THE SIMULTANEOUS MEASUREMENT OF TIME AND FREQUENCY.

R. Walker British Broadcasting Corporation, Research and Development Department
Kingswood Warren, Tadworth, Surrey, KT20 6NP, UK

INTRODUCTION.

In the relatively recent past, significant attention has been paid to the effects of early reflections in control rooms and listening rooms [1]. The need has arisen to measure a number of parameters of the acoustic signal in short intervals of time. Many years ago (1945-46), Shorter used time domain gating followed by conventional frequency response analysis to measure the short-term responses of loudspeakers [2]. More recently, but still some years ago (1978-79), the present author used a similar method to measure the time-frequency responses of rooms [3]. In both cases, the methods were clumsy, inflexible and extremely laborious to apply, to the extent that no great use was made of them in either case. Modern methods are simple to apply, using commercially available equipment. Some proponents of such measurement methods have succeeded to the point where their methods are accepted essentially as *de facto* standards for such measurements [4,5].
The conventional way of describing the characteristics of acoustic or audio systems is as functions of frequency. However, all real signals, whether acoustic or otherwise, are inherently functions of time – in the sense that the signal only exists at all as some instantaneous measure of a physical quantity, which may (or may not) vary with time. The concept of frequency and the frequency domain is a mathematical abstraction, with no physical existence. It can be derived from other representations by appropriate mathematical operations[1]. It is well known that the impulse response of a linear, time-invariant system theoretically contains all of the information necessary to specify the system response fully. However, the impulse response is a time domain function. In practice, it may or may not effectively be constrained in the frequency domain, depending on the equipment used to measure it, but there is no information in the impulse response to indicate the

[1] The mapping of the Fourier transform onto the basis set of orthogonal, trigonometric functions is only one of many (possibly an infinite) set of possibilities.

frequency content directly. For all of these reasons, it is important to understand the transformations from time domain to frequency domain, and in particular the inherent resolution limits of the joint frequency-time space.

MEASUREMENT OF TIME-FREQUENCY RESPONSES IN ROOMS.

The human hearing system is a complicated signal processing system, especially in the context of the effects involved in a stereophonic audio illusion. The following is a very brief summary of a great deal of psychoacoustic research, quoted without references because of the very large (and sometimes contradictory) body of evidence. It is included just to illustrate the ideal requirements for the performance of a measurement system.

In the interval up to about 5ms after the arrival of the direct sound, any delayed sound is entirely integrated with the direct sound. The sense of direction is governed by the arrival of the direct sound. From about 5ms to about 20ms, some information is extracted about the direction of a sound source but mainly to form some impression of the sound 'quality'. After about 20ms the extracted information is largely about the surrounding space rather than the source itself. After about 50–80ms, the sound becomes distinctly reverberant and wholly related to the surrounding space.

For reflections in a small room, the interval up to about 15–20ms is the most important. Although in the period up to 5ms reflections are not perceived directly, they can have a profound influence on the sound quality. For example, a delayed signal with a relative time offset of 1ms could produce strong cancellations at odd multiples of 500Hz by interference with the direct sound. In most cases of very short delays, reflecting surfaces would not be large enough to cause strong reflections at 500Hz, but might at 1500Hz and certainly would at higher harmonics. Such cases are frequently encountered in control rooms, where the top surface of the mixing desk usually forms an efficient reflector. The potential reflection from the mixing desk is likely to arrive at the listener about 1ms after the direct sound from the loudspeakers.

Reflections later than 5ms can disturb the subjective stereophonic imaging process, causing mislocation of images. Individual early reflections from room surfaces are likely to occur at about 3ms (from the ceiling), 5–8ms (from the side walls) and 15–20ms (from the rear wall). Thus, for the measurement of early reflections in control rooms, it is desirable to be able to resolve time differences of the order of 1ms. It is also desirable to obtain as high a frequency resolution as possible, implying longer time records, in order to obtain some idea of the frequency characteristics of any reflections. These are clearly conflicting requirements. Fortunately, the stereophonic illusion process involves mostly the higher frequencies. Although some sense of spaciousness is conveyed by frequencies between about 300 and 500Hz, the main image-forming frequencies are those from about 500Hz upwards.

These factors lead to measurement processes based on time resolutions of about 1 – 2ms, resulting in frequency resolutions of 0.5 to 1kHz. It is, as a result, conceptually possible to identify and measure reflections with time and frequency resolutions high enough to be useful for the investigation of stereophonic systems in relatively small rooms. (However, it is also likely that the ultimate performance of the human hearing system exceeds anything that can currently be implemented by instrumentation.)

There are two particularly useful measures of time domain responses – the so-called 'Energy-Time' response (ET) and the 3-dimensional Energy-Time-Frequency response (ETF). The first of these is in fact the square of the complex system impulse response. It is taken to represent 'instantaneous' energy. Although the theoretical nature of the response has been the subject of some discussion [8], it does present a view of the time domain response in which representations of discrete reflections can be observed. The measurement may include some form of frequency-domain pre-filtering and weighting of the raw impulse response in order to highlight particular features. In some implementations, that frequency response processing is, at best, unclear but it can have a pronounced effect on the results obtained. Its investigation was one of the main objectives of the work described in this paper.

The second useful measure of response is the 3-dimensional, 'ETF' plot. For this, the time domain impulse response is translated to the frequency domain by discrete Fourier transform, with a relatively short time window. The location of the window is progressively shifted to later times, to produce a series of frequency responses calculated for different time intervals. The frequency resolution is limited by the windowing process to some proportion of the inverse of the window length (depending on the type of window). It is also limited in time resolution by the temporal response of the window. Despite these limitations, a useful display indicating approximate times and frequency responses of reflections can be obtained. It can also clearly be seen that the resulting time and frequency resolutions are defined by the window functions in exactly the same predictable way as for any other Fourier transform process.

INSTRUMENTATION.

One convenient instrumental system which greatly simplifies the measurement and post-processing of impulse response functions is MLSSA (Maximum Length Sequence System Analyser)[4]. It comprises source signal generator and response recorder which works not on the impulse response directly but by using a comparatively long-duration, noise-like test signal. The actual impulse response is obtained by cross-correlation of the measured response with the original signal. That results in a significant degree of noise rejection and a more reliable measurement. The system also includes a wide range of post-processing functions.

Many other measurement methods are available. In some, the time and frequency weightings and resolutions that are effective for any particular condition are clearly evident. In others, the method of operation tends to obscure the effective settings. This article is not concerned with any particular, proprietary method of

measurement. It does, however, rely on measurements carried out using a MLSSA system to illustrate general issues.
References 1 and 9 give some examples of a large number of measurements carried out on early reflections. In those real cases, significant irregularities in the frequency responses of most reflections gave rise to the apparent discrepancies between the ET and the ETF responses that are the main subject of this article.

EXPERIMENTAL RESULTS.

In all of the following measured results, unless it is otherwise made clear, the response amplitudes are given in decibels relative to the level of the direct sound and the time delays are given relative to the arrival of the direct sound at the measurement microphone.

Overview and earlier observations.

In the initial stages of the work on early reflections in rooms [1] it was observed that the selection of filter, window, transform lengths, etc. had a pronounced influence on the results obtained. For example, in some cases where the ET response apparently showed reflections of the order of –20dB or lower, the ETF response showed narrow band reflections higher than –10dB at some frequencies.
Figs. 1 and 2 show an example from Reference 9. The reflection at about 7.5ms appears in the ET response to be at a level of about –22dB, whereas the ETF response shows it to be at about –10dB at 3800Hz (the overall system gain was such that the direct sound measured approximately +3dB). The ET response was calculated with a Blackman-Harris window, which effectively rejects just the extreme high and low frequencies, leaving most of the middle frequency range evenly weighted. The ETF response was measured with a 4.26ms transform length, with a half-Hann window, giving an effective frequency resolution of about 300 – 400Hz.
Figs. 3 and 4 show another example – one with a high amplitude and irregular reflection response. (The example was taken from a fairly typical, conventionally-designed control room, not one that had been the subject of early reflection controls). The very severe and irregular reflection pattern gave rise to large differences between the ET and the ETF responses. The level for the reflection at about 3.6 ms in the ET response, Fig. 3, appears to -6.75 dB. In the ETF response, Fig. 4, the same reflection appears to be at -0.6dB at about 4400Hz. Actually, the direct sound level at the same frequency also measured -0.6dB so that the difference was actually 0dB! In both of these cases, the same actual impulse response data was used for both types of analysis. Clearly, significant discrepancies can arise between the results given by different analysis methods.
The reason for these differences is the difference in the effective bandwidths of the two measurement processes. In general, any measurement system will produce a result which is some kind of weighted average of all of the information falling within its scope. The wideband ET response produces a value for the reflection amplitude that is the square of the sound pressure, averaged over the whole effective

frequency range. With an uneven frequency response, the amplitude of any particular, higher-level components will be reduced by the inclusion of lower-level components in the averaging. In contrast, for the relatively narrow-band ETF analysis, the different amplitudes of the constituent components will be more accurately represented because the bandwidth doesn't encompass more of the spectrum than is occupied by the isolated response feature. The same effect is also observed if the ET response is restricted to narrower frequency bands. Fig. 5 shows the same ET response as in Fig. 3, but with a half-octave wide filter centred on 4400 Hz. The apparent ET reflection amplitude was changed to 0.0 dB by the change to the narrower-band analysis.
All of these effects are, in principle, easily predictable. What is less immediately obvious, and may cause practical difficulties, is the more or less obscure way in which some of the available measurement systems operate and the resulting effective time and frequency resolutions. In some cases, the measurement system responses are so heavily weighted to the higher frequencies that they are only effectively measuring the extreme upper end of the spectrum. That may even be outside the normal audio range.

Experimental synthesis of a single reflection.

In order to illustrate the effects of measurement bandwidth under controlled conditions, a simulation was set up of a room with a single reflection. Fig. 6 shows the schematic arrangement. The system under test was a direct connection, with the addition of an electronically delayed signal. The delay was set to 5ms. The output signal was taken as the mixture consisting of the input test signal and the delayed signal at a relative level of nominally -6dB. A bandpass filter, of nominally 4kHz to 6kHz bandwidth, could optionally be inserted into the delay path. The filter actually produced a passband gain of approximately +2dB, making the filtered delayed signal about -4dB relative to the direct signal. Because of the simple passive mixer, there were also other signals corresponding to multiple passes through the system. The digital sampling frequency was 30 kHz and the overall measurement system bandwidth was 10 kHz.

Broadband delay.

Fig. 7 shows the ET response obtained without the bandpass filter using a Blackman-Harris window. The first peak, at 5ms, was measured to be very close to –6dB relative to the direct sound, as expected. Fig. 8 shows the ETF response, obtained with a 4.26ms half-Hann window. That produced an effective frequency resolution of about 400 - 500 Hz. The delayed signal was measured as about -6dB relative to the direct signal, though the results do also show a slight frequency-dependency of the delay device (between about -5 and -7 dB).
For this condition, the delayed signal had an essentially uniform spectrum. Differences in effective bandwidth or spectrum averaging in measurement systems would not be expected to cause differences in the measured results – and nothing significant was observed.

Band-limited delay.

Fig. 9 shows the Fourier transforms of the main and the delayed signals, obtained using the maximum possible time window, of just less than 5ms, with half-Hann weighting. The lowest reliable frequency and the frequency resolution are indicated by the bar at the bottom of the graph (just over 250Hz).The delayed signal can be seen to be centered on 5kHz, with a relatively broad response, 3dB down at approximately 4kHz and 6kHz. The measured value of the delayed signal at 4904Hz was 4.48 dB below the direct signal - within 0.1 dB of the measured steady-state level difference. The response irregularities at low frequencies and the trend to a dc level of about -20dB were caused by a low-frequency transient recovery effect. They are not relevant to the current discussion.

Fig. 10 shows the ET response obtained using a Blackman-Harris window. The first peak, at 5ms, appears to be at a level of –12.8dB relative to the direct sound. That was a measure of the average ‘energy’ levels of the two signals over the effective (and relatively wide) measurement bandwidth.

Fig. 11 shows the ET response obtained with pre-filtering of the impulse response, using a half-octave wide filter centred on 5kHz. In that case, the filter bandwidth of nominally 4200Hz to 5900Hz just encompassed the width of the delayed signal response in the frequency domain. The measured apparent level difference of 4.49dB corresponded closely with that obtained from Fig. 9. The significantly poorer time resolution of the narrower band filter is also evident.

Figs. 12 and 13 show intermediate conditions, for one-octave and two-octave filters respectively. The progressive change in the apparent ratio of the 5ms reflection to the direct sound is clear, with progression to -5.8 and -9.8dB respectively.

Fig. 14 shows the ETF response. It shows the delayed signal at -4.45dB relative to the direct signal at 4934 Hz, in close agreement with the true level. In that case, the effective ±3dB measurement bandwidth was about 340Hz (= $\pm 1.44\Delta f$ for a 128 sample record at 30kHz sampling frequency). It was clearly narrow enough to represent just the passband component of the delayed signal. Clearly, there was no significant inclusion of remote frequencies in the measurement process.

DISCUSSION OF RESULTS.

It is self-evident that the final result from any type of measurement will be some form of weighted average of all of the data which falls within the measurement scope, in the time domain or the frequency domain or in a complex combination of the two. From the experimental model of a single room reflection, it has been shown that the measured results of ET and ETF responses were in accordance with the theoretical expectations, averaged over whatever time and frequency ranges were in effect at the time.

Thus, if a measured reflection amplitude is obtained from either ET or ETF responses, the effective ratio obtained for reflected to direct energy depends on the frequency-domain responses of the signal and on the relative bandwidth of the measurement.

The nominally unfiltered (wideband) ET response will produce an overall value for the ratio of the two signals, averaged over essentially the whole frequency range. For a reflection (or a direct signal) which has pronounced frequency variations, the amplitudes of the highest components will be underestimated. Narrower frequency bands may be used by effectively applying a frequency-domain, bandpass filter before calculating the response.

The ETF response will always produce a relatively high frequency domain resolution (at least in comparison with the unfiltered ET response), giving the true response at each frequency, within the limitations of the Fourier and Nyquist theories.

It may be argued that the time resolution provided by either an ETF or a band-limited ET response is too poor to be of much practical use. However, theoretical considerations limit the obtainable resolutions to about the equivalent of 1ms/1kHz. That is inherent in the principle of time-frequency analysis. Any system which appears to offer much higher time resolution is inevitably restricted to high frequencies only, or at least to include high weighting factors for the higher frequencies.

In the examples shown, differences in measured levels of between 6 and 10 dB between wideband and narrowband analysis methods have been presented. Whilst those are large enough differences to be of significance, they are still less than can be observed using some types of instrumentation. In the examples given in this article, the overall system bandwidth was limited to 10 kHz. The differences between the analyses would have been greater if the ‘unlimited’ system bandwidth had been greater. If it had been, say, 30 kHz (which appears from published results to be a common setting for some swept sine-wave instrumentation) then the differences would have been 5 dB greater still, based just on the difference in effective bandwidths. Other factors, for example non-uniform frequency-domain weighting functions, may make the differences larger still.

CONCLUSIONS.

The fundamental resolution limits of time-frequency measurements have been described. It has been shown that the practical limits of time and frequency resolutions that can be obtained are just about adequate to quantify those parameters important to the perception of the stereophonic illusion in the early reflection patterns in small rooms.

Experimental measurements, using a single electronically-simulated room reflection, have demonstrated that the results obtained correspond with the theoretical expectations.

It has also been shown that a reasonably accurate knowledge of the effective measurement system bandwidths and resolutions is essential before the results can be interpreted properly. With appropriate instrument settings, the results obtained for the responses of relatively short time events can accurately represent the actual physical conditions - within the inherent limitations of time and frequency resolution.

REFERENCES.

1. Walker, R. A new approach to the design of control room acoustics for stereophony. Paper G1-1,Preprint #3543, 94th AES Convention, Berlin, March 1993.
2. Shorter, D.E.L. Loudspeaker transient response. BBC Quarterly, I, 3, October 1946.
3 Walker, R. A preliminary investigation into the measurement of time and frequency reponse of listening rooms and control cubicles. BBC Research Department Report No. 1979/9.
4. D'Antonio, P. and Konnert, J. Complex time-response measurements using time-delay spectrometry. J.A.E.S., 37, 9, September 1989.
5. Rife, D.D. "Transfer-function measurement with maximal-length sequences". J.A.E.S.,37, No. 6, June, 1989.
6. Harris, F.J. On the use of windows for harmonic analysis with the discrete Fourier Transform. Proc. I.E.E.E., 66, No. 1., January 1978, pp. 51-84.
7. Gade, S and Herlufsen, H. Use of weighting functions in DFT/FFT analysis (Part 1). Bruel and Kjaer Technical Review, No. 3, 1987.
8. Vanderkooy, J. and Lipshitz, S.P. Uses and abuses of the energy-time curve. J.A.E.S., 38, 11, November 1990, p. 819.
9. Walker, R. "Early reflections in studio control rooms: the results from the first controlled Image Design installations". 96th AES Convention, Amsterdam, 1994, Paper P12-6 (preprint no. 3853).

ACKNOWLEDGEMENTS.

This paper presents work carried out at the Research and Development Department of the British Broadcasting Corporation and is published with their authority.

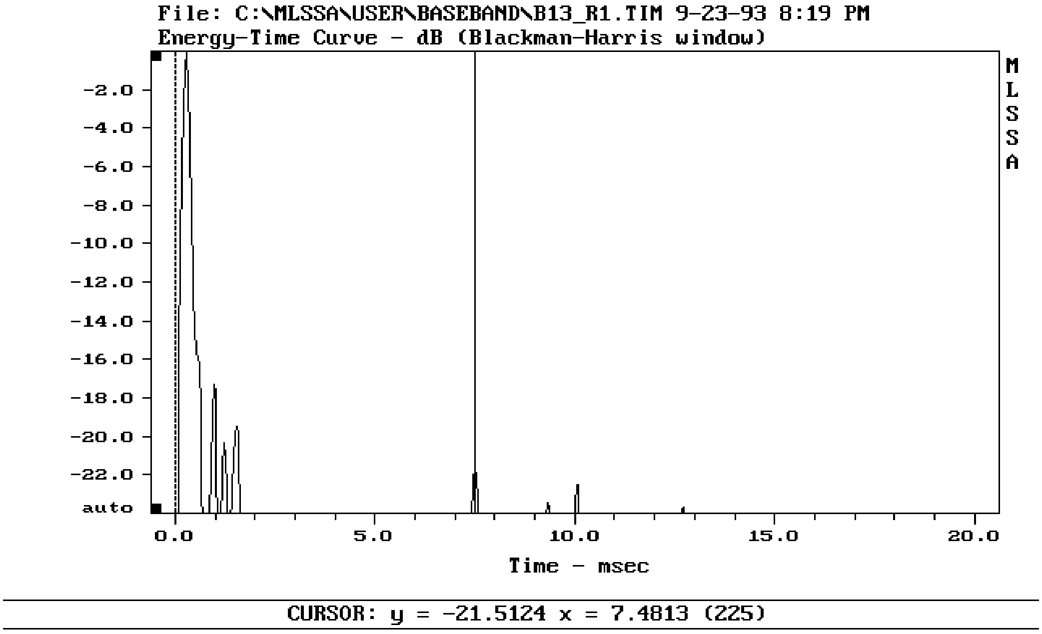

Fig. 1 ET response for controlled reflection room, showing reflection amplitude < -20dB at ≈7.4ms delay

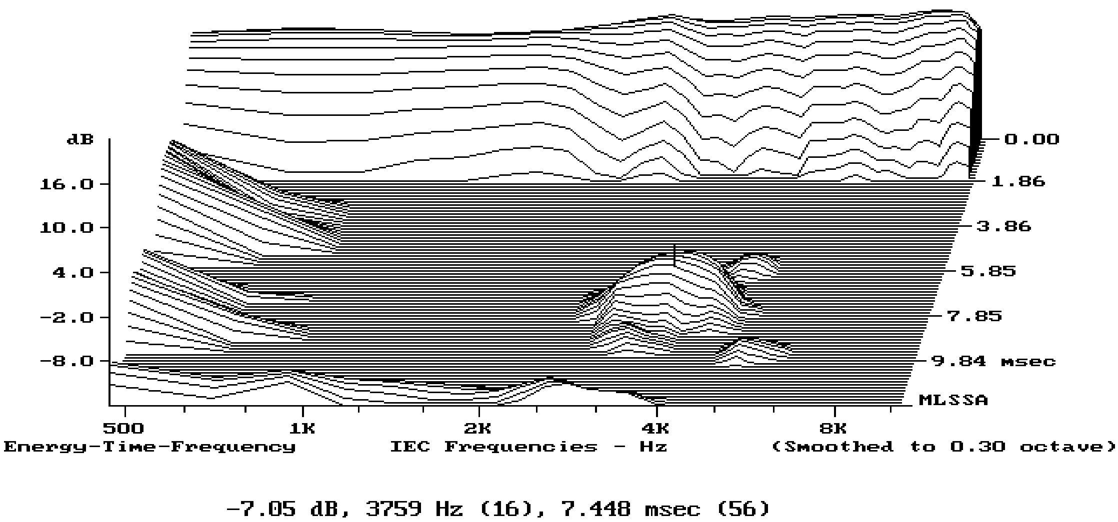

Fig. 2 ETF response for controlled reflection room, showing reflection amplitude ≈ -10 dB at 7.4 ms delay and 3800 Hz.

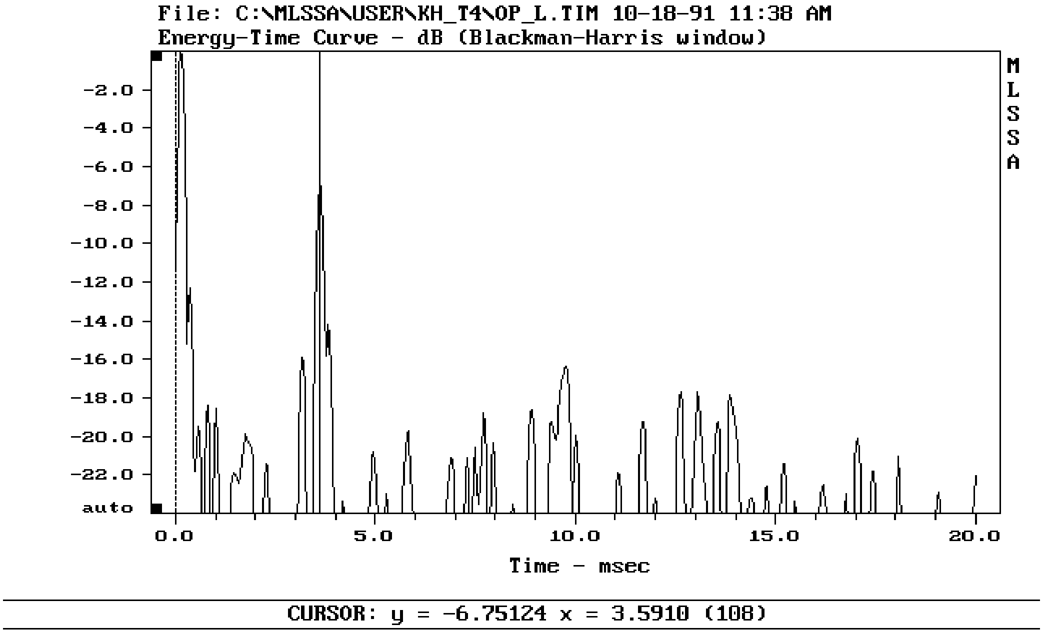

Fig. 3 ET response for conventional control room, showing -6.75dB reflection amplitude at 3.6ms delay

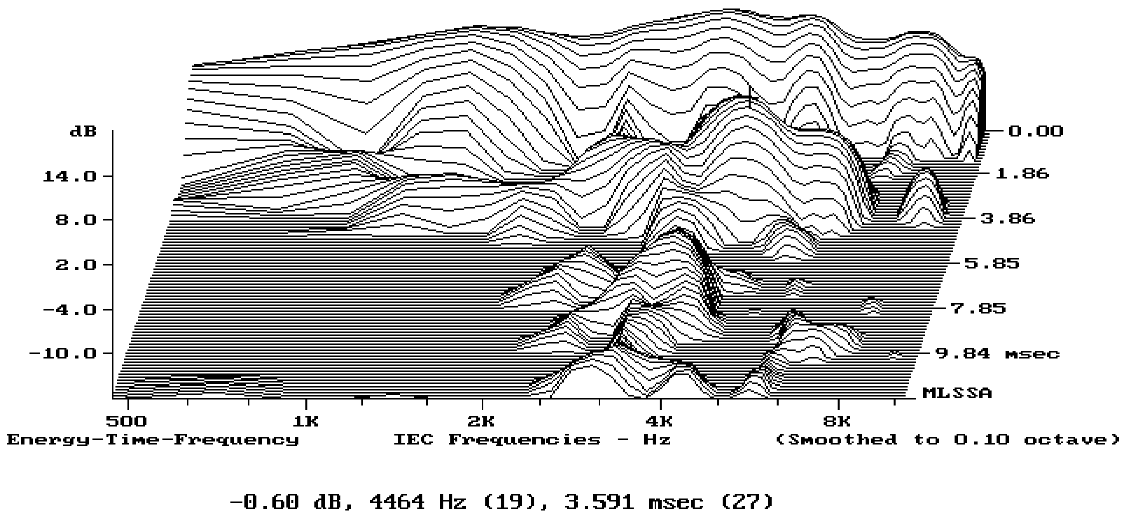

Fig. 4 ETF response for conventional control room, showing -0.6dB reflection amplitude at 3.6ms delay and 4500 Hz.

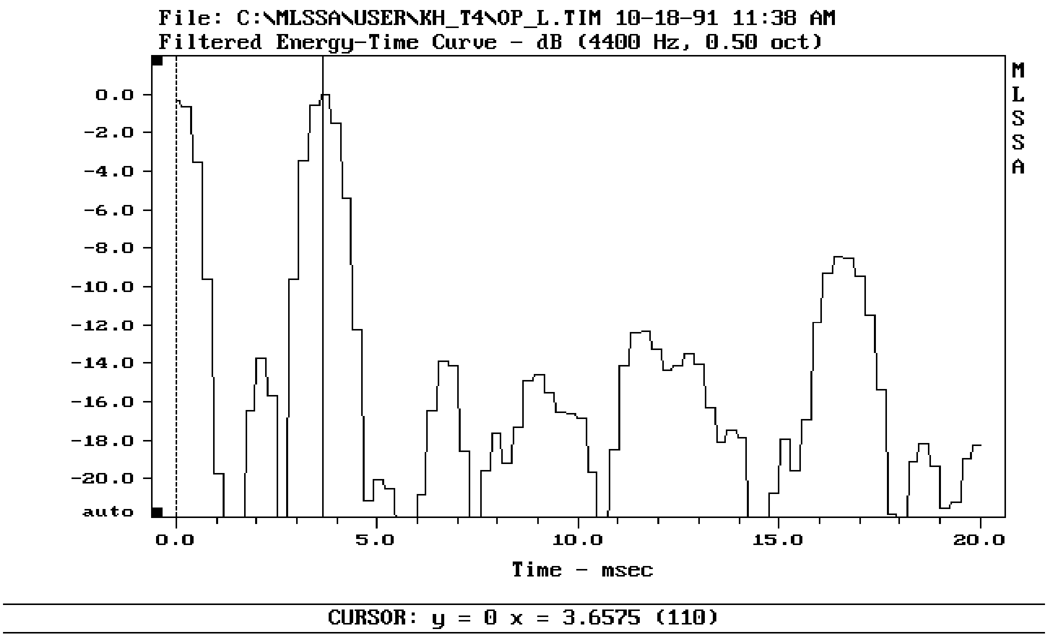

Fig. 5 Filtered ET response for conventional control room with 0.5 octave filter, showing 0.0dB reflection amplitude at 3.6ms delay.

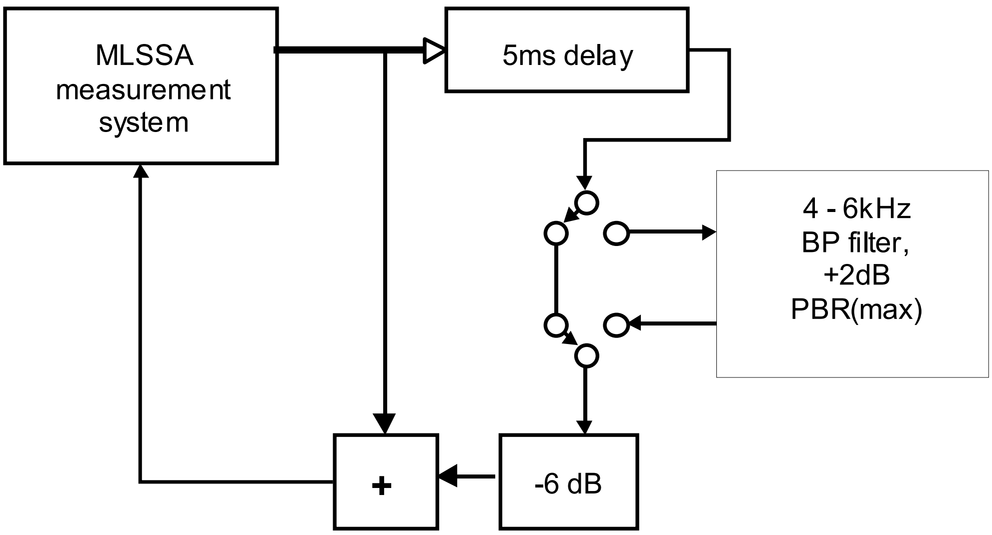

Fig. 6. Schematic of electronic reflection simulation system.

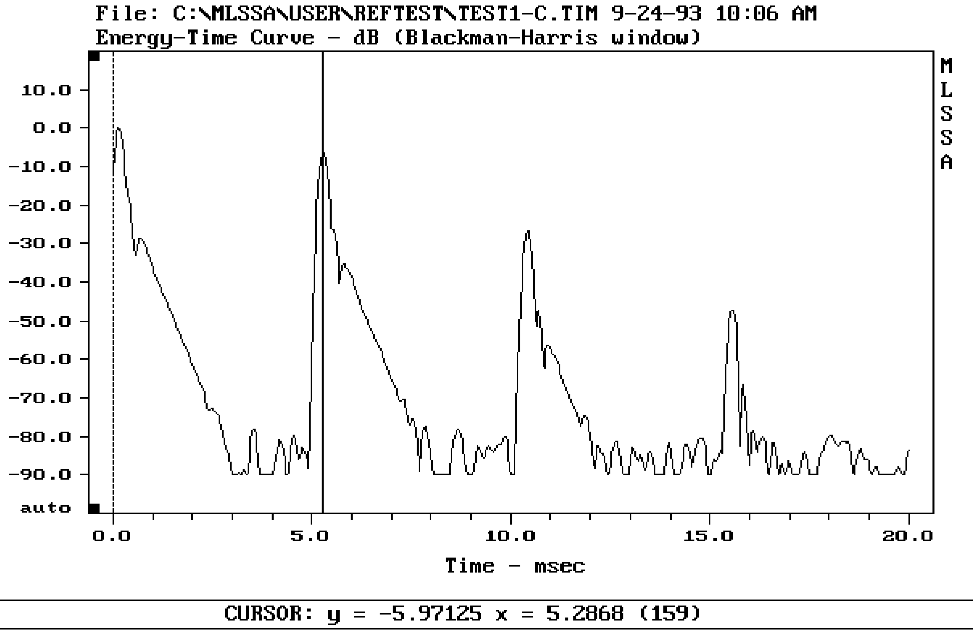

Fig. 7 Wideband ET response of unfiltered electronic reflection simulation system, showing -6.0dB reflection at 5.3ms.

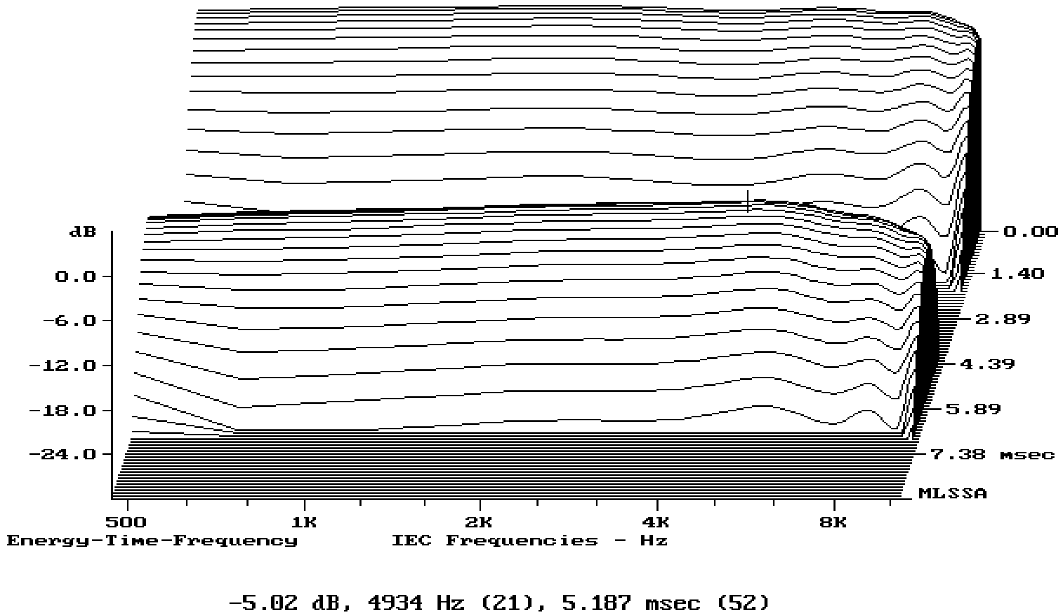

Fig. 8 ETF response of unfiltered electronic reflection simulation system, showing -5 to -7 dB reflection at 5.0ms.

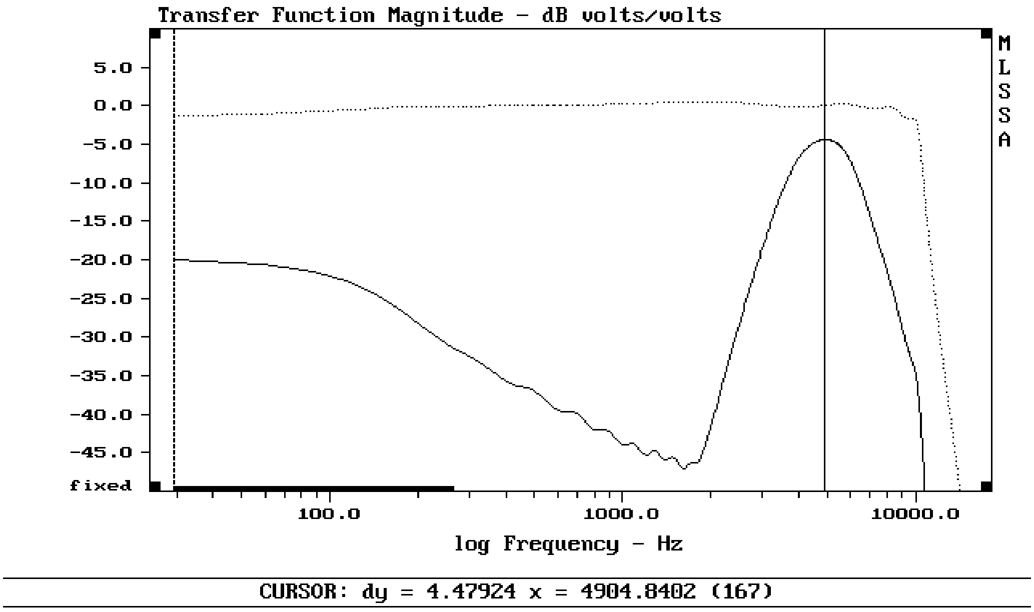

Fig. 9 Fourier transforms of direct and delayed signals for bandpass filtered electronic reflection simulation system, 4.8ms half-Hann window, showing -4.5dB response at 4904Hz.

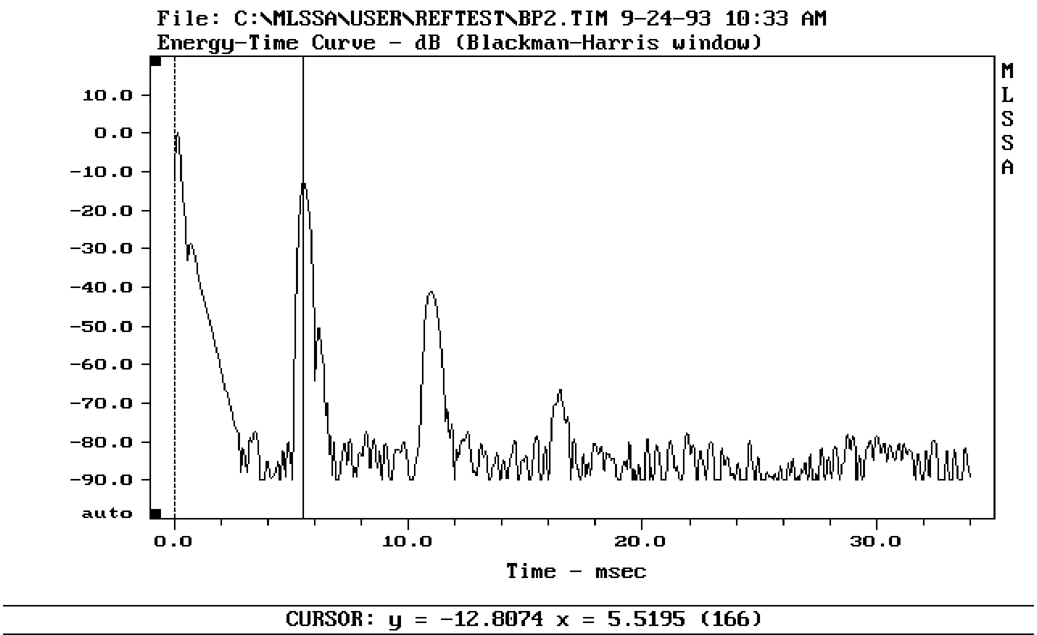

Fig. 10 Wideband ET response of bandpass filtered electronic reflection simulation system, showing -12.8dB reflection at 5.5ms.

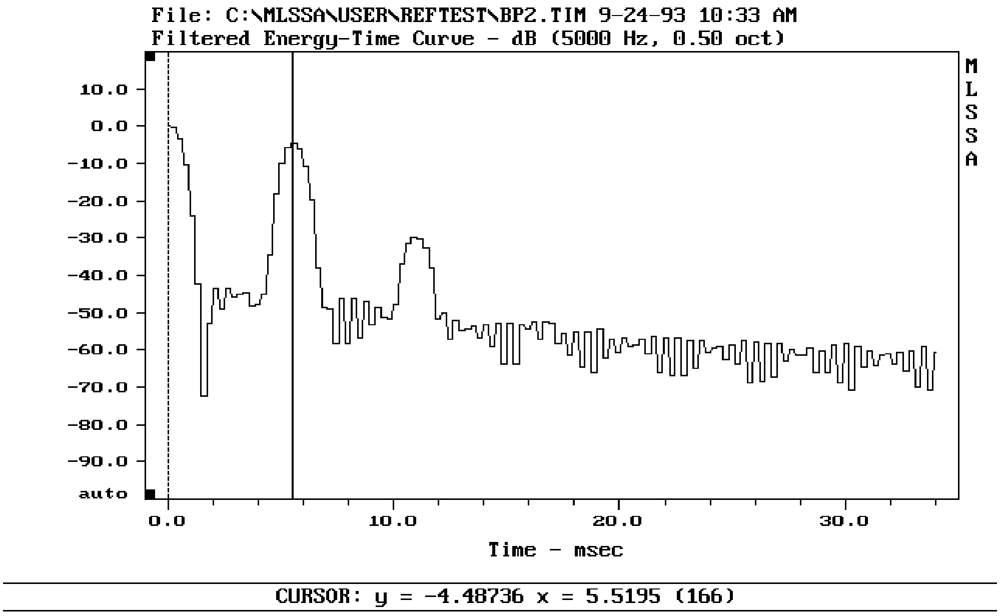

Fig. 11 Half-octave ET response of bandpass filtered electronic reflection simulation system, showing -4.5dB reflection at 5.5ms

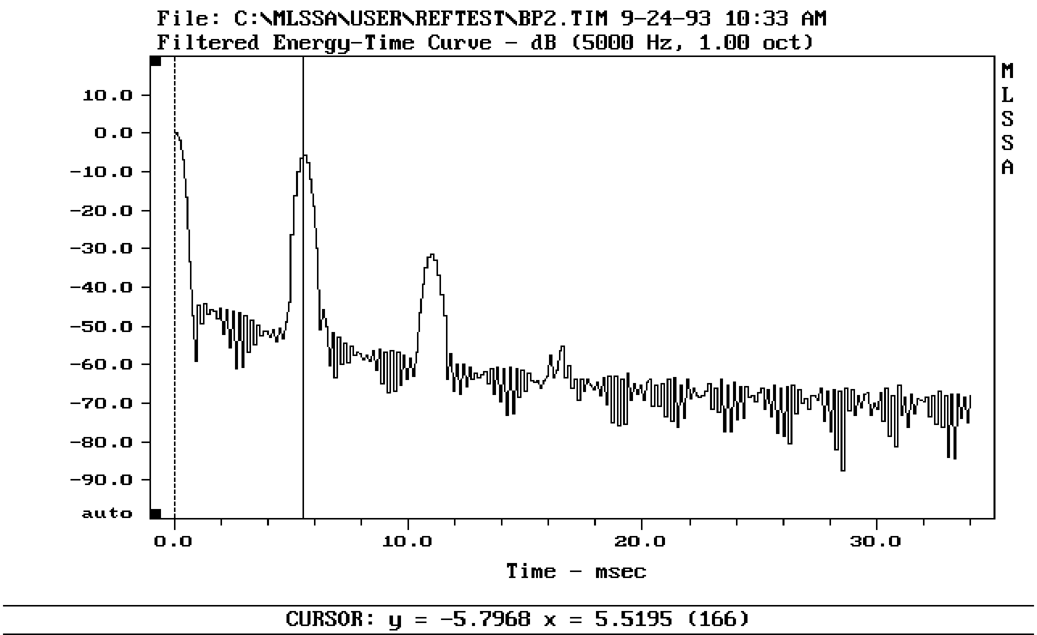

Fig. 12 One-octave ET response of bandpass filtered electronic reflection simulation system, showing -5.8dB reflection at 5.5ms

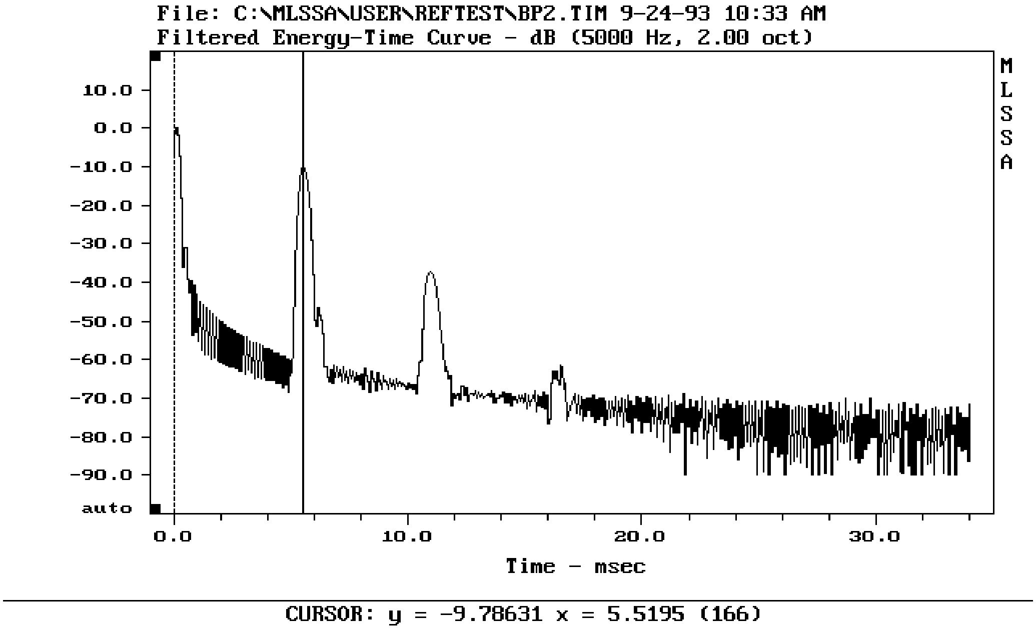

Fig. 13 Two-octave ET response of bandpass filtered electronic reflection simulation system, showing -9.8dB reflection at 5.5ms

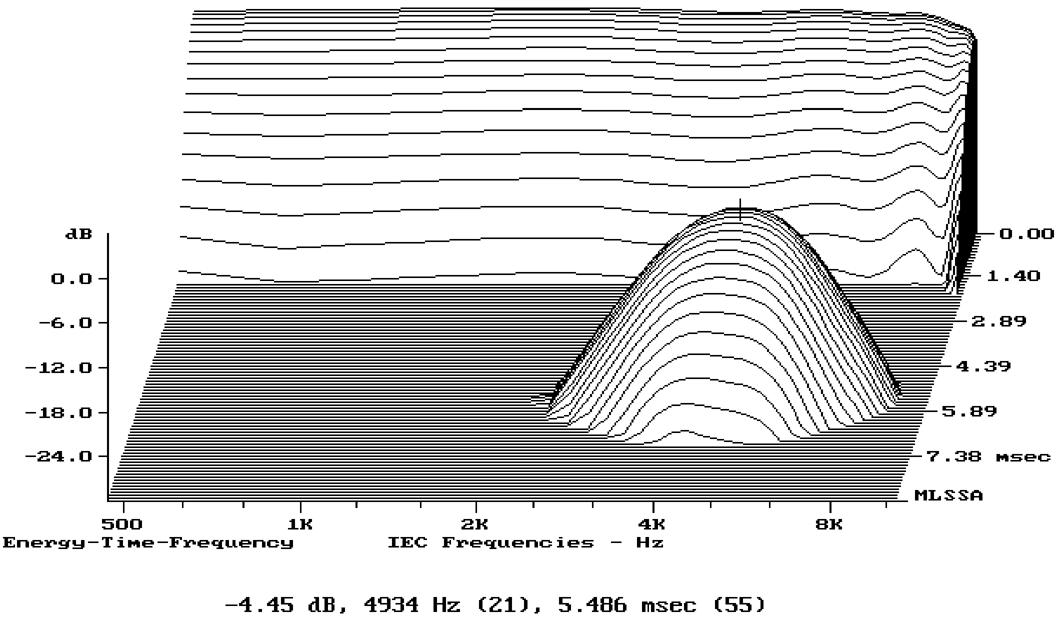

Fig. 14 ETF response of bandpass filtered electronic reflection simulation system, showing -4.5dB reflection at 5.5ms.