# PERCEPTUAL ISSUES IN MULTI-CHANNEL ENVIRONMENTS

**J. Robert Stuart**
Meridian Audio
Huntingdon
England.

The paper reviews the design of multichannel playback systems from the basic principles of psychoacoustics. Using auditory modelling as a guide, the author develops some examples on issues such as the audibility of errors which are potentially masked or unmasked in a 3-D presentation and the perceptual relationships between numbers of channels and frequency range –to clarify the subjective aspects of subwoofer splicing.
The paper opens the question of relating perceptual and cognitive responses to directional sound or to sound fields.

## Introduction

This set of notes is not a paper in full Preprint format. It simply serves as a set of notes and references to accompany the lecture.

## Perception

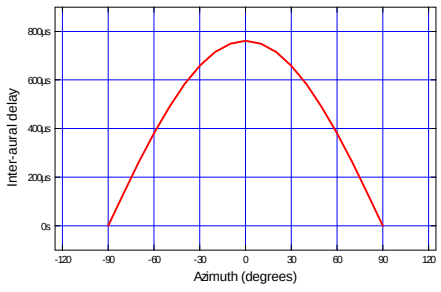
### Peripheral auditory function

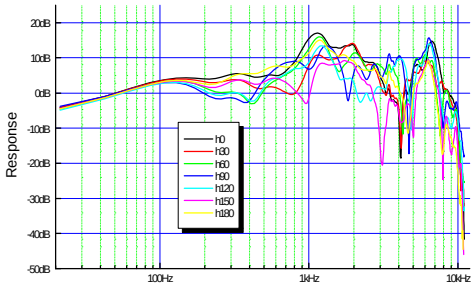
Sounds are encoded in the auditory periphery on a loudness-pitch basis. The neural code partially represents specific loudness and frequencies are dispersed in cochlea according to integral of auditory filtering function. Sounds have temporal encoding through

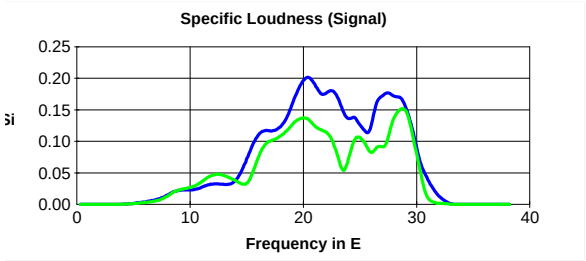
- onset and offset
- synchronously for waveforms or envelopes < 800Hz
- loudness dependency through threshold effects

### Peripheral binaural auditory function

Two spaced ears mean time arrival difference for sounds in different locations - up to 0.7ms – and intensity difference due to head ‘shadow’.
Pinna effects make important spectral modification according to angle of incidence and this filtering action combined with head diffraction is used to encode direction.




Figs of inter-aural time delay and pinna effects with azimuth. Also internal representation of white noise from two incident angles (single near ear).

### Localisation phenomena: Time-intensity trading

Cues in replay may be time, phase or intensity and may not relate ‘naturally’

### Localisation phenomena: Sound-field

The sound-field method looks at the apparent direction of a source in the absence of a listener, or his binaural capability. Localisation can be confirmed by head-turning.

### Localisation phenomena: Precedence

It is well known that sounds appear to come from the direction of first arrival, somewhat independently of amplitude – note the Haas and Franssen effects.

### Localisation phenomena: Binaural cognition

Spatial perception can be modified by ‘inaudible’ components – underlining the sophistication of the ‘binaural processor’.

## Central binaural processing

The central processor has substantially increased temporal acuity – we can perceive arrival-time differences of the order of 30us at 50 phon.

In binaural listening there are significantly modified detectability thresholds, e.g.:

- binaural masking and release
- binaural masking-level differences (≅ 12dB)
- binaural beats
- subliminal perception: e.g. Groen

### Binaural post-processing

The binaural process also significantly modifies the perceived sound. For example:

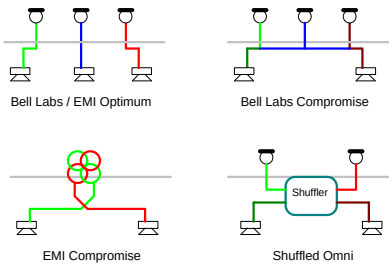
- removal of comb-filtering perceptions
- binaural pitch
- pitch - localisation
- Deutsch octave illusion

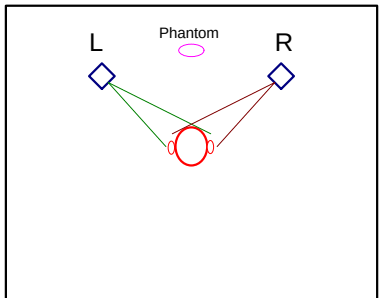
### Binaural Loudness

Loudness for binaural presentation is not simply related to mono equivalent, lateral inhibition causes loudness in each ear to grow ‘as masked’.

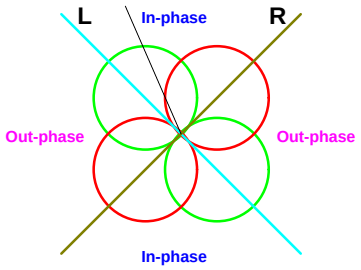
Near threshold switching from mono to binaural gives approx. doubling in Sone i.e. 10 Phon; at mid-loudness (50 Phon) we see a 4 Phon increase; at high level (80 Phon) a 3 Phon increase. Therefore multichannel makes different loudness impression to mono or fewer presentations.

## Stereo

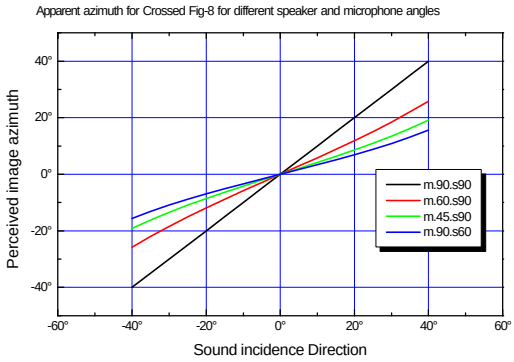



Early stereo techniques and two-speaker replay.




**Co-incident Figure-of-eight**




### Multichannel Perceptual Issues

Multichannel brings:

- Easier and more emphasised auditory object externalisation
- Simpler instrument streaming
- Changed loudness balance through the binaural process
- Changed timbre perception through location-correction
- Markedly different ambient perception
- Speaker directivity increased by 10logn
- Direct to reverberant ratio change demotes listening-room acoustics
- Increased acuity for channel or processing errors

## Cognition

### Perception of objects

The perception of music in surround depends on our ability to ‘externalise’ perceptions into ‘acoustic objects’. Some factors that effect this ‘cognitive’ process are: amplitude, fundamental frequency, timbre, envelope patterns, onset disparities, correlated changes, contrast with earlier and later sounds and spatial location. Other models used to aggregate the perceptual attributes of acoustic object formation are: constancy, similarity, auditory streaming, continuation, common fate, language, rhythm, closure – replacement of missing sounds and attention.

### Cognitive elements in sound

Regarding cognition, there are the following elements of the perceived soundfield:

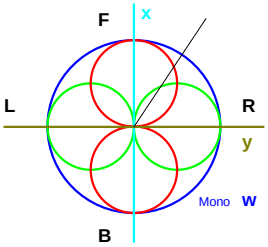
- Mono elements of sound; pitch, loudness, timbre; auditory object formation; ‘object’ grouping
- Binaural additions: auditory object location and separation, ‘object’ externalisation
- spaciousness, ambience recognition, distance perception

### Cognitive elements of Music

Cognition of music can be viewed at several levels.

- Cognition of the Sound Object
- Cognition of the Music
- Cognition of the music’s Structure
- Cognition of Content

# Encoding surround



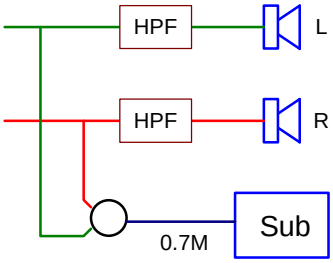
Horizontal Ambisonic

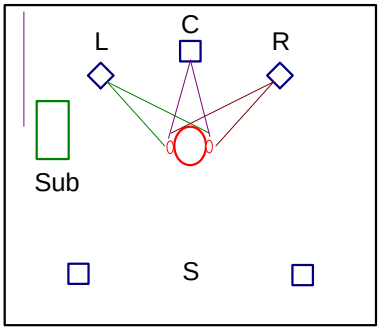
Multichannel azimuthal sampling is the general principle of sound-field encoding. Using azimuthal harmonic synthesis we can create an hierarchical structure of transmission channels for increasing accuracy of surround encoding. The transmission channels will be:

- Sum or omnidirectional channel
- Difference channel = Sum with phase shift: stereo is a special subcase with phase = 180°
- Tertiary channel = conjugate-phase version of Difference
- Quaternary channel using second harmonic

Tertiary and higher can be band-limited.

# Low-frequency factors




### Bass is directional

There is no evidence of lack of direction at LF and LF carries important ambient clues. Bass directionality can be confused in rooms because of standing-wave patterns and $R_T$ build-up.

### Bass-sharing strategies on record or decode

The bass power output is incorrect if the number of woofers is different from channels *conveyed* or *encoded.* LF loudness is foundation for 'continuity' in music

The table below illustrates that a mono subwoofer (derived as the sum of Left and Right channels, does not produce the correct in-room power for signals from each of the four cardinal directions Lt, Rt, M, S.

| | Inputs at 0dBr to: | | | |
|---|---|---|---|---|
| | **Lt** | **Rt** | **M** | **S** |
| **Pro Logic Output** | dB | dB | dB | dB |
| Left Front | 0.0 | -30.0 | -30.0 | -30.0 |
| Right Front | -30.0 | 0.0 | -30.0 | -30.0 |
| Centre | -30.0 | -30.0 | 0.0 | -30.0 |
| Mono Sub | **0.0** | **0.0** | **0.0** | **n/a** |
| Sub needed | **0.0** | **0.0** | **0.0** | **n/a** |
| | | | | |
| **Stereo Output** | dB | dB | dB | dB |
| Left Front | 0.0 | -999.0 | -3.0 | -3.0 |
| Right Front | -999.0 | 0.0 | -3.0 | -3.0 |
| Centre | -999.0 | -999.0 | -999.0 | -999.0 |
| Mono Sub | **0.0** | **0.0** | **3.0** | **-999.0** |
| Sub needed | **0.0** | **0.0** | **0.0** | **?** |

## Data - reduction and associated factors

### Difficulties with Coders

Psychoacoustic data-reduction coders present an hierarchy of problems in multichannel encoding or presentation. The following few problems are listed:

- Intensity stereo coding does not necessarily preserve phase or waveform, there is damaged potential for Soundfield systems.
- Potential for directional unmasking (binaural)
- Risk of underestimating masking through Pinna effects (perceptual)
- Compression/distortion of the stereo image

## Summary

1. Multichannel design requires good understanding of both perception and cognition.
2. Cinema methods equate to multiple-mono using logic-assisted matrix decoders or 5 discrete channels.
3. 'Music' systems need to encode the whole soundfield by representing and producing the wavefronts in the absence of a listener (i.e. not binaural methods). Head and body movement are essential for multi-listener environments.
4. Soundfield transmission does not require 5 full-bandwidth channels so better use of any data rate is possible.
5. This is an urgent problem NOW for setting standards: cinema must not drive recording or channel design!

## References

### Bibliography


1 Blauert, J. *Spatial Hearing* (MIT Press, 1983)

2 Bregman *Auditory Scene Analysis*

3 Carterrette, E.P. and Friedman, M.C. *Handbook of Perception*, **IV**, 'Hearing' (Academic Press, 1978)

4 Deutsch, D. *The Psychology of Music* (Academic Press, 1982)

5 Moore, B.C.J. *An Introduction to the Psychology of Hearing* (Academic Press, 1991)

6 Tobias, J.V. *Foundations of Modern Auditory* T*heory* (Academic Press, 1970)


### Perception


7 Buus, S. 'Release from masking caused by envelope fluctuations' *J. Acoust. Soc. Amer.*, **78**, 1958–1965 (1985)

8 Groen, J J. 'Super and subliminal binaural beats' *Acta Oto-Lar*, **57**, p224

9 Hall, J.W. 'Experiments on Comodulation Masking Release', in *Auditory processing of complex sounds*, Eds Yost, W.A. and Watson, C.S., Erlbaum and Assoc. (1987)

10 Irwin, R.J. 'Binaural summation of thermal noises of equal and unequal power in each ear' *American Journal of Psychology*, **78**, 57–65 (1965)

11 Lochner, J.P.A. and Burger, J.F. 'Form of the loudness function in the presence of masking noise' *J. Acoust. Soc. Amer.*, **33**, 1705–1708 (1961)

12 Scharf, B. 'Loudness summation between tones from two loudspeakers' *J. Acoust. Soc. Amer.*, **56**, 589–593 (1974)

13 Scharf, B. and Fishken, D. 'Binaural summation of loudness' *J. Exp. Psychology*, **86**, 374–379 (1970)

14 Stuart, J.R. 'Predicting the audibility, detectability and loudness of errors in audio systems' *AES 91st convention*, New York, preprint 3209 (1991)

15 Stuart, J.R. 'Estimating the significance of errors in audio systems' *AES 91st convention*, New York, preprint 3208 (1991)

16 Stuart, J.R. 'Psychoacoustic models for evaluating errors in audio systems' *PIA*, **13**, part 7, 11–33 (1991)

17 Yost, W.A. and Watson, C.S., (eds) of '*Auditory processing of complex sounds'*, Eds Erlbaum and Assoc., section VI (1987)

**Cognition**

18 Deutsch, D. 'The octave illusion and auditory perceptual integration' in *Hearing Research and Theory*, Eds Tobias, J.V. and Schubert, E.D., 99–142 (Academic Press 1981)

19 Terhardt, E., 'Music perception and sensory information acquisition: relationships and low-level analogies', *Music Perception* **8** no 3, 217-239, (Spring 1991)

20 Umemoto, T. 'The Psychological Structure of Music' *Music perception* **8** No 2, 115–128 (Winter 1990)

**Surround sound**

21 Boone, M. M. et al. 'The Wave Field Synthesis Comcept Applied to Sound Reproduction' *AES 96th convention*, Amsterdam, preprint 3814 (1994)

22 Cooper, D.H. and Shiga, T. 'Discrete-Matrix Multichannel Stereo' *J. Audio Eng. Soc.*, **20,** 346–360 (1972)

23 Gerzon, M.A 'Periphony: With-Height Sound Reproduction' *J. Audio Eng. Soc.*, **21**, (Jan/Feb 1973)

24 Gerzon, M.A 'NRDC Surround Sound System' *Wireless World* 36–39 (April 1977)

25 Gerzon, M.A 'Ambisonics in Multichannel Broadcasting and Video' *J. Audio Eng. Soc.*, **33**, 859–871 (Nov 1985)

26 Gerzon, M.A. 'Optimal Reproduction Matrices for Multispeaker Stereo' *AES 90th Convention*. (1991)

27 Herre, J., Brandenburg, K. and Lederer, D. 'Intensity Stereo Coding' *AES 96th convention*, Amsterdam, preprint 3799 (1994)

28 Klipsch, P.W. 'Circuits for three-channel audio playback derived from two sound tracks' *IRE Transactions on Audio* **AU7** No. 6, 161–165 (1959)

29 MacCabe, C. J. and Furlong, D. J. 'Virtual Imaging Capabilities of Surround Sound Systems' *J. Audio Eng. Soc.*, **42,** 38–49 (Jan. 1994)

30 Meares, D. J. 'Perceptual Attributes of Multichannel Sound' *Proceedings of AES 12th International Conference "The Perception of Reproduced Sound"*, 171–179 (June 1993)

31 Pedersen, J. A. et al. 'The Distribution of the Low Frequency Sound Field and its Relation to Room Equalisation' *AES 96th convention*, Amsterdam, preprint 3852 (1994)

32 Perrott, D. R. 'Auditory and Visual Localisation: Two modalities One world', *Proceedings of AES 12th International Conference "The Perception of Reproduced Sound"*, 221–231 (June 1993)

33 Scheiber, P. 'Toward a More Accurate Spatial Environment' Letter to *J. Audio Eng. Soc.*, **17,** (Dec 1969)

34 Schroeder, M. R. 'Listening with Two Ears' *Music perception* **10** No 3, 255–280 (Spring 1993)

35 Snow, W.B. 'Basic principles of Stereophonic Sound' *Journal of SMPTE*, **61** 567–589 (1953)

36 Steinberg, J.C. and Snow, W.B. 'Physical factors in Auditory Perspective' *Journal of SMPTE*, **61** 420–430 (1953)

**Data-reduction issues**

37 Gerzon, M.A. 'Problems of error-masking in audio data compression systems' *AES 90th Convention*. (1991)

38 Stuart, J.R., 'Future Coding' *HiFi News*, June 1991.

39 D. Wiese, and G. Stoll, (1990) 'Bitrate reduction of high-quality audio signals by modelling the ears masking thresholds', *AES 89th convention,* preprint 2970, (1989).