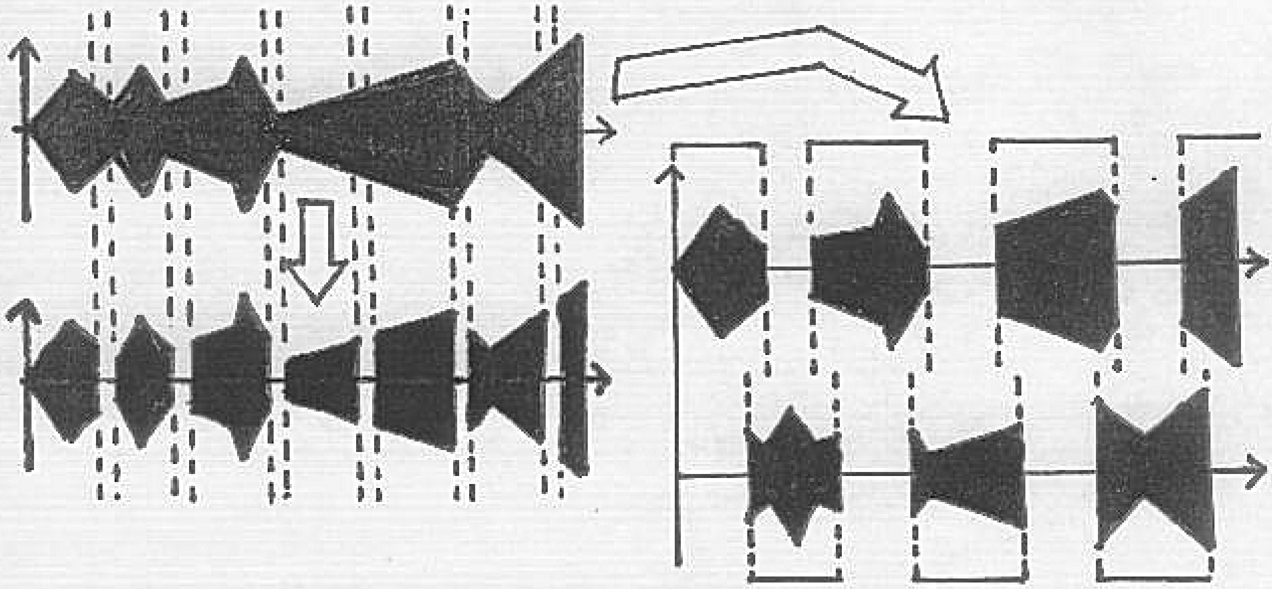
Specify goal speeds relative to goal time.



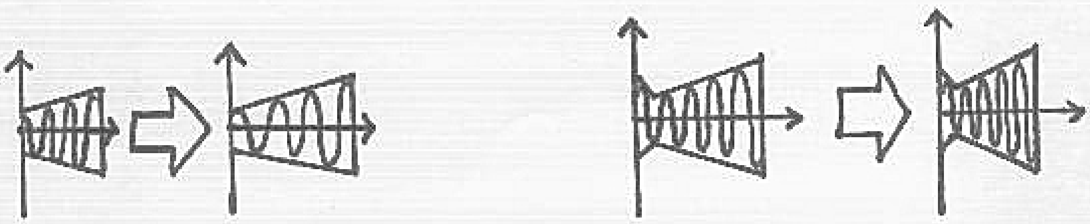
Specify end time & end speed relative to start

"TAPE-SPEED" VARIATION, DIGITALLY

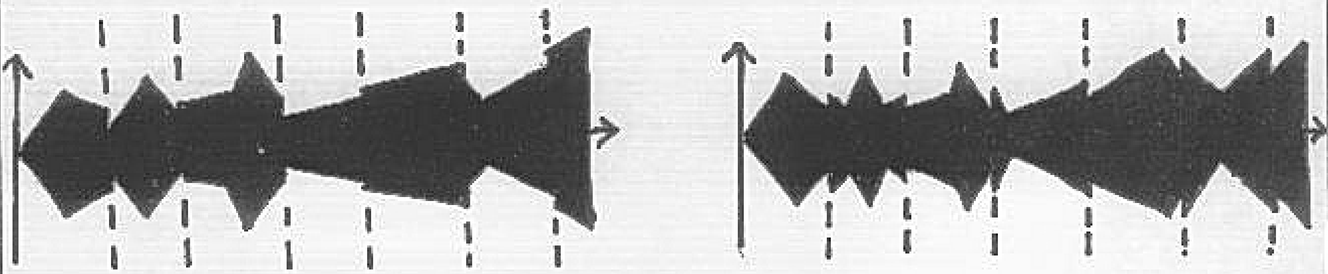




Cut sound into disjoint, or overlapping segments.
Segment length chosen so that, after they have been expanded or contracted (see below), they will rejoin into the exact duration of the original sound.

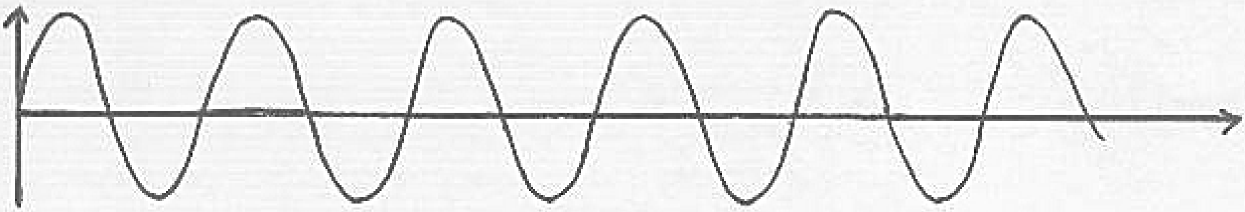


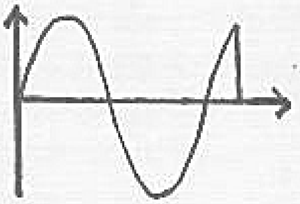
Change segment pitch by tape-speed variation, which also modifies segment lengths.



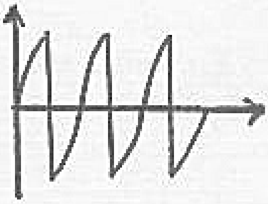
Rejoin changed-length segments, with no overlaps or gaps. They now fit exactly into the original duration.

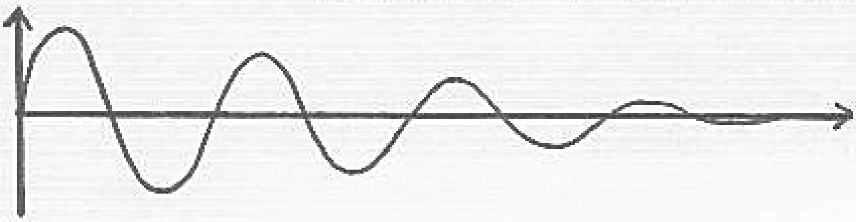
CUTTING





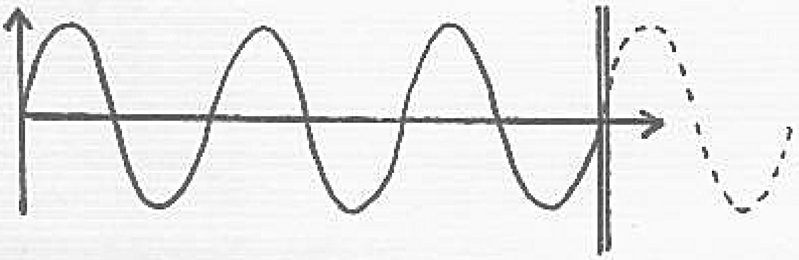
We cannot cut a sound simply by stopping it at an arbitrary point. Usually this produces a sharp edge (see left), equivalent to a fragment of a new waveform (see right) with many high partials.
In practice, we hear a click.





A normal cut is a brief fade-out to zero loudness.

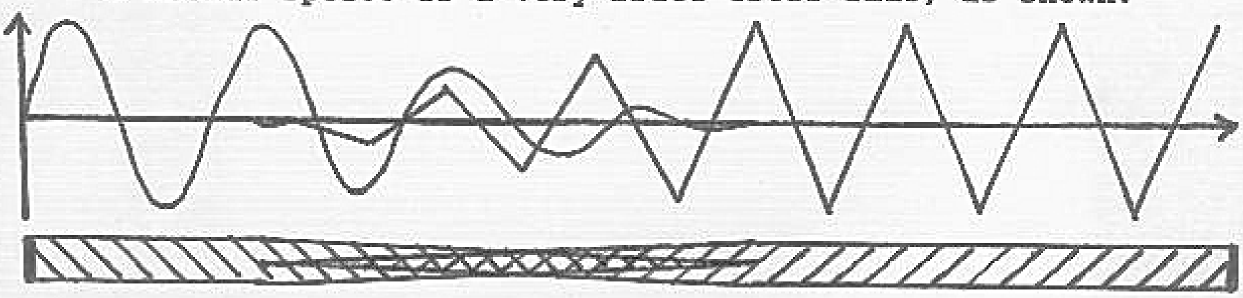
ZERO CUTTING



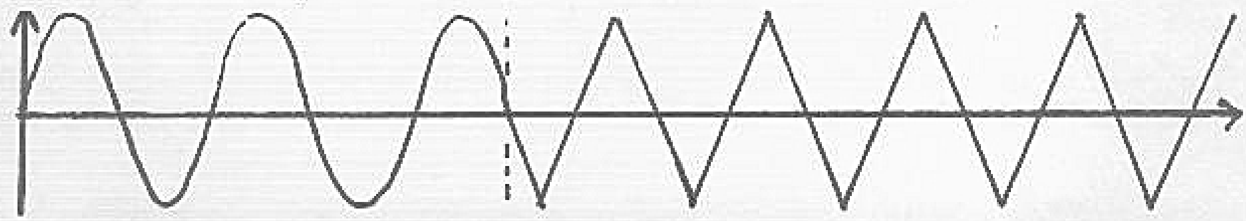
Unconventionally, we may cut a sound at a zero-crossing.

SPLICING

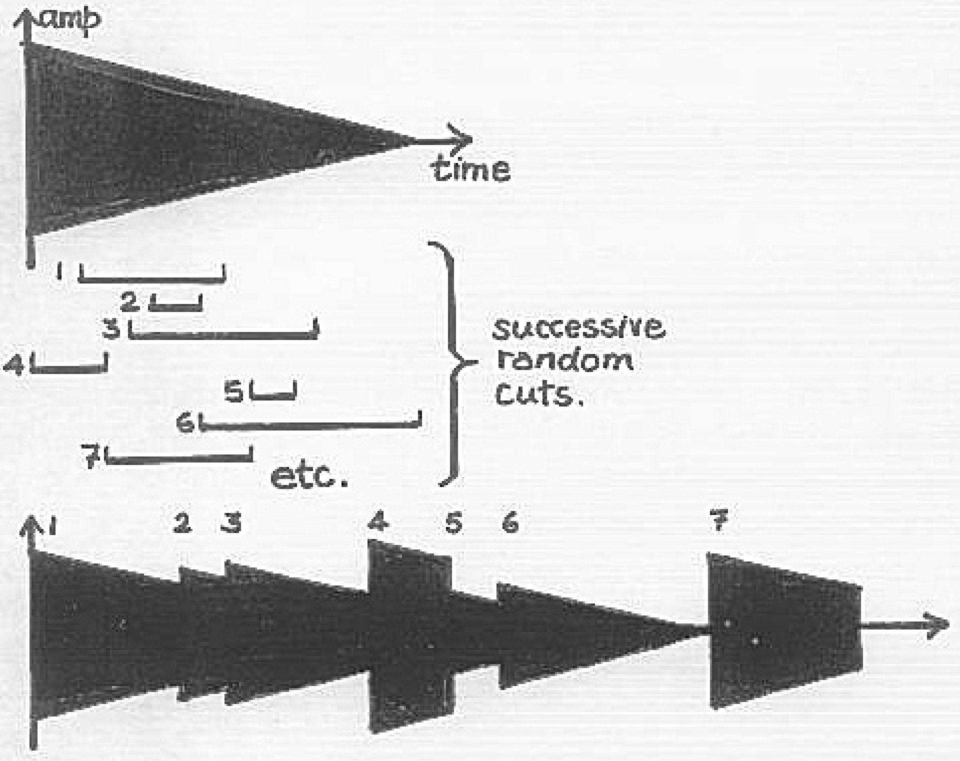
A normal splice is a very brief cross-fade, as shown.



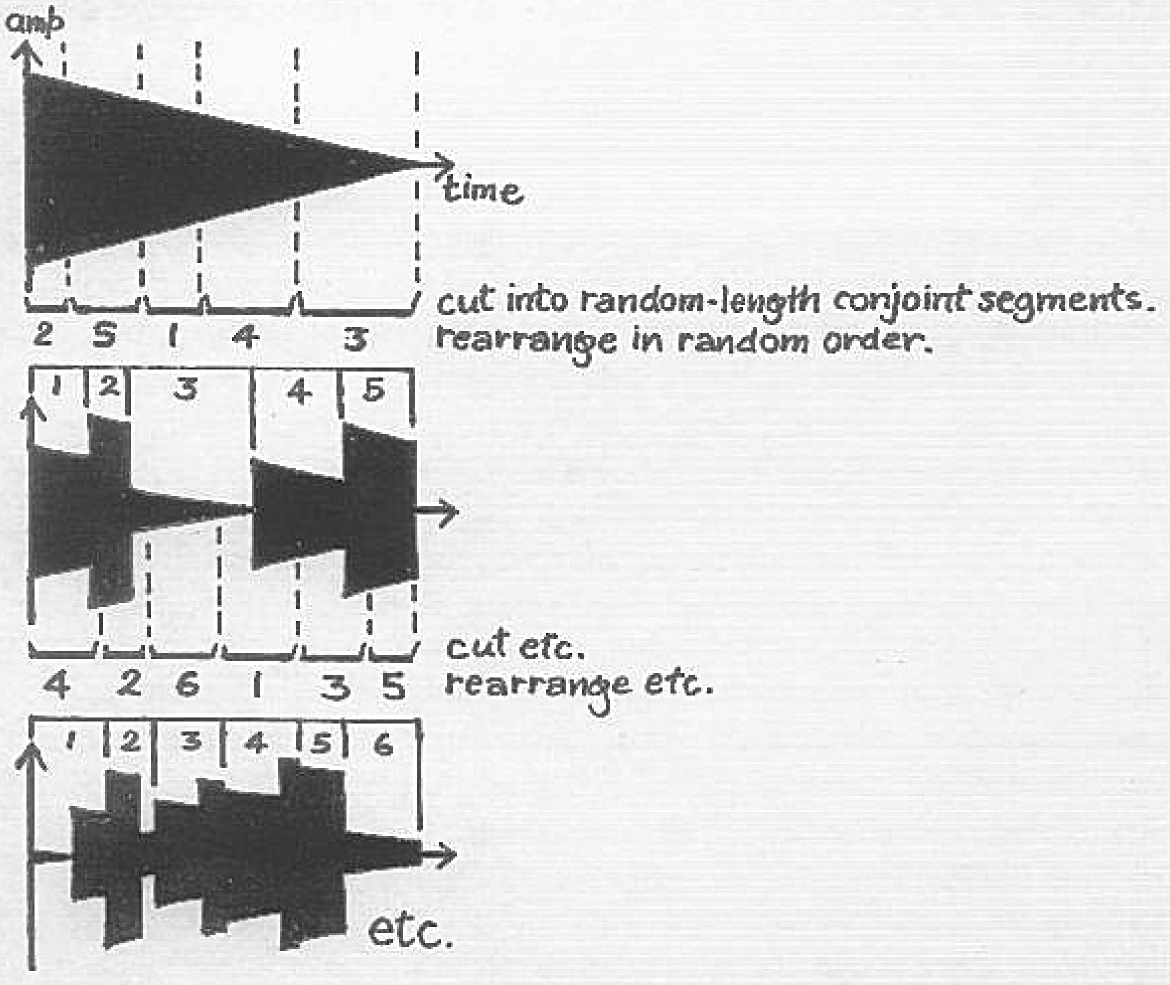
Unconventionally, we may splice at a zero-crossing in both sounds.



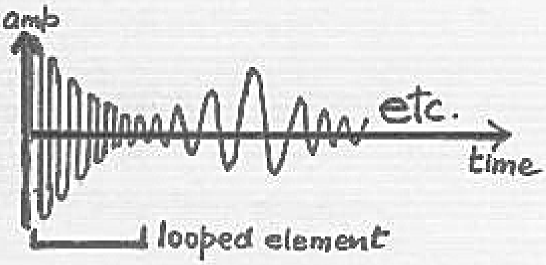
RANDOM CUTTING

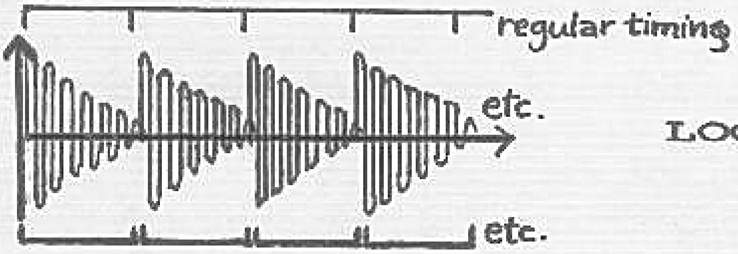


SHREDDING

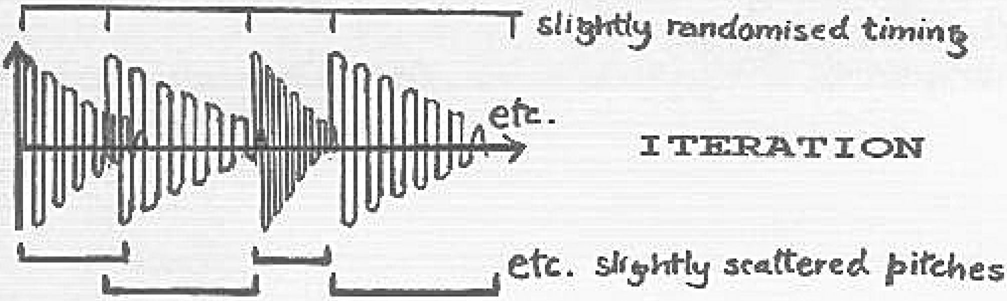


LOOPING AND ITERATION



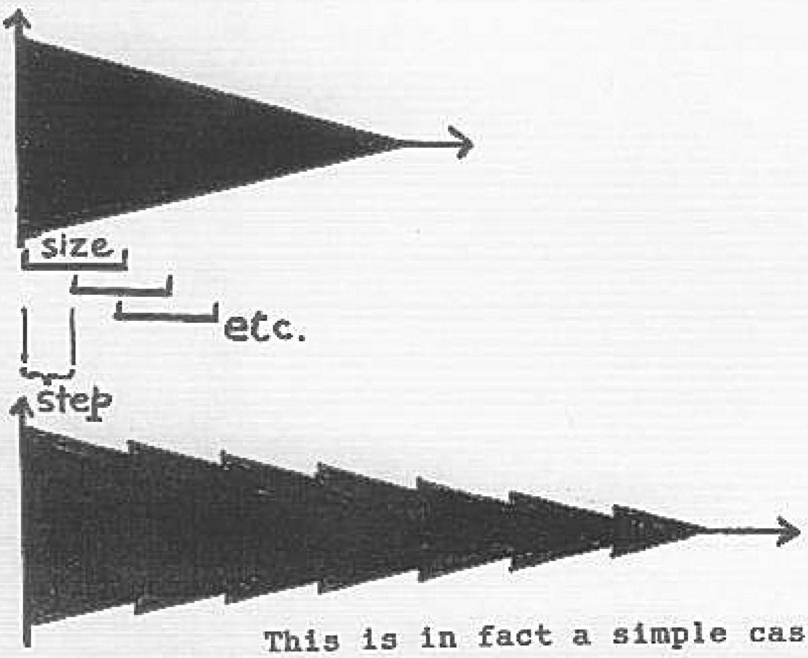


LOOPING



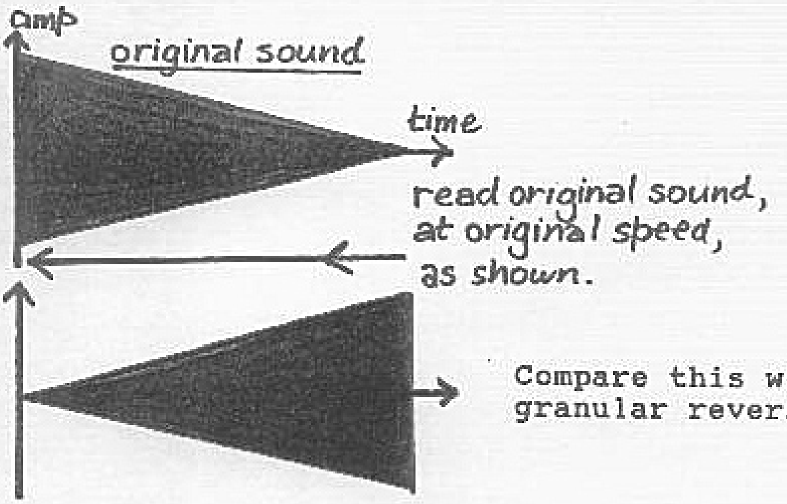
ITERATION

PROGRESSIVE LOOPING



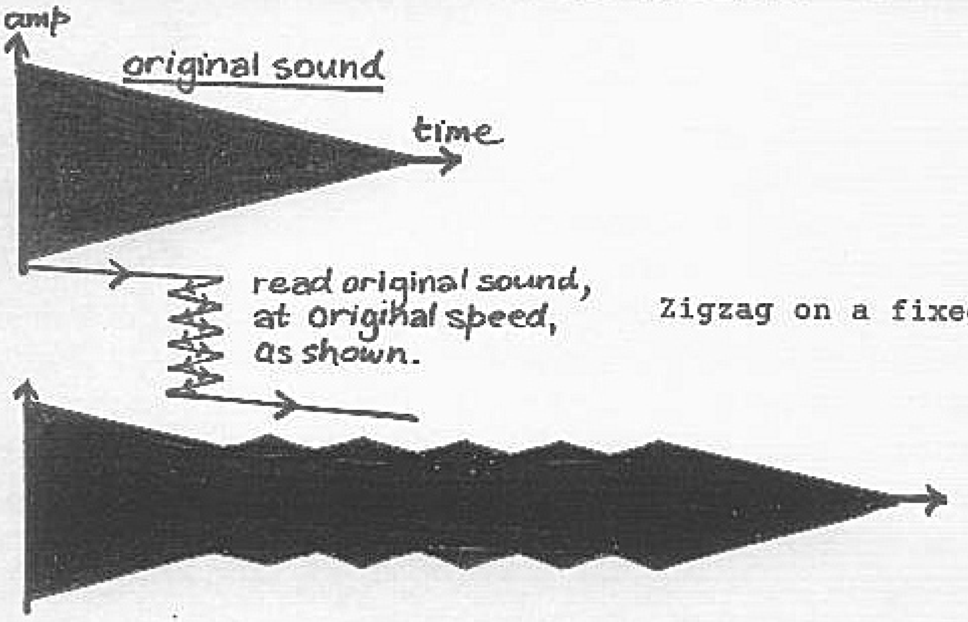
This is in fact a simple case of Brassage.

SOUND REVERSING

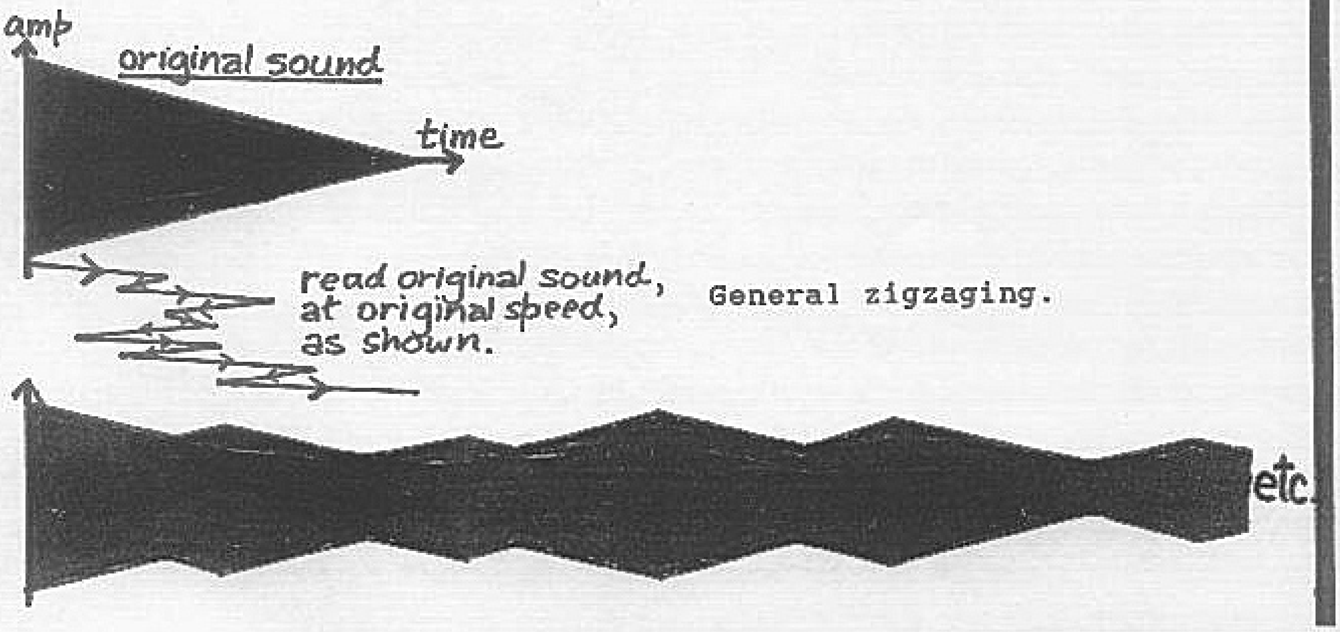


Compare this with granular reversal.

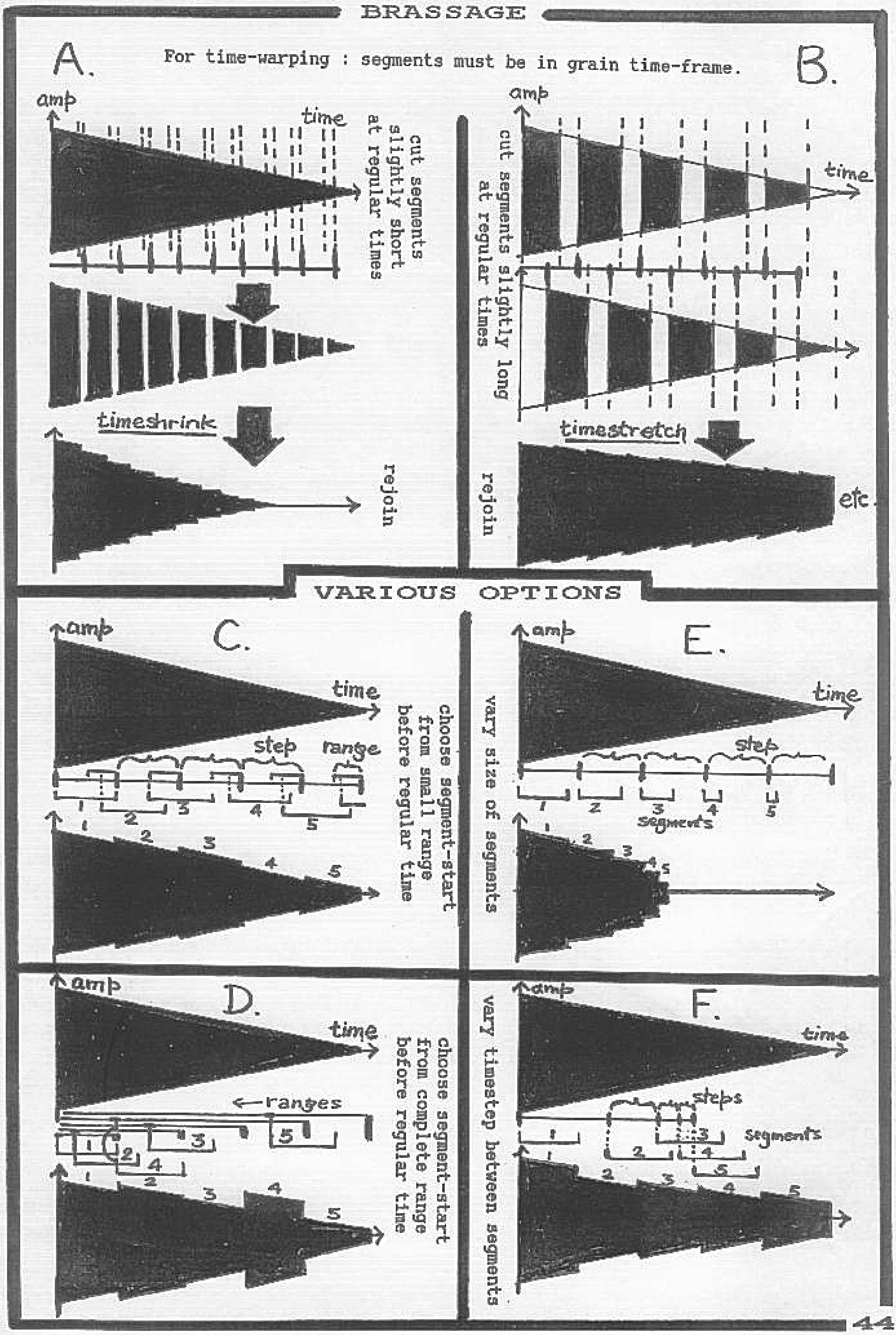
ZIGZAGING



Zigzag on a fixed element.



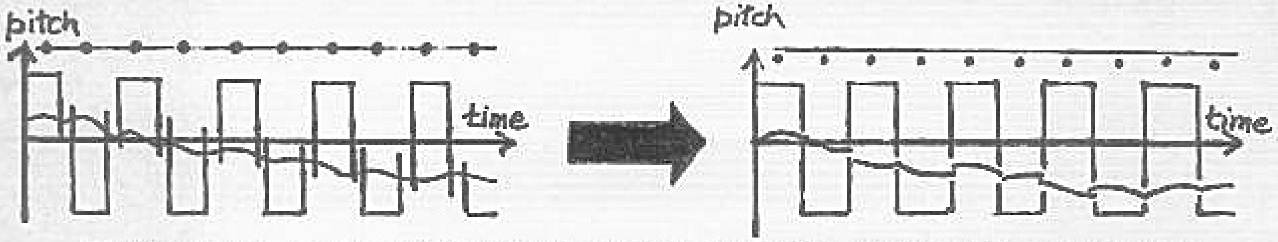
General zigzaging.



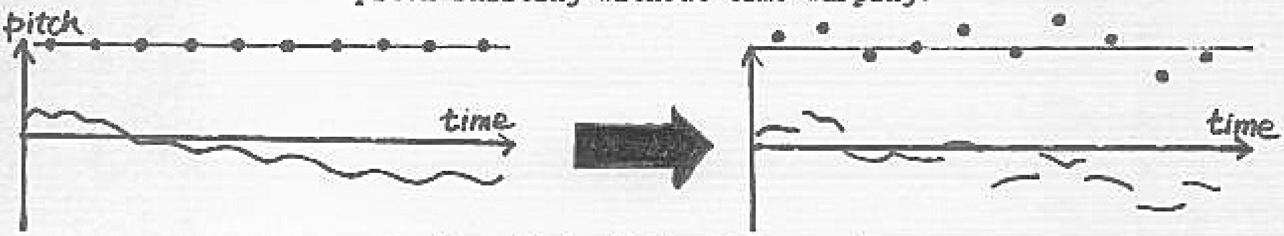
BRASSAGE
A.
For time-warping : segments must be in grain time-frame.
B.
amp
time
cut segments slightly short at regular times
timeshrink
rejoin
cut segments slightly long at regular times
timestretch
etc.
VARIOUS OPTIONS
C.
choose segment-start from small range before regular time
step
range
1
2
3
4
5
E.
vary size of segments
segments
D.
choose segment-start from complete range before regular time
ranges
F.
vary timestep between segments
steps
segments

BRASSAGE (continued)

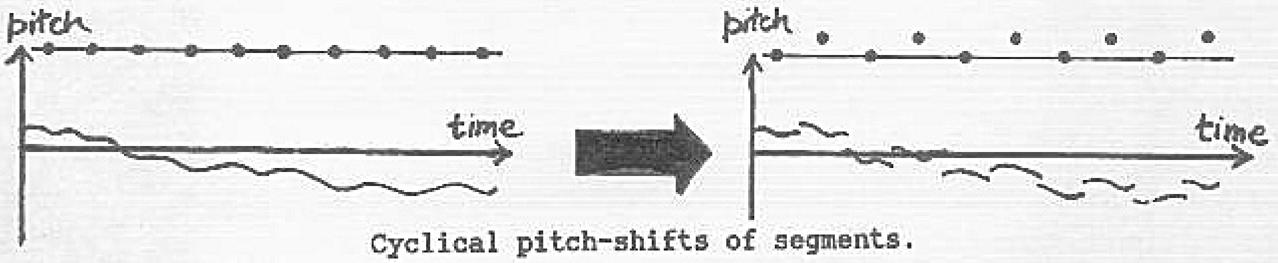
PITCH-SHIFTING OPTIONS



'Harmoniser' (only works with segments of grain dimension) or montage pitch-shifting without time-warping.

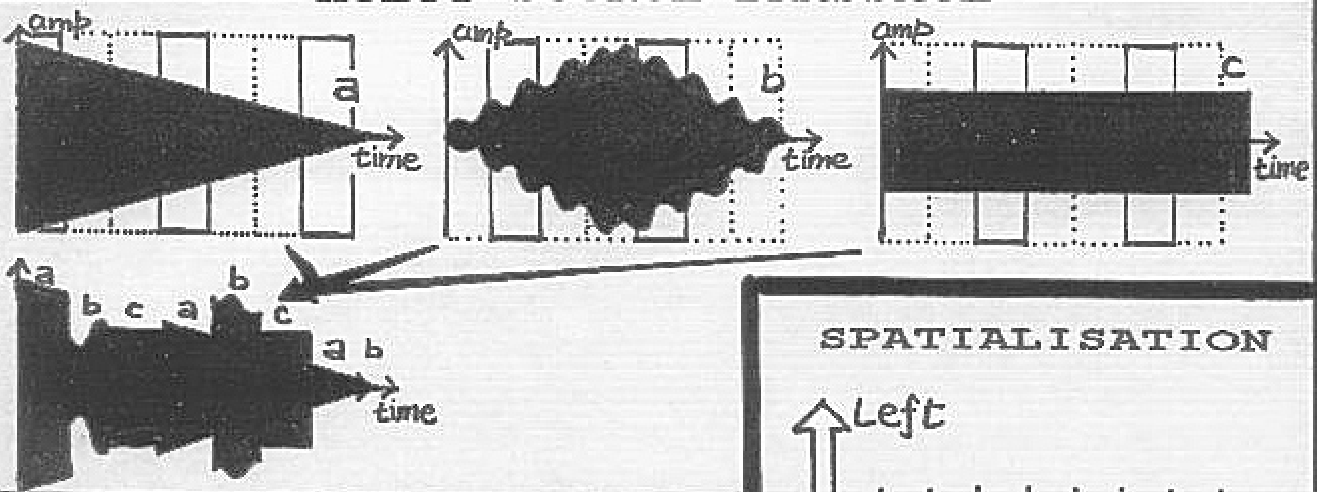


Random pitch-shifts of segments.

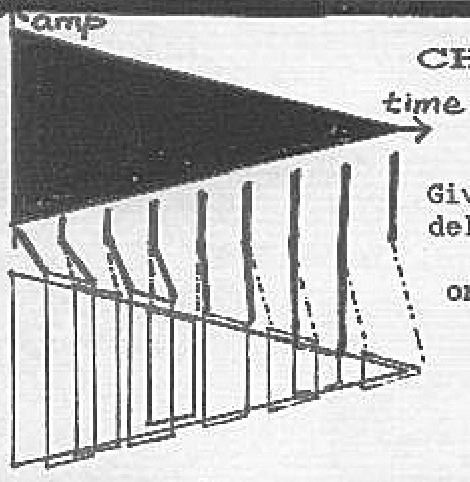


Cyclical pitch-shifts of segments.

MULTI-SOURCE BRASSAGE

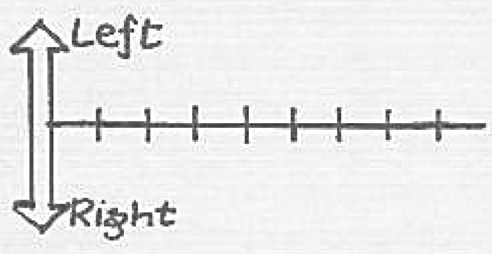


CHORUSING

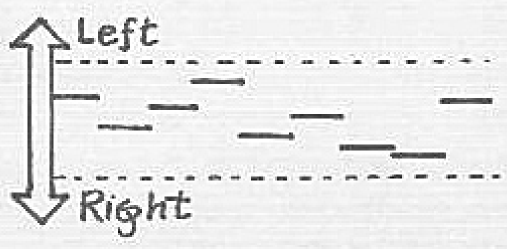


Give small random delays to segments and mix with original source.

SPATIALISATION



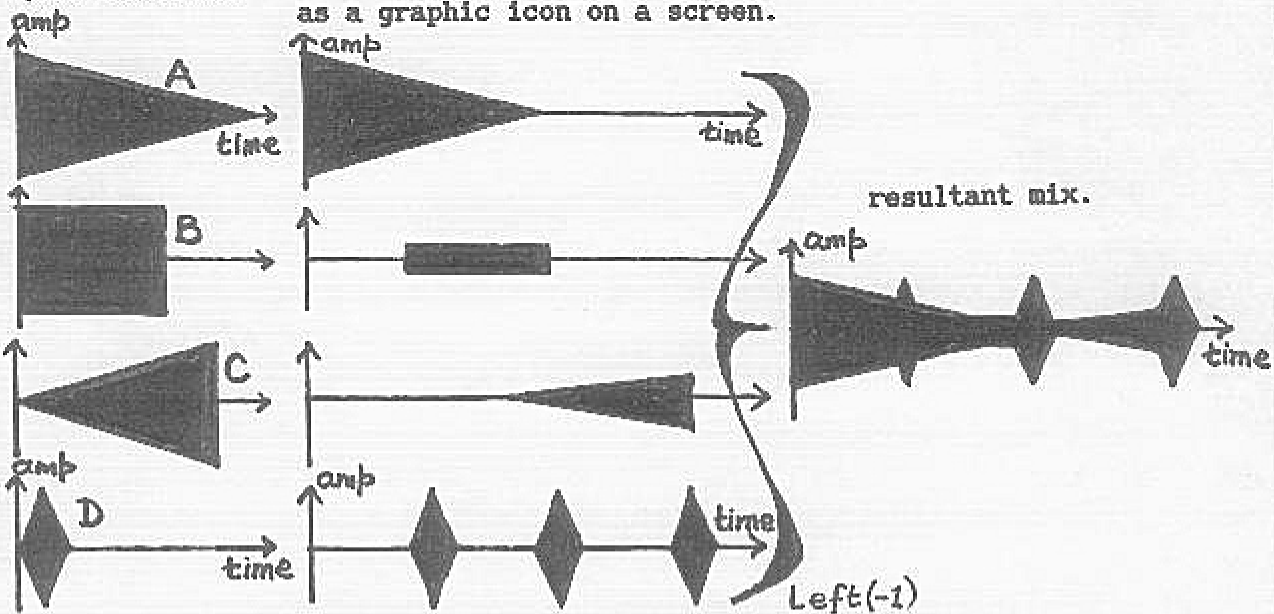
e.g. random spatialisation within a range.



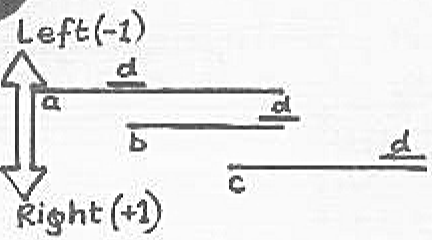
MIXING

Source sounds (mix elements)

each mix element assigned a start time and a loudness as a graphic icon on a screen.



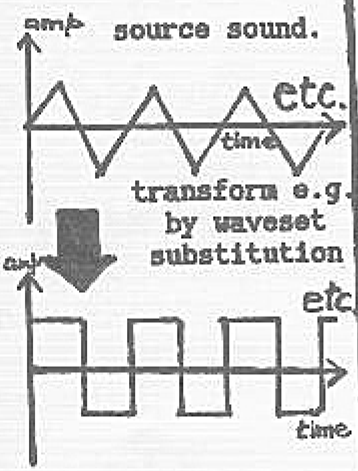
name	time (secs)	loudness	position
A	0·0	1·00	−1·0
B	0·5	0·33	0·0
D	0·625	1·0	−1·0
C	1·00	0·5	1·0
D	1·25	1·0	0·0
D	2·00	1·0	1·0

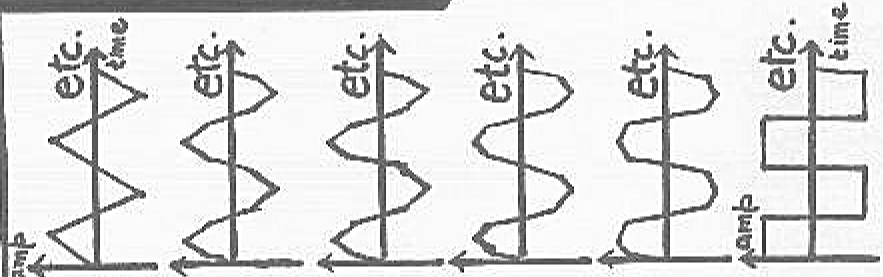


spatialisation of mix.

each mix element assigned a start time and a loudness on a line in a mixing score.

IN-BETWEENING

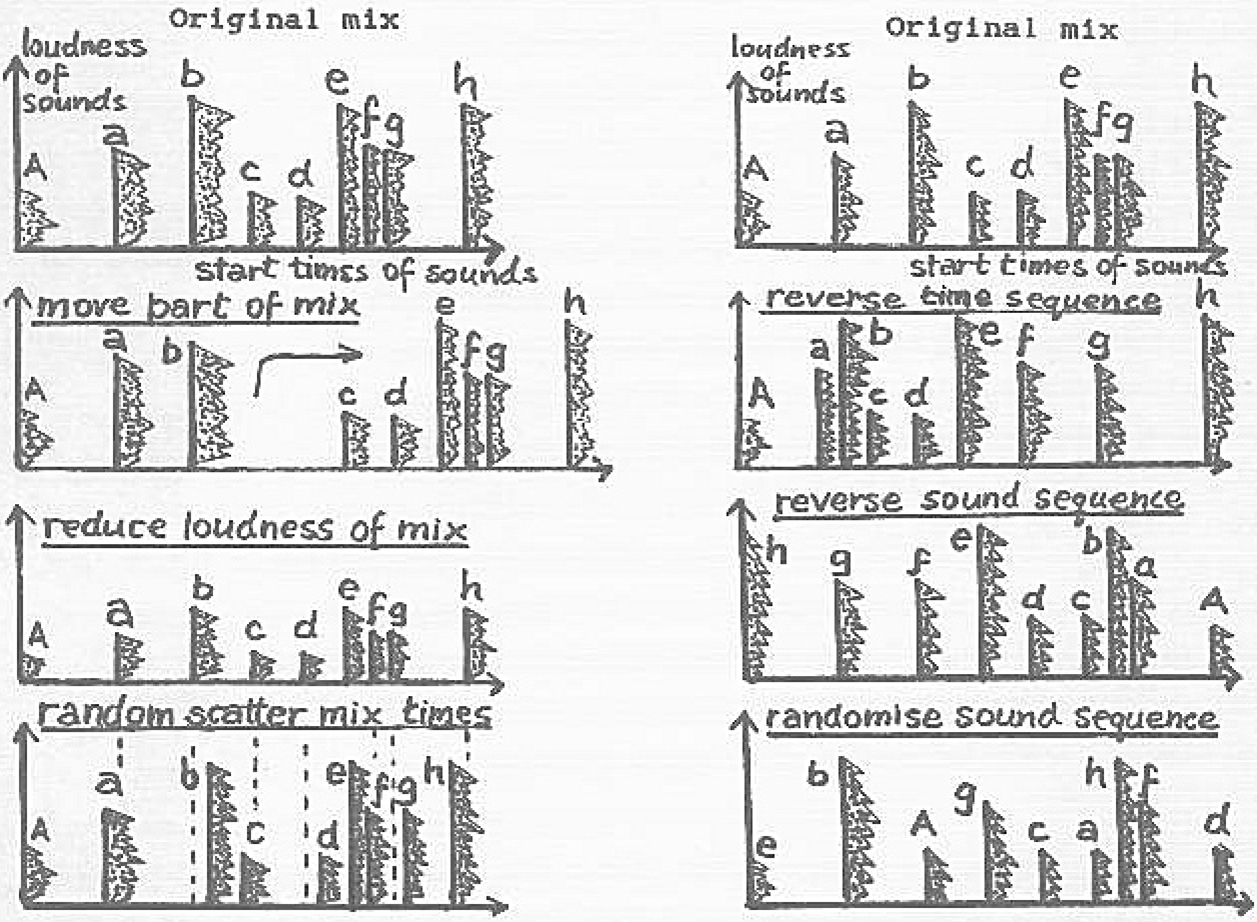




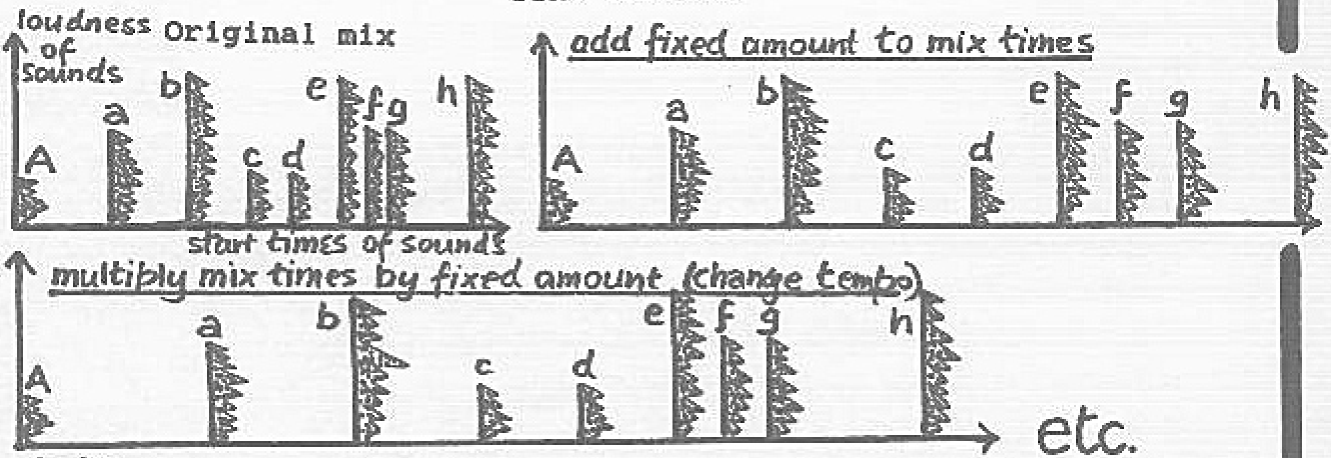
name	time	loudness
A	0·0	1·00
B	0·0	0·00
A	0·0	0·75
A	0·0	0·25
A	0·0	0·6
B	0·0	0·4
A	0·0	0·4
B	0·0	0·6
A	0·0	0·25
B	0·0	0·75
A	0·0	0·0
B	0·0	1·0

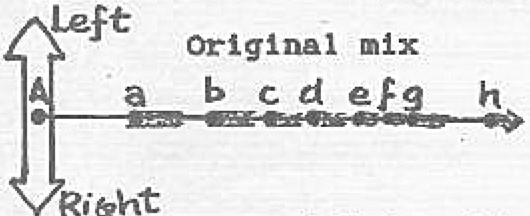
create intermediate types, by appropriate mixing.

MIX SHUFFLING

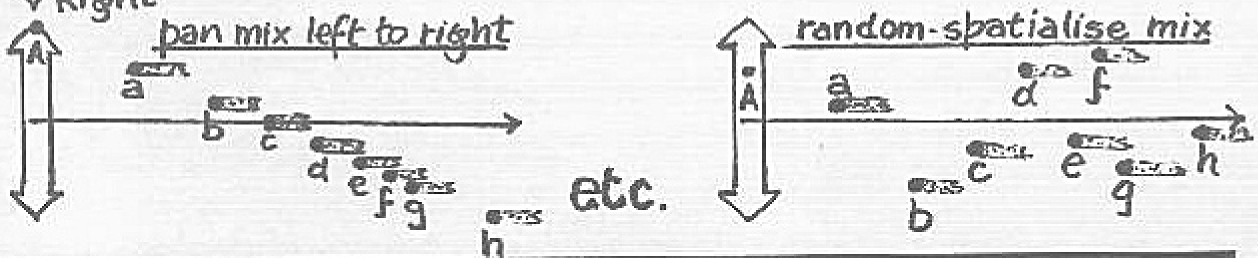


TIME WARPING



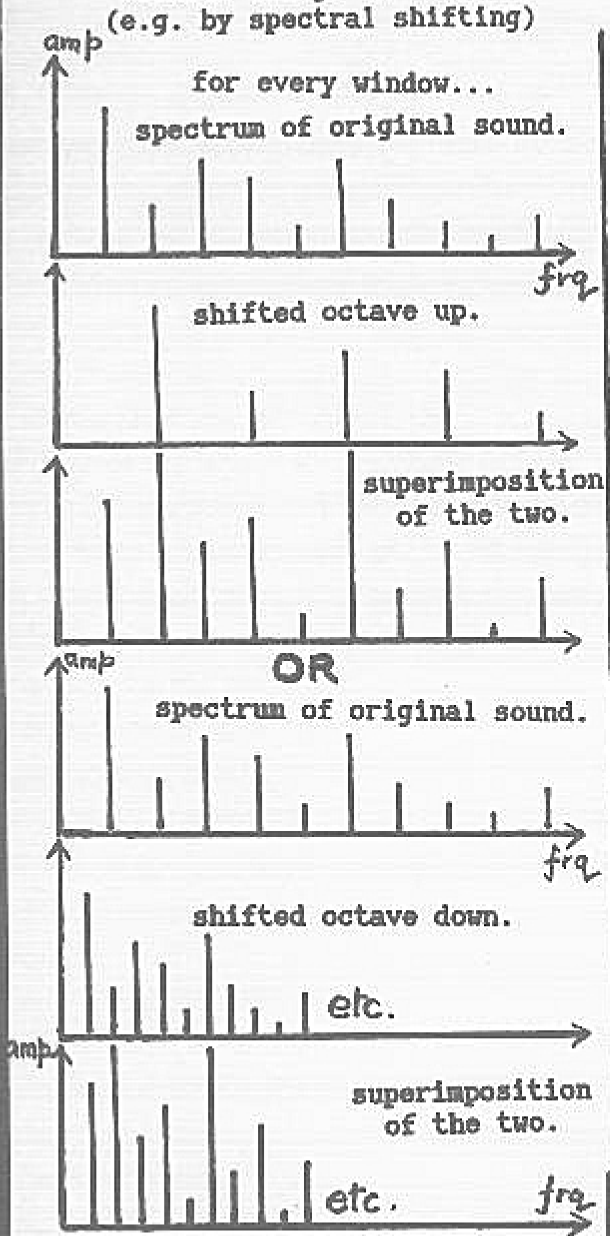


RE-SPATIALISING

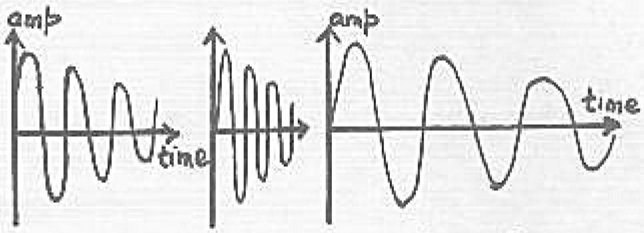


OCTAVE STACKING

Octave shift
without change in duration
(e.g. by spectral shifting)

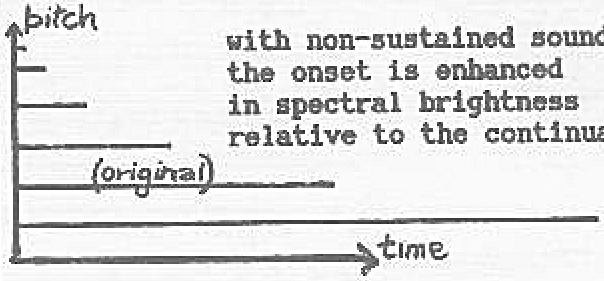


Octave shift
with duration change
(e.g. by tape-speed variation)



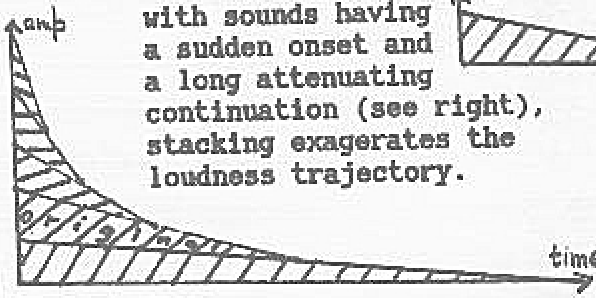
octave up
(shorter)

octave down
(longer)



with non-sustained sounds
the onset is enhanced
in spectral brightness
relative to the continuation.

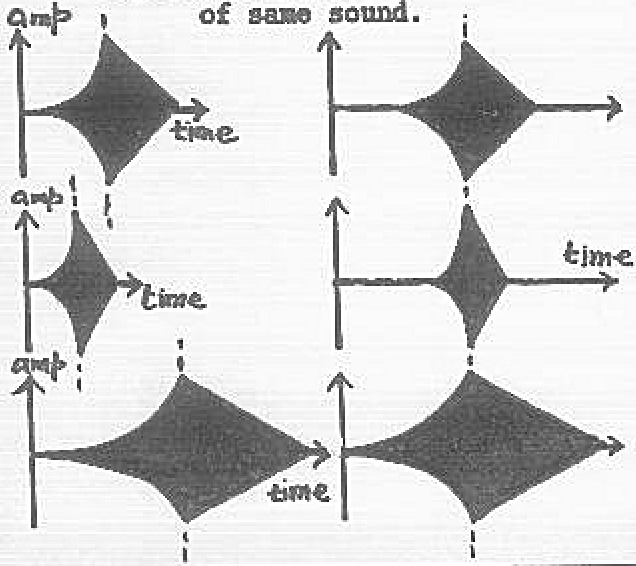
with sounds having
a sudden onset and
a long attenuating
continuation (see right),
stacking exagerates the
loudness trajectory.



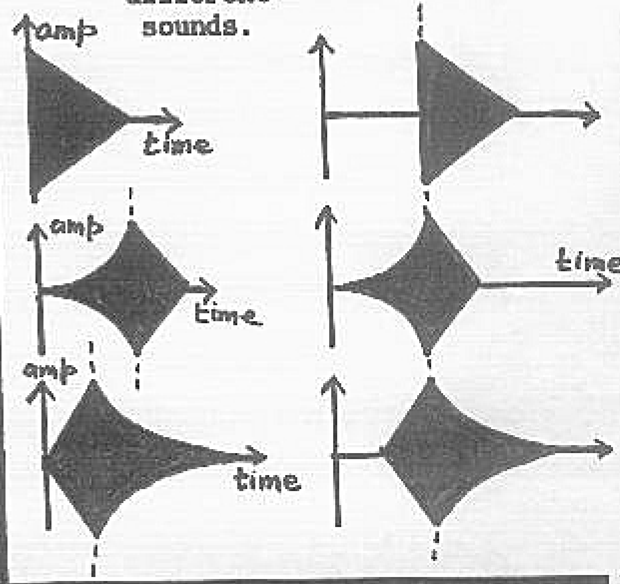
If sounds do not have a sudden onset,
attack-synchronisation might be
more appropriate.

ONSET SYNCHRONISATION

tape-speed varied versions
of same sound.

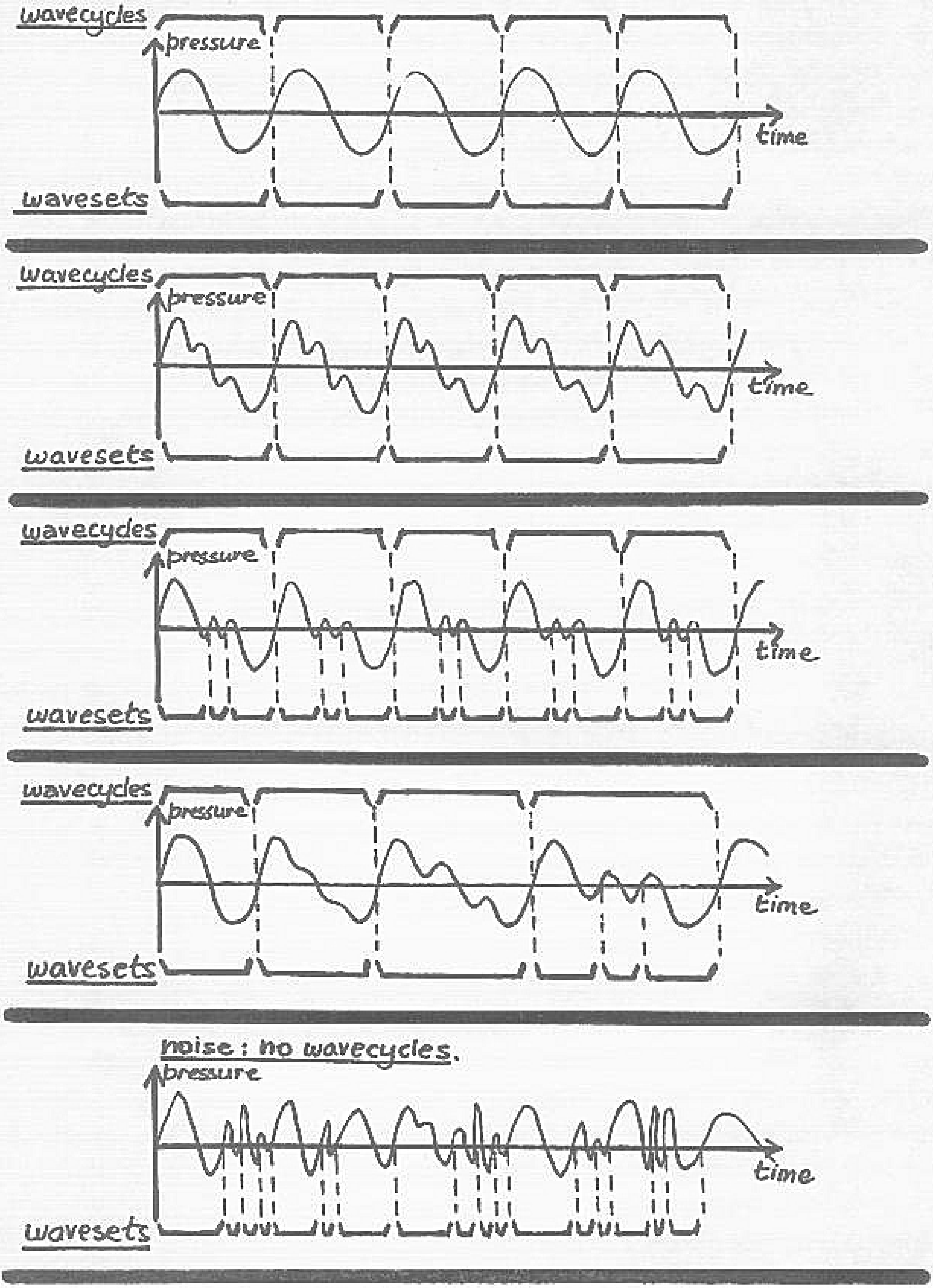


different
sounds.

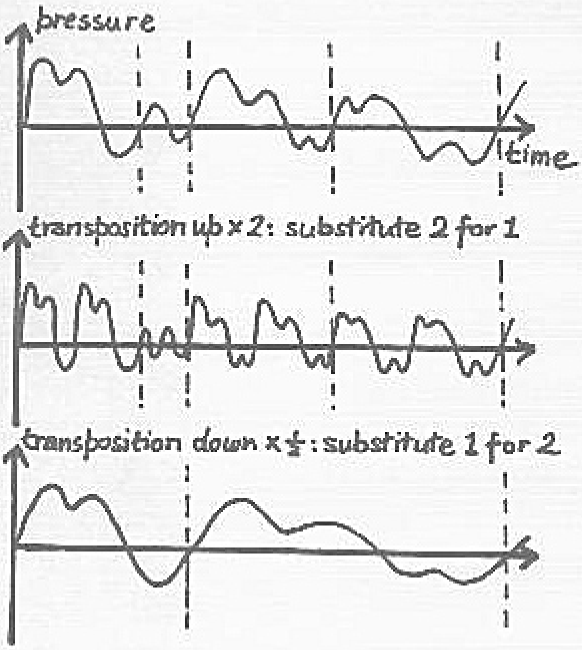


WAVECYCLES AND WAVESETS

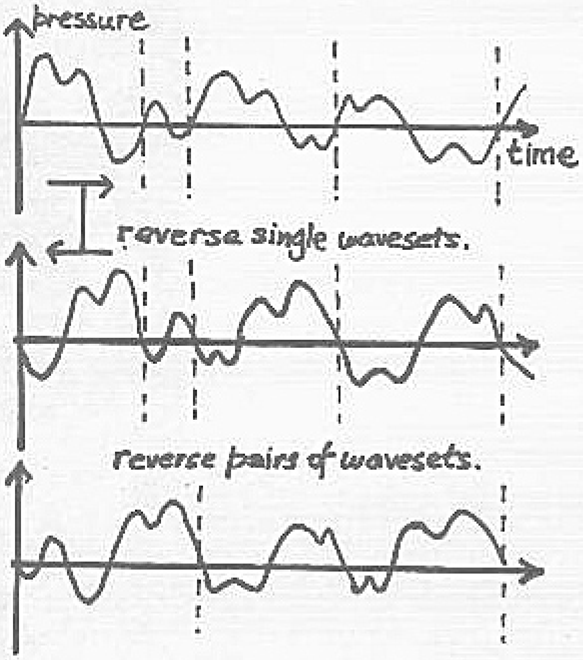
WAVECYCLE : Wavelength of sound, where clearly pitched.
WAVESET : Distance from a zero-crossing to a 3rd zero-crossing.



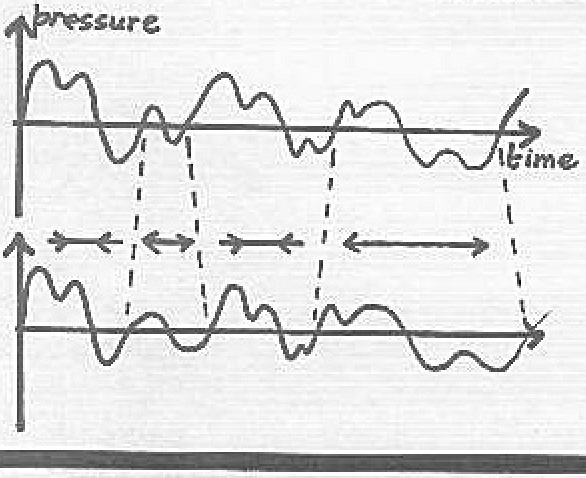
WAVESET TRANSPOSITION



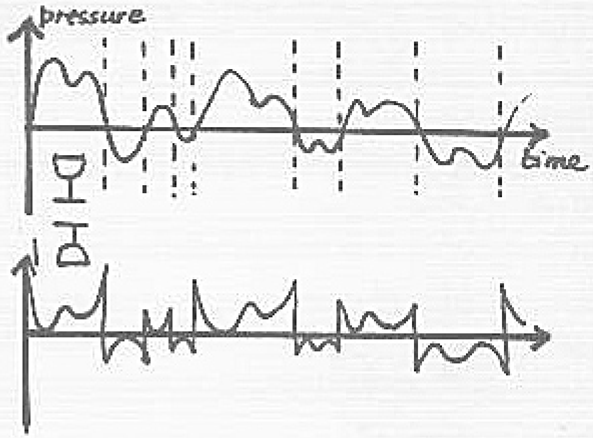
WAVESET REVERSAL



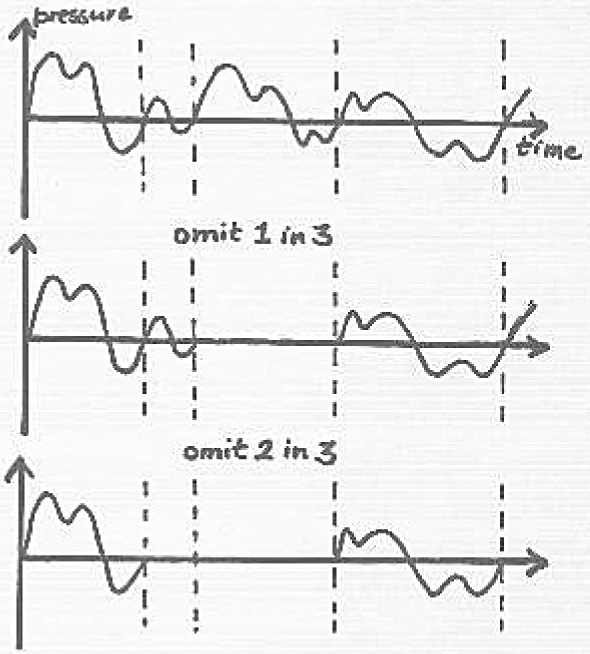
WAVESET SHAKING



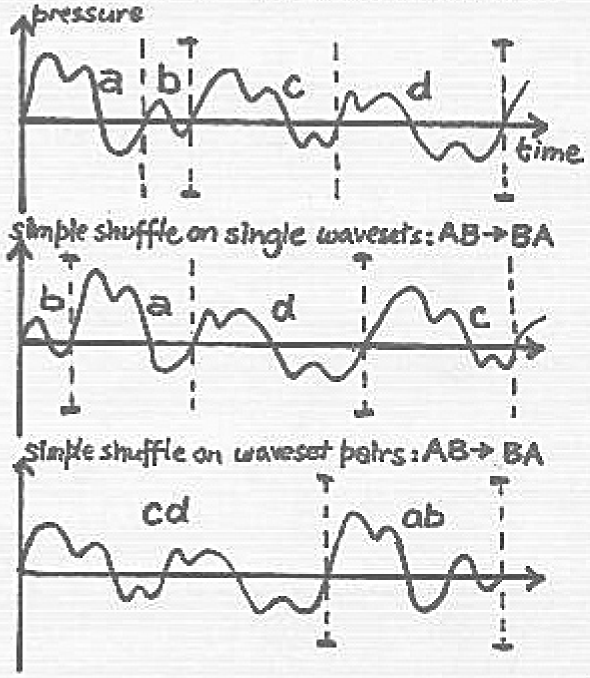
WAVESET INVERSION



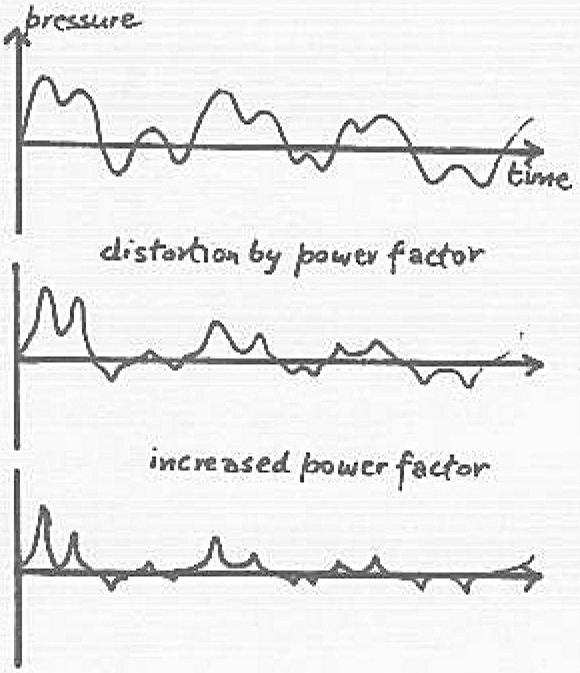
WAVESET OMISSION



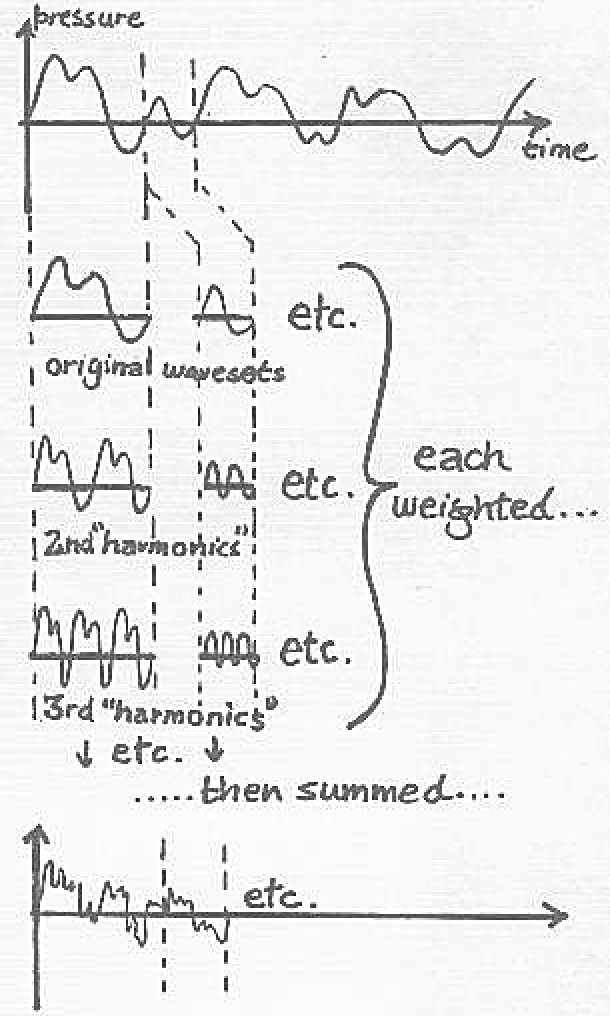
WAVESET SHUFFLING



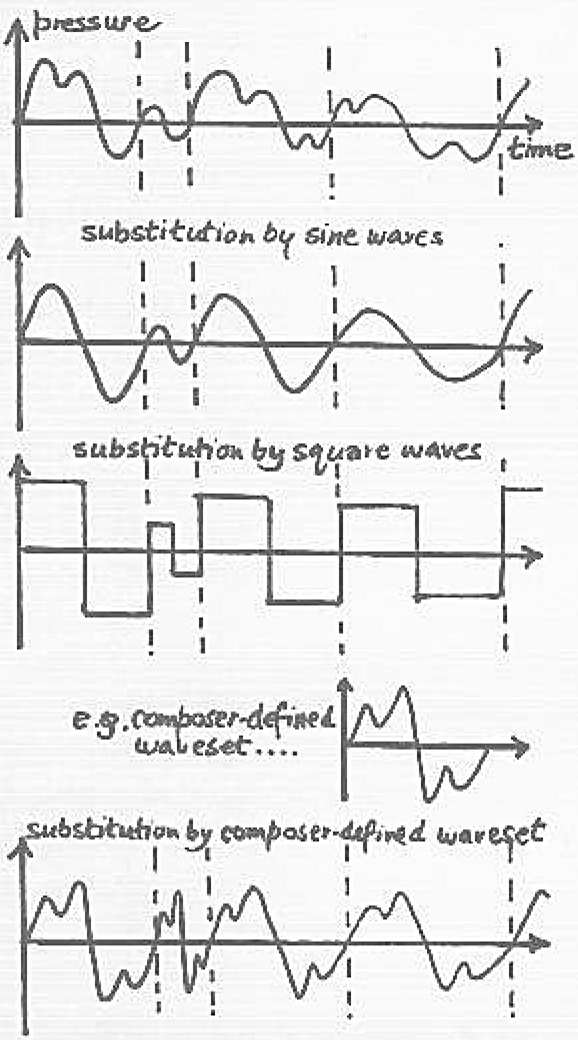
WAVESET DISTORTION



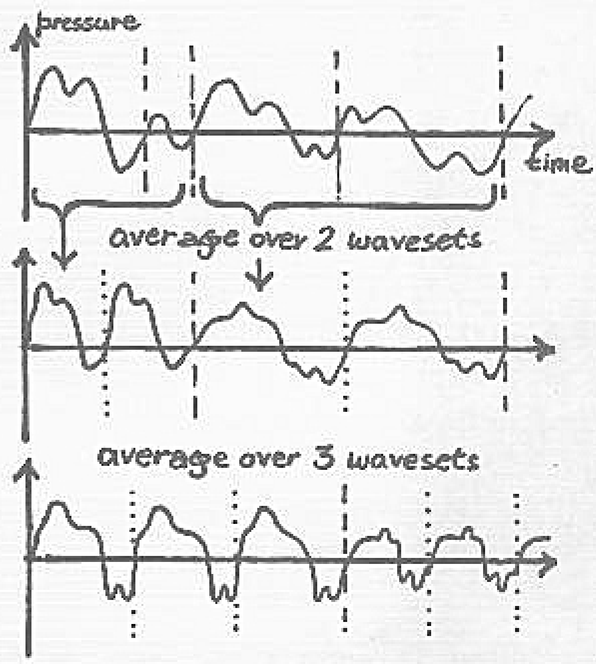
WAVESET HARMONIC DISTORTION



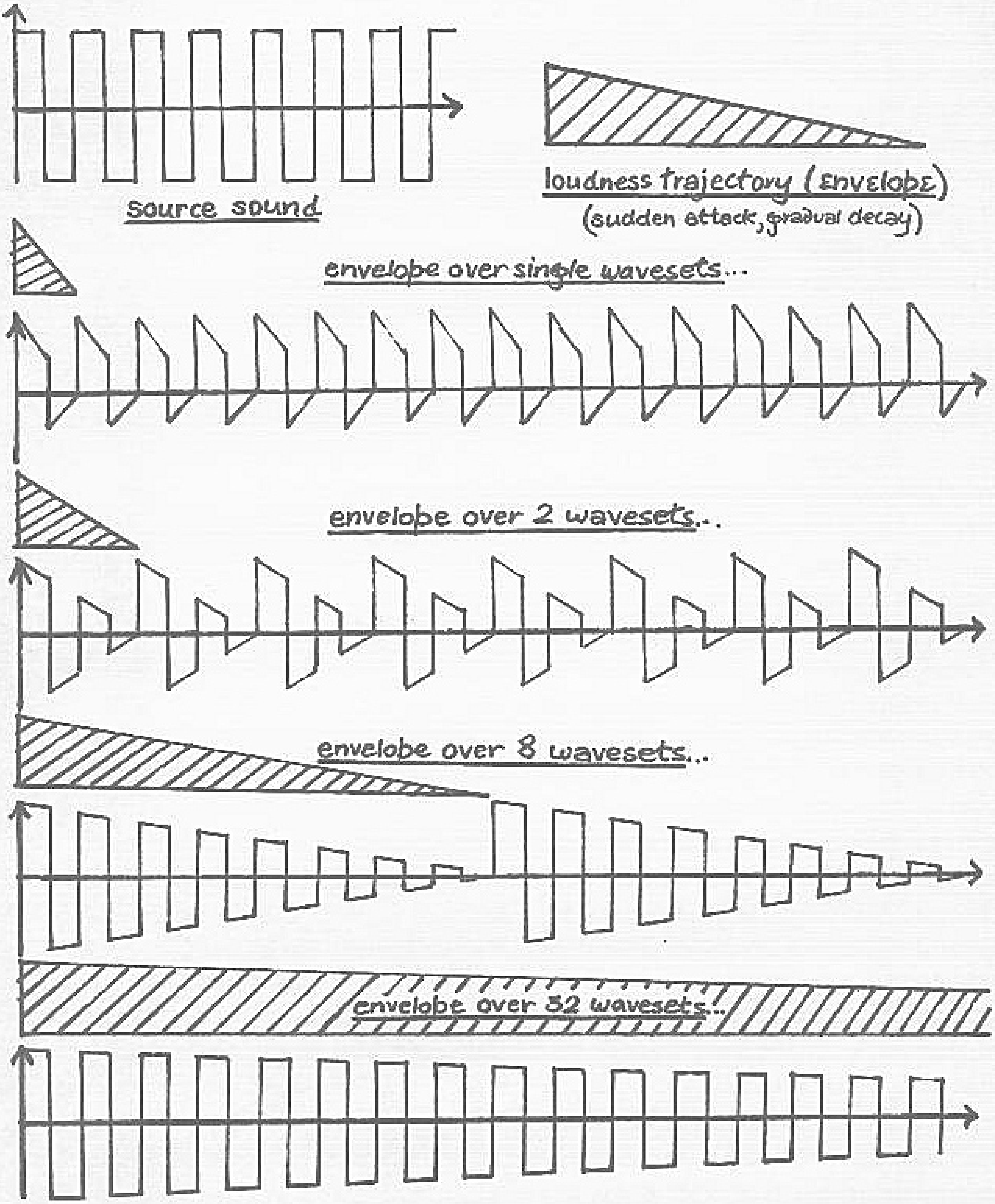
WAVESET SUBSTITUTION

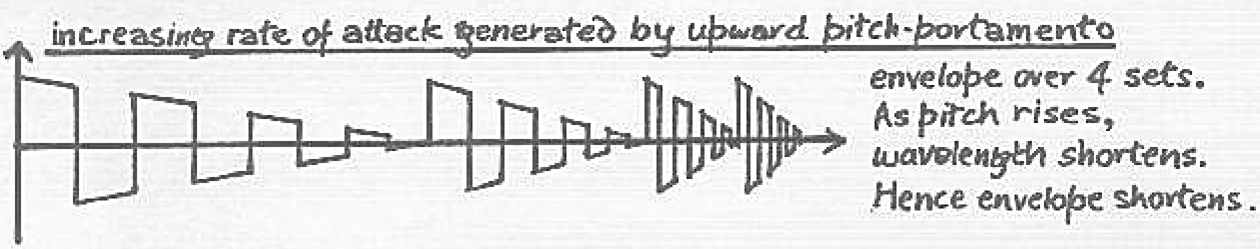


WAVESET AVERAGING

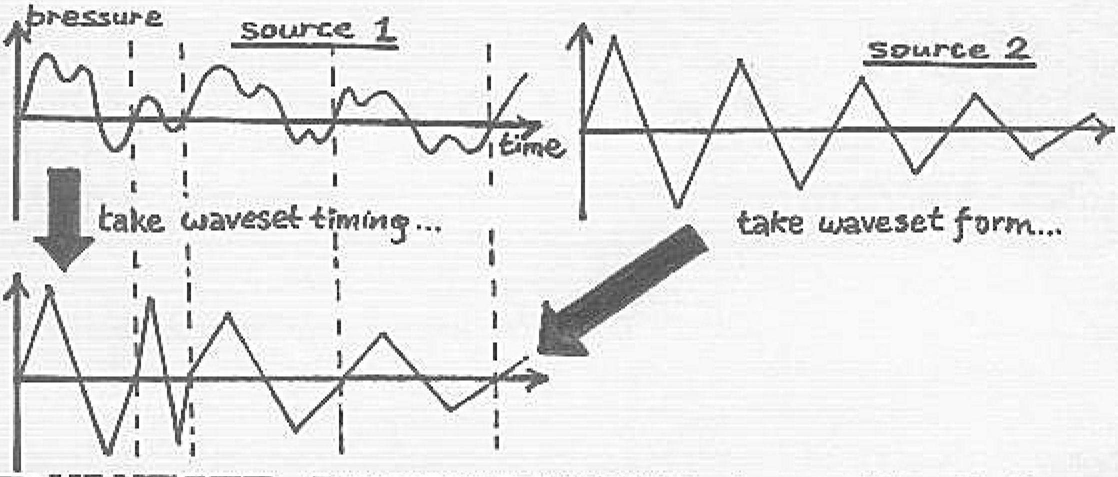


WAVESET ENVELOPING

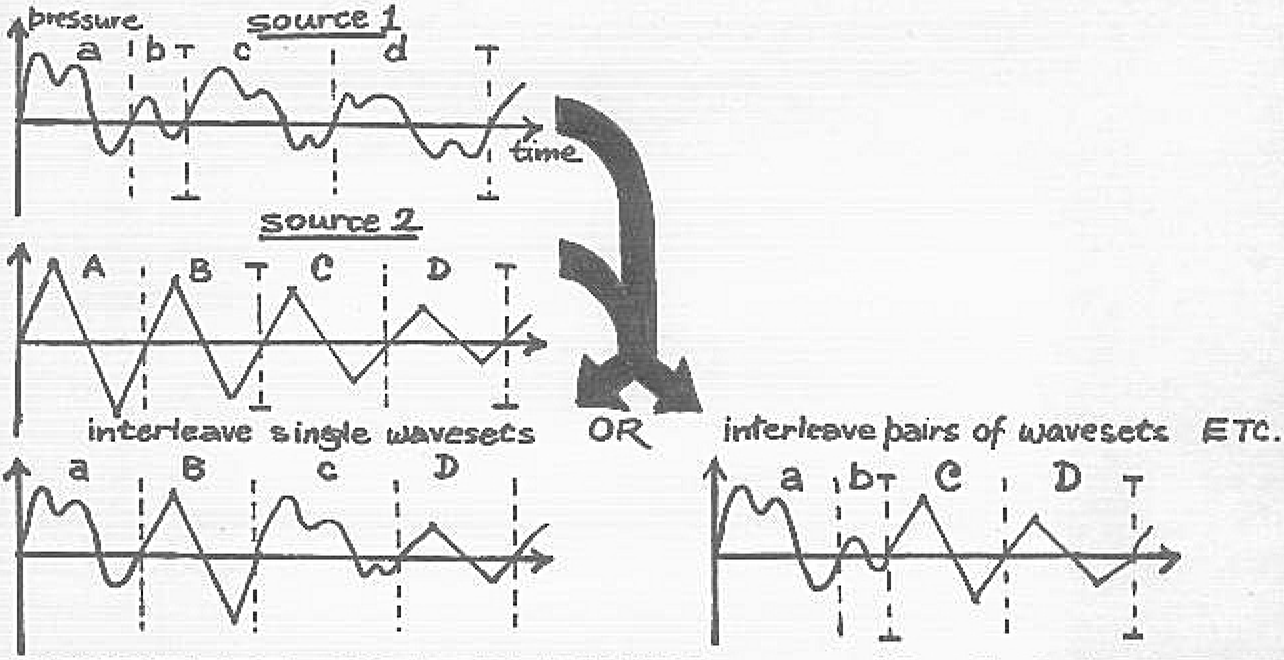




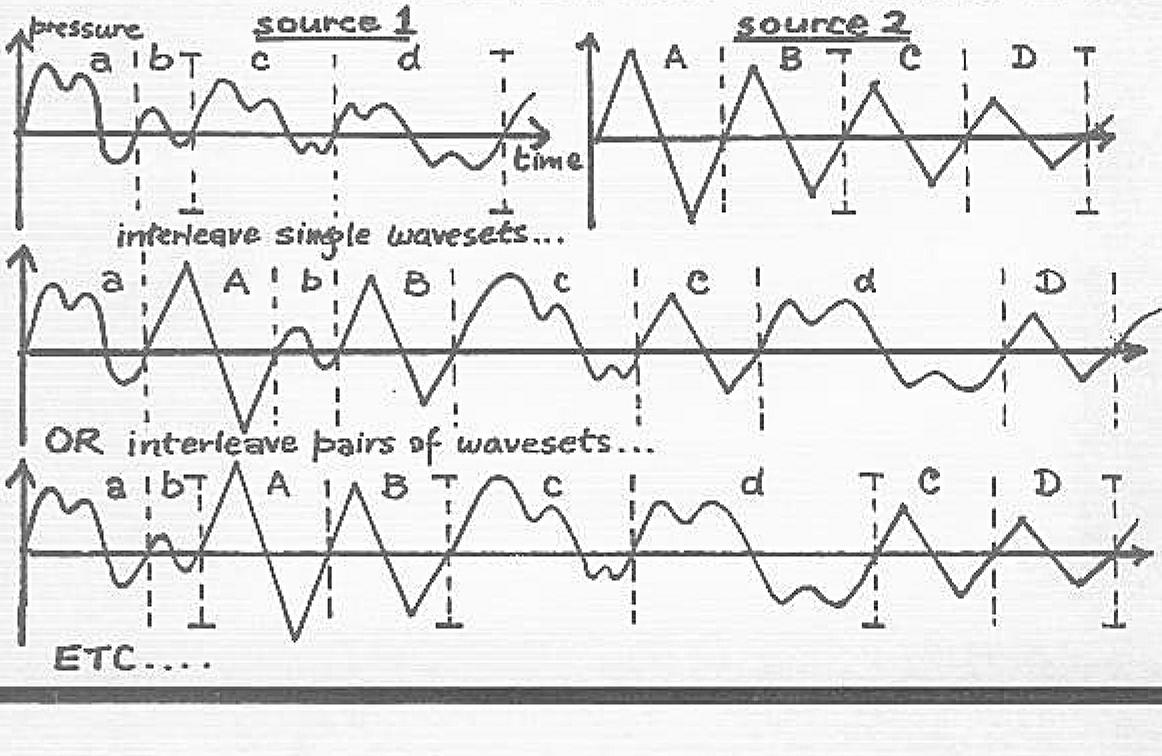
WAVESET TRANSFER



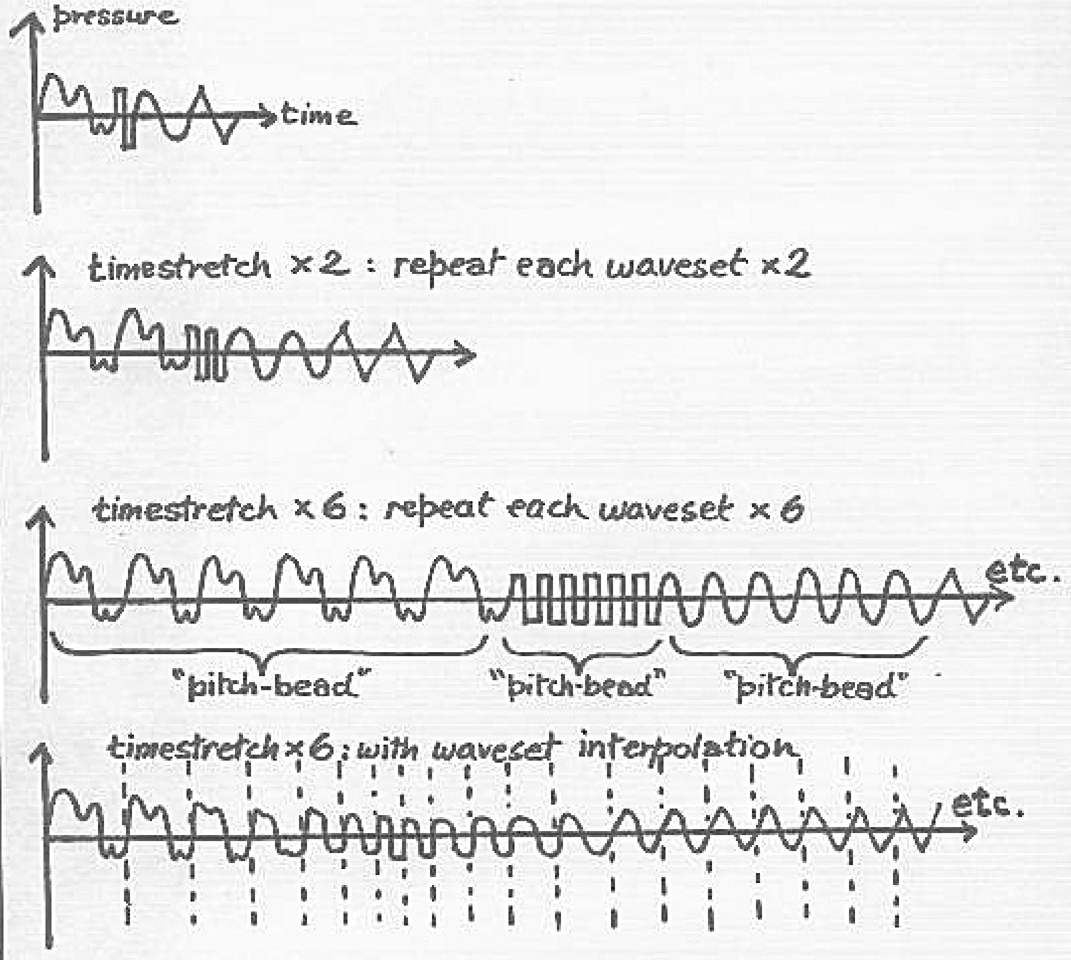
WAVESET INTERLEAVING : method 1



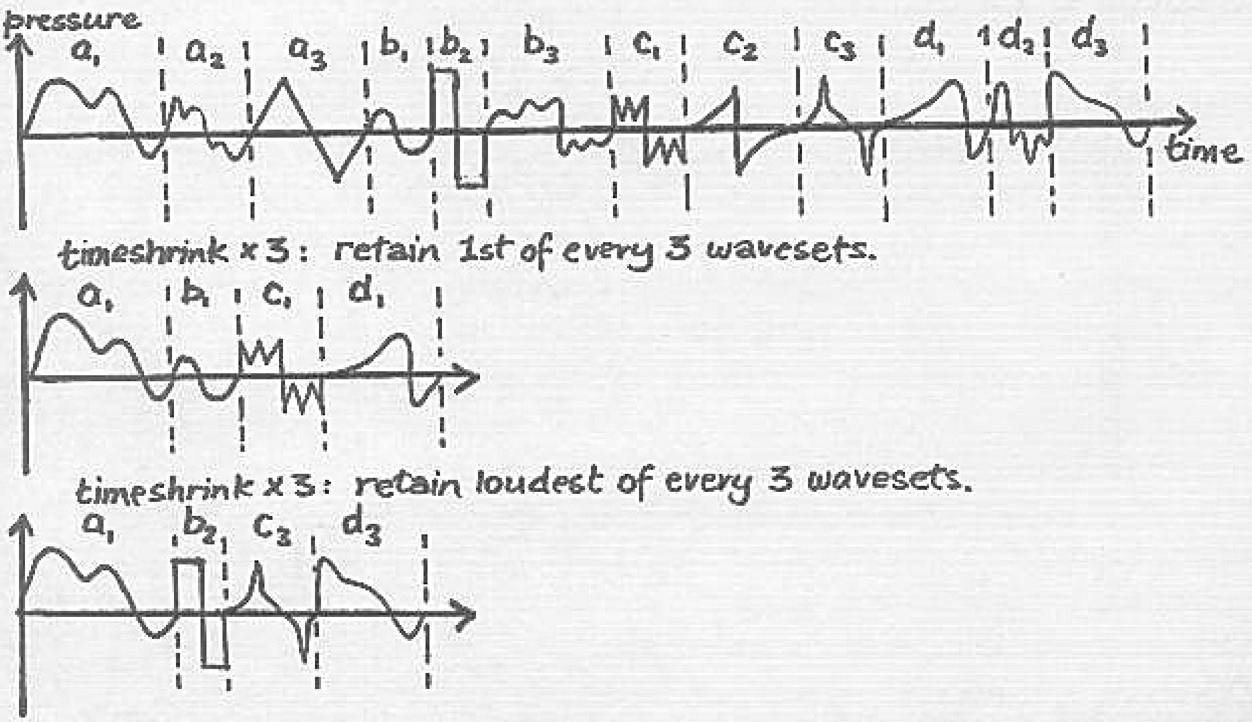
WAVESET INTERLEAVING : method 2



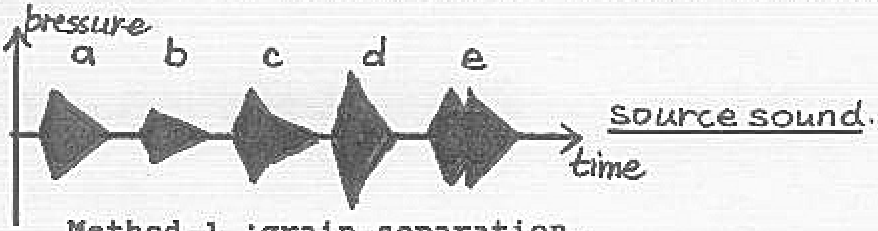
WAVESET TIME-STRETCHING



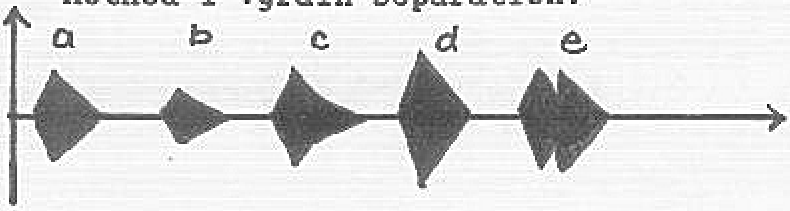
WAVESET TIME-SHRINKING



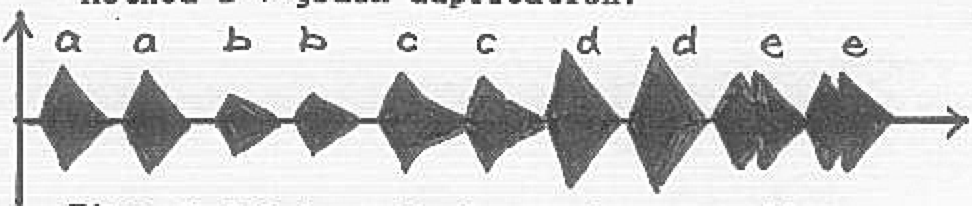
GRANULAR TIME-STRETCHING



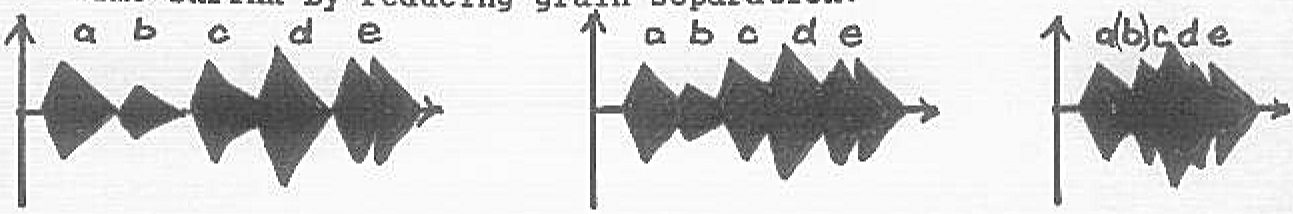
Method 1 :grain separation.



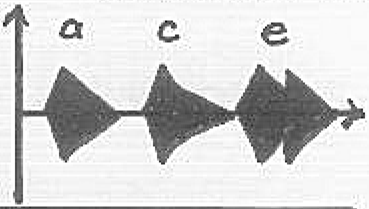
Method 2 : grain duplication.



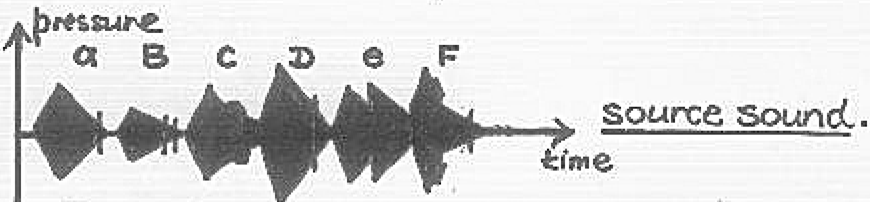
Time-shrink by reducing grain separation.



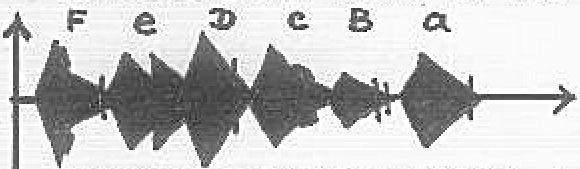
Time-shrink via grain deletion.



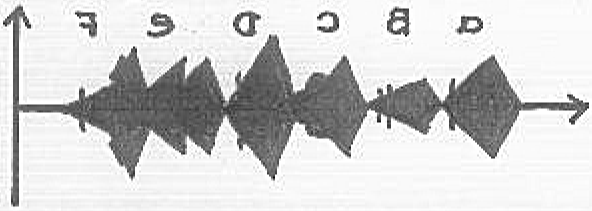
GRANULAR REVERSAL

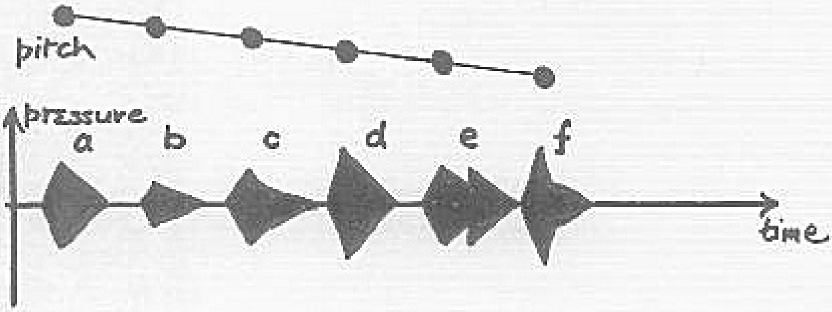


Reversing the order of the grains in the sound.

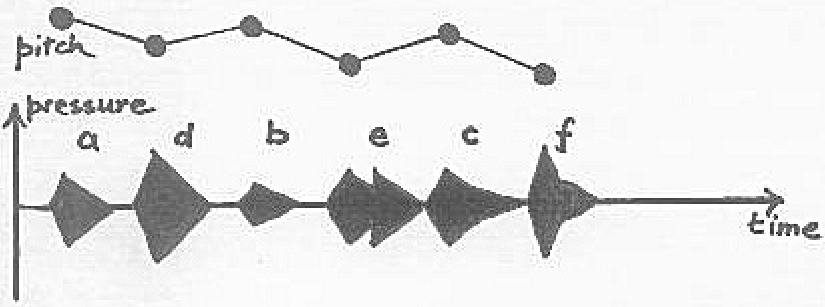


compare - SOUND REVERSAL : Reverses the grains themselves.

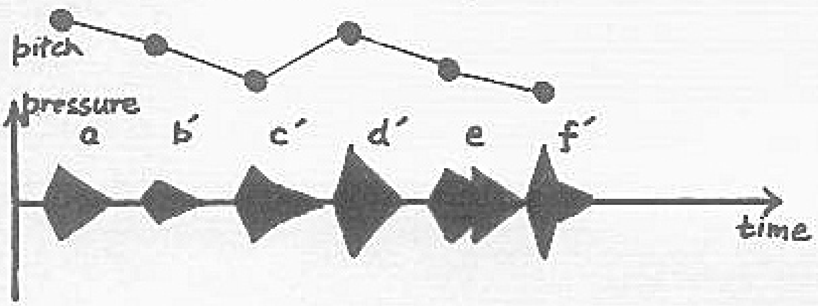




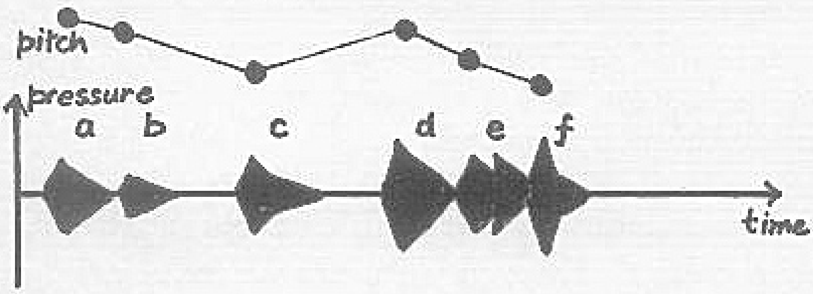
Reorder grains (pitch sequence is disturbed).

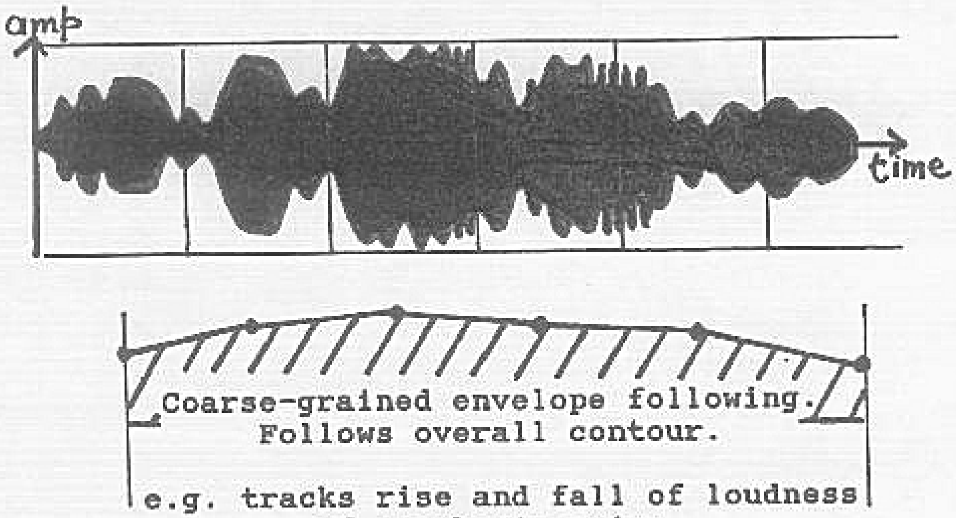


Alter pitch of grains (retain original order).



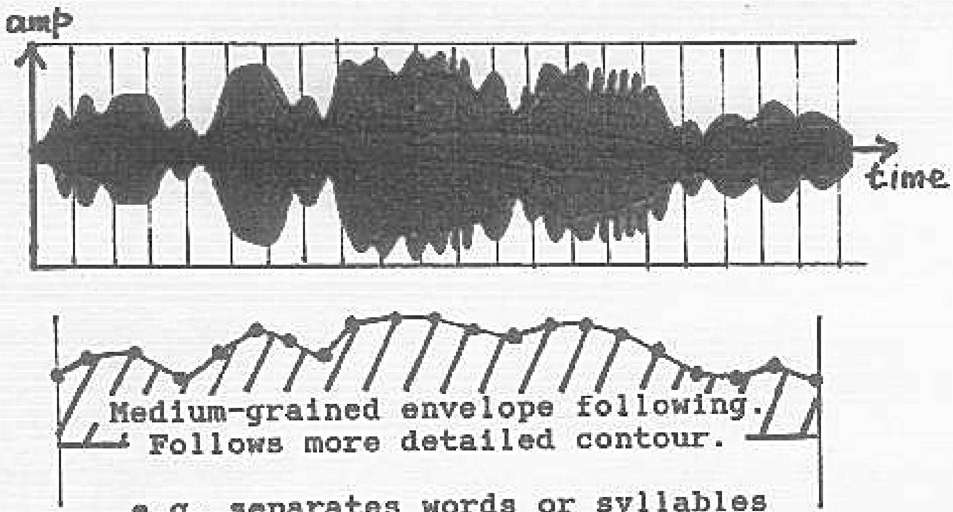
Alter pitch and scatter timing of grains.





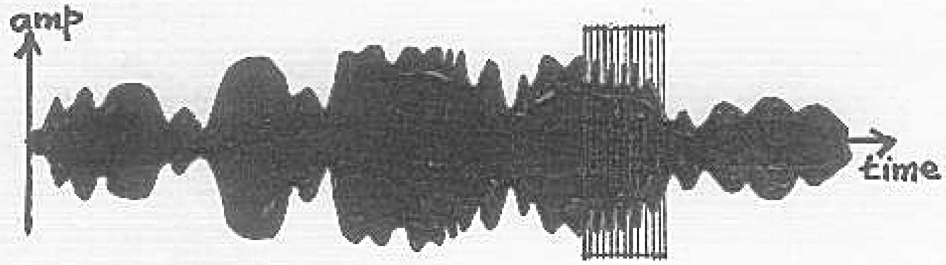
Coarse-grained envelope following.
Follows overall contour.

e.g. tracks rise and fall of loudness
of speaker's voice.



Medium-grained envelope following.
Follows more detailed contour.

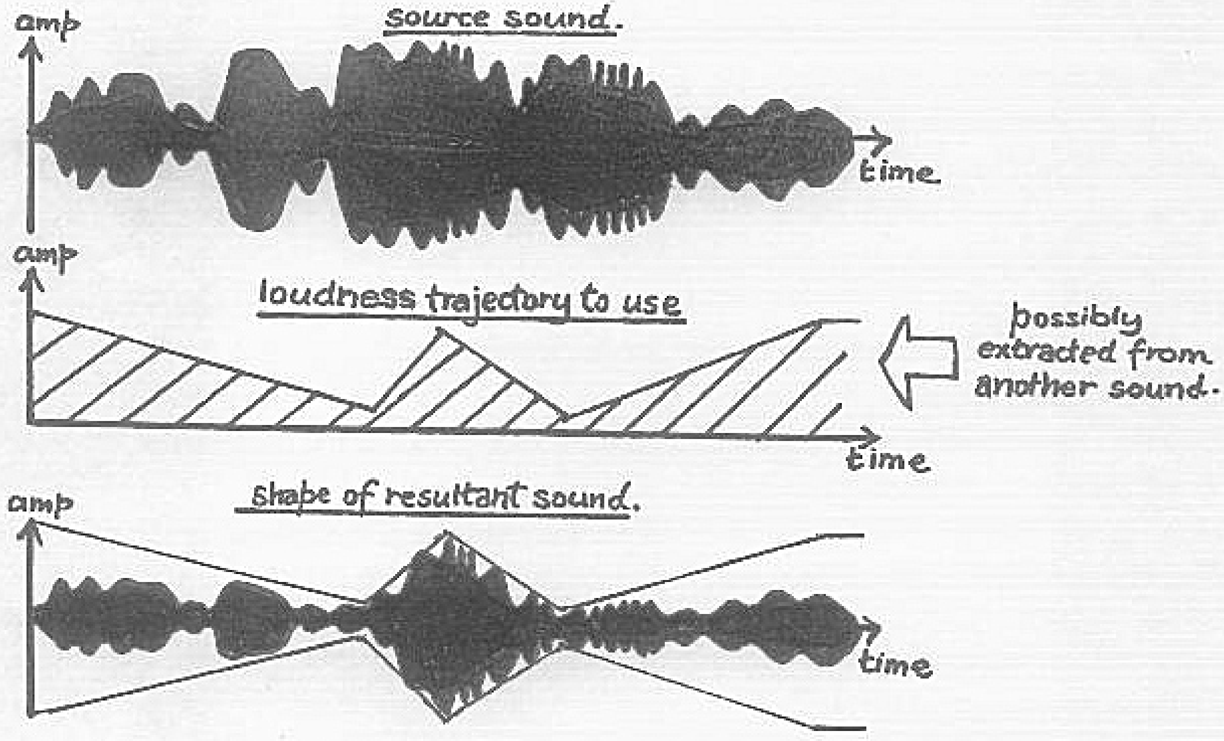
e.g. separates words or syllables
in speaker's voice.



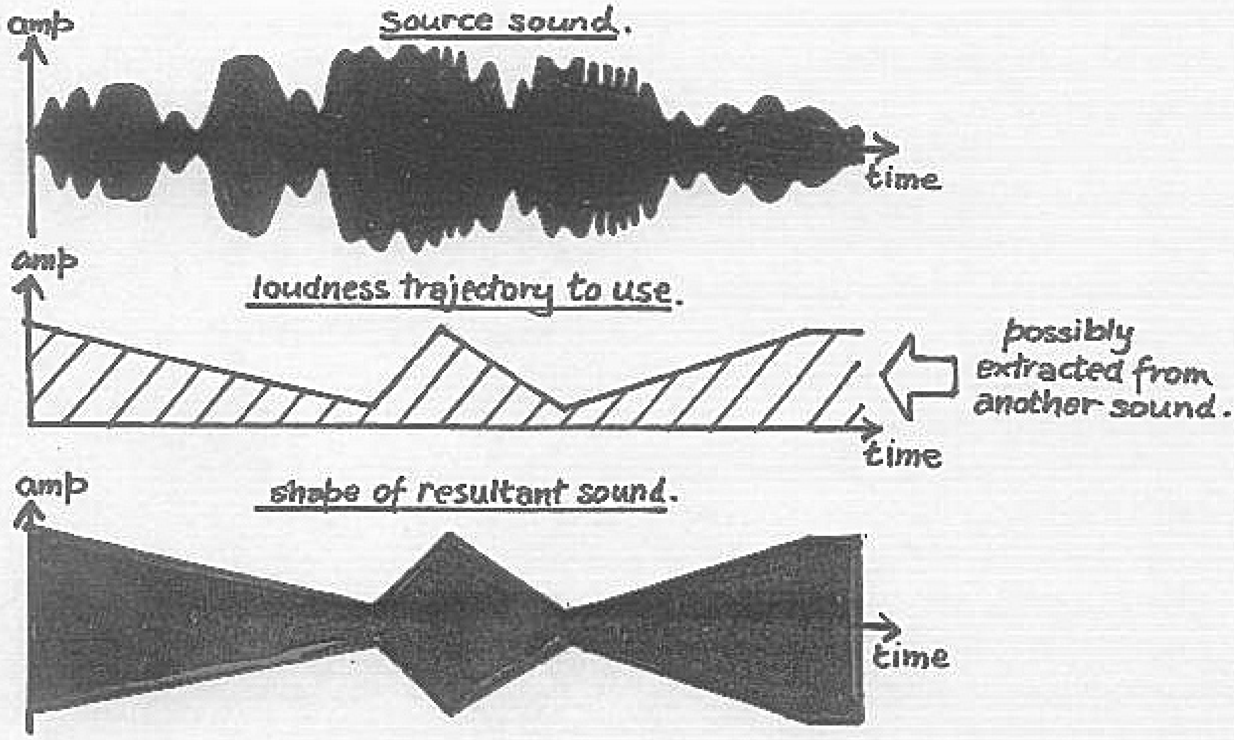
Fine-grained
envelope following.
Follows finest detail
of contour.

e.g. separates individual
grains of a rolled 'r'
in speaker's voice.

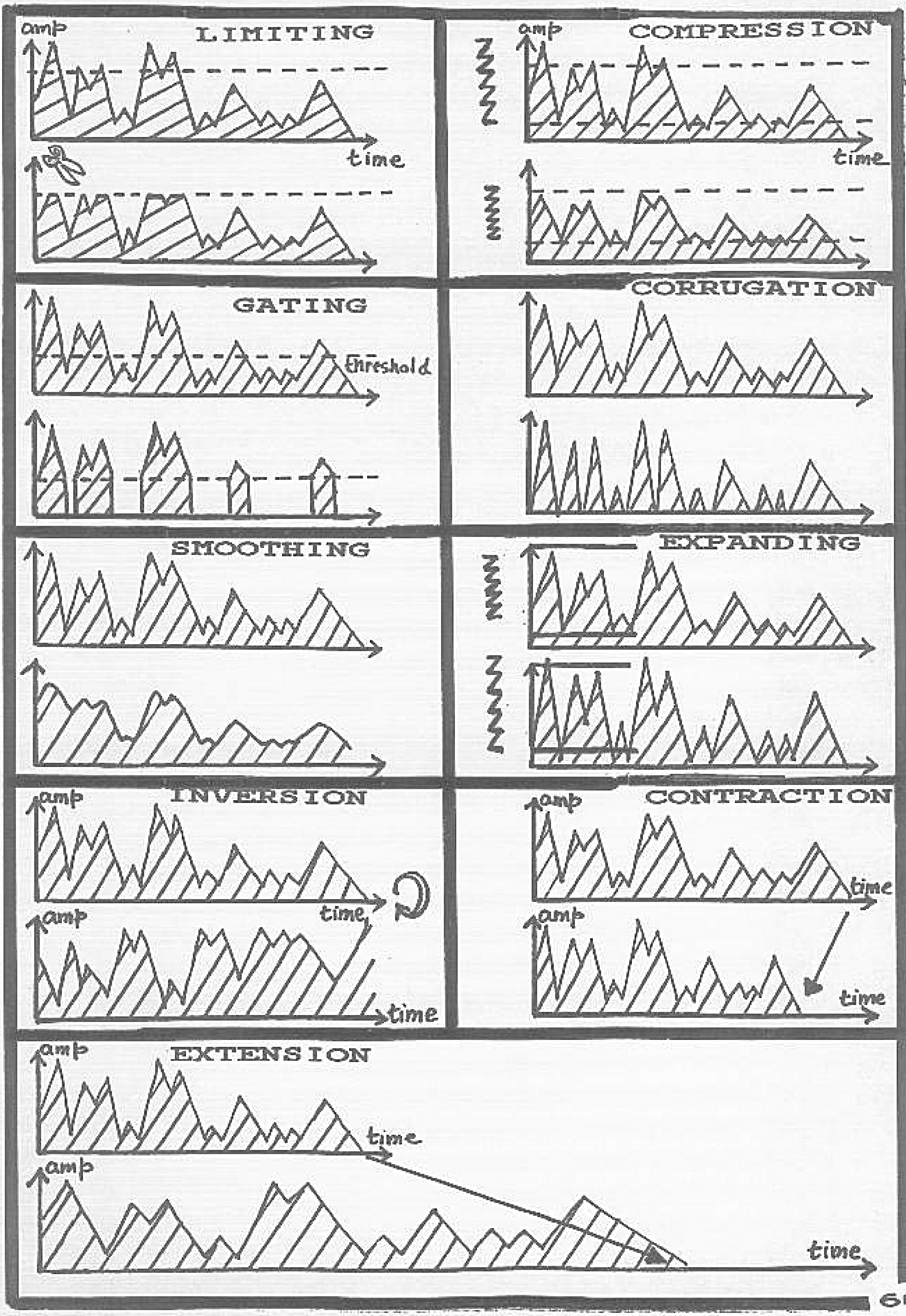
ENVELOPING



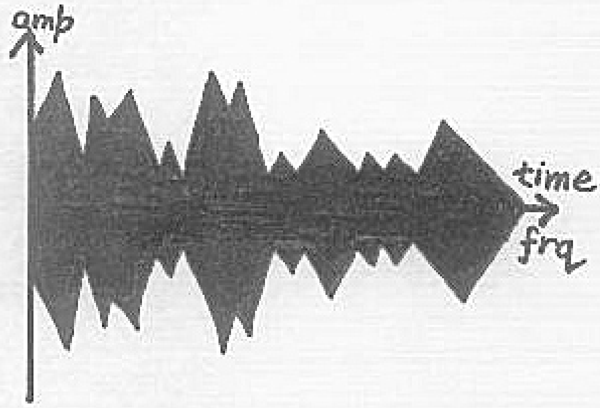
ENVELOPE SUBSTITUTION



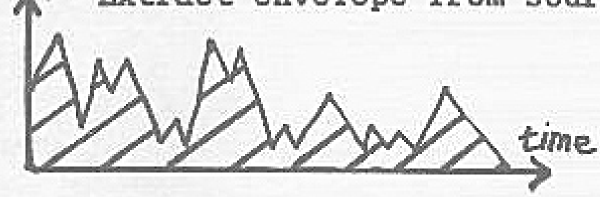
ENVELOPE TRANSFORMATION



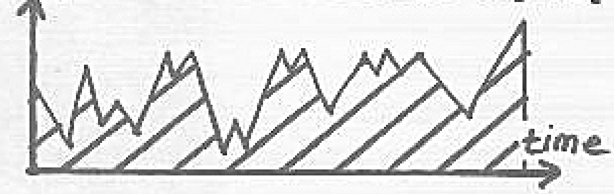
ENVELOPE TRANSFORMATION & SUBSTITUTION



Extract envelope from source.



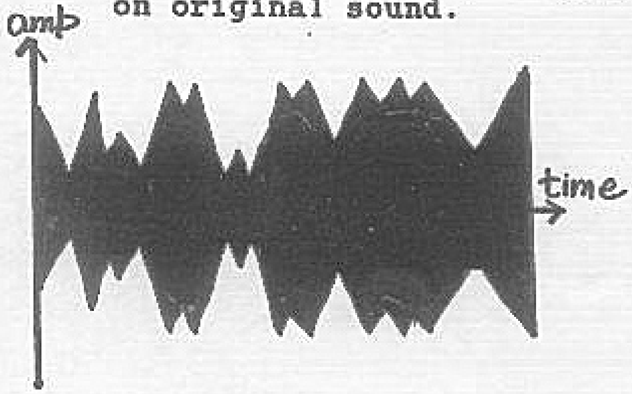
Transform envelope e.g. by inversion.

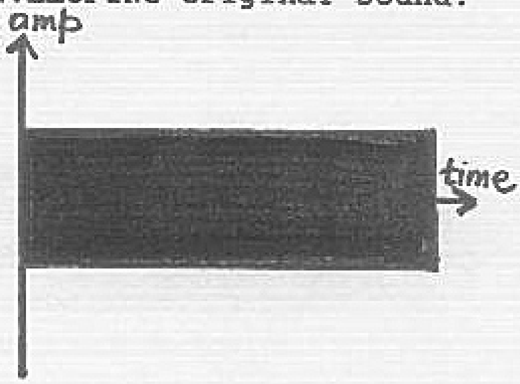


SUBSTITUTE new envelope on original sound.

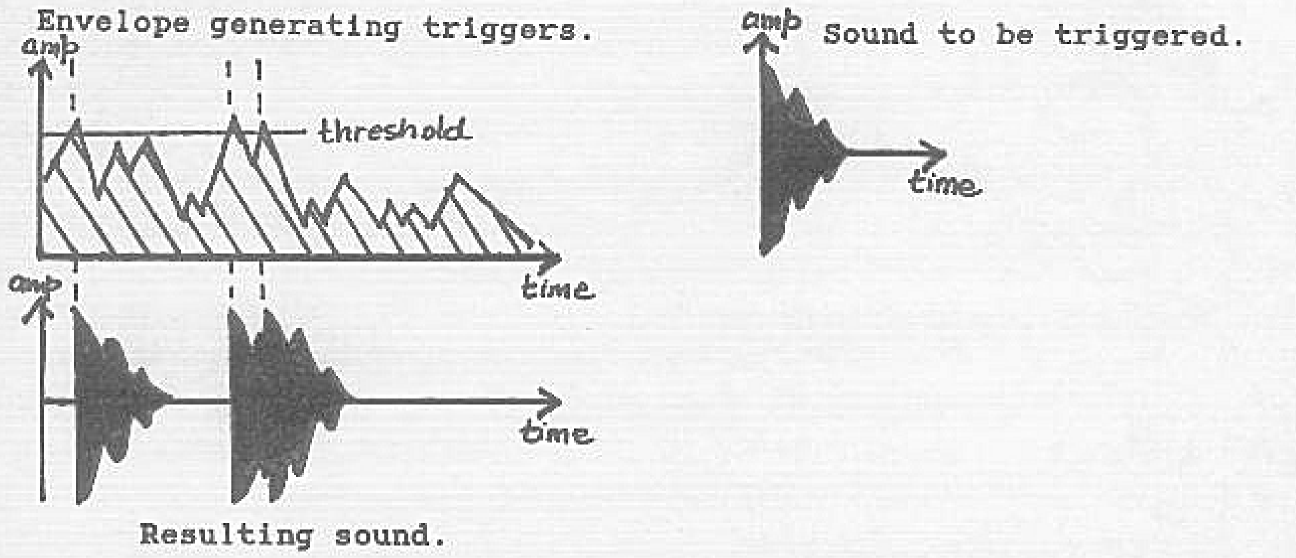
OR

Use new envelope for **ENVELOPING** original sound.

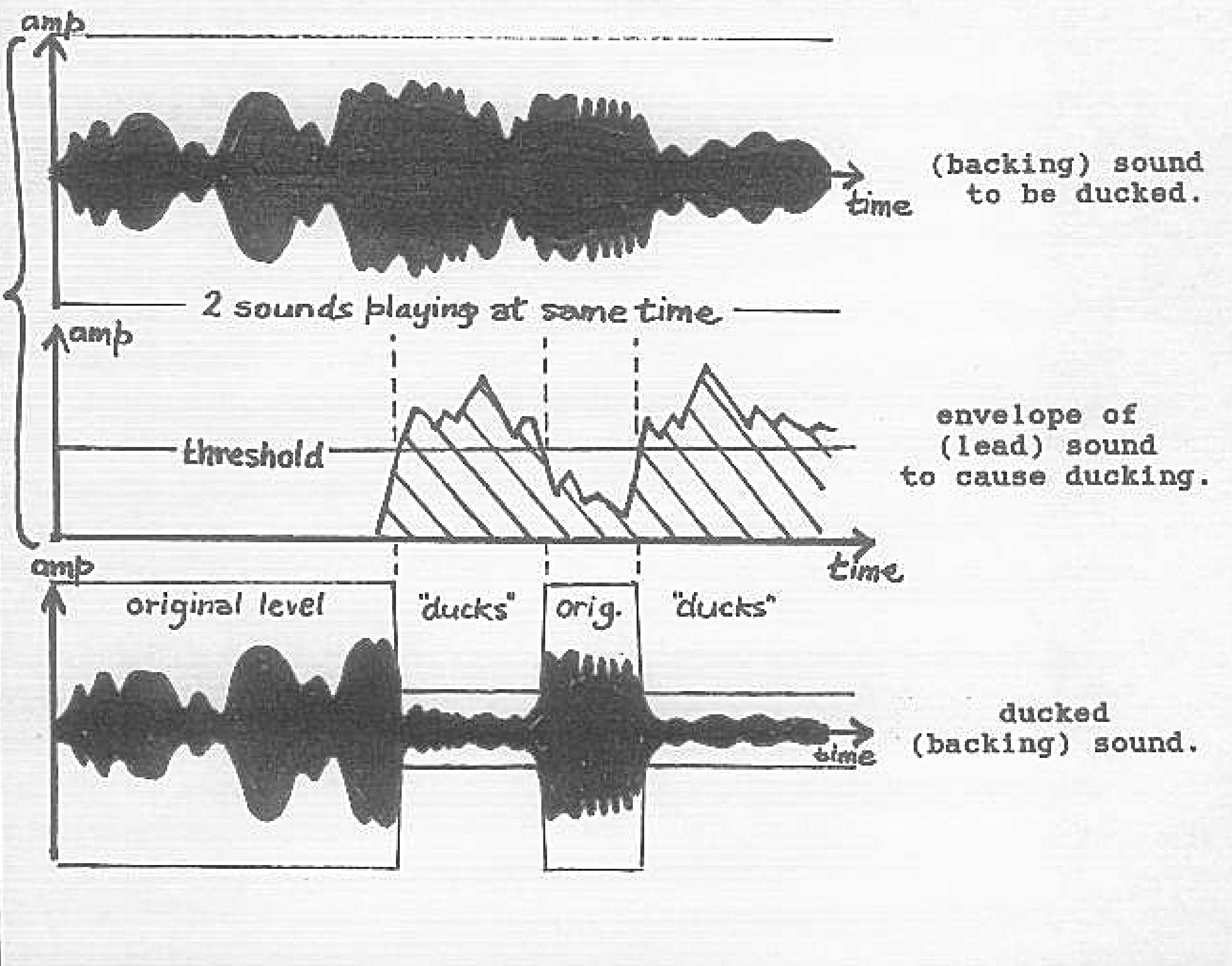




TRIGGERING

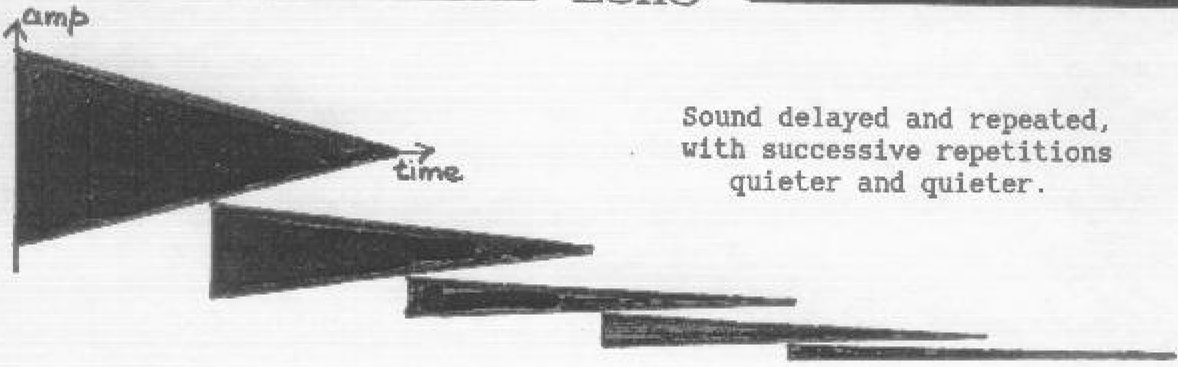


DUCKING



DELAY

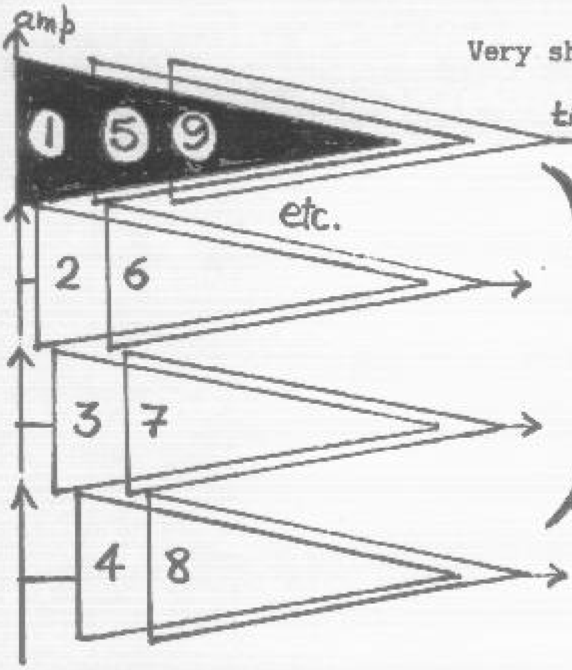
ECHO



Sound delayed and repeated,
with successive repetitions
quieter and quieter.

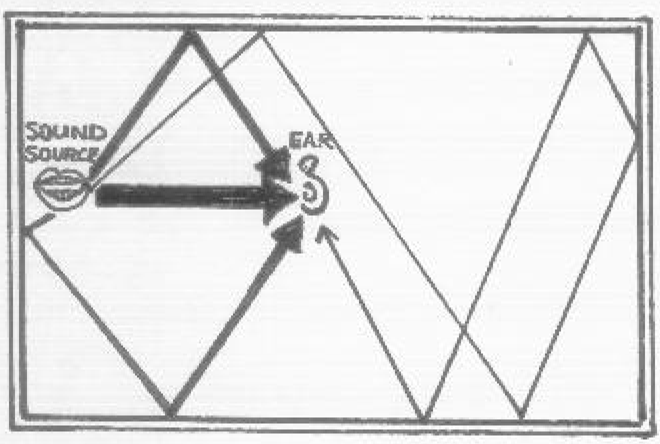
COMB FILTERING

Very short delay time.

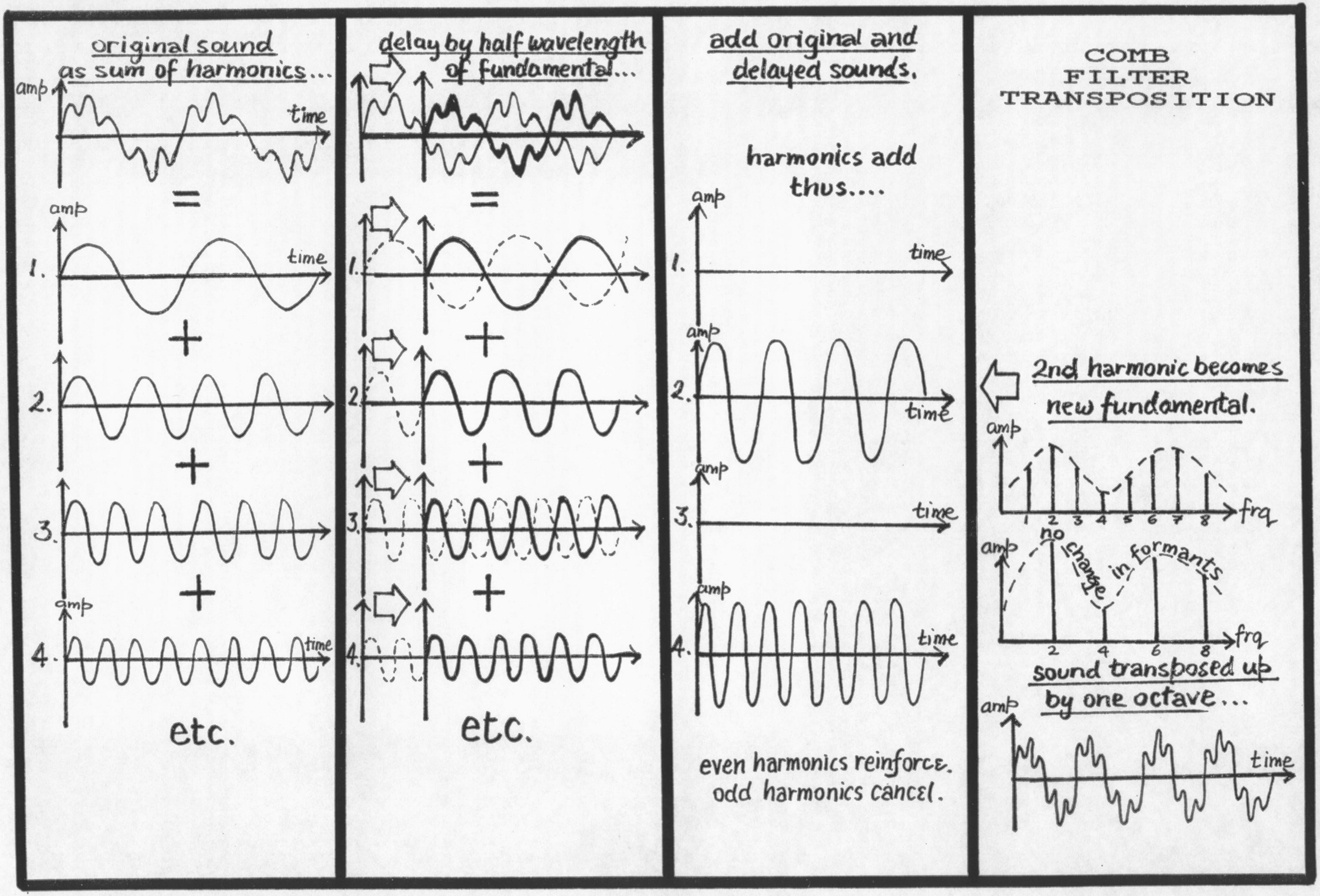


If number of repeats
per second is very high
(e.g. 16+ per second)
we hear a pitch.

REVERBERATION

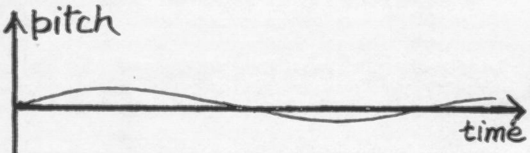


In an enclosed space,
sound reaching the ear directly
is mixed with many
differently delayed
versions of the sound.
The delays are caused by
the different distances
the sound must travel
on its various paths to the ear.
The delayed versions
are (usually) quieter
than the direct sound,
& have been changed in quality
through the filtering effects
of reflection from
different kinds of surfaces.

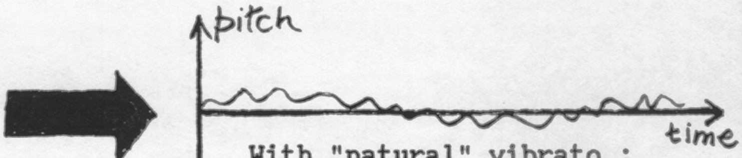


original sound
as sum of harmonics...
amp
time
=
1.
2.
3.
4.
+
etc.
delay by half wavelength
of fundamental...
etc.
add original and
delayed sounds.
harmonics add
thus....
even harmonics reinforce.
odd harmonics cancel.
COMB
FILTER
TRANSPOSITION
2nd harmonic becomes
new fundamental.
frq
1 2 3 4 5 6 7 8
no change in formants
2 4 6 8
sound transposed up
by one octave...

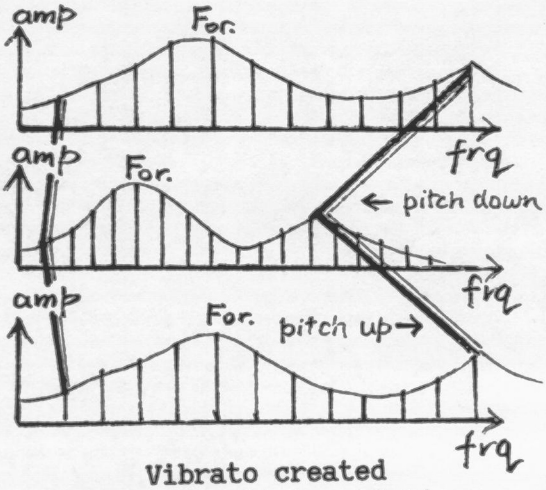
VIBRATO



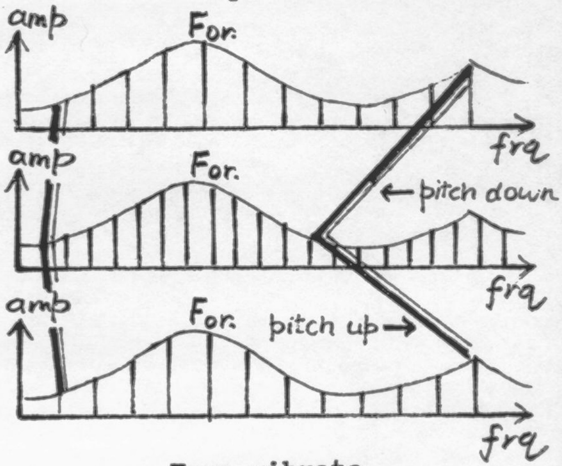
Original pitch trajectory.



With "natural" vibrato :
Slightly irregular
cyclical fluctuations
of pitch.

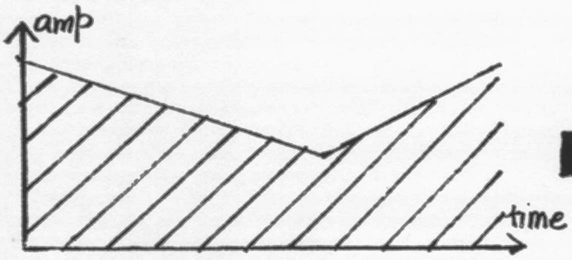


Vibrato created
merely by pitch-variation
does not preserve
formant position

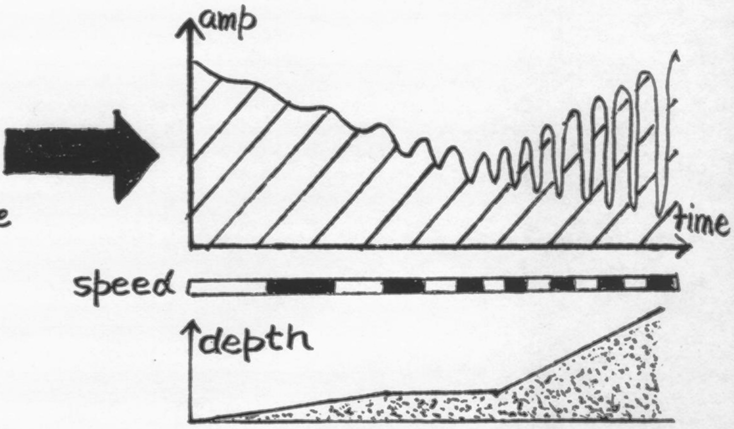


True vibrato
retains original
formant position
as pitch moves

TREMOLO

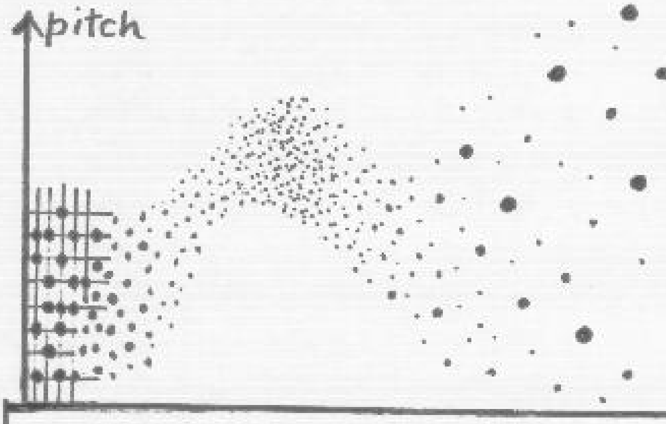


Original loudness trajectory.



With (one example of)
time-varying tremolo.

TEXTURE



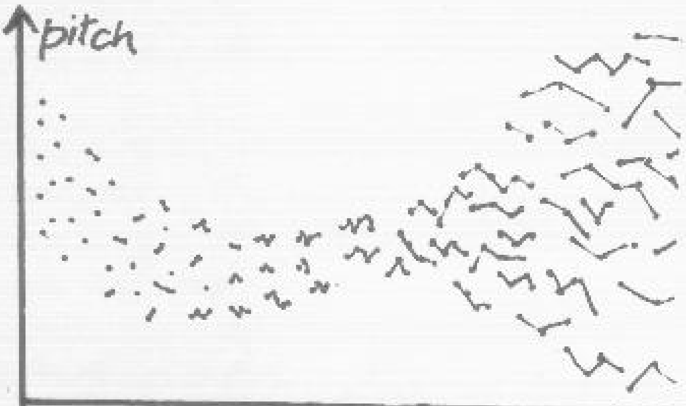
<u>specify variation of</u>....

event density
• event regularity
time quantisation grid
number of sound sources
hi & lo limits of loudness
hi & lo limits of event duration
hi & lo limits of pitch
harmonic field (or not)
range of event spatial position

medium ⇨ high ⇨ low	density
loud ⇨ quiet ⇨ mixed	loudness
low ⇨ high ⇨ full range	pitch range
quantised ⇨ random	time
harmonic field ⇨ random	pitch

· quiet event
• loud event

TEXTURE OF GROUPS

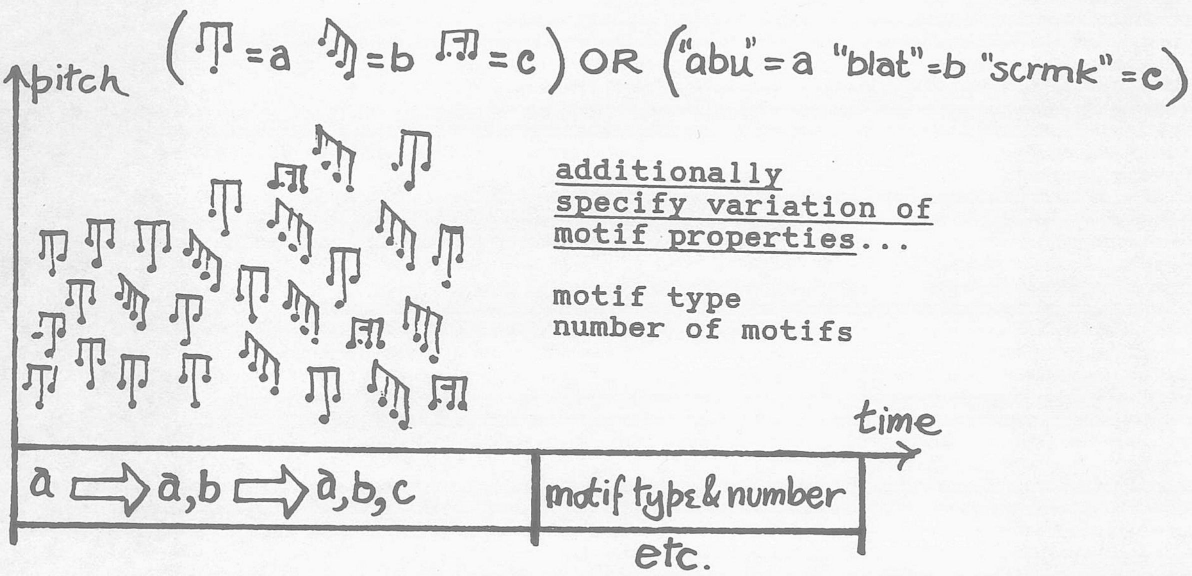


<u>additionally</u>
<u>specify variation of</u>
<u>group properties</u>...

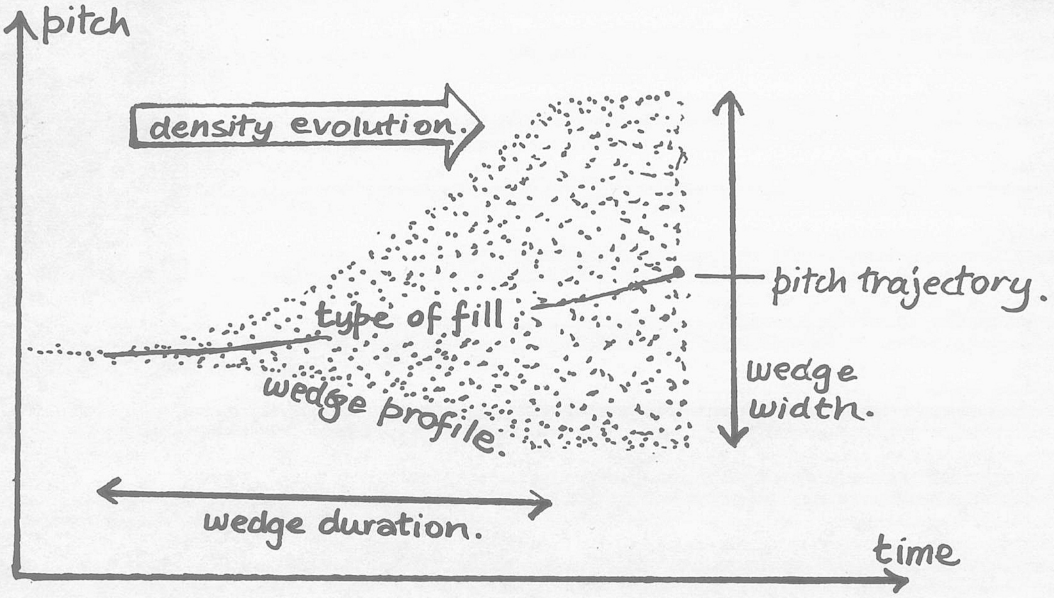
loudness trajectory
hi & lo limits of size
hi & lo limits of speed
hi & lo limits of pitchrange
harmonic field (or not)
spatial form (e.g. moving left)

1 ⇨ 1-3 ⇨ 2-4 ⇨ 3-6	group size
fast ⇨ slow	group speed
	etc.

TEXTURE OF MOTIFS

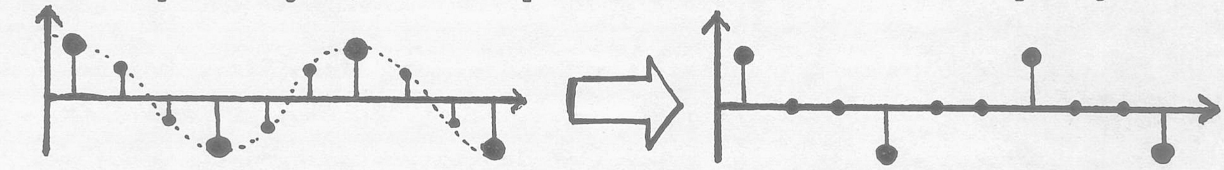


WEDGEING

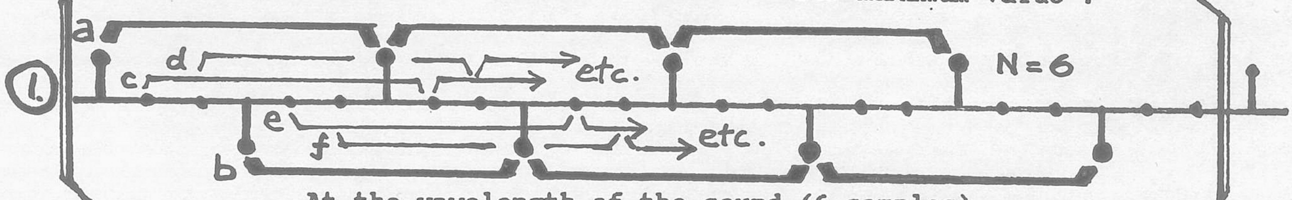


PITCH-TRACKING BY AUTO-CORRELATION

Example of an elementary auto-correlation procedure.
First clip the signal, so local peaks remain, but all other samples go to zero.

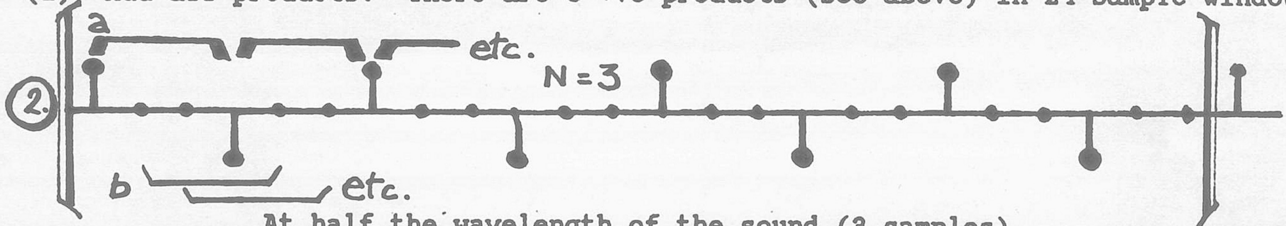


Add : the products of all pairs of samples separated by a step of N samples.
For what value of N does this sum have its maximum value ?



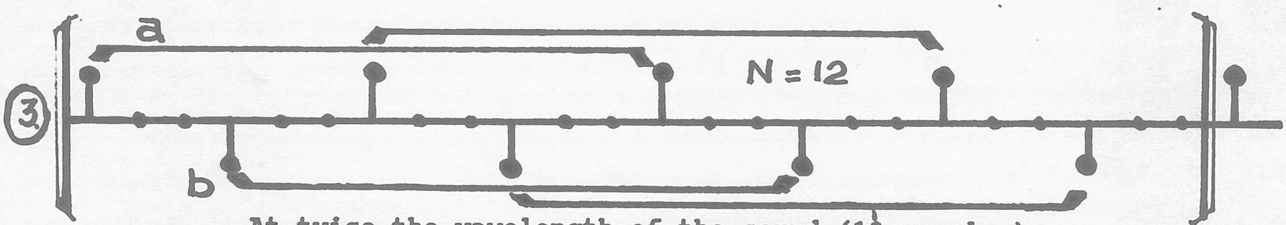
At the wavelength of the sound (6 samples).

(1) Multiply each sample value by value 6 samples ahead, within the window.
  (a) All these products are [+ve value] X [+ve value] = [+ve value].
  (b) All these products are [-ve value] X [-ve value] = [+ve value].
  (c-f) All these products are zero X zero = zero.
(2) Add all products. There are 6 +ve products (see above) in 24-sample window.



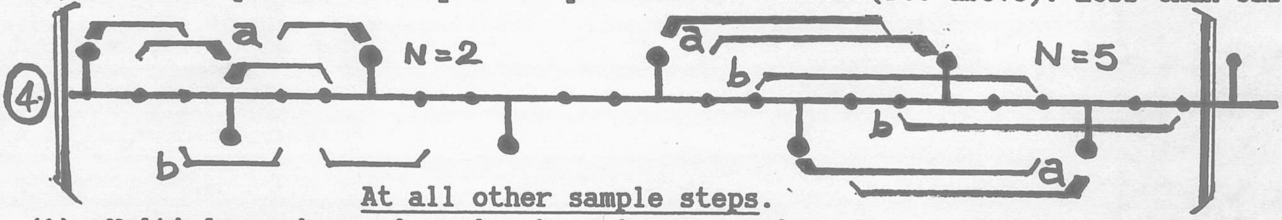
At half the wavelength of the sound (3 samples).

(1) Multiply each sample value by value 3 samples ahead, within the window.
  (a) All these products are [+ve value] X [-ve value] = [-ve value].
  (b) All these products are zero X zero = zero.
(2) Add all products. Total must be a -ve value, & hence less than case 1.



At twice the wavelength of the sound (12 samples).

(1) Multiply each sample value by value 12 samples ahead, within the window.
  (a) All these products are [+ve value] X [+ve value] = [+ve value].
  (b) All these products are [-ve value] X [-ve value] = [+ve value].
  (c) All other products are zero X zero = zero.
(2) Add all products. Only 4 +ve products in window (see above). Less than case 1.

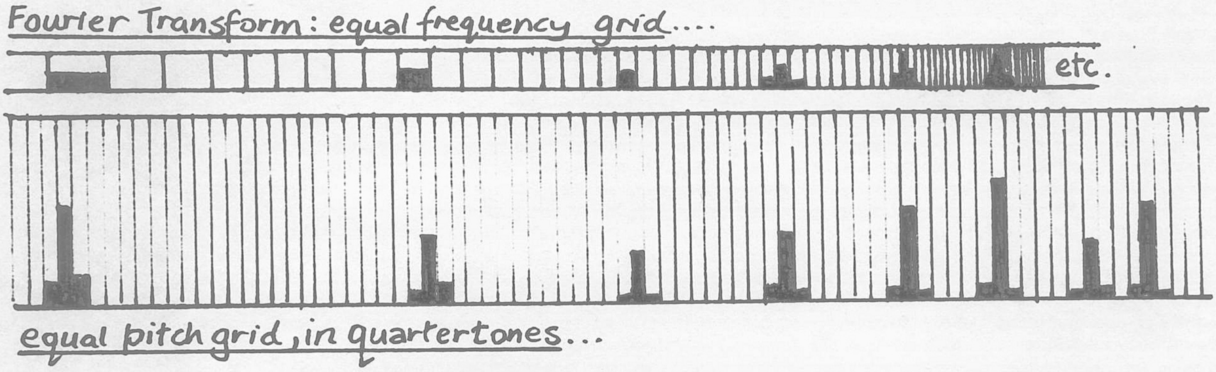


At all other sample steps.

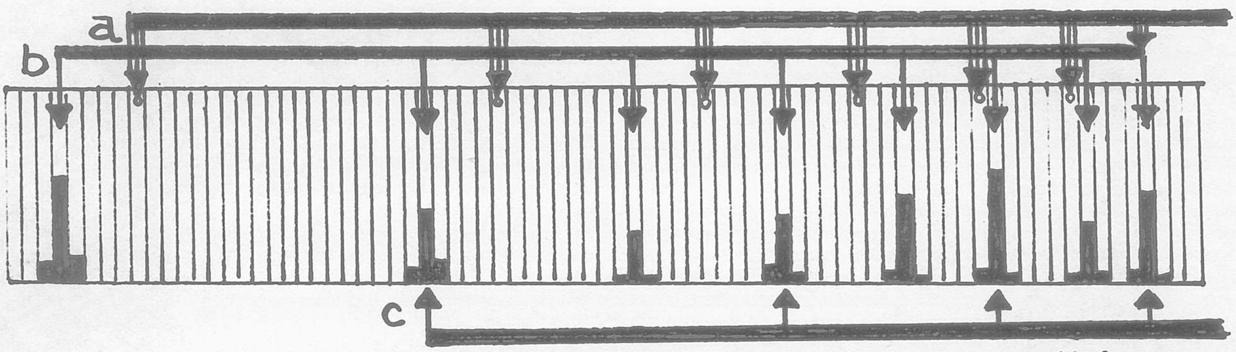
(1) Multiply each sample value by value N samples ahead, within the window.
  (a) All these products are [+ve or -ve value] X zero = zero
  (b) All these products are zero X zero = zero.
(2) Add all products. Sum is zero, and therefore less than case 1.

Hence sum has maximum value when step is wavelength of sound (case 1). We can thus track wavelength, & hence frequency. With more complex waveforms, transient subpeaks tend to cancel each other in the sum, over a large window, but as pitch may be changing, window can't be too big. Wrong-by-an-octave errors are a problem.

PITCH—TRACKING BY PARTIAL ANALYSIS



For each time-window
transfer spectral data from the equal frequency grid of the Fourier Transform,
to an equal-pitch grid, spaced in quarter-tones.



We cannot merely take the lowest peak in the spectrum as the pitch.
For example, there may be no pitch present in the particular window
we are considering, so that the position of the minimum spectral peak
has no pitch-tracking significance.

To find a pitch effectively we must proceed as follows.

For each quarter-tone channel on the pitch-grid:

(1) Find the value in that channel.
(2) Find values in those channels which fall on the harmonic-series template above that channel (a,b,c).
(3) Form a weighted sum of these values.

Most sums of this kind will miss all or most partials in spectrum
(e.g. 'a' in diagram)

Templates which fall on the partials (e.g. 'b' & 'c' in diagram)
will yield significant values for the sum.

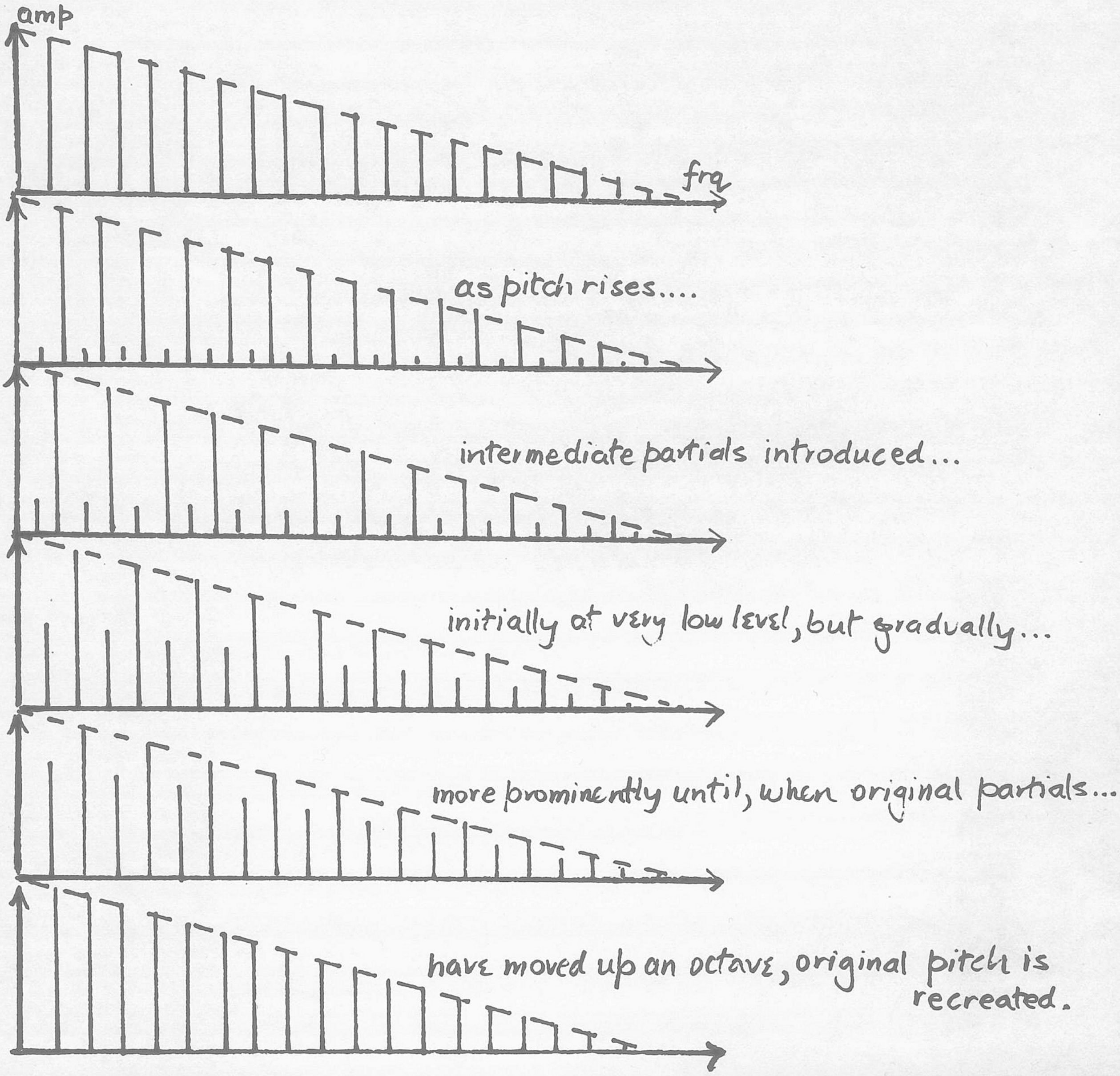
The sums given by 'b' & 'c' can only be differentiated if appropriate
weightings are given to the contributions of each partial in
the (suspected) spectrum.

Weightings for specific instruments (or other sound sources)
may be well known.

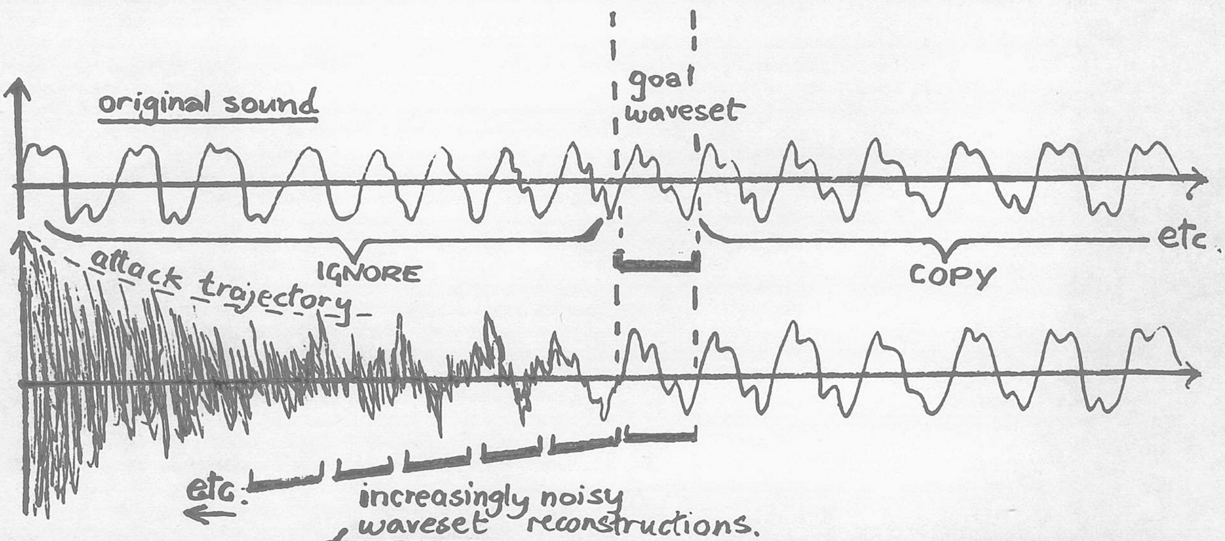
Difficulties arise with sources, like the voice,
which can change their spectral contour from window to window.

SHEPARD TONES

In this example, a tone of rising pitch always stays in the same tessitura

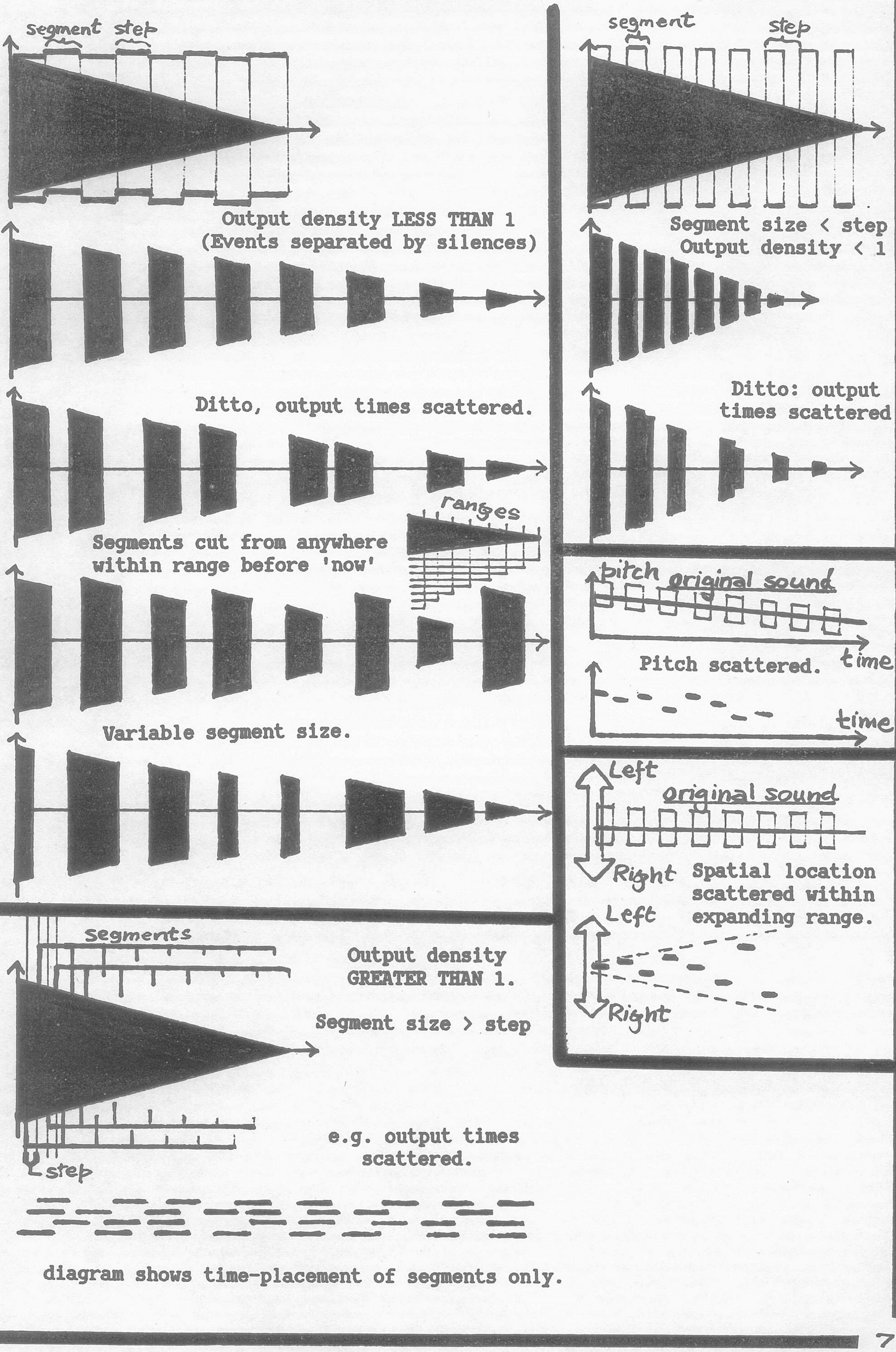


SOUND PLUCKING



GRANULAR RECONSTRUCTION

Step is (average) distance between segments in source sound.
Output density = 1 joins segments end to end as in Brassage.



Audible Design : Sound Examples Index

Chapter 1

1.0 Progressive waveset time-stretching of a noise-based source. (***Tongues of Fire***).

1.1 Examples of Brassage with different segment sizes …
(a) Source sound.
(b) Segment size less than wavelength, hence wavecycle extended & pitch falls.
(c) Segment size larger, but in grain time-frame. Source time-stretched.
(d) Segment size larger than grain time-frame. We hear echo effects.
(e) Segment size larger : begin to produce collage of source constituents.

1.2 (a) Short event followed by large time-stretch of that event.
(b) Vocal source followed by successive time-contractions.
Details of source lost and it shrinks to a grain.

1.3 Trill (articulated in speed and loudness) as a musical motif.

1.4 (a) Vocal source.
(b) Ditto, time-stretched.
(c) Sound derived from (b) by sequence of transformation processes.

Chapter 2

2.1 Sound of voice with odd harmonics on one loudspeaker, and even harmonics on the other. We hear a single source in mono.

2.2 As 2.1, but different vibrato added to the 2 sets of harmonics. Voice splits into 2 voices, one an octave higher than the other.

2.3 (a) Band of pitched events not confined on a HArmonic field.
(b) Band of pitched events, over same range, confined on HArmonic field.
(c) Band of pitched events approximating to a HArmonic field.
(d) Same band confined strictly to the HArmonic field.

2.4 Band of pitched events. Pitch-bandwidth of band gradually narrows until we perceive it as a single (animated) pitch.

2.5 Portamenti textures, for comparison….
(a) Start-weighted portamenti rising FROM notes of HArmonic field.
(b) End-weighted portamenti rising TO notes of HArmonic field.
(c) Rising portamenti with no start or end weighting.

2.6 3 examples, as in 2.5, using falling portamenti.

2.7 Rising then falling arch of sound, itself constructed from distinct portamenti, in the manner of Xenakis.

2.8 Pitch of speaking voices bring tracked by synthetic instruments. From ***Odd Evening*** by Paul de Marinis, from CD ***Music as a Second Language*** (Lovely Music LCD 3011) with permission of composer and Lovely Music.

2.9 (a) Vocal source.
(b) Double speed (octave up) using ‘tape-speed’ transposition.
(c) Half speed (octave down) using ‘tape-speed’ transposition.

2.10 Same vocal sound time-varyingly transposed using ‘tape-speed’ transposition.

2.11 (a) Steady vocal tone, and its (octave upward) waveset-transposition.
(b) Articulated vocal source, its octave-upward waveset-transposition and its octave-downward waveset-transposition.

2.12 (a) Vocal source.
(b) Transposed up octave, then down octave, using brassage/harmoniser.

2.13 (a) Vocal source.
(b) Transposed up octave, then down octave, in the frequency domain.

2.14 Excerpt ***VOX 2*** (4 voices and tape). Tenor sings through harmoniser, which parallels voice at tritone below actual sung pitch. Both pitches heard.

2.15 (a) Vocal source.
(b) The same source in which the spectrum is split into two bands and the upper band transposed upwards, suggesting two separate pitches.

2.16 Same source. Part of spectrum now slides towards its new transposition.

2.17 (a) Vocal source.
(b) Same source with 'chorusing'.

2.18 Sound rising in Hpitch yet falling in tessitura.

2.19 Reinforced harmonics produced entirely vocally.

2.20 Vocal 'ko→u' to bell (***VOX 5***).
Vocal source spectrally-stretched by greater and greater amounts. Note how pitch of sound tends to rise, even though the fundamental does not move.

2.21 (a) Vocal noise source.
(b) Same source waveset-transposed upwards, then downwards.

2.22 (a) Unvoiced 'a', 'e', 'i' , 'o'.
(b) The same transposed down 2 octaves. Note how aspects of the 'e' are retained, but the other vowel formants are altered.

2.23 Vocal noise-source 2.21, varispeeded to give pitch-band gliding noise.

2.24 Various filterings of a noise-band with progressively higher Q. so the noise moves towards pitchedness.

2.25 (a) Noise band.
(b) The noise band passed through a stack of filters.
(c) The noise band passed through the same filter-stack, with filter Q increasing from 'zero' to full value. Loudness contour also added.

2.26 (a) Vocal source.
(b) Same source with delayed signals added, firstly with audible delay time, but through the sequence the delay time becomes shorter and shorter, and the result manifests itself as pitchness in the sound.

2.27 Water-like sounds (in fact shredded voices) mixed with delayed copies of itself such that the copies resynchronise at the end of the example.
Result is rising portamento as sound concludes. (***Tongues of Fire***).

Chapter 3

3.1 ‘ko→u’ to bell. (***VOX 5***). Vocal syllable ‘ko→u’ spectrally stretched by increasing amounts. Fall in pitch at item 10 due to stretch parameters chosen. Fall to lower octave at item 17 probably due to combination of partials implying such a lower pitch. Sound also slightly (and increasingly) time-stretched on each repeat.

3.2 Examples of spectrally-stretching a vocal sound: each item (except first) has a total stretch of 2, but with a different type of stretching.

3.3 Slightly spectrally-stretched sound.

3.4 Inharmonic sounds with percussive loudness trajectories, (***Tongues of Fire***).

3.5 (a) Sung note. (b) Same sound with inharmonic-to-harmonic vibrato.

3.6 (a) Voiced speech. (b) Unvoiced (whispered) speech,

3.7 (a) The sea. (b) The sea vocoded.

3.8 (a) Source sound.
(b) Same sound but, after a moment, freeze the spectral amplitudes.
(c) The same, but freeze the spectral frequencies, instead of the amplitudes.

3.9 Noise portamenti with inharmonic spectral coloration produced by filtering.
The filtering is gradually being removed from one layer to reveal the dense textural source of the noise-band. (***Tongues if Fire***).

3.10 Examples of complex noise sources :
(a) Radio interference.
(b) Twigs breaking.
(c) Small stones on tiles.
(d) Vocal ‘gh’.
(e) Vocal texture produced by brassage.
(f) Time-contracted speech-stream.
(g) Water.

3.11 (a) Water sound. (b) The same sound spectrally-traced.

3.12 (a) 5 individual sounds.
(b) 5 dense textures made from each of these sounds,

3.13 (a) A sequence of source sounds.
(b) Same sequence spectrally enhanced :
Transposed versions superimposed on originals in same time-frame.

3.14 (a) Source sound.
(b) Same sound spectrally enhanced with stereo-reversed transposed copies, in the same time-frame.

3.15 (a) Source sound.
(b) Random perturbation of *amplitudes* of spectral components (spectral shaking).

3.16 (a) Random perturbation of *frequencies* of spectral components.
(b) Ditto, greater random scatter.

3.17 (a) Complex source sound.
(b) Same sound through narrow bandpass filter, Articulation retained.

3.18 (a) Source sound.
(b) Narrow filter focuses source on a specified pitch.
(c) Filterbank focuses source onto HArmonic field. Filters are ‘brushed’ like an aeolian harp.

3.19 Single vocal sound. Even harmonics given different vibrato to odd harmonics and sound image splits in two (we hear two different voices).

3.20 (a) Source.
(b) Source with spectral arpeggiation (3 versions).

3.21 (a) Source sound (time-stretched voice).
(b) The same sound subjected to increasingly sparse spectral tracing.

3.22 Successive time-stretches of a sound revealing its internal spectral qualities.

3.23 Source sound progressively waveset time-stretched, (***Tongues of Fire***).

3.24 Freeze both amplitude AND frequency data briefly after start of source.

3.25 (a) Source sound (crackling).
(b) The same sound after Spectral Blurring (scraping).

3.26 Three examples of Spectral Shuffling.

3.27 (a) Struck “wood”-like sources.
(b) Wavesets replaced by square waves.
(c) Interpolation between the 2 using In-Betweening.
(d) (e) (f) Same sequence, this time using sine wave substitution.

3.28 (a) Source sound. (b) Same sound with waveset inversion.

3.29 (a) A set of sounds with waveset averaging over increasing number of wavesets.
(b) Set of sounds with waveset omission. Proportion of omitted wavesets increases over the set.

3.30 Various types of destructive distortion on source suggesting bouncing object. (***Tongues of Fire***).

Chapter 4

4.1 (a) Time-stretched vocal sound.
(b) Same sound with ‘struck’ trajectories imposed.

4.2 Transition from ‘wood’-lime to drumlike, using destructive distortion and in-betweening. (***Tongues of Fire***).

4.3 Sound with ‘rubbed’ or ‘coaxed’ trajectory.

4.4 ‘Scraped’ or ‘forced’ trajectory.

4.5 Sequence of trajectories moving from struck-inelastic-object to Struck-resonating-object, through ‘coaxed’ to ‘singing’ trajectory of bowl-gongs.

4.6 (a) Source sound.
(b) Onset enhanced with octave-upward transpositions of source.
(c) Plus octave-downward transpositions.

4.7 Metallic attack(s) generated by octave stacking and imposing struck onset trajectory onto source. Source revealed to be fluid pitched material (in fact, spectrally-traced voices, further revealed as music evolves). (***Tongues of Fire***).

4.8 (a) Vocal source.
(b) Same source with plucked onset.

4.9 (a) Source sound.
(b) Same sound spectrally time- stretched so that onset is smeared out.
(c) Sound (a) used as the element of a dense texture.

4.10 Vocally initiated event, zero-cut into a different sound, forming a perceptually smooth transition ??

4.11 (a) Sound made by mixing 3 vocal sounds with onsets precisely aligned.
(b) Sequence made from similar mixed sounds but, as sequence proceeds, vocal source are gradually spectrally stretched. (***VOX 5***).

Chapter 5

5.1 Attack-dispersal continuation.

5.2 Accumulation.

5.3 Sequence of accumulation-dispersals approaching closer and closer to original onset of the sound, eventually leading to a new event,

5.4 Examples of undulating continuation…..
(a) Voice with vibrato.
(b) Voice with diaphragm flutter.
(c) Voice with ululation.
(d) Voice with head-shake flutter.
(e) A wobble-board.

5.5 (a) Synthetic tone.
(b) Same tone with vibrato with a ‘natural’ evolution;
(c) Same tome with a completely regular vibrato.

5.6 Examples of time-varying vibrato in vocal music from Japan and the USA.
5.7 (a) Power-drill. (b) Power-drill with vocal-like vibrato.
5.8 (a) Source 1. (b) Source 1 with extremely wide vibrato at end.
(c) Source 2. (d) Source 2 with extremely deep tremolo at end.
(***Tongues of Fire***).

5.9 Vibrato slowing to pitch-glide.
5.10 Accelerating tremolo, on which a second (slower) tremolo is imposed.
5.11 Hand-yodelling, ‘iuiuiu’ and ‘iaiaia’ vocal articulations.
5.12 Very rapid spatial movement of short sounds.

5.13 (a) Source 1. (b) Expansion into phrase time-frame, by time-stretching.
(c) Source 2. (d) Progressive contractions of the source to a grain.

5.14 (a) Source sound.
(b) Source sound with reverb.
(c) Source sound filtered.
(d) Filtered sound with reverb.

5.15 Integration of initator-reverberator as a single sound-event.
(a) Source sound.
(b) Source sound with reverb.
(c) Source (b) with vibrato added at end.
(d) Source (c) with new loudness trajectory.
(e) Sound(d) with variable ‘tape-speed’ transposition.
(f) Source sound filtered.
(g) Sound (f) with reverb.
(h) Sound (g) with vibrato added at end.
(i) Sound (h) with variable ‘tape-speed’ transposition.

5.16 (a) Sung source. (b) Sung source with loop,
(c) Spoken source. (d) Spoken source with loop.

5.17 (a) Source.
(b) Zigzagging over short segments, and advancing through source.
(c) Source (a) time-extended by granular reconstruction.
(d) Source (c) extended by zigzagging over long segments.

5.18 (***Tongues of Fire***). Foreground vocal sound progressively time-stretched by brassage, introducing granular quality to it. Grains increasingly separated by corrugation and made to decelerate. (NB vibrato & tremolo articulation added to some of repeats).

5.19 (a) Source A.
(b) Source A time-stretched by brassage, using grain time-frame segments.
(c) Slightly larger segments : rapid-pulsed echoes within sound.
(d) Even larger segments : clear and rhythmic echoes within sound.
(e) Segments of variable length : collage of elements of source.

5.20 (a) Source A brassage : grain-timeframe segments of randomly varied size.
(b) Similar, using segments above grain time-frame size.
(c) Source B.
(d) Source B brassaged with segments above grain time-frame size.
(e) Similar : segments selected from small time-range before 'now'.
(f) Similar : segments selected from large time-range before 'now'.

5.21 (a) Melodic source.
(b) Largeish-segment braqssage over entire time-range, develops melodic line into a harmonic field.
5.22 (a) Source.
(b) Using similar brassafge procedure to smear qualities of the onset over the entire sound. (***Tongues of Fire***).

5.23 (a) Monotone speech source C.
(b) Brassage : random pitch variation of grain time-frame segments over small range.
(c) Ditto, with segment size above grain time-frame.

5.24 Brassage same source …. With cyclical pitch variation over small pitch band.
5.25 … with widening pitchband of random pitchshifts of (grain time-frame) segments.
5.26 As last, but moving from mono into full stereo spread brassage.

Granular reconstructions of Source C …..
5.27 …. Using long overlapped segments diverging in pitch, and spatially,
5.28 (a) …. Using much overlap of grain time-frame segments (overlap >> 1).
(b) …. Using grain time-frame segments separated in time. (overlap << 1).
5.29 Using large time-range for segment choice, large pitch-range for segment transposition, long time-stretch, & segments slightly separated (overlap slightly less than 1).

Chapter 6

6.1 (a) Natural grain-stream sound.
(b) Artificially regular grain-stream made from single grain of (a).

6.2 (a) Natural grain-stream sound.
(b) …. add attack-dispersal trajectory, and exaggeratedly wide vibrato at end.

6.3 Examples of naturally-produced grain-streams ….
(a) Rungs of metal basket.
(b) Voiced lip-flabber.
(c) Ribs of plastic washboard.
(d) Voiced rolled-rr.
(e) Gliss keyboard without depressing keys.
(f) Flutter-tongued flute.
(g) Strum stair-banister uprights.
(h) Vocal diaphragm flutter.
(i) Unvoiced lip-flabber.
(j) Vibraphone glissando.

6.4 (a) Granulatio tied to wavecycle. (b) Simple granulation (same source).

6.5 (a) Source. (b) Looped (exact repetitions), (c) Iterated.

6.6 (a) Source.
(b) Iterated via mixing score : greater grain-placement randomisation.
(c) Iterated similarly, with deceleration of grain-placement.
(d) As (c) but with greater randomisation of grain-placement.

6.7 Examples of grain-streams assembled from sequences of grain sounds which are progressively modified, using spectral transformation techniques.

6.8 (a) Continuous portamento vocal source.
(b) Granulate and time-extend by brassage.

6.9 (a) Source : then 9 progressive ‘tape-speed’ speed-ups (pitch emerges).
(b) Source : then 4 progressive speed-ups using spectral time-shrinking.

6.10 Source ; then progressive sequence of time-contractions which reduce the grain-separation without shortening the grains themselves.

6.11 (a) Granulated source.
(b) Reverb.
(c) Much reverb (granularity lost).

6.12 (a) Source 2 : Granulated stream (granular reconstruction of continuous vocal source).
(b) , (c) & (d) Similar sound but with greater and greater grain overlap.

6.13 Source 2 with ritenuto of grain time-placements.

6.14 (a) Source.
(b) Part of source with increasing time-stretching.

6.15 (a) Source.
(b) Source with grains transposed semi-randomly.

6.16 Granular time-warping …
(a) Source.
(b) Slowed, with accelerando.
(c) Slightly randomised.
(d) Regular rhythmed.

6.17 (a) Grain-stream.
(b) Reversing grain-sequence.
(c) Reversing sound.

6.18 (a) Source grain-stream.
(b) 3 ornamentations of the grain-stream, made by replacing each grain by a pitch/rhythm motif produced from transpositions of the grains.

Chapter 7

7.1 Three examples of sequences.
7.2 Portamento, when cut and rearranged, becomes a sequence.

7.3 (a) Source (text). (b) Text completely reordered by shredding.
(c) Text reordered by random cutting such that we hear repetitions of particular elements with the resulting stream
7.4 Same text cut into segments, each treated with different filterings, the respliced back together again.

7.5 (a) Source. (b) Brassage, very small search-range.
(c) Larger search-range.
7.6 (a) Source. (b) Cyclically pitch-varying brassage.

7.7 Brassage with spatialisation.
7.8 (a) Sequence contracted to granular event.
(b) Sequence expanded to phrase timeframe.
7.9 (a) The same sentence in a texture in which it is still recognisable.
(b) The same sentence in a very dense texture.

7.10 Excerpt from ***Tongues of Fire*** : principal layer is a string of vocal utterances, shredded progressively by greater amounts until the side becomes waterlike.

7.11 (a) Sequence over clear (HArmonic) filed. (b) Sequence without field.
7.12 Sound poetry focusing on small set of syllables, and order-sequencing.

7.13 (a) Melody ….dispersed in texture, retaining (HArmonic) field properties of melody
(b) Percussive sequence ….. repeated *but* field properties evolve through time.

7.14 (a) Field evolves by pitched elements becoming more similar.
(b) Field evolves by elements themselves evolving.

7.15 Blurring the boundaries between the elements of a sequence ….
(a) Two examples of reverberation.
(b) Extensive granular reconstruction.
(c) Two examples of Spectral Blurring.

7.16 Focus in on pitch-set of arbitrary chunk of spoken text, through repetition.
7.17 (No Sound Example!) Loop a short (2 secs) piece of text. Listen for 10 minutes.
7.18 English 8-bell change-ringing pattern played on electronic Kalimba.

7.19 From ***VOX 3*** : 4 singers in 4 different tempi chosen such that certain stressed beats coincide between the 4 parts. (performed by *Electric Phoenix*).

7.20 Pattern in 3, changes to 5 (loudness stress), and then to 4 (sustain-note stress).

7.21 (a) Most probably 3:4 meter. (b) Most probably 6:8 meter.
(c) Ambiguous : either 3:4 or 6:8 (d) Undecidable meter.

Chapter 8

8.1 (a) Measuredly perceived set of events. (b) Texturally perceived set of events.

8.2 Random-set of pitches, with 1 note-pair repetition. Can you spot it ?
8.3 (a) Sequence of random pitches. (b) Part of this sequence, looped.
8.4 Dense, pitch-and-time-random texture of events.

8.5 (a) Texture which gets denser.
(b) Source sound for texture (c).
(c) Texture which spreads in pitch-band width.
(d) Texture in which constituents get shorter.
(e) Random-time texture becoming metrically regular.

8.6 Texture, with falling and widening pitch-band, generated from spectral variants of a single staccato sound.

8.7 (a) Grain-stream.
(b) Ditto randomised in order and time-placement.
(c) Texture created from similar such elements.

8.8 (a) Descending continuum. (b) Texture made from it by granular reconstruction.

8.9 (a) Rapid stream of vocal consonants, mono.
(b) The same, with alternate events placed to left and right.
(c) …. And with individual events placed randomly across the stereo stage.

8.10 (a) Random events, time-quantised over a 1/30th second pulse.
(b) The same random sequence, but with no quantisation.

8.11 Transforming textural layer from ***Tongues of Fire*** (see text).

8.12 (a) Change of HArmonic field by substituting elements.
(b) Change of HArmonic field by gradually retuning elements.
(c) Destruction of HArmonic field by pitch-spreading.
(d) Destruction of HArmonic field by pitch-gliding.

8.13 Texture of cyclically varying density.
8.14 Density fluctuation in grainy texture, itself becomes grain-like.
8.15 Vocal ‘LIS’ → birdsong, from ***Red Bird***.
8.16 Vocal ‘SSS’ → crowd, from ***Red Bird***.

8.17 Voices texture which whites out: the resulting noise-bands pitchband-slide as they are gradually (increasing Q) filtered to create pitchness. Pitchness recedes, and texture finally thins to reveal massed voices again. (***Tongues of Fire***).

8.18 (a) Motif with specific internal pitch relations and rhythm.
(b) Same motif distributed in texture where motif begins at random points in the time and pitch continua.

Chapter 9

9.1 (a) Texture over modal minor scale.
(b) Texture over a semitone grid.
(c) Texture over quartertone grid.
(d) Texture over the continuum.

9.2 (a) Rigorous mechanical rhythm ….
followed by increasing degrees of randomisation of event onset-times producing..
(b) More natural fell.
(c) Sloppy rhythm.
(d) Bad rhythm.
(e) Complete loss of rhythm.

9.3 Event-set moving towards rigorous rhythm.

9.4 (a) Strict rhythmic pattern in ratio 3:1.
(b) Rhythmic pattern varying randomly between 2:1 and 3:1.

9.5 Strict rhythmic pattern in ratio 37:26

9.6 (a) Tempi in ratio 2 to 3 (see Diagram 5) (synthetic reconstruction from ***Vox 3***, variation 8).
(b) Ditto, with common *larger*-time-frame pulse emphasized.

9.7 (a) Tempi in ratio 2 to 3 (see Diagram 5) where grouping of beats in the 2 streams confounds sense of a common pulse. (synthetic reconstruction from ***Vox 3***, variation 8).
(b) Ditto, with common *smaller*-time-frame pulse emphasized.

9.8 In quick succession …..
(a) 2 parts in the tempo ratio of 3 to 5. (b) Same, with common smaller pulse.
(c) 2 parts in the tempo ratio of 7 to 10. (d) Same, with common smaller pulse.
(f) 2 parts in the tempo ratio of 7 to 10. (f) Same, with common smaller pulse.
Note that eventually the common pulse itself becomes pitch-like as it repeats so quickly that we can no longer perceive it as a pulse.

9.9 (a) 3-part pattern in conflicting meters, as in Diagram 13.
Common larger pulse enters halfway through.
(b) Similar 3-part pattern but with de-emphasized common beats, (Diagram 14).
Common larger pulse enters halfway through.

Chapter 10

10.1 (a) A sequence, a texture and a grain-stream.
(b) These 3 combined to make a musical event, with random-peaks loudness trajectory on sequence, crescendo trajectory on grain-stream and attack-dispersal loudness trajectory on texture.

10.2 (a) Two sounds abutted.
(b) The same, but with anacrusial trajectory leading to attack-dispersal trajectory.

10.3 (a) Sustained tone.
(b) Same tine with attack-dispersal trajectory.
(c) A sequence given an attack-dispersal loudness trajectory (plus reverberation).
(d) The original sequence.

10.4 (a) Sound with highly characteristic loudness trajectory.
(b) 2nd sound.
(c) Loudness trajectory of (a) superimposed on (b).

10.5 (a) Source sound.
(b) Brassage time-stretched, introducing graininess.
(c) Greater brassage time-stretch, with graininess.
(d) Even greater brassage time-stretch, with graininess (start portion only).
(e) The previous sound, corrugated.
(f) The previous sound with grains decelerating.

10.6 (a) Source sound.
Same shape loudness-trajectory imposed over wavesets grouped in sets of …
(b) 1 (c) 4 (d) 8 (e) 16 (f) 32 (g) 64.

10.7 (a) Source sound.
(b) Loudest sounds in source trigger reverberation on themselves.

10.8 (a) Struck-‘wood’-like sound
(b) Drum-like sound produced by waveset distortion of sound (a).
(c) Transition from sound (a) to sound (b) by (waveset synchronised) mixing of sounds (a) and (b) in different proportions (In-Betweening).

10.9 (a) 3 complex sources mixed in mono.
(b) The same 3 sources mixed in stereo.

10.10 (a) Dense screeching with a low bubbling sound mixed in the middle.
(b) The same using inverted envelope of the bubbling on the screeching sound.

10.11 Recording of struck metal object, where 6 zoom-ins by the microphone reinforce mid and high frequency partials in the spectrum.

10.12 Versions of the same sound with more and more time shaved off the onset of the sound, so the attack softens.

Chapter 11

11.1 (a) Source.
(b) Source ‘tape-speed’ transposed down 2 octaves. Durarion grows (times 4), formants shift (from long ‘e’ as in’hAIr’ to long ‘a’ as in ‘cAR’), and the granularity of the sound on a small time-scale is revealed.

11.2 (a) Complex vocal source.
(b) The same, ‘tape-speed’ transposed down 2 octaves (duration grows times 4).

11.3 (a) Extract from a dense texture of voice-derived material.
(b) The same texture accelerated rapidly.
(c) Extract from a texture of struck metallic sounds.
(d) The same texture accelerated rapidly.
These accelerated examples create similar noise-portamenti.

11.4 (a) Source A.
(b) Harmoniser-type time-stretch times 4 (tunnel-resonance artefact).
(c) Spectral time-stretch for comparison.

11.5 Source A time-stretched 16 times …
(a) Harmoniser-type time-stretch. (Granular and tunnel-resonance artefacts).
(b) Spectral time-stretch. (Metallic ringing artefacts).
Here, the spectral-stretch artefacts seem to me to be less coherent with the gritty vocal source than the harmoniser artefacts.
(c) Source 2.
(d) Harmoniser-type time stretch by 4, of source 2.
(e) Spectral time stretch by 4, of source 2.
(f) Harmoniser-type time stretch by 16, of source 2.
(g) Spectral time stretch by 16, of source 2.
Here, the spectral time-stretch seems to retain the body of the sound better.
In later examples (Hear 11.13 & 11.14). I’ve removed onset portions of stretched sounds, as these are radically transformed by the time-stretch, in different ways.

11.6 (a) Source.
(b) Source time-stretched by 2 by repetitions of single wavesets,
(c) Source time-stretched by 2 by repetitions of wavesets-grouped-in-16s,

11.7 (a) Source.
(b) Wavesets, grouped in 16s, repeated times 2, 3, 4, 5, 6, and 8.

11.8 (a) Source.
(b) Times 4 time-stretch, wavesets grouped in 16s.
(c) Times 4 time-stretch, wavesets grouped in 4s.
(d) Times 4 time-stretch, single wavesets.
(e) Times 6 time-stretch, single wavesets.
(f) Times 6 time-stretch, single wavesets, with interpolation.
(g) Times 16 time-stretch, single wavesets.
(h) Times 16 time-stretch, single wavesets, with interpolation.

11.9 (a) Source.
(b) Spectrally time-stretch by 8, by synthesizing longer amounts of sound from each analysis window.

11.10 1/20th of a second of same source, time-stretched by 64 by interpolating windows.

11.11 (a) Noisy and spectrally changing source.
(b) Spectral time-stretch by 20 (interpolating windows).

11.12 3 time-varying spectral time-stretch versions of the same source : maximal time-stretch is at a different point in the source, in each case.

11.13 (a) Bell-like events made from voices (***VOX 5***).
(b) One of these events with its onset much time-stretched (***VOX 5***).

11.14 (a) Source sound.
(b) Source sound spectrally time-stretched.
(c) Source sound time-variably spectrally time-stretched. In this case, stretch factor rises from 1 (no stretch) to full value only after onset portion of sound.

11.15 (a) Source sound. (b) Extreme extension of attack moment.
11.17 (a) Grain-stream source. (b) Decelerated by grain-separation.
11.18 (a) Separate synthetic tones. (b) Granular time-shrunk to grain-stream.

11.19 (a) Grain-stream source.
(b) Source time-stretched by 2 by grain duplication.
(c) Source time-stretched by 4. Duplication now heard as pitch-stepping of sound.

11.20 (a) Source. (b) Source spectrally time-shrunk, then expanded.
(c) Sound (b) reprocessed similarly, using a different shrink-ratio.

11.21 (a) Grain-stream source.
(b) Successive spectral time-contractions (times 2, 4, 8, 16).
(c) Attack detail lost in the shrinking process.

11.22 (a) Grain-stream source.
(b) Successive granular time-contractions (times 2, 4, 8, 16).
Contracted sounds retain a grittiness from sharp spectral peaks in the attack of the original grains.

11.23 (a) Vocal source.
(b) Waveset time-contractions by 2. First the simple case, and then the case where we retain the *loudest* waveset in each group.
(c) As (b), but contract by 4.
(d) As (b), but contract by 8.
(e) As (b), but contract by 16.
(f) As (b), but contract by 32.
(g) As (b), but contract by 64.

11.24 (a) Source.
(b) Spectral time-stretch by 2, 3, 4, 5, 6, 7, 8 and 16.
At X3, downward-glissando feature in extended onset becomes evident.
At X5, fluid nature of extended onset becomes clear, then persists.

11.25 (a) Source.
(b) Original onset characteristics of source restored to the X16 time-stretched example from 11.24(b). This is done by making an onset-preserving (time-varyingly) spectrally time-stretched version of the source, and mixing it onto the start of the transformed sound.

11.26 (a) Source sound.
(b) Grand extension using spectral, tape-speed-transposition, onset-reinforcement, sound reversal etc. (***Tongues of Fire***).

11.27 (a) Source.
(b) Spectrally stretched. Time elongation without sense of rhythmic slowdown.
(c) Source (a) granulated.
(d) Gradual spectral stretching of (c). Rhythmic ritenuto and spectral transformation of sound (sounding like added reverberation).
(e) Time-stretching (d) by grain-separation. No spectral transformation.

11.28 (a) Source sound
(b) Spectrally time-stretched by 8,revealing inner morphologies.

11.29 (a) Source sound with vibrato.
(b) Part of source spectrally time-stretched to become slowly gliding.

11.30 Accelerating stream and decelerating stream converge on common pulse.

11.31 Ululation section from **VOX 5** (stereo reduction of 4-channel original).

11.32 End-synchronised delay gives upward pitch portamento. (***Tongues of Fire***).

11.33 (a) Dense Texture-A, with rising tessitura, and ebb-and-flow loudness trajectory.
(b) Spectrally time-stretched by 6. All properties extended by 6. (At this density, difficult to grasp the spectral transformation of source : but listen to 11.38).

11.34 Texture-A waveset time-stretched by 2. Manifests as a spectral transformation.

11.35 Texture-A brassage time-stretched (regular grainsize).
(a) No grain overlap (stretch by 2). Time-regular grain-pattern imposed.
(b) (Smaller grainsize), more overlap, *Small range*. Stretch by 3.
(c) As (b) but *Large range.* Loudness trajectory destroyed.

11.36 (a) Texture-A.
(b) Loudness-flat version of Texture-A, spectrally time-stretched by 6, then loudness contour add at same undulation rate as original. (Start portion only).

11.37 (a) Texture-A
(b) Ditto, internally time-stretched (manifests as reduced density).

11.38 (a) Texture-A, but generated for longer time, with time-stretched (by 6) loudness contour imposed. (Start portion only).
(b) Texture-A spectrally time-stretched by 6, with time-stretched (by 6) loudness contour imposed, for comparison. (Same start portion only).
Compare these two to hear the transforming effect of the spectral stretch.

Chapter 12

12.1 (a) Time-stretched vocal source.
(b) Source (a) with hard-edged loudness trajectories imposed, towards end.

12.2 Voice → bees → Voice (***VOX 5***).

12.3 Voices onset-synchronised to give bell-percept. (***VOX 5***).

12.4 Struck-‘wood’-like sound transforming to drum-like sound by waveset-synchronous in-betweening.

12.5 Source ’ko→u’, and spectrally-stretched bell-like result. (***VOX 5***).
(For full sequence of stretchings, listen to Sound Example 3.1).

12.6 (a) Vocal and orchestral source mixed.
(b) The same sources spectrally interleaved, analysis window by window.
(c) As (b) but taking 8 analysis windows at a time.

12.7 The same two sources intercut using a brassage process.

12.10 (a) Set of short sounds made from filtered noise, where filter Q increases through the set, so that the final items are clearly pitched.
(b) Texture made from these elements, progressively substituting later ones for earlier ones, and pitchband narrowing towards (almost) a single pitch.

12.11 (a) Vocoded sea. (b) Sea without vocoding (for comparison).

12.12 Spectral masking, In this example, made by Rob Waring, high partials in the voice interact with those in the flute, producing subtle resonances in the sound.

12.13 (a) A poor example of spectral interpolation (transition too fast).
(Compare sound example 12.2)
(b) Single window to single window interpolation. Works almost as well here, as the 2 sounds do not have complicated micro-fluctuations.

Chapter 13

13.1 Excerpt 1 from ***Tongues of Fire***,
13.2 Excerpt 2 from ***Tongues of Fire***,
13.3 Excerpt 3 from ***Tongues of Fire***,

13.4 (a) Two portamenti settling on a goal pitch.
(b) Two linear portamenti.
(c) Two portamenti moving away from a base pitch.

13.5 (a) Leading From. (b) 3 examples of Leading To. (c) Stable. (d) Unstable.