DO WE NEED AN ULTRASONIC BANDWIDTH FOR HIGHER FIDELITY SOUND REPRODUCTION?

M Colloms Colloms Electroacoustics, London

1 INTRODUCTION

Reproduction bandwidths greater than 20kHz are now on offer to the consumer, promoted on grounds of higher fidelity. Recording, transmission/storage means, amplification and sound radiation aspects must all be consonant and capable. Wideband SACD, DVDA and now HD DVD and Blu-ray HD sound carriers deserve review in conjunction with the generally limited bandwidth of available sound reproducers. Not least, human perception deserves further examination in respect of a greater high frequency bandwidth. The objective is to deliver a practical perspective on the widely promoted, higher bandwidth audio proposition.

2 HIGH FREQUENCY PERCEPTION

2.1 Classic data and the established consensus

From the beginnings of audio recording our industry has fought for greater fidelity in sound reproduction. Progressively, the application of technology improvement has resulted in lowered noise and distortion, extended playing times, increased dynamic range and greater bandwidth. While the accepted audible range of typically 15kHz has long been a defining parameter, in recent years this limit has been subject to doubt. Claimed fidelity improvement is now said to result from greatly expanding the high frequency limit for the audio chain. Some industry sectors are advising up to two and a half more octaves more, to 100kHz. Potential reproduction bandwidths greater than 20kHz are now offered to the consumer, promoted on grounds of superior sound quality.

Standard references inform us that the ear is sensitive to variations in air pressure. If described as cyclical, the rate of variation that may be detected under ideal conditions, by a young healthy adult, is extended towards the extremes of 20Hz in the low bass and up to 20kHz in the highest treble. Certainly some humans and children can hear to higher frequencies and animals have shown to be particularly sensitive, dogs to 45kHz (ultrasonic dog whistles) and cats to 80kHz.

Above 10kHz, sound is perceived more as a high-pitched whistle. There is an increasing loss of information, the sensation moving towards that of noise. We are informed that it very soon becomes inaudible. Beyond this point we enter the region of ultrasound. Via bone conduction, sound is still perceived up to 100kHz, but as a single, noise-like pitch. High intensity sound above 23kHz may also be perceived as pain. Propagation in air becomes increasingly lossy at higher frequencies, calling into question the evolutionary need for more than 20kHz sensitivity in a natural environment. Benefiting from the efficient mechanical conduction of sound vibrations in water, dolphins may respond to stimuli up to 200kHz.

2.2 Established practice

Historically, when developing high fidelity stereo for FM broadcast, audio engineers felt, almost without hesitation, that 15kHz was a sensible bandwidth limit. Thus they confidently placed the synchronising pilot tone, required for decoding the two-channel transmission, at a quite low 19kHz. Placed in the reproducing path, the accompanying notch filter limited the overall bandwidth to 16kHz. A classic and highly respected series of BBC monitoring loudspeakers, subject to extensive live sound comparisons by experts, was designed using an otherwise excellent high frequency unit whose output happened to decay sharply above 14kHz. This was not considered important. An audiologist, who by profession diagnostically tests hearing, will likely venture 18kHz as a sensible upper limit for significant human response to natural sound levels. 20kHz was proposed long ago as a convenient end-stop for engineering the audible range, leaving some margin for error and equipment tolerances.

2.3 Recent audio developments

Pressure to increase sales encourages a continual striving for improvement. In addition designers do honestly desire to improve performance. The limited emergence of SACD source material, some with genuinely extended bandwidth content, has prompted the sale of costly, accessory super tweeters on the hi-fi market, devices with greater than 20kHz frequency responses. In addition loudspeaker designs have been introduced, some at substantially higher prices, that are claimed to have a greatly extended treble performance. The proposition made is that better sound quality will inevitably result from extending the audio reproduction range.

At its best, SACD is capable of 100kHz, though in practice the level of ultrasonic noise which accompanies the one-bit noise shaping conversion ideally employed, is so great as to have been known to destroy a noted audiophile amplifier before the launch demonstration got underway. Consequently the first player, released to the market from the SACD technology principal Sony, had a screwdriver-locked low-pass filter switch fitted to the rear panel. This restricted the bandwidth to 50kHz despite the company claiming a reference grade classification for the much lauded design.
196kHz sampling for multi-bit formats, DVD-A, PCM, is also capable of providing a 100kHz replay bandwidth and while in theory the ultrasonic noise problem is much diminished, regrettably nearly all of today’s consumer DACs are of the noise shaping, high over-sampled types, of which many have significant ultrasonic noise contributions, regardless of the media source or digital format. These decoders generally require strong filtering before the audio signal continues down the chain.

Sadly the industry now accepts that, DVDA and SACD have made little market impact in the quest for higher fidelity replay, despite the bandwidth and resolution advantage over CD, this mainly for commercial rather than technical reasons. Imminently there are two new disc carriers, HD DVD and Blu-ray. These are essentially video directed but have such enormous storage capacity that the potential for multiple wide band audio channels is embedded in their standards. Whether these become the super audio carriers of the future partly depends on the outcome of the format war now in progress between them.

2.4 High Frequency Reproducers

If it is of value, we still need to get this extra bandwidth to the listener. Considering elements of the replay channel, such as amplifiers and decoders, many examples have low-pass filters at 20 kHz. Ironically, many of the latest switching technology power amplifiers also have low pass output filters. These are located at 25kHz or so, to combat their propensity to radiate serious EM. Loudspeakers are also an important issue. Taking sensitivity as a key parameter we know that significant diaphragm area is beneficial. Conversely, a reduced area is essential for maintaining directivity and power response at higher frequencies. If the radiator is pistonic, we need very small sources, ideally 10mm or less in diameter, to generate useful sound power up to 100kHz.

The alternative choice of larger, louder but intentionally resonant diaphragms, of potentially wider directivity, may well result in a wildly discontinuous frequency response. If operated down in the normal audio band such would be dismissed without question. I have measured famous brand ‘wide band’ drivers whose 20kHz to 50kHz operation was based on just a few high Q modes, these emanating from a resonant 25 mm metal dome.

Room absorption continues to increase at higher frequencies reducing the potential impact of an extended response. In the higher range the ear itself is more directional, as HRTF work has shown. Any possible benefit from those higher frequencies would require a reasonably direct path to the ear canal entrance. Thus we must question the benefit from the increased investment required if it seems so difficult to deliver the wide band proposition to the listener’s ear.

2.5 Anecdotal Consumer Press Reporting

So far I have considered that informal, anecdotal reports from reviewers concerning the advantage of add-on super tweeters to be rather suspect since I feel that in general the basis of their evaluations has generally been so flawed that their results are of little value. The difficulties encountered are several and only a few need be considered to explain my negative viewpoint.

For example, a subjective test for response extension will only be valid if the extended response is achieved without affecting the performance in the existing known ‘audible’ range. Testing for a subtle effect, which may be barely audible, is nonsense, if it changes the uniformity and loudness in the already operative treble range. Yet this is what is happening in these reported tests. So far, no commercially available add-on tweeter and matching crossover avoids this fundamental error.

In addition, when such a driver and crossover is patched on to an existing audio chain, as it often is, it will inevitably change the loading on both the cable and the amplifier, and thus very likely impart another audible difference. Such electrical loading will generally extend into the audible range.

The test ‘super tweeter’ typically operates in parallel with the existing tweeter over about an octave of shared bandwidth, and may thus destructively interfere with the primary tweeter output. Thus there is potential to impair as well as alter the results. Sadly, some critics are so pleased to have heard a difference they are tempted to judge the change as an improvement.

Often the crossover is a simple capacitor feeding a metal dome tweeter, perhaps with a beryllium composite diaphragm. Such an arrangement is something of a disaster since a calculated 20kHz ‘crossover’, comprising a single capacitor for a nominal 5ohm rated tweeter, provides the response shown in Figure 1. The intrinsic output should be compared with the crossover objective, which is seen to be markedly different from the practical result generally obtained with a single element filter. The problem lies in the complex impedance presented by the high frequency driver compared with a simple resistive termination. Even with more complex, higher order filters, the practical crossover points for super-tweeters are often placed well into the audio band.

It is not surprising that many audio professionals dismiss such published subjective results, which often seem to be produced in support of media and equipment marketing.

Tested in my audio system, subjective results for two noted super-tweeters provided a loss of perceived image focus and depth, with poorer sense of musical timing, some emphasis in distortion where it was present in historic programme, and finally a subtle yet audible non-uniformity in the high treble range. This was accompanied by a mild increase in treble power, not necessarily to the benefit of the overall timbre of the designed loudspeaker. There was felt to be a little more air and sparkle dispersed in the room acoustic, this merely an observation, not value call.

2.6 Industrial Ultrasonics

With audiologist's data considered well accepted and documented, the industrial use of sound is firmly established. Referring to health and safety literature, there is much teaching on audible bandwidth. Worldwide, the authorities concerned with hearing damage and health note that the rate of sensitivity falloff with frequency for the general population is very rapid, typically 100dB per octave above 20kHz, that is about 100dB down by 40kHz. Acoustical energy above 20kHz is typically termed ultrasound and is by such definition deemed inaudible.

Such data has lead to legislation regulating the use of ultrasound, for example authorising the use of sound sources that on the face of it might be a cause of concern even shock for an audio specialist. An example is the use of industrial cleaning baths capable of 110dB SPL at 35kHz. In fact ultrasound is widely exploited while relying on accepted grounds of inaudibility and aural safety, the latter consideration because it may in some circumstances cause tissue damage while it remains inaudible, and the former on the basis of annoyance and distraction. At very high intensities sound may cause damage since the high kinetic energy levels resulting may produce sufficient molecular agitation for cavitation to occur. In fluids this promotes the formation of micro-bubbles, which collapse explosively, capable of damaging cell membranes, even DNA. Intentionally destructive high power ultrasound is the basis of the lithotripter whose energetic, multiple, directed and focused sound beams can smash kidneys stones in situ, hence avoiding invasive surgery. It seems to be safe in other respects for patients and operators when properly administered. Doctors confidently beam ultrasound at a foetus for diagnostic imaging, and assume that it is safe, (that incidentally it not unpleasant), and that it is inaudible. The unborn seem calm enough under such examination, while they may readily respond to conducted audible sound such as dance music delivered at a relatively low level from the environment.

2.7 Further Perception Tests

A series of tests for subjective hearing response was conducted at the 1980 London AES convention by Laurie Fincham[1], then Technical Director at KEF Electronics. Although the results were presented at an accompanying seminar, the work was not published. Laurie, now at THX Ca., had kept a copy, and on my request kindly sent his material by return. Laurie had succeeded in inveigling the participation of a galaxy of audio specialists, a veritable 'Who's Who' of industry illuminati. Many have since risen to prominence in the audio field and the AES and included Subir Pramanik, Floyd Toole, Peter Mapp, Dick Heyser, Paul Messenger, Bent Hertz, Ragnar Lian, John Vanderkooy, Stan Lipshitz, John Borwick, Peter Baxandall, and David Meares. Of course, confidentiality is maintained in respect of their listening sensitivities.

The main question posed was audible bandwidth. During that busy convention he and his KEF team captured over 480 man-hours of subjective testing. Looking over the Fincham procedure, he chose a stimulus of near maximum aural sensitivity, namely a synthetic, wide-band, repetitive pulse train. This was reproduced via a wide directivity, high frequency driver, a 20mm dome radiator selected for a uniform, 40kHz frequency range and good directivity. Care was also taken so that all participants enjoyed an aural line-of-sight relationship to the source. The reproduction bandwidth for this easily learned signal could readily be altered in stages, the chosen limits provided by specially designed, 9th order elliptical filters of 0.18dB pass band amplitude ripple, judged to be aurally innocuous. 80dB or more attenuation was achieved by 1.25X the nominal cut-off frequency. In contrast to many hearing tests that have employed sine wave or band-limited signals, here a stimulus closer to a complex natural transient was chosen.

Summarising the results, which had been subjected to mean and standard deviation processing, and remembering the considerable auditory experience and eminence of the participants, the test showed that a filter limiting at 10 kHz was reliably audible for almost everyone. There was a sigh of relief all round! However for the 16kHz filter setting there was obvious difficulty in identification. This limit was considered to be weakly identified, and only by some participants. Reviewing the results overall, the subjective impact of a 16kHz low pass filter was so small that it may be judged to be on the verge of inaudibility. This result correlated with the empirical, contemporary BBC work on loudspeaker development, which had involved live source comparisons. When the results for the final 20kHz Fincham bandwidth test were assessed, it was clear that statistically at least, all the tested listeners, though working under less than ideal conditions, could not hear this applied limit. These results confirmed conventional wisdom.

More recent tests have been published on this subject. Oohashi et al[2] worked very hard to nail the issue in a listener friendly 'Hi-Fi' context, and chose the gamelan for the source, verified as harmonically rich to 100kHz. A special digital recorder of the similar bandwidth was employed. In addition a loudspeaker pair was specially constructed for stereo reproduction. A custom-made diamond dome tweeter provided 80kHz extension. A youthful and sensitive panel was engaged, and the participants were assessed for their responses. The stimuli comprised firstly CD, then 20kHz nominal bandwidth material, followed by the gamelan reproduction when limited to 24kHz, then the 'sound' of the gamelan when reproduced only in the band above 24kHz, and finally the gamelan with the full reproducing bandwidth of about 50kHz.

In addition to conscious scoring by the participants, physiological and brain wave analysis, including PET scanning, was also undertaken. The latter had the ability to assess changes in 'pleasure' response even if the panelist is not consciously aware of it. The results revealed some fascinating contradictions. While full bandwidth was generally preferred to CD bandwidth, the test for the ultrasonic music band alone showed that this band was inaudible. Exceptionally, exposure to the full bandwidth stimulus seemed to additionally confer some kind of conditioned subjective improvement for subsequent reproduction at a reduced bandwidth, as if there was some kind of beneficial short-term memory effect.

Finally, from the physiological data, the full bandwidth replay was judged to give the greatest listening pleasure. This data showed good correlation with the conscious, written down subjective reporting. Since the ultrasonic band energy had been shown to be inaudible when assessed alone, it was considered that this component had become audible only when meaningfully associated with the sub-ultrasonic range. Much debate has followed this report.

Lenhardt et al[3] showed that speech could be heard in the ultrasonic range, if it was modulated onto an ultrasound carrier and presented via bone conduction. It was shown that the inner ear has some mechanism for demodulating the speech signal from the carrier. This implied non-linearity raises further issues concerning how such listening tests are devised.

Boyk[4] generally supports the wider bandwidth proposition and provided a useful assessment of the bandwidth of many musical instruments, showing that they radiate to at least 40kHz. This aspect was not widely accepted until Boyk's material was published.

Black[5] has considered whether for some tests, reports for ultrasound audibility were in fact due to another spurious mechanism. His investigation showed that in some cases nonlinearity in the system and/or a reproducer could result in intermodulation products where a difference product is formed from ultrasound signals and is then falsely presented within the accepted audible band.

In 2,000 Howard[6] usefully reviewed the many issues to date, prompted by consideration of the increasing availability of commercial 'super tweeters', but did not venture a conclusion concerning the value of increased reproduction bandwidth.

Oohashi[7] has since conducted extensive brain scanner investigations with ultrasound stimuli, 'hypersonic sounds'. It was noted that physiological effects were quite complex, especially in respect of the observed memory effect, and further investigation was proposed in connection with human responses to ultrasound and changes in brain chemistry.
Seeking to investigate further, in part focusing on the Oohashi gamelan results, and also aiming to thoroughly review extant material, Nishiguchi et al[8] of NHK Laboratories, set up yet more tests. A

particular concern was the steepness of the filters used to separate the notionally audible band from the ‘inaudible’, this considered to be a possible weakness of the original Oohashi test. They ran a classical, forced decision type of test, contrasting with Oohashi and his single presentation method. For my part I have found forced decision methods tend to impair subjective response sensitivity though they can be time efficient. However it is abundantly clear that great care was again taken by these researchers to try and deliver a representative result.

Taken overall, and with some polite hesitation from the authors, the Nishiguchi results would appear to contradict Oohashi. In practical terms they conclude that 20kHz is entirely sufficient as a nominal limit for sound reproduction. Their result is attributed in part to the use of very steep band filtering in the experiment, to more firmly separate the ‘audible’ energy band from the ‘inaudible’.

Advance notice has now been given of further work by Oohashi et al[9] which suggests that his previously reported phenomenon actually requires that the body and not just the ears be exposed to the ultrasonic sound field. If the body is shielded from the ultrasonic component, then the perception of an extended bandwidth is no longer reported.

Considering the above evidence, it would seem that the case for a wider sound reproduction bandwidth is not proven.

3 EXTENDED BANDWIDTH, MASTERING AND REPLAY

For origination, regardless of the audibility issue, you may still argue advantages for the proposition. A wider operational bandwidth may result in improved performance, or the equivalent of increased resolution, benefiting the primary audible range. At present, strong low pass filtering is often employed to fit a 20kHz bandwidth to many digital audio chains, especially MP3, broadcast, and CD. The use of a recording medium of extended bandwidth makes possible more gentle filtering, with improved phase and impulse characteristics in the filtered range.

For loudspeakers the effort to extend bandwidth may result in improved quality in the lower frequency range. Designing for better bandwidth should provide superior power response to higher frequencies, improving off-axis responses and consequently, sound quality. In addition, diaphragm resonances that may have been closely proximate to the ‘audible’ band will now be located well out of range. In Figure 2 the typical low loss, multi-resonant output of a rigid ‘piston’ metal dome driver is compared with that for a wideband pure diamond type, clearly moving such problems well out of the way.

4 CONCLUSIONS

As an industry we need to maintain a healthy skepticism concerning marketing based performance claims. Human perception is notoriously difficult to measure, especially when differences are small, and some reported results may well be misleading. I believe that elements of the audio industry have exaggerated the necessity for the consumer to invest in an extended reproducing bandwidth.

Considering this review of the subject, the case for an ultrasonic reproducing bandwidth is not proven, and 20kHz remains a practical and sensible limit for the chain as a whole.

However within reason, for origination and post-production, we should be encouraged to use the widest bandwidth possible to preserve the greatest information content. This will maximise archival quality.

5 REFERENCES

1. Laurie Fincham, 'The Golden Ear Room, Listening tests on high frequency bandwidth and clipping conducted at the AES 1980 Convention, London' (unpublished)
2. Oohashi et al, 'High–Frequency sound above the Audible Range affects Brain Electric Activity and Sound Perception' Preprint 3207 91st AES Convention NY
3. Martin L. Lenhardt, Ruth Skellett, Peter Wang, Alex M. Clarke, 'Human Ultrasonic Speech Perception'. Science, Vol. 253, 5 July 1991, pp. 82-85. Abstract, p. 82.
4. James Boyk, 'There's Life Above 20 Kilohertz! A Survey of Musical Instrument Spectra to 102.4 KHz' 1997 www.cco.caltech.edu/~boyk/spectra/spectra.htm
5. Richard Black, 'Anti-alias Filters: the invisible distortion mechanism in digital audio', AES 106th convention Munich, May 1999 (see WW.musaeus.co.uk/aespaper.htm)
6. Keith Howard, 'The How and Why of Super Tweeters', HFN June 2000
7. Oohashi, Tsutomu, et al, 'Inaudible high-frequency sounds affect brain activity: hypersonic effect', Journal of Neurophysiology, Volume 83, No. 6, June 2000, pg 3548, http://jn.physiology.org/cgi/content/abstract/83/6/3548
8. Nishigichi et al, 'Perceptual Discrimination between Musical Sounds with and without Very High Frequency Components', NHK Laboratory Note No 486. AES 115th Convention 2003 NY October
9. Oohashi T, Kawai N, Nishina E, Honda M, Yagi R, Nakamura S, Morimoto M, Maekawa T, Yonekura Y, Shibasaki H.
   'The role of biological system other than auditory air-conduction in the emergence of the hypersonic effect'. (Pubmed preprint announced no date yet)
   Department of Research and Development, Foundation for Advancement of International Science, Tokyo 164-0003, Japan; National Institute of Information and Communications Technology, Koganei 184-8795, Japan

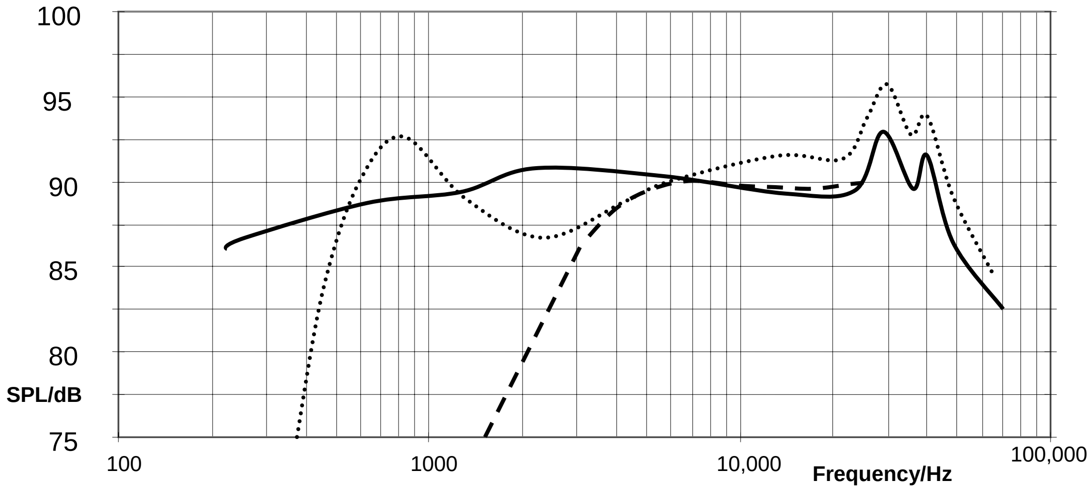

Figure 1, Behavior of 1st order, single capacitor, high-pass crossover.

Key : Frequency response of typical driver, solid line.
Ideal crossover performance, dashed line.
Effect of single capacitor crossover, dotted line.

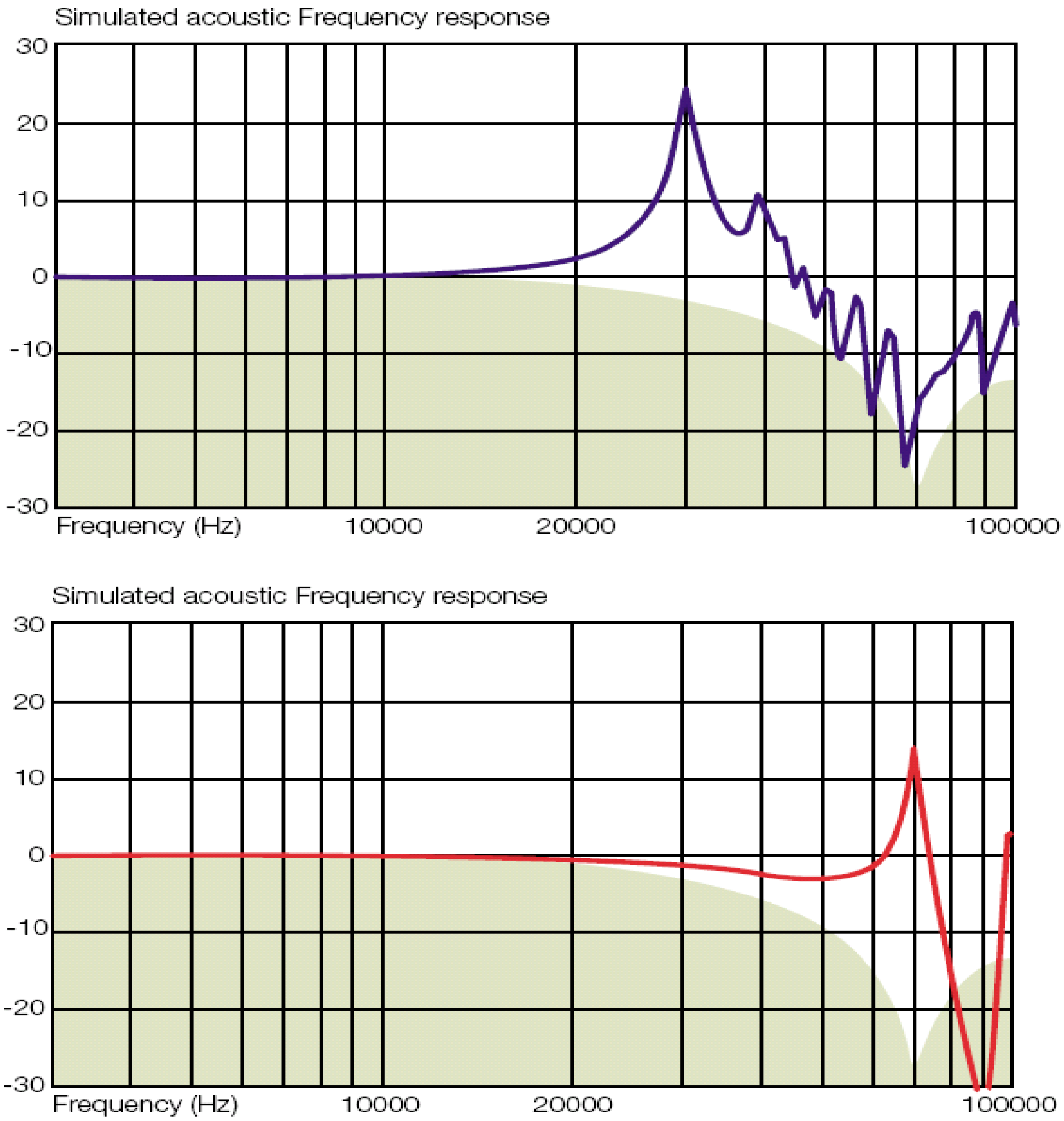

Figure 2 , FEA for axial radiation from a 25mm aluminium dome, upper graph, and for an equivalent in ion deposited crystal diamond dome, 50um thick, lower graph (courtesy B&W).