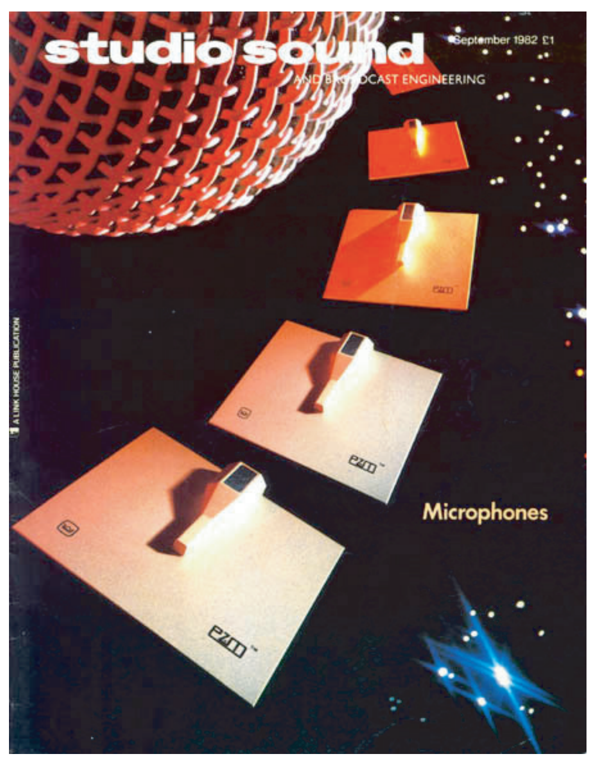
Developments In Recording Acoustics & Control Room Acoustics

By

Andy Monro

Studio Sound

October 1980
July 1982
Septenber 1982



Developments in recording and monitoring acoustics

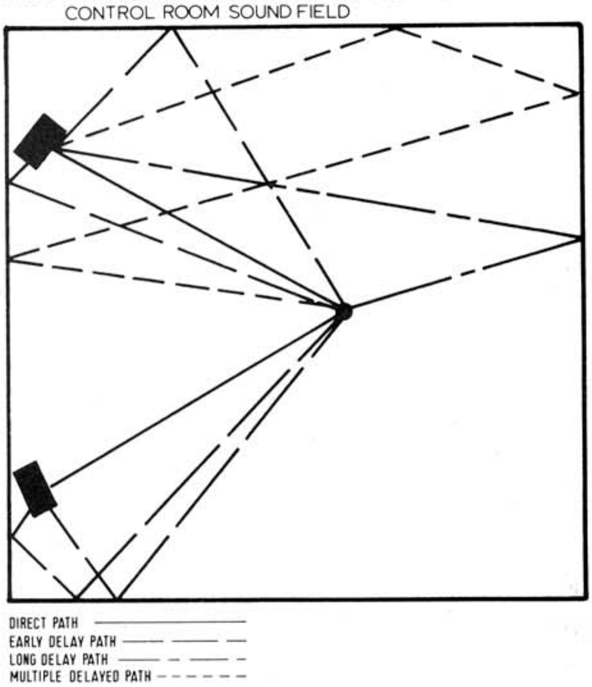
Andy Munro (Turnkey Two)

As the professional recording chain becomes less and less distorted, the performance of the acoustic environment has increased in significance to the point where many traditional design concepts can be shown to be inadequate. It is the purpose of this article to bring attention to some of the latest developments in acoustic technology.

THE FOCUS of attention in the modern recording studio is the control room, since it is here where the final sound is 'created', manipulating a multitrack master tape into the recorded product the discerning public has come to expect. In order to judge the quality of the sound originating in the studio, it is necessary that the monitor system maintains total neutrality and it is in the attainment of this that the greatest effort should be expended.

The monitor system itself consists of every element between the console output and the engineer/producer's ears which does not exist in the console/tape machine chain. The loudspeakers have always been considered the weak point in the hardware part of the chain, being the source of distortion and coloration far in excess of anything else. However, new developments in power handling and phase correction have eliminated many deficiencies to the point where the control room itself comes into question as a source of signal degradation.

FIG.1 DEVELOPMENTS IN RECORDING AND MONITORING CONTROL ROOM SOUND FIELD



Acoustic environment

Fig 1 shows a typical sound field development in a control room and it can be seen that many path lengths must occur between loudspeaker and listener. These path lengths may be defined as follows:

(a) direct path, effectively anechoic;
(b) early path lengths, arriving up to 10ms later than the direct sound;
(c) long path lengths, 10 to 20ms behind (a);
(d) reverberant sound field created by multiple path lengths and defined in terms of the RT_{60} of the room.

The reflections having the most disruptive effect are (b) which cause phase cancellation at frequencies defined as $F = \frac{500}{\Delta t}$ where Δt is the path difference in ms.

Path lengths in excess of 10ms can, of course, produce similar phase cancellations, but they have much less disruptive effect for the following reasons. Their amplitude is considerably lower as they approximate inverse square law decay, for at least some of their journey. Their intensity may be reduced by two or more reflections from absorbent surfaces, and they are also subject to the 'Haas effect' which is the ability of the ear to discriminate between direct sound and echoes. In the region of 15ms to 20ms of delay, human hearing is particularly aware of this effect which leads it to ignore such delayed information in the presence of the direct signal. This can be demonstrated by placing such a delay in one channel of an otherwise equally aligned stereo sound system. The listener will be aware only of the direct channel, almost totally ignoring the delayed sound. When the delayed channel is switched off, however, the ear detects a change in the ambient sound field and usually finds the sound less interesting.

It becomes apparent therefore,

that the ideal monitoring sound field consists of the direct sound path free of early interfering reflections followed with a highly diffuse ambient field consisting of multiple reflective paths with the first path difference greater than 10ms.

Time delay spectrometry

Much of the work involved in the measurement of the effects of room reflections has been made possible by time delay spectrometry (TDS), a technique developed by Richard Heyser of the California Institute Research Foundation. Basically the system consists of a swept filter which tracks the output of a synchronous swept sinewave generator with a specified lag time. By adjusting the lag time (hence TDS) it is possible to measure the direct signal or any successive reflection with no interference. The time window can be adjusted to measure direct and early reflections together and it is this analysis which shows severe phase cancellations in traditional monitor systems and rooms.

Another useful tool in this aspect of time related measurement especially where rapid analysis is required, is the Ivie *30/17A* microprocessor analyser. This extremely powerful device is well documented elsewhere, but mention of its time related measurement capability is warranted here.

An external signal generator provides a steady sinewave which can be gated to provide a tone burst. The analyser may be set to open a measurement window which corresponds to either the direct or reflected pulse. This enables successive delay paths to be studied and corrective acoustic treatment provided as required. Alternatively an internally generated snare pulse allows wideband spectrum analysis at specific time intervals allowing the measurement of both frequency response and delay time simultaneously. **Fig 2** shows a typical setting Ivie *17A* in the gated time mode in order to determine the amplitude of the finest discrete reflection arriving at the measuring microphone. Care must be taken to allow for the initial response of the speaker system to the snare pulse especially with large bass drivers.

Successive reflections can be measured at different frequencies to construct a time v energy curve for the monitoring environment.

Such techniques would not be viable in a commercial sense were it not for the speed with which such measurements can be obtained. In terms of the construction time of the average commercial studio, time related analysis more than justifies itself as the acoustic environment can be pieced together bit by bit, eliminating the need for costly alterations when the final decorations are in place.

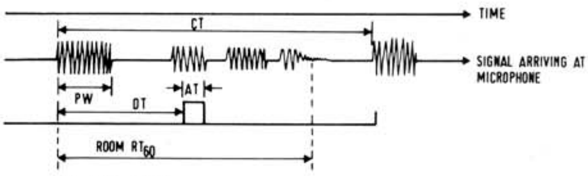
Live end-Dead end

In America such techniques have been employed for several years and a new generation of control rooms are emerging. Hardly a month goes by without news of the latest 'Live end-Dead end' (LEDE) studio opening in LA or New York. **Fig 3** illustrates the basic layout of a typical LEDE control room.

The concept of the LEDE room was evolved by Don Davis of Synergetic Audio Concepts and is the registered trade mark of this organisation. In order to qualify as an LEDE room a monitoring environment must satisfy the following criteria. There should be a solid non-resonant outer shell which is asymmetrical and large enough to develop extended bass frequencies. This should contain a symmetrical inner shell with a crossover frequency between the shells of $fx = \frac{3v}{1}$ where v is velocity of sound and 1 is smallest room dimension. There should be a direct monitor path free of early reflections for between 2 to 5ms (early reflections 15dB or less than the direct sound level). A highly diffused sound field should exist generated by the rear of the room with time delay gaps ranging from 10 to 20ms. A comb effect must be produced by careful spacing and narrow band interference with constant RT_{60} frequency. Also the monitor speakers should be mechanically isolated to eliminate structure-borne transmission which can result in sound arriving before the direct path as sound travels at greater speed through solid material. The speakers must be flush mounted and time aligned to the wall plane to avoid any diffractional effects.

By correct geometric design it is possible to eliminate early reflections simply by the exclusion of particular surface angles—it is true that a wall area cannot be truly dead if a window exists as is the case in most control room designs.

FIG. 2 DEVELOPMENTS IN RECORDING AND MONITORING GTM SETTING FOR IVIE 30/17A



PW = PULSE WIDTH
CT = PULSE CYCLE TIME
DT = MEASUREMENT DELAY TIME (ANALYSER OFF)
AT = APERTURE TIME (ANALYSER ON)

The essential theory of LEDE, however, stands all argument and there is no doubt that time delay spectrometry is capable of providing more than sufficient proof. The most obvious subjective effect of LEDE room design is the awareness of what is happening to sound in the studio itself. By removing the coloration of the monitor system, every aspect of the recorded sound is enhanced making microphone placement and instrument screening far easier. Phase cancellations and anomalies in the recording chain become painfully obvious in the absence of 'masking effects'. It is logical therefore, that time compensated monitor speakers should find greater acceptance and enthusiasm in such control rooms and this has proved to be very much the case.

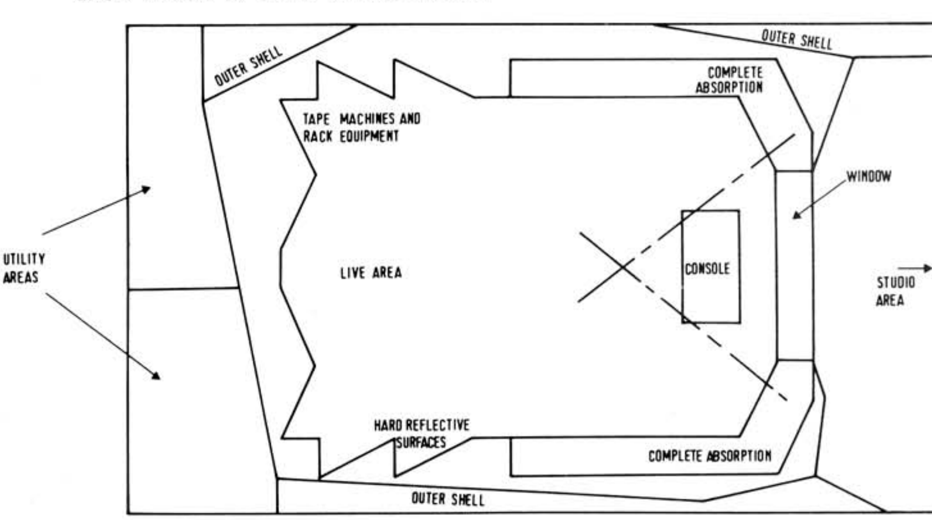
Time alignment

The concept of time alignment or linear phase compensation is not new, but it has long been argued that phase shift between discrete frequencies in a musical spectrum is not detectable to human hearing. There is now an opposing argument that transient response at high frequencies is very dependent on phase, but the most important aspect of phase (time) differences between drivers in a 2- or 3-way monitor speaker is the response anomalies that occur near the crossover points. Here, if two drivers are radiating equal energy, but are displaced by a small physical distance then phase cancellations of up to 15dB can occur. This can be proved very simply using time delay spectrometry and many demonstrations have taken place to indicate just how dramatic the effect can be on the overall system response.

The first investigations into time alignment which produced meaningful results were made by Ed Long in 1977. He experimented with a well proven and accepted monitor driver, the Altec *604*, which by virtue of its dual concentric construction has spatial differences between low and high frequency drivers. The effect of correct phase alignment of the two units produced remarkable results. Certainly any A/B comparison of musical material shows interesting effects depending on the phase integrity of the recording and playback system (including the rooms involved). One fact emerges —certain pieces of music sound truly miserable in such conditions,

52 ▶

FIG. 3 DEVELOPMENTS IN RECORDING AND MONITORING BASIC LAYOUT OF LE DE CONTROL ROOM



therefore it is inevitable that an ever widening quality gap will develop in the commercial recording world which will be immediately apparent when time aligned speakers are available to the domestic hi-fi market.

Having attained the most accurate monitor source one should consider exactly what is the optimum acoustic environment for monitoring. By definition the purpose of a monitor system is to: (a) allow the engineer to assess both technical and musical quality as material is recorded; and (b) provide the producer with an artistic image which correlates not only to the control room but also to the outside world of domestic hi-fi and car radios.

A control room, which is virtually anechoic, may satisfy requirement (a) but with two disadvantages which preclude their use. Music has no natural timbre under such conditions and therefore must have artificial reverberation added which later under normal listening conditions, will appear excessive. The absence of a reverberant sound field with the high listening levels demanded by today's producers would require speakers and power levels on the limit of available technology in a room large enough to be anechoic at low frequencies, especially if boost equalisation is used.

The desired reverberation time for rooms of given volume has been assessed emperically over almost a century and within the limits of personal preference figures of around 0.25 to 0.4s for control rooms have been judged to be ideal. (The Fitzroy RT_{60} equation should be used for small control rooms.)

What has not been considered so much in the past is the development of the sound field, which is equally important and provides the ear and brain with the information it requires to assess the acoustic environment. For example a bathroom and a medium sized concert hall may show the same RT_{60} at mid range frequencies, but it is perfectly obvious that their acoustic characteristics are fundamentally different. The critical elements in room ambience are the time gaps between early reflections and more importantly the time and amplitude of the initial delay path in relation to the direct signal.

Berenak states that in a concert hall initial delays in the region of 20ms and 10dB below the direct level produce the most desirable ambience with near perfect articulation of the original sound. It is no coincidence that the Haas effect appears at its greatest in these conditions.

Control room of Regal Sound, Hitchin, a recent project undertaken by Turnkey.

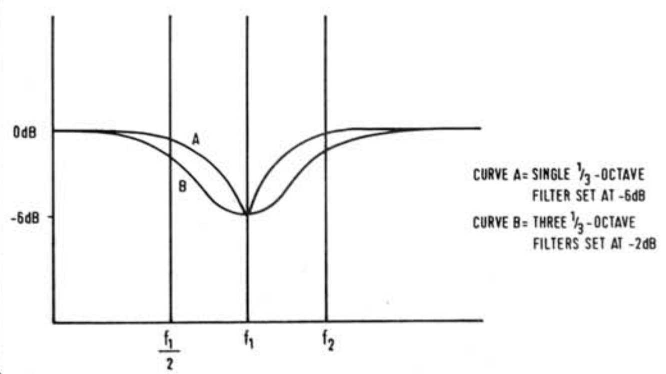
In the control room, initial time delays of 10ms or more with highly diffused 'comb effect' reflections following, create the psychoacoustic effect of a much larger room without any increase of the reverb time.

In the design of the rear of such a control room, great thought must be given to the development of the reverberant sound field. Again, time related spectrometry is invaluable in measuring the density of reflective energy as it is the evenness of the 'time v energy' relationship which is important. The low frequency performance of the control room causes problems when the wavelength becomes greater than a third of the smallest room dimension (invariably the height). In order to maintain even sound field development at lower frequencies it becomes necessary to construct an outer shell which is hard, rigid and asymmetrical. Such a shell must often be constructed for isolation purposes anyway so the geometrical consideration is the only new parameter. The low frequency sound field can only be smooth to the limit of bass audibility if the shell is large enough to develop wavelengths of 30ft and angular enough to generate complex standing wave patterns. The low frequency reverb time can be controlled by the inner control room walls which, if constructed correctly, will provide wide band absorption free of undamped, high 'Q' resonances. The front of the control room must be considered very carefully in this respect as no early low frequency reflections must take place or, more correctly, their amplitude must not be significant in respect of the directly radiated energy.

Having considered the importance of time/energy relationships, the more conventional concept of frequency/energy curves—frequency response to the layman—seems almost insignificant. However, the ability of the ear is remarkable in detecting response anomalies so the monitor system must be flat to within a 3dB envelope to be considered acceptable.

The perfect LEDE room, in theory, would need no equalisation for the reasons previously stated—a time aligned monitor speaker listened to without early reflections would be perfectly flat regardless of the ambient spectral density which the brain would dismiss as almost insignificant in terms of the original data. This is why again when listening in a good concert hall one is blissfully unaware of the frequency response of the room—the brain is only interested in the direct information.

FIG. 4 DEVELOPMENTS IN RECORDING AND MONITORING FULLY COMBINING FILTERS



It is at this point where one of the most misused and misunderstood elements in the recording chain comes into question—equalisation.

To quote Richard Heyser, "A monitor speaker which has been equalised to a room should exhibit minimal phase shift. A minimum phase speaker is one which when equalised using standard RC or RLC networks has the minimum possible phase shift over the frequency spectrum."

It has since been established that minimum phase shift in a speaker is achieved by time alignment of the drivers to produce what is effectively point source radiation. It has further been established that a correctly designed equalising filter set will automatically balance the phase response of such a speaker.

Frequency response anomalies in non-time aligned speakers cannot be corrected simply by graphic equalisation because phase and time related effects cannot be compensated purely by a change in amplitude.

It can further be shown that phase shift in a filter is proportional to slope rate and that filter bandwidths of less than ⅓-octave can produce excessive transient distortion. This is further compounded by narrow band filters 'ringing' due to their long time constant.

The ideal graphic equaliser for studio use is a fully combining 30 or 31 band graphic giving up to 10dB boost or cut per filter. Fully combining filters are such that if two adjacent bands are cut by, for example 5dB the result will be a single ½-octave band cut of −6dB. If three adjacent filters are cut 2dB there will be a single ½-octave band cut of 6dB (**See Fig 4**). This property not only allows very smooth equalisation but also such filters are of the minimum phase type. Non-combining filters are not suitable for room equalisation.

To generalise, graphic equalisers are extremely useful for final system tuning and, set correctly, should result in identical energy/frequency curves from one room to the next.

However, they are not suitable for narrow band resonant effect corrections such as standing waves or cavity resonance. In a steady state measurement a graphic equaliser may appear to flatten such effects but the dynamic response of the room will remain poor at these frequencies. The only correct solution for such problems is treatment at source using real time analysis and preferably time delay spectrometry. By identifying the time domain of

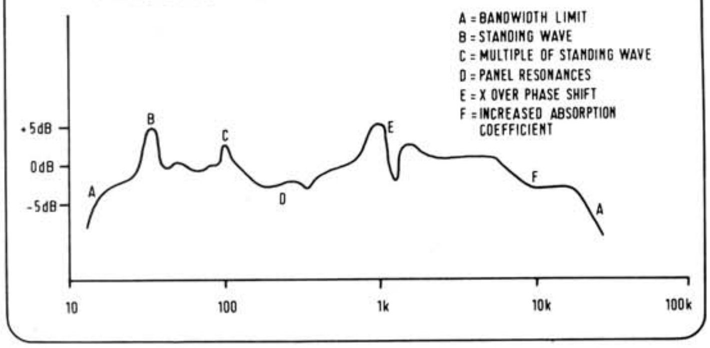
54 ▶

Developments in recording

each resonance, its cause and location can be determined and systematically tuned out by damping (see **Fig 5**). Equalisation can be extremely useful in reducing the monitor levels required by producers to accurately hear very low and high frequencies. Usually high monitoring levels are justified on the basis of the Fletcher Munsen equal loudness contours but the real problem is often that the monitor room is bass-light. The ultimate benefit of equalisation is that, used correctly, it enables tapes recorded in one room to be mixed elsewhere with little or no change in the quality of the mix.

One common problem is that monitor speakers which have been equalised to different settings on a stereo graphic, mainly because the room is not symmetrical, exhibit poor centre imaging. This is particularly noticeable on voice recordings and is a direct result of phase shift variations with different filter amplitude, emphasised probably by the asymmetrical waveform of the human voice. As would be expected these effects are less noticeable on correctly time aligned speakers and careful room design removes the problem completely.

FIG. 5 DEVELOPMENTS IN RECORDING AND MONITORING
TYPICAL CONTROL ROOM RESPONSE



Great care should be taken when applying excessive low frequency equalisation as it may result in severe amplifier clipping or at worst driver blow-out. A 3dB boost of the 50 to 100Hz octave band would warrant a 100W monitor system increasing to 200W to maintain the same operational headroom! The final comment on equalisation has to be: Two rings don't make a right!

Every element of this discussion has so far been concerned with the control room. It can be safely said that if the monitor system is accurate in terms of energy v frequency v time then what happens in the studio will be heard so acutely that many new recording techniques will evolve. In a minimum phase monitoring situation the smallest anomaly in phase becomes painfully obvious, new mic positions must be found, certain mic combinations will be stopped and a new method of recording will evolve.

It has long been understood that microphones exhibit severe response changes near reflecting surfaces. The reason has been well explained without the need for TDS but is of course due to time delay itself. Bob Shulein of Shure Bros produced an excellent paper on the correct placement of pressure gradient microphones and noted that pressure response microphones (omnidirectional of small dimensions) do not suffer nearly as badly. In fact when placed very close to a hard, large reflecting plane their response is perfectly flat. This is due to the fact that all sounds reflecting from the surface must be in phase with the direct sound. This ideal has been taken a stage further by Ed Long who has patented the Pressure Zone Microphone *PZM*. This microphone has caused much controversy and it is not the intention of this article to pass judgement, but merely to draw readers' attention to new developments in acoustic technology.

The concept of total phase integrity within the recording chain has now become a realisable objective. Digital recording and other advancements are pointless without further improvement in the understanding of sound field propagation and decay. There can be no doubt in the minds of anyone who has experienced recordings made in such LEDE studios that the next few years will see a total rework of sound recording as we know it.

It cannot be over-emphasised that the ideas discussed do work and are entirely provable, but totally unsuitable for studios unprepared to let the room dictate the final performance.

It only remains to acknowledge the work and writing of the following individuals and companies who have been the source of much of the material which has been used in this article. Don and Caroline Davis; Synergetic Audio Concepts; Richard Heyser; UREI; Ivie; FWO Bauch Ltd; Alex Garner, Tannoy Loudspeakers; and B & K Laboratories. ■

Control room acoustics

Andy Munro (Turnkey Two)

This is the first section of a two part article detailing design and acoustic techniques for small control rooms. In the second part, Andy Munro, will explore measurement techniques used to assess the performance of small control rooms.

THE last ten years has seen the growth of many small recording and broadcast studios serving an expanding and increasingly cost conscious music industry. The arrival of video and audio visual studios has also swelled the numbers considerably. Historically, acoustic measurement and design has been based almost entirely upon statistical analysis of large rooms resulting in constants applicable on an empirical basis to any geometric shape providing the volume was sufficient to produce a truly reverberant sound field.

It is generally agreed that a room which does not generate at least 100 reflections of an original sound within a 60dB decay time will not establish a steady reverberant sound field. It is the purpose of this article to examine the acoustic performance of small control rooms in the light of recent theoretical and instrumentation developments.

A control room of 5m length, 4m width and a height of 3m will be used as an example. Room volume will therefore be $60m^3$ and the total surface area of the room $94m^2$, if considering only its basic geometric shape without additional boundary surfaces. Thus, $L=5m$, $W=4m$, $H=3m$, $V=60m^3$ and $S=94m^2$.

The mean free path of a sound travelling within any room of any size is given by:

$$MFP = \frac{4V}{S},$$

$$\therefore MFP = 2.55m.$$

If the room was covered by evenly distributed absorbing material of, for example, coefficient $\bar{a}=0.41$, then traditional Sabine analysis would predict a 60dB decay time of:

$$RT_{60} = \frac{0.161V}{S\bar{a}} = 0.25s.$$

Sound travelling at 343.24m/s (Vs) will experience $\frac{Vs}{MFP}$ reflections per second and therefore the average number of reflections (NR) within the defined reverberation time is given by:

$$NR = \frac{Vs \times RT_{60}}{MFP} = 33.67;$$

ie insufficient reflections to remotely satisfy the Sabine formula requirement for statistical accuracy.

For a truly reverberant sound field (100 reflections) and an RT_{60} of 0.25s, MFP must equal:

$$\frac{343.24 \times 0.25}{100} = 0.858m.$$

An MFP of 0.858 would require in a room of volume $1,000m^3$, a surface area given by:

$$S = \frac{4V}{MFP} = 4,662m^2.$$

Now an average room with $V=1,000m^3$ has room dimensions in the order of $L=20$, $W=10$, $H=5$, giving surface area $S=700m^2$. It can be seen therefore that any approximation of an RT_{60} of 0.25s can only be achieved by devious means with a geometrically contorted boundary surface. It is also obvious that by increasing the effective surface area of our small room by geometric changes to say $300m^2$ the NR value will be:

$$NR = \frac{Vs \times RT_{60} \times S}{4V} = 107.26.$$

$RT_{60}=0.25s$, $V=60m^3$ and $S=300m^2$. satisfying Peutz' equation[1] for sound field density.

However, the absorption coefficient requirement now becomes:

$$\bar{a} = \frac{0.161V}{S \times RT_{60}} = 0.128.$$

While it is theoretically possible to achieve a room design of the above parameters, several important and undesirable effects would result unless certain modifications are carried out.

With reference to **Fig 1**, in a truly reverberant acoustic environment there is a point Dc where the sound radiated from a source at Do equals the reverberant sound field and beyond this point no further change in level will occur (B). Under anechoic conditions curve (A) will decay to zero with constant slope. However, in our small room a less predictable curve will occur (C) which falls between the definite limits of A and B.

Dc is defined from inverse square and room reverberation[2] as:

$$Dc = 0.141\sqrt{\frac{QS\bar{a}M}{N}}$$

where:
Q = the directivity of a monitor loudspeaker; $S\bar{a}$ = the room absorption; M = an 'architectural' modifier; and N = a multiple source modifier.

M is a correction factor for uneven placement of acoustic absorption $\bar{a}$ and will be discussed later.

N relates to additional loudspeaker sources not usually relevant in control rooms.

Q is a non dimensional expression

of loudspeaker directivity ranging from 1 for ominidirectional free standing bass loudspeakers to 50 for horn loaded HF drivers. Variation of Q with frequency must be taken into account in both monitor and acoustic design.

Now consider our small room with $\bar{a} = 0.128$ and $S = 300m^2$. The chosen monitor loudspeaker claims a Q of 2.5 in the 1kHz octave band.

$$\therefore\ Dc = 0.141\sqrt{\frac{2.5 \times 300 \times 0.128 \times M}{N}}$$

Let M and N = 1.

$$\therefore\ Dc = 1.38m.$$

Critical distance can be regarded as an ideal position from which to mix as the engineer will experience both direct and reverberant energy equally, but it is desirable that early reflections within the time window from Dc to +10ms should be attenuated so as not to cause comb filtering of the monitor loudspeaker response. For Dc = 1.38 this time window is given by:

$$\Delta t = \Delta t_1 + \Delta t_2 = \left[\frac{1.38 \times 1000}{343.24}\right] + 10 = 14;$$

where Δt is the total time delay of the indirect path;
Δt_1 is the direct path time;
Δt_2 is the desired delay for minimum interference (in this case chosen as 10ms).

Therefore, the earliest path to Dc by indirect means should be at least 14ms or 4.8m. This is clearly difficult to achieve in a room $5 \times 4 \times 3$m, but some compromise can be made to satisfy both initial time delay and reasonable room 'ambience'.

Some consideration must be given to the materials used to eliminate early interfering reflections as these should be reduced by a least 10dB. **Table 1** shows the sound energy level reduction of a single reflection from a given acoustic boundary.

Table 1 Acoustic energy absorption

Absorbency coefficient $\bar{a}$	Attenuation (– dB)
0.5	3.0
0.6	4.0
0.7	5.1
0.8	7.0
0.9	10.0

It is interesting to note that an acoustic absorber of 99% effectiveness (ie reflecting 1% of incident energy) will reduce level by:

$$\frac{1}{10 \log 100} = -20\text{dB}.$$

Therefore, by decibel law, sound pressure of a reflected signal will also be –20dB SPL relative to the incident signal:

$$\text{ie } 20 \text{ Log}\frac{\text{Reflected SPL}}{\text{Incident SPL}} = -20.$$

$$\therefore\ \frac{\text{SPL R}}{\text{SPL I}} = 0.1 \text{ or } 10\% \text{ SPL}.$$

Even an acoustic absorber of $\bar{a} = 0.99$ will ensure 10% sound pressure will be reflected. It is obvious, therefore, that to reduce individual reflections by 10dB or more requires specialist acoustic materials of the highest specification (No account has been taken of frequency dependence at this stage.)

By introducing enough absorbency to attenuate reflections with time delays of 10ms or less it can be seen that the overall $\bar{a}$ must be modified, but it is possible to make use of the areas directly within the coverage angle of the loudspeakers to increase apparent reverberation. The modifier Ma in the critical distance formula is defined as:

$$Ma = \frac{1 - \bar{a}}{1 - a_c};$$

where a_c is the absorbency coefficient within the direct coverage angle of a monitor loudspeaker.

Suppose the overall $\bar{a}$ has increased to 0.4.

If for example:

$$a_c = 0.01,\ Ma = \frac{1 - 0.4}{1 - 0.01} = 0.6;$$

then,

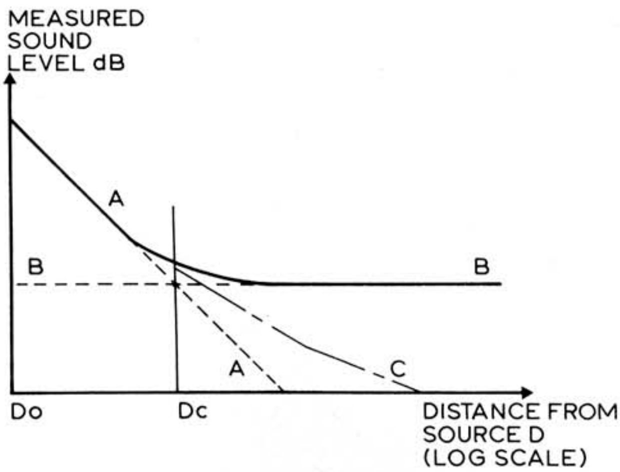
$$D_c = 0.141\sqrt{\frac{2.5 \times 300 \times 0.4 \times 0.6}{1}} = 1.89\text{m};$$

and without Ma,

$$D_c = 0.141\sqrt{\frac{2.5 \times 300 \times 0.4}{1}} = 2.44\text{m}.$$

Therefore Ma may be used to offset the apparent effects of reduced room reverberation and keep Dc to a lower value. In rooms with 'dead' rear walls Ma = infinity, and monitoring is effectively anechoic, even if some areas of the room are decorated in exotic 'reflective' stone.

FIG.1 SOUND FIELD DEVELOPMENT



Conclusion

By assuming that a reverberant sound field can be generated at least at mid and higher frequencies, it is possible to design a room capable of psychologically representing a larger volume than is actually there.

The performance of the sound field will be critical and require detailed analysis to achieve constant energy/frequency/time relationship.

The critical distance in a small reverberant room will necessarily be small, dictating a nearfield monitoring position similar to that frequently occupied by *Auratones*. This is an obvious reason why small single source loudspeakers are so popular, often used by knowledgeable engineers and producers in preference to the main 'blood and thunder' monitors. A control room doesn't have to be small to be awful. The acoustic treatment of small rooms is critical with small numbers of units capable of at least 95% absorption required instead of blanket cover. The geometric room design will dictate both S in relation to V and also LF performance. This will be covered in a following article which explores measurement techniques in relation to small room performance. ■

To be continued

References

1. V. R. Peutz modified Hopkins Stryker Equation.
2. Reference data published by Synergetic Audio Concepts is duly acknowledged as being source material for this article although the approach is, to the best knowledge of the author, original.

Control room acoustics

Part two

Andy Munro (Turnkey Two)

"The hall itself is a monument to the imprecision of acoustical science, with its battery of glass resonators suspended from the roof like so many Cona coffee machines and already gathering dust, which will presumably in due course affect their resonance." (1)

"There should be an effectively anechoic path between the monitor loudspeakers and the mixer's ears that extends for at least 2 to 5ms beyond the initial time delay gap." (2)

"The shorter the time intervals between the individual reflections and the direct sound, the less is the ear able to detect acoustic comb filter effects. Also such reinforced music is often evaluated as more pleasing simply because of this variety of phase effects." (3)

These three statements each by an acknowledged writer in his own field show the difficulty in determining a logical approach to acoustic design given that the final judgement will be totally subjective.

The acoustic parameters for the recording or performing environment itself should be clear cut – such auditoria have been built and used for the purpose for at least 3,000 years. The Greek amphitheatre at Epidauros was constructed around 400 BC, seated 14,000 people and gave birth to the legend of 'perfect acoustics'. Let us analyse the sound field propagation of such a theatre **(Fig 2)**. The open air theatre cannot sustain a reverberant sound field and will be considered anechoic. (In fact discrete reflection may occur depending on the architectural features within the area.)

A sound source such as a single voice in raised conversation would have an approximate level of 73dB SPL at 1m.

Using inverse square law decay the level at a distance of 30m would be 43.5dB or 29.5dB less! This could account for the fact that most Greek amphitheatres were situated on a hillside facing towards a nearby forest in order to avoid environmental noise levels such as passing chariots. Certainly it shows why sound reinforcement is used in the same auditoria today.

The sound level achieved by the same voice in an enclosed space is given thus:

for a room volume $V = 43{,}000\text{m}^3$,

$RT60 = 3\text{s}$

$S\bar{a} = 2{,}300$ (using classical Sabine analysis).

Taking the mid frequency Q of the talker as 5 we find:

$$\Delta\,Dx\,(1\text{m}) - \Delta\,Dx\,(30\text{m}) = 26.5 - 4 = 22.5\text{dB}$$

$\therefore$ level at 30m from stage would be $73 - 22.5 = 50.5$dB.

$\therefore$ room gain as it is often called $= 50.5 - 43.5 = 7.0$dB. (4)

Considering equivalent power requirement, if the same level were to be achieved outdoors then 50 times more electro acoustic power would be needed. However, this 'free power' has been gained at the expense of articulation in that the intellegibility of the sound heard within the reverberant sound field will be diminished. The open air auditorium will have articulation losses %AL CONS of nil whereas the large hall we have defined **will give % AL CONS = 8%, not bad in fact, as 15% is regarded as the** accepted limit for loss of articulation in speech. (Dc = 8.2m).

$$\%\text{AL CONS} \propto \frac{D2^2\,RT60^2}{VQ}$$

$$\frac{VQ}{1000} = \frac{215{,}000}{1000} = 215$$

Suppose the hall is much smaller, say $15{,}000\text{m}^3$, at 30m max distance area

% AL CONS = 20%

$$\frac{VQ}{1000} = 75$$

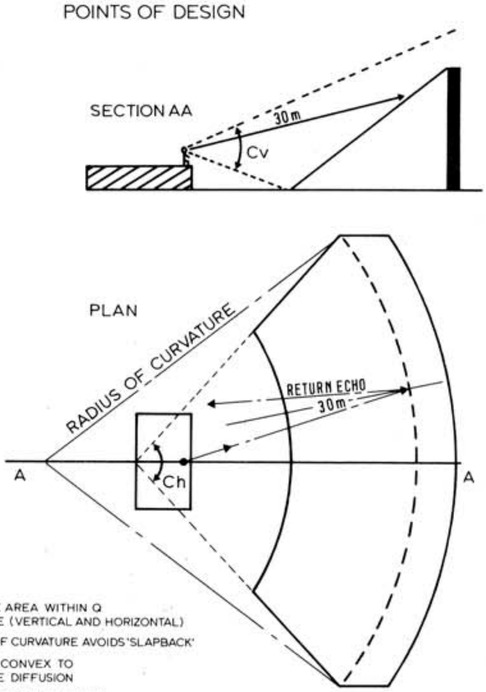
(VQ/1000 is a multiplier used in the articulation formula.)

Therefore for a given RT60 a large hall or studio will give articulate sound coverage far more readily and small halls must be less reverberant for the same value of %AL CONS.

It can be seen that a long reverberation decay in a smaller auditorium or recording studio will give rise to considerable masking effects which will totally change the subjective quality of any sound source heard or recorded within that space.

This fact was recognised in the early '30s by BBC engineers who published preferred reverb time curves for rooms of any given size. The curves differ for speech and music indicating a subjective recognition that articulation losses can be a positive feature in environments intended for music reproduction. For example a typical medium sized studio of 100m^2 would be specified by BBC designers as requiring an RT60 of 1.4s for music and 0.85s for general purpose speech and drama. They also determined that an RT60 of less than 0.3 was extremely undesirable for either speech or music unless under special circumstances. The articulation loss in a music studio of the above specification would be 10% at worst for a maximum performer-to-microphone distance of 10m. Other work carried out in the same period compared reverberation with frequency and determined that an increase of RT60 with decreasing frequency was desirable although difficult to define as the type of music recorded was an overriding consideration.

FIG.2 CLASSICAL AMPHITHEATRE – POINTS OF DESIGN



NOTES:
1) THEATRE AREA WITHIN Q OF VOICE (VERTICAL AND HORIZONTAL)
2) RADIUS OF CURVATURE AVOIDS 'SLAPBACK'
3) SEATING CONVEX TO INCREASE DIFFUSION
4) EVIDENCE OF HELMHOLTZ RESONATORS UNDER SEATS

How does the monitoring environment relate to the foregoing descriptions of acoustic performance? The most important factors for a monitoring environment can be defined as follows:

1. To give an accurate representation of the recorded or transmitted programme material in terms of:
   (a) amplitude – frequency linearity;
   (b) phase linearity;
   (c) dynamic linearity;
   (d) psycho-acoustic integrity;
   (e) acoustic compatibility with end user.
2. To be sufficiently isolated from the recording environment with minimal ambient noise.
3. To be environmentally conducive to intense concentration for long periods in terms of:
   (a) Temperature and air quality;
   (b) Lighting quality;
   (c) 'Creative comfort';
   (d) Absence of distraction.
4. To be technically and ergonomically interactive with involved personnel.

For the purposes of this article factor 1 only shall be considered, the remainder forming the basis for future discussion.

The monitor system comprises a pair of louspeaker units radiating into an enclosed space. The purpose of the system is to:

(i) Determine the optimum recording parameters (track laying) ie microphone position, use of screens etc;

(ii) Quality control the recording process ie to check what goes in comes out as it went in. (A/B monitoring);

(iii) Provide the means to produce a product compatible with and pleasing in a domestic playback situation (Mixdown).

Although seemingly compatible, purposes (i) and (ii) are distinctly separate from (iii).

Relating (i) (ii) and (iii) to factor 1 in our list of monitoring parameters we find some interesting conflicts.

During the initial recording process it is vital that the engineer hears an uncoloured impression of the recording environment. Any important reflections and early reverberation characteristics should not be masked by a stronger early sound field in the monitoring environment. However, in the final analysis the finished products will almost certainly be heard in an extremely coloured early reverberant sound field (Walkman users excused!)

Another important factor in rock studios at least, is distortion created by the monitor speakers themselves, not to mention certain amplifiers driven to and beyond their linear capabilities. The constantly increasing demand for sound levels peaking at 120dB SPL has resulted in a breed of horn loaded compression drivers delivering anything up to 20% harmonic distortion. It should be duly noted that the amplifier power required to produce such levels with linearity must be in excess of 4kW for a system of sensitivity 93dB/1W/1m in an effectively anechoic room using horn loaded speakers. The cynical may at this point question the validity of such levels but there is a case for high level monitoring in that psycho-acoustically the level of performance in the studio should be accurately reproduced in the control room in order to interpret its musical value. Rock music has brought about the means by which the recording industry has grown as a whole and in the process has produced both the best and the worst in every facet of that industry.

Phase linearity within the monitoring environment falls into two categories: linearity within the monitor chain – replay heads, console, crossover, amplifiers, speakers; and interference patterns caused by discrete reflections interfering to produce comb filter effects within specific time zones.

Those in the first category are relatively straightforward aspects of electronic design but it is remarkable how little attention seems to be paid by both manufacturers and users to publication of phase response data of the equipment they use. (5)

The second aspect of phase response only became apparent when time delay spectrometry was developed in the late '60s. This form of measurement is well documented although still financially beyond the reach of almost all but the larger acoustic consultancies and industrial users. The results of TDS, however, have changed several well entrenched attitudes to the design of listening environments and there is a slow, almost embarrassed movement towards many of the concepts originally dismissed by traditional purists as "just another trendy American marketing exercise".

Specifically, TDS enables acoustic performance to be measured in exactly the same way the ears/brain do, as a Fast Fournier Transform function operating in both time and energy domains.

The early reflected energy in a room can be shown to distort the amplitude-frequency curve severely at the specific points in time before an overall masking effect can be applied by the remaining energy in the room.

Conventional acoustical measurement such as impulse and pink noise give no indication of such early effects although they are useful in many steady state assessments of a room performance.

Much can be determined by the use of geometric and wave analysis and control rooms lend themselves to such treatment by virtue of their smaller size. Part one of this article discussed reasons why a small room is incapable of sustaining a true reverberant sound field without considerable modification of its acoustic parameters and even then will not completely satisfy statistical analysis beyond a few basic criteria.

It is important to realise that all rooms determine the development of the sound field by virtue of their shape, size and acoustic absorption. At low frequencies sound radiates as a spherical wave from a point source and any point on that wave may be considered a point source of secondary radiation (Huygens Law). This phenomenon accounts for the ability of sound to travel around corners and to fill every corner and crack of a room with alarming ease.

At higher frequencies sound propagates more and more in a predictable geometric fashion and obeys 'ray' theory to a useful degree. The crossover between geometric and wave propagation is determined by the smallest dimension of the room, which in the case of control rooms, is almost invariably the ceiling height. Below this frequency acoustic energy is distributed as a series of harmonically structured peaks and nulls relating to resonant standing wave conditions.

As each standing wave corresponds to a half wavelength room path (or multiple thereof) it is obvious that a large number of different path lengths are required to give even energy distribution. This is most easily achieved in an asymmetrical shell, with overlap of each of the fundamental dimensions. The shell should be rigid enough to act as a reflector to the lowest frequency given by the particular room. The actual crossover from geometric to wave frequencies is given by;

$$fc = \frac{3v}{d}$$

where d is smallest room dimension and v is sound velocity. (6)

This gives our original room (see part one) a crossover at 343.25Hz, but it should be realized that the changeover is gradual and not fully accomplished until two octaves lower in the case of 'wave theory' acoustics at 86Hz.

As control rooms are invariably used for the monitoring of stereo material the higher frequencies must be radiated into the room in such a way that any anomalies in the geometry of one side of the room are matched by the other, in order to preserve phase integrity. Failure to achieve this will result in poor centre image and at worst a noticeable difference in left-right monitor response. This effect is most critical for early reflections and becomes insignificant for reflections arriving within the "Haas Zone".

This requirement obviously conflicts with the low frequency geometry of the room and therefore an inner symmetrical shell must be constructed with a crossover frequency to the outer structural shell as defined. For the faint-hearted (and cramped) it is possible to create a combination of these two shells providing the original room dimensions are within certain preferred ratios. It is possible to make the outer shell symmetrical but of such structural geometry as to create the correct overlapping standing wave pattern.

The alternative approach to the low frequency shell is to absorb virtually all bass frequencies after they have passed the monitoring position. Although sometimes a necessity due to design constraints (extremely small rooms, mobiles etc), this approach has inherent disadvantages:

1. Higher monitor power needed;
2. Loss of space using large traps;
3. 'Unreal' or non existent reverberant field.

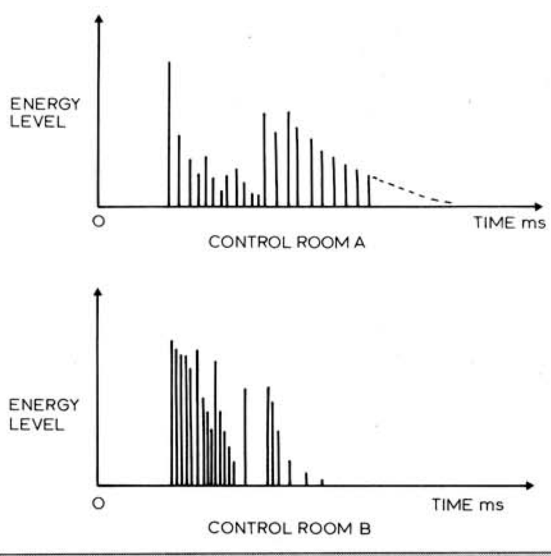
To compensate for bass traps several designers make the front and side walls hard and reflective causing many early order reflections. **Fig 3** shows a TDS comparison of the two types of room.

Each vertical line represents a complete frequency sweep of a heterodyne analyser tracking its own oscillator with respect to time. The FFT analyser displays time and level as 400 discrete bands of information with, in the case of B & K at least, the facility to 'window' by a further magnitude of 10. In other words each individual line may be viewed by integration as a complete 20 to 20kHz frequency *v* level spectrum.

The great advantage of TDS and a time related approach to design in general is that any control room may be significantly improved almost overnight by use of specialist materials requiring a minimum of down time and building work. This is particularly true of frequencies above fc and by skilful use of Helmholtz resonators even low frequency problems can be eliminated or improved.

Considering the high capital cost of the recording studio and the competitive nature of the business, improved acoustic performance should be considered a vital investment particularly as most other forms of potential signal distortion are slowly being removed from the chain. ■

FIG. 3 TYPICAL TDS SPECTRUM



REFERENCES

(1) Sunday Times review of Barbican concert.
(2) Don Davis *AES Journal* 1980.
(3) M. Rettinger *dB* magazine June 1980.
(4) Sound System Engineering: Don and Carolyn Davis

$$\Delta\, Dx = 10 \log \left[\frac{Q}{4\pi Dx^2} + \frac{4}{S\bar{a}} \right]$$

(5) Richard Heyser, *AES Journal* 1969.
(6) *Syn Aud Con* Vol 9 Number 3.