NRC
TT - 608

NATIONAL RESEARCH COUNCIL OF CANADA

TECHNICAL TRANSLATION TT - 608

THE MATHEMATIC - ACOUSTICAL FUNDAMENTALS OF ELECTRICAL SOUND COMPOSITION

BY

WERNER MEYER - EPPLER

FROM
TECH. HAUSMITT. NWDR, 6: 29 - 39, 1954

TRANSLATED BY

H. A. G. NATHAN

OTTAWA

1956

NATIONAL RESEARCH COUNCIL OF CANADA

Technical Translation TT-608

Title: The mathematic-acoustical fundamentals of electrical sound composition*.

(Mathematisch-akustische Grundlagen der elektrischen Klang-Komposition).

Author: Werner Meyer-Eppler.

Reference: Technische Hausmitteilungen des Nordwestdeutschen Rundfunks, 6: 29-39, 1954.

Translator: H.A.G. Nathan, Translations Section, N.R.C. Library.

Translated with permission.

*Paper No. 8 of a special collection of twelve papers on electronic music published by the Northwest German Broadcasting System. These have been translated by the National Research Council and issued as TT-601 to TT-612.

THE MATHEMATIC-ACOUSTICAL FUNDAMENTALS OF ELECTRICAL SOUND COMPOSITION

Summary

The inclusion of electric tone sources in the process of musical production forces the composer to tackle to an unusual degree certain problems of a psycho-acoustical nature. For dealing with the manifold relations between the sound and allied sensations, a simplified, descriptive formula is described. This has properties which take account of the behaviour of the human ear, and by means of which the expected sound effects may be defined in advance with some degree of certainty.

The musical composition which can be produced with electric tone sources differs so much from the conventional that only in exceptional cases will it be possible to extrapolate some of the "assets" of traditional orchestration methods into the new regions of sound. Anyone entering the new field of electronic music will be confronted with entirely different conditions and unexpected as well as unfamiliar phenomena. In order to avoid confusion owing to the steadily increasing number of forms of sound, it is necessary to understand the relationships between the physical, psychological and physiological aspects of the problem. Furthermore, in order to avoid getting entangled in a mass of conflicting opinions, a well-defined terminology is required. Insofar as the various committees of experts have recommended definitions and certain terms it is suggested that they be taken into account.

1. Definitions

The German Committee on Acoustics recommends the following definitions*:

Simple tone	-	Sound of sinusoidal curve form.
Tone mixture	-	Sound composed of tones of arbitrary frequency.
Simple musical sound	-	Sound composed of harmonic partial overtones.
Musical-sound mixture	-	Sound composed of musical sounds with fundamental tones of any frequency.
Noise	-	Tone mixture to which a continuous spectrum corresponds, or which is composed of a great number of individual tones whose frequencies are not related to each other as integral numbers.
Report	-	Sound impulse, chiefly of great sound intensity.

Of course, this list must seem strange to a musician since in places it is contradictory to his own idiom**.

In order to avoid misunderstandings, the sensations caused by sound are defined here by the word "sensation"; hence "simple-tone sensation", "tone-mixture sensation", "musical-sound sensation", "noise sensation", etc. However, it should be noted that the physiological-psychological qualities cannot be defined even approximately in as simple a manner as the physical objects of the same names.

---

*DIN 1320

**The explanations contain the following sentence: "It is a known fact that the definition of the terms 'tone' and 'musical sound' in acoustics differs from those in music. The musician denotes a simple musical sound as 'tone' but a musical-sound mixture as 'musical sound' (e.g. a triad)".

The American Standards Association attempted to supplement the physical definitions by defining the qualities of sensation*:

Tone	-	A tone is a sound sensation having pitch.
Simple tone (pure tone)	-	A simple tone is a sound sensation characterized by singleness of pitch.
Complex tone	-	A complex tone is a sound sensation characterized by more than one tone.

Surprisingly enough, the American terminology does not contain a definition of what might be denoted as "noise sensation". But a term for timbre, which is absent in the German list, has been included:

Timbre (Musical quality)	-	Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.

2. <u>Relationship Between Stimulus and Sensation</u>

The human ear does not uniformly respond to all acoustic stimuli. Nor do stimulus and sensation correspond to one another in a clear and reversible way. However, while the realization of psychological optics that "colours" are not properties of objects but exclusively qualities of sensation has become general knowledge, little attention was hitherto paid to the respective relationships in the field of acoustics.

With respect to the definitions alone there is a widespread lack of competence in making distinctions between the physical-acoustic phenomena (and the abstracted representation in the form of written music) and the qualities of acoustical sensation. The pattern presented by notes and the sound sensation are often too closely correlated. For example, for this reason

---

*Acoustical Terminology. New York, 1951.

it is possible that a diminished fifth (or a fourth) blown by horns in the unaccented octave is frequently assumed to be functionally equivalent to a diminished fifth (or a fourth) produced on a xylophone in the thrice-accented octave*, although there is no such equivalence in the sphere of sound sensations.

Therefore, by analogy with the terminology of the psychological optics the property of a stimulus depending on similarity and dissimilarity of the sensation is designated here as a valence of the stimulus. If the same sensation is caused by objectively (i.e., physically measurable) different stimuli, the respective valences are designated as "conditionally equal" and the sensations as "metameric"**.

The valences may be presented in a multidimensional space, where the components of the valences, e.g. amplitudes, frequencies, coordinates in space and time, etc., have the function of coordinates. To every sensation there corresponds a point of sensation in the valence space. However, these points are not in arbitrarily close proximity; on the contrary they are separated by increment thresholds. The cellular structure thus resulting is the metric field of the valences. As a rule, this field is not constant but depends on the rate of change of the stimuli***. For example, physical sounds thus carry along the

---

*Translator's note: The octave starting two octaves above middle C.

**The term "metameric" was coined by W. Ostwald for colours having the same appearance but an objectively different remission spectrum.

***In simultaneous presentation the number of distinguishable qualities of sensation is much smaller than in successive presentation.

metric field over great distances. Only if they follow one another directly is a distinction made between them[12]. The phenomenon is designated as "restimulation". The adaptation, i.e., the restimulation of the cells of the auditory organ to another absolute range of sensitiveness[9,11], is well known. For example, this restimulation causes a fortissimo immediately following a pianissimo or a pause to be perceived with greater loudness than the same fortissimo after a forte.

The effect of adaptation on the sensation intensity for an initial mf passage with a sudden increase of the stimulus intensity to f after two minutes is shown in Fig. 1 after v. Békésy[11 p.132].

In an isolated presentation the valence points are only roughly distinguishable. The cellular structure becomes more clearly defined when the stimuli are in close succession, but at the same time it shrinks.

The cellular structure of the valence field renders all changes in stimulus which are smaller than the increment threshold ineffective. Hence, an arbitrarily fine graduation cannot be attained with respect to either pitch intenisty or sound intensity. For example, at a sound intensity of 80 phon. the number of pitch stages that can be distinguished in the entire range of audibility at successive presentations scarcely exceeds 2,000. The ability to distinguish the frequency changes of formants is even less well developed. In the optimum case, i.e., that of sonant vocals, a change in the frequency of the centre of gravity of a formant by 6% can just barely be perceived by the human ear. In addition to the ability to perceive differences in frequency in successive presentation, where the two frequencies follow one another at not too great time intervals, persons having an "absolute pitch" can distinguish the pitch of isolated musical sounds just as persons with a sense of colour can distinguish between the optic colour tones. In the range from 50 to 4,500 c.p.s. such a person can

accurately distinguish between more than 70 different frequencies (approximately at half-tone distance[17]). The average observer, who does not have absolute pitch, can distinguish only five groups of frequencies[10].

3. Morphology of Acoustic Stimuli

(a) Classification of sound phenomena

Any sound phenomenon is completely determined from the space and time slopes of its velocity potential, whence the acoustic pressure is deduced by differentiation with respect to time and the volume current by differentiation with respect to the coordinates in space in a known manner[16]. Although the volume current is very important for the formation of plastic sound impressions[15] we shall confine ourselves below to a consideration of the acoustic pressure distribution at a fixed point of observation. Hence the acoustic pressure p is assumed to be a function of time alone:

$$p = p(t).$$

The acoustic pressure distributions that are physically possible may be classified according to a great many different view points. However, if it is required that the result of such a classification be in a clear relationship with the sound sensations, a number of special classification methods may be used. It is characteristic of these methods that they attempt to simulate the sound-analysis mechanism of the ear as far as it is known. Such a combination of morphological classification methods and the state of our knowledge of the hearing mechanism at the time of investigation is not very convenient, of course, since it might be necessary to dispense with a classification method hitherto applied as soon as new discoveries in the field of physiology and psychology may require it. However, this cannot be avoided altogether or else it may be impossible to gain greater insight into the nature of sounds and noises.

(b) Harmonic analysis

The most elementary and best known classification method, which is almost exclusively applied in the literature of musical science, is based on the "harmonic analysis". In terms of mathematics this means resolving of periodic acoustic pressure distributions into sinusoidal components ("harmonics") whose frequencies are related to one another as integral numbers. The term "periodic" implies, as a condition of the admissibility of the method, that the oscillation phenomenon is completely uniform and has neither a beginning nor an end. Hence, the harmonic analysis can be used as an approximate method of classifying sound phenomena, since it does not explicitly contain the time as the most essential element of all sound phenomena. This important restriction is frequently overlooked and, as an unavoidable result, the data thus obtained do not show agreement with the psycho-acoustical observations.

(c) The time-frequency analysis

Very good agreement between the result of resolving and auditory sensations is obtained from the time-frequency analysis (cf. appendix), in which the transition from the Fourier series (harmonic analysis) to the Fourier integral is carried out. If it were possible to express the result of a harmonic analysis by a number of discrete spectral amplitudes independent of time (line spectrum), then the time-frequency spectrum to be calculated by means of the Fourier integral is continuous and variable with time as well. Since time, frequency and spectral amplitudes are independent coordinates, two-dimensional symbolization is not possible here, as opposed to the line spectrum of the harmonic resolution. Either a method of perspective representation must be used or the third coordinate must be expressed by optical characteristics (density graduations) as in the "sonograms" of the well-known visible-speech method(5). In Fig. 2 the time-frequency spectrum is shown as a perspective drawing and is also expressed by the variable-density method.

In the deduction of the time-frequency spectrum (by calculation or determination by apparatus) another very essential fact must be taken into account. While only one type of harmonic analysis is admissible for a periodic sound process, any non-periodic process may be represented by a time-frequency spectrum in infinitely many ways. This is due to a freely selectable parameter, which is designated as the "analysis interval" of the spectral analysis(7). The smaller the analysis interval the finer will be the subdivision of the spectrum with respect to time, and at the same time the rougher will be the spectrum analysis. A wide analysis interval results in only a rough localization of the processes in time but in a more detailed spectral analysis, i.e., the "selectivity" is high. The limiting case of the infinite analysis interval, which while it gives the greatest possible resolution of frequency does not admit any localization in time, is merely the harmonic analysis mentioned above.

The other limit is the infinitely small analysis interval, where no spectrum analysis takes place, but the way the acoustic-pressure function varies with time (i.e., the time course of the acoustic pressure function) is faithfully reproduced. This is the case in ordinary oscillograph recording.

These internal relations between the time resolution and the frequency resolution of a sound process satisfy a mathematical law, which combines the width, Δt, of the analysis interval per unit time and the width, $\Delta\nu$ of the finest structures relative to frequency. In mathematics this is known as Schwarz's inequality in "Physics as the Uncertainty Principle", and it says:

$$\Delta\nu \times \Delta t \geqq 1. \qquad (1)$$

4. The Time-Frequency Spectrum and Properties of the Human Ear

The width of the analysis interval is of decisive importance for the degree of agreement between the time-frequency

spectrum of a sound phenomenon and the corresponding acoustical sensation. From numerous investigations it is well known that the order of sound sensations following one another with a period of less than 25 milliseconds* is no longer perceived. For example, echoes can only be perceived if their transition time is longer than 25 milliseconds. Only for time intervals up to this value does the often misunderstood phrase "phase-insensitiveness of the ear" apply[(2)].

At a tape speed of 76.2 cm./sec. the "critical length of tape" is thus 1.9 cm. Pieces of tape up to this length can be played forwards and backwards without change in the sound sensation.

However, the interval uncertainty relative to frequency, $\Delta\nu \geqq 40$ c.p.s., which may be deduced from the analysis interval $\Delta t \geqq 25$ milliseconds, by relation (1), only applies for simultaneous presentation. The successive threshold for simple tones or sounds lies in the lower range of audibility at approximately 0.5 c.p.s. and increases as the frequency increases (at 10,000 c.p.s. it has the value 40 c.p.s.). These data apply to an acoustic-pressure level of 60 db. above the audibility threshold.

Whether at an interval length of 25 milliseconds an observer can actually hear all the spectral details that may be determined mathematically depends to a high degree on his acoustic-gnostic capabilities, the direction of his attention and his familiarity with the sound being presented. Therefore, as far as the average listener is concerned, the structure of a sonogram with respect to frequency should be considered an upper limit of the analytical resolution. It would be useless to go beyond this limit**.

---

*In psychology this finite time interval is known as "presence density".

**Provided the sensation qualities which can be distinguished in a single presentation remain mnemically available, repeated presentation may result in a considerably finer pattern of the time-frequency structure.

Since the perception volumes due to a given acoustic pressure distribution are centralized phenomena, they cannot be read from mathematical spectral representations at all. Nevertheless, the opinion of P. Schaeffer that time intervals of less than 0.1 sec. provide the sound material seems justifiable, although sound forms only occur in longer time intervals[13].

A second aspect also is important. In first approximation the human ear has the properties assigned to it by the time-frequency analysis*. However, more specifically, there seem to be processes at work which tend to blur the simple pattern. It is thus due particularly to the overriding effect that spectral regions of low intensity are attenuated or obliterated by those of higher intensity.

Furthermore, in contrast to usual means of analysis the ear has no analyzing properties which are invariable with time. For example, if a pure sound is presented, the ability to determine this sound vanishes after a few seconds, e.g. the capability of recognizing the vowel ö or a sound on the oboe with certainty[12].

This very complex phenomenon, which presumably is due to the central processes, is known as "fusion"**.

---

*Approximately such that two phenomena having the same time-frequency spectrum for the same analysis interval cannot be distinguished acoustically.

**The following experiment is characteristic. With the aid of two formant filters a two-formant vowel, e.g. ö or e, (pronounced ay in German) is produced electronically. After a while, the vocal quality disappears and a sound is heard that can no longer be recognized. If this sound is briefly cut out and then cut in again, the impression of a vowel is instantaneously restored. However, if only one of the two formant filters is cut out and cut in again, the impression of a vowel is no longer obtained. Instead the cut-out and cut-in formant is heard as an isolated quality.

It is difficult to described processes involved in the fusion, since influences due to the personality type of the observer become apparent here.

Furthermore, it was found that the human ear is capable of detecting periodicities within the sound phenomenon even if these periodicities are not shown by the time-frequency analysis*. Therefore, it was suggested that other methods of analysis be used in which the frequency is replaced by a quantity which may be obtained from the sound process. The spectrum itself remains three-dimensional, i.e., in addition to the new quantity, the spectrum contains the latter's amplitude and the time as coordinates. The following should be mentioned here: phase analysis(3), exhaustion analysis(8) and correlation analysis(1). Each of them adds new features, which are in good agreement with the sound sensations to the present pattern, although it is not yet possible to define the most suitable method of analysis. Therefore the time-frequency spectra, as the manner of representation which is the least controversial at present, are tentatively used here.

5. Sounds and Noises in the Physical and Psychological Regions

To begin with, the definitions on (physical) sound and noise given in section 1 are translated into a simplified time-frequency spectrum. For mathematical calculation the reader is referred to the appendix.

Tones, sounds and noises are steady processes, i.e. practically speaking, their spectral structure does not depend

*All the sound components which do not appear in a harmonic analysis have been mistakenly termed "subjective", although the only admissible interpretation of the facts would have been to consider the basic analysis method unjustified.

on time when an analysis interval similar to that in a human ear is used. Of course, owing to the finite length of interval there are no sharply-defined spectral lines here such as would be obtained from the harmonic analysis. This means that certain deviations from the position of these poorly-defined spectral lines may be admitted without changing the character of the spectrum; the ear is tolerant. Hence a simple sound with the frequency components 300, 400 and 500 c.p.s. remains a simple sound, even if the frequencies are slightly changed, e.g. to 306, 398 and 502 c.p.s., respectively. However, the change is perceived not as impurities of the tone but as beats of the timbre. As a surprise to a musician, the addition of physically simple tones, i.e., sine oscillations, gives an entirely different result from what would have been expected on the strength of musical experience with pure sounds (i.e., those sound processes which produce conventional music). If simple sounds, not simple tones, are chosen as components for the above experiment, then considerable roughness effects occur even in the above deviations. The reason for this different behaviour of tones and sounds is due to the fact that not only do the barely noticeable beats of timbre appear in the sounds but also the much more annoying beats of amplitude. Relative to the cited example this means that the fourth harmonic of 306 c.p.s. (1224 c.p.s.) and the third harmonic of 398 c.p.s. (1194 c.p.s.) carry out beats of amplitude having a frequency of 1224-1194 = 30 c.p.s., causing a very perceptible acoustic roughness. For the higher the competing harmonics the more annoying will be the beats of amplitude (Fig. 3). Therefore, the more overtones the sounds contain the more disturbing will be the deviations from the integral frequency ratio. A tone mixture at a "distant position", i.e., with frequency spacings never less than approximately 100 c.p.s., sounds perhaps strange but never harsh, noisy or "dissonant". Noise disturbances due to the combination formed in the ear occur only at intensitites of sounds above approximately 80 phon.

On the other hand, tempered diads* and triads, which according to conventional musical terminology would have to be considered "consonant" often produce a very unpleasant sound effect when the lowest harmonic is suppressed electrically.

The following sensations occur particularly when two simple tones (i.e., sinusoids) of moderate intensity are presented simultaneously. If the frequency spacing is less than approximately 15 c.p.s. (the value doubles or trebles for the upper range of audibility), beats of amplitude can be heard. The additional sound sensations occurring may be transcribed approximately as follows: "ringing" (thrice-accented octave), "twittering" (five-stroke octave) and "chirping" (six-stroke octave). If the frequency difference exceeds the above value, some sort of "rumbling" is heard and above the five-stroke octave a "metallic buzzing". This range of sensation steadily changes to the accordant sensation of the minor third as soon as the frequency ratio approaches the value 5:6. Hence the frequency range is greater for high frequencies than for low ones. In contradiction to the opinions held on most of the interval theories to date it must be noted that the specific character of a musical interval (third, fifth, sixth) is most probably not due to "subjective" differential tones but to the capability of the ear to perceive periodicities as a special sound quality.

Hence it may be said that the customary distinction between "harmonic" and "non-harmonic" partial tones or between "rational" and "irrational" frequency ratios becomes null and void as long as beats of amplitude are disregarded.

As stated above, the sum of two or more sine oscillations, whose frequencies are entirely within the frequency interval $\Delta\nu$, differs from a simple tone only by the regular or irregular modulation of frequency and amplitude. The sound

---

*Translator's note: As in an instrument tuned to the equitempered scale.

sensation remains that of a simple tone. However, if the frequencies of sine components which are very close together are distributed over a continuous frequency range which is considerably wider than the frequency interval of the ear, the tone or sound assumes the character of noise. The pitch of this noise may nevertheless remain sufficiently distinct for the ear to form melodies. It was found that noises whose frequency interval does not greatly exceed a third can still be used for forming melodies. Physically and psychologically it does not make any difference whether such a noise is composed additively of individual sine oscillations sufficiently close together or whether it is selectively separated by a filter from a broad-band noise (e.g. white noise). Both result in metameric sensations. On the other hand, a filter of sufficiently narrow pass-band (less 20 c.p.s.) may be used to produce a simple tone from white noise. The variations of frequency and amplitude observed in the oscillogram of the respective phenomenon of oscillation then occur in periods longer than 50 milliseconds hence they do not affect the character of the tone.

Physically Pure Sounds Without Fundamental Tone

The electronic sound producers permit the production of sounds, which can be perceived by mechanical acoustics only with difficulty. The filtering of a sound, for example by means of a third- or octave-filter, is one of the simplest, but nevertheless most effective, measures. A sound containing many overtones, i.e., a periodic impulse sequence or a sawtooth oscillation, would be the most suitable initial tone.

In sounds filtered in this manner a sensation phenomenon, which normally is not noticed in conventional musical instruments, although it can be detected here in some cases, is particularly distinct.

It was found that a sound, which passes a third- or octave-filter, has two coincident different pitch qualities. One

pitch (the "residual tone") corresponds to the fundamental frequency of the initial material used (e.g. the impulse frequency) while the other pitch (the "formant tone") is determined by the pass-band of the filter. If a series of impulses with a fundamental frequency of 440 c.p.s. (a_1) is passed through a third-filter having a pass-band of approximately 2,250 to 2,850 c.p.s., the sound thus perceived has the quality of both the a_1 and the e_4 (2,640 c.p.s., i.e., the strongest harmonic of the filtered sound). Which of the two tone qualities is heard more intensely depends on the musical context. It is difficult to determine the octaval position of such filtered sounds having two tone qualities. The musical context is of decisive importance here as well. The distinctly accordant effect of monophonic series of filtered sounds is very remarkable indeed. The inexpert listener is readily inclined to attribute major and minor harmonizations to such a monophonic series.

6. Non-Steady Processes

Simple tones whose amplitude frequency of phase changes abruptly are now dealt with.

The reasonable assumption that a simple tone also releases a simple tone sensation when its amplitude, frequency or phase changes abruptly is not true. A simple tone, which is suddenly initiated at the time t = 0 (i.e., without physical building-up processes), and which then continues for e.g. 1/10 sec. with undiminished intensity (Fig. 4a), has not- as might be assumed - the time-frequency spectrum shown in Fig. 4b. Although the manner in which this process is produced suggests the spectrum shown in Fig. 4b, the tone has nevertheless the spectrum shown in Fig. 4c. The sudden start and failure of the oscillation impulse brings about a considerable widening of the time-frequency spectrum, which can be heard as a "click" (i.e., as a physiological transient phenomenon). However, owing to the overriding of the preceding stationary part of the tone and because of the blur caused by the

reverberation of the room in which the tone is produced, the final click is heard less intensely than the initial click.

Hence a certain time* passes after the initiation of a simple tone before the ear can distinguish the pitch with certainty. This time constitutes a more or less large fraction of the analysis interval corresponding to the definition of the "distinguishing threshold". Particularly the investigation of K. Schubert(14) showed that a constant analysis interval can only be expected in a mean frequency range of approximately 2,000 c.p.s. For higher and lower frequencies the duration of the interval increases. The table below (after Schubert) gives the time required for distinguishing a simple tone at various frequencies. The values are mean values and the variation between different observers is considerable.

Frequency in c.p.s.	100	200	500	1000	2000	3000	4000	6000
Time in m.sec.	45	30	26	20	13	14	14	18

The above observations may easily be extended to cover any sound phenomenon. It is found that any sudden change of an oscillation parameter (e.g. of the amplitude frequency or phase) results chiefly in the widening of the time-frequency spectrum. The more abrupt the transition from one state of oscillation to another the more marked will be the widening of the time-frequency spectrum. Therefore, the effect of such a step ranges from the blurring of the pitch to the production of noise.

In order to reduce the proportion of noise the transitions must be sliding ones. When the band is mounted the ends

---

*Therefore, as G.A. Miller assumes, a "short pure tone" is a contradiction in itself. A short tone cannot be pure and a pure tone cannot be short.

of the bands are cut not at right angles to the band edge but obliquely. Crackling at the initiation of electrically controlled tones or sounds may be prevented by employing special circuit parameters. If a hard starting and failing process is passed through a filter whose pass-band is narrower than the spectrum, which has been widened owing to the discontinuous changes in oscillations, then the physiological building-up and dying-out processes are suppressed physically. Hence the tone and sound sensations start and fade gently.

There still remains the question of how fast an oscillation parameter may change without resulting in blur and noise.

O. Sala(12) solved this problem as far as the amplitude of oscillation is concerned. If the logarithmic decrement Λ* is used as the criterion, it is found that it is possible to determine the pitch with accuracy when $\Lambda = 0.26$, but precise pitch data can no longer be given when $\Lambda = 0.96$.

7. Reproduction Diagram and Spectrum

The above statements concerning the physiological transient phenomena and the tone and sound sensations coordinated with them touch upon a field that is of decisive importance for electrical sound production, i.e., the relationship between the "material", which is recorded on the sound track in accordance with the directions (score, etc.) of the composer, and the time-frequency spectrum of this material. Only the time-frequency spectrum is directly related to the sound sensations.

---

*If two successive amplitudes of oscillation in the same direct-tion are denoted by a_1 and a_2, then

$$\Lambda = \ln (a_1/a_2).$$

A logarithmic decrement

$$\Lambda = 0.115\ A/\nu$$

(where ν is the frequency of the oscillation) corresponds to a constant variation of amplitude of A db./sec.

Both the material and the relevant spectrum may be represented by a plan system of coordinates if the blacking method (Fig. 2b) is used for the characterization of the amplitudes. The time slope of the sensations and perceptions may possibly be represented in a similar manner. However, it is in the nature of things that only a vague idea exists of how to represent the psychic correlative of a time-frequency spectrum graphically. Therefore, we confine ourselves to discussing the well-defined mathematical relations between the reproduction diagram of a composition, i.e., the graphically fixed "score" (in the most general sense), and the time-frequency spectrum, which permits suitable characterization of this reproduction diagram as soon as a certain interval has been decided on.

The most important relationships are shown by means of a number of examples. Regarding the mathematical formulation the reader is referred to the appendix.

1. The manner in which a hard start and failure of a sine oscillation is reflected in the spectral region is shown in Fig. 4. If two sine oscillations of different frequency alternate in rapid succession (frequency modulation, "alternating tremolo") - this can easily be accomplished with an electron switch - then the more the frequency of alternation exceeds the characteristic frequency interval of the ear the more effective will be the blur (Fig. 5). The same applied to periodic variations of amplitude at maintained frequency (amplitude modulation, vibrato). Hence frequency modulation and amplitude modulation may be used for making sounds diffuse with respect to frequency. If the modulation frequency is higher than the so-called "phase-limiting frequency", then the sensations caused by frequency modulation and amplitude modulation with small modulation depth are both metameric. The phase-limiting frequency depends on the mean frequency of the modulated sine oscillation ("carrier frequency") approximately as follows (after E. Zwicker[18]):

Carrier frequency	60	250	1000	4000	8000 c.p.s.

Phase limiting frequency	30	40	80	400	1000 c.p.s.

2. The time-frequency spectrum of a sine oscillation, which changes its frequency rapidly but continuously (glissando), by no means faithfully reflects the curve form of the frequency of the oscillation; on the contrary, the result is again a blurred, noisy frequency band. If the frequency is permitted to rise or drop rapidly, the result will be a crackling noise. The same result is obtained if only a narrow frequency range (approximately up to octave width) is filtered out from the time-frequency spectrum of the sliding note. If the filter range is low (e.g. between 75 and 150 c.p.s.), a downward glissando sounds like a dull thump.

3. If a small piece of a sound track, on which a sine oscillation is stored magnetically, is cut out and the two cut ends are stuck together again, the result usually is a shift in phase, causing a widening of the spectrum at this point and a noise sensation.

8. The Effect of Room Acoustics

Insofar as the electronic compositions are not heard with the aid of headphones or in a non-resonant room, there will be an additional temporary blur whose extent depends on the acoustical properties of the reproduction room. For example, the reverberation blurs all the sudden transient phenomena during a sustained or falling intensity of sound. The longer the reverberation period of the room the greater will be the extent of this blurring. Natural or artificial reverberation is the correct method of eliminating the technical rigidity of an electronically produced sound and of maintaining this sound diffuse in the time coordinate. Inasmuch as the processes involved show not only dynamic, but also pitch and timbre structure, the temporary blur brings about a blur with respect to frequency and thus noise sensations. Hence reverberation may be utilized for

converting sounds into noises.

9. Compression and Expansion of Frequency

The increase and decrease of the track speed presents one of the most characteristic shaping possibilities in electronic sound track composition. If the track speed is increased by a factor γ, all the frequencies increase by the same factor*, while the period of the recorded phenomena decreases to $1/\gamma$.

A decrease of track speed has the opposite effect.

However, the time-frequency spectrum by no means changes in a similarly simple way, since the analysis interval of the ear remains invariant to all transformations of frequency and time. Therefore, a compression or expansion of frequency is associated with changes in sensation, which exceed the effect of a mere frequency transposition (in the sense of musical science) and changes in tempo. For example, an increase of track speed results not only in higher frequencies and shorter periods of time, but also in a stronger accentuation of the noise component and in increased blending. Non-steady processes of the rhythmic type may thus become steady by increasing the track speed, i.e., they may cause a noise sensation which is no longer differentiated rhythmically. Finally, if the increase of track speed is great enough, practically all sound processes change to a tweeting noise. Inversely, a decrease of track speed has the result of causing steady tones, sounds and noises to dissociate into non-steady individual processes, since the analysis interval of the ear does not suffice, in this case, to form from the individual phenomena a mean value independent of time.

---

*The muscial intervals (third, fourth, octave, etc.) remain unchanged here.

10. The Spectrum in the Case of Multiplicative Mixture

A score of the conventional type always implies the unwritten rule: "The recorded voices must be superimposed, i.e., they must be combined by mere addition of the respective acoustic pressure distributions." No doubt for acoustic musical instruments no other type of combination is possible, but in an "electronic" score any type of combination between the recorded voices may be expected, since this can be accomplished with the aid of communications technique. For example, if the composer specifies: "Voice 3 to be mixed multiplicatively with voice 5, then this is a directive for the acoustical engineer to combine the two voices, using a suitable method of multiplicative mixture (e.g. a ring modulator).

However, the sound effect of this multiplicative mixture is not obtained directly from the score recordings but only by means of the time-frequency spectrum. A graphical method, which can easily be used in practice, is described below. By means of this method, and without any calculation, a spectrum of multiplicatively mixed simple forms of oscillation may be visualized(6).

Two steady sine oscillations of 600 and 240 c.p.s. are under consideration. Question: What are the frequencies resulting from the multiplicative mixture of these two oscillations? First, the two values 600 and 240 c.p.s. are marked on the scales ν_1 (first frequency) and ν_2 (second frequency) in the nomogram shown in Fig. 6, both on the left-hand side and the right-hand side of the scales, which are numbered symmetrically. Then the four plotted points are joined by straight lines and extended up to the intersection with the ν_r scale (resulting frequencies). The result can be read in this scale: The initial frequencies have been replaced by two new frequencies of 360 and 840 c.p.s., respectively. Of course, this result might have been obtained quite as easily by calculation, for 840 c.p.s. = 600 c.p.s. + 240 c.p.s. (the summation frequency of the two components) and

360 c.p.s. = 600 c.p.s. - 240 c.p.s. (the differential frequency of the two components). However, the superiority of the graphical method immediately becomes evident in the case of tone mixtures and sounds. An example of the multiplicative mixture of two tone mixture, each consisting of two components, is shown in Fig. 7. If all the connecting lines possible are drawn with care, errors occurring in calculations are practically impossible.

The tone mixture obtained from the multiplicative mixture of two sounds or of a pure tone and a sound is usually quite involved. However, sounds of remarkable and unusual spectroscopic structure are obtained if a rational ratio of the fundamental frequencies (i.e., of the first harmonics) of the two sounds to be mixed is selected. The harmonics of the resulting oscillation occurring at the lowest rational frequency ratios (1:2, 1:3, 2:3, etc.) are listed in the table below, assuming that a sine oscillation of the frequency ν_1 is multiplicatively mixed with a periodic oscillation of the fundamental frequency ν_2, which contains all the harmonics uninterruptedly from the first harmonic (fundamental oscillation). As assumed, the frequency ratio $\alpha = \nu_1/\nu_2$ is rational, i.e., it may be expressed in the form of a reduced fraction, viz.,

$$\alpha = m/n,$$

where m and n are integers. The fundamental of the resulting frequency, i.e., the frequency of the periodicity, but not the lowest Fourier component, then equals ν_1/m and ν_2/n, respectively, and only harmonics of the order

$$h = [kn \pm m] \quad (k = 1, 2, 3, 4\ldots)$$

occur.

Table

Listing the harmonics occurring when a simple tone and a sound are mixed multiplicatively

Fre-quency ratio	Musical interval starting from a simple tone	Only harmonics of the following ordinal numbers occur:														
Frequency of the simple tone lower than the fundamental frequency of sound:																
1:2	octave up	1		3		5		7		9		11		13		15
1:3	duodecimo up		2		4	5		7	8		10	11		13	14	
2:3	fifth up	1			4	5		7	8		10	11		13	14	
1:4	double-octave up			3		5		7		9		11		13		15
3:4	fourth up	1				5		7		9		11		13		15
2:5	tenth up			3				7	8				12	13		
3:5	major sixth up		2					7	8				12	13		
4:5	major third up	1					6			9		11			14	
Frequency of the simple tone higher than the fundamental frequency of sound:																
2:1	octave down	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3:1	duodecimo down	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3:2	fifth down	1		3		5		7		9		11		13		15
4:3	fourth down	1	2			5		7	8		10	11		13	14	

If the sound is kept constant, e.g. at a fundamental frequency of 300 c.p.s. while the tone traverses all the frequencies (starting from 0 c.p.s.) in glissando, the time-frequency spectrum of the product of the two oscillations shows a structure made up of intersecting components (Fig. 8). The more rapidly the glissando is performed the more closely will the result approach a noise.

The multiplicative mixture of a simple tone and filtered noise may also be evaluated easily with the aid of a nomogram (Fig. 9). It is sufficient to draw the noise

diagrammatically as a broad spectral band without inner structure. The mixture results in a doubling of the number of noise bands together with a frequency distortion (transposition). No difficulties are encountered as long as the noise band is not shifted, owing to the transposition, close to the zero point of the frequency, or even beyond it, i.e., as long as the frequency of the tone does not fall within the frequency range of the noise. As soon as the noise bands overlap, the present simplified considerations no longer apply.

By mixing the noise and a sliding tone multiplicatively a sliding noise (howling) with two spectral regions is obtained. If required, one of these regions may be eliminated by suitable filters.

The multiplicative mixture of two noise bands gives similar results as long as the resulting bands do not overlap. However, the latter bands will be broader than the initial noise bands even without overlapping.

The Deutsche Forschungsgesellschaft supplied the apparatus for the experimental investigations described above. The results obtained were taken into account when the Cologne studio for electronic music was equipped.

Appendix (Mathematics)

In order to obtain a concrete idea of the time-frequency analysis, it is assumed that the electric phenomenon of oscillation F(t) to be analysed is simultaneously distributed over a series of band filters whose pass-bands overlap slightly. The widths of the pass-bands is assumed to be constant and is denoted by $\Delta\nu$. The points of maximum transmittance of the individual filters, i.e., the analysing frequencies coordinated to these filters are at ν_1, ν_2,...ν_K... etc. The extent of the analysis interval definition determines the band width of the filters

according to equation (1). If the voltage variation F(t) is conducted to the filter with the analysing frequency ν_K at the input, then it undergoes a deformation such that the output voltage E(t) in the main contains only frequencies close to ν_K. Mathematically this deformation is expressed by the following convolution of integrals

$$E_{\nu_K}(t) = \int_{-\infty}^{t} F(\tau)H(t - \tau) \sin 2\pi\nu_K(t - \tau)d\tau. \tag{A1}$$

Here the function H indirectly characterizes the variation of the filter transmittance and hence also the band width. It is called the system function of the filter. In order to bring the system function into conformity with the measured properties of the ear, it is expedient to choose an exponential set-up:

$$H(t - \tau) = \begin{cases} \Delta t \cdot e^{-\frac{(t - \tau)}{\Delta t}} & \text{for } t > \tau \\ 0 & \text{for } t < \tau \end{cases} \tag{A2}$$

As mentioned in paragraph 4, the analysis interval Δt has a length of approximately 25 milliseconds.

The oscillation function $E_{\nu_K}(t)$ represented by equation (A1) contains the time t as the variable and the analysing frequency ν_K as the parameter.

Now, assuming increasing multiplication of the filters covering the audibility range following one another in correspondingly closer succession, then there will ultimately be a corresponding filter for any analysing frequency ν so that a continuous spectrum replaces the discrete filter output voltage $E_{\nu_K}(t)$. This continuous spectrum is the time-frequency spectrum of the process F(t). Therefore, taking into account equation

(A2), equation (A1) is generalized to

$$E(\nu,t) = \Delta t \int_{-\infty}^{t} F(\tau) e^{\frac{\tau - t}{\Delta t}} \sin 2\pi\nu(t - \tau) d\tau$$

$$= \Delta t \cdot \sin 2\pi\nu t \int_{-\infty}^{t} F(\tau) e^{\frac{\tau - t}{\Delta t}} \cos 2\pi\nu\tau \, d\tau \qquad (A3)$$

$$- \Delta t \cdot \cos 2\pi\nu t \int_{-\infty}^{t} F(\tau) e^{\frac{\tau - t}{\Delta t}} \sin 2\pi\nu\tau \, d\tau$$

If the right-hand side integrals are abbreviated to

$$f_c(\nu,t) = \int_{-\infty}^{t} F(\tau) e^{\frac{\tau - t}{\Delta t}} \cos 2\pi\nu\tau \, d\tau \qquad (A4)$$

and

$$f_s(\nu,t) = \int_{-\infty}^{t} F(\tau) e^{\frac{\tau - t}{\Delta t}} \sin 2\pi\nu\tau \, d\tau \qquad (A5)$$

then equation (A3) assumes the following simplified form

$$E(\nu,t) = \sqrt{f_c^{\,2}(\nu,t) + f_s^{\,2}(\nu,t)} \sin\left(2\pi\nu t - \tan^{-1} \frac{f_s(\nu,t)}{f_c(\nu,t)}\right)$$
$$= f(\nu,t) \sin (2\pi\nu t - \Psi(\nu,t)), \qquad (A6)$$

i.e., with respect to the time coordinate, the complete time-frequency spectrum presents an oscillation with modulated amplitude and phase[7]. These conditions can readily be observed by means of sonograms which were obtained by K. Imahori[4] with a photographic raster analyser. Fig. 10 gives such an example.

If the "carrier oscillation", which is periodic with the analysing frequency ν, is separated, the time-frequency spectrum may be represented by a complex envelope curve:

$$f(\nu,t) = f(\nu,t)e^{i\Psi(\nu,t)}. \qquad \text{(A7)}$$

For present purposes it is sufficient to take into account its absolute value, i.e., the function $f(\nu,t)$ alone, as in Figs. 2 to 5. This representation may be called "simplified time-frequency spectrum".

References

1. Bennett, W.R. The correlatograph. A machine for continuous display of short term correlation. Bell Syst. Tech. J. 32: 1173-1185, 1953.

2. Gabor, D. Problems of speech analysis. U.R.S.I. Xth General Assembly, Sydney, 1952, Doc. No. 314 Comm. VI.

3. Huggins, W.H. A phase principle for complex-frequency analysis, and its implications in auditory theory. J. Acoust. Soc. Amer. 24: 582-589, 1952.

4. Imahori, K. Analysis of varying sounds. J. Fac. Sci. Hokkaido Imp. Univ., Ser. II, 3: 57-90, 1940; 3: 103-126, 1941.

5. Koenig, W., Dunn, K.H. and Lacy, L.Y. The sound spectrograph. J. Acoust. Soc. Amer. 18: 19-49, 1946.

6. Meyer-Eppler, W. Ein Abtastverfahren zur Darstellung von Ausgleichsvorgängen und nichtlinearen Verzerrungen (Scanning method for the representation of transient phenomena and nonlinear distortions). A.E.Ü. 2: 1-14, 1948.

7. -: Experimentelle Schwingungsanalyse (Experimental vibrational analysis). Ergeb. exakt. Naturw. 23: 53-126, 1950.

8. -: Die Exhaustions-Schwingungsanalyse und ihre Beziehung zur Theorie des Hörens (The exhaustion analysis and its relationship to auditory theory). Physik. Bl. 7: 355-361, 1951.

9. Michels, W.C. and Helson, H. Man as a meter. Physics Today, 6: 4-7, Aug. 1953.

10. Pollack, I. The information of elementary auditory displays. J. Acoust. Soc. Amer. 24: 745-749, 1952.

11. Ranke, O.F. and Lullies, H. Gehör - Stimme - Sprache (Hearing - voice - speech). Berlin/Göttingen/Heidelberg. Springer-Verlag, 1953.

12. Sala, O. Psycho-physische Konsequenzen elektro-akustischer Klangsynthesen (Psycho-physical consequences of electro-acoustic sound synthesis). Frequenz, 5: 13-20, 1951.

13. Schaeffer, P. A la recherche d'une musique concrète. (Editions du Seuil) Paris, 1952.

14. Schubert, K. Über die Prüfung des Tonhöhenunterscheidungsvermögens. (Tests of the capacity of the ear to distinguish pitch). Arch. Ohr- usw. Heilk. u. Z. Hals- usw. Heilk. 159: 339-354, 1951.

15. Skudrzyk, E. Die Bedeutung der Ausgleichsvorgänge für Musik und Tonübertragung. (The importance of transient phenomena for music and sound transmission). E. u. M. 67 (9 and 10), 1950.

16. Trendelenburg, F. Einführung in die Akustik (Introduction into acoustics). Second edition Berlin/Göttingen/ Heidelberg. Springer-Verlag.

17. Ward, W.D. Information and absolute pitch. J. Acoust. Soc. Amer. 25: 833, 1953.

18. Zwicker, E. Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones (The limits of audibility of the amplitude modulation and frequency modulation of a tone). Acustica, 2: AB 125-AB 133, 1952.

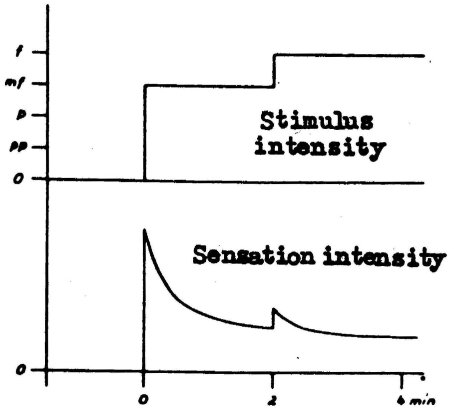

Fig. 1

Diagram showing stimulus intensity and sensation intensity for an initial mf passage with a sudden increase of the stimulus intensity to f.

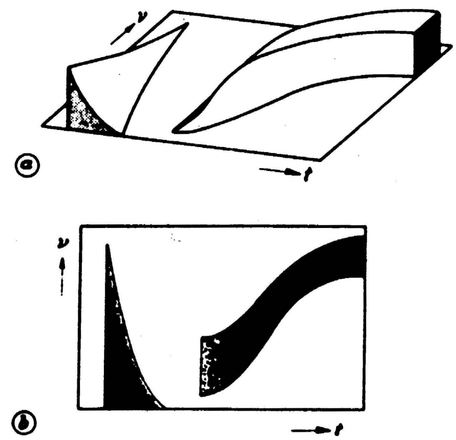

Fig. 2

Time-frequency spectrum:

(a) As a perspective drawing

(b) Represented by the variable-density method.

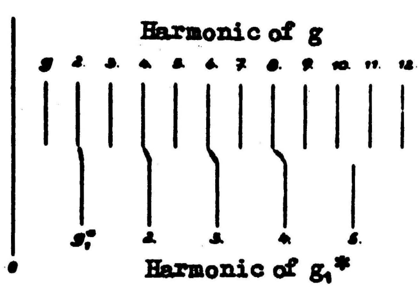

Fig. 3

Beats of amplitude (hatched lines) between the harmonics of two sounds g and g_1* (= dissonant g_1).

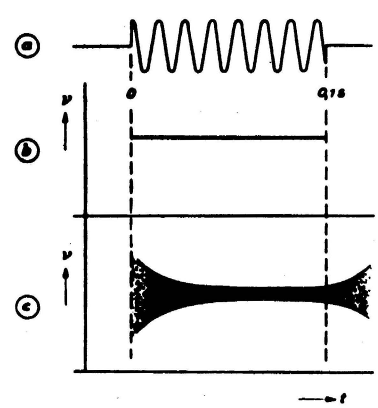

Fig. 4

Short tone:

(a) Oscillogram

(b) Reproduction diagram and presumed time-frequency spectrum

(c) Correct time-frequency spectrum (simplified).

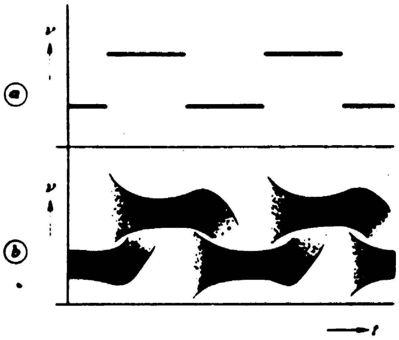

Fig. 5

Alternating tones
(alternating tremolo):

(a) Reproduction diagram

(b) Relevant time-frequency spectrum (simplified).

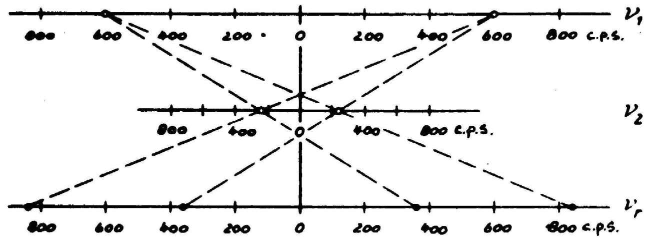

Fig. 6

Nomographic determination of the frequencies occurring in multiplicative mixture of two tones.

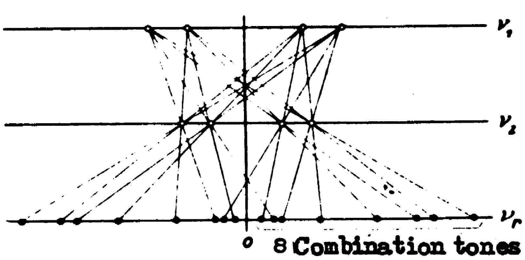

Fig. 7

Nomogram of the multiplicative mixture of two two-component tone mixtures.

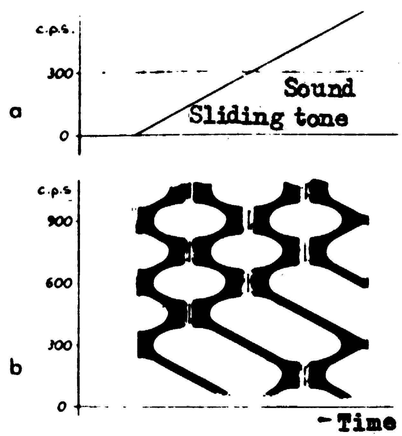

Fig. 8

Multiplicative mixture of a steady sound and a sliding tone:

(a) Reproduction diagram

(b) Spectrum (simplified).

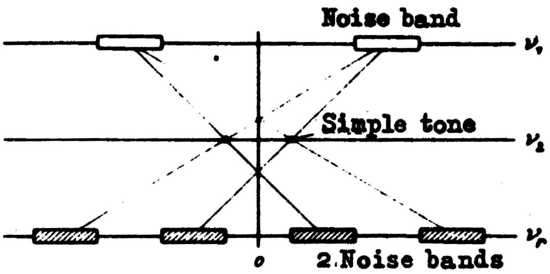

Fig. 9

Nomogram of the multiplicative mixture of filtered noise and a simple tone.

Fig. 10

Complete time-frequency spectrum (after Imahori).