# Acceptable Imbalance of Sound-Object Levels for Off-Center Listeners in Immersive Sound Reinforcement

Franz Zotter, Stefan Riedel, Lukas Gölles, Matthias Frank

*Institute of Electronic Music and Acoustics, Universit of Music and Performing Arts, 8010 Graz, Austria, Email: zotter@iem.at*

## Introduction

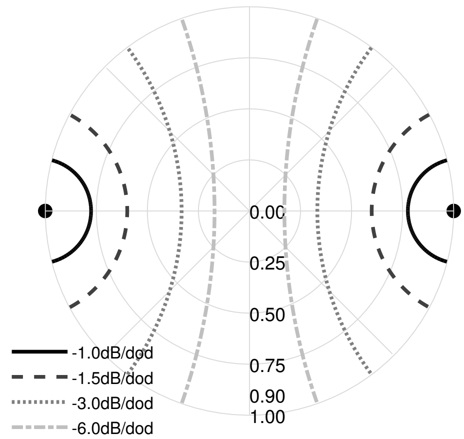
For a balanced immersive sound reproduction, ISO 2969 [1] and Dolby ATMOS specifications define a rectangular area of $\pm^1/_5$ times the width by $\pm^1/_6$ times the depth of a hall, in which the loudspeaker sound level must be flat within $\pm 3\,$dB. This rectangular area is centered at $^2/_3$ of the hall's depth in the back and at half its width. This or such a rule is a practical design target for immersive sound reinforcement systems. One goal could be maintaining the mix balanced for listeners off-center. The above criterion limits level imbalance to stay within $\pm 6\,$dB. An off-center listener at the sixth of the depth / fifth of the width roughly sits at a radius of $r = {}^1/_3$ when dimensions are normalized. For a pair of direct-sound objects rendered at the angles $\varphi = \pm 90^\circ$, off-center shifts towads one of the objects by $r = {}^1/_3$ would cause distances of $r_1 = 1 \pm {}^1/_3$, and with a distance decay $^1/_{r_{1,2}}$ of point-source loudspeakers an imbalance $g = {}^{r_1}/_{r_2} = 2$ of $\pm 6\,$dB.

Knoll et al [2] investigated the optimal lead-to-accompaniment mix in mono and while ratings depended on presentation sequence, listeners could be shown to accept a range of $\pm 3\,$dB mixing imbalance around the optimum, *so should it rather be* $\pm 3\,$dB *than* $\pm 6\,$dB*?*

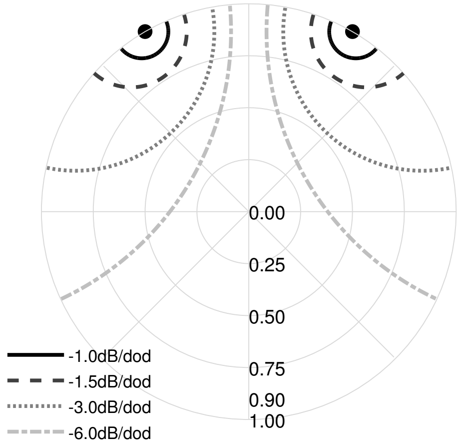
Sound systems using line-source arrays [3] allow to customize the decay with distance for every loudspeaker to $^1/_{r^\beta}$, or $-6\,\mathrm{dB} \cdot \beta$ per doubling of distance with the design parameter $\beta < 1$. Recently, Riedel et al [4] showed for horizontally surrounding loudspeakers that envelopment of diffuse sounds is best maintained within a large audience area when $\beta = {}^1/_2$; *would this also ensure a balanced mix of direct-sound objects?*

We present a small study below with different direction pairs in which listeners had the task to balance a mix composed of two sound objects. We will propose a $\pm 3\,$dB criterion, which yields listening areas for a balanced mix in Fig. 1 for the design alternatives $\beta \in \{{}^1/_6, {}^1/_4, {}^1/_2, 1\}$.

**Figure 1:** Audience areas when accepting $< 3\,$dB mixing imbalance for a mix of two sound objects, in contours for surround rendering on loudspeakers with $-6\,\beta\,$dB direct-sound decay per doubling of distance and alternative values for $\beta$.




(a) $\pm 3\,$dB limits for objects separated by 180°




(b) $\pm 3\,$dB limits for objects separated by 60°

## Experiment

15 participants with experience in audio engineering participated and solved the given mixing tasks (average age 30 years, male, average duration 19 min) in the IEM anechoic room with 5 Genelec 8020 loudspeakers laid out at 45° spacing in a frontal semi-circle at ear height Fig. 2.

Sound-object stems of these audio loops had to be mixed:

(i) a 12.14 s loop of piano and female singer from Clara Berry and Wooldog's *Waltz For My Victims* from MUSDB18 [5] between 37.73 s to 49.870 s.

(ii) a 5.57 s loop with two intertwined electric guitar riffs from Lenny Kravitz' intro to *Always On the Run* (guitar riffs played and recorded by first author),

The audio material was chosen as it contains typical pop music with singer's voice and accompaniment that is stationary enough in level by its repeatedly played chords, and because it contains a pair of equally important and

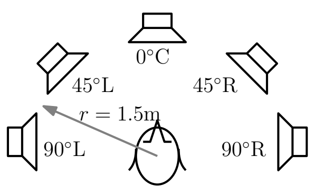
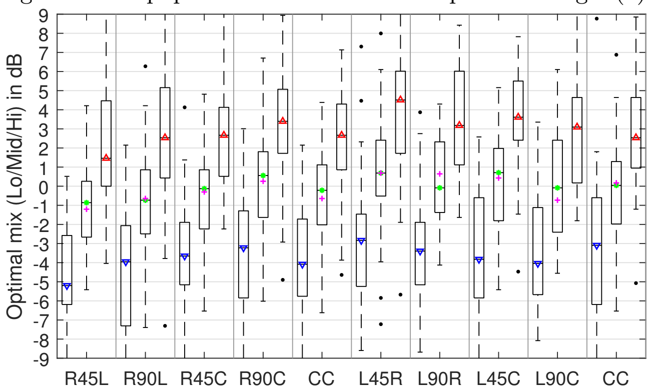



**Figure 2:** Experimental setup in anechoic room with Genelec 8020 loudspeakers at a radius of $r = 1.5\,$m.

**Figure 3:** All Lo/Mid/Hi mixing balance ratings over all pieces and listeners, after compensating the piece-related mixing bias with population median of the experiment Fig. 4(a).




rhythmically complementing riffs with same dynamic and spectral range. The directional mapping and audio loops for the 20 different trials are listed in Tab. 1. A neutral mixing balance was pre-adjusted by ear.

Listeners should adjust the sound-object mixing level by a slider (whose implicit range was $-15\cdots+15\,\mathrm{dB}$) and whose right side favored the lead/beat voice versus the accompaniment/offbeat voice. The slider adjusted the player interactively to use the RMS-normalized gains

$$g_{\mathrm{lead}} = \frac{10^{\frac{\mathrm{slider}}{20}}}{\sqrt{1+10^{\frac{\mathrm{slider}}{10}}}}, \qquad g_{\mathrm{acc}} = \frac{1}{\sqrt{1+10^{\frac{\mathrm{slider}}{10}}}}. \tag{1}$$

All 20 trials were presented individually random-permuted for trial $1\dots10$, followed by individually random-permuted trials $11\dots20$, to start with a more familiar task. In every trial, listeners were asked to provide a mix in 3 versions, and sliders were uncovered sequentially to focus on a sequence of mixing tasks:
(i) set slider for balanced mix (Mid) and confirm,
(ii) set slider 2 to lower balance limit (Low) and confirm,
(iii) set slider 3 to upper balance limit (Hi), and listeners could re-adjust all sliders before proceeding to their next trial. The initial value of slider 1 was 0 dB, of slider 2 it was $-15\,\mathrm{dB}$, and of slider 3 it was $+15\,\mathrm{dB}$, in every trial. Listeners were also informed about the purpose to find allowable mixing imbalances occurring in sound reinforcement in the audience, from the view of a sound engineer.

## Results and Discussion

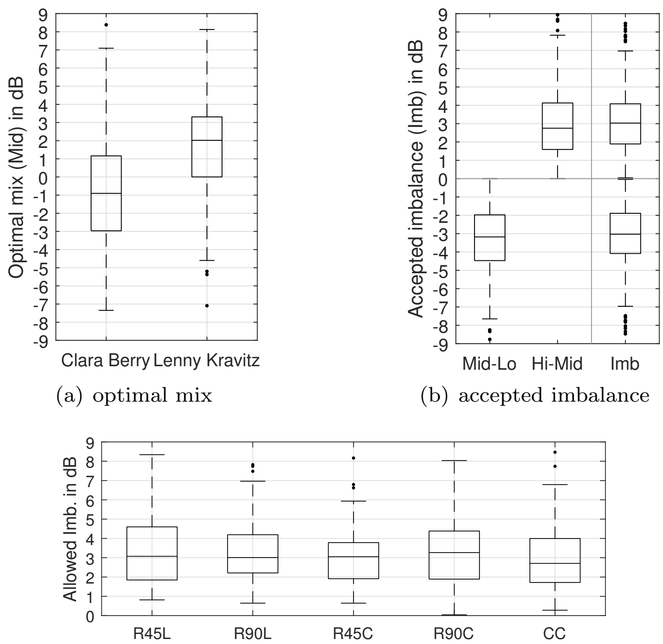
As the pre-adjusted optimal mix balance needs not necessarily match the preference of the overall population, the median preference of $-0.9\,\mathrm{dB}$ towards piano accompaniment in piece 1 and $+2\,\mathrm{dB}$ towards the beat riff piece 2 in Fig. 4(a) were subtracted from the experimental ratings.

Fig. 3 gives a statistical overview (boxes with 25% percentile, median, 75% percentile) of all these ratings with the Lo (blue), Mid (optimal, green), and Hi (red) ratings given by the participants for the various playback mappings. The medians of the Lo and Hi ratings are shown as pink markers, for reference, and their deviation to the median of Mid ranges between $-0.6\,\mathrm{dB}\cdots+0.7\,\mathrm{dB}$ with an average of $-0.1\,\mathrm{dB}$. One might feel tempted to read off

**Table 1:** Signals and directional mappings in the experiment.

| | | mapping | |
|---|---|---|---|
| trial | piece | acc./offb. | lead/beat |
| 1 | Waltz For My Victims | 90°L | 90°R |
| 2 | Waltz For My Victims | 90°R | 90°L |
| 3 | Waltz For My Victims | 45°L | 45°R |
| 4 | Waltz For My Victims | 45°R | 45°L |
| 5 | Waltz For My Victims | 0°C | 0°C |
| 6 | Waltz For My Victims | 0°C | 0°C |
| 7 | Waltz For My Victims | 45°L | 0°C |
| 8 | Waltz For My Victims | 45°R | 0°C |
| 9 | Waltz For My Victims | 90°L | 0°C |
| 10 | Waltz For My Victims | 90°R | 0°C |
| 11 | Always on the Run | 90°L | 90°R |
| 12 | Always on the Run | 90°R | 90°L |
| 13 | Always on the Run | 45°L | 45°R |
| 14 | Always on the Run | 45°R | 45°L |
| 15 | Always on the Run | 0°C | 0°C |
| 16 | Always on the Run | 0°C | 0°C |
| 17 | Always on the Run | 45°L | 0°C |
| 18 | Always on the Run | 45°R | 0°C |
| 19 | Always on the Run | 90°L | 0°C |
| 20 | Always on the Run | 90°R | 0°C |

loudspeaker level imbalances from the given Mid ratings for R45L compared to L45R, L45C, R90C, however this is not relevant to the current study. In fact, the task to rate Mid was relevant to always ensure listeners were focusing on their currently preferred balances before finding the tolerated imbalances to both sides Lo/Hi. Variations of Mid in Fig. 3 may not only be speculated to relate to loudspeaker gain mismatches of up to a dB, but maybe more so relate to the respective listeners' HRTFs or most




(a) optimal mix

(b) accepted imbalance

(c) Left-right symmetrized analysis of Imb.

**Figure 4:** Boxplot analysis of the accepted mixing imbalance.

likely to the finding by Knoll et al [2] that the optimal mix tends to be context-dependent, i.e. time-variant.

For the purposes of this study, it is justified to focus on the Lo and Hi ratings relative to the respective instantaneous optimal mix Mid. The ranges Mid−Lo, Hi−Mid are less affected by context and time variation, as their ratings stem from the same task and time at which the optimal mix was set. We compare the relative measures Mid−Lo, Hi−Mid to their average

$$\mathrm{Imb} = \frac{\mathrm{Hi} - \mathrm{Lo}}{2} \tag{2}$$

in Fig. 4(b). While there seems to be slight a bias towards piano accompaniment/offbeat riff ($-0.3\,\mathrm{dB} \ldots -0.2\,\mathrm{dB}$), the observed dB-value range is practically irrelevant.

The justified assumption that there are mappings with smaller mixing tolerance because their support for binaural unmasking is weak or absent can be investigated when regarding Fig. 4(c) with the left-right data symmetrically pooled for the tolerance range Imb. The increase of the tolerance Imb for R90C/L90C by $+0.2\,\mathrm{dB}$ from the average $3\,\mathrm{dB}$ might be explained by this, as well as the reduction of CC by $-0.3\,\mathrm{dB}$, which is small. As these variations of the tolerance across mappings are not close to reaching integer dB-values, yet, the average of $3\,\mathrm{dB}$ could be seen as independent of the directional mapping of the sound objects rendered.

## Deriving Constraints for Immersive Sound Reinforcement

Fig. 5 shows the distances to a pair of sources at $\pm\alpha$. Considering that the loudspeakers/sources could be designed to yield a direct-sound amplitude decay with $1/r_i^\beta$, we get an imbalance factor $g = \left(\frac{r_1}{r_2}\right)^\beta$, and with the law of cosines,

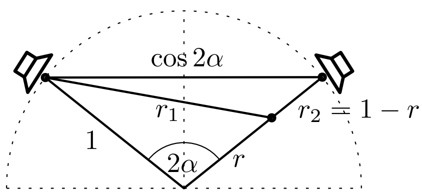
$$g^{\frac{1}{\beta}} = \frac{r_1}{r_2} = \frac{\sqrt{1 + r^2 - 2r\,\cos 2\alpha}}{1 - r}. \tag{3}$$

We can calculate the required $\beta$ for a desired radial range $r$ as target audience area:

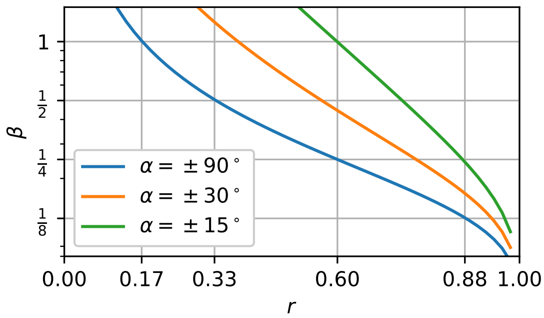
$$\beta = \frac{\ln g}{\ln \frac{\sqrt{1+r^2-2r\,\cos 2\alpha}}{1-r}}. \tag{4}$$

Assuming $g = \sqrt{2}$ defined the $\pm 3\,\mathrm{dB}$ imbalance tolerance to be representative for audio engineering preferences, we can derive constraints for a listening area in multichannel surround sound, concerning the preservation of a balanced mix of direct sounds for such a population.

Fig. 1 shows different countours indicating different listening areas limited by the tolerated $\pm 3\,\mathrm{dB}$ mixing imbalance, for 4 alternative $\beta$ values providing $-1\,\mathrm{dB}$, $-1.5\,\mathrm{dB}$, $-3\,\mathrm{dB}$, or $-6\,\mathrm{dB}$ per doubling of distance (dod). The figure shows two examples of a pair of sound objects played with an angular separation of either 180° as most demanding case, or one of 60° as typical stereo case; naturally, a 0° separation would yield a mono mix whose direct sound can't get out of balance.




(a) distances $r_1$, $r_2$ for sound objects at $\pm\alpha$ and listener at $r$




(b) source design exponent $\beta$ for listener at $r$ and sound-object spread $2\alpha$

**Figure 5:** Trigonometry of sound-object pair layout on normalized-radius surround system and off-center listener at $r$, and resulting values for $\beta$ if allowable imbalance is $g = \sqrt{2}$, according to eq. (4).

The listening area is rather restricted for typical point-source loudspeakers with $-6\,\mathrm{dB/dod}$. For a relative diameter that is nowhere constricted to less than 75%, decays should be limited to $-1\,\mathrm{dB/dod}$, or if the main directional voices are frontal: probably $-1.5\,\mathrm{dB/dod}$.

## Conclusion

This contribution presented an experiment simulating distance-related mixing imbalances that arise depending on the listening position within an extended audience area, when playback considers sound objects played back surrounding loudspeakers in a typical multi-channel audio system or a large, immersive sound reinforcement layout with surrounding line-array systems.

In the mixing task of the experiment, listeners would permit a $\pm 3\,\mathrm{dB}$ mixing imbalance with regard to the optimal mix that they would define at the same time instant. Between the two sound objects to be mixed, several spacings and directional mappings were investigated. While the tolerated imbalance varies slightly across different directional mappings, or also slightly to one of the two sound objects, the amount of the variation is still irrelevant to the practice, as it is noticeably smaller than one decibel.

Moreover, we used the tolerated mixing imbalance to plot usable audience areas sizes or *sweet areas* fulfilling this tolerance, when surrounding loudspeakers are line arrays designed to fulfill specific dB level decays per doubling of the distance (dB/dod). Simulations showed that for a reasonably large sweet area, $-1\,\mathrm{dB/dod}$ for strict demands or $-1.5\,\mathrm{dB/dod}$ for predominantly frontal direct-sound objects could be reasonable design targets. We could moreover show a design equation for listening areas of size determined by the maximum off-center ratio $r$, the maximum acceptable mixing imbalance $g$, and the source separation $2\alpha$.

This is somewhat complementing our previous findings on enveloping diffuse-sound objects, for which a $-3\,\mathrm{dB/dod}$ design target works best in horizontally surrounding loudspeaker system.

## Acknowledgment

The authors were supported by the Austrian Science Fund (FWF) under project number P 35254-N, titled envelopment in immersive sound reinforcement (EnImSo).

## References

[1] ISO 2969:2015, "Cinematography — B-chain electroacoustic reponse of motion-picture control rooms and indoor theatres — Specifications and measurements," 2015.

[2] A. L. Knoll and K. Siedenburg, "The optimal mix? presentation order affects preference ratings of vocal amplitude levels in popular music," *Music & Science*, pp. 2059–2043, 2022. [Online]. Available: https://doi.org/10.1177/20592043221142712

[3] F. Straube, F. Schultz, D. A. Bonillo, and S. Weinzierl, "An analytical approach for optimizing the curving of line source arrays," *J. Audio Eng. Soc*, vol. 66, no. 1/2, pp. 4–20, 2018. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=19372

[4] S. Riedel and F. Zotter, "Surrounding line sources optimally reproduce diffuse envelopment at off-center listening positions," *JASA EL 2*, no. 094404, 2022. [Online]. Available: https://doi.org/10.1121/10.0014168

[5] Z. Rafii, A. Liutkus, F.-R. Stöter, S. I. Mimilakis, and R. Bittner, "The MUSDB18 corpus for music separation," Dec. 2017. [Online]. Available: https://doi.org/10.5281/zenodo.1117372