

From Line Arrays to Mid/Side and Ambience

A personal journey toward realistic stereo reproduction

Ronald van der Meulen

Independent researcher and DIY audio enthusiast

The Netherlands

www.vandermill-audio.nl

Version 1.1

June 2026

"The loudspeakers serve as the first 'early reflections' of a phantom sound source whose direct sound we didn't hear. Because our brain is good at filling in the missing blanks, it infers where that phantom source must be and that inference is what we actually perceive."

— David Moulton

Brave New World: Loudspeakers to the Left
Moulton Laboratories (archived via the Internet Archive)

Contents

1. Introduction	1
2. Context and influences	1
3. The loudspeaker as a room tool, not just a sound source	3
4. Cleaning the first 20 milliseconds	4
5. Why the array platform mattered	5
6. When the room gets out of the way, stereo reveals its limitations.....	7
7. Mid/Side as a practical correction tool.....	8
7.1 Mid EQ: the phantom problem comes first	9
7.2 Side EQ: restoring body to the sides	9
7.3 Phase linearity matters here	11
7.4 Stereo bass and low-frequency Side energy	12
8. Ambience channels: support first, spectacle second.....	12
8.1 Stream A: support ambience for the crosstalk holes	12
8.2 Stream B: ambience for space, boundaries and envelopment.....	13
9. Why Lexicon, and why not convolution.....	14
10. Real-room compromises still matter.....	14
11. A note on honesty and limits	15
12. Closing thoughts	15
Appendix A – References and Influences	17

1. Introduction

This paper describes a playback chain that grew out of a fairly simple question: how far can stereo reproduction be improved in a normal living room when the loudspeaker, room and DSP are treated as one system rather than as separate problems?

The answer, at least in my room, did not come from a single breakthrough or a single technology. It came from a long chain of decisions, experiments, refinements and compromises. Some of those choices were deliberate from the start. Others were consequences of earlier choices that only became obvious once the room had been cleaned up far enough for new limitations to reveal themselves.

The playback chain described here spans roughly a decade. The first Mid/Side experiments, the first ambience channels and the first FIR filtering all date back to 2015. In 2016 the system reached its first clearly phase-coherent result at the listening position, and by late 2016 / early 2017 the broader framework was largely in place: a phase-conscious stereo playback chain built around FIR correction, Mid/Side processing and ambience support. What followed after that was mostly refinement rather than reinvention. The frequency-shaded array described here is one of the later loudspeaker refinements in that longer story, not the beginning of it.

I have always treated the system as an experimental playground, but not an aimless one. The broader concept was already taking shape before the arrays were even built: a low-diffraction full-range array system, FIR filtering to improve phase behaviour, damping at the first reflection points, and ambience support that would work with the recording rather than simply decorate it. Over the years I compared different forms of ambience, different approaches to crosstalk cancellation and different ways of balancing direct sound against room contribution. Some ideas were discarded. Others survived and gradually became part of a stable framework that I still depend on today.

This is not a prescription for how everyone should listen to stereo. It is simply an honest account of what I ended up building, why I built it that way, and what I learned along the way.

2. Context and influences

Like many DIY projects, this one grew through a mixture of curiosity, stubbornness and a willingness to learn from people who had already thought deeply about related problems. But while the project was certainly influenced by the work of others, the implementation described here is not a borrowed recipe assembled from existing parts. Most of the system was developed within the project itself over many years, by trial, simulation, measurement and listening.

The loudspeakers themselves are my own designs. That applies not only to the later frequency-shaded arrays, but also to the earlier full-range line arrays with their deliberately

low-diffraction cabinet shape, and to the active subwoofers that were later integrated into the system. The Mid/Side EQ work was also developed within the project itself over a long period of listening and refinement, and the ambience channels were already functioning as part of the playback chain years before the shaded-array concept was added as a later loudspeaker refinement.

So while this project clearly sits in a wider tradition of ideas about diffraction, directivity, phase behaviour, spatial perception and room interaction, it would be misleading to describe it as a simple collection of borrowed concepts. The practical implementation, and many of the specific solutions, are very much my own. What others mainly provided was not a ready-made answer, but perspective: useful ways of thinking about the problems, and in some cases the confidence that a particular line of experimentation might be worth pursuing.

Several people influenced the direction of the project, each in a different way.

David Griesinger's work became especially important once the system started to move beyond loudspeaker building and into perception. His writing on phase relationships, localisation, the role of early sound and the distinction between a more direct, urgent stream and a more diffuse background stream gave language to things I had already been hearing but had not yet connected as clearly. In hindsight, some of the low-frequency Side-energy work and the later ambience strategy make much more sense through Griesinger's lens than they did when I first stumbled onto them by ear.

Earl Geddes was one of the people who made me pay attention to diffraction behaviour and the importance of avoiding loudspeaker shapes that keep reminding you where the boxes are. He also helped shape the broader idea that a loudspeaker should not be judged only on-axis, but by how it behaves as a system in a room.

Floyd Toole and Sean Olive provided useful information on listener preference and directivity behaviour. I do not see eye to eye with all of Toole's published thinking about rooms, and my own room strategy deliberately pushes farther toward a reflection-light early window than that work would normally encourage. But their work was still useful as a reference point, especially in reinforcing the value of controlling off-axis behaviour and optimising directivity rather than treating it as an afterthought.

Tom Danley and John Dunlavy are part of the background too. If I had not ended up with arrays, Danley's horn solutions or Dunlavy's phase-coherent MTM thinking would probably have been the directions I would have explored more deeply.

Roger Russell also deserves a quiet mention. His full-range line-array articles in *AudioXpress* in 2005 and 2006 were among the early pieces that made me look seriously at arrays as a domestic playback solution. That line of thought later connected more directly with David Smith's work on arrays during his time at McIntosh.

David Smith's work on expanding arrays is another obvious point of reference. My own frequency-shaded arrays resemble his expanding arrays in some ways, although the design

goal here was somewhat different: to create a frequency-shaded array that remains phase coherent across the band it is asked to cover, both on its own and when combined with the subs. Edgar Choueiri's work also served as a source of inspiration during a long period of experiments around stereo crosstalk and image formation.

The DIY community mattered as much as any individual paper or article. Forums were not just a source of information; they were a place to test ideas, get useful criticism, bounce possibilities around and have my own assumptions challenged. That grounding was important. It kept the project from drifting too far into theory without contact with reality.

3. The loudspeaker as a room tool, not just a sound source

One of the practical advantages of DIY is that a loudspeaker does not have to be designed for every room. It can be designed for your room, or at least with your room in mind. A commercial loudspeaker has to survive in a very wide range of spaces and placements. A DIY project can be much more opportunistic. It can choose a topology that suits the geometry, the listening distance, the room limitations and the intended balance between direct and reflected sound.

That was a large part of the appeal of the array concept here.

The original loudspeakers were designed in 2011/2012 and built in 2013/2014. From the beginning, the cabinets were deliberately given a rounded, low-diffraction shape because I did not want the loudspeaker itself to keep advertising its physical outline once the image started detaching from the boxes. In other words, low diffraction behaviour was not a late refinement. It was part of the original loudspeaker concept from the start.

The later frequency-shaded array work was not about tonal balance or output capability. The major goal of the shading work was to optimise vertical behaviour as well as I could, so that floor and ceiling contributions would be reduced even further than with the full-range array I was using before. A closely related aim was to create a directivity index that rises as gradually and smoothly as possible across the working band. In other words, the shaded array was designed to act as part of the room-control strategy, not as a separate loudspeaker exercise. That links indirectly to directivity-index thinking and some of the listener-preference work of Toole and Olive: not because there is a single target to chase, but because controlling where energy goes is often more useful than endlessly correcting the energy after the fact.

Line arrays often have a reputation for sounding larger than life, with voices and images appearing oversized rather than naturally scaled. I heard that too before the system was fully under control. In my case, though, it turned out not to be an inherent feature of the array at all. Once room treatment and tuning were working together properly, that character disappeared entirely and the presentation returned to a much more natural scale. I would therefore treat the familiar "large array sound" less as a fixed property of the topology and more as something that depends heavily on implementation. Simple EQ in an untreated

room can easily cause an array to present itself as larger than life; a more elaborate approach can change that completely.

Another important practical point is diffraction. If the goal is to let the image detach from the loudspeaker positions, then the speaker should not keep drawing attention to its physical outline. Rounded shapes and low diffraction behaviour matter here. That is one of those details that is easy to dismiss in isolation but becomes more important once several other parts of the system are already behaving well.

4. Cleaning the first 20 milliseconds

From the outset, the project was not only about building an interesting loudspeaker but about cleaning up the first part of what reaches the listener's ears. That objective shaped the design long before the system reached its present form, and a great deal of time went into developing the ideas needed to get there.

I had set myself the goal of reducing the influence of the room during roughly the first 20 ms after the initial arrival. At the time I was somewhat naive about where that would lead. I understood that a stronger direct field and fewer early reflections would probably help clarity and imaging. I did not yet understand how much it would expose the limitations of stereo itself once the room stopped masking them.

The loudspeaker was one part of that work. The room still had to cooperate.

Three large absorbing panels were needed to tame the first reflection points well enough for the rest of the idea to work. By February 2015 all three panels were in place. That was essential. Without those panels, the room would simply have remained too active in the first part of the response for the later phase-coherent and crosstalk-related work to matter in the same way. Once the panels were in place and the loudspeaker behaviour was under control, the direct sound became much more dominant at the listening position.

That result is visible in the measured decay plot shown in Figure 1. It is a graph that tends to be misunderstood, partly because it is unusual to see this kind of behaviour in a normal living room rather than a studio or a heavily treated dedicated space. It does not mean the room has vanished, nor that the system is somehow reflection-free. It simply shows that the direct arrival and its subsequent decay are allowed to dominate the first part of the response much more clearly than before.

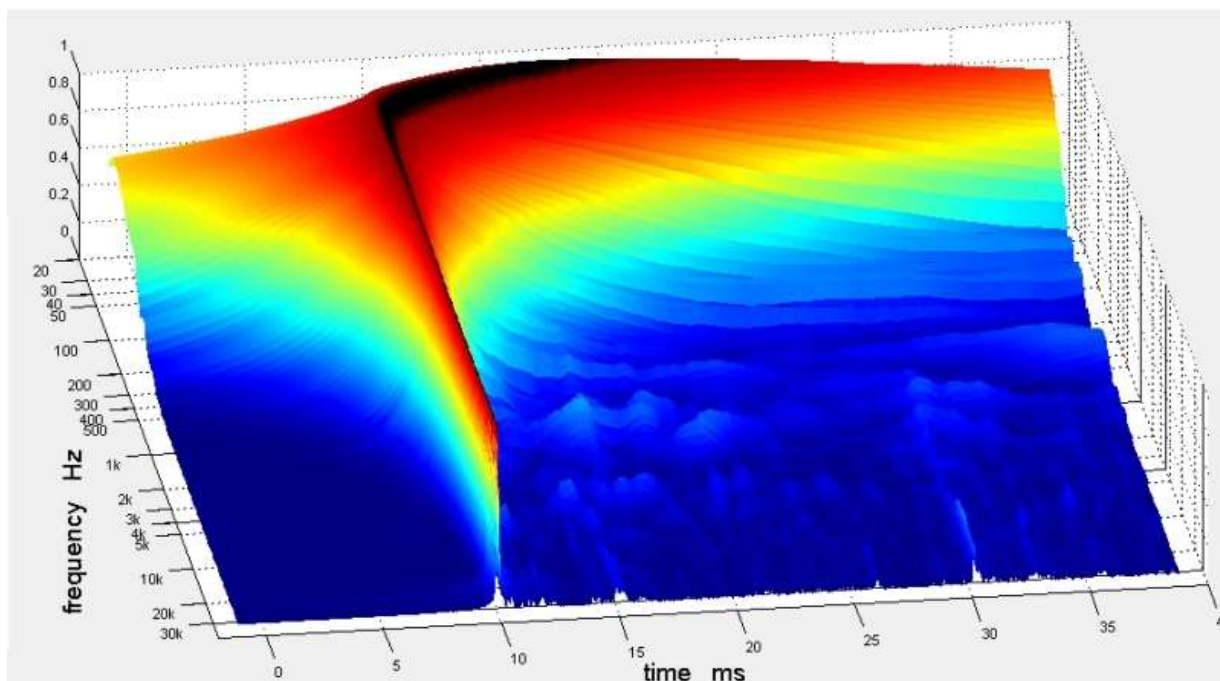


Figure 1 — Measured cumulative spectral decay / early-time behaviour at the listening position

The important consequence was not simply “better imaging.” Once the room had been pushed back far enough, the playback chain stopped being dominated by room clutter and started revealing other weaknesses much more clearly. One of those weaknesses was stereo crosstalk.

I should be clear about one thing here: the crosstalk problem I later started addressing at this level of intensity is largely a problem of my own making. Most people allow the room to remain part of the experience, either by preference or because it is the practical reality of a domestic system. In such systems, floor, ceiling and wall reflections continue to blur and partly mask some of the tonal and spatial damage caused by crosstalk. By deliberately reducing the room’s influence, I removed part of that mask. That was a choice. I do not regret it, but it is not a universal default and it should not be presented as one.

The sequence, in hindsight, was simple: clean up the first 20 ms, and stereo crosstalk becomes much harder to ignore.

5. Why the array platform mattered

The array was helpful here not because arrays are magical, but because it made the room-control part of the problem easier to tackle. Compared with a more conventional single-driver source in the same geometry, the shaded array keeps the direct sound more

dominant in the first milliseconds and reduces the immediate influence of floor and ceiling reflections.

That is the point of the comparison shown in Figure 2.

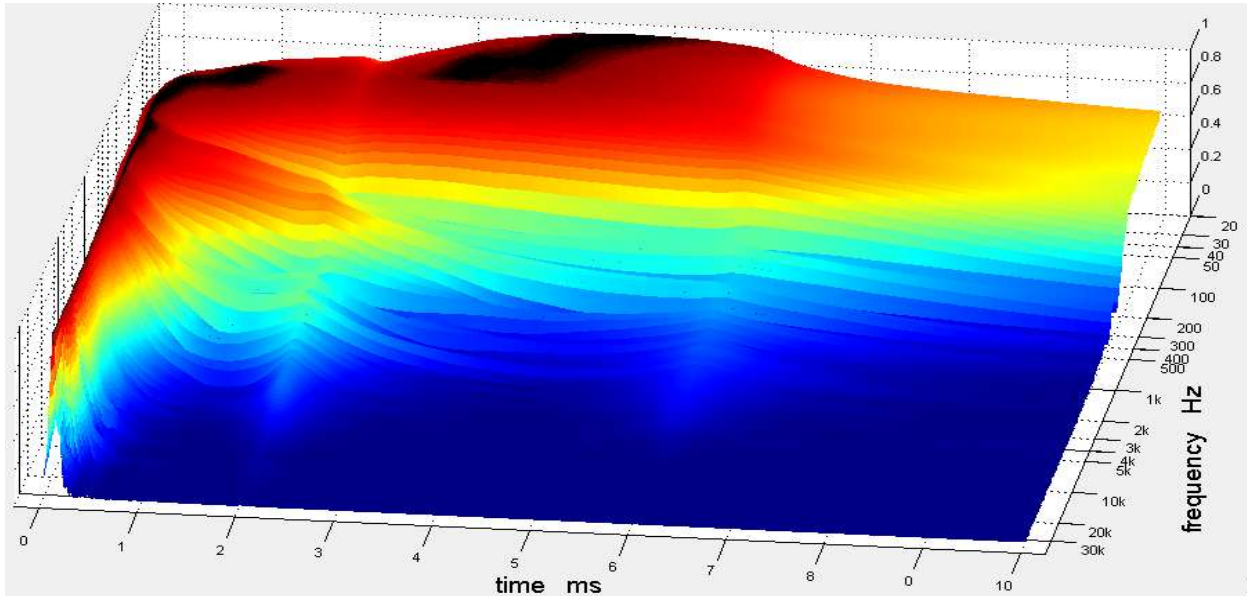


Figure 2a Single driver simulation with floor and ceiling reflections

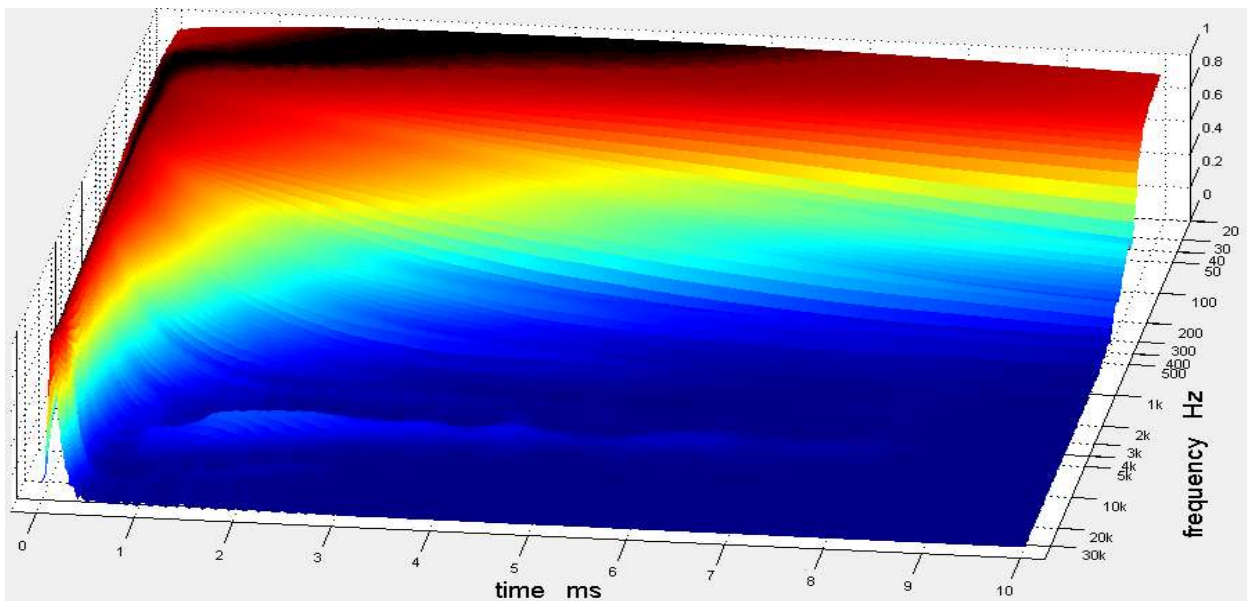


Figure 2b Shaded Array simulation with floor and ceiling reflections

Figure 2 — Simulated early-time behaviour at the listening position for a single-driver source and a shaded line array under similar floor and ceiling reflection conditions

The purpose of the comparison is not to argue that one topology is universally “better.” It is simply to show why the array was a useful platform for this project. Its narrower vertical behaviour and frequency shading helped reduce early vertical room interaction, which in turn made it more realistic to pursue a stronger direct field in a normal living room. The large radiating area also helps: with 25 drivers sharing the work, excursion demands remain low for a given SPL, which likely contributes to the sense of cleanliness. That cleaner early window was an important prerequisite for everything that followed.

6. When the room gets out of the way, stereo reveals its limitations

The more phase-coherent and reflection-light the direct field became, the more obvious it was that conventional stereo does not preserve tonal balance and spatial stability as gracefully as we often pretend. A phantom image between the speakers is not simply a centre loudspeaker that happens to be invisible. It is a psychoacoustic construction assembled in the presence of crosstalk, level differences, arrival-time differences and room contributions. Once the room contribution is reduced, the artefacts of that construction become easier to perceive.

The first thing that stood out was a lack of depth in the presentation, but tonal consistency was just as problematic. If a centrally panned voice had the right weight and body, a similar voice panned hard left or right could lose some of that body and start sounding lighter, thinner or brighter than it should. Background vocals could become unnaturally lean. The phantom voices were not without problems either. Instruments moving off centre would not simply move laterally; they would change character as they moved.

This is where the Mid/Side work started. Not because I wanted to add a spatial effect, but because the system had become transparent enough to reveal a tonal and perceptual problem that I no longer wanted to ignore.

That tonal problem was only part of the story, though. The larger aim was always to make stereo imaging and staging feel as realistic and believable as possible within the limits of two-channel playback. That was also the reason for chasing phase correctness of the first wavefront reaching the ears in the first place. Once that first arrival became cleaner and less masked by early room clutter, the stereo illusion itself became easier to judge, but also easier to break. In my experience, tonality is central to that judgement. If a voice or instrument changes character as it moves across the stage, the illusion of a stable event collapses. But once tonal consistency is restored, the spatial side of the presentation improves as well: phantom images become more solid and palpable, depth becomes more convincing, and the stage is less likely to remain trapped as a flat left-right construction between the loudspeakers.

That is also why I spent time experimenting with more explicit anti-crosstalk approaches over the years. They can be very effective at increasing spatial realism and image

specificity, but in practice they are often tied too tightly to a narrow centre listening position. The Mid/Side EQ work proved more useful here because it captures part of that effect while remaining more tolerant, more tonally integrated and ultimately more practical in daily use. In that sense, the Mid/Side work mattered to me precisely because it affected both at once. It was a tonal correction, but one with clear consequences for the realism of the stereo event itself.

7. Mid/Side as a practical correction tool

The Mid/Side section of the playback chain is often the part that attracts the most curiosity, but in practice it was not designed as a “trick” to widen the stage or throw sounds around. It began as a way of restoring tonal continuity and image stability once crosstalk had become impossible to unhear.

The main stereo signal is first encoded into Mid and Side, processed separately and then decoded back to left/right before the final main EQ stage. In itself, that is not unusual. The unusual part is the reason for the EQ choices and the way they were arrived at. Before looking at the individual parts of the processing, it may be helpful to look at the signal flow of the playback chain as a whole.

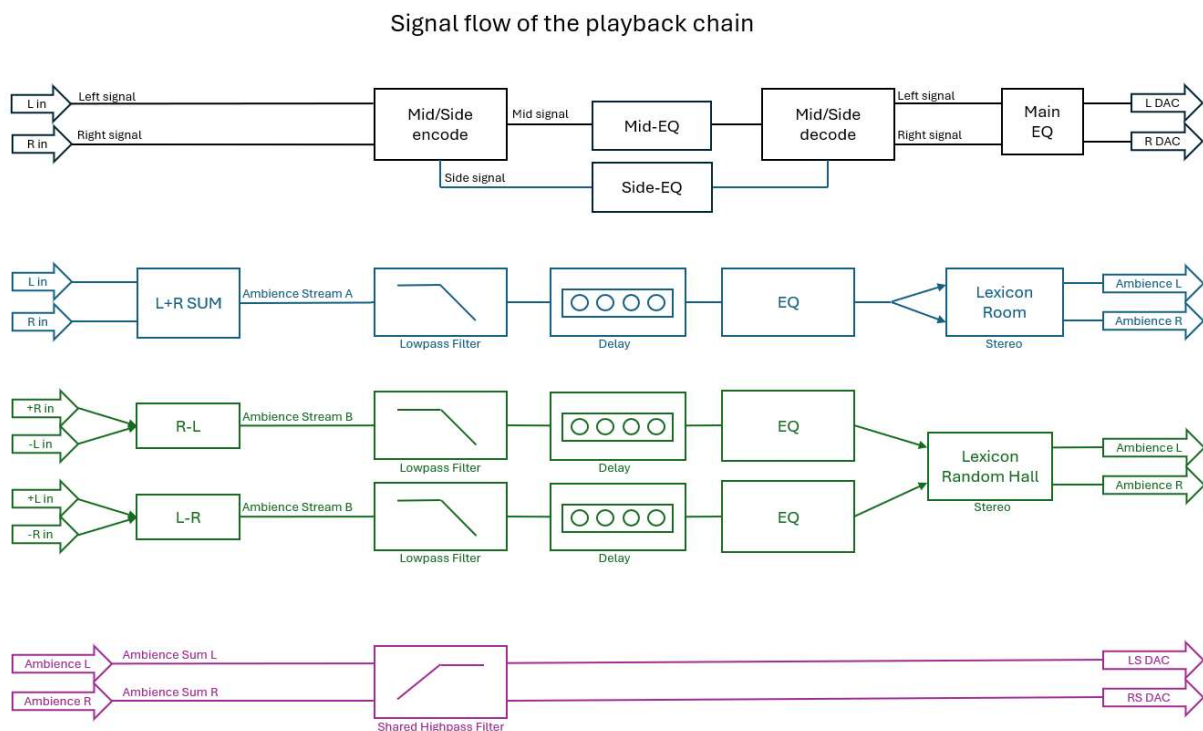


Figure 3 — Simplified signal flow of the playback chain

7.1 Mid EQ: the phantom problem comes first

It is important to note that Mid EQ is not isolated from the rest of the stereo picture. In an M/S representation, the Mid stream still contains information that also belongs to left and right events, because it is the sum component of the stereo signal. When Mid and Side are decoded back to left/right, the two streams fuse again. That means Mid EQ is not some private correction applied only to phantom-centre content. It inevitably influences the tonal result across the stage, and part of why it works as well as it does is precisely because it is not completely separate from the side information.

The practical goal of the Mid EQ was simple: restore tonal continuity and solidity to the phantom presentation, so that centrally placed voices and instruments no longer felt like a special case with their own spectral signature.

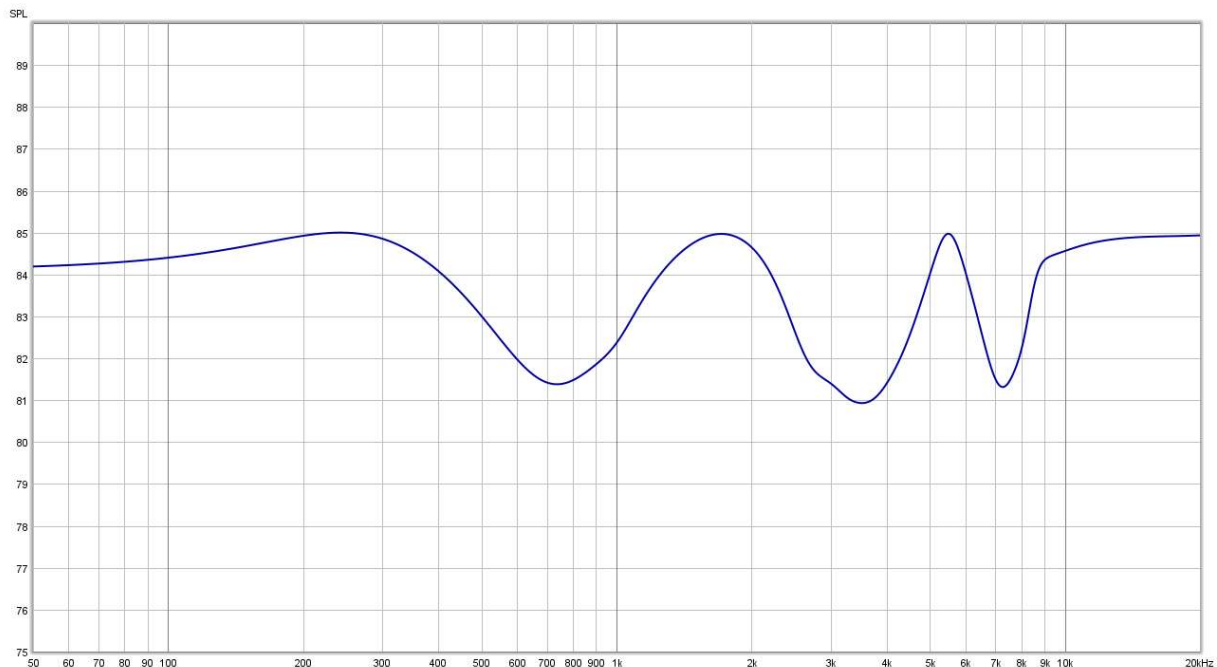


Figure 4 — Illustrative Mid EQ contour

The amount of Mid correction shown here is specific to my system. Because the loudspeakers, room treatment and phase-linear processing produce a strong, coherent direct field, the summed Mid component approaches the theoretical behaviour of two coherent sources more closely than is typical in domestic listening rooms. The graph should therefore be read as an illustration of the correction strategy rather than a transferable EQ preset.

7.2 Side EQ: restoring body to the sides

The Side EQ came from a different observation. Once the phantom tonality had been improved, laterally panned voices and instruments still tended to sound too lean and too

bright compared with the phantom equivalent. That was the real reason for the Side shaping.

In broad terms, the Side channel needed more weight at lower frequencies and less energy toward the top, so that panned vocals would keep the body they should have rather than turning into a thinner, more “chipmunk-like” version of the same singer. The goal was not to make the soundstage wider for its own sake. It was to restore believable tonality to left and right events when compared with the phantom centre.

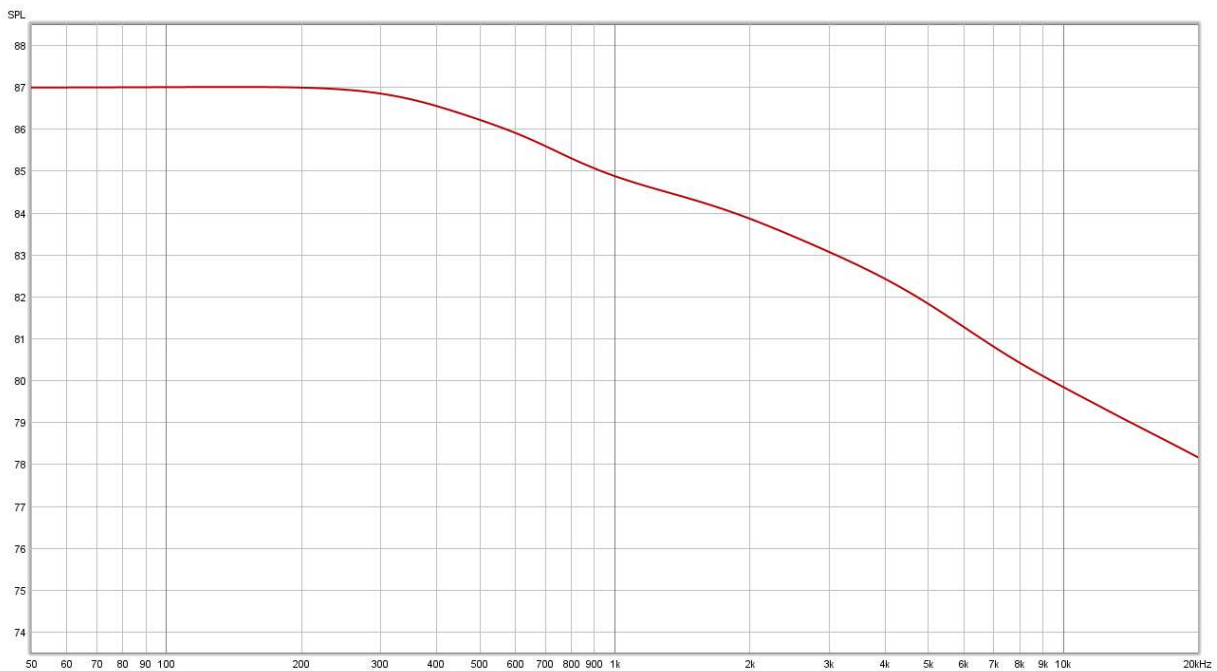


Figure 5 — Illustrative Side EQ contour

That tonal correction does have a useful side effect. Because Side processing changes how the left and right channels recombine, it can create a partial negative contribution in the opposite channel — effectively a mild form of anti-crosstalk. Used carefully, that improves image specificity and can make the stage project more convincingly in front of the speaker plane rather than remaining glued between the cabinets. Used too aggressively, it falls apart and starts sounding artificial. The important point is that the tonal correction and the imaging result are not really separate in listening. The Side EQ was developed because the tonal mismatch was unacceptable, but once that mismatch was reduced it also became clear that the same correction could make the stereo stage behave more convincingly as a stable event in front of the listener rather than as a flatter left-right construction between the loudspeakers.

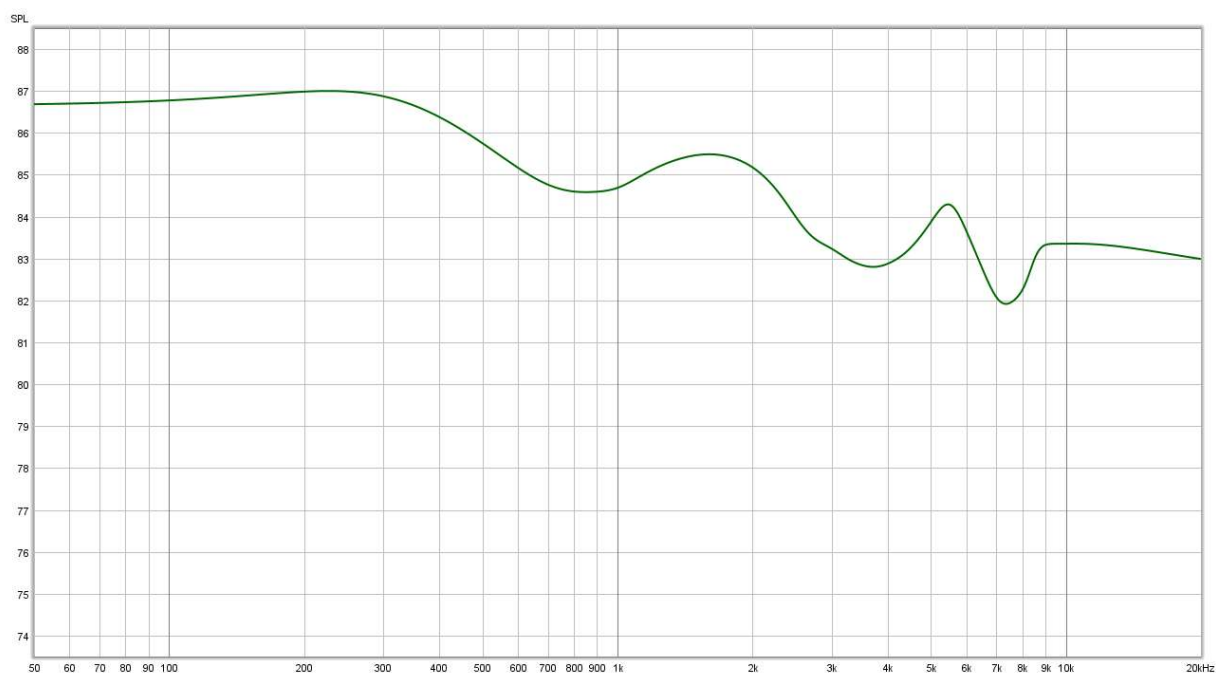


Figure 6 — Effective lateral response after Mid/Side recombination

The previous two curves should not be interpreted independently. The listener never hears the Mid and Side channels in isolation; they are recombined during decoding. The effective lateral response shown here illustrates how the contributions of both paths combine in practice and why the explicit Side EQ curve, when viewed on its own, can appear more extreme than the final perceptual result.

Once those tonal faults and discontinuities were reduced, the spatial change felt natural, and together with the restored phantom images the presentation became much more holographic. In a way, I believe that removing those tonal faults also freed up the natural space stereo is capable of.

7.3 Phase linearity matters here

This Mid/Side approach only behaves as intended when it is implemented in a phase-linear way. That matters not just for neat signal processing, but because the entire project was aimed at preserving the realism of the first wave of sound reaching the ears. If the phase relationships through the processing chain are not preserved, the tonal and spatial interactions between Mid and Side do not recombine cleanly and the illusion becomes less stable. That is part of why the earlier work on loudspeaker phase behaviour, room treatment and time-domain cleanliness matters so much: the M/S correction is not floating in isolation, but sits on top of an already phase-conscious playback chain.

If you look carefully at the impulse response of a corrected phantom-centre stream, you can often see side features around the main pulse that resemble a mild anti-crosstalk pattern, with energy appearing slightly before and after the dominant arrival.

7.4 Stereo bass and low-frequency Side energy

One of the more interesting side roads in this process was the discovery that some of the Side work at lower frequencies can have a very strong perceptual effect when the recording contains the right information. A controlled boost of low-frequency Side energy can make the bass wrap around the listener and contribute to a sense of envelopment rather than simply sitting between the speakers. I only later found that David Griesinger had written about very similar effects. That was one of those moments where a listening experiment suddenly found a theoretical home years after the fact.

It also reinforced the importance of preserving stereo bass deeper into the low frequencies than many playback systems bother to do. The perceptual advantages are real when the recording carries that information.

8. Ambience channels: support first, spectacle second

Once the mains were behaving more coherently and the M/S processing had restored much of the tonal continuity across the stage, I started experimenting with ambience channels. These are not surround channels in the home-theatre sense, nor are they there to paint an artificial room over every recording. Their purpose is more modest and more conditional: to support what the mains are already doing, and to do so in a way that follows the recording rather than fighting it.

The ambience goal is twofold, and it is important to separate those two functions because they do different jobs.

8.1 Stream A: support ambience for the crosstalk holes

The first ambience stream exists mainly to support the direct stereo presentation where crosstalk has created perceptual holes. It is derived from an L+R support feed, delayed, band-limited and processed with Lexicon Room before being mixed into the left and right ambience channels at a level far below the mains.

The signal-flow figure in Figure 3 shows the simplified logic. The support stream is created from taps of the left and right channels, summed as an L+R feed with level management, and then processed separately from the main stereo path. In practice this can be implemented as a single L+R sum presented to both stereo inputs of the Lexicon Room processor. Even when both inputs carry the same material, the stereo algorithm still produces differing synthetic reflection streams over time, which is useful because the ambience channels themselves are also subject to normal stereo crosstalk at the listening position.

The role of Stream A is deliberately modest. It is not there to announce itself. It is there to partially fill the crosstalk-related gaps left by the mains, especially in the region where a little support can improve intelligibility and restore some ease to the presentation without sounding like an added effect. It uses Mid-derived information because that is where much of the “missing support” resides once the room has been pushed back and the phantom construction is left more exposed.

It sits well within the Haas limit, so it will not be noticed as a separate event. Add too much and it will start to sound like coloration, much like a room does. It can fill in the blanks because it arrives from lateral angles and is timed differently at the ears than the mains.

8.2 Stream B: ambience for space, boundaries and envelopment

The second ambience stream is different in purpose and in source material. It is not about filling crosstalk holes. It is about strengthening the sense of space and helping room boundaries, depth cues and envelopment emerge more clearly when the recording already contains that information.

This stream is derived from difference information: R–L and L–R feeds, processed separately and routed to the ambience outputs in a cross-coupled way. In the simplified signal-flow diagram, those difference-derived streams feed Lexicon Random Hall, band-passed and at a lower level than the mains, and are then combined with Stream A at the ambience outputs.

The cross-routing matters. The ambience contribution that ends up in the left ambience speaker is derived from R–L, and vice versa. This is one of those implementation details that looks minor on paper but changes the perceptual result. Because Stream B is later mixed together with Stream A in the ambience outputs, the polarity and channel assignment of the difference-derived stream affect how the final ambience field relates to the direct stereo presentation.

The reason for using a Lexicon algorithm here rather than an impulse-response room simulation is straightforward. A fixed IR tends to paint the same room over every recording. The Lexicon algorithms are more adaptive in the sense that they preserve and enhance the spatial cues already present in the music rather than imposing one recognisable room signature on top of everything. If a recording contains no useful ambience information, Stream B does not magically create a believable acoustic world out of nothing. If the recording does contain those cues, it can strengthen them in a way that feels connected to the source rather than detached from it.

That distinction matters in listening. Some recordings barely change. Others open up dramatically. **Malia & Boris Blank’s “Celestial Echo”** from the album *Convergence* is one example of a production that already contains multiple perceptual streams and can benefit strongly from this kind of support. **Steven Stills’ “Treetop Flyer”** from the album *Stills Alone* is a different kind of example: it creates a beautiful space in front of the listener, with guitar harmonics bouncing off walls, but never turns into an enveloping wrap-around

presentation. The ambience system should not try to force it to do so. It should stay true to what the recording is actually offering.

Older hard-panned stereo recordings are another useful reminder of the limits. A track like **“Helplessly Hoping”** by **Crosby, Stills & Nash** from the album *Crosby, Stills and Nash* remains largely glued to left, centre and right because that is how it was mixed. The ambience system may soften the edges slightly, but it does not rewrite the production.

9. Why Lexicon, and why not convolution

This deserves its own short section because it is easy to misunderstand the role of the ambience channels if the reverbs are treated as interchangeable effects.

An IR-based reverb can be excellent when the goal is to recreate a specific acoustic space. That was not the goal here. The aim was not to impose a room on the music, but to support the cues already present in the recording and let those cues breathe a little more in the listening room.

That is where the Lexicon algorithms proved more useful. They seem to preserve the “native” spatial character of the recording more gracefully while still decorrelating and enriching the support channels enough to make them perceptually useful. In Stream A, Lexicon Room works well as a decorrelating support mechanism for the Mid-derived hole-filling role. In Stream B, Lexicon Random Hall is better at strengthening spaciousness and room-boundary cues without sounding as though a fixed room has been stamped onto the music.

10. Real-room compromises still matter

It would be misleading to present the system as though it exists in a laboratory. It lives in a fully functional living room, and some of the choices are compromises imposed by that reality.

The ideal loudspeaker placement for stereo imaging is not necessarily the placement that looks acceptable in a domestic space, especially with tall line arrays. For the widest stable imaging across multiple seats, I would normally prefer the speakers to cross in front of the listening position. That would make stronger use of intensity trading and help listeners away from the sweet spot maintain a more favourable geometric relationship to the opposite speaker. In my room, and with the visual impact of the arrays, that toe-in looked too strange to live with. The practical compromise was to cross the speakers behind the sweet spot while using angles that still gave the best subjective result.

That compromise has consequences. Off-centre listeners do not get the same phantom stability as they would with a more aggressive toe-in strategy, and the phantom centre can

shift somewhat for those away from the hot seat. I do not want to hide that. It is part of the honesty of the project: some of the choices were guided as much by living-room reality as by theory.

11. A note on honesty and limits

By the time this system reached its present form, I had effectively chosen an extreme: a highly direct, phase-conscious stereo system in a relatively small, asymmetrical living room, with the room deliberately silenced as far as practical during the first part of the response. That choice produced a result I like very much, but it also made stereo's own flaws more exposed and therefore more difficult to ignore.

In that sense, I created part of the problem I later set out to solve.

I do not see that as a failure. It was simply the next step in a chain of cause and effect. Once the room became less dominant, crosstalk and phantom instability moved into the foreground. Once they did, Mid/Side correction became worth pursuing. Once the direct stereo image was repaired with Mid/Side EQ, it made sense to move on to the next step and support it with ambience channels that worked with the recording rather than against it.

The important point is that none of these steps should be taken as universal doctrine. They are solutions to a specific set of priorities in a specific room. For me, the journey was worth it, not only for the result but for the knowledge gained along the way. The system serves music much more like an event than a song. That is what this journey was all about, and still is.

12. Closing thoughts

If there is one lesson I would take from this project, it is that loudspeaker design, room treatment and signal processing are much more powerful when they are allowed to inform each other instead of being treated as separate hobbies. The line array was not “the answer” on its own. FIR processing was not “the answer” on its own. Absorption panels, Mid/Side EQ and ambience channels were not the answer on their own either. The result came from combining them in a way that gradually reduced one limitation, revealed the next, and then addressed that one too.

That process also reinforced something I value deeply about DIY audio. We are not limited to choosing between finished products that have to work reasonably well in thousands of unknown rooms. We can build systems that are shaped around the room we actually have, the listening habits we actually have, and the compromises we are actually willing to make. That does not guarantee a better result, but it does open doors that commercial design often cannot.

This project is simply one example of where that freedom can lead: from arrays, to room control, to phase coherence, to crosstalk correction, and finally to a playback chain that aims to preserve not only tonal accuracy and imaging, but also the sense that music can occupy more than one perceptual stream at once — a direct stream that feels urgent and intelligible, and a background stream that can wrap around you when the recording invites it.

I owe a great deal of that journey to the people who have thought about these subjects before me, and to the DIY community that gave me a place to learn, question, test and refine the ideas. If this paper gives anything back, I hope it is not a blueprint to copy blindly, but a useful account of one route through the maze.

Appendix A – References and Influences

The references below should not be interpreted as direct sources for every technique described in this paper. Rather, they represent the work that most strongly influenced the author's thinking about spatial hearing, loudspeaker behaviour and stereo reproduction over roughly a decade of experimentation.

1. Psychoacoustics, Spatial Perception and Hearing

- Blauert, J. (1983). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press.

Conceptual foundation for: ear-brain processing of phantom images, stereo crosstalk, and the role of interaural time and level differences (ITD/ILD).

- Griesinger, D. (1996). *Spaciousness and Envelopment in Musical Acoustics*. Lexicon Inc. / Audio Engineering Society paper.

Conceptual foundation for: listener envelopment and the auditory system's separation of the direct, urgent sound stream from the diffuse background stream.

- Griesinger, D. (1998). *Objective Measures of Spaciousness and Envelopment*. Audio Engineering Society Conference Paper.

Conceptual foundation for: the psychoacoustic importance of stereo bass and the specific perceptual benefits of low-frequency Side energy around the listener's head.

- Griesinger, D. (2010). *Phase Coherence as a Measure of Acoustic Quality*. Proceedings of the International Congress on Acoustics (ICA).

Conceptual foundation for: the rationale behind cleaning the first wavefront to preserve clarity, intelligibility and image solidity.

- Moulton, D. (c. 2001). *Brave New World: Loudspeakers to the Left*. Moulton Laboratories. Archived via the Internet Archive (Wayback Machine).

Conceptual influence for: viewing stereophonic reproduction as a perceptual inference rather than a direct acoustic event, and understanding phantom images as constructs formed by the auditory system from the signals arriving from the loudspeakers.

2. Loudspeaker Design, Directivity Control and Room Interaction

- Dunlavy, J. (1997). *Loudspeaker Accuracy*. Dunlavy Audio Labs, Inc.

Conceptual foundation for: defining audible accuracy through time-domain integrity, the importance of phase linearity, and the use of symmetrical driver arrangements to preserve transient behaviour while reducing the influence of early boundary reflections.

- Geddes, E. & Lee, L. (2002). *Premium Home Theater: Design and Construction*. GedLee Publishing.

Conceptual foundation for: controlling cabinet diffraction through rounded geometry and preventing the loudspeaker itself from drawing attention to its physical position.

- Russell, R. (November 2005; July 2006). Line Arrays for Home Audio. AudioXpress, two-part article series.

Conceptual foundation for: the practical feasibility, low-profile footprint and domestic applicability of full-range line arrays.

- Smith, D. L. (1997). Discrete-Element Line Arrays—Their Modeling and Optimization. *Journal of the Audio Engineering Society*, Vol. 45, No. 11, pp. 949–964.

Conceptual foundation for: the mathematical and technical basis of multiple-driver arrays, including amplitude-shading and frequency-shading techniques used to control vertical dispersion, reduce sidelobes and smooth response behaviour.

- Toole, F. E. (2008). *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*. Focal Press.

Conceptual foundation for: directivity index (DI), smooth off-axis behaviour and evaluating loudspeakers and rooms as a unified system.

3. Stereophonic Localisation, Crosstalk and Spatial Behaviour

- Blumlein, A. D. (1931/1933). British Patent 394,325: Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems.

Conceptual foundation for: Mid/Side theory, sum-and-difference processing and variable stereo width.

- Gerzon, M. A. (1986). Stereo Shuffling: New Approach – Old Technique. *Studio Sound*, July 1986.

Conceptual influence for: frequency-dependent Mid/Side processing, low-frequency Side enhancement to improve spaciousness and listener envelopment, the interaction between phase and localisation, and the importance of subtle, natural-sounding spatial enhancement that remains satisfying during long-term listening.

- Sengpiel, E. (undated). Frequenzabhängige Hörereignisrichtung bei Stereo-Lautsprecherlokalisation [Frequency-Dependent Apparent Direction in Stereo Loudspeaker Localisation]. Sengpiel Audio.

Conceptual foundation for: the frequency dependence of stereo localisation, the tendency of low frequencies to collapse toward the phantom centre and high frequencies to localise too far outward, and the potential benefits of increasing low-frequency Side energy to improve spaciousness and apparent source width.

- Choueiri, E. Y. (2008). Optimal Crosstalk Cancellation for Binaural Audio with Two Loudspeakers. Princeton University / Audio Engineering Society.

Conceptual foundation for: the analysis of phantom-centre tonality issues, spectral coloration caused by crosstalk and the artefacts that become exposed when the room contribution is reduced.

- Vickers, E. (2009). Fixing the Phantom Center: Diffusing Acoustical Crosstalk. Audio Engineering Society Convention 127.

Conceptual foundation for: the destructive comb-filtering effects inherent to phantom-centre imaging and the potential benefits of phase-decorrelated ambience and frequency-dependent Side energy in stabilising the central soundstage.

- Goldfarb, B. S. / BSG Technologies (2012–2013). qøl Signal Completion Stage and associated patents and reviews.

Conceptual influence for: the use of sum-and-difference signal processing to alter the perception of space, ambience and image palpability, and the idea that conventional stereo reproduction may leave some aspects of spatial information underrepresented.

About the Author

Ronald van der Meulen is a mechanical engineer by training and a lifelong DIY audio enthusiast. His interest in loudspeaker design, room acoustics and psychoacoustics led to a decade-long exploration of phase-conscious stereo reproduction, culminating in the playback chain described in this paper. Although developed outside academia and industry, the work draws heavily on published research and extensive practical experimentation.

Website: www.vandermill-audio.nl