IMPROVED LOCALIZATION FOR BINAURAL RECORDINGS AND STEREO PROGRAM MATERIAL USING 'BLUMLEIN SHUFFLING'

JJ Digby   Electro-Acoustics Research Lab, University of Derby, UK   iD
AJ Hill   Electro-Acoustics Research Lab, University of Derby, UK   iD
BJ Wiggins   Electro-Acoustics Research Lab, University of Derby, UK   iD

1   INTRODUCTION

Near-coincident two-channel 'stereo' microphone techniques can provide realistic and useful representations of an auditory scene direct to binaural earphones (left and right). However, if the microphone signals are reproduced by a spaced pair of loudspeakers the 'binaural illusion' is disturbed by a progressive narrowing of 'stereo' image width as frequency decreases (amongst other possible effects). Alan Dower Blumlein's far-reaching 1931 patent describes a process which principally addresses this situation.[1]

This paper is born out of an ongoing research project which includes investigations into a satisfactory two-channel near-coincident microphone technique for: measurement, concert recording, convolution, and reproduction. A near-coincident baffled microphone technique was developed to suit these purposes.

For loudspeaker reproduction of the recordings a simple digital implementation of a Blumlein shuffling network was assembled using two types of built-in plugin in the Digital Audio Workstation *Reaper*.[2] Its effectiveness was surprising, and immediately apparent. Therefore, the following needed to be understood: how and why does this work? can the subjective and objective results withstand a detailed investigation? are there reasons to avoid it – a possible explanation for its general lack of familiarity? or has it been overlooked?

In this paper we provide: an overview of the 'Blumlein shuffling' process; a detailed description of the baffled near-coincident microphone arrangement; and demonstrations of Blumlein shuffling emulations using application software.   Included are audio demonstrations of processed and unprocessed recordings using this microphone arrangement, and examples of the reciprocal Blumlein shuffling process applied to commercially available program material for earphone reproduction.

2   BACKGROUND

Blumlein used the terms 'binaural transmission', 'binaural illusion', 'binaural reproduction', and 'binaural effect' with reference to both earphone and loudspeaker reproduction.
In a typewritten memo Blumlein describes the earphone presentation ('head receivers') of a pair of spherical response microphones (pressure, omnidirectional), near ear-spaced apart (200 mm is noted), and separated by a wooden baffle (perhaps of balsa)[3]:

> *Suppose in a rather live studio two microphones are arranged on each side of a block of wood roughly representing the human head, and suppose the outputs of these two microphones are combined and taken to a pair of head receivers in another room.*

When the two microphones are connected separately to the two receivers, the echoes are still heard by the observer, but he mentally discounts them and focuses his attention on the source of sound to which he is listening. The room does not sound dead, but the echoes are heard as such and do not worry the observer. It is this effect that it is desired to obtain by the proposed system of binaural reproduction. Of secondary importance, it is desired that the apparent position of a sound source shall be clearly indicated by the reproduction.

If the telephone receivers are replaced by two loudspeakers situated one on each side of the listening room, the binaural effect is lost.

A. D. Blumlein, "Binaural Reproduction – typewritten memorandum to Mr Isaac Shoenberg", Director of Research at Electric and Musical Industries Ltd (EMI). 1932[4]

In this instance, an oblique source's arrival time information is a result of path length differences between near and far microphones. In addition, an appropriate frequency-dependent level difference is provided by shadowing from the head-sized baffle. The use of earphones ensured the encoded directional cues from left and right microphones were independently assessed by each ear (binaural). However, the two loudspeakers placed in the front corners of the listening room are not prevented from reaching both ears of 'a central observer' (interaural crosstalk).[4]

The duplex theory of sound localization shows that: 'high-frequency sounds are localized through interaural level differences (ILD), and low-frequency sounds are localized through interaural time differences (ITD).'[5] In the above example the oblique positioning of the loudspeakers is important. Above approximately 1000 Hz the listener's head provides a shadowing of the near loudspeaker to the far ear.[6,7] However, for wavelengths comparable and above the size of the human head (500 Hz and below) diffraction allows the signals from both loudspeakers to reach both ears at a similar level.[8,5] Therefore, the timing differences encoded between left and right microphone are no longer presented *independently* to each ear.

2.1 'Change over' frequency

Beginning with the publishings of John William Strutt (Lord Rayleigh) in the early 1900s, attempts have been made to identify a boundary between ILDs and ITDs.[8] In recognition of the uncertainty, Strutt proposed a 'tentative' 500 Hz, and later stated that ITDs could not be used much above 400 Hz.[5] Blumlein recognized 700 or 750 Hz as a 'change over frequency', with important caveats:

This may vary within quite wide limits in different circumstances and from person to person, and that in any case the transference is not sudden or discontinuous but there is considerable overlap of the two phenomena so that over a considerable frequency range differences of both phase and intensity will to some extent have an effect in determining, the sense of direction experienced.

A. D. Blumlein, 1931[1]

Hartmann, Rakerd, and Crawford (2025) now state that regardless of the room environment, 'the perceptual effects of ITD and ILD become equivalent in the frequency range from 400 to 600 Hz'.[5]

2.2 'Binaural illusion' from loudspeakers: Blumlein's elegant solution

One method of obtaining a binaural illusion (with loudspeakers) is to convert the low frequency phase differences of the pressure microphone outputs into amplitude differences. Thus an oblique low frequency sound would produce phase differences in the microphone outputs, which … would be electrically converted to include amplitude differences, thus producing differences in output intensity of the two speakers.

The modification of microphone output described above may be called "shuffling".

A. D. Blumlein, 1932[4]

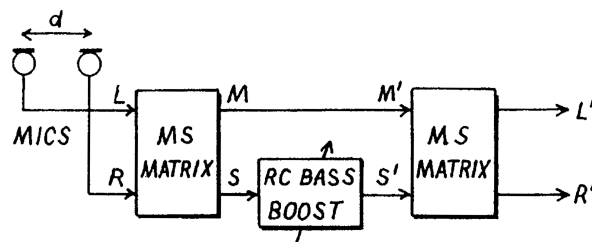A pictorial representation of Blumlein's sum-and-difference shuffling network is shown in Figure 1.

Figure 1: *Basic Blumlein shuffler for a pair of laterally spaced identical microphones pointing in the same direction – Gerzon (1992).[9, fig 3]* $\big(M \text{ (Mid)} = L + R \text{ (sum)}, S \text{ (Side)} = L - R \text{ (difference)}\big)$

At low frequencies Blumlein shuffling converts arrival-time differences between the two microphones into amplitude differences. 'Shuffled' low-frequency amplitude differences produced at the listener provide the low-frequency timing differences (ITD) necessary to localize a two-channel encoded source at a unified direction. In this application, Blumlein shuffling fulfills its original purpose of enhancing intelligibility: by maintaining directional cues which allow a listener to focus upon what has captured their interest ('directional-'[10] or 'binaural unmasking'[11]).

2.3 Blumlein shuffling implementation

The single-pole filter on the *difference* channel provides two effects: a flattening of the 20 dB/decade cut which occurs at low frequencies for an oblique source (arrival time difference), and a phase shift toward -90°.

Figures 2a and 2b show an FFT magnitude and phase analysis of an **unfiltered** *difference* channel: Left channel is a Dirac impulse, Right channel is the identical impulse delayed by 0.33 ms. Note the 20 dB/decade slope at a corner frequency of approximately 750 Hz, and a phase shift toward $+90^{\circ}$.

Figures 2c and 2d show the **filtered** *difference* channel; *Reaper's* `ReaEQ` VST plugin is used to create a -20 dB/decade low shelving filter with a boost of 20 dB: `Frequency` 300 Hz, `Gain` 20 dB, `Bandwidth` 1.23 octave.

Equation 1 is used to estimate ITD (c = sound velocity; 2*a* = head width, or microphone spacing; b = path difference).[12] Therefore, with a sound velocity c = 343 m/s, a delay of 0.33 ms approximates the ITD of a distant source at azimuth $\theta = 40^{\circ}$ with an ear spacing of 170 mm. Also, 0.33 ms corresponds to a microphone spacing of 200 mm with a distant source at 33.3°.

$$T = \frac{b}{c} = \frac{a}{c}\big(\theta + sin(\theta)\big) \qquad (1) \qquad\qquad F_{\text{transition}} = \frac{0.3}{T} \qquad (2)$$

(a) MAGNITUDE: **Unfiltered** *difference* channel for two Dirac impulses: Right channel delay: 0.33 ms.

(b) PHASE: **Unfiltered** *difference* channel for two Dirac impulses: Right channel delay: 0.33 ms.

(c) MAGNITUDE: **Filtered** *difference* channel for two Dirac impulses: Right channel delay: 0.33 ms.

(d) PHASE: **Filtered** *difference* channel for two Dirac impulses: Right channel delay: 0.33 ms.
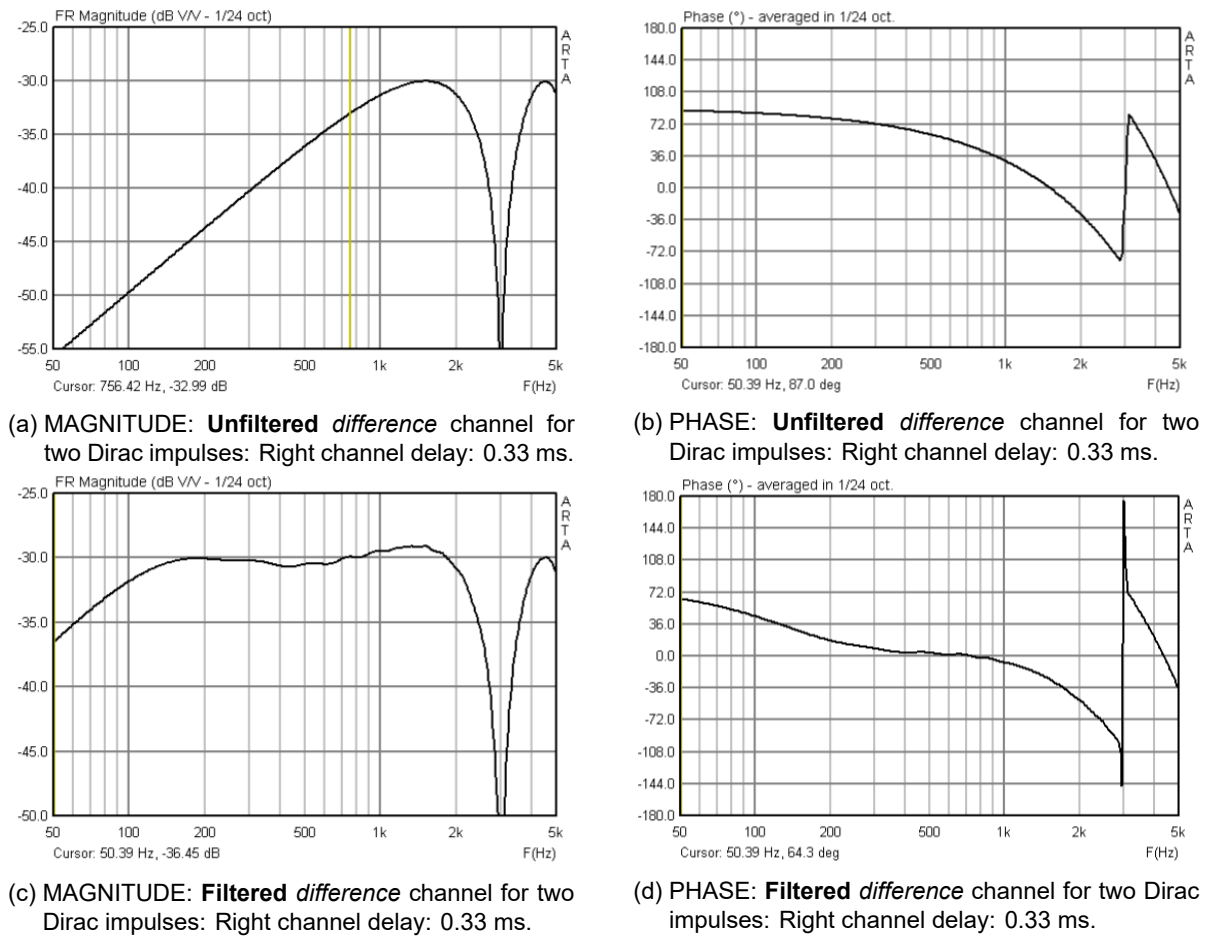
Figure 2: Magnitude and Phase plots of an **unfiltered** and **filtered** *difference* channel using Dirac impulses. Frequency range: 50 – 5000 Hz

Gerzon (1992) contains computational proofs which approximate a sum-and-difference network's '"transition frequency" between the low frequency amplitude stereophony region and the high frequency time delay stereophony'.[9, eq. 10–12]

Equation 2 shows that a transition frequency of 750 Hz occurs at $T = 0.4$ ms; an approximate ITD for a distant source at azimuth $\theta = 50°$ (T = arrival time delay between L and R).
This correlates with a recognized upper limit for detecting ITDs at 1500 Hz[7], and an ILD maximum for 1500 Hz at this azimuth[6].

Digital implementations of an 'alt-Blumlein shuffling' method used in this paper coincide with Gerzon's description of an 'alternative method of Blumlein shuffling' using first order low shelving filters.[9] The implementation includes the corrective filter on the *sum* channel in the range of 1–2 dB (microphone dependent). This maintains an averaged flat frequency response, and provides a small phase correction above the transition frequency.[9]

The Dirac impulse figures and audio demonstrations were prepared in the Digital Audio Workstation *Reaper*, using two instances of `JS: Mid/Side Encoder` and `VST: ReaEQ (Cockos)`.

2.4 Signal analysis

Signal transformations through each stage of an alt-Blumlein shuffling matrix can be illustrated with the use of sine bursts at stepped frequencies. A MATLAB implementation has been used for the plots in Figure 3: the Left channel signal is fed to the Right input with a delay and level adjustment.[13] This approximates the ILD and ITD offsets for a distant oblique source at azimuth $\theta = 50^O$.

Difference channel filter: gain + 20 dB; 1st order low shelving filter, corner frequency 500 Hz.
Sum channel filter: gain: −1.5 dB; 1st order low shelving filter with the frequency adjusted to provide an averaged flat response (see Gerzon[9]).

Result for 1000 Hz: the Left and Right input signals of our oblique source pass to the corresponding outputs with negligible amplitude gain. At 500 Hz, the Left output shows an amplitude gain, and the Right output has been reduced. For 250 Hz, the trend continues; there is a greater amplitude gain at the Left output, and further reduction at the Right.

At low frequencies, the shuffler has the effect of converting a time delay between the input channels into amplitude gains. The frequency-dependent amplitude gains correspond to consistent directional information for bass and treble frequencies of a reproduced source between the loudspeakers.

3   PRACTICAL APPLICATIONS

3.1 A near-coincident omnidirectional microphone configuration

A search was made for a near-coincident microphone technique suitable for acoustic measurements, and for reproduction using earphones and loudspeakers. Omnidirectional microphones of measurement quality have useful qualities for recordings: neutral response, lack of proximity effect, and ability to represent reverberant spaces. Typically, they may not be truly omnidirectional at all frequencies; however, this may suit a purpose. In Macauley, Hartmann, and Rakerd (2010)[6] interaural level differences (ILD) are plotted as a function of source azimuth:

- for tones of 1000, 1500, and 2000 Hz their maximum ILDs occur between a source azimuth (incident angle) of 40–55 degrees[6, fig 1]
- at 1500 Hz the approximate difference between near and far ears at an azimuth of 40 degrees and beyond is $>$ 6 dB[6, fig 2]

A narrowing of an omnidirectional microphone's polar response at frequencies above 1000 Hz may provide appropriate inter-channel level differences. For a typical spherical response microphone with a capsule diameter of 20 mm, a near-coincident pair at $\pm$ 45–50 degrees (a 90–100 degree included angle) may provide useful frequency-dependent level offsets for the central area. However, for a source at an azimuth above approximately ± 25–30° an additional effect is needed.

A centred medial baffle disc of appropriate diameter and thickness can be used to provide a shadowing effect. The statistical software R[14] was used to generate the graph plots for various microphone spacings and baffle sizes (see Figure 4). Variations of microphone array using discs of acoustic foam (without a rigid centre) were first optimized in a hemi-anechoic chamber using Time Delay Spectrometry (EASERA[15]), and then deployed for recording projects (see Figure 5). Acoustic measurements verified appropriate levels of frequency-dependent inter-channel level differences; and, the absence of unwanted colouration that may be caused by the baffle.
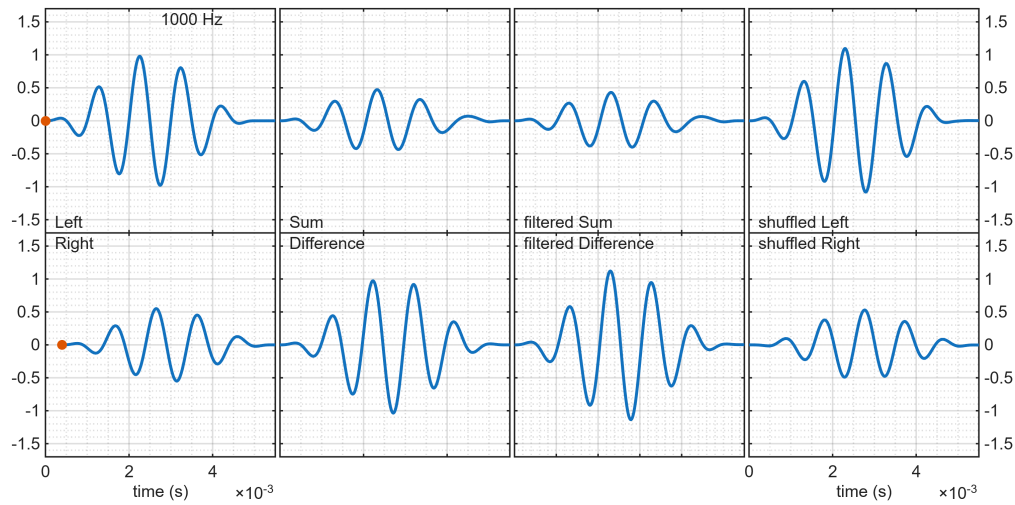
(a) 1000 Hz raised-cosine envelope sine burst; $\theta = 50°$ (R) = (L) $- 5$ dB, 0.4 ms delay
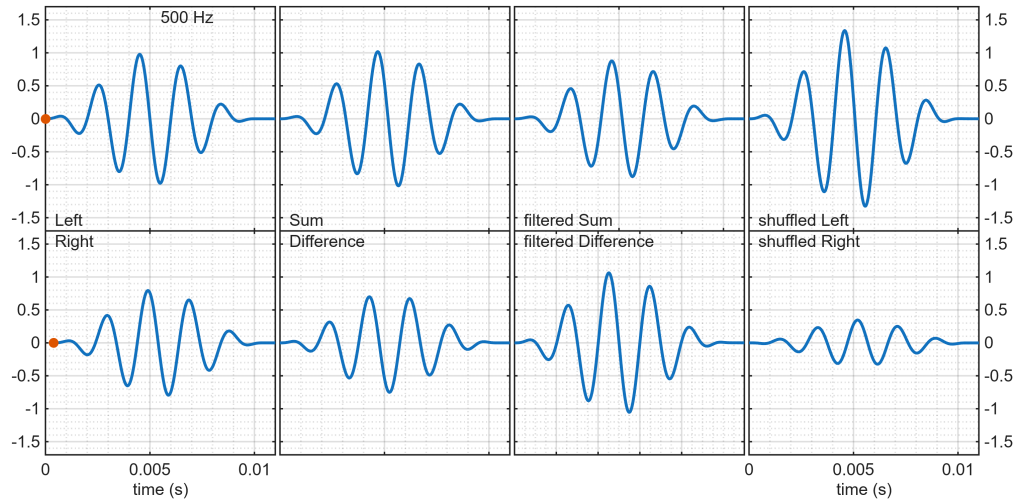
(b) 500 Hz raised-cosine envelope sine burst; $\theta = 50°$ (R) = (L) -1.8 dB, 0.4 ms delay
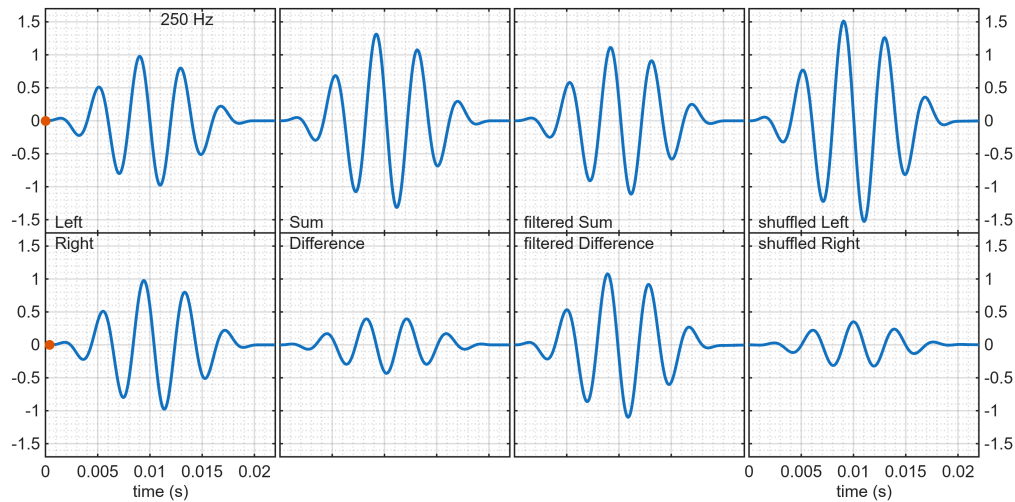
(c) 250 Hz raised-cosine envelope sine burst; $\theta = 50°$ (R) = (L) -0 dB, 0.4 ms delay

Figure 3: Sine bursts, raised-cosine envelope ; (R) input uses ITD / ILD offsets for a distant source $\theta = 50°$
Filtered *Difference*: 500 Hz, low-shelf, $+20$ dB; Filtered *Sum*: offset frequency, low-shelf, -1.5 dB
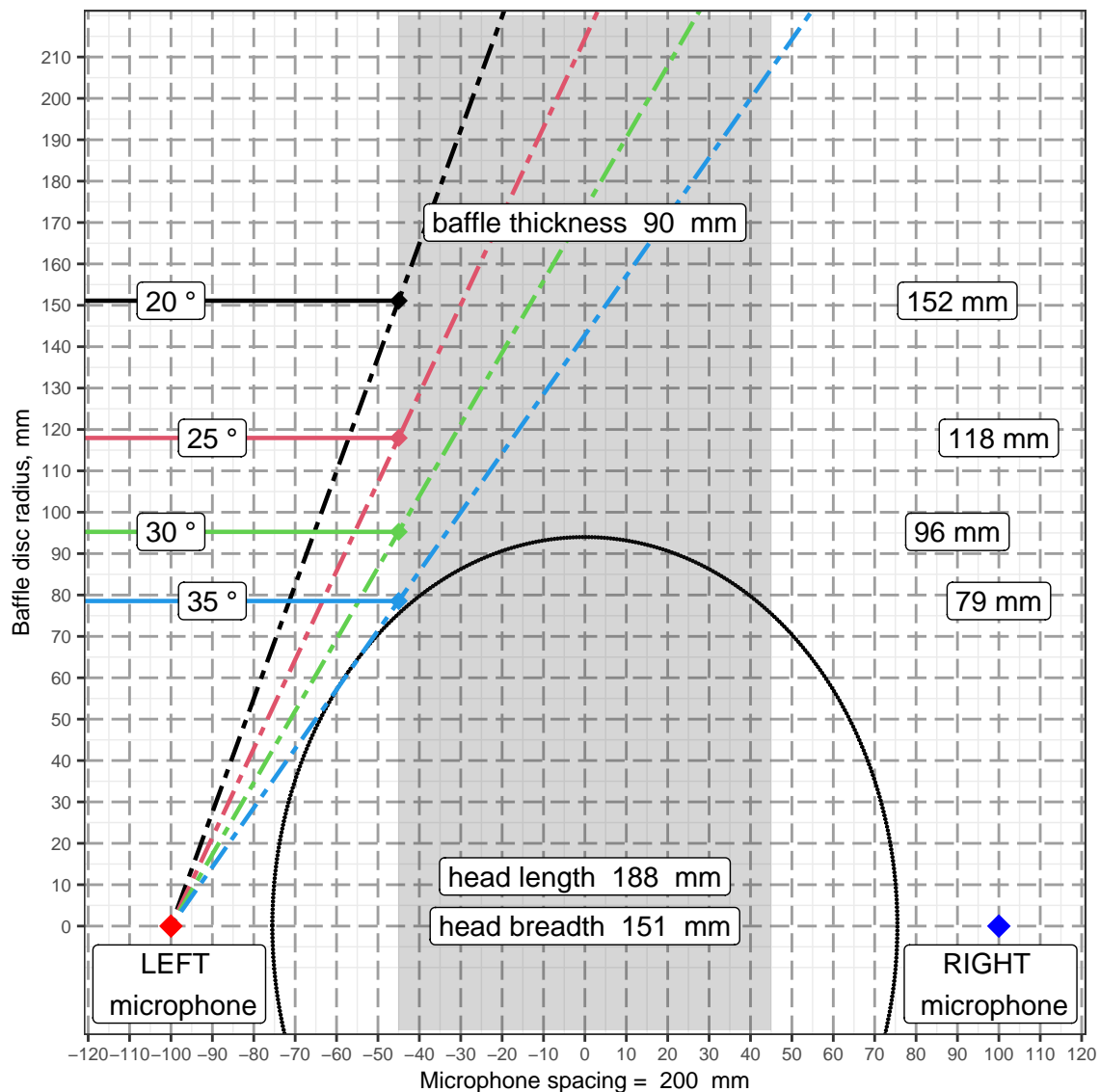Timescales (*x*-axis) adjusted to match signal frequency.

Figure 4: Plot representing a baffle disc positioned vertically between two near-coincident microphones (plan view). Azimuths represent distant oblique sources directed toward the LEFT receiver (zero elevation). A baffle disc radius(y) is indicated for each interference angle. Variables: receiver spacing, baffle thickness.

For recordings with the spherical response microphones (SCHOEPS MK2, matched pair) a subjective preference was found at an included angle of $\pm 45°$. This agrees with Macaulay *et al.* above; also, with Head Related Transfer Function (HRTF) measures at 2066 and 3962 Hz found in Braren and Fels (2020).[19]

Audio demonstrations are available at this URL:
https://digbyphonic.com/research/rs2025/RS2025supplements.html

Short excerpts from a selection of live recordings are provided: an unprocessed version, and an alt-Blumlein shuffled version. To emulate Blumlein's original demonstration, the listener may first audition the unprocessed version using earphones: a realistic spatial representation is expected. Then, using a pair of loudspeakers in the standard 'stereo format', the listener is encouraged to

(a) SCHOEPS MK2[16] spherical response microphone, ー45°; a baffle disc of acoustic foam[17]

(b) Forward-facing Sennheiser 8030[18] figure-8's ('phased array'), and SCHOEPS MK2 ±45° (baffled)

Figure 5: Example near-coincident microphone arrays used in the investigation

first audition the unprocessed version, followed by the shuffled version.[20] A consistent directional cue for the full-frequency range of a sound source is immediately recognized: the binaural image reproduced by the loudspeakers has been 'unshuffled'. It is not necessary to listen to each excerpt on earphones.

3.2 Reciprocity: Inverse shuffling for earphone reproduction of loudspeaker material

Unequal effects and conditions exist between 'stereo' loudspeaker and earphone reproduction. These include: interaural crosstalk, spectral balance, echoic vs. anechoic, externalized localization vs. in-the-head, and so forth.[21]

Two-channel stereo program material optimized for loudspeakers may be processed for earphone reproduction using an inverse alt-Blumlein shuffler. The low-frequency inter-channel *amplitude* differences are converted into *timing* differences. Therefore, loudspeaker-optimized content is somewhat modified into a binaural earphone presentation of the 'stereo' input. In this instance, the *difference* filter gain is inverted to ー20 dB, and *sum* filter gain ＋1.6 dB.

A best-fit approach is used for the *difference* filter's corner frequency, with the following criteria: the linear ＋20 dB/decade slope of the low shelving filter response should be established from 500 Hz and below; above 1200 Hz there should be ＜ 1 dB change in level. The corner frequency of the *sum* filter should be chosen to match the response of the *difference* filter at the same magnitude – this stipulation applies to both normal and inverse Blumlein shuffling.[9]

A selection of commercially available two-channel stereo program snippets are provided in processed and unprocessed pairs at the above URL. This includes a template to replicate results using *Reaper*. The demonstrations for this section may be carried out using earphones, only. The listener is encouraged to switch between the two versions, and to indentify the original (if known) or preferred. A list is provided to identify the A and B version for each pair, along with details of the recordings.

The expected effect may be described, again, as an unshuffling of the sound stage. A physical sensation in the ears may be experienced. The process is idealized for loudspeaker optimized two-channel 'stereo' program; especially that which incorporates amplitude panning of multiple microphones or direct sources. However, the process also performs well with classical or jazz recordings using a minimal number of microphones to capture a live performance.

A primary aim of post-processing of program using so-called 'binauralization' is to externalize the earphone presentation: to rectify an 'in-the-head' perception. Spatialization is commonly achieved with the introduction of new signals to augment the original presentation (virtual reflections); however, a not ideal impact on timbral quality may be inevitable.[22] In comparison, inverse alt-Blumlein shuffling is relatively benign.

This is an exploratory investigation at this stage; therefore, formal listening experiments have not been carried out. It is of interest to know if other's experience of this process is improved by a personalized adjustment of the *difference* channel's corner frequency. The author's preferred filter setting may be related to head dimension (HRTF).[7] However, a general fit for this method within the criteria of duplex theory (400–600 Hz) may be sufficient.[5]

4   DISCUSSION

4.1 An absorbent baffle's influence on arrival times

Gerzon notes an expected increase in the 'effective acoustic separation at low frequencies' for a central baffle. Measurements recorded in the hemi-anechoic chamber were used to verify frequency-dependent inter-microphone arrival time variations. Octave filtered energy time curves (ETC) did not show an appreciable difference in arrival times for the absorbent baffles of different thickness and diameter.[17] This may be due to the baffle's non-reflectivity. A future investigation is planned using Left/Right comparisons of Gauss or raised-cosine windowed sine bursts at suitable frequency intervals. It may be useful to measure a rigid baffle for comparison purposes.

4.2 Compatibility

Principal features in the effectiveness of Blumlein shuffling technique applied to near-coincident microphones may restrict its usage: compatibility within the two-channel 'stereo' format.

There is a traditional need for two-channel 'stereo' program to conform to physical limitations in analogue media. For example: a vinyl disc-cutting engineer needs to make decisions for the physical medium, alongside the artist and producer's original intent.[23] In part, the restrictions placed on the inter-related channels of a vinyl disc contribute to a principal legacy requirement of 'stereo' media: mono compatibility. This coincides with, or maybe a feature of, 'stereo' media's need to transfer equally well within any sound reproduction situation. Digital storage media removes inter-channel dependencies between the two channels; but, the legacy restrictions and limitations remain.

The Blumlein shuffling technique with a near-coincident microphone pair has poor mono compatibility. However, its effectiveness highlights the possibilities of allowing two-channel 'stereo' to be experienced as Blumlein intended: as inter-related channels to be signal-processed and reproduced as a whole.

5   CONCLUSIONS

Blumlein shuffling of a near-coincident microphone pair provides a realistic binaural effect with two-channel loudspeaker reproduction. A simple digital implementation of an alt-Blumlein shuffler can provide impressive results; using first order low shelving filters within the sum-and-difference network.

In addition, an inverse Blumlein shuffling process can provide subjective localization improvements for earphone reproduction of two-channel media optimized for loudspeakers (primarily, commercially available two-channel 'stereo' program material). A simple digital implementation of an inverse alt-Blumlein shuffling process may be included in personal music players and portable devices, or applied to in-ear stage monitoring.

The effectiveness of Blumlein shuffling highlights the dichotomy between opposing sound reproduction formats: a pair of loudspeakers in a triangulated listening arrangement, and earphones.[20] Differing psychoacoustic evaluations between these two situations result in a fundamentally unequal presentation of the original work. In other words, by encoding and optimizing content for two-channel 'stereo' loudspeakers there is a serious degradation with earphones for the identical content, and vice versa – a 'stereo compromise'.

With few exceptions, existing commercially available 'stereo' program material may be acknowledged as optimized for loudspeaker reproduction.[21] A possible solution is for two-channel 'stereo' media to be produced and mastered for differentiated playback: a *two-channel loudspeaker version*, and *earphones version*. It is not uncommon for music albums to be released, and re-released, in multiple versions; therefore, this could present further commercial opportunities.

Metadata in digital media could be used by listening devices to automatically choose an appropriate playback version. This would provide a predictable forecast of the listener experience, and increased artistic freedom, for the original artists and producers. The metadata solution is equally applicable to differentiated earphone and loudspeaker versions of broadcast media in general (movies, television programmes, podcasts, etc.).

The artistic and creative potential that exists in Blumlein's original concepts for two-channel binaural transmission may yet be fully appreciated.

6 REFERENCES

1. Blumlein, A. D. 394,325: Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems. (1933).
2. REAPER - Digital Audio Workstation. cockos Incorporated (2025).
3. Burns, R. W. *The Life and Times of A D Blumlein*. (Institution of Electrical Engineers in association with the Science Museum, London, 2000).
4. Blumlein, A. D. Binaural Reproduction - typewritten memorandum to Mr Isaac Shoenberg. (1932).
5. Hartmann, W. M., Rakerd, B. & Crawford, Z. D. Localization of sound in rooms VI: Duplex theory. *The Journal of the Acoustical Society of America* **158**, 2048–2061 (2025).
6. Macaulay, E. J., Hartmann, W. M. & Rakerd, B. The acoustical bright spot and mislocalization of tones by human listeners. *The Journal of the Acoustical Society of America* **127**, 1440–1449 (2010).
7. Carlini, A., Bordeau, C. & Ambard, M. Auditory localization: a comprehensive practical review. *Front. Psychol.* **15**, 1408073 (2024).
8. Strutt (Lord Rayleigh), J. W. XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **13**, 214–232 (1907).
9. Gerzon, M. Applications of Blumlein Shuffing to Stereo Microphone Techniques. in *AES 93rd Convention* (San Francisco, 1992).
10. Gerzon, M. A. Signal Processing for Simulating Realistic Stereo Images. in *AES 93rd Convention* (San Francisco, 1992).
11. Culling, J. F. & Lavandier, M. Chapter 8 - Binaural Unmasking and Spatial Release from Masking. in *Binaural Hearing: With 93 Illustrations* (eds. Litovsky, R. Y., Goupell, M. J., Fay, R. R. & Popper, A. N.) vols Volume 73 (Springer, ASA Press, Cham, Switzerland, 2021).
12. Bowers, J. S. *The Subjective Effects of Interchannel Phase-Shifts on the Stereophonic Image Localisation of Narrowband Audio Signals*. https://downloads.bbc.co.uk/rd/pubs/reports/1975-28.pdf (1975).
13. MATLAB. The Mathworks, Inc. (2025).
14. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (2025).
15. EASERA - Electronic and Acoustic System Evaluation and Response Analysis. AFMG Technologies GmbH (2024).
16. SCHOEPS Mikrofone. MK 2 Omnidirectional Microphone Capsule. https://schoeps.de/en/products/colette/capsules/omnis/mk-2.html (2024).
17. EQ Acoustics. Square 60 cm Acoustic Foam Tile _ EQ Acoustics. *EQ Acoustics* https://eqacoustics.com/products/square-60-tile (2025).
18. Sennheiser electronic. MKH 8030 Microphone: Quick Guide. (2024).
19. Braren, H. S. & Fels, J. Head and Torso HRTF Computation. RWTH Aachen University https://doi.org/10.18154/RWTH-2020-06760 (2020).
20. European Broadcasting Union. *Tech 3276 Listening Conditions for the Assessment of Sound Programme Material: Monophonic and Two–Channel Stereophonic*. (EBU, Geneva, Switzerland, 1998).
21. Toole, F. E., Olive, Sean. & Welti, Todd. *Sound Reproduction : The Acoustics and Psychoacoustics of Loudspeakers, Rooms and Headphones.* (Taylor & Francis Group, Oxford, 2025).
22. Morell, P. A. & Lee, H. Binaural Mixing of Popular Music. in *Listening tests and case-studies* vol. 665 (Audio Engineering Society, Online, 2021).
23. Borwick, J. *Microphones: Technology and Technique.* (Focal Press, 1990).