

4655
WISHART, T.
1996

Chapter 3

PYTHAGORAS, FOURIER, HELMHOLTZ: TOWARDS A PHENOMENOLOGY OF SOUND

Pythagoras, that grave and venerable personage, reprov'd all judgement of Musick which is by the eare, for he said that the intelligence and vertue thereof was verie subtile and slender, and therefore he judg'd thereof, not by hearings, but by proportionall harmonic; and he thought it sufficient to proceed as farre as to Diapason, and there to stay the knowledge of Musick. [Plutarch]

Pythagoras: stable vibrations and the *Harmony of the Spheres*

The preceding chapter sought to account for the ideology of musical praxis which sees pitch and duration as primary musical qualities, timbre as a distinct and secondary musical quality and takes instrumental streaming and the generation of music on a lattice for granted. The philosophy of the musical practice based upon establishing elementary relationships between stable vibrating systems has a very ancient and respectable pedigree. Pythagoras himself is believed to have first noted the fact that there is a simple relationship between the lengths of vibrating strings and the perceived quality of musical consonance between them. Given two strings² of the same material at constant tension then if one is stopped exactly half-way along its length it will produce a note sounding one octave above the other string (which is not stopped). The octave itself is qualitatively perceived to be the most stable or consonant of music intervals. With a length ratio of 3:2 the musical interval of a fifth is perceived which is also very stable and consonant. In general, the relative consonance of an interval is seen to be directly relatable to the simplicity of the ratio of lengths of the strings (or columns of air) which produce it (see Figure 3.1). This was not only the first important contribution to music theory but also had a significant role to play in the development of a scientific view of the world. It was the first clear demonstration that qualitative aspects of nature

¹ The author can no longer recall where he came upon this quaint translation into Elizabethan English (Ed.).

² The argument applies equally well to the lengths of columns of air.

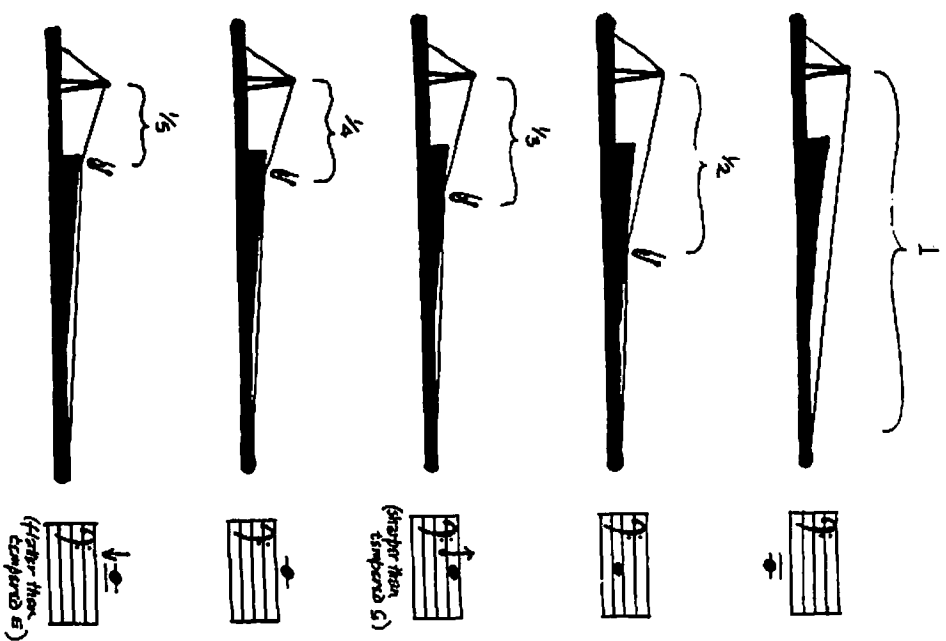


Figure 3.1 Relation between string length and 'consonantly' related tones.

could be reduced (apparently) to simple numerical relationships. This general view has been exceedingly fruitful but on occasions misleading. It led indirectly to the concept of the *Harmony of the Spheres*, one of the most persistent and misleading conceptions ever to animate the human mind: the heavenly bodies were assumed to be transported around the earth on giant spheres whose motion generated a heavenly music, governed in some way

by the Pythagorean laws of proportion. The desire to find 'celestial harmony' in the world of nature persists even into the work of the astronomer Kepler, who was obsessed by the desire to fit the (assumed) spheres of planetary motion (Figure 3.2) around the Platonic solids (Figure 3.3).

It is important to point out that, important though these elementary physical relationships are in the underpinning of various musical languages, music in general is a cultural construct. Although the primacy of the octave and the fifth is preserved in most musical cultures (though not everywhere), we would have great difficulty in explaining away all the subtle ramifications of the North Indian scale system in terms of Pythagorean interval theory. More significantly, perhaps, for the subject of this book, the Western tempered scale in fact preserves only the true octave in its structure. All other

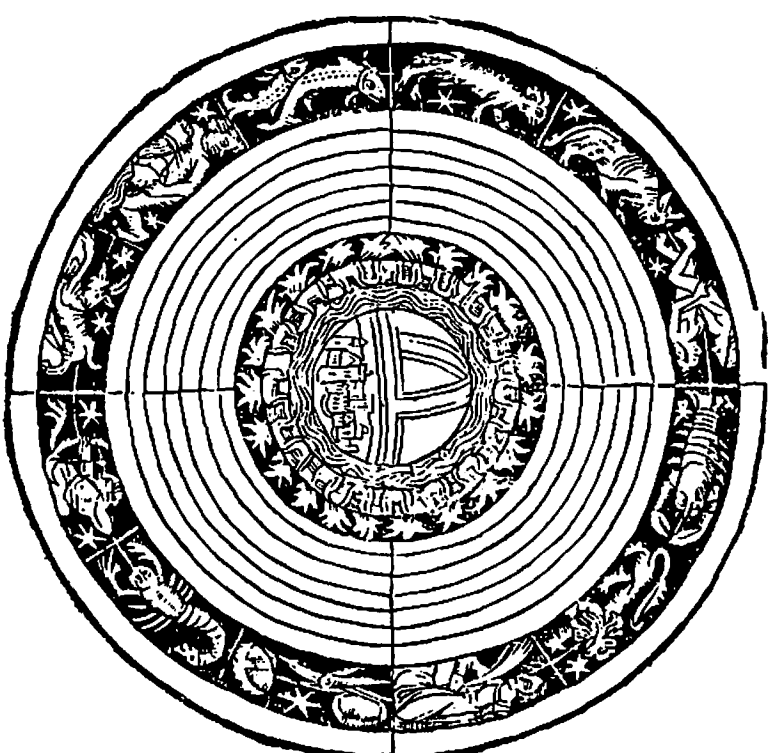


Figure 3.2 The planetary and celestial spheres.

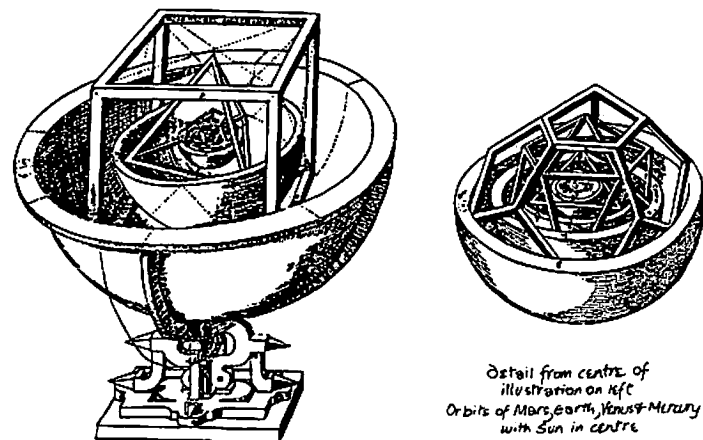


Figure 3.3 Kepler's attempt to fit the Platonic solids within the planetary orbits (assumed spherical).

apparently simple-ratio intervals (such as the fifth, major third and minor third) are mere approximations to the simple Pythagorean ratios, the actual ratios used being governed by the rationality of tempering and the twelfth root of two! Although in Western tonal music the fifth plays a central role (next to the octave), this cannot be put down merely to its Pythagorean simplicity. The well-tempered ratio $((^{12}\sqrt{2})^7 (= 1.48):1)$ is in no sense simple and it is difficult to see in what sense it approximates 'simplicity' — are we to say that it is simpler than the ratio 4:3 because it is closer to 2:3? — in that case, would not $2:\pi$ be simpler than 4:3?

One further point: Pythagoras' theory essentially establishes relationships between simple, stable vibrations. Current (and in fact ancient) musical practice is not solely concerned with such sound phenomena and the advent of computer analysis and synthesis permits us to understand and control much more complex sound phenomena.

Helmholtz, Fourier and Standard Musical Practice

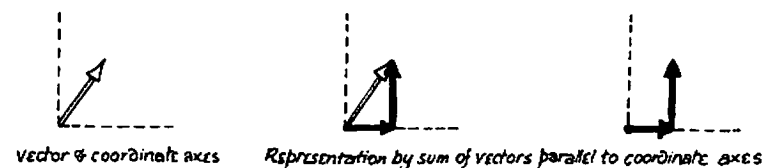
The next major breakthrough in our physical understanding of the nature of sound came with an important discovery by the mathematician Fourier. While attempting to solve various problems relating to the conduction of heat in solids, Fourier discovered that it was possible to represent an arbitrary mathematical function by a sum (possibly infinite) of simpler

functions. The simpler functions which he chose for this representation were the elementary sine and cosine functions with which we are all now familiar from work in acoustics or electronic music. It is possible to give an approximate description of the Fourier method without going into mathematical details to anyone familiar with the concept of *vector*.

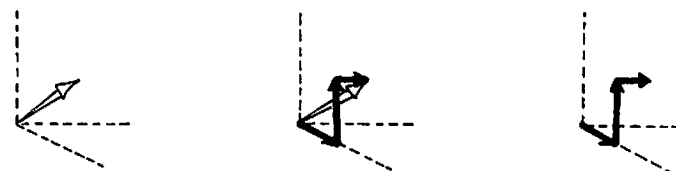
A vector may be regarded as a line of a particular length and a particular direction. In three-dimensional space we may affix to one end of this line a set of three lines at right angles to one another, a system of coordinates. The point where these lines meet is called the origin. If our vector starts at the origin no matter where its other end point is, it is always possible to reach that end point by proceeding from the origin a certain distance parallel to one axis, a certain distance parallel to the second axis and a certain distance parallel to the third axis (see Figure 3.4). These three new vectors are called the components of the original vector. Roughly speaking the fact that the three components are at right angles to one another means that they are independent of one another (one cannot express any of the components in terms of the others). The components are then said to be *orthogonal*.

If we now take an arbitrary function (see Figure 3.5), this associates with every point along the horizontal axis a point on the vertical axis. We

2-DIMENSIONS:-



3-DIMENSIONS:-



4-DIMENSIONS et al.

This cannot be physically drawn or modelled, but the mathematical representation is precisely parallel.

Figure 3.4 Representation of vectors in a coordinate system.

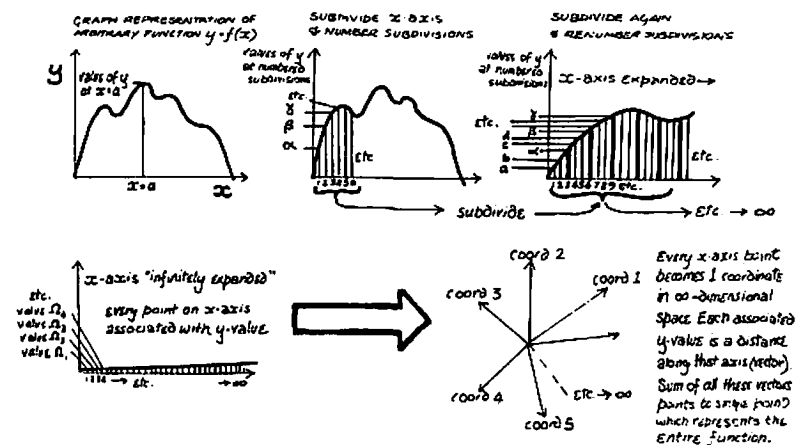


Figure 3.5 Representing an arbitrary function as a point in infinite-dimensional space.

now need to make a leap of the imagination and imagine that every infinitesimally small point along the horizontal axis corresponds to a dimension in an infinite-dimensional space and that the corresponding value on the vertical axis corresponds to a distance along that particular dimension. We can now, at least conceptually, represent the entire function by a single point in this infinite dimensional space. If now, just as in the case of the vector in three-dimensional space, we can set up a system of orthogonal coordinates and define a set of related components in the space (which in this case will turn out to be other mathematical functions) then we can make a representation of the original arbitrary function. The set of sine and cosine functions used by Fourier, can be shown to fulfil the criterion of orthogonality. Hence we have discovered a very powerful mathematical tool. It soon became clear that Fourier's method was especially applicable to the description of sound phenomena. As is now well-known, any sound phenomenon, no matter how complex, is carried by variations of pressure within the air and can be represented as a function of air-pressure against time. Such functions are, at least in principle, directly amenable to Fourier's method of analysis.

Helmholtz and others, working with what they called 'musical' tones, i.e. the sounds of conventional musical instruments, proposed a simple theory of pitch and timbre perception. A sound perceived as a single pitch was found to be made up of various sine wave components (through Fourier analysis). These bore a simple harmonic relationship to

one another. The frequency of the higher sine tones were integral multiples of the frequency of the lowest (which for the moment we will assume to be the fundamental) frequency. The pitch of a sound corresponded directly with the frequency of the fundamental, the timbre was the result of the presence (relative amplitude) or absence of the other sine tones (the partials). Before going on to criticise and comment upon this theory, we should note that it seemed to absolutely confirm the ruling musical ideology that pitch was primary and timbre secondary. Pitch could be seen as fundamentally related to frequency and timbre as merely a secondary phenomenon arising from the combination of the frequencies of the constituent sine tones. Timbre appeared to be thus almost a fused chord over a fundamental pitch. However, the fact that Helmholtz's theory appeared to confirm the ruling musical ideology should come as no surprise. It was framed within a culture which took that system of musical thinking for granted. Helmholtz confined himself to the analysis of what he arbitrarily defined to be 'musical' tones, i.e. sounds forced onto the pitch-timbre-duration lattice by preconceptions of the musical and their realisation in instrument technology. Furthermore, the assumptions firstly that timbre was a unitary phenomenon and secondly that pitch and timbre were clearly separable qualities, were taken for granted directly from preconceptions of music tradition.

Walsh functions, indeterminacy and the missing fundamental

The first question we must ask about Fourier analysis is, although it is clearly a very powerful mathematical tool, does it bear any relationship at all to our perception of sonic reality? Is the Fourier analysis of sonic events into sine tones unique, or is there any other alternative analytic breakdown? It turns out that other systems of orthogonal functions can be defined and used to represent arbitrary mathematical functions. One such system is that of *Walsh functions* illustrated in Figure 3.6. With the advent of digital technology some programs have already been developed for the synthesis of sounds using Walsh functions rather than the more usual sine tones. However, whereas the Walsh function analysis of a sound-object seems to bear no clear intuitive relationship to our aural experience, Fourier analysis relates very clearly to what we hear. It has been shown in fact that the human ear is a kind of Fourier analyser so that we may assume that up to a point the mathematics of Fourier analysis has some direct relationship with our perceptual experience.

The result of Fourier analysis is what is called a *Fourier transform*. The mathematics of the Fourier analysis convert information about the variation of amplitude with time (*time-domain information*) into information about

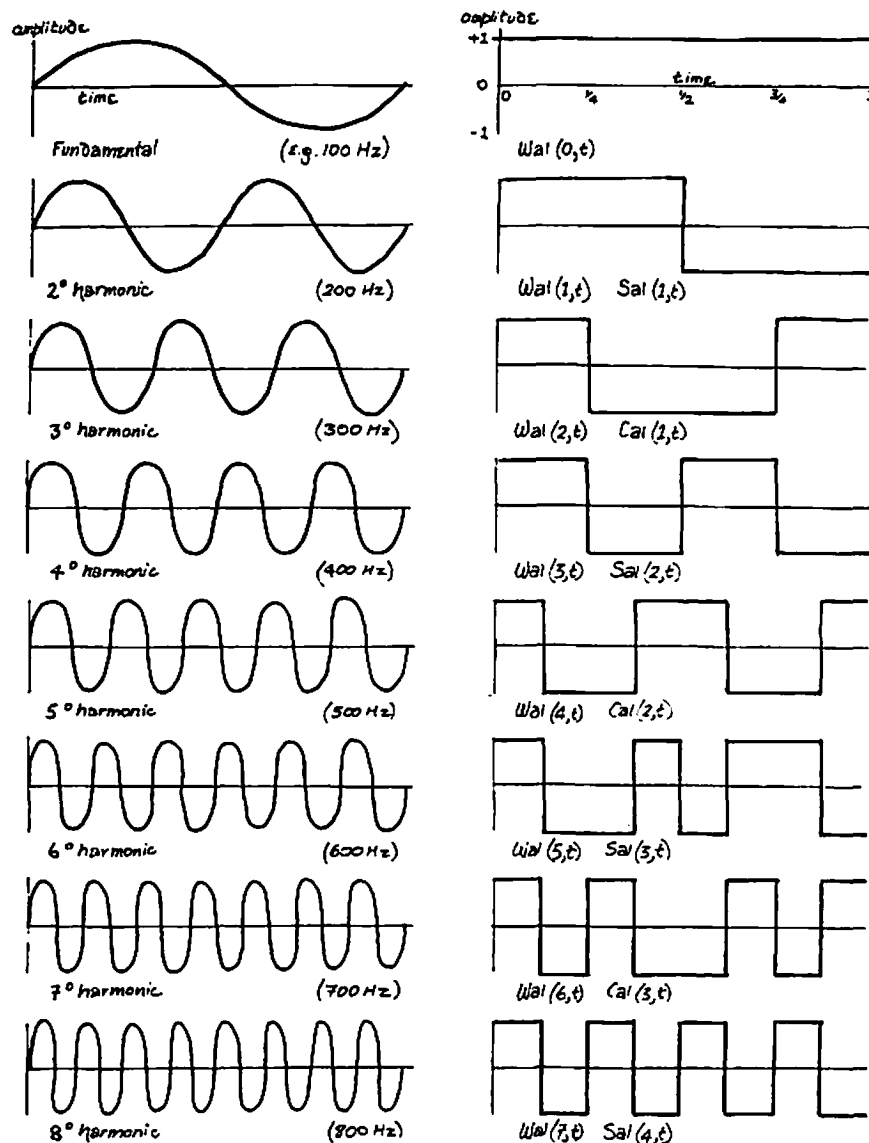


Figure 3.6 The first eight harmonics (sine waves) and the first eight Walsh functions.

the variation of amplitude with frequency (frequency-domain information). Simply put, we start off with a graph of amplitude against time and we end up with a graph of amplitude against frequency (see Figure 3.7). The *inverse Fourier transform* performs the opposite function, turning information about frequency and amplitude into information about amplitude and time.

It must be said immediately that the notion that somehow frequency (periodicity) is more physically real than spectral information is hard to justify. In a simple instrumental tone, periodicity can certainly be more easily seen from a graph of amplitude against time than can any spectral information. However, this is partly the nature of the beast being analysed and when we consider more complex musical objects (see the section on noise below) we will find that the graph of frequency against amplitude (i.e. the spectrum of the sound) is far more lucid than the amplitude against time graph (which may be totally aperiodic). In fact, to be entirely reductionist for a moment, all that really exists is the amplitude of displacement of the air or the ear-drum and its variation in time. Both periodicity and spectral information are higher-level derived entities.

A more important problem is simply that the ear is unable to function as a Fourier analyser above frequencies of around 4,000 Hz

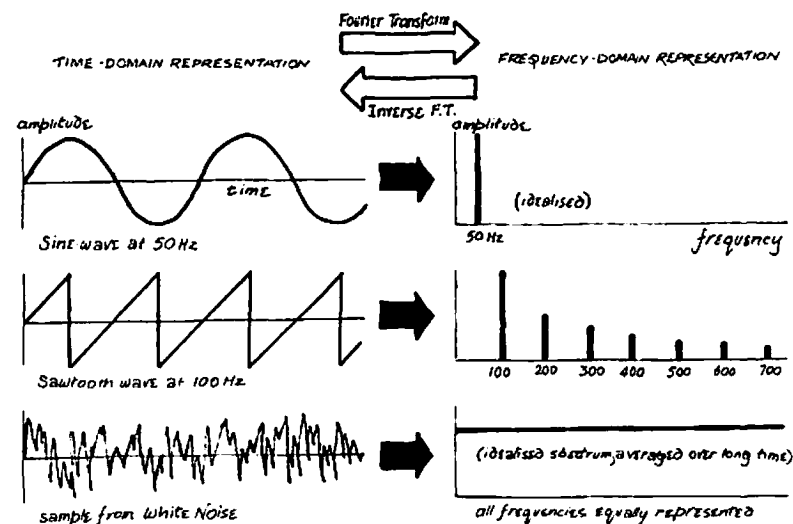


Figure 3.7 The Fourier transform and its inverse.

although frequency information above this threshold can be very important in our perception of timbre. In relation to these I quote from Schouten, Ritsma and Cardozo:

[...] there may exist one or more percepts (residues) not corresponding with any individual Fourier component. These do (however) correspond with a group of Fourier components. Such a percept (residue) has an impure, sharp timbre and, if tonal, a pitch corresponding to the periodicity of the time pattern of the unresolved spectral components.
(Schouten, Ritsma and Cardozo 1962: 1419, emphasis added)

Put simply, above 4,000 Hz we hear timbre simply as timbre!

A more serious problem with the ideologically natural view which equates sine tones with 'pure pitches' and their combination with timbre is related to the principle of indeterminacy. Towards the end of the nineteenth century a serious problem arose in the theory of the absorption and emission of radiation by heated bodies. Put simply, conventional theories seemed to predict that any body in thermal equilibrium should radiate infinite amounts of energy in the ultraviolet region of the spectrum and this was consequently known as the *ultraviolet catastrophe*. The problem was eventually solved by Planck's introduction of the *quantum of action*. The Quantum Theory is now central to contemporary physics and its central assumption is that energy can only be emitted in small finite packets (known as *quanta*). The size of these energy quanta (E) is directly related to the frequency (f) of the oscillator emitting or absorbing radiation by the formula $E = hf/2\pi$ where h is Planck's constant.

An important cornerstone of quantum mechanics is Heisenberg's *Uncertainty Principle*. This states simply that the position and momentum of a particle cannot both simultaneously be known exactly. In fact, the more accurately one is known the less accurately the other *must* be known. A similar relationship holds between energy and time, i.e. it is possible to know the energy of a particle with great accuracy only if one does not know exactly at what time it has this energy. The Uncertainty Principle is often presented in elementary books about modern physics as a result merely of the interference of the instruments of observation with what is being observed. It is after all more natural for us to assume that the particle does have a definite energy at a particular time and it is merely a problem of getting to know what that is. The conventional view among physicists, however, is that the Uncertainty Principle is intrinsic to the nature of reality rather than a mere accident of experimental design. It can be shown that the mathematics resulting from an assumption of the truth of the Uncertainty Principle has observable physical consequences which would not be expected

if the uncertainty were not an intrinsic part of nature (for example, so-called *exchange phenomena* between electrons in chemical bonding).

Energy/time indeterminacy (which has a direct relationship to our discussion of the structure of sounds) can be understood fairly simply. First of all, we must remember that in Quantum Physics the energy of a system is related to its frequency. How can we therefore measure the frequency of a system at a particular instant in time? The answer is simply that *we cannot* because frequency is a property of the system dependent on its actual evolution through time. We can say that a system vibrates five times in a second but we cannot talk about how many vibrations it undergoes at an instant in time. Hence the instantaneous energy of a system is not definable. Fourier analysis of a signal into its spectral components is similarly limited. The mathematics of Fourier analysis assumes that the signal persists for an infinite length of time. If it does not, even if it appears in the time domain to be a pure sine tone, in the frequency domain it will be found to have some spectral colouration. The reader may intuitively grasp why this is so by looking at Figure 3.8. Thus when only a very small part of the curved edge of a sine wave is present we have no way of predicting that this will continue as a sine wave or that it will prove to be the leading edge of

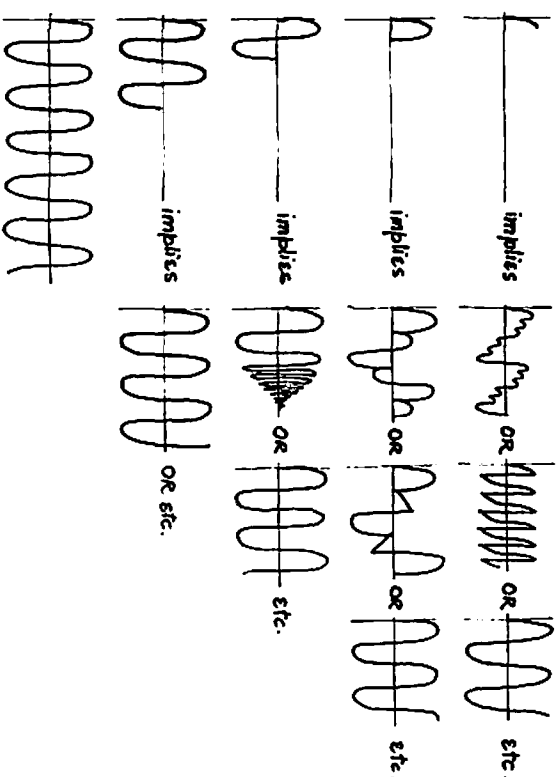


Figure 3.8 Indeterminacy in the analysis of brief signals.

an essentially stepped signal with sinusoidal rounding (for example). Even when we have a complete cycle of a sine wave we do not know that this is one complete cycle of the wave. It may be only the opening formation of a more complex pattern. Even when we have two cycles of a sine wave, though now we can perceive a regularity in the structure, it is still not certain that we have the complete picture. As more and more cycles of the sine wave are taken into our sample, the uncertainty in the nature of the signal reduces rapidly but it only reduces to zero when the sine wave persists to infinity.

The physical consequence of this is that if we produce a single cycle of a sine wave (no matter how 'pure' it may be) what we will in fact hear is a click, a sonic impulse whose frequency is maximally indeterminate. This is not as some of the *Groupe de Recherches Musicales* writings seem to suggest just a limitation of the ear. It is intrinsic to Fourier analysis itself and therefore any physical instrument performing a spectral analysis on the signal would register a similar indeterminate result. Hence a sine tone, no matter how 'pure', if sufficiently brief has no definite pitch! This phenomenon is illustrated in Examples 3.1 and 3.2. The same melody is played in both examples but in the first the individual elements are less than five milliseconds long, whereas in the second example they are ten milliseconds long. In the first example no sensation of pitch is conveyed.

For similar reasons, if we suddenly switch on a pure sine tone and then suddenly switch it off we will experience (and analysis will confirm) a spreading of the spectrum at the start and end of the sound. In this case we are not talking about switching transients in the apparatus which generates the sine tone (which we may assume we have eliminated) but the inevitable results of Fourier analysis of a finite signal. This frequency spreading can be reduced by using a cosinusoidal envelope on the attack and decay of the sample (see Figure 3.9).

It will no doubt have occurred to the reader that no real sounds last for an infinite length of time! Furthermore, few sounds exhibit constant periodicity for any length of time. Also, a sound does not have to be random to have a constantly changing wave-shape. A simple example would be a portamento on any instrument. Any practical Fourier analyser (including the human ear) must do a kind of piece-wise analysis of sound and link up the results in order to gain any reasonable conception of the nature of the complex sound-world surrounding us.

Having discovered that 'pure' sine tones may have no definable pitch, we may also discover that clear pitches may not be associated with the appropriate sine tone! The simplest example of this is illustrated by first recording a low note on the piano (Example 3.3) and then filtering out

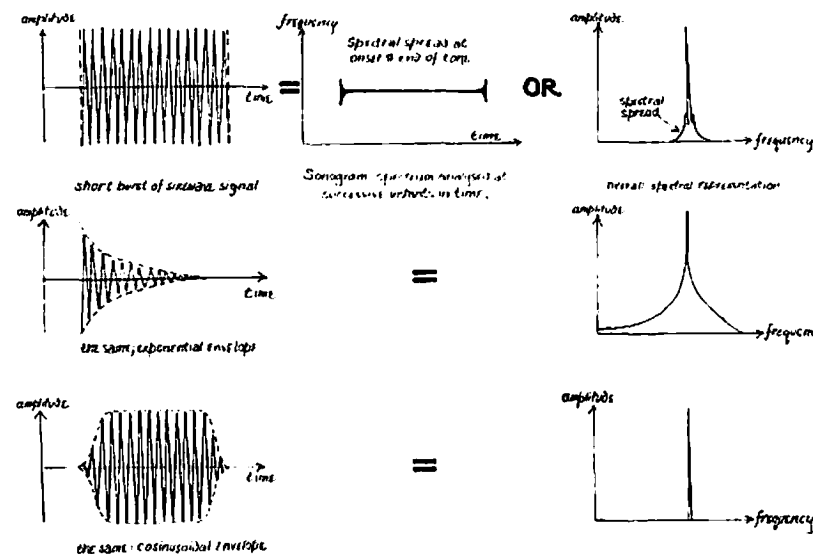


Figure 3.9 Spectral spreading in finite signals.

all frequencies below the second harmonic,¹ i.e. filtering out what we assume to be the fundamental pitch (Example 3.4). Amazingly the filtering has no effect on the sound whatsoever and we are led to conclude that there is no single component in the sound corresponding to what we hear as the fundamental pitch! The fundamental is in fact a mental construct extrapolated from the information contained in the higher partials. This effect is not peculiar to piano tones and can be demonstrated for other types of sound.

Most spectacular of all (Example 3.5), using extremely precise control of the overtone structure of a sound through computer synthesis, we can trade off the increase in perceived 'fundamental frequency' against the weighting of high and low partials in the sound in such a way that the pitch of a sound glissandos upwards at the same time as the sound descends gradually to the lower limit of the perception of pitch! The sound appears to ascend and descend simultaneously. Such experiences lead us inexorably to the conclusion that pitch is an aspect of the perception of

¹ Strictly 'second partial' as the piano spectrum is substantially inharmonic (Ed.).

timbre, and not vice versa. At the very least, we should become aware that the strict separation between pitch and timbre is an artefact of the way we have constructed our musical reality.

The inharmonic and the non-periodic: a dual conception of pitch

So far we have confined ourselves largely to Helmholtz's material, i.e. sounds of stable tessitura in which the partials are integral multiples of the frequency of the real (or imaginary) fundamental. When the relationships between the partials of a sound are not of this kind (i.e. when they are inharmonic) the Fourier analytical attempt to extract a single pitch characteristic from the sound breaks down. The object appears to perception as an aggregate of various pitches, more fused than a typical chord in instrumental music but definitely not a singly-pitched note. Many bell and bell-like sounds are of this type. Composers working in electro-acoustic music seem to be particularly attracted to these kinds of sounds and one speculates whether it is not due to their similarity to chord structures in conventional music. The sounds, in fact, are not radically different from the normal sound-objects found in conventional musical practice. They lend themselves to similar (lattice-based) modes of organisation and do not challenge our conception of what is and is not a musical object (as sounds with dynamic or unstable morphologies will do). At another perceptual extreme we have entirely non-periodic signals. In the architecture of typical analogue synthesisers (and in much discussion of electronic music) such sounds, usually referred to as 'noise', are often treated as entirely separate entities from materials with clearly defined spectra usually generated from simple oscillators. In fact, as we shall discuss, there is no simple dividing line between periodic and non-periodic signals, but in fact a multidimensional array of complex possibilities between the two extremes. Noise is not something to be treated separately from other materials, either compositionally or conceptually, but an alternative way of perceiving and relating to sound phenomena which we shall now discuss.

As discussed earlier, the result of Fourier analysis is to transform information about a sound from the time domain into the frequency domain. The most immediately striking thing about noise-type sounds is that the time domain criteria apparently cease to apply. For example, if we take a recording of a typical periodic pitched sound or even an inharmonic sound and we play back the recording at double speed the time domain information passes us at twice the rate and we hear the sound transposed up an octave. If, however, we try the same experiment with white noise, we experience no change in the frequency domain information (Example 3.6).

More complex signals may yield even more startling results (for instance in Example 3.7 where a complex sound is played first at normal speed and then at double speed. The sound appears to shift by about a third).

The explanation for this lies in the characteristics of noise-type sounds and the way we perceive them. For noise-type sounds the time domain information is essentially random. The amplitude of the signal varies randomly with time and, for example, for Gaussian noise, the probability that the amplitude will have any particular value at any particular time is maximum for zero amplitude and dies away smoothly to a probability of zero as the amplitude increases. A Fourier analysis of a small sample of a noise signal will yield a somewhat arbitrary array of frequencies and amplitudes. Analysis of another similar sample will yield a quite different but equally arbitrary array of frequencies and amplitudes. The only way that we can achieve a coherent analysis of this signal is through a statistical averaging process. We can show the average amplitude of the various frequencies in the signal over a long period of time. For white noise the frequency domain analysis reveals a completely flat spectrum. All frequencies are present with equal probability. This explains the phenomenon illustrated in the example. When the tape is played at double speed, all the frequencies are shifted up an equal amount and those that were lost near the bottom of the spectrum are replaced by even lower frequencies (which, due to the physiology of hearing we might not even have been able to perceive previously).

There are two things we must bear in mind when considering the nature of noise-based signals. First of all by declaring that a typical noise-based signal contains all possible frequencies at random amplitudes distributed randomly in time, the typical text book on acoustics tends to imply that noise is essentially an aggregate of more elementary sounds. This conception is an artefact of our perception of the analysis. As anyone who has switched on a synthesiser will know, noise is a perfectly coherent source, no less coherent than a typical sine wave oscillator and need have no particular granular characteristics. What we are saying here applies particularly to white noise but the noise concept (involving time-averaging of a spectrum) may be applied to a great many sound-objects, and noise-like objects may certainly be produced through the aggregation of simpler objects. For example a dense cluster in the lower register of the piano has a fused noise-like quality quite different from a similar cluster played in the high register. Secondly, and most important, noise involves a different mode of listening. In fact, it might be more accurate to suggest that the noise conception is a property of the way we hear rather than of the object itself. This distinction might seem arbitrary when we consider only oscillators and

noise-sources on a typical synthesiser but when we begin to consider sounds with extremely complex evolving and semi-irregular spectra, the distinction comes strongly into play.

With sounds therefore that we describe as noise, or noise-based we appear to respond directly to information in the frequency domain as there is little information to be gained from the time domain. This fact may be illustrated in two ways. In Figure 3.10 we see time domain representations of Gaussian and binary noise. The Gaussian noise has the random amplitude fluctuations we might expect. The binary noise, on the other hand, flips randomly backwards and forwards between two very definite

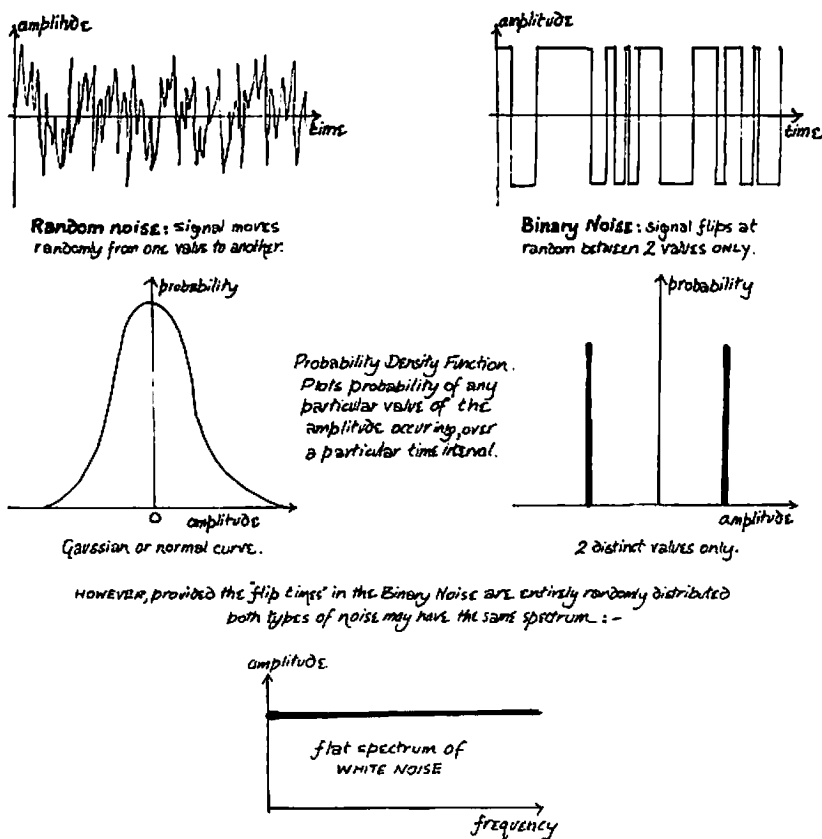
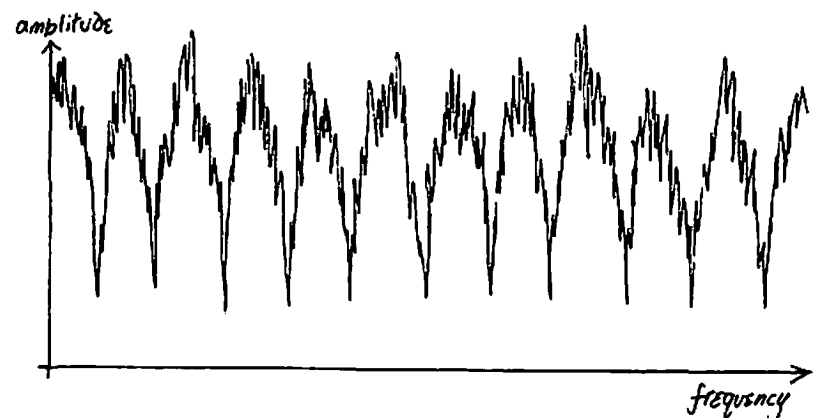


Figure 3.10 Random and binary noise (independence of frequency-domain representation).

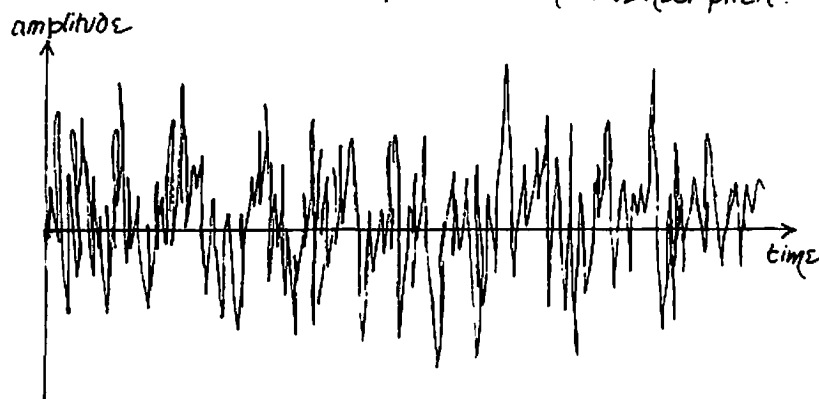
amplitudes. However, the frequency domain information can be the same in both cases and we can perceive both as the familiar 'white' noise. In Examples 3.8 and 3.9 we hear first of all the effect of transposing white noise by an octave (no effect on the frequency domain information) and secondly a 'melody' created by filtering variously the white noise to produce bands of different 'coloured' noises. Again the effect of the latter on the time domain information would be barely perceptible but can be clearly heard in the frequency domain. Perhaps the most striking example of the independence of frequency domain hearing from time domain information in noise perception is the experience of comb-filtering. In this particular case a brief delay is imposed on the noise signal and the result mixed with the original signal. The time domain representation presents us with no perceivable patterning but in the frequency domain we discover that various regular peaks and troughs appear in the spectrum, the spacing of these being related to the duration of the time-delay. Perceptually speaking, a pitch or particular spectral characteristic is imposed upon the previously undifferentiated noise source (see Figure 3.11).

Filtering noise can be used to produce broad bands which can be perceived to be higher or lower than one another but are not perceptibly pitched or to produce narrower bands which are perceived as being more or less clearly pitched. This suggests that there are at least two conceptions of pitch involved in our perception of sound-objects. The first type of pitch is related to periodicity and arises from the real or implied fundamental frequency of vibration of the source (or several of these in the case of an inharmonic sound). The second results from the imposition of a spectral envelope through some sort of filtering or resonance procedure acting on a noise-averaged spectrum. Just as the perceived fundamental pitch may be multiple (in an inharmonic spectrum) the spectral envelope itself may be quite complicated and we will normally refer to this as a *formant structure*. These two conceptions of pitch (in a sense one related more strongly to the time domain and the other more strongly to the spectral domain) meet most strikingly in the sound production of the human voice where the pitch of standard musical practice is defined by the fundamental vibrations of the larynx while the articulation of vowels in speech (and in song) is defined by varying formant structures.

We may now imagine very complex tones in which the two concepts of pitch are traded off against each other. For example, an inharmonic tone in which the various partials are in constant and rapid motion might be given a focused pitch sense by the strong accentuation of a narrow formant band using a filter. Conversely and more commonly heard, the time domain pitch of a sung tone is clearly maintained through the most extreme variations of



Comb-filtered noise: patterning of spectrum. Amplitude peaks are in a harmonic relationship to one another. We hear pitch.



Comb-filtered noise: lack of patterning remains in time-domain.

Figure 3.11 Comb-filtered noise (perception of pitch direct from frequency domain).

formant structure details. The question of what happens when both of these change at once will be dealt with in the following chapters!

Spectral evolution: harmonic and dynamic timbre

A much more fundamental break with Helmholtz may be made. Simple experiments with sound-objects demonstrate that timbre is not merely

dependent on spectral information but also upon the way that information evolves through time. The most striking illustration of this fact is given in Examples 3.10 and 3.11. In Example 3.10 we hear first the sound of a piano note whose envelope has been smoothed and flattened. Following this is the sound of a flute playing the same note. The two are virtually identical. In the second example (3.11), we first hear the sound of a flute on which has been imposed the amplitude envelope of a piano note. This is followed by a piano note at the same pitch. Again, the two sounds are indistinguishable. The fact is that in this particular register the spectral content of piano and flute sounds are very similar. What differentiates these sonic objects is the temporal evolution of the amplitude of the event. The flute remaining relatively constant whilst the piano decays linearly¹ from a sharp attack. Thus dynamic aspects of the spectrum enter into our perception of timbre, so we see that timbre is at least a two-dimensional entity contrary to the conventional wisdom.

Furthermore, the characteristics of the evolution of the amplitude envelope are of fundamental importance. If we take a sound of relatively constant envelope and edit away various parts of it, we experience no noticeable change in timbre (Example 3.12). Extending this notion we need not keep the amplitude constant but merely maintain a constancy in its rate of change. (A constant amplitude is simply an amplitude envelope with constancy of change, the rate of change being zero.) With a piano, the sound actually dies away but the rate of this decay is constant (linear decay). If we therefore edit off various amounts of time from the beginning of the piano note, its timbre will not be noticeably changed (Example 3.13). If we try the same experiment with an instrument whose amplitude envelope varies non-linearly, we discover that editing off the beginning of the sound changes the timbre distinctly (for instance, Example 3.14 using a bell). This mode of argument applies equally well to unpitched sound and may be heard by listening to Example 3.15 which uses a cymbal having a linear decay with noise-type material.

To be even more precise we must take into account the evolution of amplitude of each component of the spectrum in our sound. If we repeat the same experiment using a vibraphone (Example 3.16), we will find that when we lose the very start of the sound it is altered significantly, but editing off more of the sound's beginning has no further appreciable effect. This is simply because there is a rapid spectral change during the initiating moment of a vibraphone note, due to the metallic sound of the hammer striking the

¹ 'Linearly' in terms of the perception of loudness (roughly speaking dBs); not amplitude which is decaying exponentially with time. (Ed.)

key but then the subsequent resonance dies away linearly (as with a piano note) because of the material and design of the instrument.

Even a conventional instrumental sound may contain significant noise components. In Examples 3.17 and 3.18 we first of all hear a flute sound without its initial 50 milliseconds (followed by the original recording) and the same experiment repeated with a trumpet. The effect is very much more marked with the flute because the flute sound is partly characterised by a noise-based breath-sound which initiates the resonance of the tube. Thus, even with conventional instruments, we begin to see a sequential breaking up of the characteristics of the sound-object which we will explore more fully in the section on multiplexing and which is more typically characteristic of speech-streamed sound-objects. The discovery that timbre itself is partly dependent upon the evolution of spectral characteristics is our first real link with sounds of dynamic morphology, i.e. sounds in which the perceived pitch spectrum, amplitude envelope etc., all evolve through time.

Coherence of sound-objects; aural imaging

Having discovered that sound-objects may be exceedingly complex and that our perception of them may involve processes of averaging and attention to spectral evolution, an obvious question presents itself; how are we ever able to differentiate one sound-source from another? As all sounds enter our auditory apparatus via a single complex pressure wave generated in the air why do we not just constantly hear a single source with more or less complex characteristics? We might ask the same question in reverse: how is it that a complex sound does not dissociate into a number of separate aural images? Much research has already been done and much is still being carried out on the problem of *aural imaging*. The following observations are drawn from the work of Steven McAdams at IRCAM in Paris (see McAdams 1982).

To phrase the question a little more technically, once our auditory mechanism has dissociated the incoming sound into its constituent sine wave components, or at least generated some kind of spectral analysis, how can it then group the various components according to the separate sources from which they emanated? There appear to be at least four mechanisms in operation here. These are:

- (1) components having the same (or very similar) overall amplitude envelope, and, in particular, components whose onset characteristics coincide will tend to be grouped together;
- (2) components having parallel frequency modulation (either regular in the form of vibrato or irregular in the form of jitter) will be grouped together;

- (3) sounds having the same formant characteristics will be grouped together;
- (4) sounds having the same apparent spatial location will be grouped together.

Any or all of these factors may enter into the process of separating one aural image from another. The importance of onset synchrony is demonstrated in Example 3.19 where the various constituents of a sound are separated by increasing time-intervals and then the time-intervals successively reduced until there is again complete synchrony. The sound-image will be heard to split into its component parts and then recombine. The importance of frequency-modulation information has been most eloquently demonstrated in work by Roger Reynolds⁵ (Example 3.20). Data from a phase vocoder analysis was used to resynthesize an oboe tone, elongating it as well. The regenerated oboe tone was projected from two loudspeakers, the odd harmonics on one side, the even on the other. These two groups of partials were each coherently, but differently, frequency modulated. Because the even set was modulated at a rate corresponding to vocal vibrato, and the odd necessarily had a clarinet-like sound, the listener experiences a distinctive composite as the amplitude of modulation increases: clarinet on one side, voice on the other at the octave, and the sum, an oboe sound, in the centre. In this way we can contemplate playing with the aural imaging process and not merely destroying the convention of instrumental streaming.

Conversely, we may use these aural imaging factors compositionally to integrate sound materials which might otherwise not cohere into objects. Thus, by imposing artificial attack and amplitude characteristics on a complex of sounds (e.g. a small group of people laughing), by artificially synchronising the onset of two or more normally quite separate sound-objects, by artificially synchronising the vibrato and jitter on two or more normally quite separate sound-objects we may create coherent composite sound-objects. A recently popular example of this approach is the use of the vocoder where the evolution of the formant characteristics of a speaking voice is imposed on an otherwise non-coherent source (e.g. the sounds of a large group of people speaking before a concert as in Michael McNabb's *Dreamsong*). This further opens up our conception of what might be considered a coherent musical object.

With all these potential sound-materials at our disposal a further problem arises. Music is normally concerned with establishing relationships

⁵ Working with Steven McAdams and Thierry Lancino at IRCAM. This example is based on band 2c on the IRCAM LP 0001 which is described somewhat differently on the sleeve note. It is from his work *Archipelago*. (Ed.)

between various kinds of material. The question is what determines whether we perceive a particular piece of sound-material as related to another. Speaking of sound organisation in the broadest possible sense, the answer to this question will clearly depend partly on context (upon which aspects of sonic organisation are being focused upon — pitch, spectral type, formant streaming etc.). But whichever approach we take there will be a point beyond which the manipulated sound-material will cease to have any audible relation to its source. This can already be perceived in conventional music where in some types of complex serial organisation, the concept of the derivation of material from a source set becomes meaningless. In the studio it is seductive to assume that a sound derived from another by some technical process is hence derived from the original in a *musical* sense. A simple example of this may be given as follows. Suppose that we start with a sustained orchestral sound, the sound of a large crowd and the sound of a stable sine tone. Let us now take each sound, put it on a tape recorder and switch the tape recorder onto fast wind so that the sound accelerates from speed zero to very fast and as the tape recorder reaches its maximum speed fade out the sound to nothing.⁶ Having done this with all the sounds, let us now speed them up to at least sixteen times their original speed. In what sense are the resultant sounds related to the originals? What we perceive in each case is a brief, high frequency glissando. Furthermore, and most striking, all three sounds now appear very closely related whereas the sounds from which they originated have no relationship whatsoever. At this distance of derivation it is the overall morphology of the sound-structures which predominates. We may learn two lessons from this. First of all, with sound-objects having a dynamic morphology, it is this morphology that dominates our perception of relatedness- unrelatedness, rather than spectral or even more general timbral considerations. Secondly, if the organisation of our music is to be based on the audible reality of the listening experience, the music must be organised according to *perceived* relationships of materials or *perceived* processes of derivation of materials. In order to accomplish the former we need an analysis of sound-materials based upon their perceived properties, a phenomenological analysis of sounds.

The phenomenology of sound-objects

The pioneering work on the development of a phenomenological description of sound-objects and an aesthetic based upon it was done by Pierre Schaeffer and the *Groupe de Recherches Musicales*. Certain aurally-

⁶ Unedited from 1983, this phenomenon may be simulated in digital systems (Ed.).

perceived characteristics of sound-objects have already been discussed above and this section is intended to complement what has already been said. It draws largely upon the GRM research incorporating ideas drawn from the writings of Robert Erickson and Iannis Xenakis. Perhaps the most important concept advanced by Pierre Schaeffer was that of the *acousmatic*. This term was originally applied to initiates in the Pythagorean cult who spent five years listening to lectures from the master, delivered from behind a screen (so that the lecturer could not be seen) while sitting in total silence. Acousmatic listening may therefore be defined as the apprehension or appreciation of a sound-object independent of, and detached from, a knowledge or appreciation of its source. This means not only that we will ignore the social or environmental origins or intention of the sound (from a bird, from a car, from a musical instrument, as language, as distress signal, as music, as accident) but also its instrumental origin (voice, musical instrument, metal sheet, machine, animal, larynx, stretched string, air column etc.). The idea of acousmatic listening is easily appreciated by anyone who has worked with recorded sound-materials in the electro-acoustic music studio. When working with large numbers of sounds from different sources and particularly when this material has been transformed, if only slightly, it becomes difficult to remember from where the various sounds originated and from a compositional point of view such origins need have no special significance. The transformation of the flute tone into the sound of a piano (above) illustrated this thesis, though a truly acousmatic approach would demand that we forget not merely that the sound derived from a flute but also that after its transformation it appeared to derive from a piano! We should concern ourselves solely with its objective characteristics as a sound-object. Example 3.21 illustrates the need for the separation of sound-object description from any reference to the source. The various sounds here are all derived from the same source-object (a tam-tam excited by a variety of objects).

From our discussion in the previous section we became aware of a distinction between sounds which are transposed by a change of replay-speed (time-domain transformation) and sounds (like white noise) which do not change their pitch, or change pitch in some unpredictable way under the same transformation. We also discovered that various properties of a sound may be altered by filtering. With certain sounds, however, though qualitative changes can be perceived as a result of filtering, we do not feel that the underlying sound-object has been fundamentally altered. The instrumental tones of conventional musical practice are typical examples of sounds which can be transposed by time-domain transformation and are resistant to filtering. They are not the only ones, however, and it is necessary

to define a more general characteristic of sound-objects having this property. Following the French terminology, we will refer to this as *mass*. An example of a complex sound having a definite mass and illustration of its resistance to filtering and its transposition are given in Example 3.22. In actual practice, of course, there will be a large, grey area where time-domain-based perception of complex timbres and spectral perception of formants (see previous section) meet.

A second perceived characteristic of sound-objects is *grain*. If we take a slowly repeating pulse and gradually speed it up beyond about 20 Hz, we begin to hear a definite pitch and as the speed increases further we begin to lose any sense of the original individual impulses. In between the extremes of impulse perception and pitch perception we perceive a pitched object with a certain amount of 'grittiness' as the individual impulses are still apparent in some sense to our perception of the sound-object. This internal 'grittiness' is the *grain* of the sound-object and is illustrated for an electronic impulse (Example 3.23) and for a bassoon note (Example 3.24) in the sound examples. In this particular case we are talking about a regular, periodic grain, but it is possible to define grain in a more general manner. If we have a sound made up of a large aggregate of brief impulses occurring in a random or semi-random manner, we can talk about the statistical average rate of occurrence of these impulses and this particular parameter will have an effect on the perceived characteristic of a sound. Erickson (1975) refers to this characteristic as *rustle time* and might also be thought of as *aperiodic grain*. It has an important role in our perception of particular types of percussion sounds. For example, we may note the perceptual difference between sounds of types of rattles filled variously with sand, seeds and shot. Note, however, that the spectral characteristics of the individual impulses in all these cases will also contribute (probably in a statistically averaged way) to the perceived character of the resulting sound. Aperiodic grain also has a bearing on the particular sonority of sizzle-cymbals, snare-drums, drum-rolls, the lion's roar, and even the quality of string sound through the influence of different weights and widths of bows and different types of hair and rosin on the nature of bowed excitation.

It is important to realise that there is a perceptual threshold at which we cease to perceive individual events as individual events and begin to experience them as contributing to the grain of a larger event. As a result, at sufficiently high speed, any sequence of sound-objects may become fused into a larger object with grain. Example 3.25 illustrates this first for a descending scale and then for an irregular melodic pattern. Incidentally, if we applied the same process to a string of speech sounds (Example 3.26) we approach the conception of a *multiplex* to be discussed in the next chapter.

Combining the concept of rapid rustle-time with multiplicities of brief sound-objects of various spectra we begin to define another huge class of possible sound-objects (e.g. the sound of rain, of poured pebbles, etc.) hinted at in Xenakis (1971/1992). This book is a very interesting early analysis of a generalised notion of sound-objects as evolving groupings of

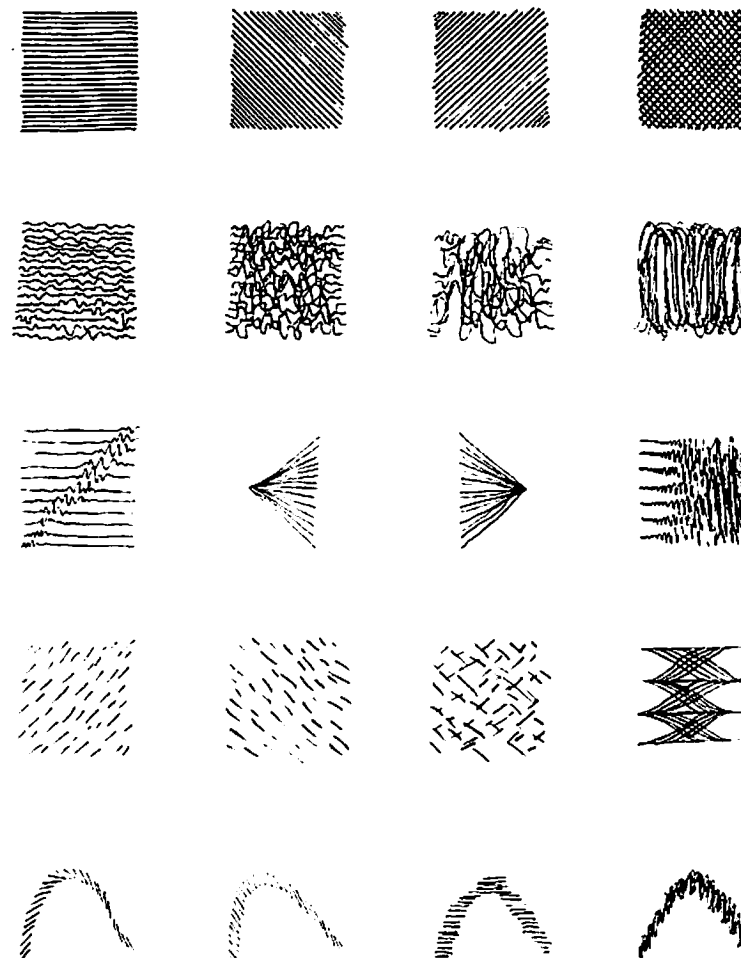


Figure 3.12 Examples of filament structure.

elementary particulate sounds. Unfortunately, the musico-descriptive potential of the approach gets rather lost in Xenakis' absorption in the particular mathematical methodology (stochastic processes and Markov chains) leading the composer off in a rather specialised aesthetic direction. In the book, however, Xenakis does deal with the *grain structure* and *filament structure* of dense sound-objects. Just as *grain structure* may be thought to apply to sounds made up of elementary impulses, filament structure applies to sounds made up of elementary sustained units. Figure 3.12 expanding on Xenakis' descriptions, illustrates various possible filament structures. If we now begin to discuss the temporal evolution of such concepts as aperiodic grain and filament structure a whole new world of complex sound-objects begins to open up before us. The GRM classification goes on to discuss the concepts of *complex-note*, *web* and *eccentric* sounds but I will discuss sounds of this and related types in a somewhat different manner in the following chapters. A more complete description of the GRM methodology can be found in the *Solège de l'objet sonore* (Schaeffer, Reibel and Ferreyra (1983)).

IN SONIC ART

⁷ The *Solège* is still available on three cassettes from the INA/GRM (Paris) at time of press but the trilingual printout of the recorded French commentary which accompanied the LP version appears to have been discontinued (E/f.).