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

Martin Colloms Colloms Electroacoustics, London

1 INTRODUCTION

Reproduction bandwidths greater than 20kHz are now offered to the consumer, promoted for higher fidelity. Recording, transmission/storage means, amplification and sound radiation aspects must all be consonant and capable. 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, perception deserves further examination, especially ultrasonic sensitivity. The objective is to deliver a practical perspective on the present, wider bandwidth audio proposition.

2 HIGH FREQUENCY PERCEPTION

2.1 Classic data and the established consensus

From the beginnings of audio recording our industry has strived 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, more recently this limit has become questioned. 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, 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 respond up to 200kHz.

2.2 Established practice

Historically, audio engineers when developing high fidelity stereo for FM broadcast, felt, almost without hesitation, that 15kHz was a sensible bandwidth limit and thus confidently placed the synchronising pilot tone, required for decoding the two-channel transmission, at a quite low 19kHz. The accompanying notch filter in the reproducing path limited the bandwidth to 15kHz. A classic and highly respected series of BBC monitoring loudspeakers, subject to extensive live sound comparisons by experts, were designed with 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 engenders a continual striving for improvement. Not least, designers do wish to improve performance. The limited emergence of SACD source material, some with genuinely extended bandwidth content, has prompted on the hi-fi market the sale of costly, accessory super tweeters, devices with above 20kHz frequency responses. In addition loudspeaker designs have been introduced, some at substantially higher prices, claimed to have a greatly extended treble performance. The proposition 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 DAC ideally employed, is so great as to have been known to destroy a noted audiophile amplifier before the symphony got going. 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 restricts the bandwidth to 50kHz despite the company claiming a reference grade appendation 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 despite the bandwidth and resolution advantage over CD, DVDA and SACD have made little market impact in the quest for higher fidelity replay, mainly for commercial rather than technical reasons. Imminently there are two new disc carriers, HD DVD and Blu-ray, which while essentially video directed have such enormous storage capacity that the potential for multiple wide band audio channels is embedded in their standards. Will these become the super audio carriers of the future?

2.4 High Frequency Reproducers

If of value, we nevertheless need get this extra bandwidth to the listener. Considering elements of the replay channel, such as amplifiers, 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 located at 25kHz or so to combat their propensity for serious EM radiation. Loudspeakers are an important issue, and taking sensitivity is a key parameter we know that diaphragm area is contributory. Reducing this area is essential to maintaining directivity and power response to much higher frequencies but this requirement conflicts directly with sensitivity. If pistonic, we need very small sources, ideally 10mm or less in diameter, for useful output to 100kHz.

The alternative of larger, louder but intendedly resonant diaphragms of potentially wider directivity, may well result in a wildly discontinuous frequency response. If such were provided down in the normal audio band it would be dismissed without question. I have measured famous brand 'high range' drivers whose 20kHz to 50kHz range was confined a few high Q resonances, these emanating from a 25 mm metal dome.

Room absorption continues to increase at higher frequencies. In addition, in the higher range, the ear itself is more directional, as HRTF work has shown, and any possible benefit of higher frequencies would have a better chance arriving on a reasonably direct path to the entrance of the ear canal. Can we really benefit from the investment if it seems to be so difficult to deliver the proposition to the listener's ear?

2.5 Anecdotal Consumer Press Reporting

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 generally the basis of their evaluations has generally been so flawed that the results may be discounted. The difficulties encountered are manifold and only a few need be considered to confirm my negative view. Firstly the test for response extension benefit will only be valid if the extended response is achieved without affecting the performance in the existing 'audible' range.

Testing for a subtle effect, which may well be barely audible, is a manifest nonsense if it changes the uniformity and loudness in the already operative treble range. Yet this is what is happening in these tests. So far, no commercially available add-on tweeter and matching crossover can avoid this fundamental error. Note that 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.

These 'super tweeter's typically operate in parallel with the existing tweeter over about an octave bandwidth, and may destructively interfere with the primary tweeter output. Thus there is a potential to impair as well as alter the results. Sadly, some critics are so pleased to have heard a difference they are tempted simply to judge it as an improvement. Often the crossover is a simple capacitor feeding a metal dome tweeter, perhaps with a beryllium composite diaphragm. Such a crossover 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 is compared with the crossover objective, which is seen to be markedly different from the practical result generally obtained with such a single element filter. The cause is the complex impedance presented by the high frequency driver compared with a plain resistor load. 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 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 demonstrable super-tweeters illustrated 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 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 audiologists data considered well consolidated, the industrial use of sound has become firmly established. Some at first sight shocking information may be gleaned from the industrial arena, which employs high power sound energy in the field of ultrasound. Referring to health and safety literature, there is much solid teaching on audible bandwidth, and 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. This is about 100dB down by 40kHz. Acoustical energy above 20kHz typically, is 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 which on the face of it would cause concern to an audio specialist. An example is the use of industrial cleaning baths capable of 110dB spl at 35kHz. In fact ultrasound is widely exploited 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 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. 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 safe otherwise for patients and operators. Nevertheless we happily beam ultrasound at a foetus for diagnostic imaging, and assume that it is safe, and 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 from the outside environment.

2.7 More Recent Perception Tests

A series of tests for subjective hearing response were conducted at the London AES convention 1980 by Laurie Fincham[1], then Research Director at KEF Electronics. Although the results were presented at an accompanying seminar, the work remained unpublished. Laurie, now at THX, had kept a copy, and on my request kindly sent a stack of paper by return. Laurie had succeeded in inveigling a galaxy of audio specialists to participate, 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 has been 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 unit 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, aurally innocuous. 80dB or more attenuation was present by 1.25 times the nominal cut-off frequency. In contrast to many hearing tests which employ sine wave or band-limited signals, here a stimulus closer to a full and natural transient was chosen.
Summarising the results, which had been subject 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 nearly everyone. There was a sigh of relief all round! However for the 16kHz filter setting there was obvious difficulty. This limit was only 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 20kHz Fincham bandwidth test were assessed, it was clear that statistically at least, all the tested listeners, if working under somewhat less than ideal conditions, could not hear this applied limit. These results confirmed conventional wisdom.

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

In addition to the conscious scoring from the participants, physiological and brain wave analysis, including PET scanning, was also carried out. The latter has the ability to assess change in 'pleasure' even if the panelist is not directly aware. The results revealed some fascinating contradictions. While full bandwidth was preferred to CD bandwidth, the test for the ultrasonic music band showed that it was inaudible when reproduced alone. Further, exposure to the full bandwidth stimulus seemed to additionally confer some conditioned improvement for subsequent reproduction at reduced bandwidth, as if there was some kind of beneficial short-term memory effect. Much debate has followed this discovery. Finally the full bandwidth replay was judged from the physiological data to give 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 it had become audible only when meaningfully associated with the audible, sub-ultrasonic range.

Lenhardt et al[3] showed that speech could be heard in the ultrasonic range, if modulated on to an ultrasound carrier and presented via bone conduction. The inner ear has some mechanism for demodulating the speech signal from the carrier. This non-linearity raises further issues concerning how 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 at least to 40kHz. This was not widely accepted until this material was published.

Black[5] has considered whether for some tests the matter of ultrasound audibility was in fact due to another possible mechanism. 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 2000 Howard[6] usefully reviewed the many issues to date, partly from consideration of the increasing availability of the commercial 'super tweeter', but did not arrive at a firm conclusion on audible 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 of brain chemistry was proposed in connection with human responses to ultrasound.

Seeking to investigate further, 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. On my part I have found forced decision methods tend to impair subjective response sensitivity. However it is abundantly clear that great care was again taken by both these researchers to try and deliver a representative result.

Taken overall, and with some polite hesitation from the authors, the Nishiguchi results contradict Oohashi, and in practical terms reiterate that 20kHz is entirely sufficient as a nominal limit for sound reproduction. This later result is attributed in part to the use of very steep band filtering in the experiment, to 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 shielded from the ultrasonic component, then the perception of an extended bandwidth is no longer reported.

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

3 EXTENDED BANDWIDTH, MASTERING AND REPLAY

Regardless of the audibility issue you may still argue advantages for the proposition.
For origination a wider operational bandwidth may result in improved performance, or the equivalent of increased resolution, benefiting the established audible range. At present strong low pass filtering is often employed to fit a 20kHz bandwidth to many digital audio chains especially MP3 portables, broadcast, and CD. The use of a recording medium of extended bandwidth makes possible more gentle filtering, with improved phase and impulse characteristics. For loudspeakers the effort to extend bandwidth again may improve 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, may now be located well out of range. In Figure 2 the typical low loss, multi-resonant output of a rigid 'piston' 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 of all kinds. Human perception is notoriously difficult to quantify especially when differences are small, and some results may be misleading. I believe that elements of the audio industry have exaggerated the necessity for the consumer to enjoy an extended reproducing bandwidth. Judging from this review of the subject, the case for an ultrasonic reproducing bandwidth is not proven, and 20kHz remains the practical limit for the chain as a whole.

However for origination and post-production, within reason, 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

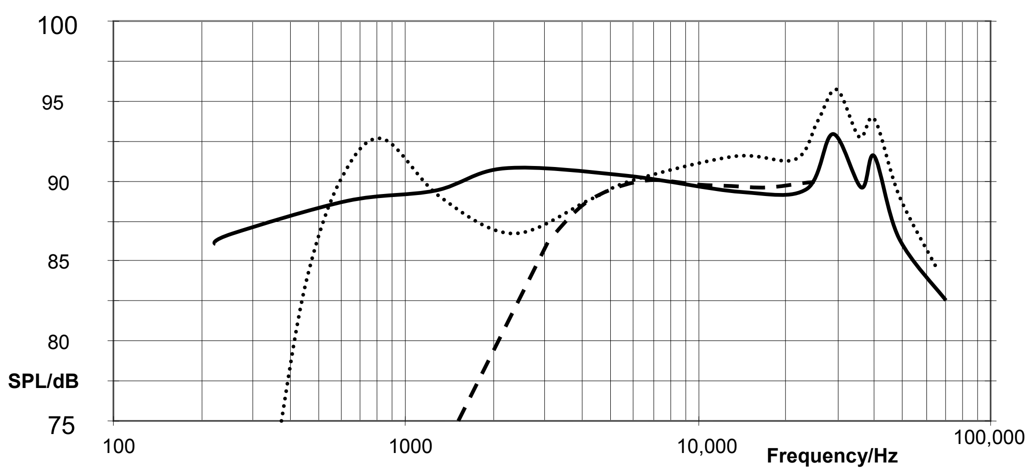

Figure 1, Behavior of 1st order, capacitor, high pass crossover.
Key ; Frequency response of typical driver (solid line)
Ideal crossover performance (dashed line)
Effect of single capacitor crossover (dotted line)

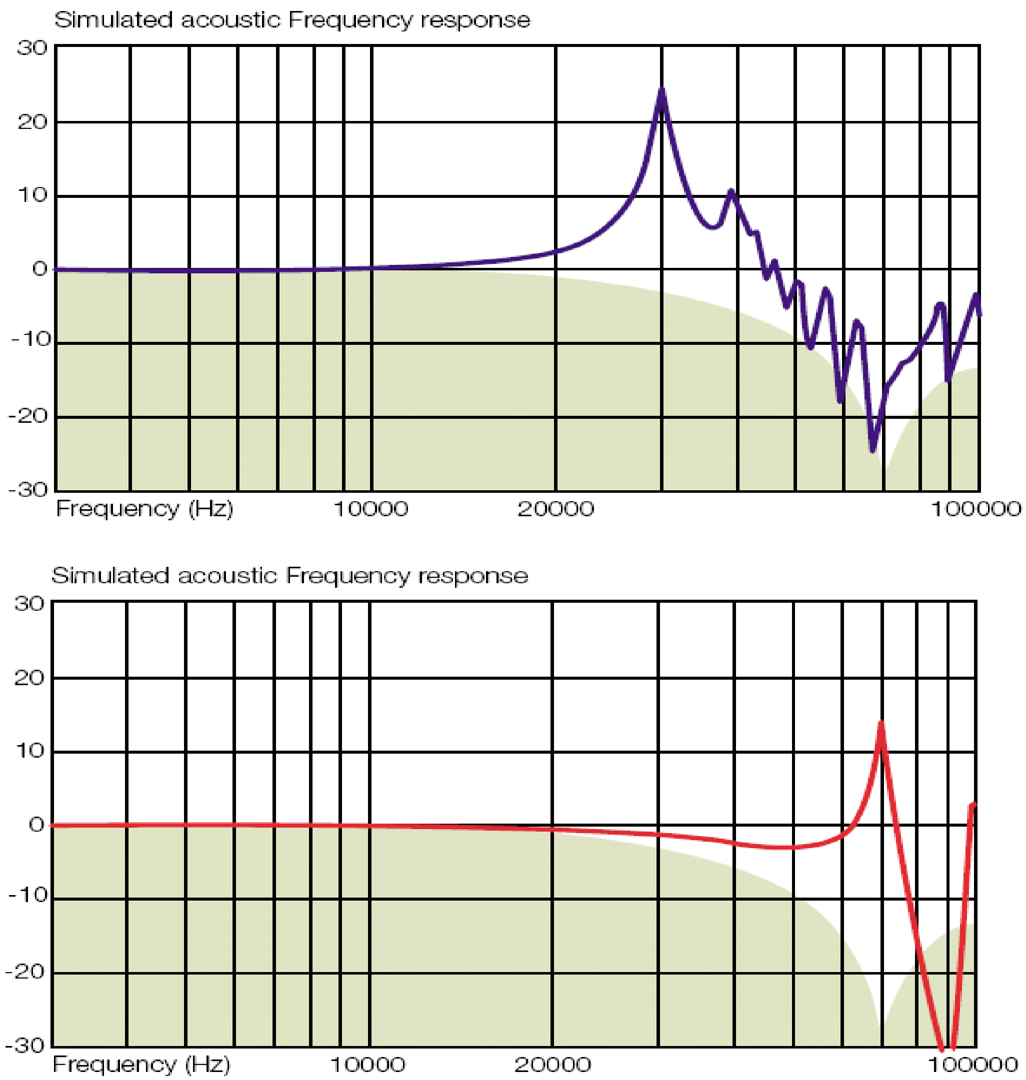

Figure 2 , FEA for axial radiation from a 25mm aluminium dome, upper graph, and for an equivalent in ion deposited crystal diamond dome, 50um thick, lower graph (courtesy B&W).