



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication: **23.07.2014 Bulletin 2014/30** (51) Int Cl.: **H04R 3/00 (2006.01)**

(21) Application number: **13152209.6**

(22) Date of filing: **22.01.2013**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**  
 Designated Extension States:  
**BA ME**

(72) Inventor: **Christoph, Markus**  
**94315 Straubing (DE)**

(74) Representative: **Patentanwälte Westphal, Mussgnug & Partner**  
**Herzog-Wilhelm-Strasse 26**  
**80331 München (DE)**

(71) Applicant: **Harman Becker Automotive Systems GmbH**  
**76307 Karlsbad (DE)**

(54) **Modal beamforming**

(57) A method and system for generating an auditory scene that comprises: receiving eigenbeam outputs, the eigenbeam outputs having been generated by decomposing a plurality of audio signals, each audio signal having been generated by a different microphone of a microphone array, wherein each eigenbeam output corresponds to a different eigenbeam for the microphone array; generating the auditory scene based on the eigenbeam outputs and their corresponding eigenbeams, wherein generating the auditory scene comprises apply-

ing a weighting value to each eigenbeam output to form a steered eigenbeam output; and combining the weighted eigenbeams to generate the auditory scene, wherein generating the auditory scene further comprises applying a regularized equalizer filter to each eigenbeam output or steered eigenbeam output, the regularized equalizer filter(s) being configured to compensate for acoustic deficiencies of the microphone array and having a regularized equalization function.

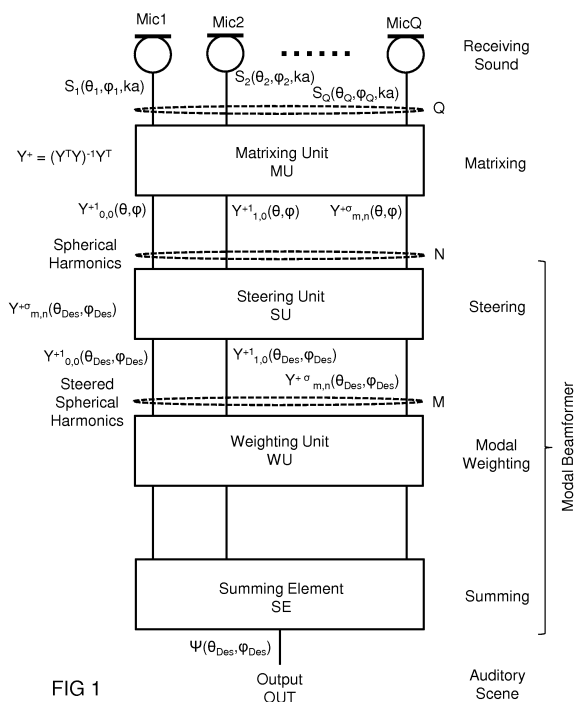


FIG 1

**Description**

## FIELD OF TECHNOLOGY

5 **[0001]** The embodiments disclosed herein refer to sound capture systems and methods, particularly to sound capture methods that employ modal beamforming.

## BACKGROUND

10 **[0002]** Beamforming sound capture systems comprise at least (a) an array of two or more microphones and (b) a beamformer that combines audio signals generated by the microphones to form an auditory scene representative of at least a portion of an acoustic sound field. Due to the underlying geometry, it is natural to represent the sound field captured on the surface of a sphere with respect to spherical harmonics. In this context, spherical harmonics are also known as acoustic modes (or eigenbeams) and the appending signal-processing techniques as modal beamforming.

15 **[0003]** Two spherical microphone array configurations are commonly employed: the sphere may exist physically, or may merely be conceptual. In the first configuration, the microphones are arranged around a rigid sphere made of, for example, wood or hard plastic. In the second configuration, the microphones are arranged in free-field around an "open" sphere, referred to as an open-sphere configuration. Although the rigid-sphere configuration provides a more robust numerical formulation, the open-sphere configuration might be more desirable in practice at low frequencies where large spheres are realized.

20 **[0004]** Beamforming techniques allow for the controlling of the characteristics of the microphone array in order to achieve a desired directivity. One of the most general formulations is the filter-and-sum beamformer, which has readily been generalized by the concept of modal subspace decomposition. This approach finds optimum finite impulse response (FIR) filter coefficients for each microphone by solving an eigenvalue problem and projecting the desired beam pattern onto the set of eigenbeam patterns found.

25 **[0005]** Beamforming sound capture systems enable picking up acoustic signals dependent on their direction of propagation. The directional pattern of the microphone array can be varied over a wide range due to the degrees of freedom offered by the plurality of microphones and the processing of the associated beamformer. This enables, for example, steering the look direction, adapting the pattern according to the actual acoustic situation, and/or zooming in to or out from an acoustic source. All this can be done by controlling the beamformer, which is typically implemented via software, such that no mechanical alteration of the microphone array is needed. However, common beamformers fail to be directive at very low frequencies. Therefore, modal beamformers having less frequency-dependent directivity are desired.

## SUMMARY

35 **[0006]** A method for generating an auditory scene comprises: receiving eigenbeam outputs generated by decomposing a plurality of audio signals, each audio signal having been generated by a different microphone of a microphone array, wherein each eigenbeam output corresponds to a different eigenbeam for the microphone array; generating the auditory scene based on the eigenbeam outputs and their corresponding eigenbeams, wherein generating the auditory scene comprises applying a weighting value to each eigenbeam output to form a steered eigenbeam output; and combining the weighted eigenbeams to generate the auditory scene, wherein generating the auditory scene further comprises applying a regularized equalizer filter to each eigenbeam output or steered eigenbeam output, the regularized equalizer filter(s) being configured to compensate for acoustic deficiencies of the microphone array and having a regularized equalization function.

40 **[0007]** A modal beamformer system for generating an auditory scene comprises: a steering unit that is configured to receive eigenbeam outputs, the eigenbeam outputs having been generated by decomposing a plurality of audio signals, each audio signal having been generated by a different microphone of a microphone array, wherein each eigenbeam output corresponds to a different eigenbeam for the microphone array and the microphones are arranged on a rigid or open sphere; a weighting unit that is configured to generate the auditory scene based on the eigenbeam outputs and their corresponding eigenbeams, wherein generating the auditory scene comprises applying a weighting value to each eigenbeam output to form a steered eigenbeam output; and a summing element configured to combine the weighted eigenbeams to generate the auditory scene, wherein the weighting unit or the summing element are further configured to apply a regularized equalizer filter to each eigenbeam output or steered eigenbeam output, the regularized equalizer filter(s) being configured to compensate for acoustic deficiencies of the microphone array and having a regularized equalization function.

BRIEF DESCRIPTION OF THE DRAWINGS

**[0008]** The figures identified below are illustrative of some embodiments of the invention. The figures are not intended to be limiting of the invention recited in the appended claims. The embodiments, both as to their organization and manner of operation, together with further object and advantages thereof, may best be understood with reference to the following description, taken in connection with the accompanying drawings, in which:

FIG. 1 is a schematic representation of a generalized structure of a sound capture system that employs modal beamforming;

FIG. 2 is a schematic representation of a possible microphone array for the sound capture system of FIG. 1 ;

FIG. 3 is a schematic representation of a more detailed structure of a sound capture system that employs modal beamforming;

FIG. 4 is a schematic representation of an arrangement for extracting ambisonic components with which an arbitrary sound field can be coded and/or decoded;

FIG. 5 is a schematic representation of an arrangement for measuring a sound pressure field;

FIG. 6 is a schematic diagram illustrating the radial function of a spherical microphone array;

FIG. 7 is a schematic diagram illustrating the magnitude frequency response of the equalizer filter corresponding to the radial function illustrated in FIG. 6;

FIG. 8 is a flow chart illustrating the process of calculating the equalizer filter referred to above in connection with FIG. 7;

FIG. 9 is a schematic diagram illustrating the regularization parameter over frequency for an improved 4<sup>th</sup>-order modal beamformer with a given minimal white noise gain of -10 [dB];

FIG. 10 is a schematic diagram corresponding to the flow chart of FIG. 8 and the diagram of FIG. 9, and illustrating the white noise gain for a 4<sup>th</sup>-order modal beamformer utilizing a regularized equalizing filter;

FIG. 11 is a schematic diagram corresponding to the flow chart of FIG. 8 and the diagram of FIG. 9, and illustrating the directivity index for a 4<sup>th</sup>-order modal beamformer utilizing a regularized equalizing filter;

FIG. 12 is a schematic diagram illustrating the magnitude frequency response of the improved regularized equalizing filter;

FIG. 13 is a schematic diagram illustrating the corresponding phase response of the improved filter of FIG. 12;

FIG. 14 is a schematic diagram illustrating the magnitude frequency response of an improved, regularized equalizing filter;

FIG. 15 is a schematic diagram illustrating the corresponding phase frequency response of the improved filter of FIG. 14; and

FIG. 16 is a schematic diagram illustrating the cylindrical view of the directional pattern of the improved 4<sup>th</sup>-order modal beamformer over frequency.

DESCRIPTION

**[0009]** FIG. 1 is a block diagram illustrating the basic structure of a beamforming sound capture system as described in more detail, for instance, in WO 03/061336. The sound capture system comprises a plurality Q of microphones Mic1, Mic2, ... MicQ configured to form a microphone array, a matrixing unit MU (also known as modal decomposer or eigen-beam former), and a modal beamformer BF. In the system of FIG. 1, modal beamformer BF comprises a steering unit SU, a weighting unit WU, and a summing element SE, each of which will be discussed in further detail later in this

specification. Each microphone Mic1, Mic2, ... MicQ generates a time-varying analog or digital audio signal  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  corresponding to the sound incident at the location of that microphone.

**[0010]** Matrixing unit MU decomposes (according to  $Y^+ = (Y^T Y)^{-1} Y^T$ ) the audio signals  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  generated by the different microphones Mic1, Mic2, ... MicQ to generate a set of spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$ , also known as eigenbeams or modal outputs, where each spherical harmonic  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$  corresponds to a different mode for the microphone array. The spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$  are then processed by beamformer BF to generate an auditory scene that is represented in the present example by output signal OUT ( $=\Psi(\theta_{Des}, \varphi_{Des})$ ). In this specification, the term *auditory scene* is used generically to refer to any desired output from a sound capture system, such as the system of FIG. 1. The definition of the particular auditory scene will vary from application to application. For example, the output generated by beamformer BF may correspond to one or more output signals, e.g., one for each speaker used to generate the resultant auditory scene. Moreover, depending on the application, beamformer BF may simultaneously generate beampatterns for two or more different auditory scenes, each of which can be independently steered to any direction in space. In certain implementations of the sound capture system, microphones Mic1, Mic2, ... MicQ may be mounted on the surface of an acoustically rigid sphere or may be arranged on a virtual (open) sphere to form the microphone array. Alternatively, weighting unit WU may be arranged upstream of steering unit SU so that the non-steered eigenbeams are weighted (not shown).

**[0011]** FIG. 2 shows a schematic diagram of a possible microphone array MA for the sound capture system of FIG. 1. In particular, microphone array MA comprises the Q microphones Mic1, Mic2, ... MicQ of FIG. 1 mounted on the surface of an acoustically rigid sphere RS in a "truncated icosahedron" pattern. Each microphone Mic1, Mic2, ... MicQ in microphone array MA generates one of the audio signals  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  that is transmitted to matrixing unit MU of FIG. 1 via some suitable (e.g., wired or wireless) connection (not shown in FIG. 2). The continuous spherical sensor may be replaced by a discrete spherical array, in particular when the subsequent processing is digital-signal processing.

**[0012]** Referring again to FIG. 1, beamformer BF exploits the geometry of the spherical array of FIG. 2 and relies on the spherical harmonic decomposition of the incoming sound field by matrixing unit MU to construct a desired spatial response. In beamformer BF, steering unit SU generates (according to  $Y_{m,n}^{+\sigma}(\theta_{Des}, \varphi_{Des})$ ) steered spherical harmonics  $Y_{0,0}^{+1}(\theta_{Des}, \varphi_{Des})$ ,  $Y_{1,0}^{+1}(\theta_{Des}, \varphi_{Des})$ , ...  $Y_{m,n}^{+\sigma}(\theta_{Des}, \varphi_{Des})$  from the spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$ , which are further processed by weighting unit WU and summing element SE. Beamformer BF can provide continuous steering of the beampattern in 3-D space by changing a few scalar multipliers, while the filters determining the beampattern itself remain constant. The shape of the beampattern is invariant with respect to the steering direction. Beamformer BF needs only one filter per spherical harmonic (in the weighting unit WU), rather than per microphone as in known beamforming concepts, which significantly reduces the computational cost.

**[0013]** The sound capture system of FIG. 1 with the spherical array geometry of FIG. 2 enables accurate control over the beampattern in 3-D space. In addition to pencil-like beams, the sound capture system can also provide multi-direction beampatterns or toroidal beampatterns giving uniform directivity in one plane. These properties can be useful for applications such as general multichannel speech pick-up, video conferencing, and direction of arrival (DOA) estimation. It can also be used as an analysis tool for room acoustics to measure, e.g., directional properties of the sound field. The sound capture system of FIG. 1 offers another advantage: it supports decomposition of the sound field into mutually orthogonal components, the eigenbeams (i.e., spherical harmonics) that can also be used to reproduce the sound field. The eigenbeams are also suitable for wave field synthesis (WFS) methods that enable spatially accurate sound reproduction in a fairly large volume, allowing for reproduction of the sound field that is present around the recording sphere. This allows for all kinds of general real-time spatial audio.

**[0014]** A circuit that provides the beamforming functionality is shown in detail in FIG. 3. The modal beamformer circuit of FIG. 3 receives the Q audio signals  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  provided by microphones Mic1, Mic2, ... MicQ, transforms the audio signals  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  into the spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$ , and steers the spherical harmonics. The circuit of FIG. 3 may be realized by hardware (and software) components that (together) build matrixing unit MU and the modal beamformer, which includes steering unit SU, modal weighting unit WU, and summing element SE. Matrixing unit MU and steering unit SU include coefficient elements CE that multiply the respective input signals with given coefficients and adders AD that sum up the input signals multiplied with coefficients so that the audio signals  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka) \dots S_Q(\theta_Q, \varphi_Q, ka)$  are decomposed into the eigenbeams, i.e., the spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$ , which are then processed to provide the steered spherical harmonics  $Y_{0,0}^{+1}(\theta_{Des}, \varphi_{Des})$ ,  $Y_{1,0}^{+1}(\theta_{Des}, \varphi_{Des})$ , ...  $Y_{m,n}^{+\sigma}(\theta_{Des}, \varphi_{Des})$ . Modal weighting unit WU includes delay elements DE, coefficient elements CE, and adders AD, which are connected to form FIR filters for weighting. The output signals of these FIR filters are summed up by summing element SE.

**[0015]** Matrixing unit MU in the modal beamformer of FIG. 3 is responsible for decomposing the sound field, which is picked up by microphones Mic1, Mic2, ... MicQ and decomposed into the different eigenbeam outputs, i.e., the spherical harmonics  $Y_{0,0}^{+1}(\theta, \varphi)$ ,  $Y_{1,0}^{+1}(\theta, \varphi)$ , ...  $Y_{m,n}^{+\sigma}(\theta, \varphi)$ , corresponding to the zero-order, first-order, and second-order spherical

harmonics. This can also be seen as a transformation, where the sound field is transformed from the time or frequency domain into the "modal domain". To simplify a time-domain implementation, one can also work with the real and imaginary parts of the spherical harmonics. This will result in real-value coefficients, which are more suitable for a time-domain implementation. If the sensitivity equals the imaginary part of a spherical harmonic, then the beampattern of the corresponding array factor will also be the imaginary part of this spherical harmonic. To compensate for this frequency dependence, weighting unit WU may be implemented accordingly. Steering unit SU allows for steering the look direction by the angles  $\theta_{Des}$  and  $\varphi_{Des}$ . Weighting unit WU compensates for a frequency-dependent sensitivity over the modes (eigenbeams), i.e., modal weighting over frequency, to the effect that the modal composition is adjusted, e.g., equalized. Equalizing is used to compensate for deficiencies of the microphone array, e.g., self-noise of the microphones, location errors of the microphones at the surface of the sphere, and other electrical and mechanical drawbacks. Summation node SE performs the actual beamforming for the sound capture system by summing up the weighted harmonics to yield the beamformer output  $OUT = \psi(\theta_{Des}, \varphi_{Des})$ , i.e., the auditory scene.

**[0016]** Due to self-noise amplification, the order of a modal beamformer has to be reduced toward low frequencies, leading to a gradually decreasing directivity pattern with decreasing frequency. Regularization of the radial filter is configured such that, for example, the white noise gain will not fall below a given limit (e.g.,  $WNG_{dBMin} = -10$  [dB] ( $\pm 3$  [dB])) to keep the robustness, i.e., the self-noise amplification, within a tolerable range, and a constant directivity in look direction over frequency, such as 0 [dB], will be reached. By doing this, an optimum balance between robustness and directivity will result, leading to a modal beamformer with enhanced properties in which the directivity of the modal beamformer is enhanced by keeping the transfer function in look direction at a frequency-independent constant value and a minimum threshold of robustness. Regularization may be achieved by adapting the weighting coefficients of the FIR filters in weighting unit WU to an optimum.

**[0017]** But before going into detail on the regularization process, some general issues are discussed, in particular issues with regard to the measurement of the acoustic wave field via a rigid spherical microphone array. In general, sound pressure values  $p_a(\theta_q, \varphi_q)$  can be described by way of the Fourier-Bessel series truncated to the  $M^{th}$  order at positions  $\theta_q, \varphi_q$  of the  $Q$  microphones located at radius  $a$ , in which  $1 \leq q \leq Q$ , as follows:

$$p_a(\theta_q, \varphi_q) = \sum_{m=0}^M W_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^\sigma Y_{m,n}^\sigma(\theta_q, \varphi_q)$$

in which:  $p_a(\theta_q, \varphi_q)$  is the sound pressure measured by the  $q^{th}$  microphone located at position(s)  $\theta_q, \varphi_q$  at the surface of a sphere having a radius  $a$ ;  $W_m(ka)$  is the radial function that describes the acoustic wave field in the vicinity of the sphere center, i.e., in a certain distance from the center; and  $B_{m,n}^\sigma$  is the complex,  $m^{th}$  order,  $n^{th}$  degree, ambisonic component that completely describes wave fields up to the  $M^{th}$  order.

**[0018]** The above equation can be rewritten as:

$$\begin{aligned} p_{a \ Q \times 1} &= Y_{Q \times N} W_{N \times N} B_{N \times 1}, \\ W_{N \times N} &= \underset{(2m+1)}{\text{diag}\{W_m(ka)\}} \underset{(2M+1)}{\text{diag}\{W_0(ka), \dots, W_m(ka), \dots, W_m(ka), \dots, W_m(ka)\}}, \\ \text{diag}(\cdot) &= \text{diagonal matrix} \end{aligned}$$

**[0019]** By rearranging the previous equation, the ambisonic components up to  $M^{th}$  order can be calculated from the  $Q$  microphone signals:

$$\begin{aligned}
 B &= W^{-1}(Y^T Y)^{-1} Y^T p_a, \\
 Y^+ &= Y^T Y^{-1} Y^T = \text{pseudo inverse of } Y, \\
 B &= \text{diag}\{W_m^{-1}\} Y^+ p_a, \text{ in which} \\
 &\hspace{15em} (2m+1) \\
 \text{diag}\{W_m^{-1}\} &= \text{diag}\{EQ_m(ka)\} = \text{diag}\left(\left[\frac{1}{W_0(ka)}, \dots, \frac{1}{W_m(ka)}, \dots, \frac{1}{W_m(ka)}, \dots, \right.\right. \\
 &\hspace{10em} (2M+1) \\
 &\hspace{10em} \left.\left. \frac{1}{W_M(ka)}, \dots, \frac{1}{W_M(ka)}\right]\right), \text{ and} \\
 \text{diag}\{EQ_m(ka)\} &= \text{diagonal matrix having the radial equalizing functions } EQ_m(ka), \\
 &\hspace{10em} \text{in which } 0 \leq m \leq M.
 \end{aligned}$$

[0020] An arrangement for extracting the N ambisonic components B from the wave field  $p_a$  is illustrated in FIG. 4. The room and, thus, the spherical harmonics  $Y^{+1}_{0,0}(\theta, \varphi)$ ,  $Y^{+1}_{1,0}(\theta, \varphi)$ , ...  $Y^{+\sigma}_{m,n}(\theta, \varphi)$  are sampled by way of matrix  $Y^+$  at the position(s)  $\theta_q, \varphi_q$  with the Q microphones, in which:

$$1 \leq q \leq Q,$$

and

$$M = \lfloor \sqrt{Q} - 1 \rfloor$$

so that the  $N = (M+1)^2$  ambisonic components of  $M^{\text{th}}$  order can be calculated from the samples.

[0021] Combining the Q microphone signals ( $1 \leq q \leq Q$ ), i.e.,  $S_1(\theta_1, \varphi_1, ka)$ ,  $S_2(\theta_1, \varphi_2, ka)$  ...  $S_Q(\theta_Q, \varphi_Q, ka)$ , by way of matrix  $Y^+$  into N output signals, which correspond to signals that would have been obtained when a wave field is sampled with N microphones having a certain directivity, can be seen as a transformation from the time domain into the spatial domain. By way of a radial equalizing function  $EQ_m(ka)$  the thereby generated spherical harmonic signals are then weighted to provide frequency-independent normalized-to-1 ambisonic components  $B^{\sigma}_{m,n}$  or the ambisonic signals B.

[0022] Referring now to FIG. 5, the derivation of the radial function  $W_m(ka)$  of a rigid closed sphere with microphones arranged on the sphere's surface can be described as follows: at the surface of a rigid closed sphere, velocity  $v_a$  is zero, i.e.,  $v_a(\theta_q, \varphi_q, ka) = 0$ .

[0023] Therefore, the related sound field is defined solely by the pressure distribution  $p_a(\theta_q, \varphi_q)$  on the sphere's surface, which can be easily measured by sound pressure sensors (microphones). Mathematically, the underlying, physically logical condition that  $v_a(\theta_q, \varphi_q, ka) = 0$  holds at the surface of a rigid body can be met when inner sources (i.e., sources inside the measurement sphere) and outer sources (i.e., sources outside the measurement sphere) are superposed, as illustrated in FIG. 5. For instance, the outer sources serve to model the scattered field occurring at the surface of a scattered sphere. Based on the general form of the Bessel series,

$$\begin{aligned}
 P(r, \omega) &= S(j\omega) \left( \sum_{m=0}^{\infty} j^m j_m(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta, \phi) + \right. \\
 &\quad \left. \sum_{m=0}^{\infty} j^m h_m^{(2)}(kr) \sum_{0 \leq n \leq m, \sigma = \pm 1} A_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta, \phi) \right),
 \end{aligned}$$

in which  $B_{m,n}^{\sigma}$  are the weighting coefficients (ambisonic coefficients) that relate to the spherical Bessel function of the

1<sup>st</sup> kind  $j_m(kr)$  and that describe the pervasive wave field (plane wave);  $A_{m,n}^{\sigma}$  are the weighting coefficients that relate

to the spherical Hankel function of the 2<sup>nd</sup> kind  $h_m^{(2)}(kr)$  and that describe the outgoing spherical wave field (spherical wave), eventually representing the scattered wave field at the surface of the solid sphere;  $P(r,\omega)$  is the sound pressure spectrum at the position  $r = r, \theta, \phi$ ;  $S(j\omega)$  is the input signal in the spectral domain;  $j$  is the imaginary unit for complex numbers with  $j = \sqrt{-1}$ ;  $j_m(kr)$  is the spherical Bessel function of the 1<sup>st</sup> kind,  $m^{\text{th}}$  order;  $h_m^{(2)}(kr)$  is the spherical Bessel function of the 2<sup>nd</sup> kind,  $m^{\text{th}}$  order; and based on the assumption that the outer sources provide incoming plane waves (indicated by the index "Inc"), the wave field generated by the outer sources that moves toward the center and thus toward the sphere's surface can be described as follows:

$$p_{Inc}(\theta_q, \phi_q, ka) = \sum_{m=0}^{\infty} j^m j_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta_q, \phi_q), \text{ in which}$$

$$p_{Inc}(\theta_q, \phi_q, ka) = \text{Sound pressure received by the } q\text{-th microphone at position } (\theta_q, \phi_q), \text{ and generated by the outer source.}$$

**[0024]** Furthermore, it is required that the velocity at the sphere's surface, i.e.,  $r = a$ , is zero:

$$V_{Inc}(\theta_q, \phi_q, ka) + V_{Scat}(\theta_q, \phi_q, ka) = 0 \text{ or}$$

$$V_{Scat}(\theta_q, \phi_q, ka) = -V_{Inc}(\theta_q, \phi_q, ka), \text{ in which}$$

$V_{Inc}(\theta_q, \phi_q, ka)$  = velocity at the  $q^{\text{th}}$  microphone at position  $(\theta_q, \phi_q)$  caused by the plane wave from the outer source, and

$V_{Scat}(\theta_q, \phi_q, ka)$  = velocity at the  $q^{\text{th}}$  microphone at position  $(\theta_q, \phi_q)$  caused by the spherical wave from the outer source.

**[0025]** Differentiating the previous equation with respect to  $r$  or  $a$  leads to

$$v_{Scat}(\theta_q, \phi_q, ka) = -k \sum_{m=0}^{\infty} j^m j'_m(ka) \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta_q, \phi_q),$$

$j'_m(ka)$  = 1<sup>st</sup> derivative of the spheric Bessel function of the 1<sup>st</sup> kind,  $m^{\text{th}}$  order.

**[0026]** Applying the Euler equation to this leads to:

$$v(\theta, \phi, kr) = \frac{1}{j\rho ck} \nabla p(\theta, \phi, kr),$$

$$\nabla = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \phi} \frac{\partial}{\partial \phi} \sin \phi \frac{\partial}{\partial \phi} + \frac{1}{r^2 \sin^2 \phi} \frac{\partial^2}{\partial \theta^2}, \text{ i.e.,}$$

$$\nabla = \text{Nabla operator expressed as spherical coordinates,}$$

$$\rho = \text{Medium (air) density in } \left[ \frac{kg}{m^3} \right] \rho_0 = 1.292 \left[ \frac{kg}{m^3} \right],$$

$$c = \text{Sonic speed in } \left[ \frac{m}{s} \right] v_0 = 343 \left[ \frac{m}{s} \right],$$

**[0027]** The Euler equation links the sound velocity  $v(\theta_q, \phi_q, ka)$  to the sound pressure  $p(\theta_q, \phi_q, ka)$  and the fact that sound velocity  $v(\theta_q, \phi_q, ka)$  and sound pressure  $p(\theta_q, \phi_q, ka)$  can be derived by weighting spherical harmonics according to the Fourier-Bessel series:

$$v(\theta, \phi, kr) = \sum_{m=0}^{\infty} \sum_{0 \leq n \leq m, \sigma = \pm 1} v_{m,n}^{\sigma}(kr) Y_{m,n}^{\sigma}(\theta_q, \phi_q), \text{ in which}$$

$$v_{m,n}^{\sigma}(kr) = \text{Sound coefficients of corresponding, spherical harmonics,}$$

so that the following relationship of sound velocity  $v(\theta_q, \Phi_q, ka)$  and sound pressure  $p(\theta_q, \Phi_q, ka)$  at the surface of a rigid sphere applies:

$$p(\theta, \phi, kr) = j\rho c \sum_{m=0}^{\infty} \frac{h_m^{(2)}(kr)}{h_m^{(2)'}(kr)} \sum_{0 \leq n \leq m, \sigma = \pm 1} v_{m,n}^{\sigma}(kr) Y_{m,n}^{\sigma}(\theta, \phi), \text{ in which}$$

$$h_m^{(2)'}(kr) = 1^{\text{st}} \text{ derivative of the spherical Bessel function of the } 2^{\text{nd}} \text{ kind, } m^{\text{th}} \text{ order.}$$

**[0028]** The sound velocity coefficients  $v_{m,n}^{\sigma}$  of an incoming plane wave can be calculated from the ambisonic coefficients  $B_{m,n}^{\sigma}$  as follows:

$$v_{m,m}^{\sigma}(ka) = -kj^m j'_m(ka) B_{m,n}^{\sigma}.$$

**[0029]** From the two previous equations, a simplified relationship can be provided for the sound pressure  $p_{\text{scat}}(\theta_q, \Phi_q, ka)$  that results from the sound field of the spherical waves distributing inner sound sources and that can be measured on the sphere's surface ( $r = a$ ) at the positions  $(\theta_q, \Phi_q)$  where the  $q$  pressure sensors (microphones) are arranged, thereby neglecting the constants  $j\rho c$  and  $4\pi$ :

$$p_{\text{scat}}(\theta_q, \phi_q, ka) = -\sum_{m=0}^{\infty} \frac{j^m j'_m(ka) h_m^{(2)}(ka)}{h_m^{(2)'}(ka)} \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta_q, \phi_q)$$

**[0030]** Superimposing the wave fields, e.g., the sound pressures of the inner and outer sources, leads to the sound pressures occurring at the surface of a rigid sphere having a radius  $a$ :

$$p(\theta_q, \phi_q, ka) = p_{\text{inc}}(\theta_q, \phi_q, ka) + p_{\text{scat}}(\theta_q, \phi_q, ka) \rightarrow$$

$$p(\theta_q, \phi_q, ka) = \sum_{m=0}^{\infty} j^m \left[ j_m(ka) - \frac{j'_m(ka) h_m^{(2)}(ka)}{h_m^{(2)'}(ka)} \right] \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta_q, \phi_q)$$

which can be simplified by way of the Wronskian relation:

$$j_m(kr) h_m^{(2)'}(kr) j'_m(kr) h_m^{(2)}(kr) = \frac{j}{(kr)^2},$$

to read as:

$$W_m(ka) = \frac{j^{(m-1)}}{(ka)^2 h_m^{(2)'}(ka)},$$

so that:

$$p(\theta_q, \phi_q, ka) = \sum_{m=0}^{\infty} \frac{j^{m-1}}{(ka)^2 h_m^{(2)}(ka)} \sum_{0 \leq n \leq m, \sigma = \pm 1} B_{m,n}^{\sigma} Y_{m,n}^{\sigma}(\theta_q, \phi_q).$$

5  
 10  
**[0031]** An accordingly calculated magnitude frequency response for the radial functions  $w_m(ka) = 1/EQ_m(ka)$  for a sphere radius of  $a=0.9m$  in a spectral range of 50Hz to 6700Hz for orders  $m$  up to  $M=10$  is shown in FIG. 6. The corresponding radial equalizing function  $EQ_m(ka)$  for orders  $m$  up to  $M=4$ , is depicted in FIG. 7. The equations outlined above provide a least-square solution that offers the smallest-mean-squared error, but cannot be used per se in connection with small or very small  $w_m(ka)$  values. However, this is the case at higher orders  $m$  and/or lower frequencies  $f$  so that instabilities may occur due to amplified noise of the sensors or measurement system, positioning errors of the microphones, or irregularities in the frequency characteristic, which may deteriorate the results.

15  
**[0032]** By introducing a regularization functionality, which means limiting the radial equalizing function  $EQ_m(ka)$  by way of a regularization function  $T_m(ka)$ , e.g., to a maximum gain, these drawbacks can be overcome, whereby filters known as Tikhonov filters may be used. The following applies:

$$EQ_m(ka) = \frac{T_m(ka)}{W_m(ka)},$$

$$T_m(ka) = \frac{|W_m(ka)|^2}{|W_m(ka)|^2 + \epsilon^2},$$

$$\epsilon = \begin{cases} \epsilon = 0, \\ \epsilon = \infty, \end{cases}$$

$$T_m(ka) = \text{Regularization function with } m = [0, \dots, M],$$

$$\epsilon = \text{Regularization parameter } \epsilon \geq 0.$$

20  
 25  
 30  
**[0033]** If  $\epsilon = 0$ , the system works as a least-square beamformer (ideal case as shown above, i.e., without any regularization, which leads to the solution with the highest directivity but also with the least robustness). If  $\epsilon = \infty$ , the system works as a delay-and-sum beamformer, which delivers the maximum possible robustness but the least directivity. The radial equalizing functions  $EQ_m(ka)$  can be further simplified to read as:

$$EQ_m(ka) = \frac{T_m(ka)}{W_m(ka)} = \frac{W_m(ka)^*}{|W_m(ka)|^2 + \epsilon^2},$$

$$W_m(ka)^* = \text{Complex conjugate of matrix } W_m(ka).$$

35  
 40  
**[0034]** Thus, with regularization parameters  $\epsilon(ka)$  or  $\epsilon(\omega)$  one can control the modal beamformer to exhibit a certain robustness with respect to the inherent noise that is amplified with  $w_m(ka)$ , in particular at lower frequencies.

**[0035]** In order to calculate appropriate values for the regularization parameters  $\epsilon(ka)$  or  $\epsilon(\omega)$ , a parameter called susceptibility  $K(\omega)$  or its reciprocal white noise gain  $WNG(\omega)$  may be used. For instance, white noise gain  $WNG(\omega)$  addresses most effects and problems caused by microphone noise, changes in the transfer function, and variations of the microphone positions, so that it is representative of the sensitivity of the beamformer. A white noise gain  $WNG(\omega) > 0$  [dB] characterizes a sufficient suppression of uncorrelated errors and is thus indicative of a robust system behavior, while a white noise gain  $WNG(\omega) < 0$  [dB] is indicative of an amplification of the noise and is therefore indicative of an increasingly unstable system behavior.

45  
 50  
**[0036]** The white noise gain  $WNG(\omega)$  represents the ratio of the energy of the useful signal provided by the microphone array to the energy of the noise signal provided by the microphone array and can be expressed as:

$$WNG(\omega) = \frac{|d(\theta_0, \varphi_0, \omega)|^2}{\frac{1}{Q^2} \sum_{q=1}^Q |H_q(\theta_q, \varphi_q, \omega)|^2}, \text{ in which}$$

$d(\theta_0, \varphi_0, \omega)$  = output signal of the beamformer having a look direction of  $(\theta_0, \varphi_0)$   
 $H_q(\theta_q, \varphi_q, \omega)$  = noise signal of the microphones caused by inherent noise.

**[0037]** The useful signal  $d(\theta_0, \varphi_0, \omega)$  output by the microphone array and the output signal of the beamformer having the required look direction can be described as follows:

$$d(\theta_0, \varphi_0, \omega) = \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} T_m(\omega) Y_{m,n}^\sigma(\theta_0, \varphi_0), \text{ in which}$$

$$N = (M + 1)^2 = \text{Number of spherical harmonics,}$$

$$Y_{m,n}^\sigma(\theta_0, \varphi_0) = \text{Spherical harmonic of } m\text{-th order, } n\text{-th degree}$$

$$\text{in look direction of the beamformer } (\theta_0, \varphi_0),$$

$$T_m(\omega) = \text{Tikhonov regularization filter.}$$

**[0038]** The noise signal of the  $q^{\text{th}}$  microphone of the microphone array over frequency caused by the inherent noise of the microphone is represented by  $H_q(\theta_q, \varphi_q, \omega)$ , which is:

$$H(\theta_q, \varphi_q, \omega) = \frac{1}{N^2} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} EQ_m(\omega) Y_{m,n}^\sigma(\theta_q, \varphi_q).$$

**[0039]** The frequency-dependent white noise gain  $WNG_{\text{dB}}(\omega)$  in [dB] is:

$$WNG_{\text{dB}}(\omega) = 10 \log_{10}(WNG(\omega)).$$

**[0040]** Thereby, the maximum white noise gain  $WNG_{\text{dB}}(\omega)$  for a modal beamformer is as follows:

$$WNG_{\text{dB}_{\text{Max}}} = 10 \log_{10}(Q),$$

which is, e.g.,  $\approx 15$  [dB] for  $Q = 32$ .

**[0041]** Furthermore, it has been found that best results are achieved when an array gain  $G(\omega)$  is maximum and the white noise gain  $WNG_{\text{dB}}(\omega)$  is above a given minimum value, for instance,  $WNG_{\text{dB}}(\omega) > -10$  [dB]. The array gain  $G(\omega)$  can be calculated according to:

$$G(\omega) = \frac{|\psi(\theta_0, \varphi_0, \omega)|^2}{\frac{1}{4\pi} \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\pi} |\psi(\theta, \varphi, \omega)|^2 \sin(\theta) d\theta d\varphi},$$

$$\psi(\theta, \varphi, \omega) = \text{Directivity of the microphone array.}$$

[0042] In words, the array gain  $G(\omega)$  is the ratio of the energy of sound coming from the look direction of the beamformer to the energy of omnidirectionally incoming sound.

[0043] The directivity  $\Psi(\theta_0, \varphi_0, \omega)$  for incoming sound from the look direction can be described as:

$$\begin{aligned} \psi(\theta_0, \varphi_0, \omega) &= \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} C_{m,n}^{\sigma}(\theta, \varphi, \omega) Y_{m,n}^{\sigma}(\theta_0, \varphi_0, \omega), \text{ in which} \\ C_{m,n}^{\sigma}(\theta, \varphi, \omega) &= \text{diag} \left( \frac{T_m(\omega)}{W_m(\omega)} \right) Y^+ p_{\alpha}, \text{ and} \\ C_{m,n}^{\sigma}(\theta, \varphi, \omega) &= \text{Weighting factors of the modal beamformer,} \end{aligned}$$

while the directivity  $\Psi(\theta_0, \varphi_0, \omega)$  for omnidirectionally incoming sound can be described as:

$$\psi(\theta, \varphi, \omega) = \frac{4\pi}{Q} \frac{1}{N} \sum_{m=0}^M \sum_{0 \leq n \leq m, \sigma = \pm 1} C_{m,n}^{\sigma}(\theta, \varphi, \omega) Y_{m,n}^{\sigma}(\theta, \varphi, \omega).$$

[0044] Then, the frequency-dependent array gain  $G(\omega)$  is:

$$G_{dB}(\omega) = 10 \log_{10}(G(\omega)).$$

[0045] The array gain  $G(\omega)$  is a measure for the improvement in the acoustic signal-to-noise ratio SNR, based on the directivity of the modal beamformer for sound coming from the look direction of the beamformer. The achievable maximum array gain  $G_{dBMax}(\omega)$  in [dB] is:

$$G_{dBMax}(\omega) = 20 \log_{10}(M + 1).$$

[0046] For instance, when  $M = 4$ , then the achievable maximum array gain  $G_{dBMax}(\omega)$  is approximately 14dB.

[0047] Referring now to FIG. 8, an exemplary iterative process of adapting the parameters of a modal beamformer is described in detail. In an initializing step 1, parameters required for calculation are set to a starting value or a constant value, as the case may be. The following parameters may be set to, for instance:

WNG parameter

- Minimum white noise gain threshold  $WNG_{dBMin}(\omega)$ , which is not undercut by the regularized modal beamformer; for instance,  $WNG_{dBMin} = -10$ [dB].
- Offset  $\Delta WNG_{dB}$  in [dB], by which the minimum white noise gain threshold  $WNG_{dBMin}(\omega)$  is overcut or undercut during the adaptation process; for instance,  $\Delta WNG_{dB} = 0.5$ dB.

Regularization parameter  $\varepsilon(\omega)$

- Maximum regularization parameter  $\varepsilon_{Max}$ , which is the upper limit for the regularization parameter  $\varepsilon(\omega)$ ; for instance,  $\varepsilon_{Max} = 1$ .
- Step size by which the regularization parameter  $\varepsilon(\omega)$  is increased or decreased.

Frequency  $\omega$

- Start value of the (angular) frequency for the adaptation process; for instance,  $\omega = 2\pi 1$ [HZ],
- Step size by which the (angular) frequency is increased or decreased when the adaption is completed at a

certain frequency; for instance,  $\Delta\omega = 2\pi 1[\text{Hz}]$ ,

- Maximum (angular) frequency at which an adaptation is performed; for instance,  $\omega_{Max} = \pi f_s[\text{Hz}]$ .

**[0048]** Then the adaptation process is started in step 2. In step 3, the regularization parameter is set to, e.g.,  $\varepsilon(\omega) = 0$  for the current frequency  $\omega$  under investigation. Regularization provides the ability to achieve a robust system by way of adjusting the regularization parameter  $\varepsilon(\omega)$ . This is a trade-off between a higher robustness, i.e., a higher white noise gain  $WNG_{dB}(\omega)$ , and less directivity in look direction  $\psi(\theta_0, \varphi_0, \omega)$ , i.e., a decreasing array gain  $G_{dB}(\omega)$ . If the regularization parameter is set to  $\varepsilon(\omega) = 0$ , the adaptation process begins with the maximum directivity  $G_{dBMax}(\omega)$  and is then decreased by the increasing regularization parameter  $\varepsilon(\omega)$  until the desired white noise gain threshold  $WNG_{dBMin}$  is no more undercut.

**[0049]** Steps 4, 5, and 6 serve to calculate the white noise gain  $WNG_{dB}(\omega)$ . In step 4, the regularization filter  $T_m(\omega)$  or  $T_m(ka)$ , is calculated as outlined above using regularization parameter  $\varepsilon(\omega)$ . In step 5, the transfer function  $EQ_m(\omega)$  is calculated as outlined above using the current version of the transfer function  $T_m(\omega)$  of the regularization filter or the current version of the regularization parameter  $\varepsilon(\omega)$ . In step 6, the white noise gain  $WNG_{dB}(\omega)$  is calculated as outlined above using the transfer function  $EQ_m(\omega)$  and the current version of the transfer function  $T_m(\omega)$  of the regularization filter (regularization function). Steps 4 and 5 may be taken simultaneously or in opposite order.

**[0050]** In the following step 7, the current white noise gain  $WNG_{dB}(\omega)$  is compared with the predetermined threshold  $WNG_{dBMin}$  so that according to step 8 or 9,

$$\varepsilon(\omega) = \begin{cases} \varepsilon(\omega) - \Delta\varepsilon, & \text{if } WNG_{dB}(\omega) > WNG_{dBMin} \\ \varepsilon(\omega) + \Delta\varepsilon, & \text{otherwise.} \end{cases}$$

**[0051]** In step 10, the directivity  $\psi(\theta_0, \varphi_0, \omega)$  of the modal beamformer is calculated for sound coming from the look direction using the transfer function  $EQ_m(\omega)$  provided in step 5.

**[0052]** In step 11, the transfer function of the equalizing filter, the equalizing function  $EQ_m(\omega)$ , is scaled according to:

$$EQ_m(\omega) = \frac{EQ_m(\omega)}{\psi(\theta_0, \varphi_0, \omega)}.$$

**[0053]** In step 12, the current white noise gain  $WNG_{dB}(\omega)$  is compared with the predetermined white noise gain threshold  $WNG_{dBMin}(\omega)$ , and it is checked to see if the regularization parameter  $\varepsilon(\omega)$  has reached its maximum according to  $(|WNG_{dBMin} - WNG_{dB}(\omega)| > \Delta WNG)$  and  $(\varepsilon(\omega) \leq \varepsilon_{Max})$ . If both requirements are met, i.e., if  $(|WNG_{dBMin} - WNG_{dB}(\omega)| > \Delta WNG) \& (\varepsilon(\omega) \leq \varepsilon_{Max})$ , the adaptation process is not yet finished, resulting in jumping back to step 3 and starting again with an updated regularization parameter  $\varepsilon(\omega)$ .

**[0054]** Otherwise, i.e., if the adaptation process for the current angular frequency  $\omega$  has been completed so that the current equalizing function  $EQ_m(\omega)$  has been limited to the given threshold or if the current regularization parameter has reached its maximum, the angular frequency  $\omega$  is incremented according to  $\omega = \omega + \Delta\omega$  in step 13, which is followed by step 14.

**[0055]** In step 14, the current angular frequency  $\omega$  is checked to see if it has reached its maximum value  $\omega_{Max}$ . If  $\omega < \omega_{Max}$ , the process jumps back to step 2 using the current angular frequency  $\omega$ . Otherwise, i.e., if the equalizing filter has been adapted for the complete set of frequencies, the filter coefficients are outputted in step 15.

**[0056]** Referring to FIGS. 9 through 16, measurements made with an exemplary arrangement in combination with an exemplary adaptation method are described in detail. The arrangement includes a sphere having a radius of  $a = 0.09$  [m] and the shape of an obtuse icosahedron, which is a blend of two platonic solids, i.e., an icosahedron and a dodecahedron. The number of microphones arranged on the sphere is  $Q = 32$ . The directivity characteristic of the beamformer is a 4<sup>th</sup>-order cardioid and the minimum white noise gain  $WNG_{dB}(\omega)$  used in the adaptation process is -10 [dB].

**[0057]** FIG. 9 illustrates a regularization parameter over frequency  $\varepsilon(\omega)$  for a common 4<sup>th</sup>-order modal beamformer. As can be seen from FIG. 9 with regularization, i.e., limiting the maximum directivity index for frequencies up to, for instance, 750 [Hz], values above a minimum lower threshold  $WNG_{dBMin}$  of -10 [dB] may be maintained. Above 750 [Hz], the exemplary beamformer exhibits the desired directivity of a 4<sup>th</sup>-order cardioid. FIG. 10 illustrates the corresponding white noise gain  $WNG$  for the above-mentioned 4<sup>th</sup>-order beamformer, which supports the findings in connection with the diagram of FIG. 9. The corresponding directivity index DI and the array gain  $G_{dB}(\omega)$  as shown FIG. 11 illustrate that the maximum array gain  $G_{dB}(\omega)$  is more or less below 10 [dB] depending on the frequency.

**[0058]** However, applying the adapted regularization filter ( $T_m(\omega)$ ) described herein causes a monotonic decrease of the array gain  $G_{dB}(\omega)$  down to 7.5 [dB] at 20 [Hz] as shown in FIG. 11. The magnitude frequency responses of the thereby

applied M regularization filter  $T_m(\omega)$  is shown in FIG. 12, and its corresponding frequency-independent phase characteristic is illustrated in FIG. 13.

**[0059]** Further applying the optimized radial equalizing filter ( $EQ_m(\omega)$ ) causes an improved regularized equalizing filter whose magnitude frequency response is depicted in FIG. 14 and whose phase frequency response is depicted in FIG. 15. The directivity of the corresponding improved beamformer at frequencies above 650 [Hz] is a 4<sup>th</sup>-order cardioid, between 300 [Hz] and 650 [Hz] a 3<sup>rd</sup>-order cardioid, between 70 [Hz] and 300 [Hz] a 2<sup>nd</sup>-order cardioid, and below 70 [Hz] a 1<sup>st</sup>-order cardioid. FIG. 16 depicts the resulting directivity of the beamformer outlined above in look directivity  $\psi(\theta_0, \varphi_0, \omega)$  as amplitudes over frequency.

**[0060]** While exemplary embodiments are described above, it is not intended that these embodiments describe all possible forms of the invention. The words used in the specification are words of description rather than limitation, and it is understood that various changes may be made without departing from the spirit and scope of the invention. Additionally, the features of various implementing embodiments may be combined to form further embodiments of the invention.

## Claims

1. A method for generating an auditory scene, comprising:

receiving eigenbeam outputs, the eigenbeam outputs having been generated by decomposing a plurality of audio signals, each audio signal having been generated by a different microphone of a microphone array, wherein each eigenbeam output corresponds to a different eigenbeam for the microphone array, and the microphones are arranged on a rigid or open sphere; and

generating the auditory scene based on the eigenbeam outputs and their corresponding eigenbeams, wherein generating the auditory scene comprises applying a weighting value to each eigenbeam output to form a steered eigenbeam output; and

combining the weighted eigenbeams to generate the auditory scene, wherein generating the auditory scene further comprises applying a regularized equalizer filter to each eigenbeam output or steered eigenbeam output, the regularized equalizer filter(s) being configured to compensate for acoustic deficiencies of the microphone array and having a regularized equalization function.

2. The method of claim 1 wherein the regularized equalization function is a radial equalization function that comprises the quotient of a regularization function limiting the radial equalization function and a radial function describing an acoustic wave field in the vicinity of the surface of the rigid sphere or the center of the open sphere.

3. The method of claim 2 wherein the regularization function is the quotient of the absolute value of the square of the radial function and the sum of the absolute value of the square of the radial function and a regularization parameter, the regularization parameter being set to a value greater than 0 and smaller than a maximum value that is smaller than infinity.

4. The method of claim 3 wherein the maximum value of the regularization parameter is 1.

5. The method of claim 3 or 4 wherein the regularization parameter depends on a susceptibility parameter that is the reciprocal of a white noise gain parameter, the white noise gain parameter being greater than a minimum white noise gain parameter that is not undercut by the equalizer filter.

6. The method of claim 5 wherein the minimum white noise gain parameter is -10 [dB].

7. The method of any one of claims 3 through 6 wherein the regularization parameter is adapted in an iterative process.

8. The method of claim 7 wherein, for a given frequency, the iterative process comprises:

setting at least the minimum white noise gain parameter and the regularization parameters to a starting value or a constant value; and

calculating the white noise gain, the regularization function, and the radial equalization function; and comparing the calculated white noise gain parameter with the set minimum white noise gain parameter; and calculating the directivity for sound coming from the look direction using the radial equalization function; and scaling the radial equalization function; and

comparing the calculated white noise gain with the set minimum white noise gain and checking if the regularization parameter has reached its maximum; if both requirements are met, the adaptation process is not yet finished, resulting in jumping back and starting again with an updated regularization parameter; otherwise the process for the current frequency has been completed and the frequency is incremented; and  
 5 checking if the current frequency has reached its maximum value; if the frequency has not reached its maximum, the process jumps back and starts again with another frequency; otherwise the filter coefficients are outputted.

9. The method of claim 8 wherein the iterative process comprises an offset white noise gain parameter by which the minimum white noise gain parameter is overcut or undercut at maximum during adaptation.

10. A modal beamformer system for generating an auditory scene, comprising:

a steering unit that is configured to receive eigenbeam outputs, the eigenbeam outputs having been generated by decomposing a plurality of audio signals, each audio signal having been generated by a different microphone of a microphone array, wherein each eigenbeam output corresponds to a different eigenbeam for the microphone array, and the microphones are arranged on a rigid or open sphere; and  
 15 a weighting unit that is configured to generate the auditory scene based on the eigenbeam outputs and their corresponding eigenbeams, wherein generating the auditory scene comprises applying a weighting value to each eigenbeam output to form a steered eigenbeam output; and  
 20 a summing element configured to combine the weighted eigenbeams to generate the auditory scene, wherein the weighting unit or the summing element are further configured to apply a regularized equalizer filter to each eigenbeam output or steered eigenbeam output, the regularized equalizer filter(s) being configured to compensate for acoustic deficiencies of the microphone array and having a regularized equalization function.

11. The system of claim 10 wherein the regularized equalization function is a radial equalization function that comprises the quotient of a regularization function limiting the radial equalization function and a radial function describing an acoustic wave field in the vicinity of the sphere.

12. The system of claim 11 wherein the regularization function is the quotient of the absolute value of the square of the radial function and the sum of the absolute value of the square of the radial function and a regularization parameter, the regularization parameter being set to a value greater than 0 and smaller than a maximum value that is smaller than infinity.

13. The system of claim 12 wherein the maximum value of the regularization parameter is 1.

14. The system of claim 12 or 13 wherein the regularization parameter depends on a susceptibility parameter that is the reciprocal of a white noise gain parameter, the white noise gain parameter being greater than a minimum white noise gain parameter that is not undercut by the equalizer filter.

15. The system of claim 14 wherein the minimum white noise gain parameter is -10 [dB].

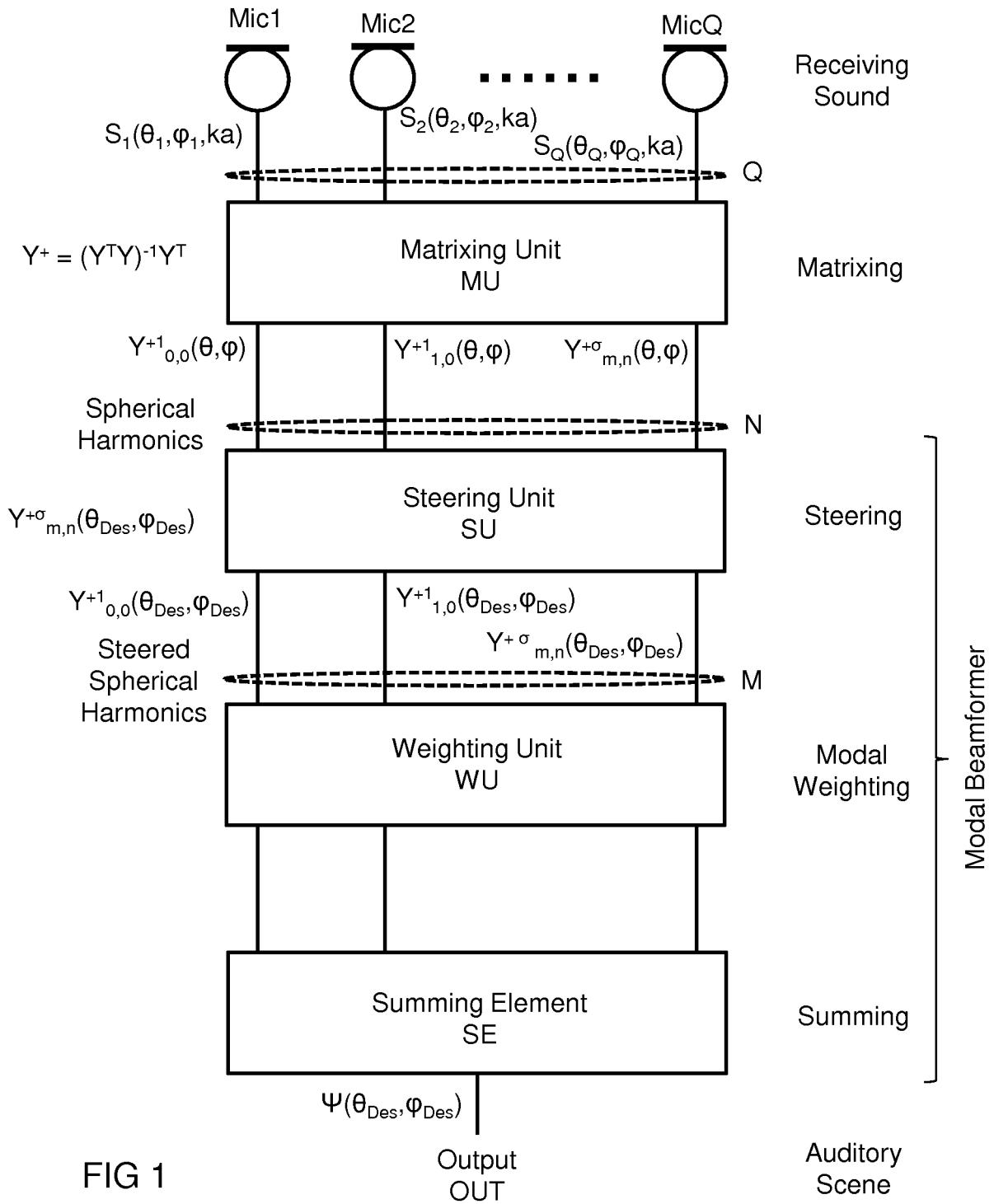


FIG 1

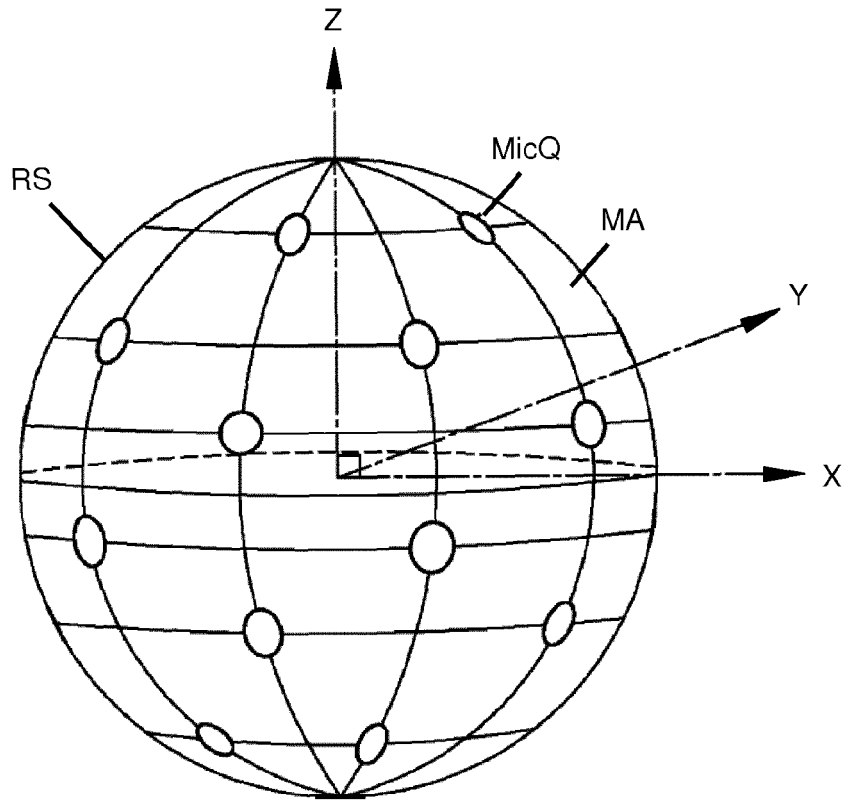


FIG 2

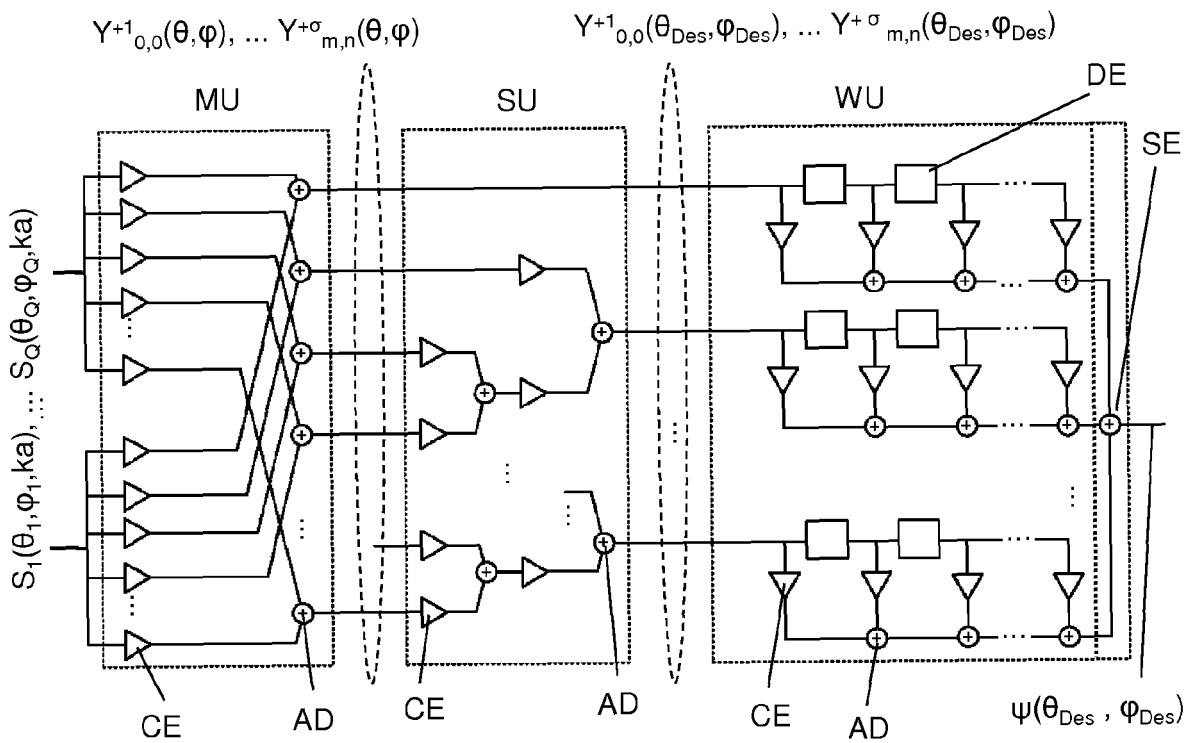


FIG 3

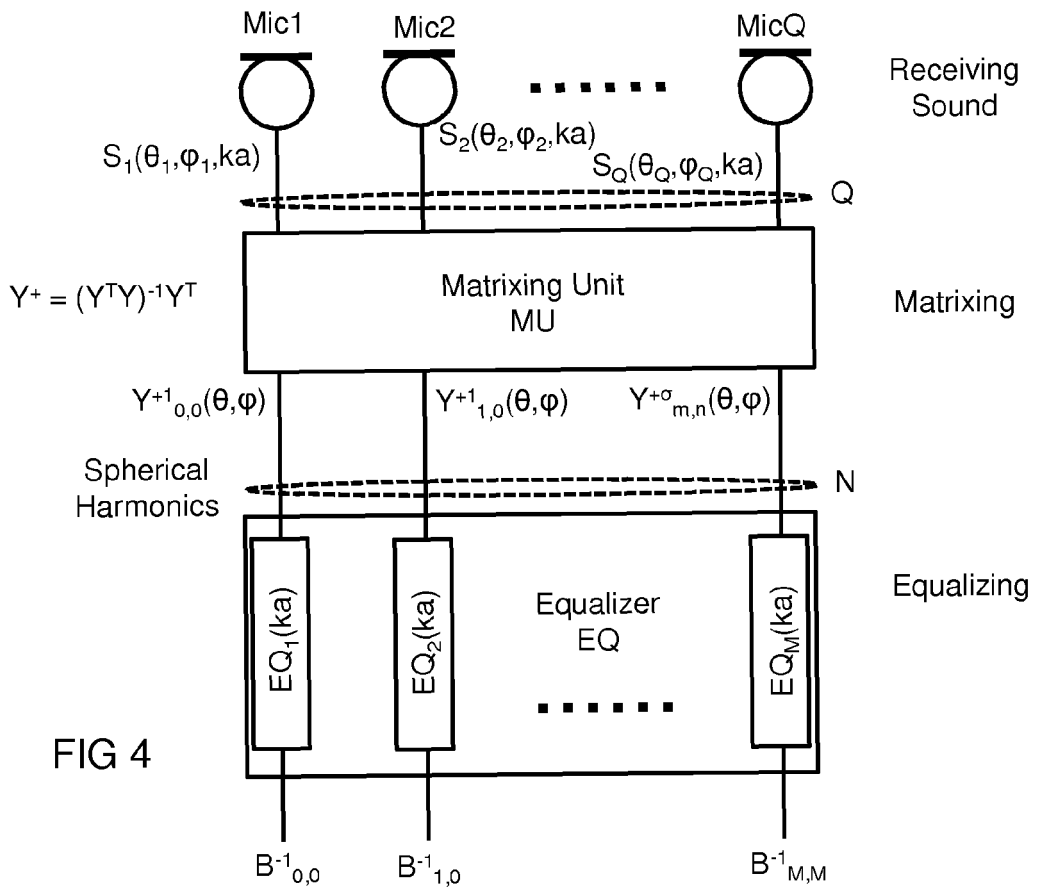


FIG 4

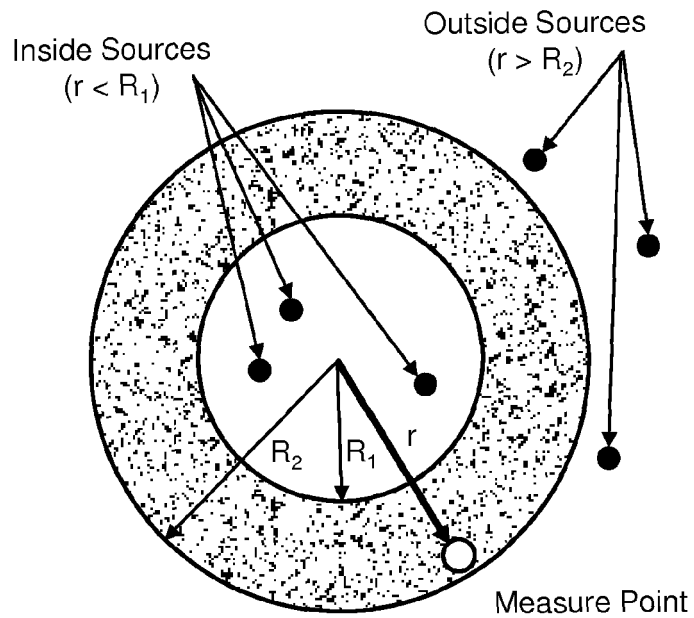


FIG 5

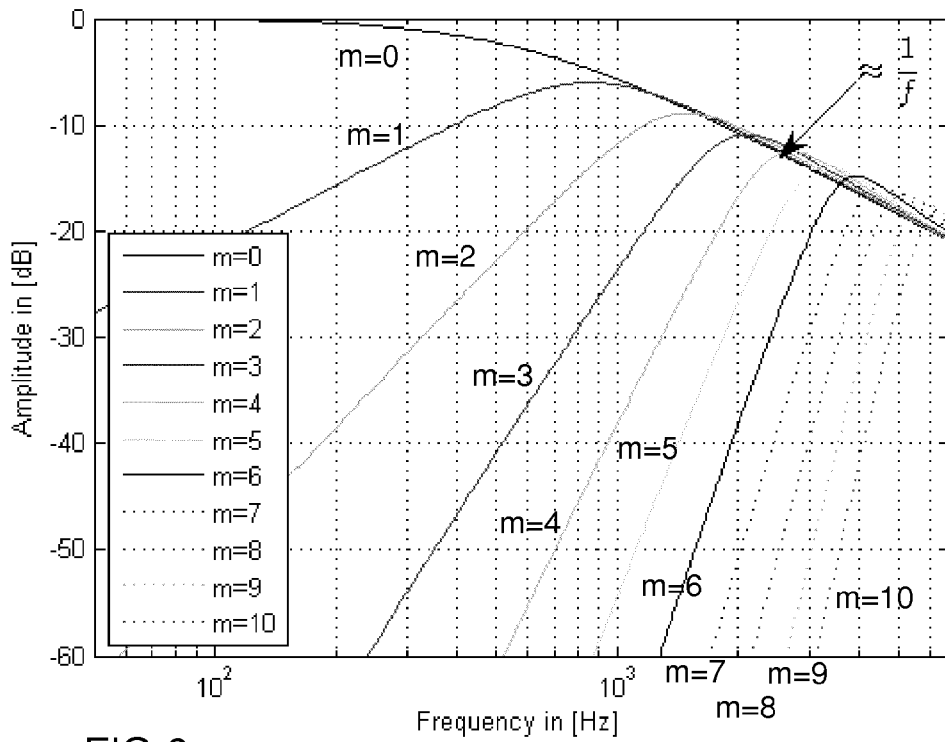


FIG 6

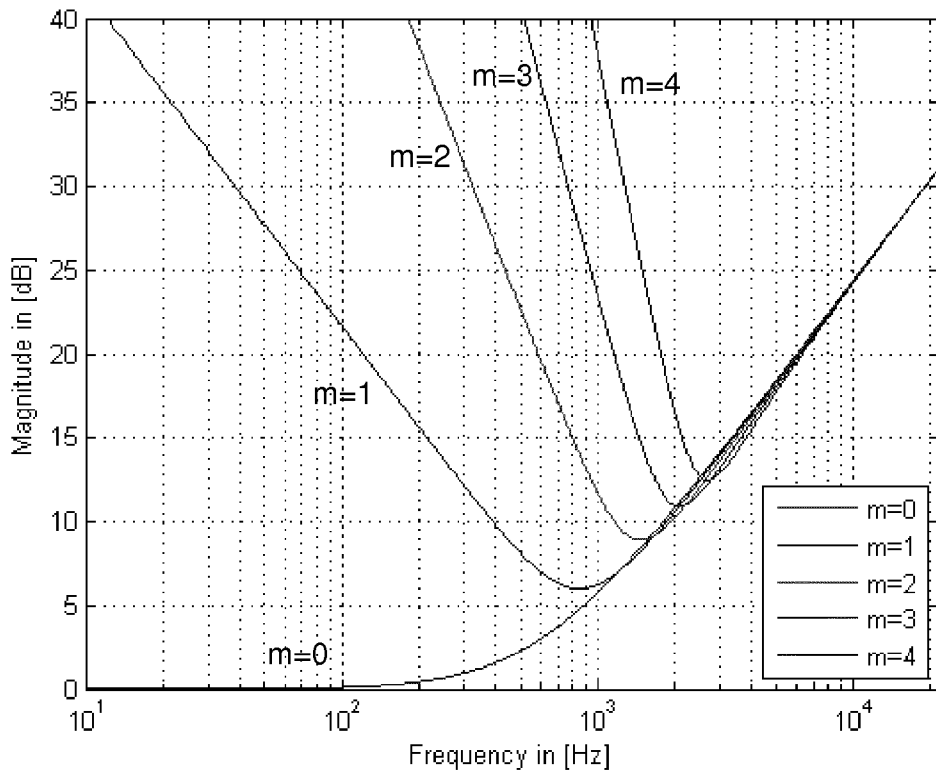


FIG 7

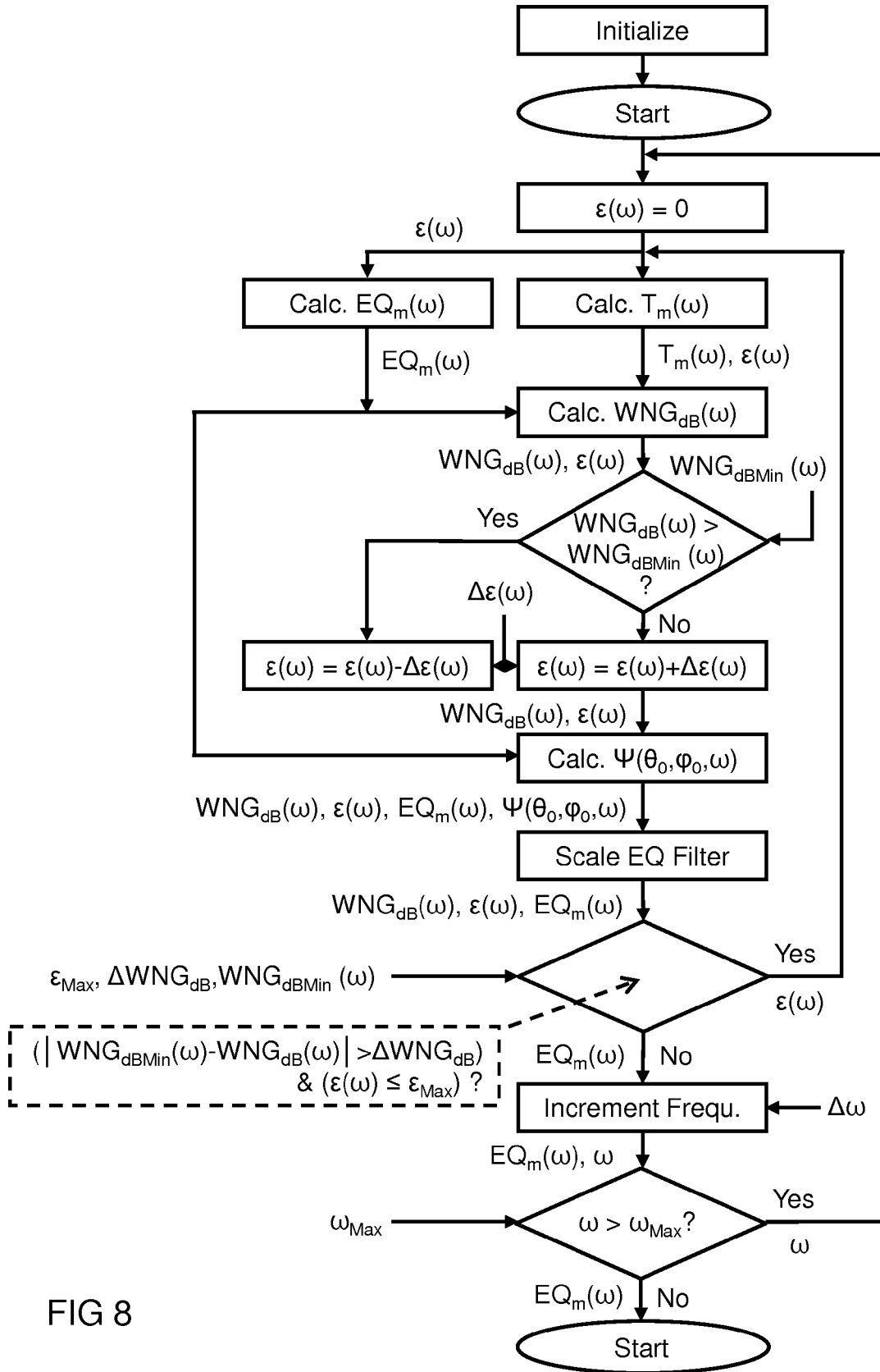


FIG 8

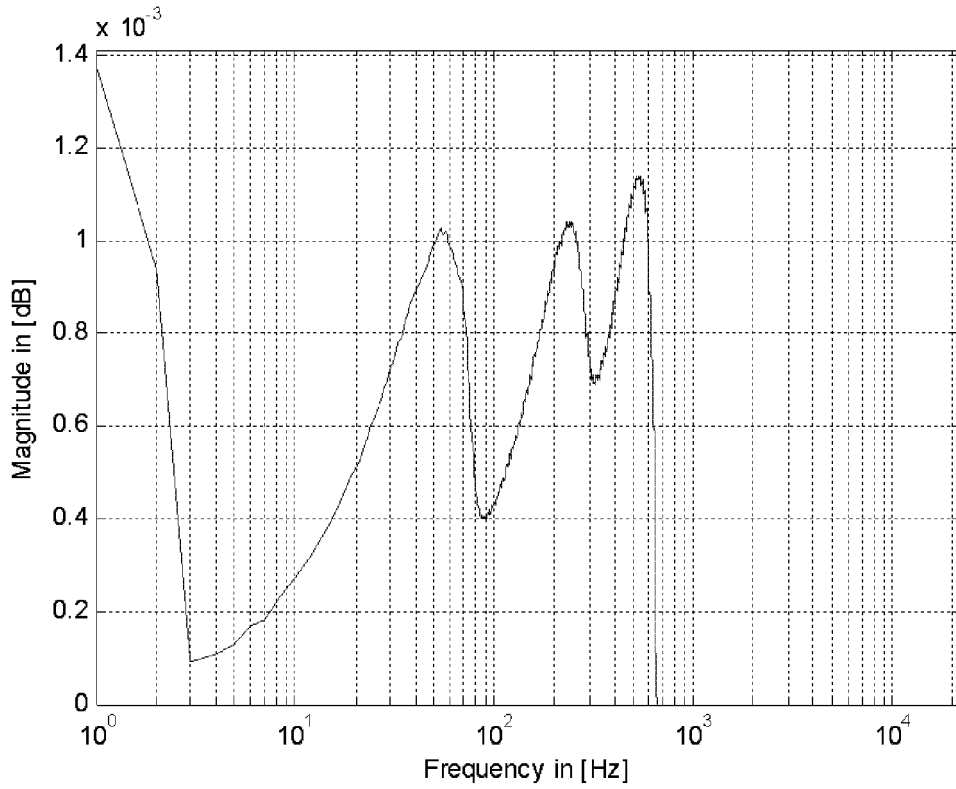


FIG 9

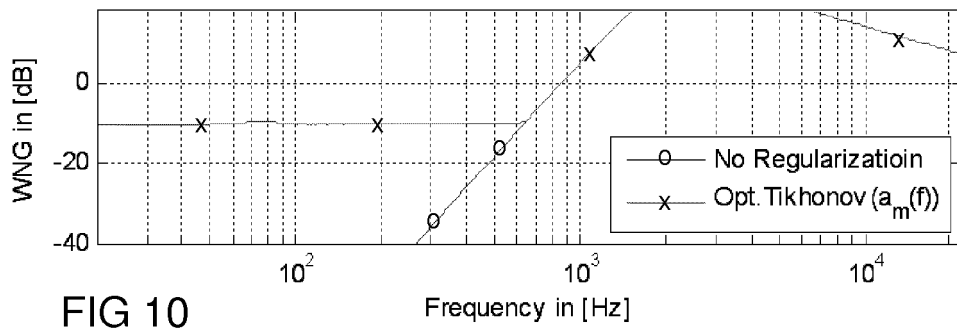


FIG 10

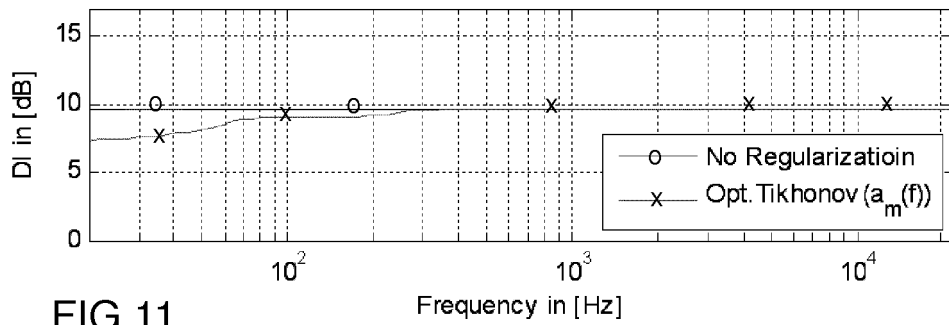


FIG 11

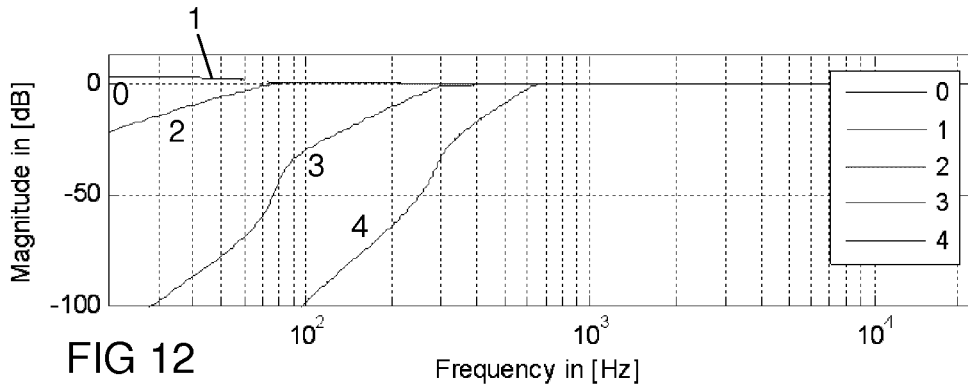


FIG 12

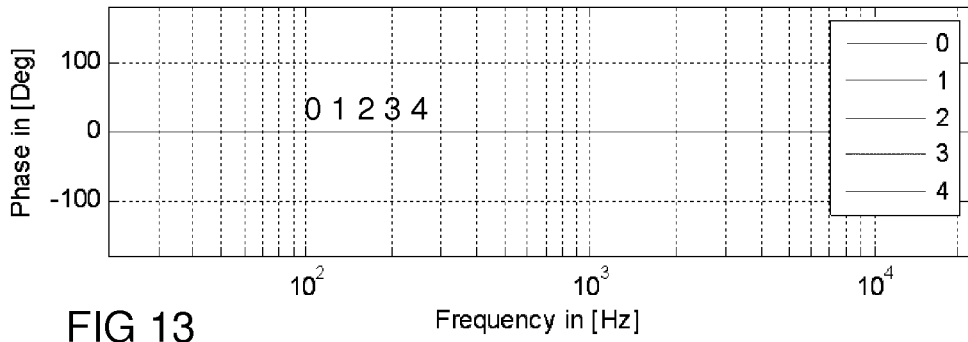


FIG 13

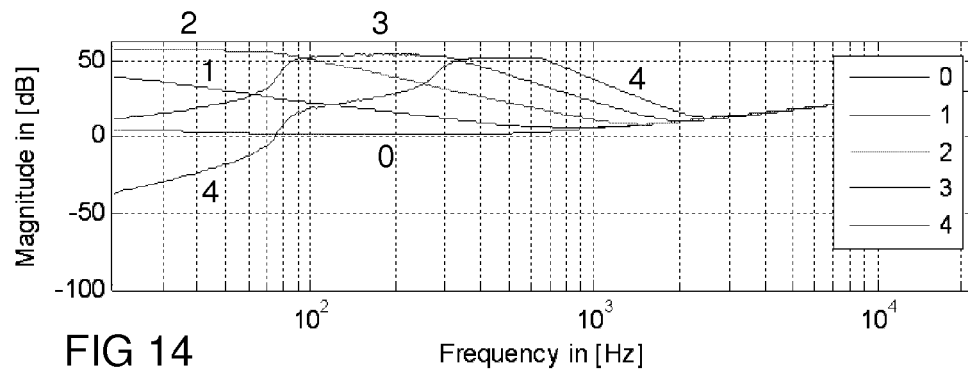


FIG 14

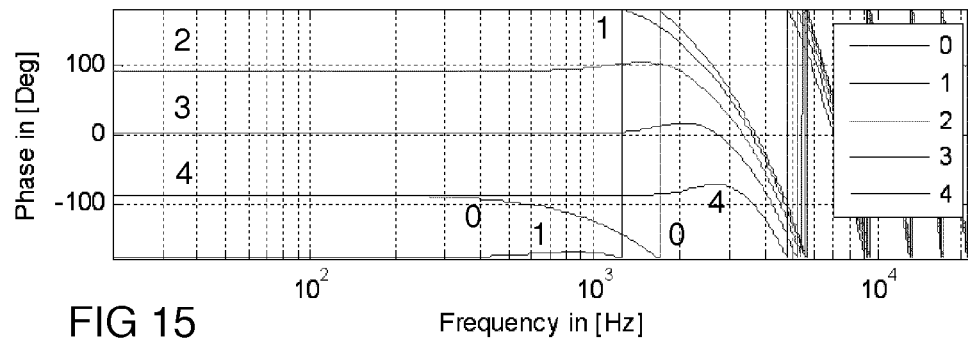


FIG 15

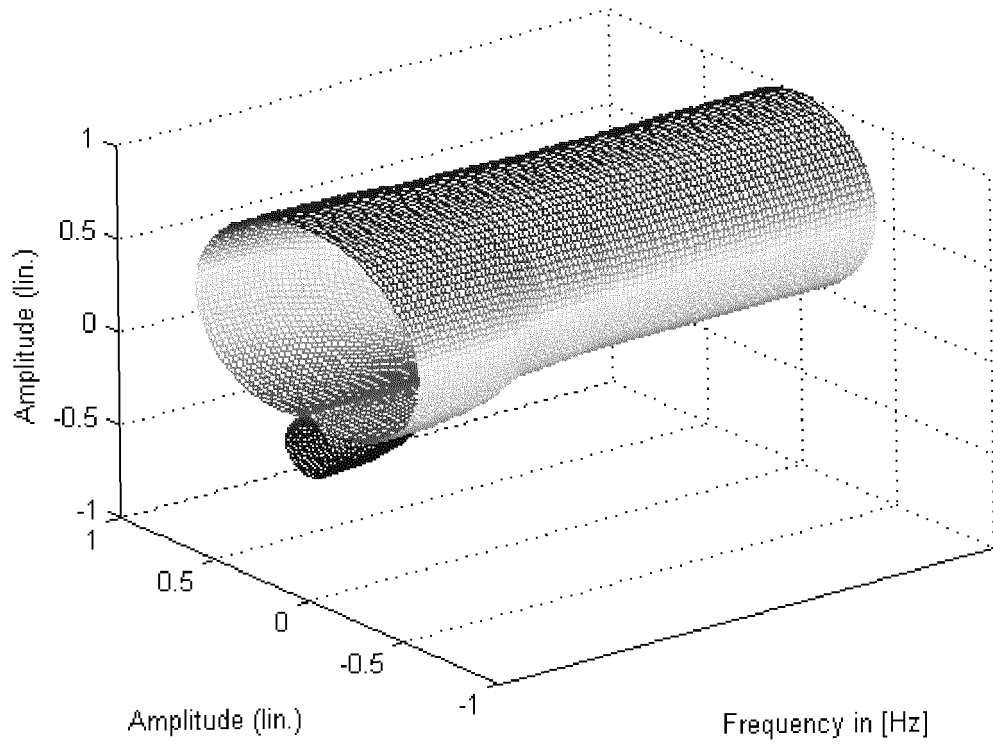


FIG 16



EUROPEAN SEARCH REPORT

Application Number  
EP 13 15 2209

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X,D	WO 03/061336 A1 (MH ACOUSTICS LLC [US]; ELKO GARY W [US]; KUBLI ROBERT A [US]; MEYER JE) 24 July 2003 (2003-07-24) * abstract; figures 1,2,17-20 * * page 5, line 14 - page 6, line 23 * * page 23, line 5 - page 26, line 20 * -----	1,2,10,11	INV. H04R3/00
A	US 2012/093344 A1 (SUN HAOHAI [NO] ET AL) 19 April 2012 (2012-04-19) * abstract; figures 18A-21 * -----	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
			H04R
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
Munich	15 May 2013	Righetti, Marco	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			

1  
EPO FORM 1503 03 82 (P04CO1)

ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.

EP 13 15 2209

5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

15-05-2013

10

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03061336 A1	24-07-2003	AU 2003202945 A1	30-07-2003
		EP 1466498 A1	13-10-2004
		US 2003147539 A1	07-08-2003
		US 2005123149 A1	09-06-2005
		US 2010008517 A1	14-01-2010
		WO 03061336 A1	24-07-2003
-----			
US 2012093344 A1	19-04-2012	CN 102440002 A	02-05-2012
		EP 2417774 A1	15-02-2012
		JP 2012523731 A	04-10-2012
		US 2012093344 A1	19-04-2012
		WO 2010116153 A1	14-10-2010
-----			

15

20

25

30

35

40

45

50

55

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- WO 03061336 A [0009]