Proceedings of the Institute of Acoustics

FROM MONO AND STEREO, THROUGH QUADROPHONY TO SURROUND - A Review of Control Room Requirements and Practices

Philip R Newell

Consultant

1. MULTIPLE-SOURCE CONSIDERATIONS

During the 1970s world-wide attempts were made to try to launch the concept of quadrophony as the successor to stereophony. Although stereo was conceived in the 1930s, at least in the form that we now so widely use it, it did not really begin to develop commercially until the late 1950s, and it was not until the late 1960s that virtually all music recordings were available in stereo. Most albums were released in stereo from the mid '60s onwards, but 45rpm singles continued to be almost universally released in mono, even into the early 1970s. The reason for the continued release of mono singles was principally due to the fact that these were the main promoters of the albums, and the bulk of the promotion was by airplay on AM, mono radio. Stereo, FM, popular music radio, in Europe at least, was not widely broadcast until the 1970s. Albums frequently continued to be released in two formats, but when they were released in both mono and stereo, they were often the products of separate mixes. The mono release was not simply the electrical summation of the stereo mix, as there lurk compatibility problems which we shall look at below.

Nowadays, stereo music reproduction systems are almost universal. Mono has perhaps now been relegated to reproduction by alarm-clock radios, and little else. However, it took quite some time for all the old, mono, "record players" and AM music radio stations to fall into disuse, so during the fifteen years, or so, that it took for stereo to effectively oust mono as a serious listening format, duality and compatibility were absolute necessities. The reason for the specific, mono mixes was because the electrical summation of the left and right stereo signals did not correspond to the acoustic summation of the stereo as perceived at the ears of the listeners. This may, at first, seem strange, as the electrical voltage is an analogue of the acoustic sound pressure level (SPL). In fact, on axis, on the centre line between the loudspeakers in an anechoic chamber, the voltage summation and SPL summation, from left and right, do equate; presuming of course, that one is listening to perfectly linear loudspeakers. However, in more usual listening conditions, no such direct equivalence exists.

1.1 Panning Considerations

Even in anechoic chambers, though, it is not only the lack of perfect loudspeakers which thwarts our mono/ stereo compatibility, there is more to it. Firstly, the fact that we have two ears, which cannot both be facing forwards and be absolutely on the centre-line at the same time, creates

some imperfection in the summation. Secondly, the acoustic phenomenon of mutual coupling, between the loudspeakers themselves, renders a frequency dependent dimension to the width of the stereo centre line. The main significance of this is that images which are panned from left to right will exhibit differing frequency responses at the listening position, dependent upon their position in the overall panorama. A centrally panned image, reproduced via two loudspeakers, will not sound the same as from a single loudspeaker, whether that loudspeaker be the left or right of the pair, or another single loudspeaker placed at the location of the phantom image. Hence the "rule", "Pan first, equalise later".

A third reason also exists which complicates the mono/stereo compatibility, and that is the effect of room boundaries. The reflexions created by two loudspeakers, both contributing to a centrally panned, phantom image, will not produce the same reflexion patterns as a single, centrally placed, monophonic source. These aspects affect the overall, perceived response, not only in terms of the reflexion patterns themselves, but also by the further effect of them loading the loudspeakers, and hence influencing their output.

It should thus be easy to appreciate that a mix which is done whilst listening to a pair of loudspeakers, whether the mix be in mono or in stereo, especially when this is done in a room with any significant acoustic life, will have the potential to differ quite considerably from a mono mix, monitored on a single loudspeaker. In the days when mono was still of significant importance, it was almost universal practice to make a mono mix on a single loudspeaker, and not on the stereo monitors with all pan-pots centralised. The use of two loudspeakers for mono (phantom) mixing highlighted the old pan-pot dilemma. A pan-pot would need to be 6dB down, electrically, in its centre position if central images were to maintain their relative balance to the other instruments when mixed into mono. On the other hand, when listening in stereo, in order to maintain the relative balance of a sound being panned from side to side when passing through centre, the pan-pot would need to be only 3dB down in its centre position. At least this is the situation which applies in non-anechoic rooms. Earlier control rooms were generally more acoustically live than most of the control rooms of today, so a 4½dB compromise was often chosen. This matter was dealt with at length in a recent AES conference paper[1]. Around 1970, this subject of pan-pot laws was something of a "hot potato", but the domination of stereo was about to become almost absolute. The widespread emergence of stereo "Ghetto blasters", FM stereo radio stations, and inexpensive, stereophonic "music centres", soon began to make the relevance of mono compatibility diminish in importance because they sounded the death knell of the mono "single".

2. EXPANSION INTO QUADROPHONY

In fact, it was about 1970, at a time when mono compatibility was still a relatively important subject, that the existing compatibility problems were compounded by the advent of quadrophonics, whose aim was to envelope the listeners in sound. When compared to mono,

stereophony had made an enormous impression in terms of realism and the involvement of the listener in the music. This was one reason why the cheaper stereo music systems sold so widely. For many people, the extra dimension of cheap stereo could only be offset by the hi-fi quality of some very expensive mono systems. It was subsequently widely expected (or, at least, hoped for by the equipment marketing people) that quadrophonics would offer a dimensional boost of the same order, only this time, over stereo.

Of course, the problem which now loomed was that of quad/stereo/mono compatibility. Although doing a mono mix for a single release was still common enough practice, as, in fact, was the separate mono mix for an album, these were the days when many recordings were still made 8-track; or 16 at the maximum. Mixing was usually seen as a relatively quick process, but it was soon to become a much more long and complicated process when 24-track arrived, and this arrival was concurrent with the rise of quadrophonics. The concept of having to do three separate mixes, for mono, stereo, and quad, from a 24-track master, was neither a satisfying prospect for the musicians and producers, nor a financial practicality for the record companies. If the multiple mixes cost too much money, then it would only be the potentially big selling albums that would have the budgets to mix in quad, and this would limit the demand for domestic reproduction systems. Mass sales of systems would need a large inventory of available recordings.

2.1 Lessons from Quadrophony

The launch of quadrophony was ill-conceived and badly handled. The marketing hype was so great, and the general knowledge of psycho-acoustics was in such short supply, that record producers were led into trying to get more out of quadrophony than the psycho-acoustics of the process would possibly support. They were also faced with a battle between the various systems being used to put a four channel recording on to a vinyl disc. Eventually, a more practicable system arrived, the Ambisonics system, made workable by the late Michael Gerzon, and closely related to the BBC's Matrix H system. However, by this time, so much money had been invested in the older systems of encoding and decoding that the multi-national companies were in a mood only to fight to the death for the superiority of their own systems. Subsequently, they all died out. Dolby continued to work on the Dolby Stereo Surround system, for cinema use, and subsequently on their improved Pro Logic system, but all of this was picture related. Quadrophony as a music-only format was effectively finished by 1980.

2.2 The Coding Systems and Their Limitations

To re-cap, there were two basic ways of encoding four channels onto the two sides of the groove of a vinyl disc: by phase encoding, or by the modulation of a 38kHz carrier frequency. These options were not available for use on cassette tapes, though, as cassettes did not have the ability to record as high as 38kHz, and nor did they, then, have the azimuth accuracy to allow stable phase encoding. In fact, both systems were fraught with problems on vinyl also. The modulated carrier frequency discs could not be played by normal stereo pick-ups, as special picks-ups were

required in order not to damage the very delicate groove modulations at the carrier frequency. The phase encoded discs were able to be played, not only by normal stereo pick-ups, but as stereo compatible recordings, but unfortunately, the coding systems with the better stereo compatibility were usually the poorer in terms of their quadrophonic response, and vice-versa. These systems also had their own specific imaging areas which were out of bounds for quadrophonic localisation if stereo or mono compatibility was required.

For example, an image placed centre-rear in the SQ system, would be out of phase in stereo or mono, and hence would be all but inaudible. What is more, the different systems had different fold-down characteristics in their quad/stereo compatibility. A sound placed right-rear in SQ would appear half-right in a stereo panorama, whilst in QS, it would appear to be wide of the right front loudspeaker. Multi-format quadrophonic releases, made from the same, discrete four-track master, would therefore produce entirely different stereophonic panoramas. What was worse, in some systems, if an image was panned to centre front in the quadrophonic array, then although it would still remain centre-front and at an appropriate level when played in stereo, the signal would have a 90° phase shift between the left and right loudspeakers. This lack of an in-phase centre image often reduced the punch and "solidarity" of the front centre location, which so often carries the most important information. Mono compatibility suffered similarly.

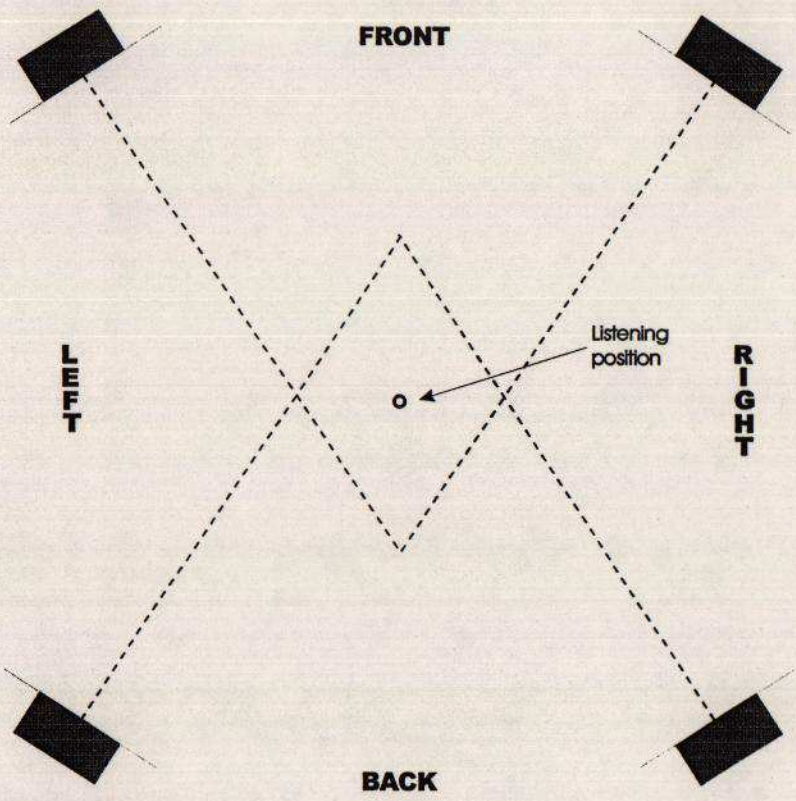
3. ADVANCES TOWARDS SURROUND-SOUND

Until here, we have been discussing quadrophonic systems with the arrangement shown in Figure 1, whereby the listener sits in the centre of four loudspeakers, facing the centre of the front pair. The "Dolby Stereo" system, concurrently with the development of Ambisonics, re-oriented the four channels. Figure 2 shows this system layout, whereby a discrete centre front loudspeaker is flanked by a left/right frontal pair. The fourth channel is divided between two loudspeakers, placed at the left and right rear of the listener. Although these rear loudspeakers were driven by a mono signal, their physical separation helped to induce a sense of greater spaciousness. Unfortunately, this vastly more sensible system arrived too late to re-kindle any interest in quadrophony from the record companies.

3.1 Psycho-Acoustic Considerations for Surround

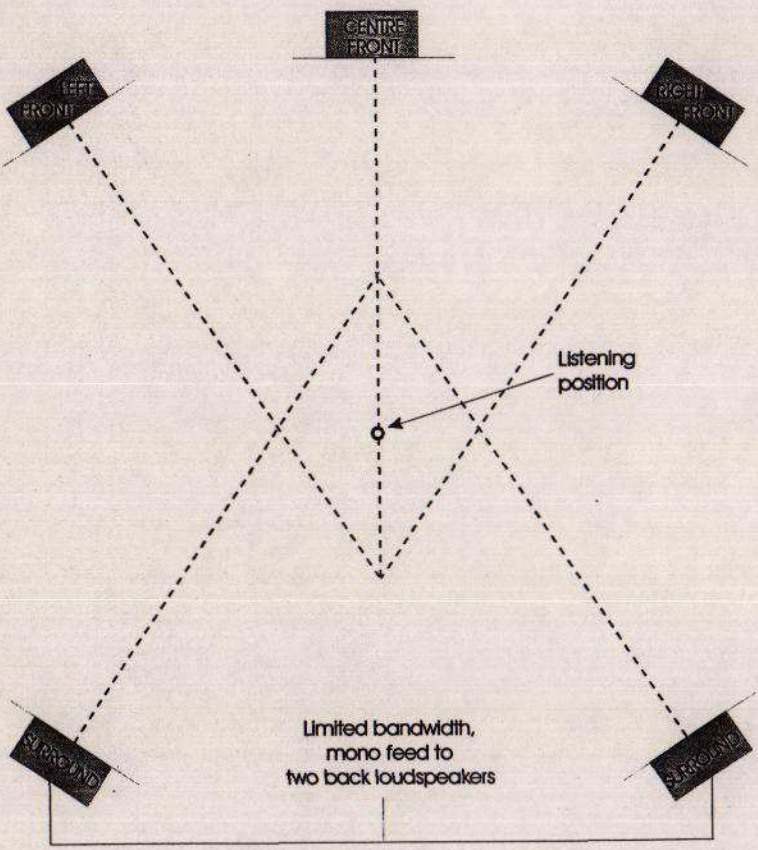
As is now well understood, the four channel quad of Figure 1 is hopelessly flawed. Stereo, in itself, is a very unstable illusion, whose psycho-acoustic principles only work well when the nose of the listener is pointing directly between the loudspeakers. With the loudspeakers facing the back of the head, not only is the perception of the high frequencies reduced, but the ability to judge the location of instruments in the panorama is also greatly reduced. Down the sides, the essentially one-eared localisation is very poor, and extremely unstable. Attempts were made to improve localisation with the so-called "Logic" decoders, which employed gain riding techniques. In some of the phase encoded systems, the basic channel separation was as little as

Figure 1 EARLY QUADROPHONIC MONITORING LAYOUT



In the quadrophonic listening arrangement of the 1970s, there are four "corner" loudspeakers supporting four phantom sound-stages: front, back, left and right. With the old matrix encoding systems, even the front sound-stage was not as stable as with conventional stereo. Although a centre front image had equal level in each of the front loudpeakers, in some systems, the signals were not phase coherent.

Figure 2 THE 4-CHANNEL DOLBY STEREO ARRANGEMENT



The above arrangement made more sense than the one shown in Figure 1. Three channels are used to create a frontal sound-stage with a discrete, centre loudspeaker. The fourth channel provides ambience at the rear by means of a mono signal distributed over a pair of loudspeakers. Various arrangements of the rear loudspeakers are possible.

3dB, and without the logic decoders, their imaging was so poor as to be almost worthless. We therefore had poor encode/decode systems working on basic principles that were poorly understood. Looking back, it is quite remarkable how many people believed that quadrophonic localisation was dependent upon the position of a pan-pot, and not on human psycho-acoustics.

Even for people working on the mixing of discrete four-channel recordings, and even in specially designed quadrophonic control rooms, stability was still poor. It was essentially only stable for a person sitting in the geometrical centre of the loudspeakers with their head in a clamp. Attempts to position sounds centre-overhead would be very differently perceived by any two people sat side by side at the mixing console. Artistes became very disillusioned when they found that what had promised to offer so much, delivered so little. What they were hoping for turned out to be an unattainable carrot on the end of a stick, frustratingly close, but just out of reach. The psycho-acoustic lessons learned from quadrophony still apply to surround, but perhaps, this time, we can use our knowledge to avoid demanding too much, and hence suffering further disillusionment. We have also learned that four loudspeakers, one in each corner of a room, is not the way to go. Figure 1 is for history only.

4. CONTROL ROOM LIMITATIONS

The compatibility between quadrophonic control rooms was another big problem, not to mention the compatibility from the control rooms to the domestic listener. Back in the days of mono, control room compatibility was not a major issue. Even when control rooms had only the most rudimentary acoustic treatment, any problem in the overall response could usually be made acceptable by the judicious movement of either the loudspeaker, or the listening position. Once stereo arrived, so did the problem of finding control rooms which not only sounded good, but which produced stereophonic results which travelled well to the outside world. It was as though the compatibility problems increased by the square of the number of loudspeakers. Obviously, with stereo, the two loudspeakers must be moved as a pair, and individual positioning for best sonic results is not feasible. The listener also must remain within more defined limits, so flexibility of movement was much less than had been enjoyed in the days of mono. However, the benefits of stereophony were enormous, so work proceeded in the direction of finding solutions to the problems.

At the time of the advent of quadrophony, the state of development of stereo control room acoustics was still in its infancy, and was generally not good. Acoustic "design" was largely an empirical process. Even the first, specifically designed quadrophonic control rooms were often mirror images of the front halves of the better stereo rooms. Unfortunately, some of these better stereo rooms were only better because of the acoustic design of their rear halves. The low frequency build-up in some of the quadrophonic rooms which were based on the above principle was such that there was great difficulty in achieving consistent musical and timbral balances.

Serious quadrophonic control room design took place mainly between 1973 and 1979, but generally with only limited success, and it is interesting to note that the more psycho-acoustically advanced concepts of stereophonic control room design did not really begin to blossom until after the limitations of quadrophonic performance had been removed, in the early 1980s.

4.1 Phantom Image Fragility

A pair of papers[1,2] presented to the September 1997, 103rd Conference of the Audio Engineering Society (AES) addressed the subject of the requirements for optimal stereophonic monitoring. The complexities of mutual coupling, were also dealt with at some length. They set out the somewhat fragile requirements for the establishment of the consistent reproduction of stereophonic phantom images. Response disturbances due to listening position, loudspeaker position, room boundaries, the acoustic coupling between the two loudspeaker systems themselves, and the coupling of the room to the individual loudspeakers, all conspire to leave conventional stereo on very unsound foundations. Although it can work extremely well under close to ideal conditions, it is not, inherently, a robust concept. When quadrophonics attempted to increase to four, the number of phantom sound stages, it is little wonder that it floundered on the rocks of its own complexity. By the end of the quadrophonic era, many mixing engineers and producers had settled for a stereo frontal stage, with only ambience, or the odd special effect, emerging from the rear. Their hard earned lessons still largely apply to the surround concepts of today.

Somewhat ironically, this technique was also ideal for the excellent Ambisonics coding system, and matched well the concept of Dolby Stereo Surround (which led the way to the Dolby Digital Surround System, now so widely used in cinemas) but the lack of the availability of a discrete, stable, convenient, four-channel storage medium for consumer use was perhaps the final nail in the quadrophonic coffin. Maybe this was just as well, as it allowed a cleaner re-start of serious interest in surround sound in the 1990s. During the intervening decade, digital multi-channel consumer systems had become a reality, and the cinema industry had had time to develop its own surround sound practices. Indeed, it had built up a very considerable body of experience in surround systems and mixing techniques. It seemed as though we finally had some new and much more practicable prospects for multi-channel listening within our reach.

5. PREPARING FOR AUDIO-ONLY SURROUND

We now have an excellent opportunity to develop surround ready for an audio-only format. Until now, surround has been almost exclusively an adjunct to a picture, whether that picture has been video, television, or film. Whenever sound is accompanied by a visual image, the visual aspects tend to dominate the senses. In the better of the current surround productions, the soundtracks are of the highest quality, but these tend to be done in very expensive, suitably large, film or video studios. These studios usually have a wealth of dedicated, highly trained

staff, such as was the case in a significantly large proportion of sound studios twenty years ago. It is, perhaps, an interesting contrast to think of what will happen when surround audio-only hits the mass market.

The sound recording industry has entered a phase whereby such luxurious conditions can only usually be afforded by some of the very top artistes. By far the majority of current stereo productions are done in rooms with minimal acoustic treatment, and which are operated by what amount to little more than skeleton staffs. Already, there are around 100 Dolby surround encoding systems in studios in the UK alone, but the majority of these are in rooms which have been adapted to surround work, rather than having been optimally designed for it. These conditions would seem to be arbitrary, at least, in terms of their ability to work to the highest standards, or to achieve optimal compatibility, either between themselves or to the domestic destination for the end product. Most of them are primarily stereo rooms, with add-ons to upgrade for surround. Let us then look at the needs for a purpose designed surround room, with full stereo compatibility, and which can be achieved within a realistic budget in terms of the state of the current audio-only market.

6. THE BASIC SURROUND LOUDSPEAKER FORMAT, AND FUNDAMENTAL MIXING PRACTICES

It is now emerging that the so-called 5.1 surround system is here to stay for the foreseeable future. This system employs three, full-bandwidth, front loudspeakers, in a left, centre, and right configuration. There are also two full-range channels for the left and right rear information, and a further channel, usually of limited bandwidth, for a single sub-woofer. Most modern stereo control rooms are symmetrical about their front-to-back axis, but are asymmetrical about their side-to-side axis. This arrangement is perfectly suitable for surround, when, in all the current working practices, the frontal sound stage is predominant. However, it should be remembered that it is not suitable for the old quadrophonic concept of a wrap-around sound stage. We shall discuss this in more detail, later.

6.1 The Importance of a Discrete Centre-Front Loudspeaker

One of the great benefits of the 5 and 5.1 channel surround formats is that they use a discrete centre-front loudspeaker. It was the use of a centre-front channel which gave Dolby Stereo its great advantage over the old quadrophonic formats. Probably the greatest weakness of two-channel stereo is the instability of the central image, which in the majority of cases is the location for the main performances, such as vocals. The centre phantom position is also usually the location for bass drums and bass guitars, which are two of the instruments which suffer most from room to room inconsistency problems. However, if these instruments are positioned in the discrete centre channel, not only does the stereo panorama remain more stable for different listening positions, but also, the room to room timbral difference problems are greatly reduced.

(See Reference[1]). Effectively, a single centre front channel is a mono source, and as mentioned previously, compatibility problems were much less evident in the days of mono control rooms.

The centre channel is of critical importance if the full benefits of surround are to be realised. It is the key to greater image stability and also offers better fidelity. It thus makes no sense to compromise in any way the quality or the output capability of centre monitoring loudspeakers. Unfortunately, current practices often do just this, by using a smaller central loudspeaker in order to more easily accommodate a control room window or a set of video monitors. There seems to be made a frequent assumption that the centre channel is dialogue only, but this greatly wastes the potential of the three channel frontal stereo. For reasons of mutual coupling between the loudspeakers, a single woofer cabinet in the centre position will not be timbrally identical to either a physically identical left/right pair, supporting a phantom centre, or to dual woofer left and right loudspeakers, either singly, or in pairs. Equalisation decisions will tend to be influenced by this fact, and they may not be appropriate when reproduced on other systems.

The critical nature of the centre front image, which may place vocals, bass guitar, bass drums, and other instruments in one, central loudspeaker, is even more reason why that loudspeaker should be of equal size to the left/right pair. In fact, rather than having a centre loudspeaker of only half the low frequency power handling (as in the case of a single woofer centre loudspeaker in between a double woofer pair) the power going into a central loudspeaker will be at least double the power going into each of a left/right pair when producing a phantom image at the same total SPL. It may even be slightly more, as the pair would develop a little extra low frequency output by virtue of the doubling of the radiating area (or quadrupling in the case of a double woofer left/right pair used with a single woofer centre), and the influence which that has on the mutual coupling. Think about it, a bass guitar reproduced by a single woofer centre channel will sound quite different to a phantom centre image of a bass guitar reproduced by the four woofers of a dual woofer, left/right pair.

6.2 The Use of Sub-Woofers

In the highest quality stereo reproduction, the use of separate, mono sub-woofers has not, in general, been well received. Unless all the loudspeakers are very carefully positioned, there can be arrival time mismatches and consequent transient smearing problems. Nevertheless, in problem situations, and by judicious siting of the cabinets, the use of a mono sub-woofer can sometimes make the difference between a monitoring situation being workable, or not. As with mono-only control rooms, if the low frequency response is uneven, a single source can be re-located for an improved overall response much easier than a pair of sources. However, when one expands to multiple sources, as can be the case with surround if things are so routed, then a single source of low frequencies can become a definite advantage. In the majority of cases, the price which one pays in stereo, is more than made up for by the extra spaciousness available from surround.

With a single source of the very low frequencies, say below 80 or 100Hz, there is far less potential for the development of room problems. Again, as in the days of mono, a single source also makes the room to room compatibility problems less serious, as there are less position dependent variations or complex interference patterns. The ideal siting for a sub-woofer is in the front wall, but it should usually also be positioned off-centre. This avoids driving the room symmetrically, so the left and right spaces ("halves") into which the woofer delivers its power will be different in terms of their modal frequencies. The modes will therefore tend to overlap, rather than superimpose as mirror images at identical frequencies. The technique tends to produce somewhat more peaks and dips in the overall response, but they have less amplitude deviation than the superimposed modes of a symmetrically driven room, and are thus less audibly offensive.

6.3 Layout Considerations

In order to avoid arrival time problems, the three main loudspeakers should be mounted on an arc of a circle, with its centre point just behind the main, critical listening position. The sub-woofer should also be positioned on this arc, typically about half way between the centre and either the left or right loudspeakers. The monitor wall should be as rigid as possible, and, ideally, the amplifiers should be mounted behind the wall, situated as close as possible to the loudspeakers to allow the use of the minimum length of loudspeaker cable. A loudspeaker wall is now almost universally required for cinema mixing and performance, and this good practice should extend to any serious, surround, music-only studios.

The practice of compromising the central loudspeaker for the benefit of having better forward vision into the studio can be resolved by the construction of mix-only rooms. This was common enough 25 years ago, although the necessary equipment to set up such a room was much less in those days. Nowadays, the cost of setting up mix-only rooms for stereo is not far short of that of a full recording facility, and the economics have tended to favour the construction of multi-use rooms. However, the film industry rarely records and mixes in the same rooms, as the differing needs are well recognised. In this instance it is the musicians who play to the picture during the recording (with a screen in the studio), and the mixing personnel who mix to the full size picture (or thereabouts) in the mixing rooms (dubbing theatres). They do not record in surround, so the recording studios can have a large, frontal, control room window, if they so desire, and a less than optimal centre monitor loudspeaker for any quick-reference purposes. High quality surround mixing, on the other hand, calls for something different, and control room windows do not comply with these needs.

6.4 Different Dynamics and Sensory Interconnections

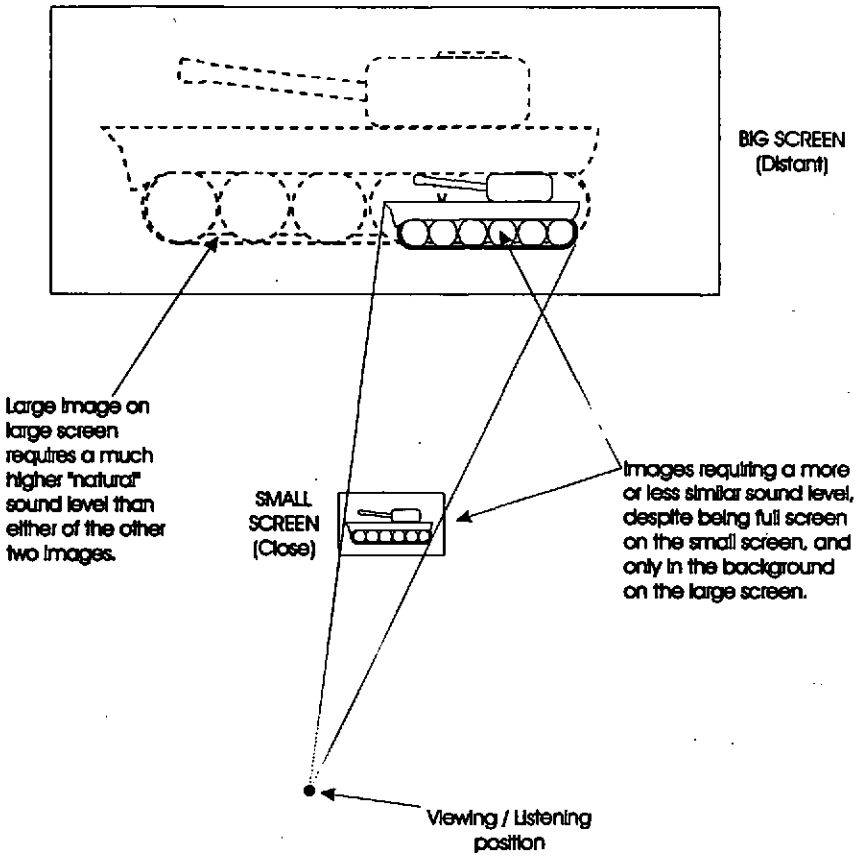
Surround mixing for cinema, video or television, all have different mixing techniques, as the sound level of the mix, and the appropriate balance, must relate not only to the actions on the picture, but also to the picture size. For example, a picture of a moving battle tank, filling the full frame of a television monitor, at a close distance to the viewer, would tend to demand its

own "natural" level. An image of a distant tank on a large cinema screen, but having the same apparent, physical image size of that which was full-frame on the closer television screen, would also call for the same sound level. See Figure 3. However, a full-frame picture of the tank on a large cinema screen (with the tank appearing enormous) would call for a much higher sound level. We have very strong impressions in our minds of the size of something we recognise, and we can relate the size of its image to the distance from us. We instinctively know what sound level we would expect to hear from the apparent size of such a visual impression. Obviously, though, life-like size requires life-like dynamic range, and these days, for maximum overall impact, that is precisely what the cinema performances are seeking to deliver.

Of course, in the listening/viewing room the peak sound levels associated with a large screen would be much higher than those associated with a small picture, such as of the size that could reasonably be expected in home cinemas. In a Dolby Digital Surround cinema presentation, the theatre noise floor could be around 35dBA, with the peak levels being over 110dBA. This yields a 75dB usable dynamic range. In a domestic environment, both the screen size and the neighbours, may only be able to support a peak level of 90dBA, and, what is more, many homes will have higher noise floors due to traffic, weather, other residents, refrigerators, washing machines, or a host of other noise sources. In such circumstances noise floor of 40dBA with a peak level of 90dBA would yield a usable dynamic range of only 50dB, 25dB less than the optimal cinema mix for the same film. Because of the complexity of the interactions of music, dialogue, and effects, and giving due respect to the fidelity of the sound track, simple compression and expansion techniques will be unlikely to yield the most pleasing results from inter-system exchanges of soundtrack. For the best results, separate mixes are necessary for separate purposes.

On the other hand, when we consider the forthcoming music-only surround, no such picture-related restrictions will apply. It would seem to be logical that the current philosophy of mixing surround sound specifically for its intended market should continue. The extra dimension of surround sound is capable of producing powerful and new sensations, but for the maximum audio-only effect, a very different mix may be required than would be optimal if the same music were picture related (dominated?). Music-only mixing will probably follow the existing stereo techniques of referencing to large and small loudspeakers during the mixing process. The overall dynamic range of the music can then be chosen to find a best-fit across a range of listening conditions and levels.

Figure 3 VISUAL IMAGE TO SOUND LOUDNESS RELATIONSHIP



6.5 Optimisation of Rear Monitors

The specifications for the actual surround (rear) loudspeakers create some dilemmas. In order to create the most diffuse sound field, smoothly covering the whole audience, the cinema industry has generally opted for the use of a large number of loudspeakers. This works well in cinemas built to the appropriate acoustic guidelines, where a large audience must be given an impression of the general ambience without any member of the audience localising the sound source to any specific loudspeaker. If, say, a single pair of rear loudspeakers were to be used, and the soundtrack were to be optimally balanced for a person in the centre of the cinema, then a person in the front row would hear a totally front dominated sound distribution, whilst a person sitting at the rear, close to one loudspeaker of the surround pair, would hear a totally exaggerated level of rear information. This problem, in large auditoria, can only be addressed by the use of many, spacially distributed loudspeakers.

In a home theatre, the situation is very different. Potential level imbalances over the listening area are not so great, nor is the size of the audience which must be evenly covered. What is more, the acoustics of the listening conditions exhibit a greater degree of variability than in purpose designed cinemas, so diktats of excessive rigidity are meaningless. These domestic conditions are also the ones which will generally apply to music-only surround. In order to introduce a more diffuse field in the home situation, THX has recommended the use of dipole loudspeakers for the rear pair, with their main axes running parallel to the side walls. This is a reproduction recommendation, though, and not something intended for use in mixing rooms. Its aim is to simulate, with two loudspeakers, the diffuse effect of the multiple loudspeakers of the cinema systems. Doubtless, these types of systems, if purchased for home cinema use, will also be used for music-only reproduction, which leads to the question as to what is the best monitoring philosophy to follow in music-only mixing rooms. The timbral differences, together with the different spacial perception, between a single pair of surround loudspeakers and a multi-speaker arrangement can be quite great. If definition is required in the rear channels, then undoubtedly the single pair is preferable.

We do seem to be faced with a situation, though, whereby recommendations as to the optimal surround arrangement may be a necessary adjunct to music-only releases if different recordings are mixed according to different concepts of what surround should be. However, as with conventional stereo, (though not necessarily with conventional thinking on the subject), it would seem to be better to standardise, as closely as possible, the recording monitoring conditions. What is more, it would also seem sensible to discuss conventions for instrument positioning, such that the choices made in the home listening environment are at least consistently appropriate from one recording to another. This requires a consistent approach to mixing techniques, the layout of the loudspeakers, and the bandwidths of the various channels. We currently have the potential to make ourselves a big mess for the future, so careful and wide-ranging discussion of surround concepts, now, would seem to be a prudent course to follow.

6.6 Desirable Loudspeaker Bandwidth and Associated Mixing Considerations

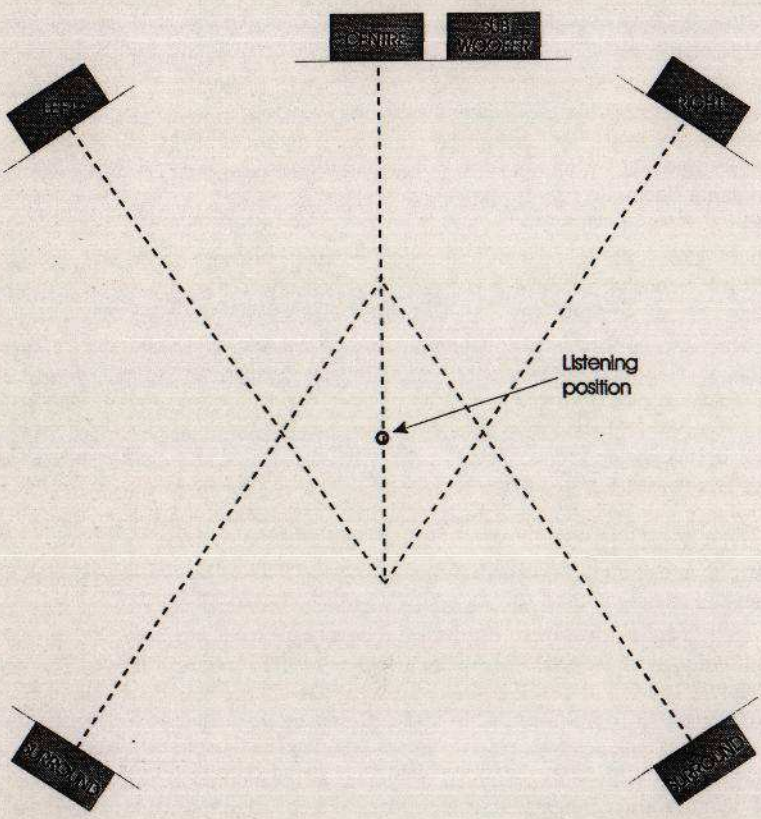
The discrete digital surround systems for domestic use will, almost certainly, provide five full audio bandwidth channels, plus a selectable sub-woofer facility. It would thus appear that what we need is a layout, as shown in Figure 4, with five identical monitoring loudspeakers, plus a sub-woofer system. Some of the older surround systems, using analogue matrixing, delivered only limited bandwidth signals into the surround loudspeakers, but there is absolutely no problem in either mixing or reproducing such encoded music on full-range loudspeakers, in fact, it is desirable. The band limiting takes place in the electronics of the coding system, and if the mix is monitored through the matrix, then the vagaries imposed by compounding the electronic bandwidth restrictions by arbitrary loudspeaker roll-offs will be avoided. The performance through a whole range of possible limited bandwidth loudspeakers in home use, with their wide differences of sonic characteristics, cannot be pre-judged on some arbitrarily chosen, limited bandwidth, mixing monitors. It would seem to be a minimum (obvious?) requirement that we should consistently monitor what we are intending to sell.

Some of the older, and, indeed, some of the current mixing philosophies do not put any full-range signals simultaneously in the front and rear loudspeakers. This is partly in deference to the older systems which could not cope very well with such a state of affairs, but it is also the result of the lessons learned from quadrophonics, as described earlier. In cinema use, there are occasions when a single, full bandwidth signal will be routed to all loudspeakers simultaneously, such as when an explosion is fed to the full system to take maximum effect of the total power output capability. There are also times when things could be panned around the room, by means of five-channel "joystick" pan-pots. However, this type of panning is usually restricted to special effects, and should not be over-used. Of course, these situations are almost always picture related, and are done after careful consideration by experienced engineers.

Unfortunately, once music-only surround enters the domain of the project-type studio, and reaches full commercialisation, there seems to be little doubt that "artistic" reasons may lead to fixed images in surround locations; or even in overhead locations, but the lessons of quadrophonics should not be forgotten: there are prices to be paid for most things. The psycho-acoustics of phantom imagery are well perceived only in the frontal sound stage, and phantom images are more prone to room-to-room timbral differences than images located in discrete loudspeakers. The new surround systems offer an excellent scope for productions of great dynamism and impact, with only minimal reliance on phantom imagery, although the "washiness" of phantom imagery can be used to great effect where a sense of spaciousness is the aim. However, they still cannot support the unsupportable, and the psycho-acoustic limitations should be respected, or else the errors of quadrophonics will surely be repeated.

Figure 4

VERSATILE 5.1 CHANNEL, FULL-RANGE SURROUND MONITORING SYSTEM, WITH MULTI-FORMAT MONITORING CAPABILITY



With full-range loudspeakers, the above system can also be used for the monitoring of full-range, discrete 5 channel (with the sub-woofer disconnected), four channel full-range (a possible HDTV format), Dolby Digital Surround, various other Dolby Surround formats (some with limited bandwidth rear channels), any of the old quadrophonic formats, stereo and mono. Best surround results will generally be achieved in a room with a very low decay time.

6.7 Further Options for Sub-Woofers

Staying with the subject of the number of sound sources, we can now consider the application of sub-woofers in the context of surround. Except when seated in the geometrical centre of the loudspeaker placement, and under free-field listening conditions, the low frequency linearity will be more consistent from room to room, and more consistent at different positions in any one room, if all the sub-bass is derived from one source. It appears that the new surround systems will allow various options in terms of sub-woofer feeds, which can be selected according to taste or for an appropriate purpose. This should prove to be a useful facility. However, there are certain types of recording techniques and certain types of music, such as using spaced microphones on pipe organs in churches, where discrete bass sources can have greatly beneficial effects in terms of realism. The ideal music surround mixing room would therefore appear to need five full-range monitors. In this context, by full-range, I mean capable as going almost as far down, even if not quite so far down, as the sub-woofer. The sub-woofer response should extend as smoothly as possible from 100Hz down to around 20Hz, or even below, and the main systems should respond down to at least 25Hz before any serious roll-offs set in. This is the only way to allow either 5.1 channel, or the option 5 channel monitoring, (as some proposed formats would appear to require) on the same monitor systems.

7. ACOUSTICS FOR SURROUND MIXING

All of the above requirements call for well designed control rooms if detailed and representative monitoring is to be achieved. As mentioned previously, modern stereophonic control rooms tend to have different acoustic characteristics to their front and rear halves, but tend to be symmetrical in the left to right sense. Two prominent philosophies are the "Live End, Dead End" (LEDE) and "Non-Environment" approaches. Within their own concepts, there is nothing to suggest that either one is not capable of supporting surround monitoring, though certain consideration should be borne in mind.

In the LEDE approach, the front half of the room is treated with absorbent materials to suppress early reflexions, and the rear of the room is designed to provide some acoustic life, though usually of a very diffused nature with a few selected late reflexions in some instances. The "Non-Environment" approach seeks to remove the room from the monitoring equation as completely as possible, but prevents the rooms from becoming oppressive for its occupants by providing a hard front wall. Actions and speech within the room are acoustically livened by the reflexions from the front wall, and by a hard floor, (which is often common to both the techniques being discussed here). The front wall cannot affect the monitoring, however, as the loudspeakers are set into it, and all the sound travels away from the wall.

If either of these techniques were to be employed for a dedicated surround mixing room, then each would suffer from the fact that the rear loudspeakers would be facing a solid monitor wall.

This wall is reflective at all frequencies in the Non-Environment rooms, and at low-mid to low frequencies in the LEDE rooms. The Non-Environment approach could benefit in such circumstances from having a highly diffusive, and less reflective front wall, such that reflexions from the rear monitors would not excessively interfere with the perception of the front monitors. The LEDE would be less coloured in terms of rear monitoring if low frequency absorption was provided in the front wall, but this would upset the acoustic loading on the front monitors, which would possibly need more low frequency drive. The use of a diffusive front wall would probably render the rooms too lively. Both concepts suffer from the fact that the front wall will be reflective at low frequencies, and so will reflect any low frequencies emanating from the rear loudspeakers. They would thus suffer from LF irregularity in the rear monitor response. In practice, however, these problems may not be too great as long as the room limitations are taken into account by the mixing personnel. The rooms can support surround as long as the rear channels carry only ambience and effects. The old quadrophonic concept of wrapping the instruments around the listeners is still only consistently supportable in anechoic conditions.

In cinemas, the screen is reflective at middle and high frequencies, and the monitor wall, behind the screen, will be reflective at low frequencies. In homes, the front wall will almost certainly be reflective, at least at some frequencies. We thus have a common situation whereby the front halves of the listening environment find great difficulty in adapting themselves to receiving a direct impact from the rear monitors, and that the rear wall techniques are unsuitable for the physical mounting and acoustic loading of the main monitors. This, as was stated earlier, was one of the gross errors made by the designers of the earlier quadrophonic rooms. Face to face, neither two front halves, nor two rear halves, of any current control room philosophy will produce a good room, even for stereo, and these problems can only be compounded when the rooms are driven from each end simultaneously. In general, the rear walls of modern control rooms are there to control sound, not to generate it, so they are not ideal as baffle extensions. As in the case of the front walls, the baffle extension concept helps to more linearily load the loudspeakers mounted in them, and hence help to contribute to an overall more linear response in the room. Modern cinema designs usually call for an absorptive rear wall, and so share the same problems.

A fact which must be faced with surround monitoring, and this was a lesson learned at great cost by the people making the old quadrophonic mixes, is that symmetrical monitoring is not possible in any room other than an enormous anechoic chamber. Mixing techniques must take this fact into account if the best results and the most consistent effects are to be realised from surround. The front sound stage should dominate, with central sounds routed to the centre loudspeaker, and not panned across left and right as phantom images. Where possible, unless there are strong reasons to the contrary, the lower bass frequencies, especially of any frontal phantom sources, should be routed to the mono sub-woofer, and what appears in the surround should normally be limited to ambience and effects.

7.1 Decay Time Considerations

For purposes of quality control, the decay time of the room should be very low in order not to mask or confuse the monitoring. In the LEDE v.s. Non-Environment debate, at least as it relates to stereophonic listening rooms, the argument for the LEDE is that the perception of the music has more of a spacial dimension, and relates more to domestic listening conditions. The argument for the Non-Environment philosophy is that it facilitates greater in-depth monitoring, allowing performing or recording flaws to be detected at an early stage, before reaching the final C.D. The claim is also made for greater room-to-room, and system-to-system, monitoring consistency. In the context of their adaptation to surround monitoring, it would seem that as the ambient information can be incorporated into the 5-channel surround system, there would be no need to try to enhance the listening space acoustically, so a very low decay time would appear to be desirable. This will allow better judgement of the amount of ambient information, and it also allows superior monitoring of the low level signals and any problems which lurk therein. Whatever limitations that the Non-Environment rooms may be perceived to have in stereo, would appear to be of no consequence in surround.

The advocates of the Non-Environment approach[3] have long argued that listening room enhancement should be left for the final reproduction rooms, and that the quality control function of the control room monitoring should not be compromised solely for the benefit of listening pleasure. On the other hand, the advocates of LEDE maintain that a sense of spaciousness is a necessary ingredient for the realistic appraisal of music. To some degree, it must be said, certain types of music or recording techniques suit certain types of rooms, but personal tastes and experiences also come into the equation. In the case of their application to surround, however, the all encompassing nature of surround signals renders the above argument somewhat academic. Low decay time and few reflexions would seem to be the requirement. The Non-Environment technique has already been applied, very successfully, in film mixing studios for Dolby Digital Surround, in many countries, and the transfer of the mixes into the cinemas has been very successful indeed.

8. SOME CONCLUSIONS

As is the case in most conventional stereophonic mixing rooms, it would seem prudent to provide a more typically domestic sized reference in control rooms used for music-only surround mixing. Large, full-range monitors, in a good monitoring room, would appear to be mandatory if the highest level of quality control is to be made possible. As an alternative reference, a small, domestic surround system, say with five satellites and a sub-woofer, could be placed at relatively close range, around the mixing console. Between the two systems, assessments could be made of dynamic range compatibility, spacial compatibility, timbral compatibility and musical balance compatibility. This is about as far as one can go with a domestic reference, as the problem of decreeing what is "typical" in world-wide domestic situations is even more complicated with surround than with stereo. There are more mutual

coupling problems between the loudspeakers themselves, and more complexities of room coupling. There is also a great range of domestic room acoustics that will be encountered, and, what is more, different manufacturers of domestic equipment are busily marketing their own concepts for their own advantages. There is currently a choice of full bandwidth or limited bandwidth surround (rear) loudspeakers; there is the dipole option, and there is the option to use multiple surround loudspeakers, as in the cinema systems. All of these possibilities, it must be realised, will produce some very different results in the compatibilities mentioned in the previous paragraph. To whatever degree is realisable in practice, it would seem beneficial, to all concerned, to try to establish control room concepts and working practices which will help to ensure a relatively consistent concept of frequency balances and spacial distributions which will best relate to home realities, yet will still retain both their recording quality and their artistic integrity. If this is not done, and the studios simply tend to follow the end-user market, then, surely, only chaos will follow.

After considering all the topics under discussion here, it will probably have become apparent that if the best results are to be gleaned from the full potential of surround, then merely adapting former stereo control rooms to surround use, by the installation of ad hoc extensions to the monitor systems, is hardly the way to go. Good surround rooms need building as surround rooms. Mixing practices need to take into account the problems of the psycho-acoustics of phantom image localisation and their stability problems. Surround sound mixes being performed by untrained personnel run the risk of producing disappointing results if they fail to translate from the mixing environment to the widely variable conditions to be found in domestic reproduction environments. Surround is not the simple concept that it may initially appear to be. Much will be learned once the widespread development of surround mixing facilities is realised, but, as we have already learned so much from the quadrophonic fiasco, it would seem wasteful not to remember the lessons which it taught at such great cost, as so many of them have great relevance to the emerging practices of surround sound. Careful planning of future surround facilities could yield a very exciting future.

REFERENCES

[1] Keith R. Holland, Philip R. Newell, "Loudspeakers, Mutual Coupling and Phantom Images in Rooms", Presented to the AES 103rd Conference, New York, September 1997

[2] Philip R. Newell, Keith R. Holland, "A Proposal for a More Perceptually Uniform Control Room for Stereophonic Music Recording Studios", Presented to the AES 103rd Conference, New York, September 1997

[3] P. R. Newell, K. R. Holland, T. Hidley, "Control Room Reverberation is Unwanted Noise", Proceedings of the Institute of Acoustics, Reproduced Sound 10, pp 365-373, Vol 16, Part 4 (1994)