A Tetrahedral Microphone Processor for Ambisonic Recording

Fons Adriaensen

Laboratorio di Acustica ed Elettroacustica, Parma, Italy
fons@kokkinizita.net

5th Linux Audio Conference
TU Berlin, 22. . . 25 March 2007

Overview

- A short intro to Ambisonics.
- The tetrahedral microphone.
- The Tetraproc application.
- Tetrahedral microphone calibration.
- The Fourmic JACK backend.

A short intro to Ambisonics (1)

- Ambisonics can be viewed on many levels:
  - A systematic way to analyse and represent the spatial structure of a wave field.
  - A systematic way to analyse and represent the directional information in sound.
  - A surround sound technology, including recording and reproduction systems.
- Developed 30 years ago by British mathematicians Michael Gerzon, Peter Craven, e.a.
- Firmly based on physics, maths, and psycho-acoustics.
- Full potential could not be realised at the time.
- Today used 'inside' many professional 5.1 and 7.1 applications.
- Much research (Higher Order Ambisonics) in recent years, mainly in France.

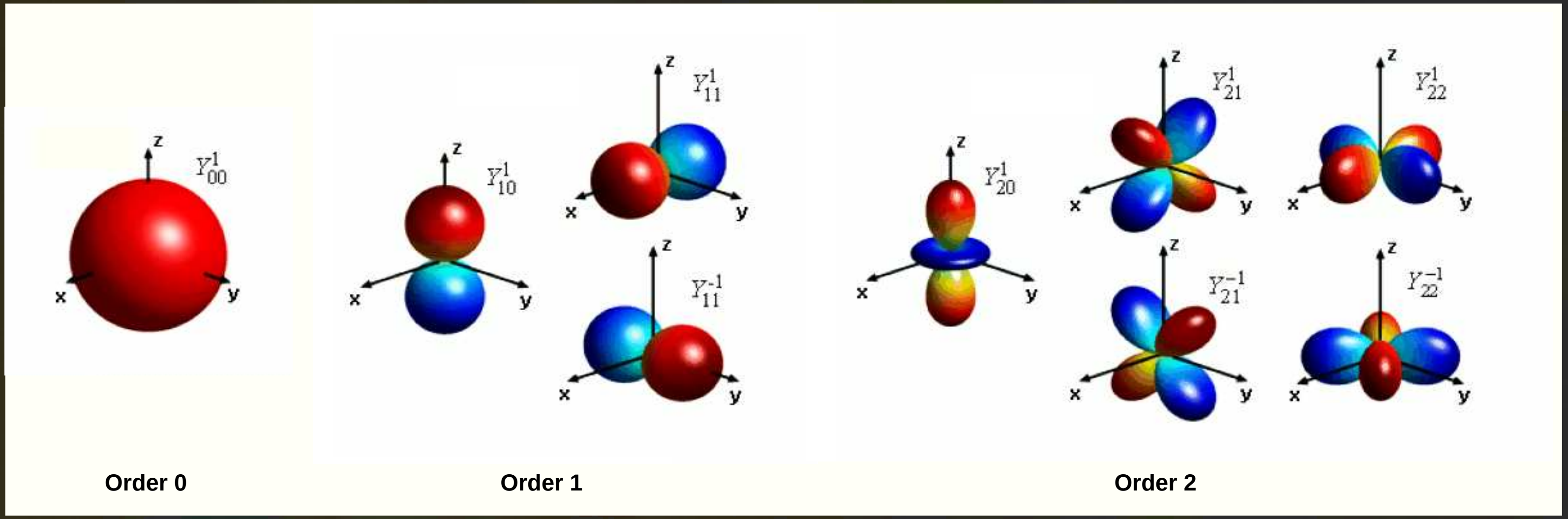
The Oxford University Tape Recording Society (1968)

L to R: Michael Gerzon, Paul Hodges, Peter Craven, Stephen Thornton
Picture by P. Hodges, used with permission

A short intro to Ambisonics (2)

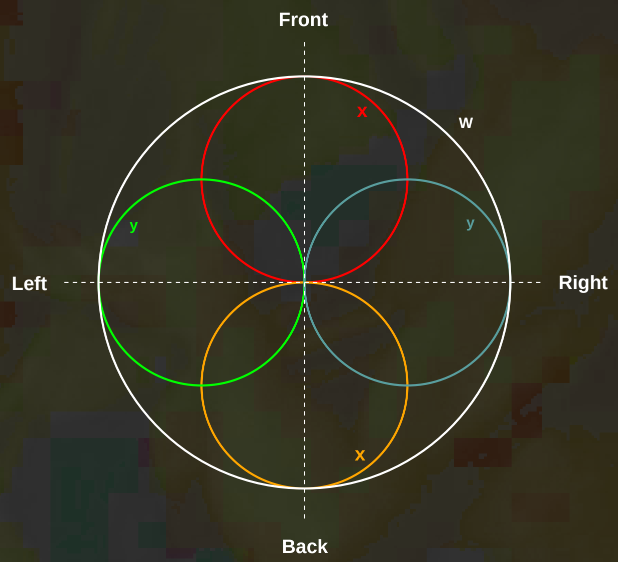
- Ambisonics encodes a sound field using a representation based on Spherical Harmonics.
  - Spherical harmonics represent the 'spatial spectrum' of a sound field.
  - Similar to Fourier analysis, but on the surface of a sphere.
- Given the 'spatial spectrum', a sound field can be reconstructed.
- There are $2M + 1$ spherical harmonics of order M.
- At low and medium frequencies, only the lower orders are required.
- Good 3-D results can already be achieved using only orders 0 and 1 (4 signals).
- 2nd order horizontal only systems (5 signals) are far superior to ITU 5.1.
- Today we have working reproduction and recording systems up to 4th order, full 3-D

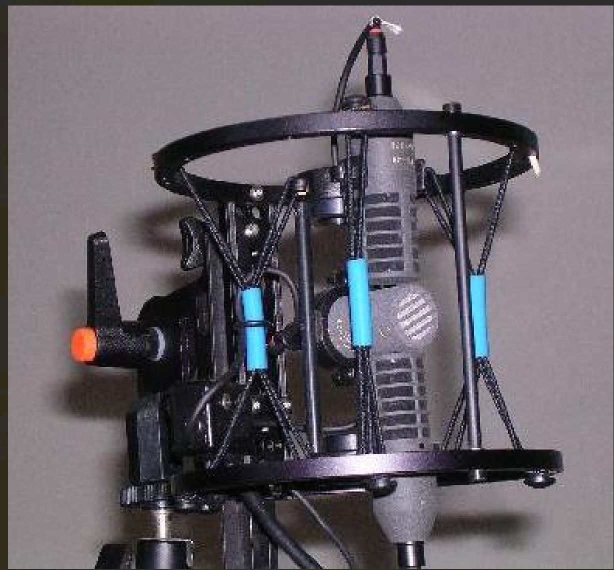
A short intro to Ambisonics (3)



- Spherical harmonics can be seen as polar patterns.
- Low order SH correspond to familiar physical quantities.
  - Order 0: pressure (W).
  - Order 1: the three components of the velocity vector (X, Y, Z).
- A set of signals corresponding to all SH up to order M is called B-format of order M. This is the standard Ambisonics format.

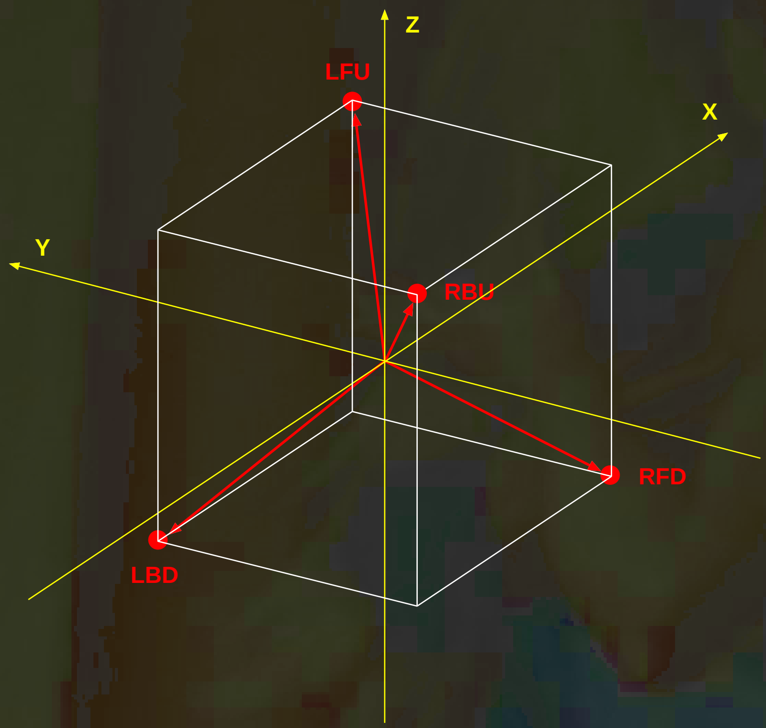
Native B-format recording





- First-order B-format signals correspond to 'real' microphones.
- An array of 3 mics can be used for 1st order horizontal recording.
- Practical solution for 3-D is a *tetrahedral* microphone.

Tetrahedral Microphones (1)



- Four near-cardiod mics at the vertices of a regular tetrahedron, pointing outwards.
- The four mics shoud ideally be as close together as possible. Practical radius is 15. . . 18 mm.
- The set of four signals (LFU, RFD, LBD, RBU) is known as A-format.

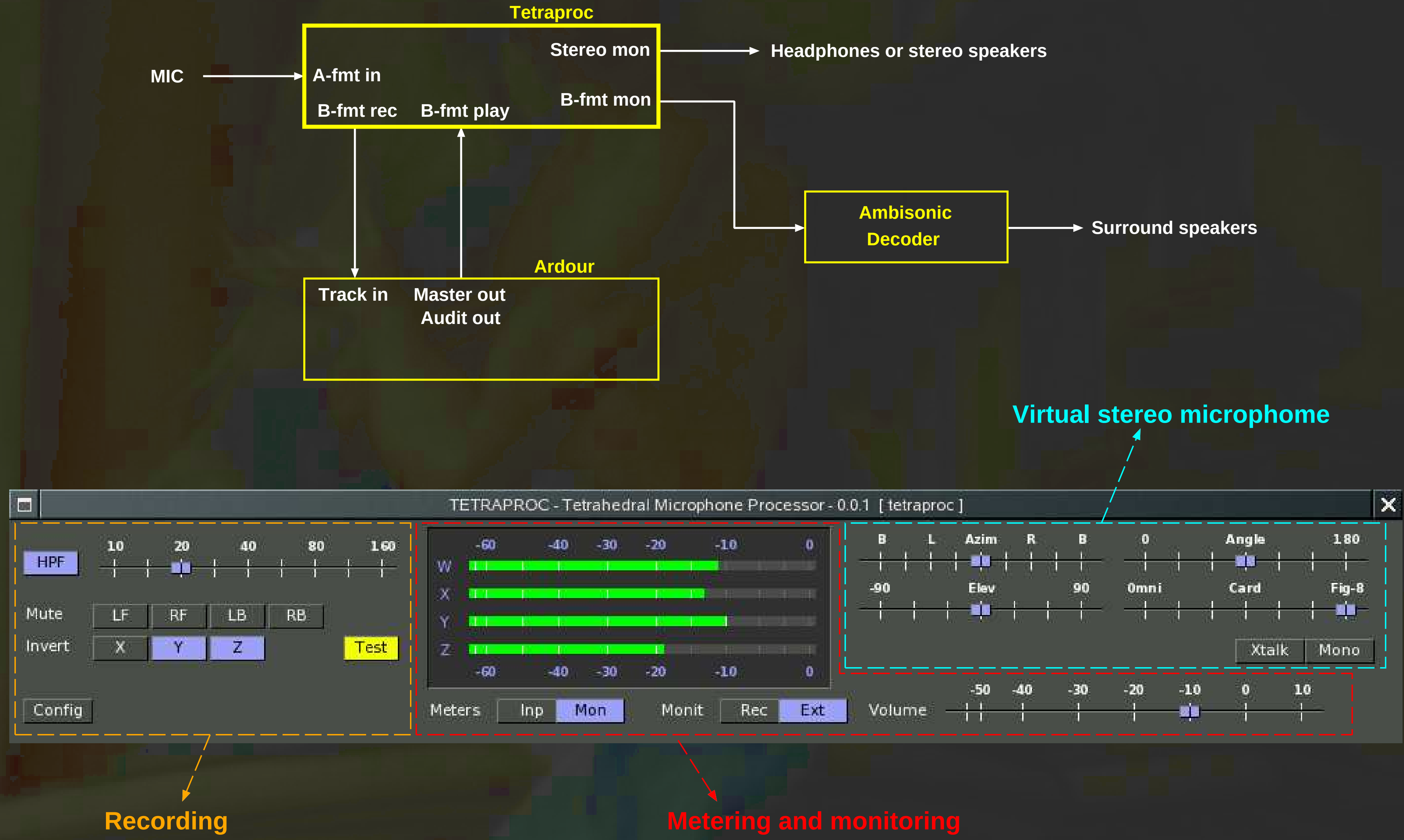
Tetrahedral Microphones (2)

- A to B format conversion is simple (in theory):

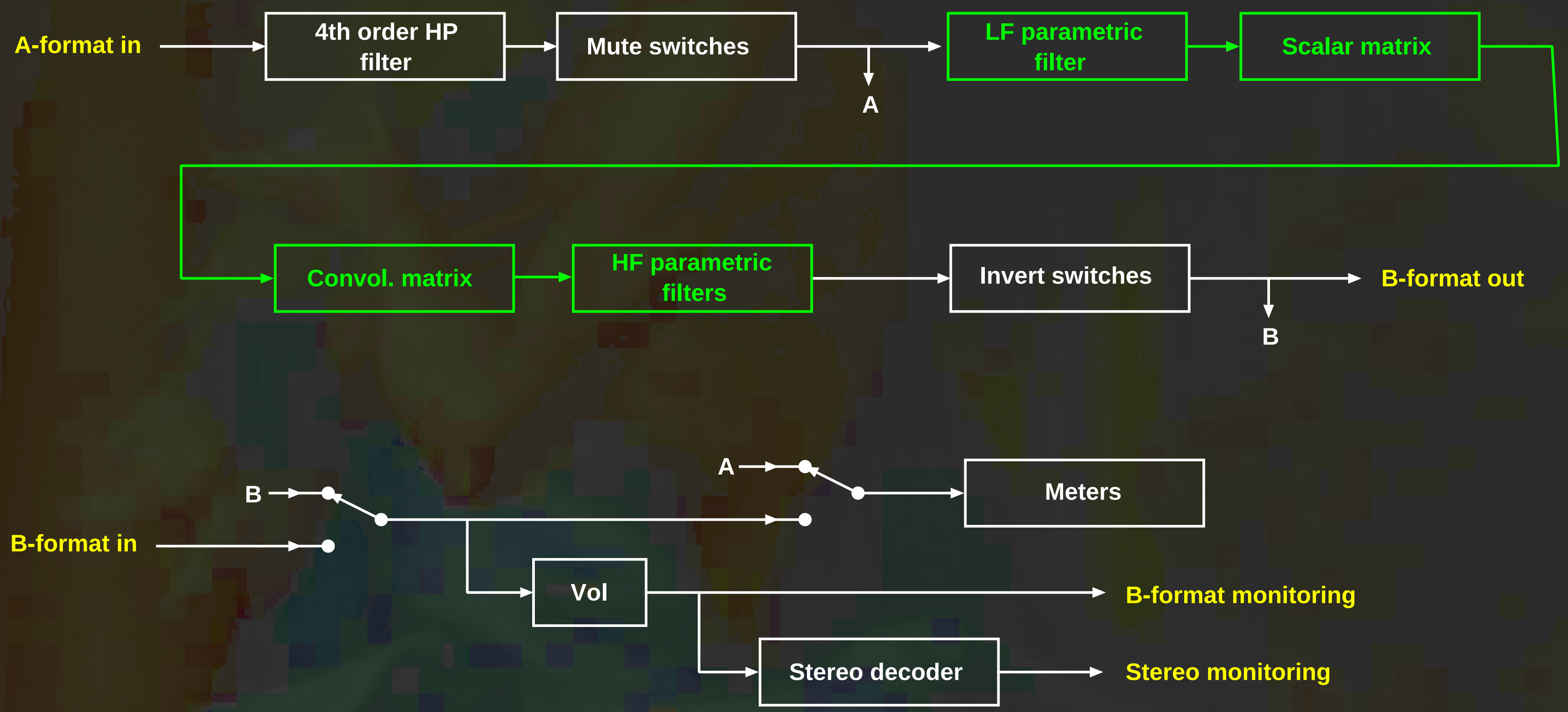
$$
\begin{aligned}
W &= F_0(\omega) \times (LFU + RFD + LBD + RBU) \\
X &= F_1(\omega) \times (LFU + RFD - LBD - RBU) \\
Y &= F_1(\omega) \times (LFU - RFD + LBD - RBU) \\
Z &= F_1(\omega) \times (LFU - RFD - LBD + RBU)
\end{aligned}
$$

- The post-matrix filters $F_0(\omega)$ and $F_1(\omega)$ are necessary because the four A-format mics are not really coincident.
- The filters can't be perfect for all directions. They are always a compromise.
- In reality the matrix must be adjusted to compensate for microphone mismatch.
- A practical processor needs some auxiliary functions: high pass filters, test controls, metering, monitoring.

Tetraproc connections and GUI



Tetraproc audio processing (1)



Tetraproc audio processing (2)

- LF parametric EQ.
  - Settings derived from measured response.
  - Optionally subjective adjustment.
- Scalar A-B matrix.
  - Also corrects gain and directivity mismatch.
  - Used only if no impulse responses are available.
- 4 by 4 convolution matrix, combines:
  - A to B-format conversion.
  - Individual microphone frequency response correction.
  - Frequency-dependent gain and directivity correction.
  - Optionally also the post-matrix parametric EQ.
- Post-matrix parametric EQ.
  - Required for correct encoding at high frequencies.
  - Function can be absorbed into convolution matrix.

Calibration (1)

- Hardware A-B processors are factory aligned for one microphone serial number.
- The same is required for a software implementation.
- The calibration requires some measurements. For each of the four A-format mics:
  - Eight impulse responses at 45 degree intervals in the horizontal plane. These are used to compute the A-B matrix (scalar or convolution), and the post-matrix equalisation.
  - (Optional) On-axis, far-field impulse response, used for individual mic equalisation.
  - (Optional) On-axis, near-field impulse or frequency response, used for low frequency equalisation.

Calibration (2)

- For an absolutely minimal calibration, the impulse responses can be replaced by 4×8 level measurements, using bandlimited noise in the 500 Hz range.
- Measurements can be done in a large room, or in open air, using e.g. Aliki.
- An omnidirectional reference microphone is required.
- A separate program is used to derive matched processing parameters, and to create a configuration file for Tetraproc.

Calibration (3)

- Calibrating a tetrahedral mic is not a simple procedure...
  - There are many options and variations.
  - Very few people have any real experience with this problem.
  - Performing acoustic measurements correctly can be tricky.
- The methods used are based on procedures recommended by Soundfield microphone designer Richard Lee, together with new possibilities offered by IR measurement software.
- More work will be required to make the calibration as easy as possible for a non-expert user.

Calibration - Microphone response correction

- MF and HF equalisation using far-field, on-axis impulse response.
  - Uses only the first 5...20 milliseconds of the IR, no LF equalisation.
  - Kirkeby inversion of IR, with manual correction.
  - Result used in convolution matrix.
- LF equalisation based on near-field measurement using a small source.
  - Corrected for near-field effects using known distance and directivity.
  - Used to set the LF parametric equaliser.
- These procedures do not compensate for directivity mismatch.

Calibration - Gain and directivity mismatch (1)

- Small errors in sensitivity and directivity of the A-format mics have a big impact on the first order B-format signals.
- Eight measurements in the horizontal plane allow to compute a cross section of the polar diagram for each of the four A-format microphones.
- P and V can be computed from 2 values, but using all 8 is much more accurate.
- The real V is $\sqrt{3/2}$ times the value in the horizontal plane.

Calibration - Gain and directivity mismatch (2)

Given P and V for each mic, a matched A-B transformation matrix can be computed.
If (x, y, z) are the direction cosines, the ouputs of the four mics can be expressed as:

$$
\begin{aligned}
LF(x, y, z) &= P_{LF} + V_{LF} \times (x + y + z) \\
RF(x, y, z) &= P_{RF} + V_{RF} \times (x - y - z) \\
LB(x, y, z) &= P_{LB} + V_{LB} \times (-x + y - z) \\
RB(x, y, z) &= P_{RB} + V_{RB} \times (-x - y + z)
\end{aligned}
$$

The required B-format signals are:

$$
\begin{aligned}
W(x, y, z) &= 1 \\
X(x, y, z) &= x \\
Y(x, y, z) &= y \\
Z(x, y, z) &= z
\end{aligned}
$$

Expressing (W, X, Y, Z) as linear combinations of (LF, RF, LB, RB) now only requires solving sets of linear equations.

Calibration - Post-matrix filtering (1)

- A complex matter. Result is a compromise and must be evaluated subjectively.
- Automatic derivation of the post-matrix filters using the measured P, V values over the entire frequency range is a research project...
  - The maths get complicated when delays are taken into account.
  - A simple first order microphone model is no longer valid.
  - Complex diffraction effects make the analysis almost impossible.

Calibration - Post-matrix filtering (2)

- Good results seem to be achieved by equalising the cardinal directions.
  - Process the IR using the computed mic EQ and A-B matrix.
  - Adjust parametric EQ using displayed frequency and phase plots.
  - Or use an automated IR inversion procedure, with manual correction.
- Alternatively, equalise the diffuse field response (average power over all directions). This requires more measurements.

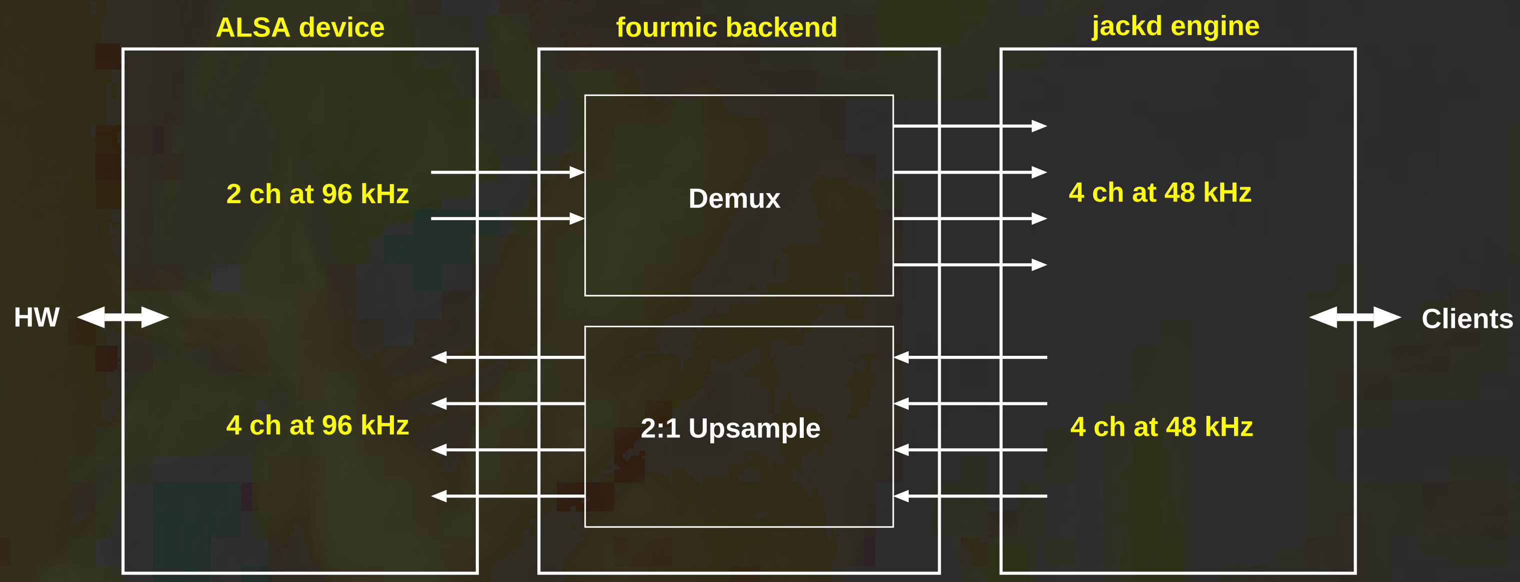
Calibration - Post-matrix filtering (3)

- If no IR measurements are available: use a 'standard' filter.
- Theoretical filters depend on the radius and the directivity.
- There is very little available literature, and no agreement.
- The following are all quite different:
  - The theoretical diffuse field response.
  - The theoretical cardinal direction response.
  - The filter implementation proposed by Gerzon.
  - The filters actually used in HW processors.
- More experimentation will be required.

The Fourmic backend (1)

- Using A-format mics requires capture channels with precisely matched gain controls.
- Core Sound's 4Mic interface multiplexes 4 channels on a stereo SPDIF stream at twice the sample frequency.
- Where to do the demultiplexing ?
  - In the ALSA drivers.
  - In the user space ALSA library.
  - In an ALSA library plugin.
  - In the application.
  - ...
  - In the JACK backend.

The Fourmic backend (2)



- The fourmic backend plugin provides:
  - Demultiplexing of the 2-channel stream to the four A-format signals.
  - Upsampling of up to four playback channels for monitoring.
- Demuxing introduces an ambiguity. Automatic resolution using the 4Mic test signal may be possible.

Tetraproc — Things to do

Lots of work . . .

- Manual and documentation.
- Improve the calibration procedure.
- Make some recordings.
- . . .

Acknowledgements

Many thanks to Richard Lee, and to the members of the surround sound mailing list, in particular:

Aaron Heller, Angelo Farina, Dave Malham, David McGriffy, Eric Benjamin, Len Moskowitz and Paul Doornbusch.

