Ecma/TC32-TG22/2013/071
Ecma/GA/2013/132

International Multimedia Telecommunications Consortium

Today and tomorrow of visual communications, building on 20 years of technology breakthroughs

October 8 – 10

Porto, Portugal

Clemens Par (Swissaudec):

GROUND-BREAKING LOW BITRATE 3D AUDIO CODING BASED ON 2D AUDIO CODECS -

INTRODUCING S5 ACTIVITIES WITHIN ECMA TC32-TG22

ABOUT MYSELF

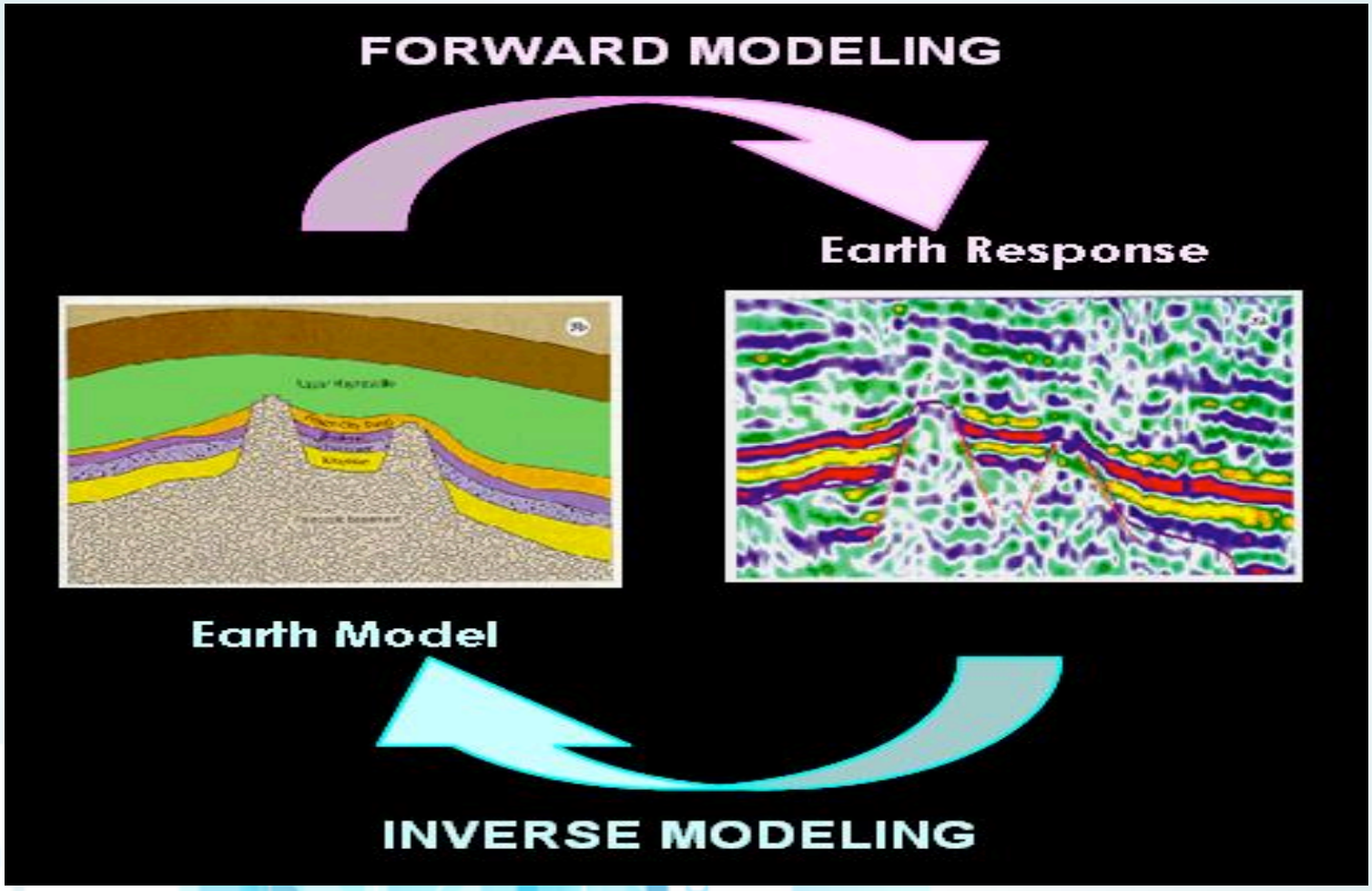
- Parallel studies in conducting at Hochschule „Mozarteum“ in Salzburg and mathematics at ETH Zurich (with R. E. Kálmán).
- Author, musician, presenter and executive producer for ARD, ORF and Schweizer Radio DRS. Translator for Insel-Suhrkamp.
- First application of inverse problems to audio engineering, and of algebraic invariants to signalling theory. WIPO Award 2009.
- CEO of Swissaudec GmbH, Switzerland
- Current standardization activities: ISO/IEC JTC1/SC29/WG11, Ecma TC32-TG22.

OVERVIEW

- V. Hambardzumyan‘s „inverse problem“ in audio engineering
- A first application of D. Hilbert‘s „Über die vollen Invariantensysteme“ to signalling theory
- An introduction to Ecma TC32-TG22
- Ecma S5 - some features and figures
- Implementing AAC and USAC with Ecma S5

V. HAMBARDZUMYAN (1908 –1996)
ABOUT HIS „INVERSE PROBLEM“:

“In the simplest case of monochromatic and isotropic elementary scattering acts the method enables to replace the search for a family of solutions of a complicated linear integral equation by a numerical solution of a single and very simple nonlinear functional equation.“


FORWARD MODELING
Earth Response
Earth Model
INVERSE MODELING

MICHAEL A. GERZON (1945 – 1996):

Another development came soon after Orban released the stereo synthesiser. It was discovered that phasiness can be reduced by using more than one all-pass network. Hence, the reduced phase algorithm uses two all-pass networks with two identical ganged gains as seen in **Figure 4**:

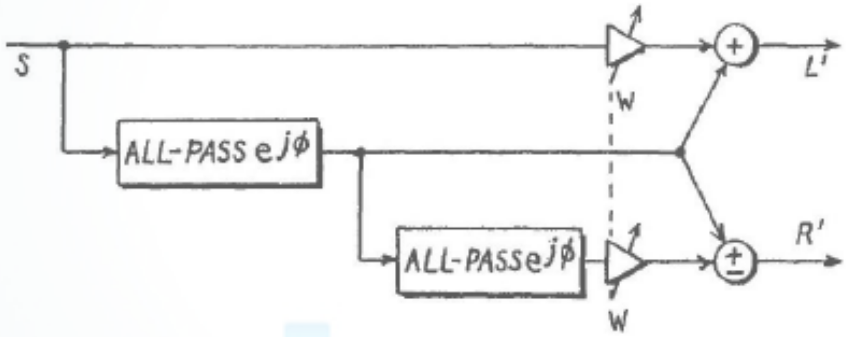
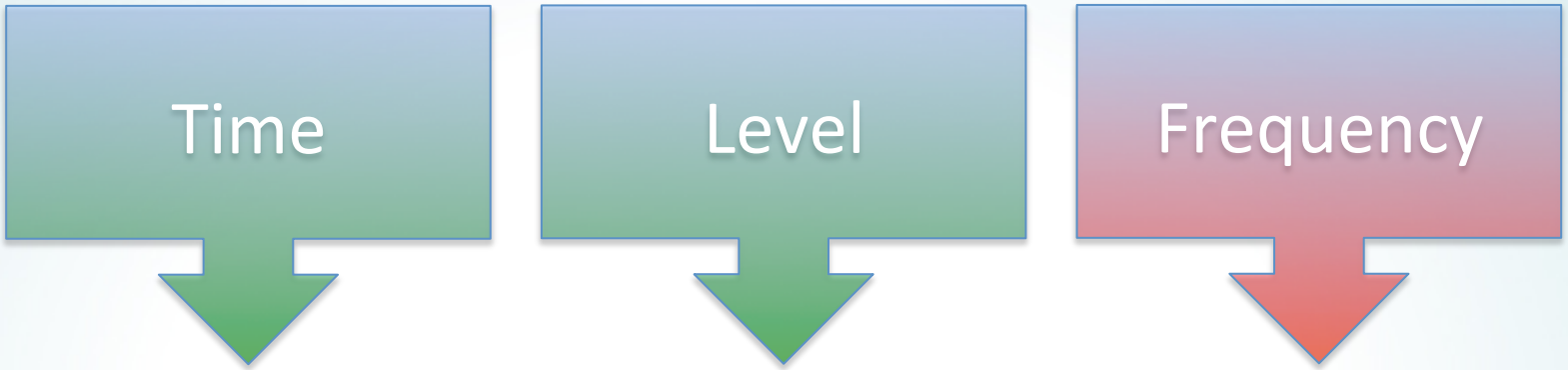


Figure 4. *Reduced-phase pseudostereo technique using two identical all-pass networks.*

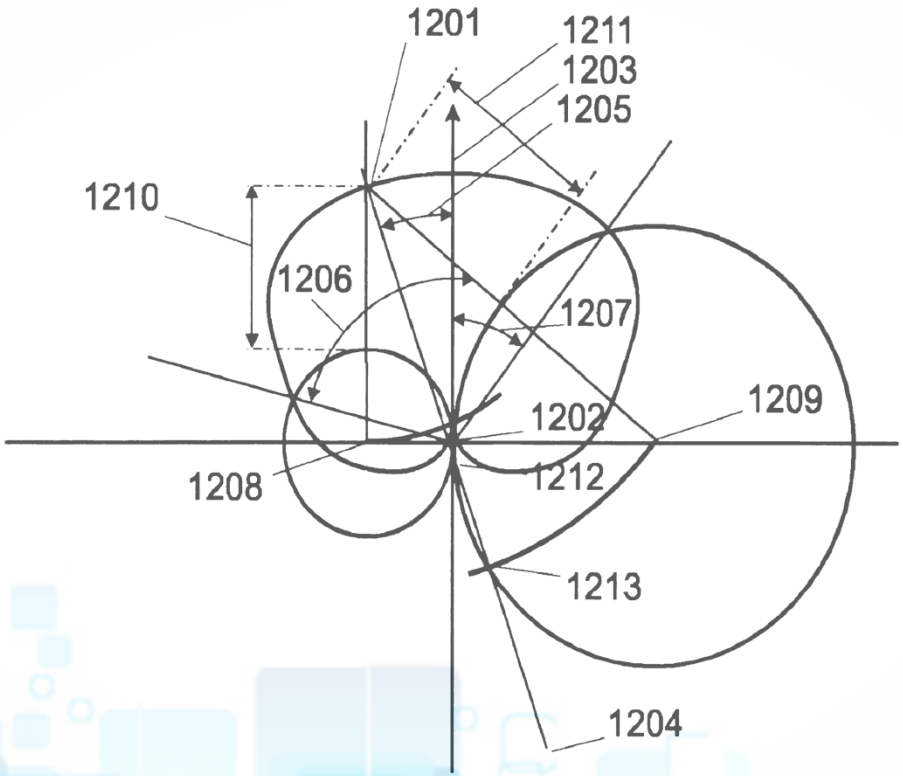
P. Montgomery: Pseudostereo Techniques. – University of Sydney 2007.



May lead to sound source separation, if appropriate.

Implies uniform localisation for all sound sources!

INVERSE CODING PRINCIPLE:



INVERSE CODING FORMULAE:

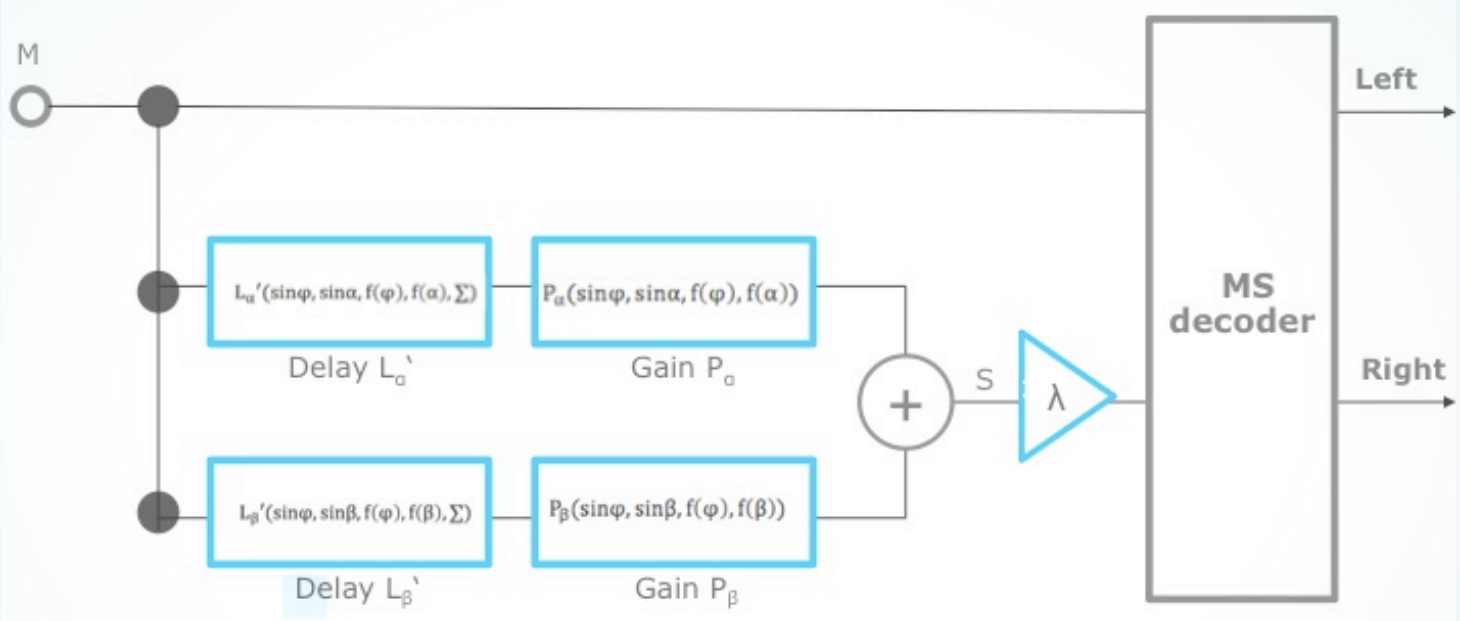
$$L_\alpha = -\frac{f(\alpha)}{2\sin\alpha} + \sqrt{\frac{f^2(\alpha)}{4\sin^2\alpha} + f^2(\varphi) - \frac{f(\alpha)}{\sin\alpha} * f(\varphi) * \sin\varphi}$$

$$L_\beta = -\frac{f(\beta)}{2\sin\beta} + \sqrt{\frac{f^2(\beta)}{4\sin^2\beta} + f^2(\varphi) + \frac{f(\beta)}{\sin\beta} * f(\varphi) * \sin\varphi}$$

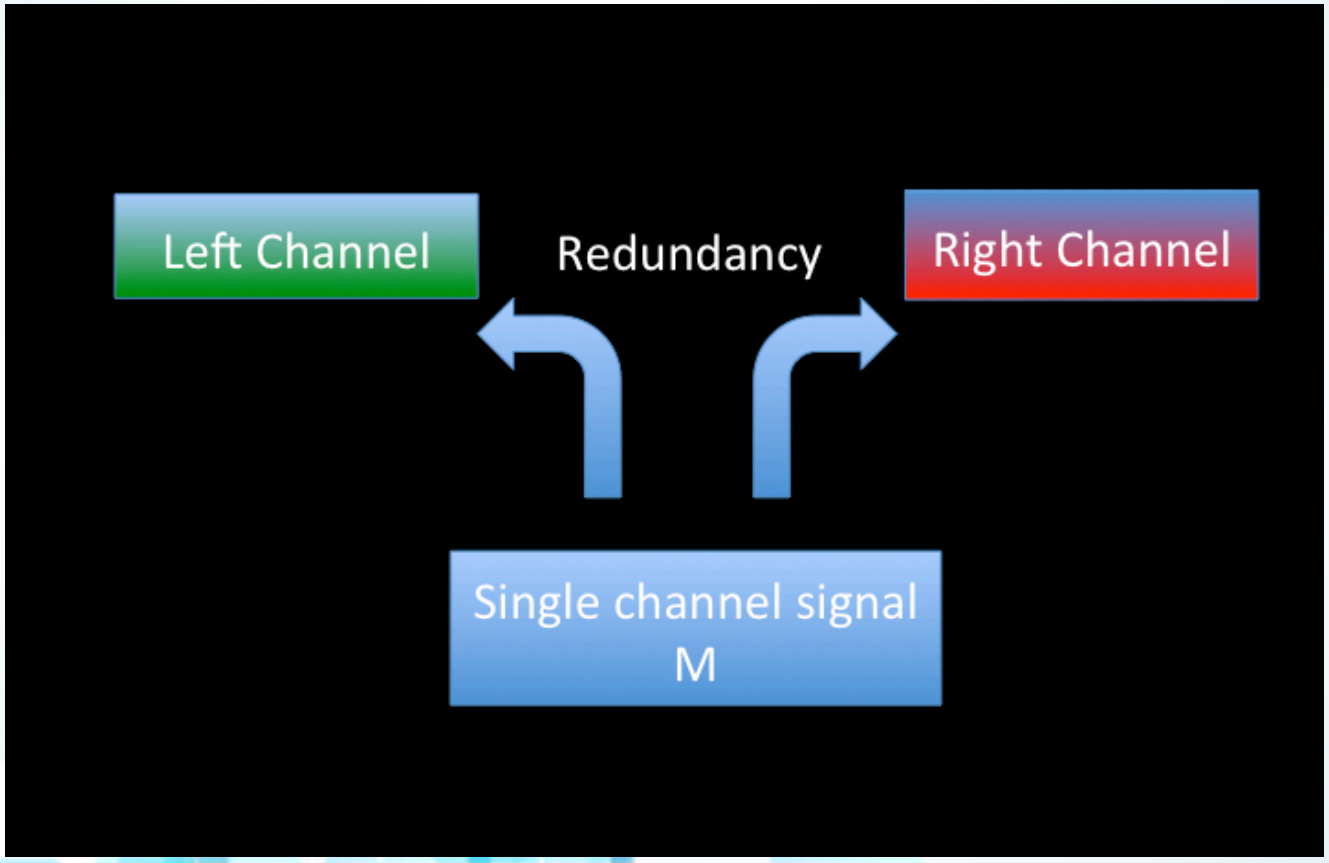
$$P_\alpha = \frac{f^2(\alpha)}{4\sin^2\alpha} + f^2(\varphi) - \frac{f(\alpha)}{\sin\alpha} * f(\varphi) * \sin\varphi$$

$$P_\beta = \frac{f^2(\beta)}{4\sin^2\beta} + f^2(\varphi) + \frac{f(\beta)}{\sin\beta} * f(\varphi) * \sin\varphi$$

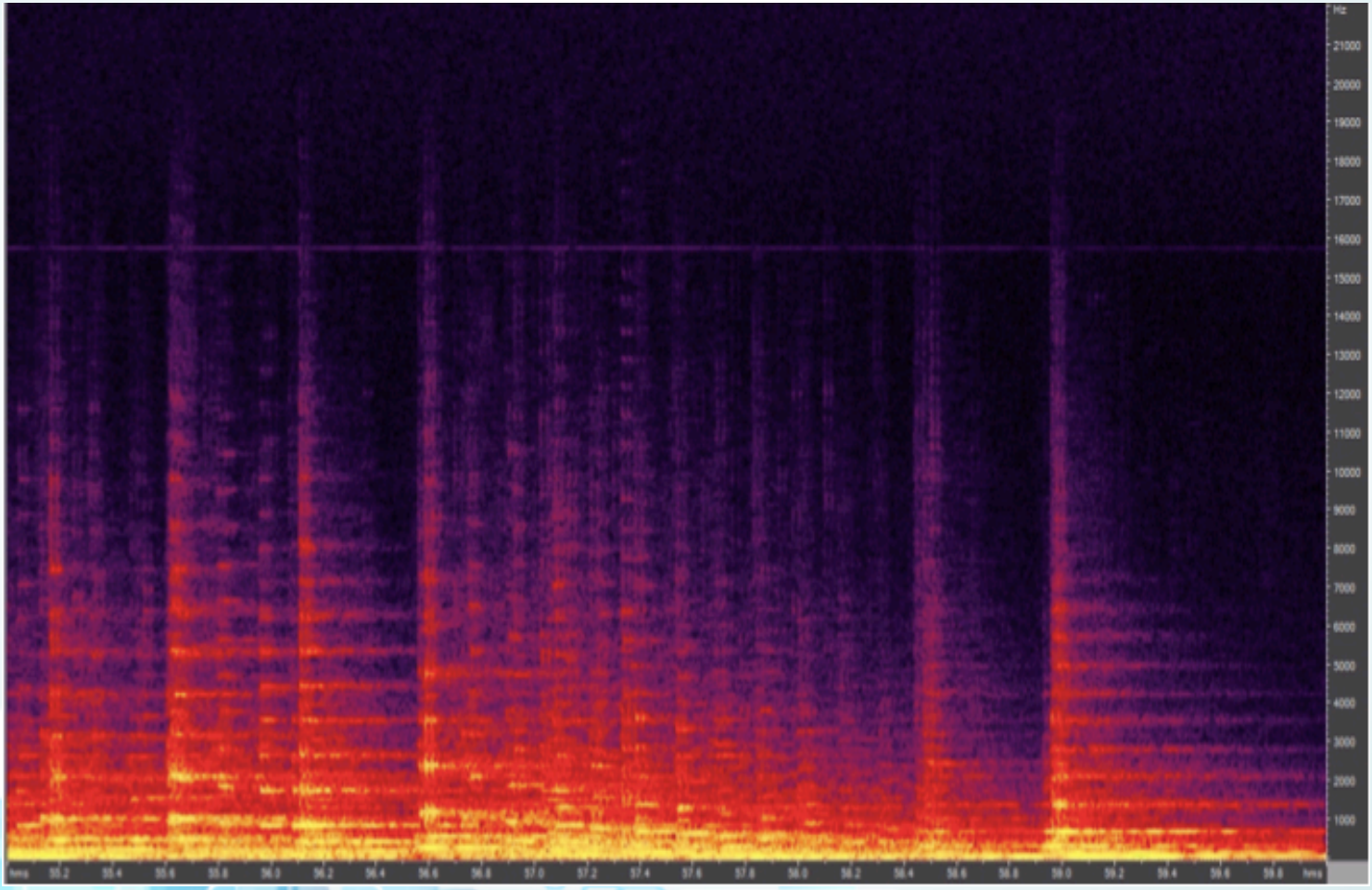
INVERSE CODING BLOCK DIAGRAM:



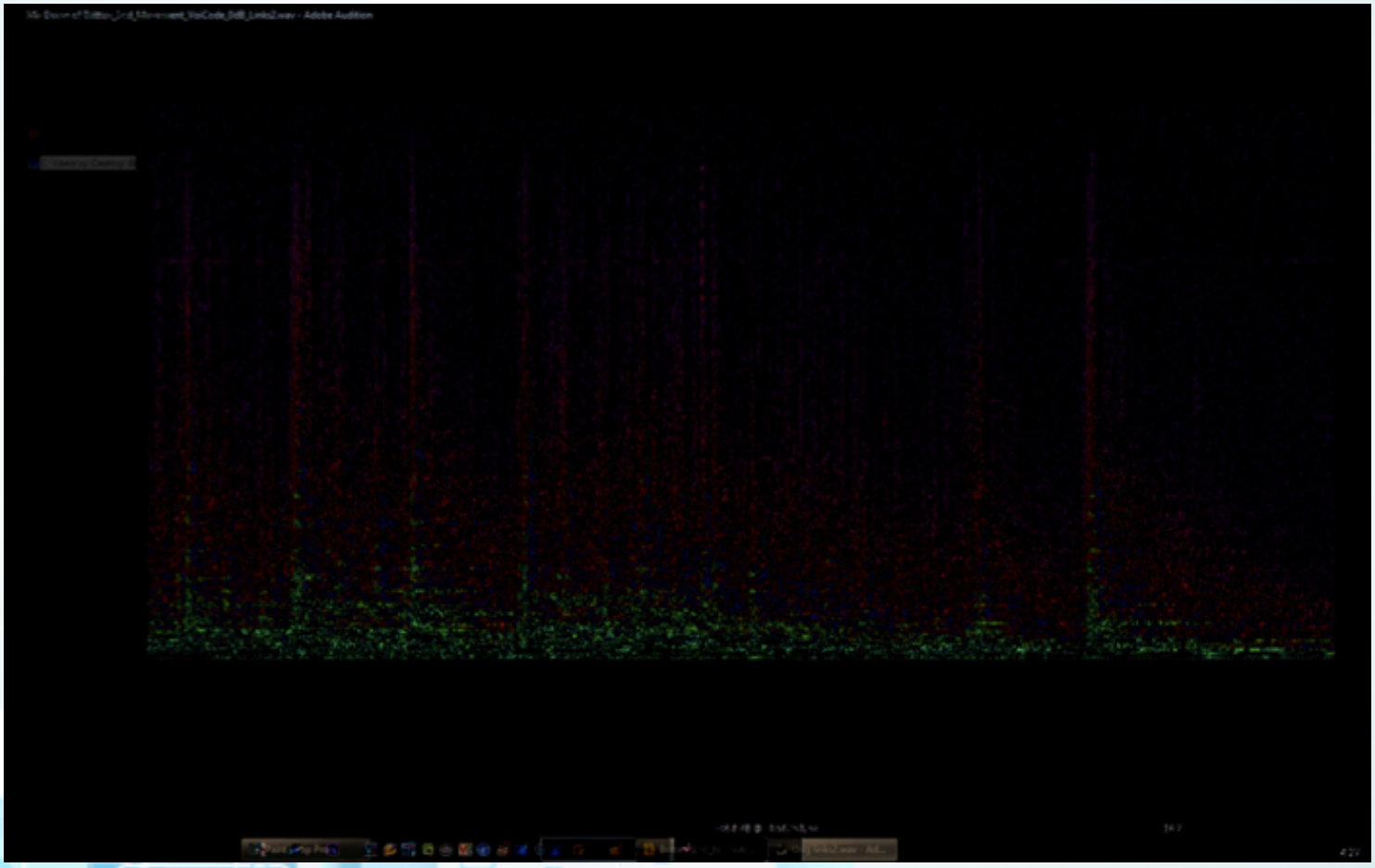
THE REDUNDANCY PROBLEM:



SPECTRAL BEHAVIOUR:



SPECTRAL BEHAVIOUR:



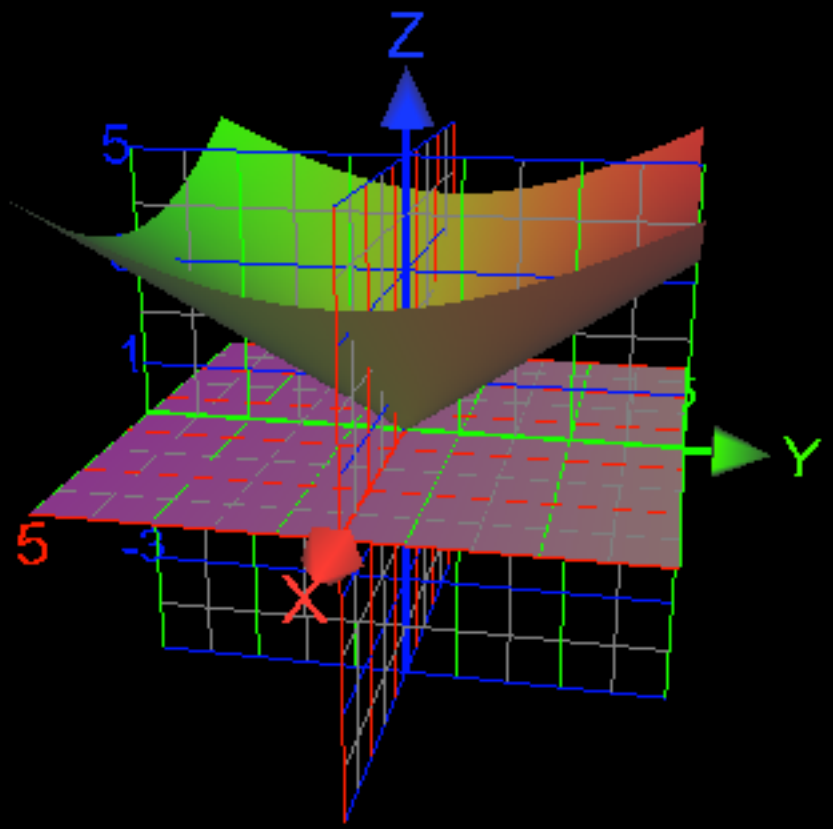
D. HILBERT (1862 –1943):

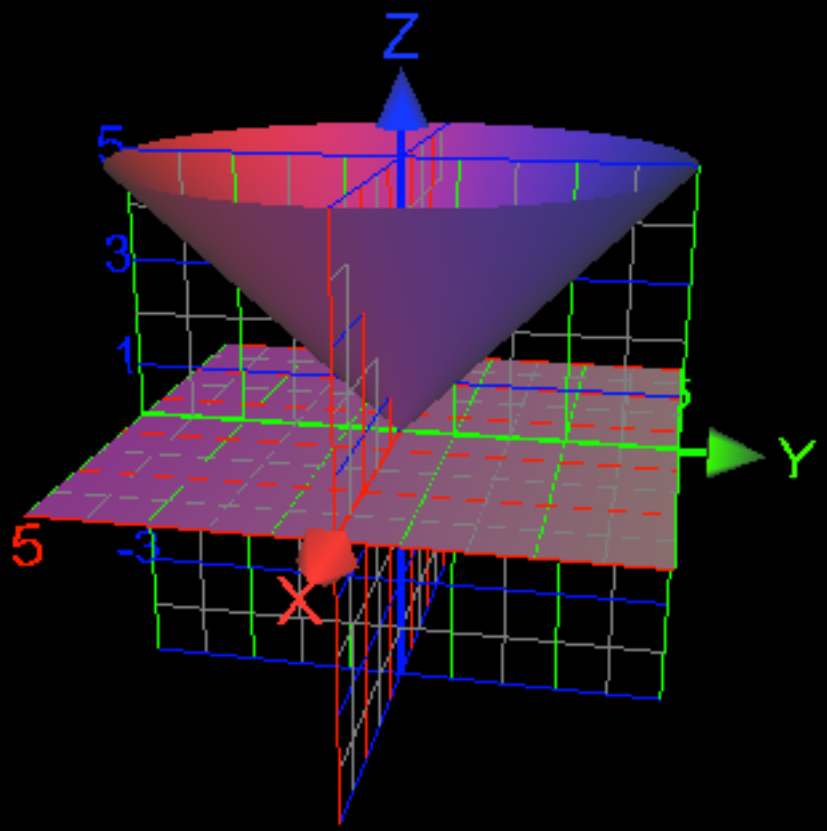
19. Über die vollen Invariantensysteme*.

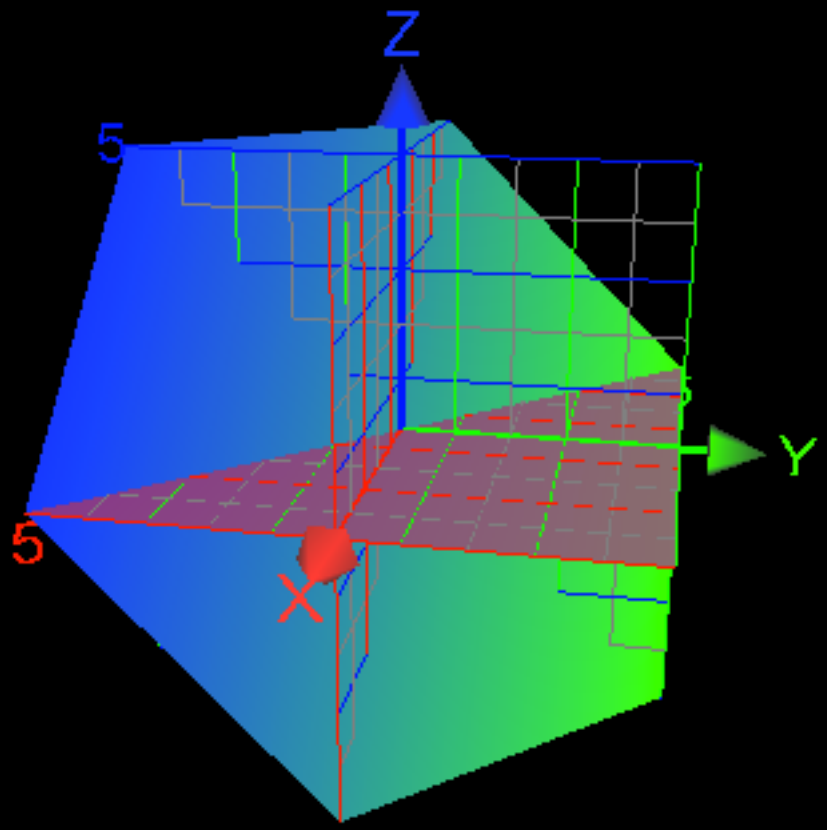
[Mathem. Annalen Bd. 42, S. 313—373 (1893).]

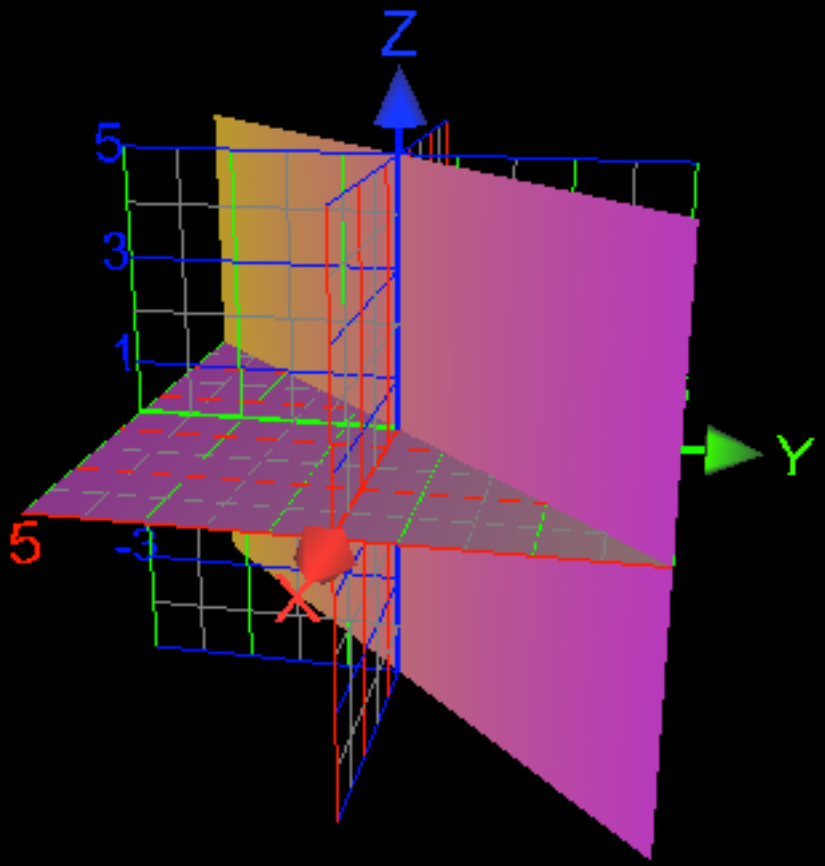
Einleitung.

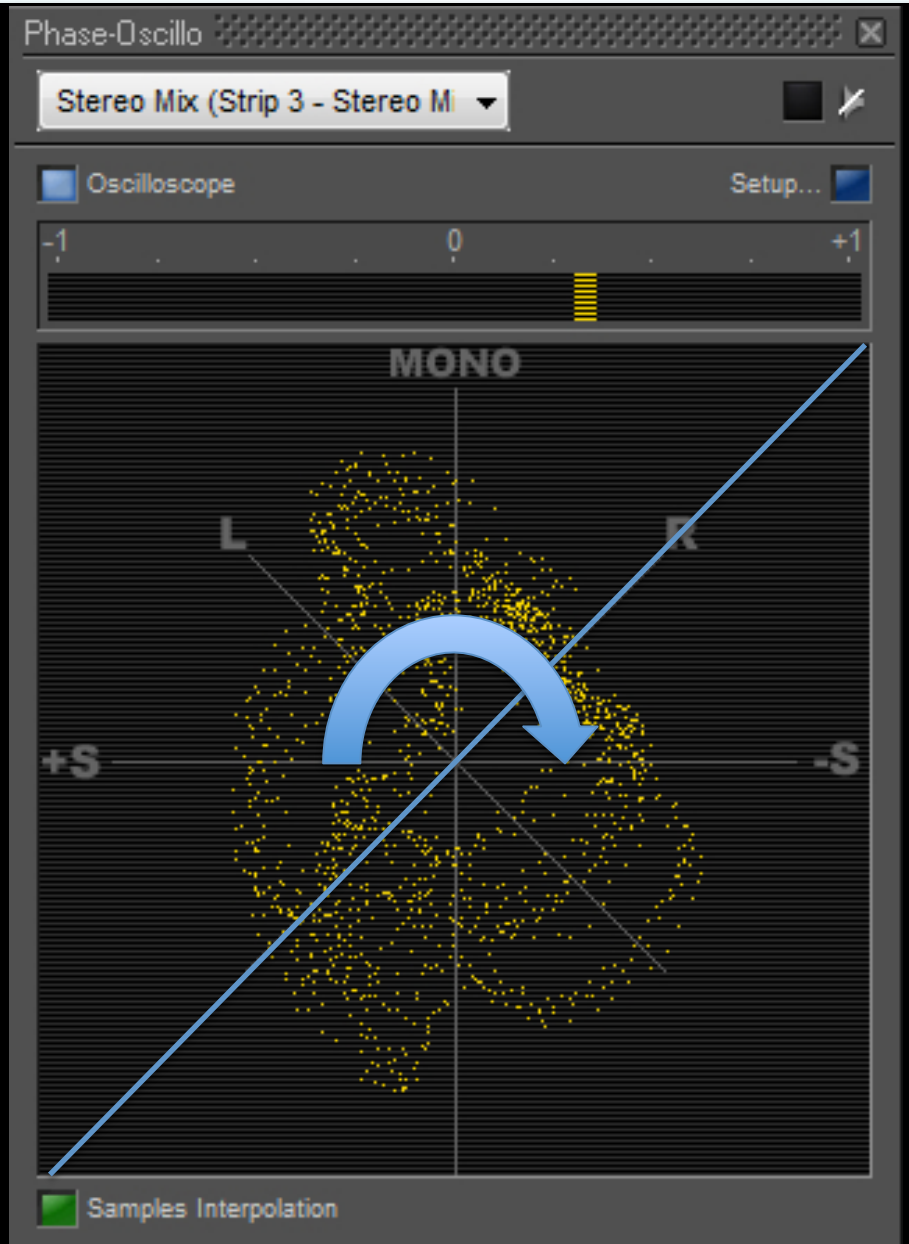
Meine Abhandlung „Über die Theorie der algebraischen Formen“[1] enthält eine Reihe von Theoremen, welche für die Theorie der algebraischen Invarianten von Bedeutung sind. Insbesondere in Abschnitt V der genannten Abhandlung habe ich mit Hilfe jener Theoreme für beliebige Grundformen die *Endlichkeit* des vollen Invariantensystems bewiesen. *Dieser Satz von der Endlichkeit des vollen Invariantensystems bildet den Ausgangspunkt und die Grundlage für die Untersuchungen der vorliegenden Abhandlung*[2]. Die im folgenden entwickelten Methoden unterscheiden sich wesentlich von den bisher in der Invariantentheorie angewandten Mitteln; bei den nachfolgenden Untersuchungen nämlich ordnet sich die Theorie der algebraischen Invarianten unmittelbar unter die allgemeine Theorie der algebraischen Funktionenkörper unter: so daß die Theorie der Invarianten lediglich als ein besonders bemerkenswertes Beispiel für die Theorie der algebraischen Funktionenkörper mit mehr Veränderlichen erscheint — gerade wie man in der Zahlentheorie die Theorie der Kreisteilungskörper lediglich als ein besonders bemerkenswertes Beispiel aufzufassen hat, an welchem die wichtigsten Sätze der Theorie der allgemeinen Zahlenkörper zuerst erkannt und bewiesen worden sind.


Z
5
1
Y
5
-3
X


Z
5
3
1
Y
5
-3
X


Z
5
Y
5
X


Z
5
3
1
Y
5
-3
X


Phase-Oscillo
Stereo Mix (Strip 3 - Stereo M
Oscilloscope
Setup...
-1
0
+1
MONO
L
R
+S
-S
Samples Interpolation

Ecma/TC32-TG22/2013

ecma INTERNATIONAL

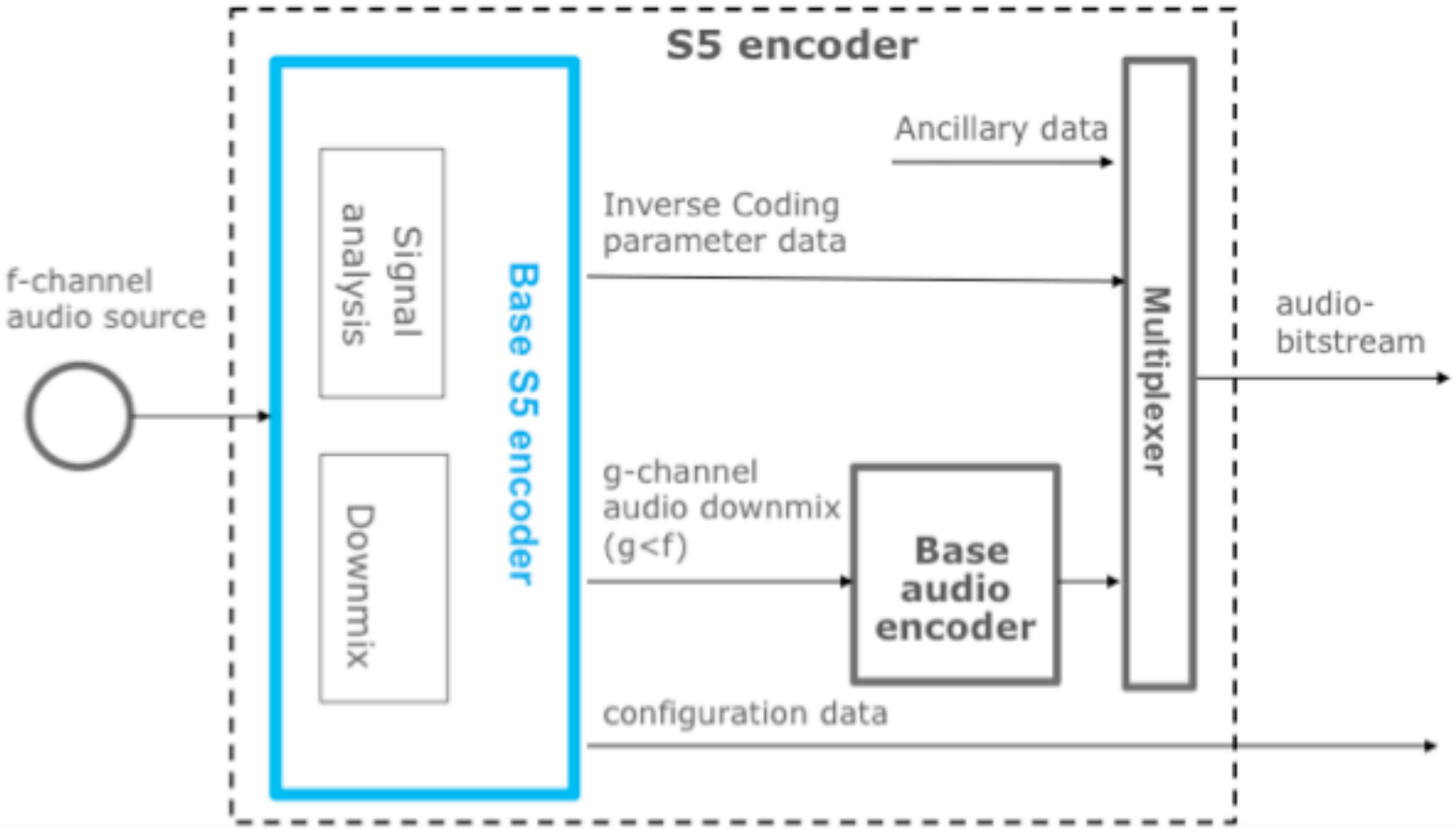
Scalable Sparse Spatial Sound System (S5)

Energy Efficient IT

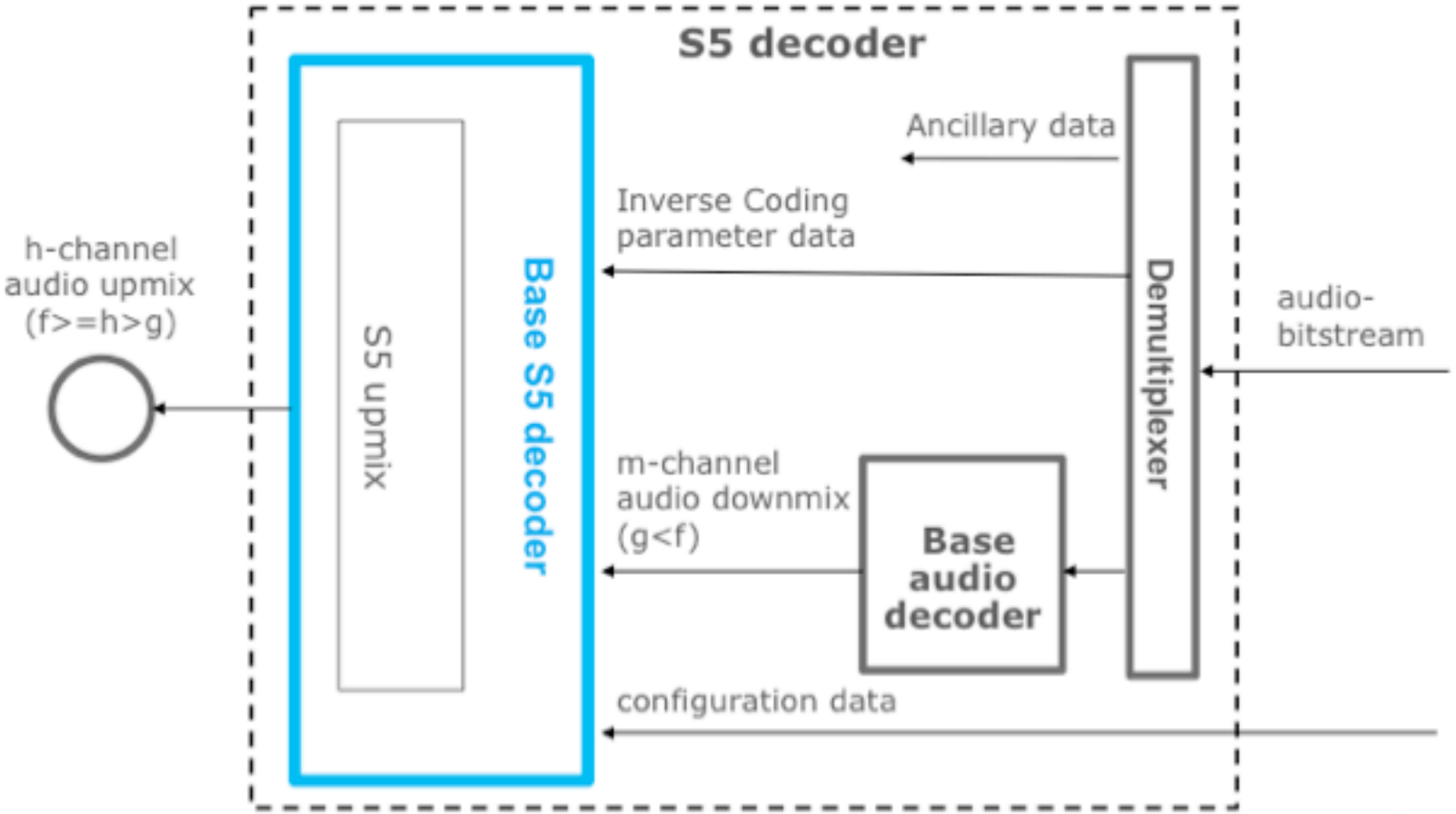
What is S5 about?

- **S5 is a channel-based digital coding system for spatial audio signals based on**
  - *State-of-the-art audio coding for presenting base audio,*
  - *Inverse Coding for retrieving spatial (localization & ambiance) audio.*
- **Inverse Coding is a new mathematical method**
  - *To derive spatial audio data from base audio using an associated sparse data set of parameter values.*
- **Compared to state-of-the-art spatial audio S5 offers**
  - *Significantly higher compression*
  - *Significantly reduced computational complexity*
  - *Low latency processing*

What is the S5 encoder architecture?



What is the S5 decoder architecture?



How does Inverse Coding work?

- ***Multichannel audio is downmixed to reduce data. Sparse Inverse Coding parameters contains the information for spatial audio reconstruction.***
- ***Inverse Coding means time-level modelling of spatial sound with delays and gains as output parameters.***
- ***Inverse Coding uses these gains and delays to reconstruct the multichannel audio signal from the downmix.***
- ***The use of algebraic invariants reduces the complexity significantly and enables real-time performance.***
- ***Unlike pseudo–stereo solutions Inverse Coding provides sound source separation even at the same frequency.***

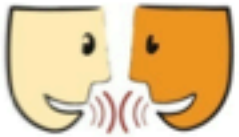
What are the benefits of S5?

- **Adds enhanced performance to state-of-the-art spatial audio coding**
  - *Severly reduced data at a given audio quality*
  - *Increased quality at a given data rate/volume,*
  - *Less energy consumption at a given quality,*
  - *Reduced hardware constraints at a given quality.*
- **S5 parameter data may be embedded into the base audio stream to achieve compatability to non-S5 decoders.**
- **The ratio of parameter data and base audio channels is scalable according to application requirements.**
- **Enables spatial audio for conversational/interactive applications due to low latency of parameter data.**

Which market opportunities does S5 provide?

- **High-quality scenarios**

  - *Broadcast/multicast applications and multimedia storage and retrieval.*

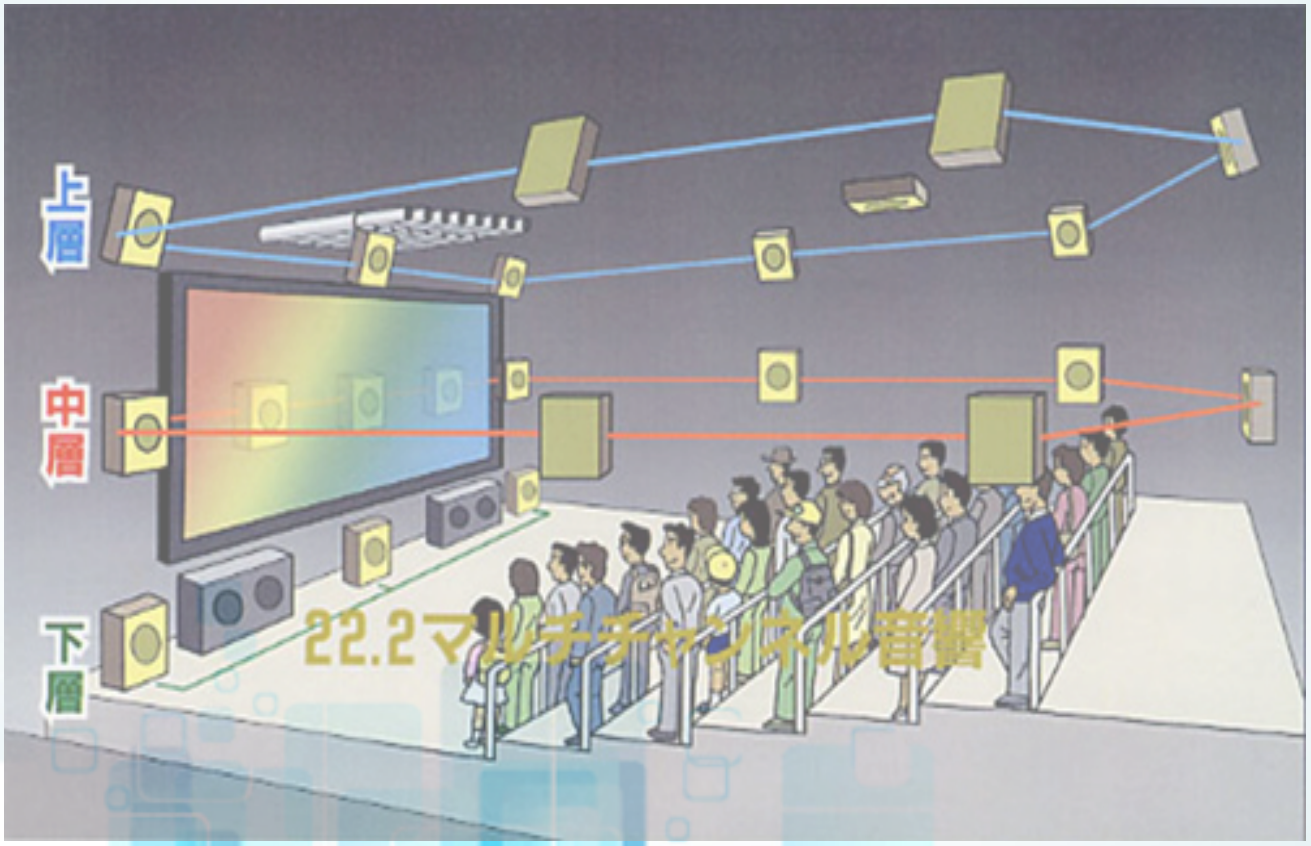


- **Low latency scenarios**
  - *Conversational/interactive communication including teleconferencing and gaming applications*

- **Low-power scenarios**
  - *Universal applicability with all stereo, surround or 3D formats including 3D up-mixing of mono content*

Rue du Rhône 114
CH-1204 Geneva
T: +41 22 849 6000
F: +41 22 849 6001

ECMA S5 may be configured for NHK 22.2 experience at 48kbps and above with shortest latency:



ECMA S5 FEATURES

- Easy integration of lossless and lossy base audio codecs
- Easy integration of audio coding libraries
- Modular and extendable configuration data
- Registry for market-relevant loudspeaker configurations
- Ecma S5 reference: Tables 7 and 8 in ISO/IEC 23001-8, *Information technology -- MPEG systems technologies -- Part 8: Coding-independent code points.*
- Modern fingerprint synchronisation
- Ancillary data option e.g. for loudness and broadcaster-defined data
- Internal and external multiplex

FUTURE ECMA TC32-TG22 ACTIVITIES

- Ecma TC32-TG22 work will focus on implementation of Ecma S5 with AAC and with USAC.
- Current liaisons: ISO/IEC JTC1/SC29/WG11, AES X-212

- **Active contribution from industry is welcome!**

THANK YOU

Name Clemens Par

E-mail mail@swissaudec.com