

THE WHY AND THE HOW OF SUPERTWEETERS

It's almost a reflex for the clued-up audiophile to respond to the question 'What's the audible frequency range?' with '20Hz to 20kHz' although, if you're a stickler, you will want to add the word 'nominally' to the answer since the upper limit in particular is subject to individual variation and gradual erosion by the ageing process (presbycusis). So well established is it that few of us can hear up to, let alone beyond, 20kHz that it formed the basis for the CD specification, whose sampling rate of 44.1kHz was chosen to provide 20kHz bandwidth plus a sufficiently wide transition band to the Nyquist frequency of 22.05kHz to accommodate steep anti-alias filtering on recording, and similar reconstruction (interpolation) filtering on replay.

Such is the enduring status of the 20kHz limit, indeed, that certain audio industry observers have questioned the whole thrust of SACD and DVD-A and their espousal of higher sampling rates. Even more controversial in this regard is the development, principally by the Japanese majors [and now of course, Townshend Audio], of a new breed of supertweeter with outputs extending to perhaps 100kHz or more, intended to exploit the new media's unprecedented ultrasonic capability. If we cannot hear above 20kHz – and nobody is seriously suggesting that we can, at least not in the conventional sense – then what is the point? Only after that question is answered is it relevant to touch on the second element of the story: how loudspeaker designers are contriving to reproduce frequencies up to two octaves higher than previously.

Ringling the changes

It's important to appreciate from the outset that the justifications for upping sampling rate and the those for increasing bandwidth into the ultrasonic region are not necessarily one and the same.

One of the more plausible, although still controversial, explanations for the desirability of increased sampling rates has been suggested by Mike Story of dCS (Ref 1) and relates to the impulse response of sharp anti-alias and reconstruction filtering. Although oversampling has allowed digital filtering to assume this task in modern A/D and D/A converters, replacing the complex analogue filters originally employed, impulse response remains a concern because of the sharp roll-offs required. Even with digital filters a rapid transition from passband to stopband causes pronounced ringing (oscillation). The exact form of this ringing depends on the precise filter characteristics

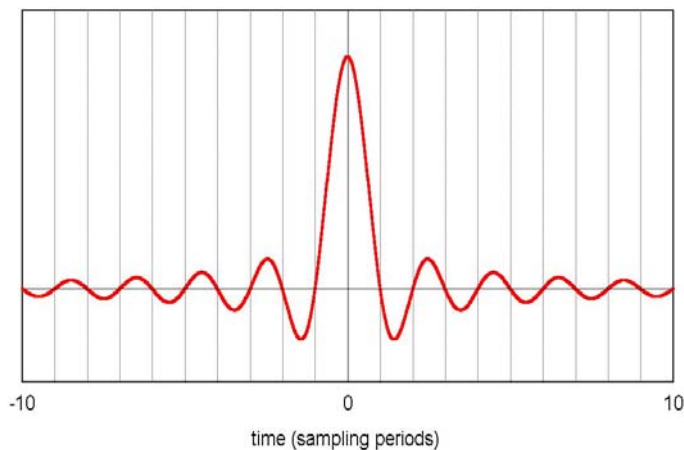


Figure 1

but Figure 1 shows a limiting example, this being the $\sin(t)/t$ impulse response resulting from a perfect (and therefore unattainable) realisation of the brick-wall low-pass filter with an infinitely steep roll-off at 22.05kHz. Practical anti-alias filters are of necessity less extreme and therefore have somewhat curtailed ringing, but this 'ideal' case serves to illustrate the principle.

It is generally held that the ringing shown in Figure 1 would be inaudible because the fre-

quency of oscillation lies just above the nominal audible range, as it would also for a typical practical filter realisation. What Mike Story has suggested is a different way of looking at this ringing, whereby its energy versus time is plotted on a logarithmic intensity scale.

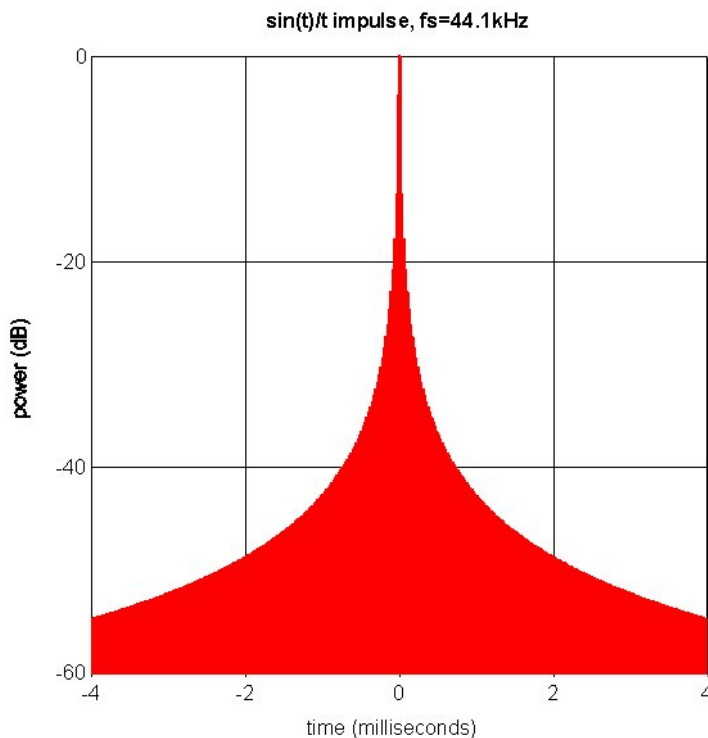


Figure 2

If you do that with the impulse response of Figure 1 you obtain the result shown in Figure 2.

On the time scale used in this second graph the fine detail of the impulse response is lost but the envelope of the power curve is clear. As you can see, the energy contained within the ringing falls off quite slowly, so that what was a short input impulse becomes smeared across quite a large time range in the output. It is this smearing, or rather its reduction, which Mike Story suggests is responsible for the improvements heard when sampling rate is increased.

Upping the sampling rate by a factor of two or perhaps four (DVD-A operating at 88.1/96kHz or 176.4/192kHz) improves matters in two ways. First, the filter's corner

frequency can be raised, thereby increasing the frequency of oscillation and contracting the associated time smear. Second, the filter's

transition band can be shaped to provide a gentler initial roll-off, which reduces the severity of oscillation. A wide range of filter alignments is possible, with different trade-offs between frequency domain and time domain behaviour, all of which can substantially reduce the extent of the time smear relative to Figure 2. SACD, with its 2.8224MHz sampling rate, banishes the problem entirely, of course.

If this contention is correct it suggests that high sampling rate sources should sound better than 44.1/48kHz equivalents even when there is no provision for realising their additional bandwidth, and there is a growing body of evidence that this is the case. As many reviewers, me included, discovered with the 96kHz DAT machines Pioneer introduced in the early 1990s, sound quality was clearly superior at 96kHz sampling rate to that of the same machine operating at 44.1 or 48kHz, even though resolution remained the same in both cases at 16-bit. dCS has also demonstrated this effect (Ref 2) and favourable reactions to its Purcell upsampler confirm it (Ref 3).

An important ramification of this is that SACD and DVD-A sources can sound better than a CD equivalent *even when using ancillaries incapable of exploiting their extended bandwidth* – a message which is sometimes lost in the advertising hype for DVD-A and SACD-ready amplifiers and loudspeakers. You don't have to equip yourself with 100kHz-capable electronics and loudspeakers to hear the benefit of increased sampling rate, not to mention the advantages which accrue from higher resolutions than 16-bit.

But this is not the same as saying you will hear all the sound quality improvement that the new media have to offer. On the contrary, there is both hard and anecdotal evidence that you won't. Despite our general inability to hear above 20kHz it seems that capturing and reproducing ultrasonic frequencies has a beneficial effect on sound quality, over and above that obtained by increasing sampling rate while keeping system bandwidth unchanged.

Music for bats?

Advocacy of bandwidth extension to ultrasonic frequencies has frequently been lampooned as satisfying only cats, dogs and flying mammals but that smug view is now the subject of a concerted challenge. The conventional view is that (a) musical instruments produce insignificant output above 20kHz, and (b), as already outlined, we cannot perceive such frequencies in any case. In the last decade both contentions have been shown to be wrong.

The idea that musical instruments have spectra largely confined to the accepted audio bandwidth of 20Hz to 20kHz probably owes its origins to the classic work of W B Snow of Bell Telephone Laboratories (Ref 4). He established that certain instruments had their timbre altered by the introduction of a low-pass filter with a corner frequency as high as 13kHz, as a result of which he arbitrarily showed their spectra extending to 15kHz – the upper limit of the reproduction system used for the tests.

The fact that Snow conducted this research as long ago as the early 1930s suggests a re-investigation using more modern equipment and analysis techniques was overdue, and this has now been undertaken by James Boyk working at CalTech (Ref 5). What his measurements have shown is that a notional 20kHz limit on musical instrument spectra is illusory. To quote the opening two sentences of the abstract from his paper, "At least one member of each instrument family (strings, woodwinds, brass and percussion) produces energy to 40kHz or above, and the spectra of some instruments reach this work's measurement limit of 102.4kHz. Harmonics of muted trumpet reach to 80kHz; violin and oboe, to above 40kHz; and a cymbal crash was still strong at 100kHz." In the case of the cymbal Boyk has shown that no less than 40 per cent of the instrument's acoustic energy is generated above 20kHz. Any notion of musical instruments having been designed or selected to conform to an upper frequency limit of about this frequency is therefore discredited.

Whether we can perceive these over-20kHz frequencies is, of course, an entirely separate issue. Testing with sine waves suggests we cannot but there is other evidence which indicates that we are both aware of and able to utilise frequencies above those we can consciously perceive. The two studies which support this (Refs 6 and 7) are commonly mentioned only in passing in articles about ultrasonic bandwidth extension, but such is their importance they deserve closer attention.

The first paper, by Martin Lenhardt and colleagues, reaches a quite remarkable conclusion about human ultrasonic perception: that we can not only detect frequencies substantially in excess of 20kHz but also analyse them in much the same way as we can signals within the accepted audible spectrum. This was demonstrated by modulating speech signals into the ultrasonic range and delivering them to test subjects via a transducer in contact with the head. Subjects not only perceived the ultrasonic signals as sound rather than vibration, they were able to understand what was being said.

If you accept the widely held view that development of all our sensory capabilities was shaped by evolutionary need, this finding strongly suggests that humans do detect and utilise ultrasonic frequencies in everyday life, and process them via the same pathways used to analyse speech. The only obvious alternative is to suggest that our ability to perceive ultrasonic sound is an evolutionary relic, rather like the human appendix – one which nevertheless remains functionally connected to our highest levels of auditory processing. It's a perverse line of reasoning to say the least.

There is a significant difficulty involved with extrapolating from the Lenhardt experiments to everyday listening experience, however. Sensing ultrasonic signals applied directly to the head, where they are subject to bone conduction, is not at all the same as perceiving them at a distance via the air. The evolutionary need argument suggests that such perception must be possible otherwise why have the facility, but is there any evi-

dence of it?

The second paper, by Tsutomu Oohashi and colleagues, provides it. By monitoring the brain electrical activity of test subjects when presented with sounds either containing or lacking ultrasonic components, they were able to show a consistent difference in EEG (electroencephalograph) response when ultrasonic frequencies were present. Moreover, they discovered that the change in EEG response caused by the presence of ultrasonic frequencies can persist after the ultrasonic components are removed - a finding which has obvious ramifications for A-B listening tests.

Subjective tests conducted in parallel with the EEG studies indicated that "the music containing high frequency components was perceived as more pleasant and rich in nuance than music from which high frequency components were eliminated." This echoes a common reaction to high sampling rate digital sources reproduced using ultrasonic supertweeters: that extension of the HF response is perceived as influencing the opposite extreme of the (audible) frequency range, with listeners often commenting on improved bass performance and greater warmth. Switched-on loudspeaker designers have known for many years that similar effects occur if you adjust high frequency balance within the audible spectrum. I well remember Max Townshend tweaking the output level of a prototype loudspeaker's supertweeter - which was rolled in around 14kHz, as I recall - in response to complaints from me that tenor voice sounded too thin. Inexplicably, reining back on the supertweeter indeed restored the missing warmth.

Transients and phase

You might suppose that the scientific evidence of human ability to perceive frequencies substantially in excess of the limit indicated by sine wave testing would be sufficient to justify ultrasonic loudspeaker bandwidth in and of itself. But such is the continuing influence of the 20kHz limit that engineers involved in developing ultrasonic supertweeters often can't resist seeking justifications which lie firmly within the established audio range. A case in point is Tannoy, whose recent white paper on the new Prestige SuperTweeter (Ref 8) suggests that improved transient ability and reduced high frequency phase distortion might be the basis of its subjective benefits.

I don't find either explanation persuasive, although that's not to say there isn't a grain of truth in them - or that I might not be plain wrong. It is correct, of course, that an extended high frequency bandwidth allows for faster signal risetime. Bandwidth determines maximum rate of change in the signal; increase one and you necessarily increase the other. But if the ear itself sets the bandwidth limit, then it sets the risetime limit also. This might be compromised somewhat by bandwidth restrictions in the signal reaching the ear, but obtaining the full benefit of the improvement in transient performance that results from increasing bandwidth to 50 or 100kHz is simply not possible unless we are able to perceive ultrasonic frequencies. In other words, citing transient response as if it were a separate issue from perceived bandwidth is misleading.

Phase distortion is an obvious target for speculation as to why ultrasonic supertweeters bring subjective improvements. If we ignore their crossovers, loudspeakers usually closely approximate minimum phase behaviour. In other words, their phase response can be calculated from their amplitude response and vice-versa (using the Hilbert Transform). So long as a minimum phase device displays a flat amplitude versus frequency response, it introduces no phase distortion. But any departure from a flat amplitude response causes an equivalent perturbation in phase behaviour. One practical consequence of this is that a loudspeaker's inherent LF and HF roll-offs introduce phase distortion. In both cases moving the roll-off to a lower or higher frequency respectively will have the effect of widening the frequency range over which phase distortion is negligible.

This suggests that adding a supertweeter to force a loudspeaker's HF roll-off to higher

frequencies should reduce phase distortion within the audible range, but there are two important caveats to that argument. If the supertweeter replaces the loudspeaker's main tweeter, working from typically 3kHz upwards, then fine: no additional crossover network will be needed. But supertweeters are often supernumerary to the main tweeter, taking over somewhere around 20kHz. Even if this crossover is achieved by exploiting the inherent roll-offs of tweeter and supertweeter (which is rarely the case – usually an electrical network is used, if only for the high-pass element), this will still introduce phase distortion due to the all-pass nature of the combined response. Far from reducing HF phase distortion, then, adding a supertweeter can actually increase it (as must be the case with the Tannoy device, which offers third-order high-pass filtering at a choice of 14, 16 or 18kHz).

This is avoided if the supertweeter is able to replace a conventional tweeter (few if any can, although Sony is working on such a design) but even then there remains good reason to question whether any audible benefit accrues. The audibility of loudspeaker phase distortion remains a contentious issue, as it has been since loudspeakers like the B&W DM6 kick-started the linear phase debate in the mid-1970s (Ref 9). It's a subject I intend to return to soon in these pages, albeit in a different context, but what evidence is there that the phase distortion associated with a typical tweeter's HF roll-off has audible effects?

The answer is: very little. To assess the subjective significance of HF phase distortion in loudspeakers it's instructive to look at early generation (i.e. analogue) anti-alias filters,

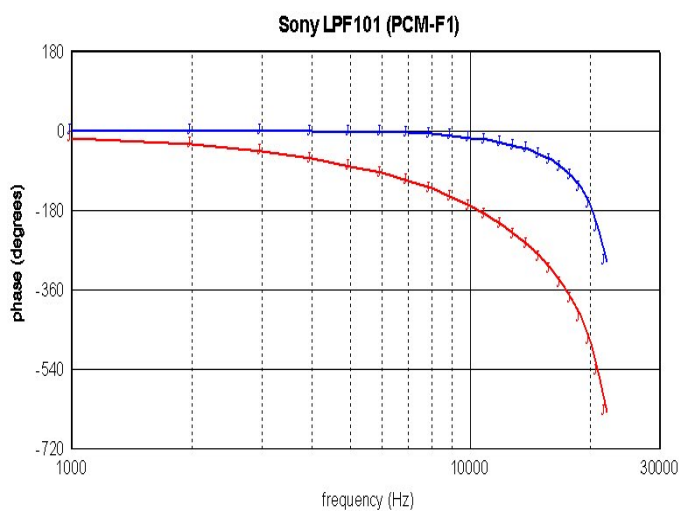


Figure 3

whose phase distortion was the subject of considerable concern in the early 1980s concomitant with the deployment of digital audio. In 1983 the late Peter Baxandall provided me with copies of the measurements he'd made on the LPF101 anti-alias filter used in the seminal Sony PCM-F1 - the recorder which first persuaded many people, Peter Baxandall included, that 16-bit digital audio at 44.1 or 48kHz sampling rate offered effective transparency to live audio signals. One of his measurements was of the phase lag introduced by the filter, which is reproduced here as the red trace in Figure 3.

As it stands this phase plot is potentially misleading because it takes no account of the fact that, over much of the audible range, the filter introduces an almost constant time delay of 43 microseconds. As a constant time delay introduces no phase distortion it can usefully be subtracted from the phase plot, to leave what we might term the excess phase (blue trace) – in this context, the phase error which does result in waveshape distortion.

Note that by 20kHz the LPF101 introduced an excess phase lag of about 165 degrees; in fact a complete record/replay cycle through the PCM-F1 resulted in twice this phase error because the same network was also used as the reconstruction filter following D/A conversion. Despite this, the phase distortion introduced by such filtering was generally held to be inaudible (Ref 10). Figure 4 (red trace) shows the equivalent graph for an idealised tweeter with a maximally flat second-order Butterworth high frequency roll-off (corner frequency 25kHz, Q=0.707). This response introduces a low frequency time delay of about 9 microseconds; subtracting this, as before, leaves the blue trace. Note the expansion of the vertical scale here relative to Figure 3: within the audible range such a

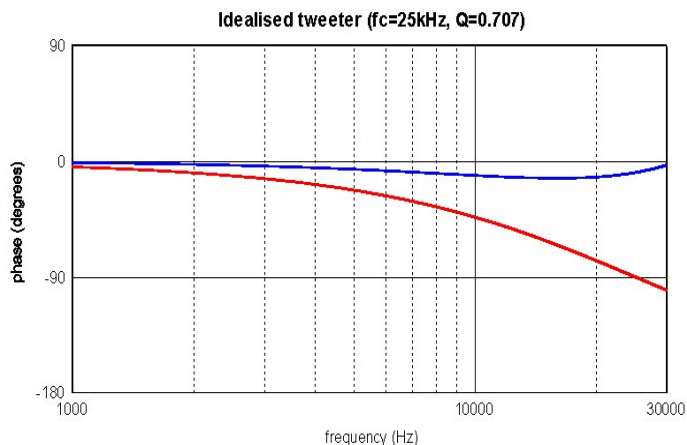


Figure 4

tweeter would introduce a maximum excess phase lag of only 13.4 degrees at around 16kHz – a tiny fraction of that introduced by the concatenated anti-alias and reconstruction filters of the PCM-F1. The conclusion has to be that phase error of this order is unlikely to be audible, so reduced HF phase distortion looks to be a weak justification for the ultrasonic supertweeter – even assuming the deployment of an additional crossover network doesn't make a nonsense of the entire claim.

If at this juncture I'm beginning to sound like a non-subscriber to the supertweeter cause, let me nail my colours to the mast and say I have heard distinct improvements when they are used. I'm not in any doubt as to their worth – I just question the value of some of the explanations offered in their justification.

Above and beyond

The output of conventional tweeters typically begins to fall off soon after (sometimes a little before) 20kHz. Extending output to around 100kHz, to provide the 96kHz bandwidth required to satisfy DVD-A's highest supported sampling rate of 192kHz, requires a substantial increase in bandwidth of around two octaves.

Tweeters which sustain output well into the ultrasonic range are not new, although they have been the exceptions. The revered Kelly ribbon and Fane Ionofane 601 ionic tweeter, for example, both offered strong output to over 30kHz in the 1960s. Further variants on the ionic theme, the Magnat Plasma MP-02 and IML Electronics Digiplasma, achieved 40kHz-plus in the early 1980s, and down the years there have been various leaf tweeters, mostly from Japanese companies, offering extended HF bandwidth to still higher frequencies. Technics manufactured a range of stand-alone units – less than succinctly designated the EAS-10TH1000, EAS-10TH400 and EAS-10TH200 - with responses to over 100kHz, which found wide application in acoustics modelling (where the wavelength of the test signals has to be reduced concomitantly with the contraction in building scale). Versions were also fitted to some mainstream Technics speakers of the period – the SB-7, SB-7K and SB-10K – although they are probably better remembered for their flat, honeycomb-reinforced bass and midrange diaphragms. Quoted frequency extension in these models was 125kHz, -10dB. I know that Pioneer has had some wide bandwidth leaf tweeters available for some years too; doubtless there have been others.

Leaf tweeters are an obvious choice for extended bandwidth applications because they combine two essential characteristics: low diaphragm mass and low voice coil inductance. Less desirable traits are a propensity to resonance in the flat, low-stiffness diaphragm - which usually has to be controlled using a thin layer of damping material (a fine mat of glassfibre, for example) - and relatively poor power handling resulting from the inherent fragility of the diaphragm and difficulty in dissipating heat from its etched 'voice coil'. (Sony (Ref 11) also cites poor sensitivity as a disadvantage of leaf supertweeters but the fact that Technics' EAS-10TH1000 combined a specified sensitivity of 95dB SPL for 2.83 volts at 1 metre with its 100kHz-plus range tends to refute this.)

There is reason for supposing that these particular disadvantages are of no great significance in an ultrasonic tweeter. It has yet to be demonstrated, for instance, that the quality criteria relevant to ultrasonic reproduction are the same as apply within the audible range. It may well be that diaphragm resonance – as well as response irregularities, nonlinear distortion, etc – is of much less significance at these elevated frequen-

cies, although the fact that a number of sources, including Tannoy (Ref 8), have reported subjects detecting the effect of ultrasonic response peaks (as commonly displayed by metal dome tweeters) suggests that familiar performance issues are still of relevance.

Power handling requirements certainly ought to be lessened at ultrasonic frequencies to judge from most of Boyk’s musical instrument spectra, but there are other factors at work here. The old problem of amplifier clipping is one: if the system amplifier is overdriven it may generate sufficient ultrasonic distortion to fry the supertweeter. More fundamentally, there is also a power handling issue inherent in wideband SACD reproduction because of the high level of ultrasonic noise generated by the one-bit DSD recording process that underpins it, which relies on noise-shaping to provide acceptable dynamic range within the audible spectrum.

This is one reason why Sony chose a different design route for the supertweeters used in its SS-M9ED and SS-X9ED wideband loudspeakers (Ref 11). A cutaway of the SS-M9ED unit – also available in stand-alone form as the SS-TW11ED supertweeter – is shown in Figure 5.

Figure 5

At first

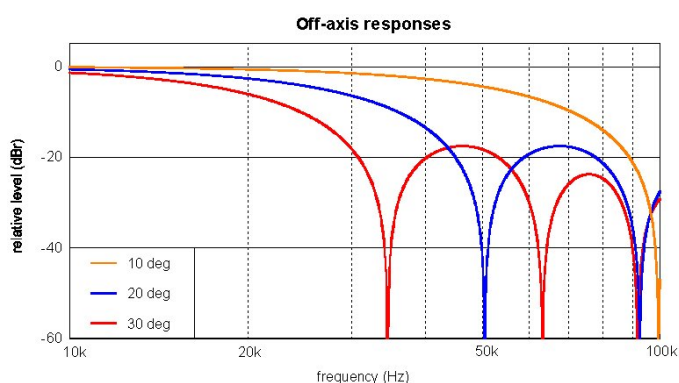


Figure 6

less impressive fre-

quency extension (specified as -6dB at 54kHz, -18dB at 100kHz, whereas Sony’s response plots for the M9ED unit show a substantially flat response to 90kHz). Note that neither design is able to operate as a rigid piston to such elevated frequencies. The Tannoy dome has a first breakup resonance at 30kHz while the Sony design relies for its extended output on “five successive resonances between 20kHz and 100kHz”.

glance an apparently conventional 1 inch dome tweeter, it actually incorporates two novel features which permit output to 100kHz: a light but stiff carbon fibre composite dome (the variant in the SS-X9ED uses a glassfibre composite which reduces the upper frequency limit to around 70kHz) and an inductive drive system. Tannoy has also chosen the dome tweeter option, albeit of more conventional design (directly driven titanium diaphragm) and

Something to bear in mind with all ultrasonic tweeters, whatever their design, is that their output is subject to pronounced beaming because of the short wavelengths involved, and also to air absorption, the extent of which varies considerably according to atmospheric conditions (both temperature and relative humidity).

Figure 6 shows idealised responses at 10, 20 and 30 degrees off-axis assuming a flat, circular, perfectly rigid diaphragm of 25mm diameter, mounted in an infinite baffle and with ruler-flat output on-axis. Off the forward axis output falls away quite rapidly above

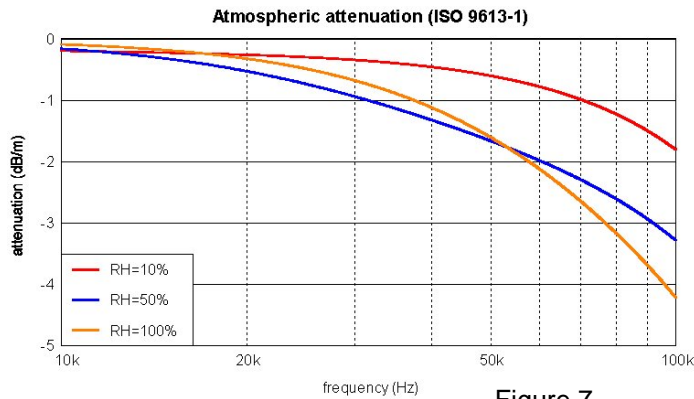


Figure 7

20kHz, suggesting an optimum listening area of even narrower confines than normal. Figure 7 shows air absorption per metre for three relative humidities, all at an ambient temperature of 20 deg C, calculated using the relationships given in ISO 9613-1 ('Calculation of the absorption of sound by the atmosphere').

At a typical listening distance of 2.5 metres the absorption is

two and a half times that shown in the graph, so even at a modest relative humidity of 50 per cent sound pressure level at the listening position will be about 10dB down by 100kHz. The degree of difference between the low and high humidity curves further suggests that the subjective benefit of extended bandwidth systems may vary from day to day. Indeed, it was suggested many years ago that ambient humidity affects system sound quality, even without HF extension to ultrasonic frequencies.

Ancillaries

Throughout this discussion of ultrasonic bandwidth extension it has been assumed that the amplifier used to drive the supertweeter has a wide enough small-signal frequency response and power bandwidth to do it justice. This is by no means a certainty. On the established principle that the wider you open the window the more muck flies in, many power amplifiers have curtailed ultrasonic output. Indeed, power amplifiers often incorporate an RC low-pass filter at their input to prevent ingress of high rate-of-change signals which might cause them to misbehave.

To exploit the new breed of supertweeters fully may therefore require an amplifier upgrade – although that's not to say no improvement will be heard using an amplifier of more modest bandwidth. The relative importance of these issues is something which will come out in the wash as we gain experience with SACD and DVD-A, and the new loudspeakers and amplifiers designed to exploit them.

Nor should it be taken for granted that DVD-A and SACD themselves will always exploit their enhanced bandwidth potential. Discs of 'legacy' recordings almost certainly won't, and neither will many created specifically with the new media in mind. The output of most studio microphones is already in decline by 30kHz. Instrument microphones with well-maintained output to 100kHz and beyond are available, but consistent response to these high frequencies demands a small capsule (usually quarter-inch) and the smaller a condenser mic's capsule is made the noisier it becomes. Aware recording engineers are going to have to decide the relative importance of noise and bandwidth, and trade them off accordingly.

References

- 1) M. Story (1998), 'The High Life', Studio Sound, p109, May 1998
- 2) P. van Willenswaard (1998), Industry Update, Stereophile, p39, March 1998
- 3) A. Harrison (1999), 'Purcell Washes Whiter', HFN/RR, p32, December 1999
- 4) W. B. Snow (1931), 'Audible Frequency Ranges of Music, Speech and Noise', J.

