Buenos Aires – 5 to 9 September 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

VA - Virtual Acoustics: Paper 801

Investigation into the role of the nonnegativity constraint in sound field reproduction problems

Filippo Maria Fazi[a], **Andreas Franck**[a]

[a] Institute of Sound and Vibration Research, University of Southampton, United Kingdom,
{filippo.fazi,andreas.franck}@soton.ac.uk

Abstract:

Given a sound field control algorithm to control the pressure and particle velocity of the field at one point located in the interior of an array of loudspeakers, it is shown that an exact solution cannot, in general, be achieved if the secondary source strength is constrained to be nonnegative. This result is put into relation with Makita's velocity vector and Gerzon's energy vector used to model human sound localisation. It is shown that, in general, Makita's vector can have the desired direction and magnitude equal or larger than one only if the non-negativity constraint is removed, and that Gerzon's vector cannot have both the desired direction and unitary magnitude.

Keywords: Spatial sound reproduction, first-order Ambisonics, energy and velocity vector, non-negativity, in-phase decoding

Investigation into the role of the nonnegativity constraint in sound field reproduction problems

1 Introduction

Many sound reproduction techniques aim at correctly synthesising the particle velocity and/or the sound pressure of a target sound field at a single, central listening position. This includes first-order Ambisonics [1], [2] but also panning approaches such as vector base amplitude panning (VBAP) [3].

In a recent paper [4] we proposed the use of a convex optimization framework to describe and compare various sound reproduction methods in a uniform way. In the present paper we focus on a specific aspect of reproduction methods, that is gain nonnegativity. Gain nonnegativity refers to a restriction of the loudspeaker driving gains in a multi-channel audio systems to nonnegative values. It is conceptually equivalent to in-phase decoding in Ambisonics [5], [6].

We show which reproduction methods comply with this criterion or can be extended to meet this condition, and the effects on the loudspeaker gain distribution. In particular, we demonstrate that exact mode matching is not possible if the gains are restricted to be nonnegative. We outline the relation between the norm of the gain vector and the magnitude of the velocity vector and of the energy vector.

2 Sound Field Model

The sound field model used here is described in more detail in [4]. It follows the conventions used in Ambisonics, e.g., [2], but also in 3D amplitude panning approaches as VBAP [3]. The sound fields generated by the $i-$th loudspeaker as well as that of a single virtual source are modelled as plane waves, represented by Cartesian unit vectors $\mathbf{l}_i$ and $\mathbf{p}$, respectively. Consequently, a loudspeaker setup consisting of L loudspeakers is described by the matrix

$$\mathbf{L} = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 & \cdots & \mathbf{l}_L \end{bmatrix}. \tag{1}$$

Sound fields are reproduced by driving the loudspeakers with the signals of virtual sources, where each loudspeaker signal is scaled by an individual gain g_i. Here we consider approaches to find gain vectors $\mathbf{g} = \begin{bmatrix} g_1 & g_2 & \cdots & g_L \end{bmatrix}$ such that the resulting sound field at the central listening position approximates the sound field of a single virtual source $\mathbf{p}$.

Omitting physical proportionality constants, the sound pressure $\tilde{p}_{sp}$ and the particle velocity $\tilde{\mathbf{p}}$ (times -1) of the reproduced sound field are given by

$$\tilde{p}_{sp} = \sum_{i=1}^{L} g_i \quad = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \mathbf{g} \tag{2a}$$

$$\tilde{\mathbf{p}} = \sum_{i=1}^{L} g_i \mathbf{l}_i = \mathbf{L}\mathbf{g}. \tag{2b}$$

By definition, the desired sound pressure of virtual sound source is 1.

2.1 Velocity and Energy Vector Magnitude

The velocity vector, also termed the Makita vector, and the energy vector proposed by Gerzon [1], [6], are two established objective measures to assess sound localisation. They are defined as

$$\mathbf{r}_{\mathrm{v}} = \hat{\mathbf{r}}_{\mathrm{v}} r_v = \frac{\sum_{i=1}^{L} \mathbf{l}_i g_i}{\sum_{i=1}^{L} g_i} \tag{3a}$$

$$\mathbf{r}_{\mathrm{e}} = \hat{\mathbf{r}}_{\mathrm{e}} r_e = \frac{\sum_{i=1}^{L} \mathbf{l}_i g_i^2}{\sum_{i=1}^{L} g_i^2} \,. \tag{3b}$$

The Makita vector $\mathbf{r}_{\mathrm{v}}$ is suited to describe localisation at low frequencies ($< 700\,\mathrm{Hz}$), while the energy vector $\mathbf{r}_{\mathrm{e}}$ is more appropriate for high frequencies ($\geq 700\,\mathrm{Hz}$).

As represented in (3), these vectors can be separated into unit vectors $\hat{\mathbf{r}}_{\mathrm{v}}$ and $\hat{\mathbf{r}}_{\mathrm{e}}$ describing the particle velocity and energy direction, respectively, and scalar velocity and energy vector magnitudes r_{v} and r_{e} to describe the length of the vectors. These magnitudes are a measure of the perceived spread of a reproduced sound source. According to Gerzon [1], the velocity vector magnitude can take arbitrary values, while the energy vector magnitude is less or equal to 1. For a natural plane wave source, both r_{v} and r_{e} are 1. Reproduction methods should approximate these ideal values.

2.2 Convex Optimisation Representation

In [4] we proposed a convex optimisation framework to describe multiple sound reproduction methods in a uniform way. Convex optimisation is a powerful way to model and efficiently solve a wide range of optimisation problems [7]. A convex problem has the general form

$$\underset{\mathbf{x}}{\operatorname{argmin}} \; f(\mathbf{x}) \tag{4a}$$

$$\text{subject to } g_i(\mathbf{x}) \leq 0 \qquad i = 1, \ldots, m \tag{4b}$$

$$h_j(\mathbf{x}) = 0 \qquad j = 1, \ldots, n, \tag{4c}$$

where (4a) is the convex objective function, and (4b) and (4c) are potentially empty sets of convex inequality and linear equality constraints, respectively. Among the family of convex functions, p-norms

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{L} |x_i|^p \right)^{\frac{1}{p}} \tag{5}$$

are of particular importance here, where $\|\cdot\|_2$ represents the Euclidean or least-squares norm, and $\|\cdot\|_1$ is the ℓ_1 norm.

For the numerical examples in this paper, we use CVX, a modelling toolkit for convex programs [8].

3 The impact of nonnegativity constraints

3.1 First-Order Ambisonics mode matching

Simultaneously controlling sound pressure and particle velocity at the central listening position can be expressed as a first-order mode matching problem [9]

$$\Psi \mathbf{g} = \begin{bmatrix} 1 \\ \mathbf{p} \end{bmatrix} \qquad \text{with } \Psi = \begin{bmatrix} 1 & \cdots & 1 \\ & \mathbf{L} & \end{bmatrix} \tag{6}$$

where Ψ is referred to as the mode matrix of the setup. However, this formulation does not incorporate nonnegativity constraints for the gain vector $\mathbf{g}$. As the mode-matching problem is typically underdetermined for multi-loudspeaker setups, additional conditions may be introduced. For mode matching with general, irregular loudspeaker arrays, a minimum ℓ_2 norm solution based on the Moore-Penrose pseudoinverse, namely

$$\mathbf{g} = \Psi^{\dagger} \begin{bmatrix} 1 \\ \mathbf{p} \end{bmatrix} \qquad \text{with } \Psi^{\dagger} = \Psi^{H} \left(\Psi \Psi^{H} \right)^{-1}, \tag{7}$$

is typically used. This provides a solution to (6) such that the ℓ_2 norm $\|\mathbf{g}\|_2$ of the gain vector is minimised. This is equivalent to a convex optimisation problem

$$\underset{\mathbf{g}}{\operatorname{argmin}} \; \|\mathbf{g}\|_2 \tag{8a}$$

$$\text{subject to } \Psi \mathbf{g} = \begin{bmatrix} 1 \\ \mathbf{p} \end{bmatrix}. \tag{8b}$$

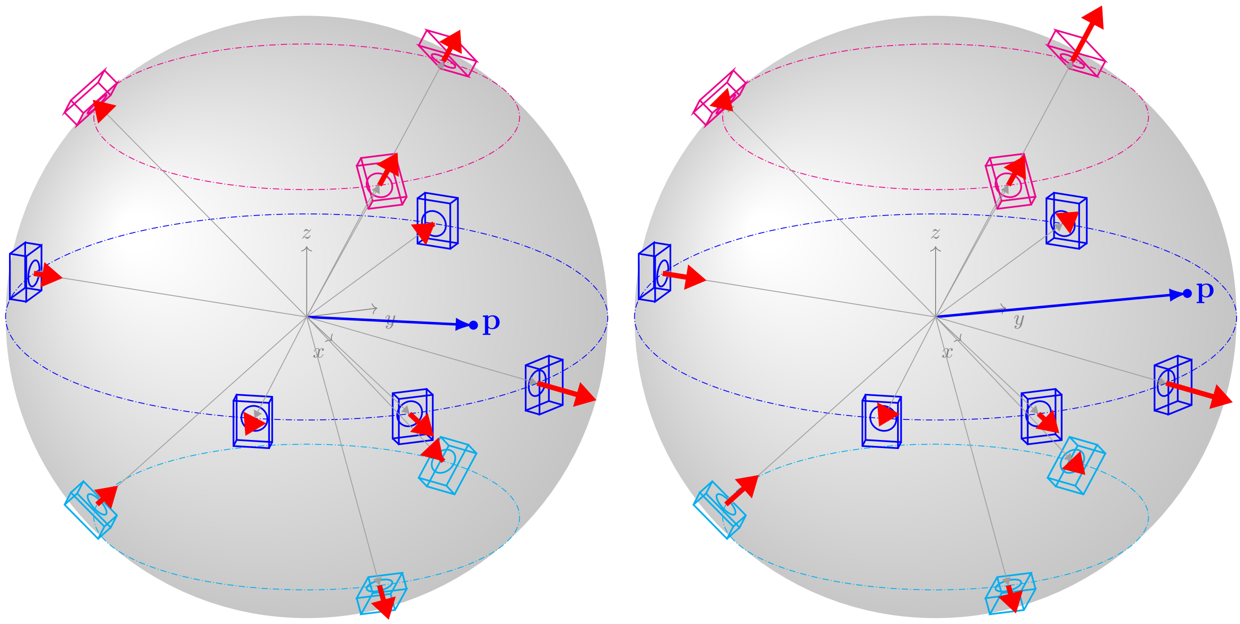
Separating the mode matrix into pressure and velocity components, this problem can be written as

$$\underset{\mathbf{g}}{\operatorname{argmin}} \; \|\mathbf{g}\|_2 \tag{9a}$$

$$\text{subject to } \sum_{i=1}^{L} g_i = 1 \tag{9b}$$

$$\mathbf{L}\mathbf{g} = \mathbf{p}. \tag{9c}$$

The resulting gains for a 3D loudspeaker setup with 11 loudspeakers and two source locations $\mathbf{p}_1 = (15^\circ, 15^\circ)$ and $\mathbf{p}_2 = (40^\circ, 15^\circ)$ are shown in Fig. 1. Obviously, the gain vectors do not comply with a nonnegativity condition. It is observed that in both cases virtually all loudspeakers are activated with significant gains and several negative values in each case.



(a) $\mathbf{p}_1 = (15^\circ, 15^\circ)$ (b) $\mathbf{p}_1 = (40^\circ, 15^\circ)$

Figure 1: Loudspeaker gains for first-order Ambisonics mode matching. Outward arrows denote positive, inward vectors denote negative gains.

3.2 Combining mode matching with gain nonnegativity

To combine the properties of the mode matching solution with gain nonnegativity, an additional inequality constraint can be added to (9), resulting in

$$\operatorname*{argmin}_{\mathbf{g}} \|\mathbf{g}\|_2 \tag{10a}$$

$$\text{subject to} \sum_{i=1}^{L} g_i = 1 \tag{10b}$$

$$\mathbf{L}\mathbf{g} = \mathbf{p} \tag{10c}$$

$$\mathbf{g} \geq 0. \tag{10d}$$

This optimisation is, however, not solvable in the general case, because no gain vector exists that simultaneously fulfils all constraints. In optimisation terminology, this is referred to as an *infeasible* problem. Applied to sound reproduction, this implies that it is not possible to correctly synthesise both sound pressure and particle velocity unless negative gains are allowed. The only exception is if the virtual source location coincides with a loudspeaker.

This restriction is equivalent to the statement that the velocity vector magnitude r_v is always less than one for nonnegative gains if the reproduced velocity vector direction $\mathbf{r}_\text{v}$ does not match a loudspeaker location.

This can be justified as follows. As the reproduced velocity vector $\tilde{\mathbf{p}}$ defined in (2b) scales with any proportionality factor applied to $\mathbf{g}$, we can enforce that $\tilde{\mathbf{p}} = \hat{\mathbf{r}}_\mathrm{v}$ without loss of generality, considering that $\|\hat{\mathbf{r}}_\mathrm{v}\| = 1$ by definition. This implies

$$r_\mathrm{v} = \frac{1}{\sum_{i=1}^{L} g_i} . \qquad (11)$$

The contribution of each loudspeaker, with its gain value, to the total velocity vector is determined by the scalar projection of the loudspeaker's particle velocity $g_i \mathbf{l}_i$ onto $\hat{\mathbf{r}}_\mathrm{v}$, that is, the scalar product $\langle \hat{\mathbf{r}}_\mathrm{v}, g_i \mathbf{l}_i \rangle$. Thus the total velocity vector magnitude is

$$r_\mathrm{e} = \sum_{i=1}^{L} g_i \langle \hat{\mathbf{r}}_\mathrm{v}, \mathbf{l}_i \rangle . \qquad (12)$$

As both $\hat{\mathbf{r}}_\mathrm{v}$ and $\mathbf{l}_i$ are unit vectors, the scalar product equals the cosine of the angle between the vectors, which is less than one unless they coincide. Consequently, any summation (12) such that $\sum g_i = 1$ and $g_i \geq 0$ is less than one unless $\hat{\mathbf{r}}_\mathrm{v} = \mathbf{l}_i$ for some loudspeaker i.

3.3 Maximising the velocity vector magnitude

As a simultaneous exact matching of sound pressure and particle velocity is not possible in general, a sensible objective is to minimise the sound pressure required to match a desired particle velocity. This can be stated as an optimisation problem as follows

$$\underset{\mathbf{g}}{\operatorname{argmin}} \sum_{i=1}^{L} g_i \qquad (13a)$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p} \qquad (13b)$$

$$\mathbf{g} \geq 0 . \qquad (13c)$$

The resulting gain distributions are displayed in Fig. 2. It is observed that the gains are indeed nonnegative, and only a few gains (three for both examples) are active. That, is the solution is sparse. Moreover, the active loudspeakers are close to the desired source location $\mathbf{p}$.

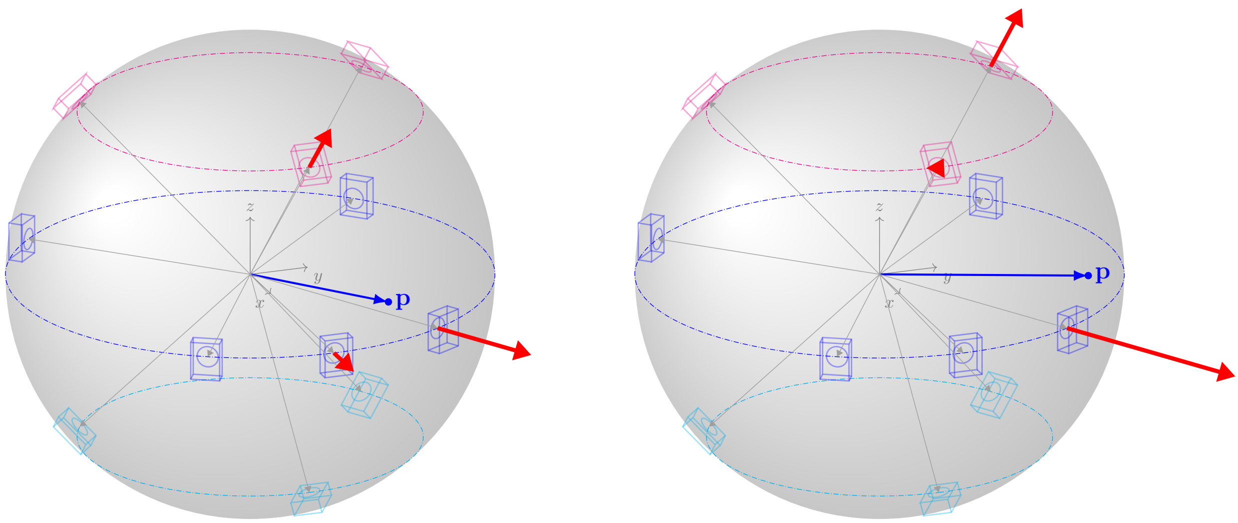
The optimisation problem (13) is equivalent to maximising the velocity vector magnitude under the constraint that the reproduced particle vector matches the desired vector $\mathbf{p}$. This follows directly from (3a), as the objective function (13b) forms the denominator of (3a). Furthermore, as the ℓ_1 norm is identical to the sum of a nonnegative vector, (13) is equivalent to the ℓ_1 minimisation problem

$$\underset{\mathbf{g}}{\operatorname{argmin}} \|\mathbf{g}\|_1 \qquad (14a)$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p} \qquad (14b)$$

$$\mathbf{g} \geq 0 . \qquad (14c)$$

In this way, the sparse gain vector for this design objective can be attributed to the sparsity-promoting nature of the ℓ_1 norm, see e.g., [10].



(a) $\mathbf{p}_1 = (15°, 15°)$ (b) $\mathbf{p}_1 = (40°, 15°)$

Figure 2: Loudspeaker gains for particle velocity magnitude maximisation with a nonnegativity constraint.

Finally, it is observed that the solutions obtained by minimizing the ℓ_1 norm subject to a nonnegativity condition are identical to the VBAP solution in the majority of cases, including the examples shown here. This equivalence is investigated in detail in a paper currently under review [11].

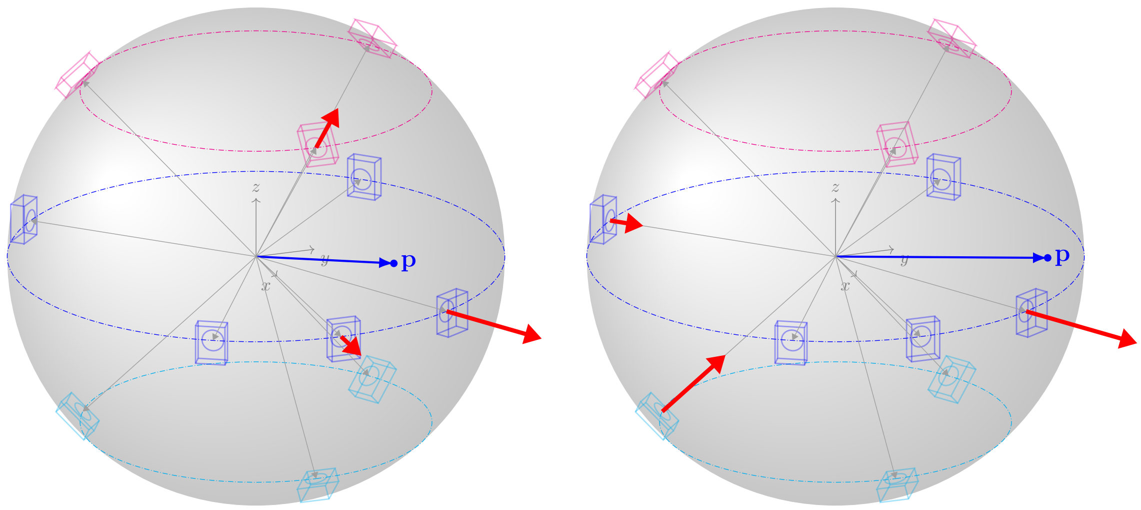
3.4 ℓ_1 minimisation without nonnegativity

Considering the equivalence between particle vector magnitude optimisation and ℓ_1 minimisation, it is worthwhile to investigate the effect of the nonnegativity constraint in this optimisation problem. Thus we drop this constraint, yielding the problem

$$\underset{\mathbf{g}}{\operatorname{argmin}} \ \|\mathbf{g}\|_1 \quad \text{(15a)}$$

$$\text{subject to } \mathbf{L}\mathbf{g} = \mathbf{p}\,. \quad \text{(15b)}$$

The loudspeaker gains obtained by this design objective are shown in Fig. 3. While for the solution is identical for position $\mathbf{p}_1$, the gain vector for $\mathbf{p}_2$ shows two negative gains on the opposite side of the array. Thus, depending on the virtual source position, this approach activates loudspeakers far off and creates anti-phase contributions to the sound field. Both effects are potentially detrimental to the perceived localisation of the source, in particular outside the central listening position.



(a) $\mathbf{p}_1 = (15°, 15°)$ (b) $\mathbf{p}_1 = (40°, 15°)$

Figure 3: Loudspeaker gains for particle velocity magnitude maximisation without nonnegativity constraint.

3.5 Relation to Energy Vector Magnitude Maximisation

In the same way as with the velocity vector in Sec. 3.3, the loudspeaker gains can be obtained such that the magnitude of the energy vector proposed by Gerzon is maximised. This design objective is used to ensure good localisation at high frequencies ($\geq 700\,\mathrm{Hz}$) and, for instance, forms the conceptual basis for $\max r_\mathrm{e}$ decoding in higher-order Ambisonics, e.g., [6], [12].

As observed in (3b), the sole difference to the velocity vector is that squared gain values g_i^2 are used instead of g_i. Thus, by substituting $\mathbf{g}$ by the vector $\mathbf{q} = \begin{bmatrix} g_1^2 & g_2^2 & \cdots & g_L^2 \end{bmatrix}^T$, energy vector magnitude maximisation can be expressed as

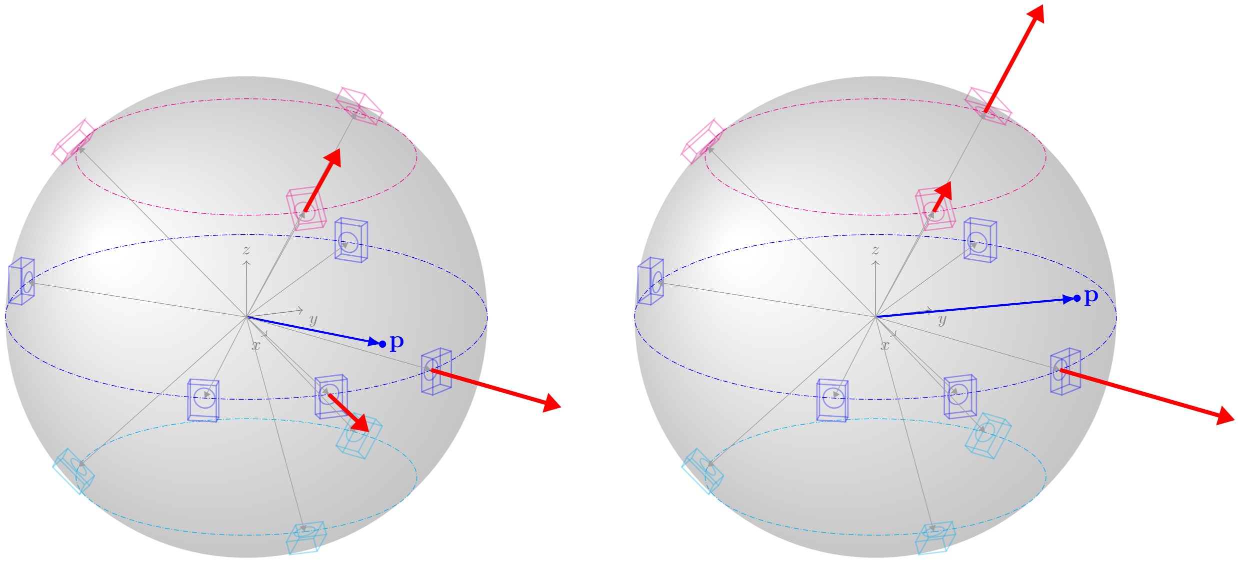
$$\underset{\mathbf{g}}{\operatorname{argmin}} \sum_{i=1}^{L} q_i \tag{16a}$$

$$\text{subject to } \mathbf{Lq} = \mathbf{p} \tag{16b}$$

$$\mathbf{q} \geq 0. \tag{16c}$$

The final gains g_i is obtained from the square roots of the elements q_i. Note that defining q_i as the square of a real-valued gain implicitly introduces the nonnegativity constraint (16c).

As observed in Fig. 4, the obtained loudspeaker gains are closely related to the velocity vector maximisation approach subject to a nonnegativity constraint described in Sec. 3.3. That is, the sparsity pattern, the nonnegativity, and the selection of active loudspeakers are identical, and the gains differ only by a square root operation. Conceptually, this relationship is similar to the relation between amplitude panning techniques for low and high frequencies described in [2].



(a) $\mathbf{p}_1 = (15°, 15°)$ (b) $\mathbf{p}_1 = (40°, 15°)$

Figure 4: Loudspeaker gains obtained by energy vector maximisation.

4 Conclusion

In this paper we have investigated the role of nonnegativity constraints imposed on the loudspeaker gains of sound reproduction methods that aim at controlling the sound pressure and the particle velocity of the reproduced field at one position. We have shown that, for a general virtual source location, sound pressure and particle velocity cannot be controlled simultaneously to match a desired target sound field if a nonnegativity constraint is active on the loudspeaker gains . Maximising the magnitude of the particle velocity vector (or Makita vector) yields the closest achievable approximation of the desired sound pressure while preserving the correct direction of the particle velocity.

We have shown that this objective can be posed as a ℓ_1 minimisation and thus expressed and solved as a convex optimisation problem. It has been demonstrated that the resulting gain distributions are sparse, i.e., with only few nonzero values, and equivalent to existing amplitude panning techniques as VBAP in the majority of possible virtual source positions. Introducing a nonnegativity constraint effectively prevents antiphase loudspeaker signals, thus improving sound quality and localisation in practical sound reproduction. Moreover, it is shown that the computation of panning gains by ℓ_1-optimisation with nonnegativity constraint is closely related to approaches that maximise Gerzon's energy vector magnitude, because this objective implicitly involves a nonnegativity constraint.

Acknowledgements

This work was supported by the EPSRC Programme Grant S3A: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L000539/1) and the BBC as part of the BBC Audio Research Partnership.

5 References

[1] Gerzon, M. A., Panpot laws for multispeaker stereo, in Audio engineering society 92th convention, Vienna, Austria, March 1992.

[2] Jot, J.-M.; Larcher, V.; Pernaux, J.-M., A comparative study of 3-D audio encoding and rendering techniques, in Aes 16th international conference: Spatial sound reproduction, Rovaniemi, Finland, March 1999.

[3] Pulkki, V., Virtual sound source positioning using vector base amplitude panning, Journal of the audio engineering society, Vol 45 (6), pp 456–466, Jun. 1997.

[4] Franck, A.; Fazi, F. M., Comparison of listener-centric sound field reproduction methods in a convex optimization framework, in Aes international conference on sound field control, Guildford, UK, Jul. 2016.

[5] Malham, D. G., Experience with large area 3-D Ambisonic sound systems, Proceedings of the institute of acoustics, Vol 14 (5), pp 209–215, 1992.

[6] Daniel, J.; Rault, J.-B.; Polack, J.-D., Ambisonics encoding of other audio formats for multiple listening conditions, in Audio engineering society convention 105, San Francisco, CA, USA, Sep. 1998.

[7] Boyd, S.; Vandenberghe, L., Convex optimization. Cambridge University Press, Cambridge (UK), 2004.

[8] Grant, M. C.; Boyd, S. P., Graph implementations for nonsmooth convex programs, in *Recent advances in learning and control*, ser. Lecture Notes in Control and Information Sciences, Blondel, V.; Boyd, S.; Kimura, H., Eds., Springer, London, UK, 2008, pp 95–110.

[9] Fazi, F. M.; Shin, M.; Olivieri, F.; Fontana, S.; Lang, Y., Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays, in Audio engineering society 137th convention, paper no. 9111, Los Angeles, CA, USA, October 2014.

[10] Candes, E.; Wakin, M., An introduction to compressive sampling, Ieee signal processing magazine, Vol 25 (2), pp 21–30, March 2008.

[11] Franck, A.; Wang, W.; Fazi, F. M., Sparse, ℓ_1 -optimal multi-loudspeaker panning and its equivalence to vector base amplitude panning, Ieee/acm transactions on audio, speech, and language processing, 2016, Submitted.

[12] Zotter, F.; Frank, M., All-round Ambisonic panning and decoding, Journal of the audio engineering society, Vol 60 (10), pp 807–820, October 2012.