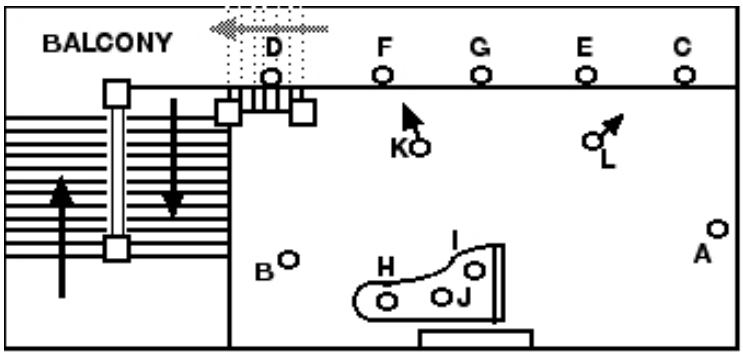
DISCRETE LAYERED SOUND

Philip R. Newell Consultant, Moaña, Spain
Shelley M. Katz Layered Sound Technologies Ltd.

1 THE CONCEPT

More than thirty years have now passed since the introduction of commercially available recordings with four or more channels. Much of the relative failure of quadrophonics during the 1970s was attributed to the limitations of domestic vinyl disc and analogue magnetic tape reproduction systems, but even in the 1990s, with the launch of multichannel digital discs, many of the old failings still seemed to be evident. Clearly, with the perfect reproduction ability of the digital discs, the medium could no longer be blamed for the lack of realistic reproduction. Something therefore seemed to be missing from the concept of the recording and reproduction systems, and not just from the distribution systems.

After some months of discussions between the authors of this paper, who both felt that a new approach was required, a recording session was arranged to take place over two days in the wood-panelled, Elizabethan Room, at Herstmonceux Castle in southern England. The first day's recordings were of a trio consisting of grand piano, flute and soprano voice, and the second day's recordings were of grand piano, double bass and a drum kit. The microphone arrangements are shown in Figures 1(a) and (b).



A,B	microphones 1m50 above floor level
C - G	microphones overhanging the balcony
H, I, J	piano microphones
K	soprano microphone
L	flute microphone

Figure 1a – Elizabethan Room, Herstmonceux Castle, microphone layout for Soprano, Flute & Piano

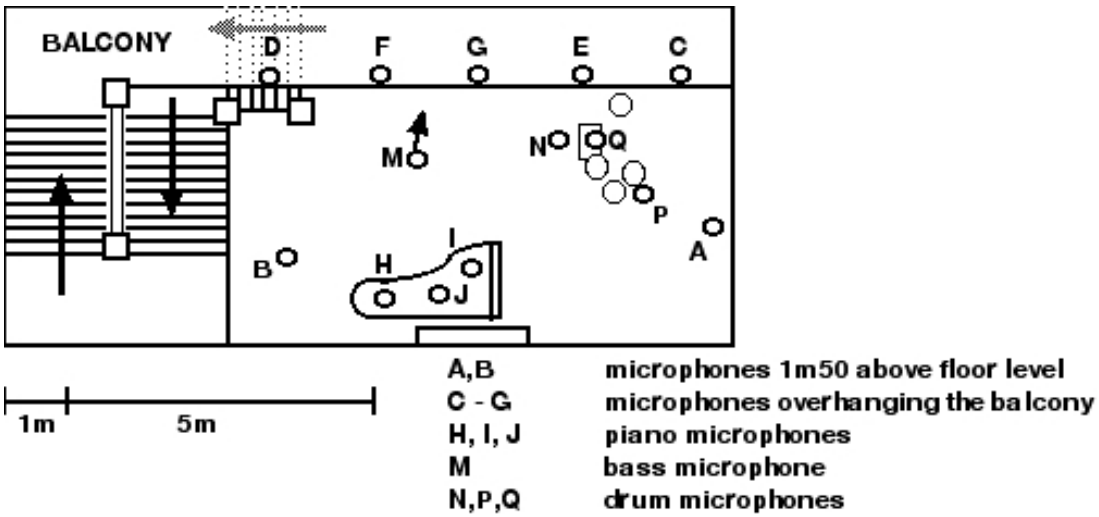


Figure 1B – Elizabethan Room, Herstmonceux Castle, microphone layout for Piano, Bass & Drums

What gave rise to the concept of these recordings was a pair of papers presented at the Reproduced Sound conferences in November 2003 and 2004. The first paper was highly critical of the current state of affairs with surround sound[1], whilst the second paper was suggesting a new means of creating much less correlated ambience in the reproduction room[2]. One of the main criticisms in the first paper was about the way in which so many surround systems and designs of surround room-acoustics were limiting the full potential of the frontal stereo sound stage, and one of the main benefits claimed in the second paper was that distributed mode loudspeakers (DMLs) could reproduce ambience better than conventional, pistonic loudspeakers.

It was decided by the authors that recordings should be made in such a way that the microphones close to the instruments would be subsequently mixed as a conventional frontal sound stage, and reproduced by conventional high-fidelity loudspeakers. However, the room ambience microphones would be placed close to the room boundaries, and would be reproduced only via DMLs, which would be initially sited in the listening room in directions corresponding to the ambience microphone directions from a chosen listening position in the recording room.

2 THE DMLS AS A DIFFUSE SOURCE

Transverse bending wave loudspeakers, of which DMLs are currently the best known commercially available examples, radiate in a manner which is spread out, both physically and temporally. At least at the higher frequencies, the radiation emanates from multiple points on the panel, and the transient response is somewhat smeared in the time domain. Another characteristic difference between pistonic sources and DMLs is that the DMLs radiate in a generally omni-directional manner which is relatively equal with frequency. The above characteristics of DMLs have led to some difficulties in their application to high fidelity music reproduction, where the radiation from DMLs has not been conducive to the traditional goal of reproducing the waveform of the electrical input signal. Smearing transients is, by definition, not high fidelity reproduction where the accurate reproduction of the impulse response is the Holy Grail.

3 THE PISTON LOUDSPEAKER AS A SOURCE

Taking as an example of a high fidelity pistonic radiator a loudspeaker such as a Quad ESL 63 electrostatic loudspeaker, its transient response, as can be seen from the step-function shown in Figure 2, is exemplary.

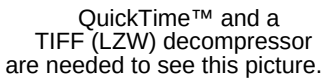


Figure 2 – Step Function of a Quad ESL63

The loudspeaker is also designed such that as the frequency rises, the source area both diminishes in size and recedes in apparent distance from the front surface of the loudspeaker, in order to emulate the radiation from the hypothetical point source, or small pulsating sphere. As conventional, two-channel stereo loudspeakers, the reproduction from such devices can be extremely lifelike, especially when reproducing close mic'd instruments. In a well-designed room, the reproduction of a frontal sound-stage via such loudspeakers can almost define the current state of the art. If distributed mode loudspeakers were to be substituted for the electrostatics in listening tests, the 'failings' of the DMLs in terms of conventional reproduction would be plainly evident. The very concept of the DML is not conducive to the reproduction of such a clearly defined frontal sound stage.

4 ARCHITECHTURAL REFLEXIONS AS SOUND SOURCES

Reflexions from architectural surfaces do not behave like the visual images in a mirrored room. Even a narrowly beamed source of high frequency sound, aimed at a flat wall, will not reflect in a manner much akin to what its continued trajectory would have been had the wall not changed its direction of propagation. Even at frequencies as high as 10kHz, a narrowly beamed incident source will reflect from the surface with a broadened conical directivity pattern, with a minimum of around 30 degrees. The concept can be visualised better by a quick look at Figure 3[3].

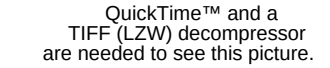


Figure 3 – Visualisation of reflexions from a directional sound source

Lower frequencies will exhibit wider reflected directivity patterns. If the reflective surface is relatively free to vibrate, such as with wood panelled walls, the re-radiation from the surface will be even more distributed, both spacially and temporally. In fact, the musical reflexions from many architectural surfaces behave rather similarly to the radiation of musical signals from DMLs.

We therefore appear to have a situation whereby the frontal sound stage of a conventional recording with ambient surround is best reproduced by pistonic loudspeakers, with fast transient responses and small source areas (commensurate with wavelength). Conversely, the surface reflexions which form the ambient rear channels of a recording would appear to be potentially better reproduced by loudspeakers radiating with the characteristics of DMLs. The recordings at Herstmonceux Castle, mentioned in Section 1, were intended to test this hypothesis.

5 IACC, AND REFLEXION DIRECTION

Barron's derivation of the LEF (Lateral Energy Fraction)[4] states that the sense of spaciousness created by lateral reflexions roughly corresponds to the cosine of the angle of arrival at the listening position. The angle is taken to the axis going through the two 'ears', in other words, it can be considered a sine measurement with 0 degrees as straight ahead. The spaciousness effect created by the lateral reflexions is optimal when coming from 90 degrees to the listening axis, and negligible when coming from 0 degrees – centre front. The cosine of 60 degrees is 0.5, which would suggest that we would lose half of the optimal effect if the reflexions came from only 30 degrees from centre front, which would be a typical location for stereo loudspeakers of a pair subtending 60 degrees at the listening position.

In a recent paper, Toole[5] reported that "In both concert halls and listening rooms reflected sounds arriving from the front or rear do not contribute to a positive impression [of envelopment]. [There is] a diminished preference for reflected sounds arriving within about 20-30 degrees of the median plane.". In general, the recorded reflexions in a stereo presentation with a 60 degree total subtended angle cannot deliver a sense of full spaciousness.

That sense of spaciousness is largely a result of low inter-aural cross-correlation (IACC). In other words, the delivery of significantly different signals to the two ears. Obviously, an acoustic signal arriving from centre front will arrive symmetrically at each ear of a listener on the central plane, so the IACC will be high. In the test recordings, therefore, the reproduction of the reflexions was initially set with loudspeakers either side of the listening room at 45degrees and 135 degrees from centre front. The inherently low IACC of the DML reproduction (as compared to pistonic sources) was also expected to aid the perception of spaciousness.

6 THE TEST RECORDINGS

The recordings were carried out in early September 2005. The instruments in each trio were mic'd in a conventional way, using AKG C414 microphones. Similar microphones were also placed in seven positions in the recording room; two at floor level (one at each side), and five along the balcony which ran along the back of the room. A photograph of the room is shown in Figure 4.

The intention was that each microphone should receive a unique perspective of the trios. Each microphone would receive the direct wavefronts from the instruments in a unique manner, with the instruments at the left hand side of the room arriving at left-located microphones before the instruments placed in the centre and the right of the room. The opposite would apply to the microphones at the right of the room, whilst the one, central microphone on the balcony would receive the acoustic signals from the left and right instruments in more or less equal timing. This was intended to be used if overall room reverberation were to be added. The signals from each microphone were recorded on individual 'tracks' of a hard disc recorded, all via good quality A to D converters, and no equalisation, compression or other processing was used.

Figure 4 – The Elizabethan Room, Herstmonceux Castle, East Sussex

7 THE FIRST REPRODUCTIONS

On the day after the recordings had been completed, the hard disc recorder and a small mixing console were transferred to a relatively large ‘hotel room’ sometimes used to accommodate guests at the castle, and which was situated in one of the towers. The room is shown in Figure 5, from which it can be seen that it had a carpeted floor, a double bed, general furnishings including a sofa, and also several large windows.

Figure 5 – Apartment 2 at Herstmonceux Castle

A stereo mix was made of the close microphones on the instruments. This was reproduced via Yamaha NS10 loudspeakers, which were used because of their familiarity to all the recording personnel involved. Four of the ambience microphones were then selected for reproduction via four Amina ACPT DMLs, positioned at roughly the aforementioned 45 and 135 degrees either side of centre front. The microphones which were selected were fed to the panels which most closely related to their angular distribution in the listening room, in the hope of reproducing via the panels the correct time and level information that would have been perceived by a listener in the centre of the recording room during the performances. Only the signals from four of the microphones were therefore used. It was hoped that the reproduction via the panels would closely mimic the nature of the architecturally reflected sound as described in Section 4.

The microphones, incidentally, were set to a cardioid response, because it was intended that only the direct sound to the boundary should be recorded, it being unknown at that time whether the rapid reflexion from the wall, if picked up by an omni-directional microphone, would confuse the ear on playback.

Obviously, the four microphones which were used for the playback of the ambient spaciousness represented only a fraction of the total reflexions within the room, but many acoustic studies have shown that, in fact, only a few reflexions are normally required by the ear to detect a sense of space. The simplicity factor also related back to a conversation in 1983 between one of the authors (Newell) and Hugo Zacharelli, the inventor of the Holophonics system of headphone reproducible surround sound. Zacharelli had said that he felt that it was best not to overload the brain, but to stimulate it just enough to give rise to the desired perception.

The big question awaiting answer at the replay stage related to whether the reflexions from the walls of the listening room, about two metres from the rear of the omni-directional DMLs, would create confusing time response anomalies, or whether they would reinforce the effect of the spaciousness.

When the four ambience microphone channels were faded up, the four people in the listening room were immediately aware of the sense of being in the recording room. This was a surprise to everybody, but especially to the two very experienced recording engineers, who had deliberately not been told about the true purpose of the recordings which they had been making. They had been told that they were making showcase recordings for the Fazzioli piano which had been specially rented for the experiments, which was at least half true!

8 SYSTEM ROBUSTNESS

At the time of the first playback, the musicians and engineers in the room had never heard such realistic reproduction of a space. Once the recording engineers had been fully informed about what was happening, questions immediately started to arise. “What’s your algorithm based on?” was one of the first questions, and they were surprised to hear that no electronic signal processing was involved. After the whole concept had been explained, attempts were soon made to upset the system, by moving the panel loudspeakers from their designated locations. They were stand-mounted, which permitted the heights to be adjusted irregularly. Their relative distances from the listening area were also varied, and their relative angles to centre front were changed.

Quite contrary to the situation with conventional surround loudspeaker systems, which should be sited on the circumference of a circle centered on the listening position, all set at the same height, and all adjusted to within a decibel of each other, the Herstmonceux system was highly tolerant to heights, angles, distances from the listeners and level differences. This bode well for domestic reproduction systems, where loudspeakers are often placed where they are convenient for domestic life rather than being positionally optimised from an electro-acoustic viewpoint.

In practice, the reflexions from the hard walls of the reproduction room also seemed to have little effect, thus relieving one of the aforementioned worries. However, this lack of effect was recently highlighted by Toole[5] who stated that “The basic audible effects of early reflexions in recordings seem to be remarkably well preserved in the reflective sound fields of ordinary rooms.” In the case of the recordings being discussed in this paper, the temporal relationships between the time of flight from the different instruments to the different ambient microphones, and the attendant level differences in the left and right hand sides, appear to have coded the acoustic sound stage in a very robust way. The whole system showed a very high degree of positional abuse tolerance, and even level changes at the mixing console of up to 5dB in various loudspeakers still yielded reasonable results.

In fact, it was even abused so far as to put one of the surround loudspeakers on the floor, behind the sofa, another on the floor in a corner, and the remaining two at different heights, angles and distances, and still the reproduction sounded recognisably like the recording room.

9 CONVENTIONAL, PISTONIC SURROUND LOUDSPEAKERS

In order to test whether the robustness was a sole attribute of the recording concept, the DMLs were replaced the next day by conventional domestic loudspeakers with approximately the same frequency range and cost. The result was abject failure. With panel loudspeakers on the front channels, the results were even worse.

The non-flat power response of the piston loudspeakers could not send the same frequency response to the listening position as was captured by the microphones whose signals they were reproducing. The ‘point source’ of the high frequencies – a dome tweeter of less than 20cm diameter – could not even begin to emulate the spacially and temporally distributed re-radiation that would be characteristic of an architectural reflexion.

10 THE OVERALL IMPRESSION

The entire system performed very realistically. With two long days previously spent in the recording room, the sensation of being there was still fresh in everybody’s minds. The realism of the reproduction was uncanny. The frontal stereo sound stage was very clear and well defined, and nothing in the surround channels detracted from its clean and tight response. It all happened in its own time and space, and no reflexions were heard before their time of flight from the instruments to the microphones, and from the DMLs to the listeners. The natural reverberation of the Elizabethan Room was, of course, captured to some degree by all of the microphones.

A further, unusual experience was perceived by all present, in that everybody was aware of a sensation of clean, deep bass, way below what any of the loudspeakers were capable of achieving. The reason behind this has still to be investigated, but it was postulated that some reconstruction of fundamentals was occurring in the brain after receiving the unusually accurate stream of harmonic, spacial and temporal information. Further tests need to be carried out to investigate this phenomenon. Mutual coupling between the six loudspeakers seems to be an unlikely candidate because of the lack of simultaneous reproduction due to the time of flight to the surround microphones.

Despite the known colouration of the DML reproduction, very little 'DML sound' was perceived in the playbacks, probably because neither the colouration of the panels nor the wall reflexions in the listening room were colouring the perception of the front, direct sound stage. Indeed, nothing else was emanating from the same direction. The primary sound stage was clean at all times, exactly as in the case when that stage consists of real instruments.

11 BEYOND ACOUSTIC RECORDING

Many modern day recordings are built up track by track, and are recorded in rooms which do not have the ambience required for the finished mix. Obviously, the recording technique described in this paper would not relate directly to such recordings, which make up the majority of current recordings. Nevertheless, surround reflexion generation could emulate a large degree of the acoustic effects. By defining the room size, the stage location, the position of each instrument on the stage, the listening position, the wall surfaces, and other such parameters, a group of instruments, recorded separately, and even in different rooms, could be electronically inserted into the desired 'room'. The virtual microphone locations for the reflexion pick-up would then be defined, and would receive the signals from the different instruments with different delays depending on the distance from each virtual microphone to each virtual instrument. Reverberation would then be applied also to the instrument microphones after an appropriate delay.

12 SUMMARY

In this paper a surround sound system has been described which requires six transmission channels. The degree of reality shown under the test circumstances, and subsequently in other locations, has been, in the opinions of the authors, a significant step forward. The frontal stage pair are traditional in every way, except that they do not carry all the ambience information as usually applied in conventional stereo. In the system described, the principal ambient information is transmitted via DMLs, whose inherently diffuse time and spacial responses are well suited to the reproduction of room reflexions and reverberation. Additionally the reflexions and reverberation reproduced via these panels are very robust in terms of their interference with domestic listening rooms.

The permutations of which and how many ambient microphones would be fed through which and how many panels are virtually endless, but the system described using four ambient loudspeakers with only one cardiod microphone feeding each panel has been shown to be both realistic and robust. A five channel system using a centre-rear channel was rejected on the grounds of the difficulty of siting a centre-rear channel in many domestic situations – typically with the settee close to a wall. A centre-rear channel also fails to enhance the IACC. A four channel system using only two lateral panels was tested and deemed to be good, but distinctly inferior to the four-panel, six-channel system.

The question of whether the optimum room sound would be better characterised by the use of cardiod, figure-of-eight or omni-directional microphones is an issue which requires further experimentation. In the light of Toole's recent work also citing other work by Beck[6] it could perhaps be that if the first arriving reflexions from the wall of the listening space do not confuse the perception of the late reflexions in the recording, then the reflexions from the walls adjacent to the microphones in the recording room may also be beneficially captured by figure-of-eight or omni-

directional microphones. It is unlikely that they would be destructive, but whether they would be useful in increasing the lateral reflexion density by time-smearing the late reflexions is something that could only be judged by the ear.

Perhaps what has been reported in this paper is only the tip of an iceberg. If this is so, then the prospects for further work are very promising indeed.

13 POSTLUDE – EXPLORING THE LIMITS OF LAYERED SOUND

In the late 1980's, a search was begun for a digital piano that would reproduce the sound of a high-quality, Steinway concert instrument. In the late 1990's, an instrument was finally built and demonstrated at St John's Smith Square which was deemed by several experts in the field to sound bigger, richer and louder than the resident Steinway concert grand. The method of combining loudspeakers to achieve the remarkable piano sound was then tested with conventional stereo recordings, and found to provide in terms of the emotional impact of the reproduced sound, an auditory experience subjectively superior to any other method of reproduction.

All material available in recorded format today contains, to varying degrees, the full content of an auditory scene, from highly defined, close mic'd instruments, to ambient noise and reverberation, whether original and acoustically recorded, or artificially inserted in judicious post-processing. The Layered Sound technology suggests that different methods of propagating sound, using fundamentally different transducers, handles the different parts of the complex audio stream better than a single method of propagating sound. By employing differing transducers, the brain processes the resultant sound stream more easily, and cognates a gestalt which is greater than the parts.

In this paper, the authors report on an experiment which maximizes the separation of parts. In addition to the standard method described as Layered Sound, in which there is a separation and distinction through the physical nature of the transducers, the authors test the total segregation of the direct and ambient sound stream, including dedicated mic placement and the maintenance of dedicated channels, with no final mix-down. As far as the authors are aware, this is the first paper to report on an experiment in which the original and unadulterated elements of the direct and ambient sound streams perceived in reproduced sound are kept entirely separate, from recording to propagation, in both space and time. The results suggest a novel sound-reproduction methodology which can be implemented using standard technology, and the qualities of which seem to provide several advantages over all other methodologies known today. Layered Sound is a technology born out of the desire to put the visceral, emotional content of live performances into the reproduction of sound. The authors are convinced that Discrete Layered Sound is an important development of the basic concept, and suggest that the future story of sound reproduction is far from known.

14 ACKNOWLEDGEMENT

The authors would like to thank Joe Fossard and Julius Newell for their work as recording engineers, and for their generosity in providing all the microphones and recording equipment used at Herstmonceux Castle in September 2005. Thanks are also due to Sergio Castro and Reflexion Arts SL for the loan of the Apogee A to D converters.

The authors would also like to thank the musicians, Diana Gilchrist, Abigail Dolan, Alastair Hume and Tristan Fry for their kind willingness to make the recordings, and their permission to use the recordings for further tests, study and distribution.

15 REFERENCES

1. P.R. Newell and K. Holland, Surround Sound – The Chaos Continues, Reproduced Sound 20, Oxford, UK, Proceedings of the Institute of Acoustics, Vol 26, Pt 8 UK (2004)
2. P. Mapp and S. Katz, Layered Sound – A New Approach to Sound Reproduction, Reproduced Sound 19, Oxford, UK, Proceedings of the Institute of Acoustics, Vol 25, Pt 8 UK (2003)
3. P. Newell, Recording Studio Design, Focal Press, Oxford, UK (2003)
4. M. Barron and A.H. Marshall, Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure, Journal of Sound and Vibration 77, pp211-232 (1981)
5. F. Toole, Loudspeakers and Rooms for Sound Reproduction – A Scientific Review, Journal of the Audio Engineering Society, Vol 54, No. 6, pp451-476 (June 2006)
6. S. Beck, Spatial Aspects of Reproduced Sound in Small Rooms, Journal of the Acoustical Society of America, Vol 103, pp434-445 (1998)