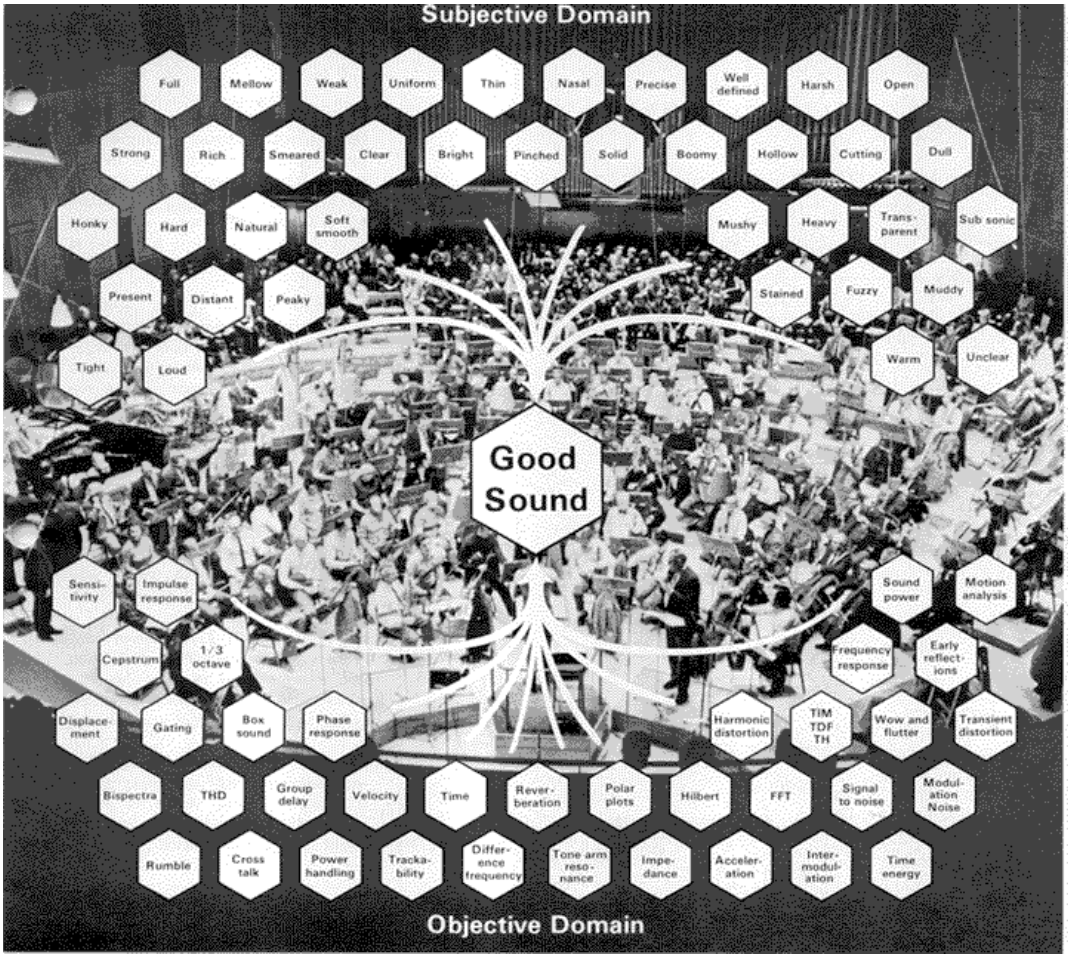


Multidimensional Audio
by *Henning Moller,* Bruel & Kjaer

1 . Introduction

What is Audio all about? Subjectively, the answer is easy. It is literally a question about good sound. In practice the human mind can tell, within seconds, if a sound picture is correct or not, just as quickly as it can tell whether a girl is beautiful or a house, a car or a landscape is impressive.

Human beings consider things in a "global" fashion - everything is registered and perceived simultaneously, but no details are clear to begin with. However, when we measure, we do exactly the opposite we describe details with extreme accuracy. We concentrate on one parameter at a time in a "local" fashion.

We could accurately measure how tall the girl is, what colour her hair is and so on, but that doesn't directly tell us how beautiful she is.

Likewise, on a Hi-Fi system we could, for instance measure frequency response and harmonic distortion, but neither does that tell us whether the system is good or bad.,

Things that are easily and intuitively perceived, like Audio, are generally extremely hard to explain. It requires many words, many measurements, foreknowledge and interpretations.

In Fig.1 [above] we have listed some of the many "local" parameters that people use today in order to describe the "global" phenomena "good sound".

Subjective Domain

Some people operate primarily in the "Subjective domain". Studio people, musicians, Hi-Fi fans etc. use a lot of words. Often it has similarities to religious sects where only the initiated themselves understand the words.

Other people operate primarily in the "Objective domain". They measure with extreme accuracy - for instance, 0,001% harmonic distortion at 1 kHz - and they insist that because of that the Hi-Fi system must be good. Others claim that because the system is phase compensated it is tremendous, while others maintain that because it is TIM-free it must be the best.

Really, both groups of people are talking about the same thing, but from different points of view. Nevertheless there is often open warfare between the "subjective people" and the "objective people". People that really have been listening intensively for years often say that measurements are absolutely useless because they judge from the few oldfashioned measurements they might know. And often the socalled objective engineers say that, for instance, reviewers are crazy because they judge contrary to the measurements.

Let's try to examine some of the words and measurements (Fig.1) that people use today.It is essential to note that none of the local parameters - let's call them one-dimensional - around "good sound" are important alone, because they are only describing a limited part of the global totality that consists of all the subjective domains as well as all the objective domains. "Good sound" must be a simultaneous combination of, in principle, an infinite number of onedimensional domains into a multidimensional meaning.

It may sound rather complex, but really it is remarkably simple since this is precisely how the human mind operates. It takes virtually no time to decide whether a sound is good, since everything is perceived and comprehended simultaneously - just like the impression of the girl

What is required is, as indicated by Richard Heyser (Ref.1), a localglobal mapping. This is a mathematical transformation from each "onedimensional" parameter to all the others. Fortunately, most people make this transformation many times every day without using a bit of mathematics.

A simplified example of this is the Fourier Transform that takes all points in time from - ∞ to + ∞ and maps them into one point in the frequency domain. Or the opposite, that a single event in the time domain, a transient, is mapped into all points in the frequency domain. In other words into a flat spectrum from - ∞ to +∞ . Likewise, each cell in the human body contains all the inherited information about that person. The chromosomes in the big toe and nose are identical.

Obviously, this global-local mapping is of extreme importance and has broad applications outside Audio. It could be considered as a law of nature or a philosophy as further described in section 10 about "Apodization".

Foreknowledge and "apodization" are essential for meaningful human evaluations as well as for relevant measurements. That we can understand things in a "global fashion" is only because just a few "bits" are required to complete an already preprogrammed picture. We could not, for instance, tell if there is distortion in a Hi-Fi system by listening if we did not know what music was supposed to sound like.

Future instruments might, therefore, also be preprogrammed with information about the object they are supposed to measure. A further step would be to make adaptive instruments whose ability to measure will improve with age and experience. Actually it would be reasonable if the Hi-Fi measuring device spent a few months in the concert hall listening to live music before it was used. 1 Mbit memories [in 1977 !!!] will soon be available on one chip, so technologically this will be possible in a few years. Various weighting functions for the individual measurements might also be preprogrammed in order to obtain meaningful measurements.

2. Audio Development and Philosophy

Before the days of so-called Hi-fi systems there were no audio measurements, there were only natural sounds with a perfect signal-tonoise ratio, unlimited power handling capability, no distortion of any kind, but the number of people with a possibility of ever listening to music was rather limited. Then came Hi-Fi. The invention of the phonograph record ruined high fidelity, but made music universally available.

In the past when the number of words and measurements was still rather limited there was no confusion, but neither was there any correlation. Today we have probably just passed the point where a correlation between objective measurements and subjective parameters is possible. This paper will conclude that six "measurement domains" today strongly correlate to the subjective perception of Audio, while obviously no single measurement is sufficient.

In the future we might - as indicated in Fig.2 - end up with multidimensional subjective domains related to subjectively weighted objective domains. This will - when properly interpreted - give good correlation, but it might create some kind of confusion to begin with.

It would be exciting if we could understand everything in the world 100% means that the complete Hi-FI directly and therefore we concenaround us. From a measurement piont oh view , it would be ideal if we could measure everything in just one measurement, preferably using a box that costs less than 10 dollars and graduated in % where 100 % means that the complete HI-FI system is perfectly natural, 90% that it is very good, 80% that is quite good and 50% that it is acceptable, etc. Unfortunately we can not do that directly and therefore we concentrate on relatively small details that we think we can understand. However different people are are concentrating on different things and therefore there is a confusion when making comparisons.

3. Steady State Distortion

Take, as an example, the discussions about total harmonic distortion (THD) and consider (as indicated in Fig.3) the soft clipping of a tape-recorder compared with the 10% THD, 1 kHz a tape-recorder compared with the cross-over distortion in an amplifier at 1 kHz. The good old THD will typically show up to 10% for the tape recorder,but as little as 0,01% for the amplifier, but as little as 0,01% for the cross-over distortion. Therefore, looking on this measurement alone,one should think that crossover distortion is 1000 times better than the soft clipping. Audibly, however,it is actually the other way around.

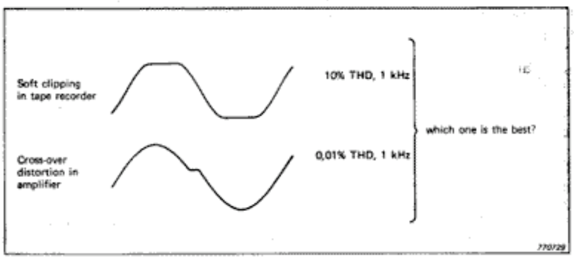


Fig.3. Is the soft clipping of a tape recorder with 10% THD better than cross-over distortion in an amplifier with 0,01% THD?

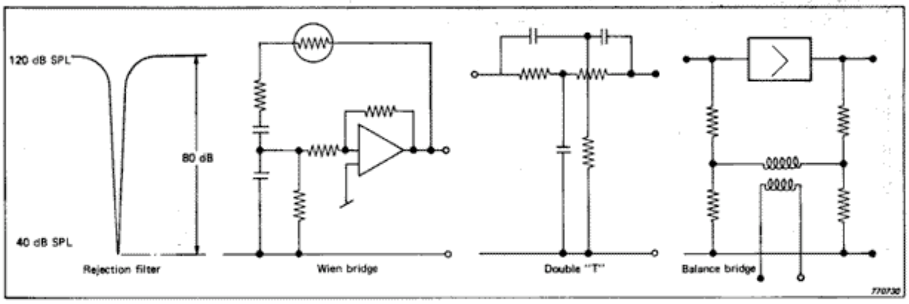


Fig.4. THD even to 0,0001% distortion is easy to measure but literally worthless from a subjective point of view

THD at 1 kHz is *very* easy to measure as indicated in Fig.4, but it normally gives completely misleading results. Making such an instrument is so easy that virtually anyone can make it himself. A Wien-bridge oscillator and a double "T" take care of the problem. The "T" typically rejects 40-60 dB (0,1%) and does not even require power. If it is required to measure lower than 60 dB it is easy as long as there is no delay in the system. Four resistors in a balanced bridge will (as indicated in Fig.4), reject the signals that are present on the input as well as on the output. Since the distortion is only present at the output it is not rejected. This rejection will most likely give 40 dB on top of the previous 60dl3, or in other words a dynamic range of at least 100 dB (0,001% distortion). It costs next to nothing, but the value of the information is also rather limited.

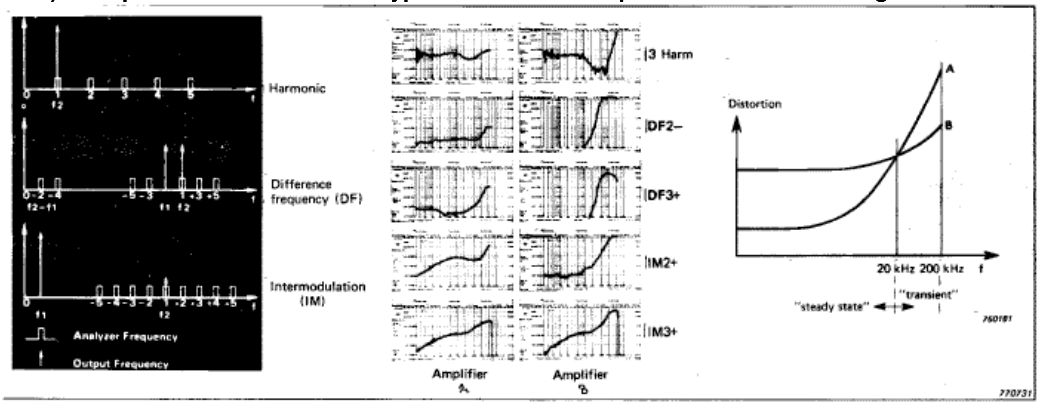
If one more dimension is added to the THD measurement it looks a little more reasonable - that is to measure THD as function of frequency. This is slightly more complex from an instrumentation point of view since it requires stable amplitude while sweeping. This is normally obtained with a beat frequency oscillator (BFO) and it requires accurate tuning of the filter which is normally obtained with a heterodyne filter. These possibilities are described in the B & K Electro Acoustic Measurements 16-035 (Ref. 2).

Unfortunately traditional swept THD measurements are difficult if there is delay in the system such as in speakers, microphones and tape recorders. Here the filter has already moved when the test frequency arrives and therefore there is almost no rejection although the filter specification might seem very nice. Moreover it is completely useless in acoustic systems as indicated in Fig.4, since the background noise level typically is 40 de which means that the loudspeaker should operate with 120dB SPL if the 80 dB rejection is to be of any use. At those levels the distortion will obviously exceed 0,01% anyway so the 80 dB "dynamic range" is useless. Therefore, also in instrumentation more dimensions are required.

Another strong disadvantage of THD, even as a function of frequency, is that the audible importance of the different components is different. Typically, the even harmonics sound quite reasonable while the odd harmonics sound pretty bad. The next step in obtaining subjectively relevant measurements is therefore to measure the individual harmonics as a function of frequency (see again Ref.2). Typically, relatively high levels from the 10th to 20th harmonic seem to correlate well with the crossover distortion previously mentioned. Still, this is very limited information relative to music since it does not measure the interference between different components in a complex music signal. Assume, for instance, that the harmonic components in the audible range are low arid therefore that means that a violin should sound perfect. A good question now is: what happens to the sound of the violin if suddenly another violin starts? In the concert hall nothing happens except that one hears two violins, but on a HiFi system the presence of the second violin often completely changes the performance of the first one.

Another good question that even the measurement of the individual harmonic distortion components as function of frequency does not tell anything about, is what happens to the sound of the violin if suddenly the bass drum starts? Again, in the concert hall, nothing happens except that a bass drum and a violin are heard, but in the Hi-Fi system the performance of the violin is normally completely changed by the presence of the bass drum.

Fortunately these effects are also easily measurable today using the B & K 1902/2010/2307 combination. This gives up to 18 different distortion curves all as function of frequency in the range 2 Hz-200 kHz and with a dynamic range of at least 80 dB (Ref.3). The possibilities and some typical results on amplifiers are shown in Fig.5.



ig.5. By overviewing the results a meaning becomes visible. Amplifier B has relatively low distortion at low frequencies while outside the traditional frequency range it is rather high relative to amplifier A. This reveals transient distortion

This is just one example of a tremendous increase in the amount of measuring data that might result in confusion compared with the simple number in % THD obtained at 1 kHz. However, the important thing when evaluating this much data is to get an "overview" of the results so a "meaning" becomes apparent.In Fig.5 the curves are deliberately rather small so it is virtually impossible to see the details. What is left is only the really essential part of the information. The two columns represent curves for two different amplifiers. The one on the left side (A) is one with rather moderate feedback of about 30dB while the one on the right side (B) has rather heavy feedback of about 70 dB. The horizontal lines represent the different kinds of distortion measured up to 200 kHz.

4. Transient Distortion

The essential thing is that the curves from amplifier A (Fig.5) have a rather high level at low frequencies, say up to 20 kHz, while it does not increase so much at higher frequencies. Amplifier B has very low - almost unmeasurable - distortion in the traditional frequency range 20 Hz - 20 kHz while above 20 kHz it increases considerably to as high as 10%. This is illustrated in the right hand part of Fig.5.

So far we have tried to describe the multi-dimensionality in Audio from the distortion point of view. If we call the THD at 1 kHz measurement "one-dimensional" we could call the swept THD measurement "two-dimensional". The swept individual harmonic measurement would then be "three-dimensional" and finally the swept individual twotone distortion curves would be "four-dimensional". Obviously, by adding more dimensions the potential value of the information increases tremendously, but some kind of interpretation is required if confusion is to be avoided.

More "dimensions" can be obtained, for instance, by expanding the dynamic range or the frequency range. For many years people have been trying to expand the dynamic range alone, by measuring down to say 120dB below the signal level or 0,0001% distortion. However, typically only in mid-frequency range around 1 kHz where practically no correlation is found with subjective results. This kind of measurement is relatively easy, as indicated in Fig.4, using the bridge arrangement, but the information is literally worthless when it is not seen together with other dimensions. The frequency range above 20 kHz is one important "dimension" and the frequency range below 20 Hz is another important one (see section 5).

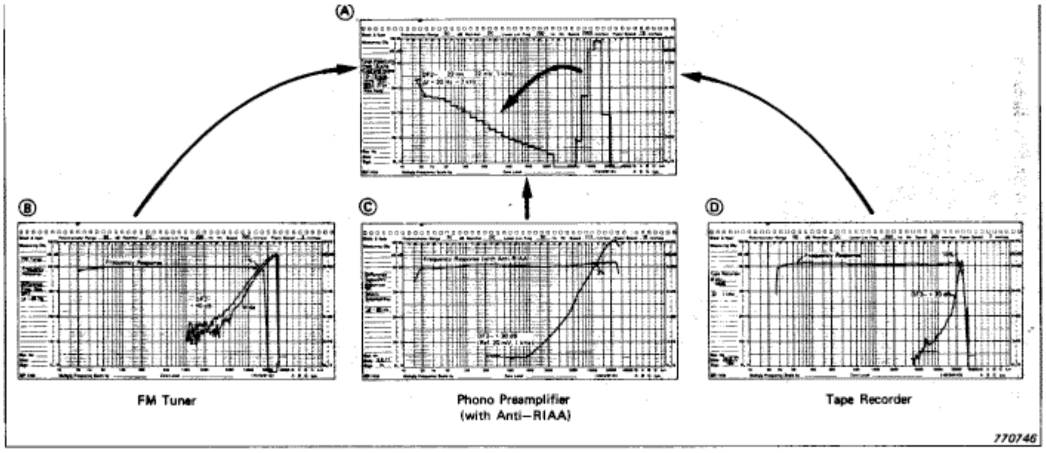


Fig.6. Transient distortion in FM tuners, phone-amplifiers and tape recorders is revealed by high frequency distortion which creates products down in the audible range. Typically 10% distortion

However the various domains interact. For example, if the input level is increased, the distortion that previously was only visible above 20 kHz, will now also be visible in the traditional 20 Hz - 20 kHz range.

Other parts of the Hi-fi system might also interact by being more sensitive to the same kind of tests. Fig.6 shows that exactly the same trend as in amplifiers is found in FM-tuners, Phono-preamplifiers and Tape Recorders, and is actually more pronounced.

A kind of total influence of this type of distortion is indicated in the upper part of Fig.6. Here a phonopreamplifier is fed from a RIAA preemphasis network simulating the conditions from a music record. The reference level is 20 mV at 1 kHz which today is quite typical, at least for transients. Some of the direct cut discs actually have peak levels up to 80 cm/s at 1 kHz which for a typical cartridge means about 80 mV at 1 kHz. The curve shows the difference-frequency distortion when a 15 kHz fixed sine is combined with a swept sine from 14,98-13 kHz. The resulting components from 20 Hz- 2 kHz are nothing but distortion - as high as 10%. This is why some preamplifiers sound "bass-heavy" and "without definition".

It is often said that frequencies outside the traditional range 20 Hz - 20 kHz are not important since they cannot be heard. It is true that they cannot be heard directly, but the effect of them is certainly important and clearly audible. Fig.5 is an example, while Fig.6 shows the same effect even inside the audible band.

The influence of the high frequency range is typically audible on transients. Intuitively this is not so strange, since transients - as known from the Fourier Theory consist of high frequencies. The higher the rise time, the wider the bandwidth. Since the Fourier Theory is only valid for linear systems it cannot be used directly when transient distortion is considered. However, if one thinks as if it is valid, remarkably good results are obtained in practice.

One could postulate a "subjective non-linear Fourier Theory" which states that the transient distortion can be seen as a combination of all the high frequency steady-state distortion curves. In other words, the high frequency part of the distortion curves is a measure of the transient distortion, popularly called Transient Intermodulation Distortion (TIM). It would probably help clarify some of the confusion around TIM to call it "Treble Intermodulation Distortion", which is what it really is. Likewise one could suspect that there is BIM (Bass Intermodulation Distortion) which is sub-audible frequencies modulating audible components. That this really is a problem is shown in the next section.

These subjects are treated in much further detail in the B & K Application Note 17-234 (Ref.4). Much of the work on these problems concerning Transient Distortion has seemingly been looking at only one of the many aspects of this more general - multidimensional - subjective description of the phenomena. For example, a combined square wave sine signal has been suggested by Otala (Ref.5). However, it lacks a dimension in that it does not sweep as a function of frequency such as the two-tone high frequency test. Therefore information on the shape and slope of the high frequency distortion curve is lost. It essentially only measures the 1 5-1 6 kHz point of the swept 2-tone curve.

Lately, another approach (Ref.6) suggested that as long as a high enough frequency range (like 100 kHz) is considered, harmonic distortion will reveal transient distortion. This is partly true when systems with sufficient bandwidth are considered, but not if the bandwidth is limited. Actually it is often seen that Difference Frequency distortion, especially DF2-, reveals the transient distortion up to 40 dB better than high-frequency harmonic distortion, simply because the harmonic components fall outside the pass band. The difference frequency components, however, fall down in the audible range as indicated in Fig.6.

The transient distortion is probably one of the more important parameters in the whole complexity around good sound. The final solution has not yet been found, but the swept individual two-tone distortion curves up to 200 kHz seemingly reveal the problems considerably better than any method previously used. Audible transient distortion primarily means a "frequency smear" so it is hard to distinguish whether there is, say, one violin or perhaps ten.

Nevertheless, we are at a very primitive state in transient distortion testing today. One significant dimension can be added by running the above measurements as a function of amplitude. We could also consider how the distortion varied as a function of time - for example to see if an amplifier "gets tired".

In addition, present test signals are handicapped in being symmetrical and steady state. For example, on AC-coupled systems, an unsymmetrical pulse train may initially cause overload, which will later disappear as the DC component stabilizes. Finally, we often assume that the devices we test do not have a memory, that is their performance is not influenced by previous signals. But it is a well known fact that semiconductors have significant thermal time constants, and that the thermal impedance of transistor cases and heat sinks can be very important.

5. Audible Effects of Wow and Flutter, Rumble, Tone Arm Resonances etc.

The Transient Distortion (section 4) was an example of how the high frequency domain influences the music domain by creating products that fall down into the audible range. This section will consider a similar effect from the low frequency domain that creates serious problems in the music domain by modulating the signals. In other words, the effect of subsonic signals folding up into the audible range (dare we call it BIM - Bass Intermodulation Distortion). Also here it is often heard that people say "I cannot hear 10 Hz, so I do not care". Again it is true that 10 Hz cannot be heard directly, but the effect of 10 Hz, however, is certainly audible. Some of these phenomena are illustrated in Fig.7.

The curve in the upper left hand corner of the figure shows a straightforward frequency analysis of the low frequency range 2 Hz - 60 Hz produced by an ordinary turntable with preamplifier. The most severe peak is produced by the mechanical resonance of the tone arm and the stylus, but motor rumble and hum is also clearly visible. Unfortunately, the tone arm resonance has a level typically only 10-20 dB lower than the signal produced simulatenously in the audible range.

This effect is indicated on the right hand side of Fig.7. Although the low frequency signals are not directly audible, they produce some clearly audible sidebands on the music signals. Also in this domain, the effect is typically 10% distortion. The most critical range of this is known from the wow and flutter weighting function which is most sensitive around 4 Hz. So really the closer the tone arm resonance is to 4 Hz, the worse the audible effect. A frequency analysis of the demodulated wow and flutter signal is also an interesting measurement of the phenomena. A typical result of this using the automatic B & K Wow and Flutter Meter 6203 is shown in the lower left hand corner of Fig.7.

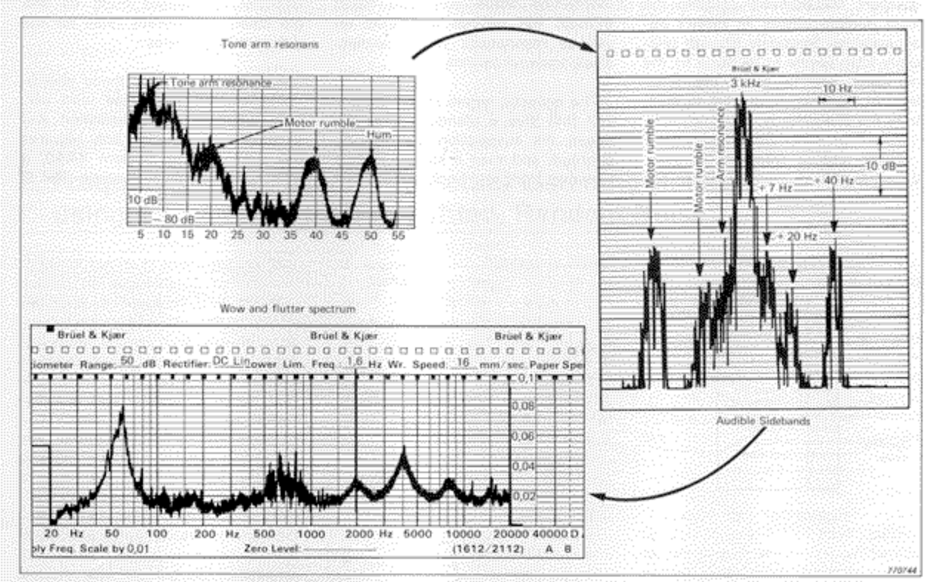


Fig.7. Low frequency distortion from tone arm resonances, rumble, wow and flutter create products that fall into the audible range. Typically 10% distortion

The pronounced resonance at 0,61-1z is due to wrong centring of the record. It is a paradox that often we think we measure wow and flutter when in reality we are measuring the influence of the tone arm cartridge combination. It does not help to improve the turntable motor mechanism when it is the tone arm resonance that is creating the problem.

Unfortunately, the mechanical resonances in tone arms are excited all the time by the warps in the records. The effect seen in the time domain is a ringing that sometimes goes on for half a revolution of the record and also affects the tracking force so it changes from near nothing to twice the "steady state' tracking force.

An interesting test of this can be made simply by making a cut in the record and offsetting the two parts. Every time the stylus passes the
"step-function" a transient is produced. A recording on a storage scope or the B & K Narrow Band Analyzer 2031 will show the time function or the time and frequency functions respectively. A typical result of different time responses for different tone arms with the same cartridge is shown in Fig.8.

The phenomena of audible effects of mechanical resonances in turntables are described in further detail in the B & K Application Note 17-233 (Ref.7). The audible effect of the phenomena again is a "frequency smear" or a "confusion effect" of the sound picture.

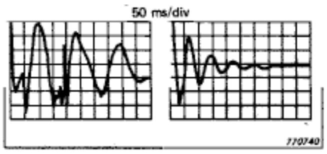


Fig.8. Transient time response of two different tone arms with the same cartridge

Gating, Early Reflections and Box Sounds

A similar phenomenon to the mechanical resonances in tone arms and turntables is creating severe problems in the other end of the HiFi system, in loudspeakers and in rooms.

Mechanical resonances in loudspeakers are probably creating the most audible effects in today's audiosystems. Strangely enough relatively little seems to be done by the manufacturers to avoid the problem. In Fig.9 we have tried to illustrate the phenomena.

Every time a transient is introduced to the loudspeaker voice coil a sound is transmitted directly, but a number of mechanical waves are also created. The wave in the diaphragm may travel several times faster than the sound in air. Therefore, the sound transmitted when this arrives at the edge of the cone will arrive before the direct sound. The mechanical wave will also travel through the cabinet and build up various resonances which successively transmit sounds.

The acoustic waves inside the box will first give a standing wave between the suspension and the diaphragm and then a standing wave between front and back, bottom and top, and side and side. All of these will, after various delays, transmit clearly audible sounds. This effect is a "time smear" that first of all means that transients are not reproduced accurately, but also gives a strong frequency dependent coloration because the various electromechanical resonances build up and die down at different rates for the different frequencies.

A measurement of these phenomena can be performed with various degrees of sophistication and expense. The simplest only requires the B & K Gating System 4440, a sine generator and a scope (Ref.8) while more advanced 3-dimensional plottings of how the frequency responses change with time also require the B & K Digital Frequency Analyzer 2131 as well as a digital calculator and plotter (Ref.9).

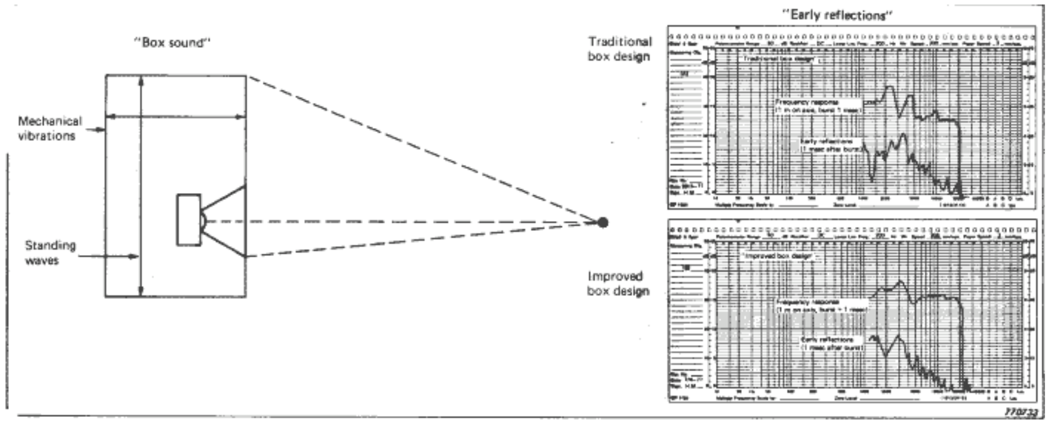


Fig.9. The sounds from different resonances arrive and continue for different times. These "Early Reflections" are clearly audible

Typical results of the 3-D plots are shown in Fig. 10, 16 and 17. This again is an example of how more dimensions increase the subjective value of the objective measurement when interpreted.

There are many reasons for these "early reflections": insufficient mechanical damping, direct coupling between the driver and the cabinet
and between the different drivers, parallel walls in the cabinets, direct coupling to the bookshelf, etc., and therefore even a rough measurement of these phenomena will reveal important information. Using a swept gated tone burst and an adjustable measuring gate curves as indicated in Fig.9 can be obtained.

The charts show the frequency response and the early reflection curve recorded 1 ms after the tone burst is supposed to stop. The upper curves are recorded for a traditional, but reasonably good box design, while the lower curves show how an improved box design - actually of the author's loudspeaker can improve the early reflections from the same loudspeaker. Typical so-called Hi-fi loudspeakers today are unfortunately, only 5-10 dB down after 1 ms. Another approach to the problem is as indicated in Ref.2, p.1 1, a measure of the mechanical vibrations using an accelerometer. However, this is only one point at a time of the higher dimensional acoustic gating measurement. Early reflections are probably one of the most pronounced problems in audio-reproduction today and a good example of an objective domain having a strong correlation to audible quality.

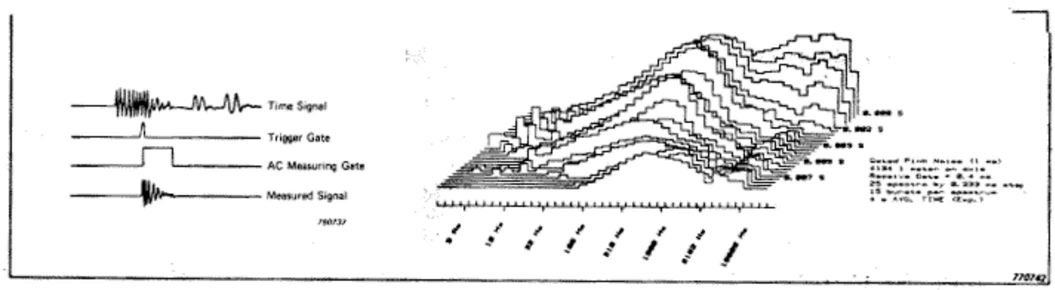


Fig.10. 3D measurement of how the various frequency responses change with time for the early sound of a loudspeaker

7 Frequency Response in the Actual Listening Room using 1/3 Octave Pink Weighted, Random Noise

One of the most fundamental "domains" in obtaining audible qualities is a measurement of the frequency response in the actual listening room using 1/3 octave pink weighted random noise. During the years there have been several investigations (Ref.10) that indicate that 1/3 octave responses at the listening position correlate strongly to subjective listening evaluations. An example is shown in Fig.1 1.

Again the curves are deliberately shown extremely small so only the "meaning" can be seen. The curves are obtained for five different loudspeakers, H1 to H5, in the same room. The upper curve, H 1, is the best, H 2 is second best and H 5 is clearly the worst. Going a little more into detail it can be found that H4 is better than H3. The important thing, however, is that the subjective listening results give exactly the same ranking.

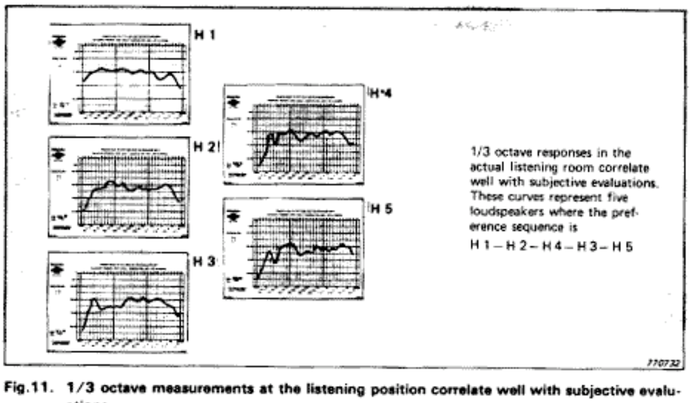


Fig.11. 1/3 octave measurements at the listening position correlate well with subjective evaluations

The most important frequency range to perform 1/3 octave measurements is from 80 Hz- 2 kHz since here the wavelength corresponds to the dimensions of normal listening environments used for HiFi reproduction. B & K has therefore also introduced a portable, low cost system directly suitable for Hi-Fi dealers and consumers. It only requires a test record QR 2011 and a Sound Level Meter 2206.

1/3 octave measurements, however, reveal only the "steady state" performance of the Hi-Fi system. Therefore the relatively popular equalizers must not be used too much. If a resonance, due to a standing wave in the room, is equalized completely, it implies that a transient is reproduced with a too low level at the frequencies where the standing waves will build up later. If, furthermore, an equalizer is used with too sharp filters, this will introduce phase distortion, as mentioned in the following section 8, and then the transient performance is degraded. Many, especially transient-oriented, people will claim that equalizers are useless, but, as usual, if they are used with care can give an improvement.

1/3 octave response in studios is an extremely important parameter since the producers listen and change the sound until it sounds good - there. However this is virtually worthless if the sound system together with the control room is not perfect. Actually, that is the main reason that most records sound so bad (Ref. 1 1).

The disadvantage of the 1/3 octave measurement is that it, in practice, requires a standard room. Sound power, however, is a slightly less valid measurement, but might be found more convenient. Sound power is a measurement of the total transmitted energy from a loudspeaker in all directions and it requires (as indicated in Ref.2, p.32) rather complex instrumentation. Sound power is also described in the B & K Technical Review No. 4, 1976 (Ref.21).

8. Phase Measurements, Transient Response and Audible Quality

The transient response of a Hi-Fi system is probably just as important a "domain" as the steady-state domain, primarily explored in section 7 about 1/3 octave measurements. When a sound is produced by a HiFi system it will first travel directly through the air and arrive at the listening position exactly as it would in the anechoic chamber or the free field. Later, the sound will be reflected and arrive from the various acoustic surfaces in the room. The "frequency response" will therefore change as a function of time. This can, as mentioned in section 6, be measured as a 3-D plot showing frequency response as a function of time (Ref.9).

The frequency response corresponding to the direct sound will reveal the "transient response" while the integrated responses after a long time will reveal the "steady state" information as obtained with 1/3 octave noise. The transient response, however, is revealed from the free-field information. Therefore amplitude and phase responses measured in the anechoic chamber or using gating techniques are certainly important when transients are considered.

In Fig.12 we have tried to illustrate the importance of Phase measurements. If the individual components in a transient are offset in time it means that the individual components in a complex music signal will not arrive at the listening position simultaneously. If, for instance, as indicated at the lowest part of Fig.12, the midrange is closer to the listener than the tweeter, which again is closer than the woofer, it means that the midrange part of the music information will arrive first, then the high frequency part and finally the low frequency part. This gives a coloration of the sound, especially audible for transients.

The right part of Fig.12 shows the result of phase response measurements on the author's loudspeakers with and without phase compensation.

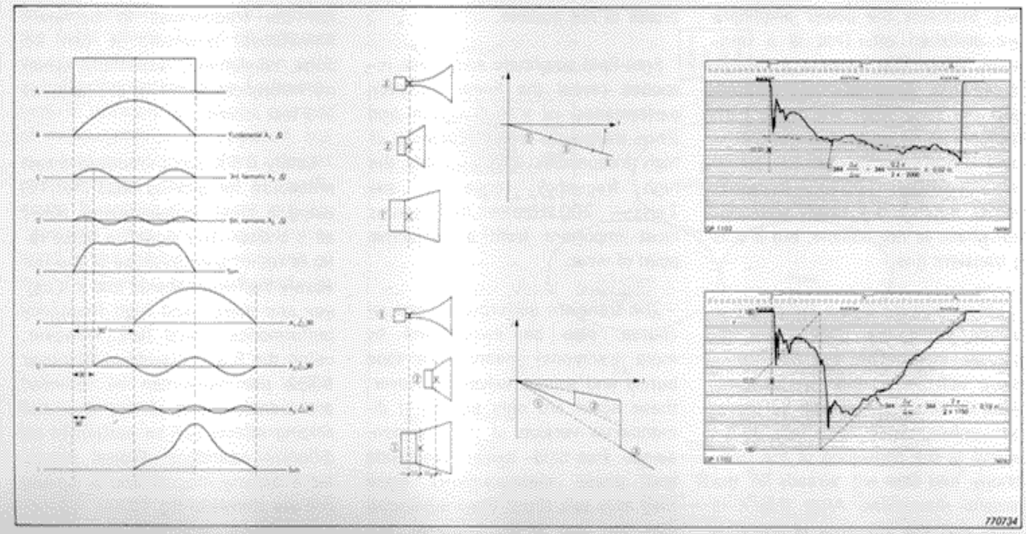


Fig.12. A non-linear phase response means that the various components in a complex music signal do not arrive simultaneously. This gives an audible coloration

A more detailed description of phase measurements can be found in B & K Application Note 17-198 "Loudspeaker phase measurements, transient response and audible quality" (Ref. 12).

The discussions about audibility of phase have been going on for many years, and have intensified since Richard Heyser introduced his first paper about the subject in 1969 (Ref.13/14). In those days, one of the few that could measure phase was Mr. Heyser. In 1973 Bruel & Kjaer introduced the Phase Meter 2971 and the Phase Delay Unit 6202 and since then the discussion has grown considerably more intensive.

It would not be reasonable to deal with all the arguments in this paper as it is definitely intended to be an overview version without too many "local" arguments. Nevertheless, let us take some of the main points in the phase discussion.

At the AES Convention in London in 1975 James Moir demonstrated the waveform change indicated in Fig.12 by offsetting a number of continuous sines relative to each other. Although everyone could see the change in the resulting waveform, no one could hear any difference. The conclusion that phase is not audible, however, was not warranted, since this was a test of steady-state signals and phase is important to transients.

Later others (for example Harwood of BBC) tried to introduce allpass systems in order to change only one parameter, phase. Most of these tests were performed for 60° and 90° non-minimum phase shift. However, loudspeakers (before phase compensated loudspeakers became common) typically display 10 x 360° non-minimum phase shift in the range 100 Hz - 10 kHz.

Various studies (such as by Rorbaek Madsen, Denmark) show that down to 15° minimum phase shift in the midrange probably is audible. This seems reasonable since the human ear has the highest time constant of 50ps since we are able to hear up to 20 kHz (Ref.15) (15° at 1 kHz corresponds approximately to 360° at 20 kHz).

The phase shift problems introduced when the listener moves his head have also been discussed many times - for instance at the AES Convention in New York, when Matsushita presented a paper about phase-compensated speakers. The answer here, is (as indicated in B & K Application Note 17-198, Fig.33) that a relative movement of the loudspeakers gives considerably more phase shift than a similar movement of the head. If not, the whole thing is useless.

At the 1977 AES Convention in Paris, Carsten Thomsen summarized the main arguments and misconceptions about phase, by pointing out that phase is an engineering unit - and thus not necessarily something we can hear. But the influences of phase errors are many and clearly audible by giving time smear, poor transient response, overload due to phase errors, distortion of distance perception, confusion of stereo image, and change of tonal quality.

Nevertheless, sometimes rather strong arguments are required in the phase discussion. Here is one of the worst ones:
Consider a 3-way loudspeaker system playing pink noise. First the tweeter is moved 1 mm back while the level is increased slightly pink noise is still heard. Then the tweeter is moved 1 cm back, the level is increased a little bit - pink noise is still heard. 10 cm back still pink noise. Now 10 kilometers back and the level is increased a little bit (it is a powerful tweeter) still pink noise. Finally, 100 km back - still pink noise (ignoring the high frequency attenuation of the air). Suddenly the power amplifiers are switched off. That is a transient. When that happens a change is audible, because the midrange and the bass stop, while the high frequencies go on for five more minutes - the time it takes the sound to travel the 100 km. In other words, when it is a steady state signal phase is not audible, but if it is a transient it is.

Let us assume that a symphony is played while the tweeter is still 100km away. The first 5 minutes there will be no high frequencies. After 5 minutes they will be there, but unfortunately they will correspond to the beginning of the symphony and we are already in the second movement. After this it is clear that the influence of phase is audible - the question is only How much phase is audible?

The discussions about phase are a good example of a local domain that does not reveal the whole global meaning alone. The discussion can go on like it does, only because the phenomenon is masked by other things. Unfortunately there are many phase-compensated loudspeakers that sound rather bad because they have ignored other domains, but that does not mean that phase is not audible.

Free-field amplitude and phase response reveal the linear transient performance of a Hi-Fi system and since transients primarily consist of high frequencies it is probably the high frequency range - say 2 kHz - 200 kHz - that seems most important from a subjective point of view.

The transient performance can, of course, also be investigated by more traditional means, like tone bursts and square waves. However, these again are only an (n-1) dimensional version of the n-dimensional free-field swept amplitude and phase measurements since they only talk about the frequencies that the test signal contains, but not the frequencies in between. Sometimes, however, it can be quite convenient, especially when no reference signal is available as in a test record.

Lately B & K has introduced a test procedure for pick-up tests (Ref. 16) using a small accelerometer 8307 as a shaker. The signal to noise ratio is rather poor because of the relatively inefficient shaker and if a bigger one were used high frequency performance would lack. However, using the B & K Waveform Retriever 6302 the noise can be removed and transient test of rise time and ringing effects can be performed on different cartridges. Typical results for a moving magnet and a moving coil are shown in Fig. 13.

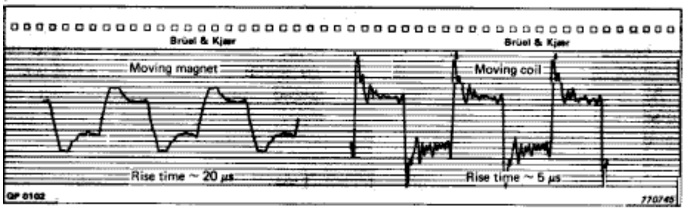


Fig.13. Transient performance of a moving magnet and a moving coil cartridge

The 6 Measuring Domains that today seem to Correlate with Subjective Evaluations

The introduction to Wireless World, August 1977 states: "Anyone who has read that curious book "Zen and the Art of Motorcycle Maintenance" will recall that the narrator apparently drove himself into a mental hospital by his obsessive attempts to discover by pure reason the essence of "quality". Even Socrates had trouble with such universals'". And later:
"Engineers certainly do follow Lord Kelvin's dictum that you can't properly understand a phenomenon until you can express it in numbers".

With this firmly in mind we will nevertheless try to correlate some of the "local" objective and subjective parameters shown in Fig.1 into a "global" meaning - good sound. The exponentially increasing amount of data today requires interpretation in order to solve the important question: What does the data mean and what is subjectively good?

Unfortunately, interpretation is to a certain extent a question of opinion, but there seems no way around it. Therefore this paper will also present the author's opinion that today we probably have just passed the point where, based on relevant measurements alone, we are able to judge the quality of an Audio system. In principle an infinite number of measurements are required, but in practice relatively few relevant measurements seem to be sufficient. It seems that there are six "domains" that are strongly correlated to the subjective perception of sound. These are indicated in Fig. 14.

1/3 octave measurements in the actual listening environment seem the most important linear parameter in the frequency range 20 Hz 2kHz. It primarily describes the "steady state" performance of the system. A standard environment ought to be introduced so this parameter could be specified by the manufacturers.

The most important domain in the linear high frequency range (2 kHz - 200 kHz) seems to be "free-field amplitude and phase measurements" that primarily reveal the transient performance since transients consist of high frequencies.

The range 200 Hz - 20 kHz could be called "the gating domain" because it describes the phenomena going on between the steady state and the transient conditions. With a long gate steady state conditions are obtained, while a narrow gate reveals the transient conditions. Early reflections and box sounds are probably one of the most important problems in today's Hi-Fi systems. Frequency response and Early reflections, for instance, 1 ms after burst, ought to be specified.

As indicated in Fig.14 [not shown] there also seem to be three "non-linear domains" that are strongly correlated to the subjective perception of Audio. The most important one in the 2 Hz - 20 Hz range seems to be the "Tone are resonance, flutter and rumble domain". This is an example of how low frequency components outside the traditional audio band create severe problems by folding up in the audible range. 10% distortion is rather typical and ought to be specified.

The most important parameters in the traditional audio range 20Hz20kHz are probably the two-tone swept difference-frequency curves - DF3- and DF2-, but also IM and Harmonic might be useful. Especially DF3- is important when narrow bandwidths are considered, as in a multiway loudspeaker system, this will normally reveal distortion considerably better than the traditionally measured harmonic distortion. The highest value or all the curves ought to be specified.

The high frequency range 2 kHz - 200 kHz seems to reveal "transient distortion". The Difference Frequency DF2- and DF3- are probably the most suited parameters. Transient distortion is (as indicated in Fig.6) not only present in amplifiers, but even more so in FM-tuners, phono-preamplifiers and tape recorders. Also here 10% distortion is rather typical, and ought to be specified although it does not look as good as 0,01%.

Fig.14 is, as mentioned, only the author's attempt at subjective evaluation of Audio measurements today. This is only a start, but it might be possible to use the modern calculators to obtain a "global" result by reasonable subjective weighting of objective "local" parameters starting with the above-mentioned parameters. With high density memories and microprocessors, it should not be beyond the capabilities of today's digital electronics to make even what Lord Kelvin asks for - a number - if that is desired.

Unfortunately, the six important parameters mentioned above are not standardized in any country, simply because standardization takes time. It might be difficult to agree on what is important, but something should be done.

10. Apodization

When all the objective and subjective local parameters are to be evaluated, "apodization" will probably play an important role. To apodize means "to remove the feet". In physics it means to remove the side lobes in the well known (sin x)/x spectrum indicated in Fig. 1 5.

From the Fourier theory we know that a pure sine, that starts at -∞ and goes on to + ∞ in the time domain, by the Fourier transform can be seen as a single line in the frequency domain. This is actually an example of the "global to local" transformation (section 1). If only a part of the sine is present - and after all that is the case in real life a "smear" is created in the frequency domain. The side lobes have a (sin x)/x nature. The shorter the tone burst the wider the frequency spectrum. Actually the relation has an extremely simple nature that T= 11/13. If the time domain gets very sharp like a transient the frequency domain gets very broad. The extreme is a unit impulse with a flat frequency spectrum from -∞ to + ∞.

The "truth" is always somewhere in between. Therefore the practical version of Apodization is to find the optimum compromise between the sharp extremes by smoothing things out.

It is always a good question to ask when a certain measuring parameter is improved: What then is getting worse? Unfortunately there has been a trend in the Audio industry to discover the extremes without mentioning what it costs. Just think of Phase, TIM, feedback, High Compliance, Noise reduction, Bass Reflex etc, as examples. A few years ago the advertisements said: "Unmeasurable distortion due to heavy feedback". Today they say: "Unmeasurable TIM due to low feedback". Obviously the optimum is somewhere in between. The "Gauss weighting" is probably the best compromise since this has the property of being the Fourier Transform of itself.

There are many practical examples of this "philosophy". In Fig.5 and Fig.11 we saw that a meaning could be seen by overview of a reasonable number of curves. Generally it means that if one is too close to the domains where things are very small it is impossible to see a meaning. And, if one is too far away, like in the infinite space, it is also impossible to see a meaning. This is obvious, but it is not much used in Audio.

Another example of how apodization, or smoothing improves things is shown in Fig.16. This is a 3D measurement made by JVC (Ref. 17) of the transient response of a softdome tweeter compared with a harddome tweeter. The soft dome will not try to move air with a sharp step function. Therefore it will ring less and have a better transient response than the stiff hard dome.

The same can be seen for the simple closing of a door. If it is closed with a bang it is not as desirable as if it is closed smoothly. When a car is stopped, especially if the road is icy, the optimum way of braking is with a Gauss function. Fortunately humans do not think about it, they just do it.

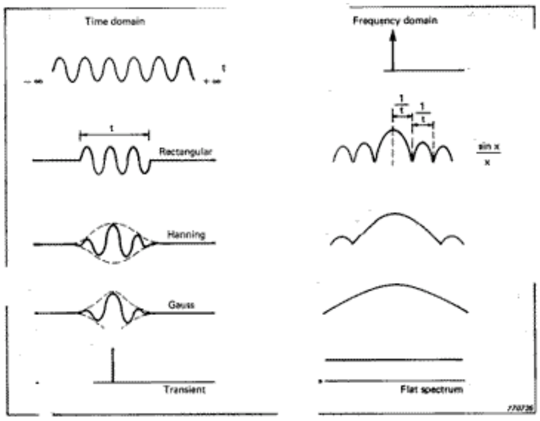


Fig.15. A broad time spectrum corresponds to a narrow frequency spectrum and vice versa

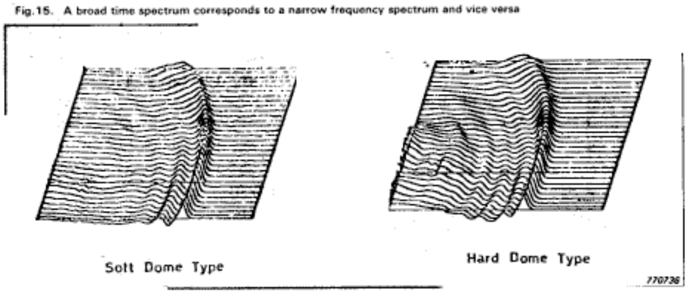


Fig.16. The more "apodized" soft-dome tweeter has a better transient performance than the hard-dome. The pictures show the acoustic response to a raised cosine pulse as function of direction

If an optical lens is blurred around the edge, the image is sharper. If the gap of a tone head in a tape recorder is rounded, the frequency response is improved. If a loudspeaker or a microphone is rounded it sounds better, etc. Apodization in Audio is very important when the various local domains are combined.

The essential thing in the "apodization philosophy" is to realize that if a parameter gets better in one domain it simultaneously gets worse in another. If, for instance, a sharp filter is used in the frequency domain it gives ringing in the time domain. If the transient distortion is improved by a low pass filter the phase response is degraded, or in other words, the transient distortion can be "improved" by taking the transients away.

This "philosophy" seems to have similarities to the general principle of uncertainty. For instance, the position and the momentum for a particle cannot be measured simultaneously with high accuracy. If one is '"clear" the other is "smeared".

Really apodization might be as general as time and space. We do not understand that an event can have happened in no time, or that the universe has been there all the time. We do not understand the infinitely big universe or the infinitely small particles, but the apodized version (somewhere in between) is intuitively easy to understand.

11 . More Dimensions in Audio

The conclusion of this paper is hopefully clear - that overview of more parameters gives more subjectively meaningful results.

Let us illustrate it with a rather popular example that unfortunately has a lot to do with Audio although it might not seem so to begin with. Consider a number of "flat animals" gliding around on the floor. They are completely flat and since they are on the same floor they can only see each other, but they cannot see up or down. Now a human being comes in and takes one of the flat animals away. If the other animals are asked "what happened to the first one?" they will say "he died". They are not able to tell why or how, but the human being who can see more dimensions, can easily tell why and how (Ref. 18). Good examples from the latest years' AES papers of how more dimensions improve the understanding are shown in Fig. 17.

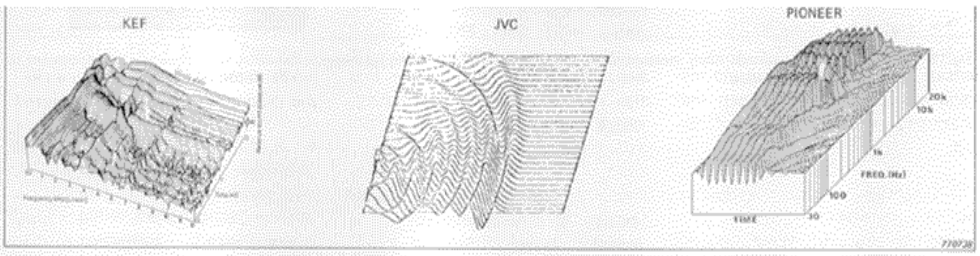


Fig.17. Examples of how the latest years' 3D plots have increased the subjective visibility

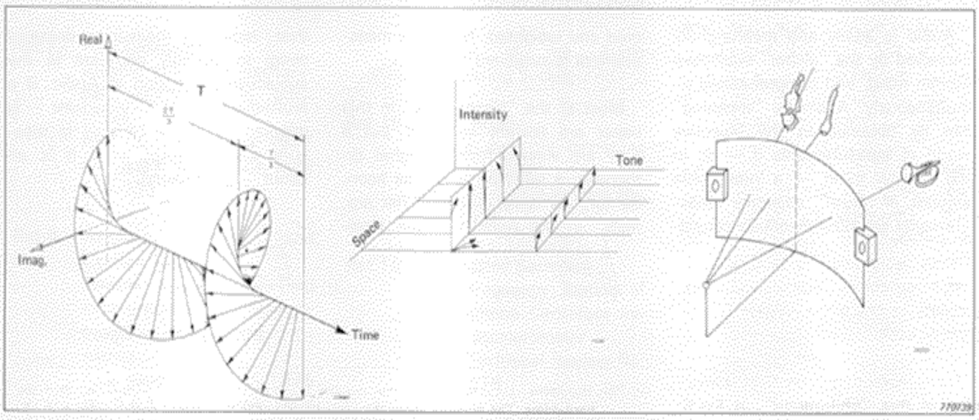


Fig.18. Examples of possible future 3D-plots

The KEF picture shows how the various "early reflections" die down for a loudspeaker as function of frequency and time (Ref.19). The JVC picture shows the response of a loudspeaker to a raised cosine impulse as function of the acoustical transmission into the environment (Ref.17). And finally, the Pioneer picture shows the time response of a tone burst as function of frequency (Ref.20).

Modern digital techniques imply lots of possibilities like this in the future - even in reasonable price ranges. Some possible examples, primarily inspired by Richard Heyser, are shown in Fig. 18.

First, phase between components in a complex signal. Second, how the acoustic position (space) is influenced by the intensity and tone of the signal. And finally, the acoustic position of the instruments relative to the loudspeakers. Many of these measurements are actually possible even with today's instruments coupled to modern computers and calculators (Ref.9).

In Fig. 19 we have shown the basic instrumentation combinations using hardware instruments and calculators connected via the digital IEC interface Bus. The limitations are largely determined by the users.

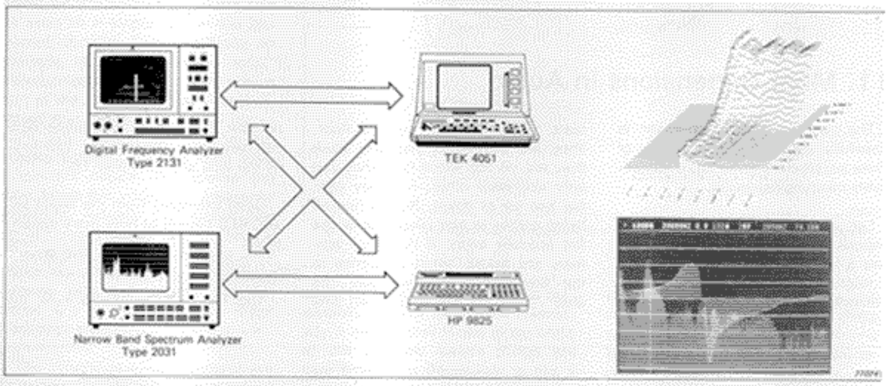


Fig.19. Today's technology permits many measurements not previously seen by using hardware instruments and calculators interconnected v the IEC digital interface Bus

12. Conclusion

Audio is easily and meaningfully perceived by the "global" subjective human mind, and comprehended simultaneously. A similar "meaning" can be obtained in the objective world of measurements if - as in the human mind - a reasonable amount of "local" objective measurements are simultaneously considered and weighted. No single measurement is sufficient.

Most of the "dimensions" in this paper are literally very old, but if they are viewed from a higher dimension a meaning might be seen. So far, we have all been "flat animals" in the Audio domain. However, today six measuring domains seem to strongly correlate to the subjective perception of Audio. If a multidimensional viewpoint is adopted we might be able to measure and interpret what it is all about - good sound.

References

Ref.1. Richard C. Heyser
Communications
Journal of the Audio Engineering Society (AES), January 1976

Ref.2. Henning Møller
Electro Acoustic Measurements
B & K 16—035

Ref.3. Carsten Thomsen and Henning Møller
Swept measurements of harmonic, difference frequency and intermodulation distortion
AES paper New York 1975 and B & K Application Note 15—098

Ref.4. Poul Ladegaard
Swept distortion measurements — an effective tool for revealing TIM in amplifiers with good correlation to subjective evaluation
AES paper New York 1977 and B & K Application Note 17—234

Matti Otala
Transient Intermodulation Distortion
IEEE 1977 No. 1, February 1977, or AES paper New York 1976

Ref.6. Jung, Stephens and Todd
Slewing Induced distortion and its effect on audio amplifier performance — with correlated measurement/listening results
AES paper Los Angeles, 1977

Poul Ladegaard
Audible effects of mechanical resonances in turntables
AES paper New York 1977 and B & K Application Note 17—233

Ref.8. Henning Møller and Carsten Thomsen
Electro Acoustic free-field measurements in ordinary rooms — using gating techniques
AES paper New York 1975 or B & K Application Note 17—196

Ref.9. Henning Møller
3-Dimensional acoustic measurements — using gating techniques
AES paper New York 1977 and B & K Application Note 17—163

Ref.10. Henning Møller
Relevant loudspeaker tests in studios, in Hi-Fi dealer's demo-rooms, in the home etc., — using 1/3 octave, pinkweighted, random noise
B & K Application Note 17—197

Ref.11. S.E. Børja
How to fool the ear and make bad recordings
AES Journal July/August 1977

Ref.12. Henning Møller
Loudspeaker phase measurements, transient response and audible quality
AES paper Los Angeles 1974 and B & K Application Note 17—198

Ref.13. Richard C. Heyser
Loudspeaker Phase Characteristics and Time Delay Distortion:
Part 1 January 1969
Part 2 April 1969
Journal of the Audio Engineering Society

Ref.14. Richard C. Heyser
Determination of Loudspeaker Signal Arrival Times: Part 1 October 1971, Part 2 November 1971, Part 3 December 1971
Journal of the Audio Engineering Society

Ref.15. Per V. Brüel
Do we measure damaging noise correctly?
B & K Technical Review No. 1 — 1976

Ref.16. Ole Brøsted Sørensen
High Frequency Testing of Gramophone Cartridges using an accelerometer
B & K Technical Review No. 2 — 1976

Ref.17. Nomoto, Iwahara and Onoye, JVC, Japan
A technique for observing loudspeaker wave front propagation
AES paper, Zurich 1976

Ref.18. Edwin A. Abbott
Flatland — a romance of many dimensions
Dover Publications Inc., New York

Ref.19. L.R. Fincham, KEF
Loudspeaker system simulation using digital techniques
AES paper London 1975

Ref.20. Suzuki, Morii and Matsumura, Pioneer, Japan
Three-dimensional displays for demonstrating transient characteristics of loudspeakers
AES paper, New York 1976

Ref.21. Holger Larsen
An easy and accurate method of sound power measurement
B & K Technical Review No. 4 — 1976

end