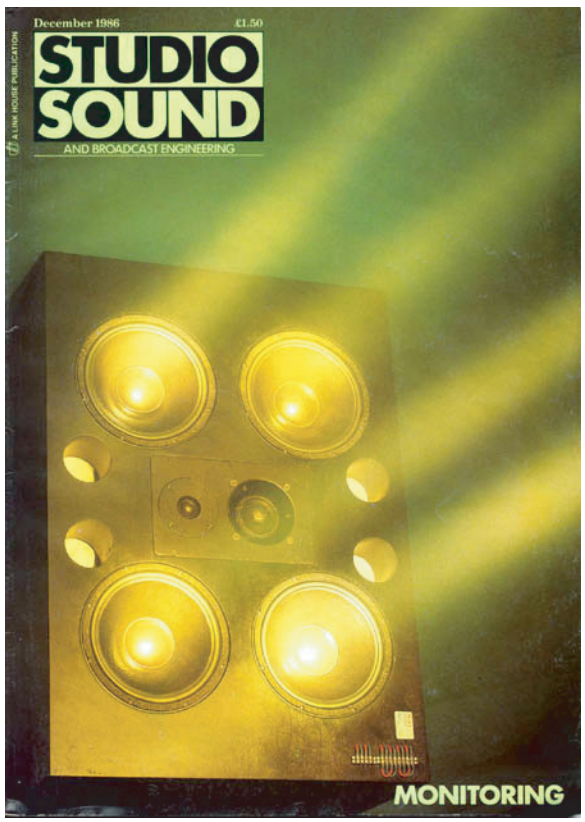
Studio Monitoring Design

By

Philip Newell



Studio Sound

December 1986

STUDIO MONITORING DESIGN

Three years ago Reflexion Arts started work on a monitor system for use in rooms of their own design and as independent systems. Designer Philip Newell describes the development philosophy which considered more than just the system components

The past few years have seen not only the advent of digital recording but also the proliferation of computer generated keyboard instruments, with the shift of personnel from the studio to the control room. The many hours of preparation and programming of these instruments, together with the necessity for close communication between the musician, programmer, engineer and producer, has inevitably led to larger control rooms, where keyboard rigs can be set up with sufficient space for everybody involved to move about freely.

Larger control rooms require greater output capability from the monitors, in order to achieve similar acoustic levels as would be produced in smaller rooms. Digital recording has preserved transients which would be lost on analogue tape, hence the dynamic range of the monitoring and its ability to handle repeated and higher transient signals must be correspondingly greater. Keyboards produce two further problems. Firstly, computer and digitally generated sounds can produce signals of a very unnatural nature, often with extraordinarily high intensity signals, concentrated in very narrow frequency bands. Secondly, with the control room now becoming the studio in which the musicians are playing, the monitoring system must be able to produce, when required, "live" volume levels. As many musicians "play off the volume" for inspiration, so these monitors must be able to recreate the levels of a concert stage. It's the frame of mind of the musician at the time when the music is recorded that dictates the overall feel of the track.

What do we need? A system with a high output capability, fast response to large transients, relative indestructibility to cope with keyboard accidents (they don't always put out the level that you were expecting), flat acoustic output to the extremes of the audio spectrum, low distortion, and a well balanced tonal character, independent of level. Let's look at these things in more detail.

Low distortion and high power handling will largely be down to the choice of individual drivers. Flat acoustic output could be achieved by equalisation, or attenuation of the more efficient units, however, these factors could work against us in other ways. If we choose a system for use with a single amplifier and high level crossover, we would be required to match the efficiencies of the drivers. A midrange driver with a 6 dB greater sensitivity than the bass driver, would produce a peak in the middle unless attenuated. Any attenuation used would waste amplifier headroom and potentially increase distortion at high levels.

If we drove each individual driver, or pair of drivers, from their own independent amplifiers, no constraints would then be placed upon our choice of drive units. The optimum units could be chosen purely for their desired characteristics. Furthermore, any components, either attenuators or crossover components, which may come between the loudspeaker and the amplifier, serve to reduce the motional feedback control which the amplifier's damping factor may exert upon any cone or diaphragm excursions. Tight control over the cone movements of the bass drivers especially, may well be severely impaired by passive, high level crossovers.

By separating the amplifiers, absolute maximum use is made of their headroom and transient handling ability. A further advantage is somewhat less apparent. Observe a sudden, low frequency peak, when fed into a system driven by one amplifier. Should the peak exceed the amplifier's output ability, harmonic distortion will be produced. The harmonics, of a higher frequency than the fundamental, will pass through the crossover into the high frequency drivers. This will not only produce unpleasant audible distortion but will introduce high level, spikey overload signals which the HF drivers may have difficulty in handling. Such repeated overloads can cause listening fatigue and also may cause premature failure of the driver.

Should a similar overload occur in a multi-amplifier system, such low frequency overloads cannot enter the HF drivers as they are not coupled to the same amplifier output. This results in three beneficial effects. One, the highs continue to be heard as clearly as ever, untainted by the LF distortion. Two, no unnecessary strain is put on the HF drivers, which helps towards long and consistent life, reducing diaphragm fatigue. Three, a bonus, the LF drivers, having a response severely limited at the higher frequencies, cannot reproduce the majority of their own distortion. In effect, the sudden LF overload passes through the system almost imperceptibly and without the risk of straining or damaging the drivers.

On the subject of flat acoustic output and well balanced, non volume-dependent tonal character, the two are to some degree linked. The aim was to choose drive units which performed effortlessly in their specified frequency ranges, with the intention of avoiding any requirement for compensation by monitor equalisation.

The whole subject of monitor equalisation can be a minefield. Unless applied in very smooth and gentle sweeps, equalisation rarely reflects the true situation. For example, suppose a room has a bump at say 58 Hz which shows at 63 Hz on the analyser (if ⅓-octave). If the nearest available frequency on the equaliser is 63 Hz, this would be pulled down till the real 58 Hz was flat on the analyser (at 63 Hz). The result of this would be a dip at the next highest frequency which would need to be boosted. This in turn would create a peak at the next frequency, which would need to be cut. . .and so on.

From a single 58 Hz hump, we end up having a flat picture on the analyser alright. However, this is achieved by the drawing of a roller coaster on the graphics. All we started off with in reality was a minor hump at 58 Hz. In no way is the picture painted by the graphics, the inverse of our original situation, and as such, it has no justification in being there. After all, when the music is playing, we are intending to listen to the speakers, not look at the analyser.

Despite the reading on the analyser, acoustically we would quite categorically not have a flat response in the room. Put the equalisers back where they belong. . . in the mixing console's effects rack! Has anyone yet seen a monitor graphic which was truly reflecting the inverse of the room/loudspeaker combination? Switch out the equalisation and see just how much more natural and clear, things sound (except possibly in extreme problem cases). You can actually hear if you're over equalising something on the console, it's not masked by unnatural, equalised monitors—try it!

Equalisers also tend to introduce phase shifts, especially when we're getting the alternating up/down pattern. This makes a mockery of achieving minimum phase shift in crossovers, or time aligning, of drivers. There's even more! Headroom at any boosted frequency is correspondingly reduced by that amount of boost for any note striking that frequency. A 3 dB boost will call upon the amplifier to double its output at that frequency, as compared with an unequalised system. A 6 dB boost would require quadruple power, hence higher distortion from the over-worked loudspeakers, and if you've got it up loud, ears too! The peaks can really be unpleasant.

When a large studio in London was completed in 1978, the rooms were fitted with equalisers and set up from scratch.

These were re-checked every few days but after a few months, strange things were noticed about the sounds. The technicians checked and double checked but the prescribed curve was still visible on the analyser. It ultimately transpired that with gradual adjustment, compensation made every few days had resulted in a totally different set of equaliser settings to those noted upon first installation. We went back to square one and started again. This time all was well. The upshot of all this is that two or more entirely different settings of the equaliser can achieve a flat response on the analysers. Clearly they can't all be right. In all probability none of them are. They never accurately correspond with what's really happening. To cap it all, even the different makes of analyser and microphones rarely correspond; or even the mics used. . . grazing, free-field, omni and cardioid. It all depends what you're measuring but they do seem to get transposed rather a lot.

In terms of tonal characteristics, analysers really don't help. Even when set flat in the same room, Altecs, Tannoys, JBLs, EVs, etc. still have their own individually recognisable sounds. You can't (as yet) make a cheap violin sound like a Stradivarius just by equalising the resonances and reverberations. Similarly, purely by electronic means, different loudspeakers cannot be made to sound the same. Analyser and graphics tell about 0.1% of the story. . .there's an awful lot more to it.

However, with a system driven from a suitable electronic crossover and with up to four individual amplifiers, smooth adjustments can be made to relatively small frequency bands by adjustment of gain controls only. To the ear, this seems to sound much more natural and lifelike, and much less fatiguing than correction by means of equalisers.

Limiting can also be dangerous on monitors. What's limiting, your monitors or your mix? It may well be imprudent to mix at ridiculous levels but occasionally, in practice, that may be what the circumstances demand. Even if it's only on peaks, monitor limiting will suppress transients and you may find yourself putting too much top on tape to help compensate for the lost peaks. Given the flexibility of the split amplifier system, careful choice of drivers and amplifiers should alleviate the need for limiters, while still not putting these ample drivers at risk.

Obviously, with excessive amplifier output capability, some damage can be done. However, with the appropriate choice, this risk can be reduced whilst still allowing for transient headroom. To achieve the full power bandwidth down to 20 Hz, and for good bass transient ability, DC amplifiers would be the first choice. Output power would depend upon the power handling and efficiency of the driver, so this choice should be left till last. The response of the crossover should also exhibit 20 dB of headroom over loud working levels, and respond down to at least 10 Hz as its 3 dB down point.

Now we have the basis of an integrated monitor system. Electronic crossover, DC amplifiers for minimum phase shift and great instantaneous LF transient surge capability. So, to investigate the specific choice of drivers, but how many and which ones?

Choice of drivers

At low frequencies, the tightest bottom end tends to be that produced by a bass reflex cabinet of suitable design, loaded with the appropriate driver(s). The choice of size of bass drivers, within normal limitations, provides the options of 12 in, 15 in or 18 in units, either used alone, in multiples of one size, or even mixed. The apparently obvious choice would be to use 18 in units, with their ability to produce great, low frequency outputs. One drawback frequently found in 18 in units, however, is the difficulty in preventing such a large cone from flexing under high level transient inputs. This serves to reduce the "punch" from the loudspeaker and, together with the resulting harmonics from the flexing cone, tend to produce a boom rather than a thud from the bass end.

By contrast, 12 in drivers, require far greater cone excursions to move the same volume of air, and there are drawbacks to having a small surface area for the generation of the low frequency sound. As the suspension systems for 12 in and 15 in cones tend to be similar, it can be appreciated that the 12 in driver receives far greater long term punishment, than the 15 in driver. Furthermore, in order to achieve a very low resonance, the 12 in cone must be either more heavily built than its 15 in counterpart or its suspension system must be even more compliant. The first option reduces efficiency and requires greater output capability from the amplifier. It also produces a lower total acoustic output per watt of power handling capacity. The second option reduces the driver's ability to cope with transient overloads without damage or strain.

Musical instruments with large, low frequency content tend to *be* large. Natural sounding low frequencies usually come from sources which produce a relatively low air pressure from a large source area. Even though it may produce a similar reading on the spectrum analyser, moving air in this way undoubtedly sounds different from the high pressure small source area, approach of 12 in or smaller cones. This once again refers back to the frequently misleading results gleaned from ⅓-octave analysers.

In practice, 15 in drivers tend to achieve the best compromise for low frequency use. The use of a 2×15 in system gives a further increase in output due to the mutual coupling achieved when the units are positioned relatively close together. This gives us a large sound source area, a relatively rigid cone, reasonable efficiency, high power handling, a good transient response, and a long, stable, life expectation.

Midrange drivers

The midrange is the most critical and most contentious area of monitor design. The choice of drivers mainly falls between cones, domes and compression horns. Cones and domes are more generally considered to produce a softer, less harsh sound than compression horns which, strictly speaking, are two separate components—the driver and the flare. When very high output levels are required, the more efficient compression horns are capable of much greater acoustic output. Thus, at high levels, they may actually be sweeter than a hard pushed cone or dome driver.

To achieve good high frequencies and fast transients, the moving mass of a midrange driver must be rigid yet light. As increase in size means increase in weight for any given material, cone and dome midrange units cannot just be made larger in order to be louder. It is also difficult for a small mass to dissipate the heat which is generated in the unit at high output levels. For a typical efficiency of say 10% for a cone unit, for every 20 W supplied by the amplifier, 18 W are turned into heat within the driver, only around 2 W being transmitted to the room as sound energy.

Doubling up the number of units obviously doubles the power handling but each doubling of output power only provides an additional 3 dB of headroom. Another problem is that unlike at low frequencies, at higher frequencies a point source is desirable as the sound wavelengths involved become shorter as ▷

235 monitor

frequencies rise. Unless the listener is equidistant from the drive units, some phase cancellation will occur, blurring the sound and impairing stereo imaging.

For high output levels therefore, the compression horn would seem to be most suitable, subject to the reduction of the characteristic harshness. Several options are available to help reduce any undesirable effects. Firstly, use cone drivers as far up as possible, whilst still on the good side of the efficiency/frequency response curve. Secondly, do not attempt to drive the horn beyond the point where the driver response begins to tail off, or the directional characteristics of the flare begin to beam and lose their even distribution over the required listening area. Thirdly, in order to maintain the output within the lowest distortion range of the compression horn, provide adequate power handling and efficiency for the desired acoustic output. Fourthly, the design of the flare itself can contribute significantly to the timbre of the sound, although frequently, this cannot be easily determined by instrumentation.

If the output of the cone drivers can be maintained up to at least 1200 Hz, this allows the use of a 1 in instead of a 2 in compression driver. In general, 1 in drivers have better transient response, improved high frequency output and generally smoother response with less tendency to the barking character of compression horns used at lower frequencies. The smaller horn throat also improves high frequency dispersion, with far less tendency to beam than larger diaphragm/horn-throat combinations.

The flare itself can be contoured to cover the desired listening area, giving a sufficiently wide dispersion to allow all concerned to be in the same sound field. Furthermore, the flare can control the dispersion to prevent unwanted reflections from points in the room where sound need not be beamed. The general requirement to achieve this would be in the order of 100° to 120° horizontal dispersion, with say, 40° vertical dispersion, presuming that most ears in the appointed listening area will be between 3 and 7 ft from the ground. The highest frequencies which can be smoothly and comfortably reached by such a combination of 1 in driver and flare, still maintaining adequate dispersion, would be around 8 kHz.

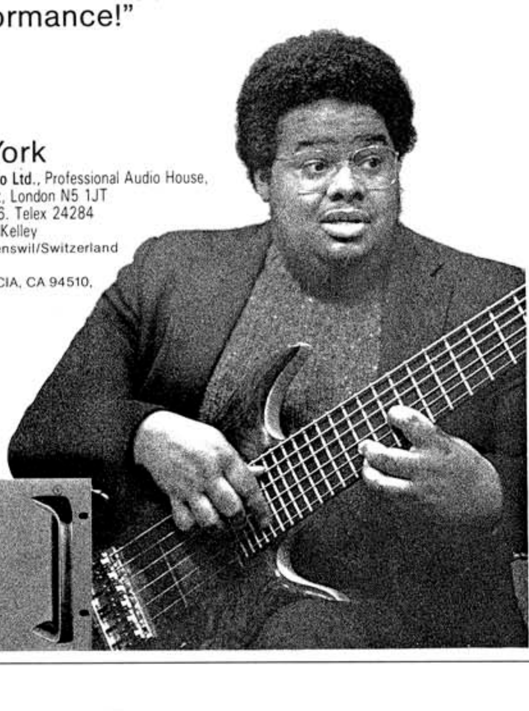
The optimum format is now beginning to unravel itself. Cones to 1200 Hz, low distortion, flat, high output, 1 in compression horn to 7 kHz and a separate unit for the top octave or so. . . 7 kHz to 20 kHz.

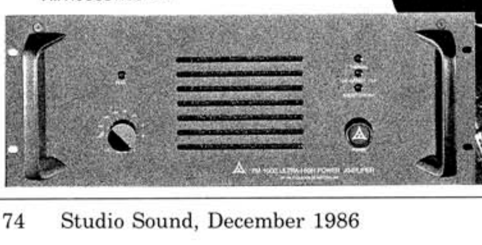
Specific choice of driver units

At the very bottom end of the frequency spectrum, two 15 in Gauss *5831F/4583F* drivers were chosen. A 4×12 in Gauss *2831* option was considered, to give a similar source area, but with 8 Ω coils, a parallel arrangement would produce 2 Ω which is unacceptably low for most conventional power amplifiers to produce their best. Individual amplifiers could be provided for each pair but this was considered cumbersome. A series parallel arrangement was considered unsuitable, for then not one of the drivers would be directly connected across the amplifier terminals as it would be in series with at least 4 Ω from the other drivers in the group. This would reduce the ability of the amplifiers' high damping factor to control cone excursions, and once again, the tightness of the bass response would be unacceptably compromised.

So, a pair of *4583F*s (or *5831F*s) would give 4 Ω in parallel which is ideal for most power amplifiers to produce full potential output power. The voice coils are rated at 400 W each, though, with the roll surround of the 19 Hz units, the RMS power handling of each driver is rated at 300 W. 600 W RMS was duly considered adequate power handling for the bottom end of each speaker system. . . 1200 W RMS down to 20 Hz for a stereo pair. Unfortunately, by 800 Hz the Gauss 15s are giving out, so 300 Hz was chosen for the upper crossover point of the bass drivers.

Whilst still delivering full output at 800 or 1000 Hz, many 15 in drivers have a distinct lack of life in their reproduction above 500 Hz. This is not easily measured, and may be connected with the mass of the moving parts, together with the cabinet linings optimised for bass response. Quite probably, many maligned drivers have been unfairly judged when being used beyond their ideal ranges. Some of this can be down to the driver manufacturers themselves, publishing measured responses within specified limits, then quoting usable frequency range even ▷





beyond these limits. This usable frequency range can probably be interpreted as that range in which there is still some audible output. In reality, this range ought to be less than the published frequency response, as frequently, towards the upper end of the range, the response becomes ragged and the tonal characteristics are no longer desirable. Once again, the provision of monitor equalisation has allowed drivers to be pushed beyond desirable operating envelopes, and well outside the range in which they can produce a natural timbre. Many monitor systems employing 15 in drivers and 2 in midrange drivers, crossing over at around 800 Hz, suffer the most in this area. The bass driver performance is compromised by the choice having to be made for a unit performing reasonably in the lower mid area. This is then asked to meet a midrange horn, operating below a frequency range which would be considered optimum for studio purposes.

It was for the above reasons that 300 Hz was chosen for the first crossover point. This leaves a gap of two octaves before the midrange horn takes over. Consistent with the overall design philosophy, a cone unit was required to bridge the gap. The JBL *2121* was considered to be probably the finest unit available for this purpose, with a very smooth frequency response, high efficiency, low distortion and 75 W RMS power handling. This 10 in unit was also considered to give an improved attack compared with similar 12 in drivers, the natural yet startling response to a snare drum appearing to confirm this. The relatively lightweight cone, roll surround, and prodigious magnetic flux, all contribute to the excellent transient response and high acoustic output.

So, at 1200 Hz we arrive at the compression horn. The primary function of the flare is to deliver the output from the compression driver to the listening area, with controlled polar pattern and frequency response. Materials which have been used range from metal to glass fibre, wood, plywood, carbon fibre and many other materials. Although all can be shown to produce satisfactory responses, without question they impart substantially different tonal characteristics to the sound. These are subtle differences, very difficult to determine with instrumentation, yet having significant effect on the overall sound of the system.

The flares chosen were a modified ASS design, manufactured from a glass fibre/resin mix with a high loading of powdered slate. The cavities in the moulding are filled with a resin/silica sand mix, of high density, the whole of the rear surfaces then being coated with a rubbery application to further damp any potential resonances. This flare was chosen for its exceptionally flat frequency response over the range of its intended use. Although constant directivity designs were considered, *and* offer a wider range of polar pattern control, by their nature, they do not have a flat frequency response. Their subsequent reliance upon equalisation precluded their use with this design philosophy.

The compression driver itself can become the source of endless, quite unfruitful argument on the subject of what is considered to be correct. This is probably the area which produces the most intense likes and dislikes. The two options initially offered for this design were the Emilar *EK175* and a combination of Coral driver with JBL titanium diaphragm. Both drivers can produce a substantially flat response from 1200 Hz to 8000 Hz. Both also have similar 50 W RMS power handling capacity. The choice is entirely down to personal preference. The Emilar produces a slightly softer, more cone-like sound, whilst the Coral/JBL combination produces somewhat more exhilarating transients. A third possible option is the TAD *TD2001*. This driver has a beryllium diaphragm, with characteristics somewhere between the aluminium diaphragm of the Emilar, and the titanium diaphragm of the JBL. It is down to individual producers and engineers to decide precisely which one most helps them to achieve their best end results. Remember, studio monitor loudspeakers are a means to an end, a tool to achieve the optimum, overall, final mix. They are, themselves, sometimes tailored to take into account the human aspects of life in a control room. Although other drivers could, no doubt, be used, the three mentioned above appear to give the smoothest transfer and closest match to the units chosen for the frequencies immediately below and above the mid horn.

By 8 kHz, the natural response of these mid horns is beginning to tail off. Once again, according to the overall design philosophy, no attempt will be made to equalise this falling response, but it will ▷

STUDIO MONITORING DESIGN

within the enclosure and prevent them from striking the acoustically, partially transparent loudspeaker cone, and passing through into the listening area. It is in the lower mid region where the timbre of the sound is most readily affected by internal damping, and where more careful consideration must be given to any such damping materials.

One advantage of crossing over at 300 Hz is that, without the compromises which are usually required, the two ends of the bass spectrum can be given their optimum enclosure treatments. Thus, the 10 in lower mid driver was placed in a separate chamber of around ½ ft^3. This chamber was mounted on the front baffle and consisted of a roughly 10 in cube. The primary function of this sub-enclosure was to prevent the 10 in cone from going into orbit when the two bass drivers punched inwards. The separate acoustical treatment facility was a further advantage. Although a cube may at first sight, not be the ideal shape from a standing wave point of view, by the time that the cone and magnet assembly had been introduced, the box became far from cubic. This break-up was further enhanced by the addition of a diagonal half-width sub-divider. This small enclosure was then lined with ½ in foam rubber, suppressing undesirable resonances, whilst not absorbing all of the life from the output. The outside of this 10 in box was treated with the same underseal/sand mixture as the main enclosure, in order to reduce any resonant tendency in this wooden sub-assembly. Lead in wires were sealed with silicone rubber.

The compression horn and slot were then mounted in the same vertical plane as the 10 in driver. These were placed as close together as practically and aesthetically possible, in order to maintain the closest approximation to a coincident sound source. Although aesthetics may initially seem a somewhat peripheral subject sight is the sharpest of our senses. It has the ability to distort our other perceptions by overriding them. A great number of the hours worked in a control room, are spent staring towards a pair of loudspeakers. A symmetry of the loudspeaker's physical layout, conditions the brain to expect symmetrical sound sources. It's a point of psychology rather than pure acoustics but these aspects cannot be ignored just because they have no bearing on the measured results.

Finally, the boxes were wired up with cable of sufficient gauge to easily pass the high transient currents associated with such a system. These systems tend to be most effective when mounted so that the front baffles are flush with the front wall of the room. This allows no areas around the cabinet sides for the bass to take any but the direct path to the ears. Enclosures designed for flush mounting have front baffles ½ in proud of the edges of the cabinet sides. Those designed for free standing installations have their front baffles recessed 1 in. This 1 in lip is of purely visual origins, and whilst the purist may maintain that cabinet edges should chamfer backward from the baffle, extensive listening tests on cabinet designs of this size can determine no audible difference whatsoever. On purely aesthetic grounds, the lip remains.

Crossovers

The lynchpin of the entire system is, of course, the crossover. 12, 18 and 24 dB/octave crossovers were tried and used extensively. The 12 dB/octave units were eventually rejected, as the encroachment of one driver upon the territory of another was distinctly noticeable. The overall design philosophy of this system called for each drive unit to handle its own range effortlessly. The crossover frequencies were chosen to cover the optimum ranges of the selected drivers. One octave beyond the crossover points, however, some of the drivers exhibit some irregularity as the response tails off. With a 12 dB/octave roll-off, a 4 dB peak in the drivers' response, one octave above the 3 dB down crossover point, is only 11 dB below the system's smooth response at that frequency. This could definitely be detected as colouration in the sound, particularly on certain instruments. This problem was all but removed with the use of 24 dB/octave crossovers. Unfortunately though, despite the drive units having been chosen to have characteristics complementing and matching each other as closely as possible, the 24 dB/octave slope was somewhat abrupt. This was noticeable especially at the change from cone to diaphragm at 1200 Hz. It was by no means distressing but a slight change in timbre was noticeable particularly on rising sections of strings. The crossovers with 18 dB/octave were finally adopted, as upon listening, they were generally considered to offer the smoothest and most pleasing overall performance. They also had the advantage over the 12 dB/octave units in their capacity to reduce the fatigue on the compression horns and slots. Indeed, the extra 6 dB increase of slope reducing by 75% the power delivered into the drivers one octave below the crossover point.

Further developments

Accepting that not all control rooms could accommodate monitor loudspeakers of the size of the *235*s, two variants were produced. The model *233* is a similar, 4-way system but with one 15 in bass driver; and a more compact model, the *238*, dispenses with the 10 in speaker and operates as a 3 way system.

As a further option, a cone midrange unit can be fitted to the systems. The cone option box is fitted to a front plate, identical in size and fixing holes, to the midrange horn. The cone driver is less efficient than the compression horn but by means of preset amplifier gain control levels, the system can be set up for rapid interchange. For orchestral music, or situations where the full output potential of the system is not required, this option may be considered desirable should cone drivers be preferred.□