

# Surround Sound

## Grundbegriffe, Methoden, Aufnahmetechniken

Schwerpunkt: Ambisonics

Jörn Nettingsmeier, Essen

[nettings@stackingdwarves.net](mailto:nettings@stackingdwarves.net)  
<http://stackingdwarves.net>

# Grundbegriffe

- Schallfeld
- Lokalisation
- Lokalisationsreize
- Koordinaten
- Dimensionen

# Grundbegriffe: **Schallfeld**

Eine Klangsituation läßt sich physikalisch durch ein Schallfeld beschreiben.

# Grundbegriffe: Schallfeld

Eine Klangsituation läßt sich physikalisch durch ein Schallfeld beschreiben.

## Kirchhoff-Helmholtz-Integral:

***„Wenn man den Schalldruck und die Schallschnelle auf jedem Punkt der Oberfläche eines quellen-freien Volumens kennt, hat man eine vollständige Beschreibung des inneren Schallfeldes.“***

# Grundbegriffe: Schallfeld

Konkret: unendlich viele Mikrofone um die Hörposition herum, unendlich viele Lautsprecher um den Zuhörer.

## Kirchhoff-Helmholtz-Integral:

***„Wenn man den Schalldruck und die Schallschnelle auf jedem Punkt der Oberfläche eines quellen-freien Volumens kennt, hat man eine vollständige Beschreibung des inneren Schallfeldes.“***

# Grundbegriffe: Lokalisation

Die Hörerin soll Schallquellen verorten können, entweder durch

- genaue physikalische Simulation des entsprechenden Schallfeldes, oder
- mit Hilfe einer überzeugenden Illusion.

# Grundbegriffe: Lokalisation

Die Höre  ten können,  
entweder durch

- genaue physikalische Simulation des entsprechenden Schallfeldes, oder
- mit Hilfe einer überzeugenden Illusion.

# Grundbegriffe: Lokalisation

Die Lokalisation kann mit geringem Aufwand, aber: können,  
entweder inter-individuell unterschiedliche  
Wahrnehmung, Abhängigkeit vom Hörplatz

- genaue physikalische Simulation des entsprechenden Schallfeldes, oder
- mit Hilfe einer überzeugenden Illusion.



# Grundbegriffe: Lokalisation

reale Situation



# Grundbegriffe: Lokalisation

physikalische  
Simulation, z.B.  
durch einen  
Lautsprecher am  
Ort des Flügels

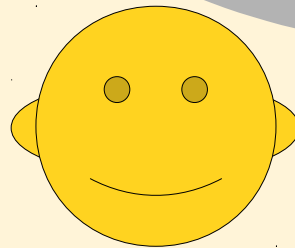


(nur bei geringer  
Quellenanzahl praktikabel)

# Grundbegriffe: Lokalisation



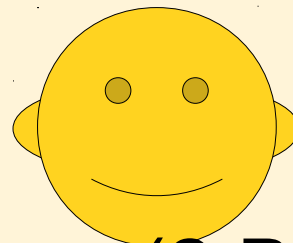
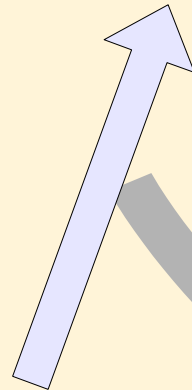
physikalische Simulation  
durch Wellenfeldsynthese



(für beliebig viele Quellen)

# Grundbegriffe: Lokalisation

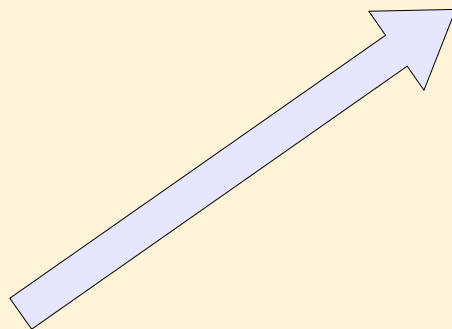
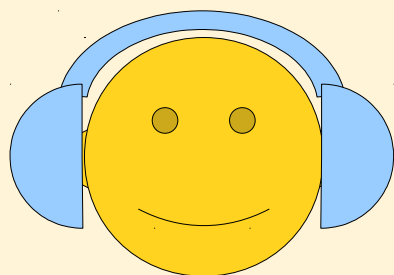
*Phantom-  
schallquelle*



psychoakustische  
Illusion durch  
zwei unterschied-  
lich laute Boxen

(2 Boxen erzeugen beliebig  
viele Phantomquellen)

# Grundbegriffe: Lokalisation

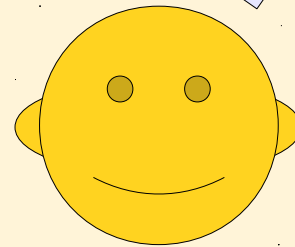


*virtuelle  
Schallquelle*

psychoakustische  
Illusion durch  
binaurale Wiedergabe  
über Kopfhörer

# Grundbegriffe: Lokalisation

Simulation im  
Bassbereich,



Illusion im Mittel-  
hochtonbereich:



= **Ambisonics**

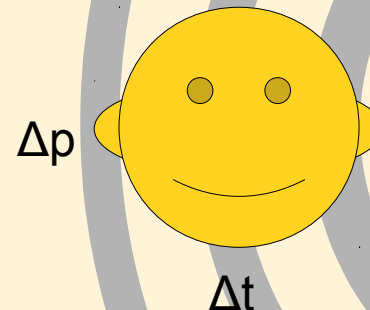


# Grundbegriffe: Lokalisationsreize

Das Gehör verarbeitet zur Lokalisation verschiedene Hinweise (*cues*):

- *interaural time difference* (ITD)
- *interaural level difference* (ILD)
- Klangverfärbungen durch frequenzabhängige Abschattung, Beugung, Absorption und Reflexion an Kopf und Ohrmuscheln (*head and pinnae cues*)

Der stärkste Reiz dominiert die Wahrnehmung.



# Grundbegriffe: Lokalisationsreize

ILD: Intensitätsstereophonie (XY, MS)

ITD: Laufzeitstereophonie (AB)

Kopf: binaurale (kopfbezogene) Stereophonie

Alle 3 Reize lassen sich durch *head-related transfer functions* (HRTFs) simulieren.



# Grundbegriffe: **Koordinaten**

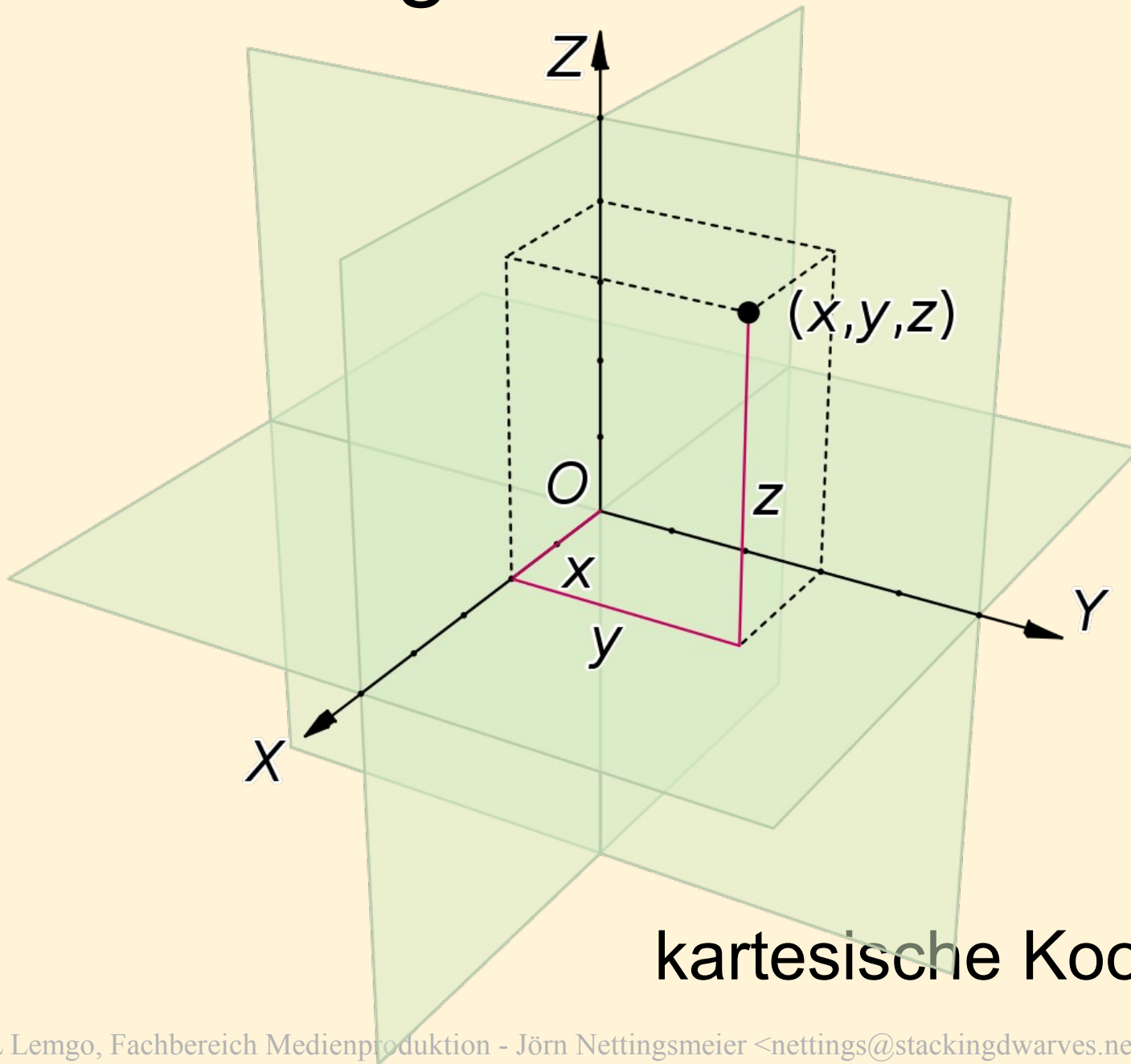
Beschreibung des Ortes einer Schallquelle:

# Grundbegriffe: **Koordinaten**

Beschreibung des Ortes einer Schallquelle:

- Distanz in x-, y- und z-Richtung:  
**kartesische Koordinaten**

# Grundbegriffe: **Koordinaten**



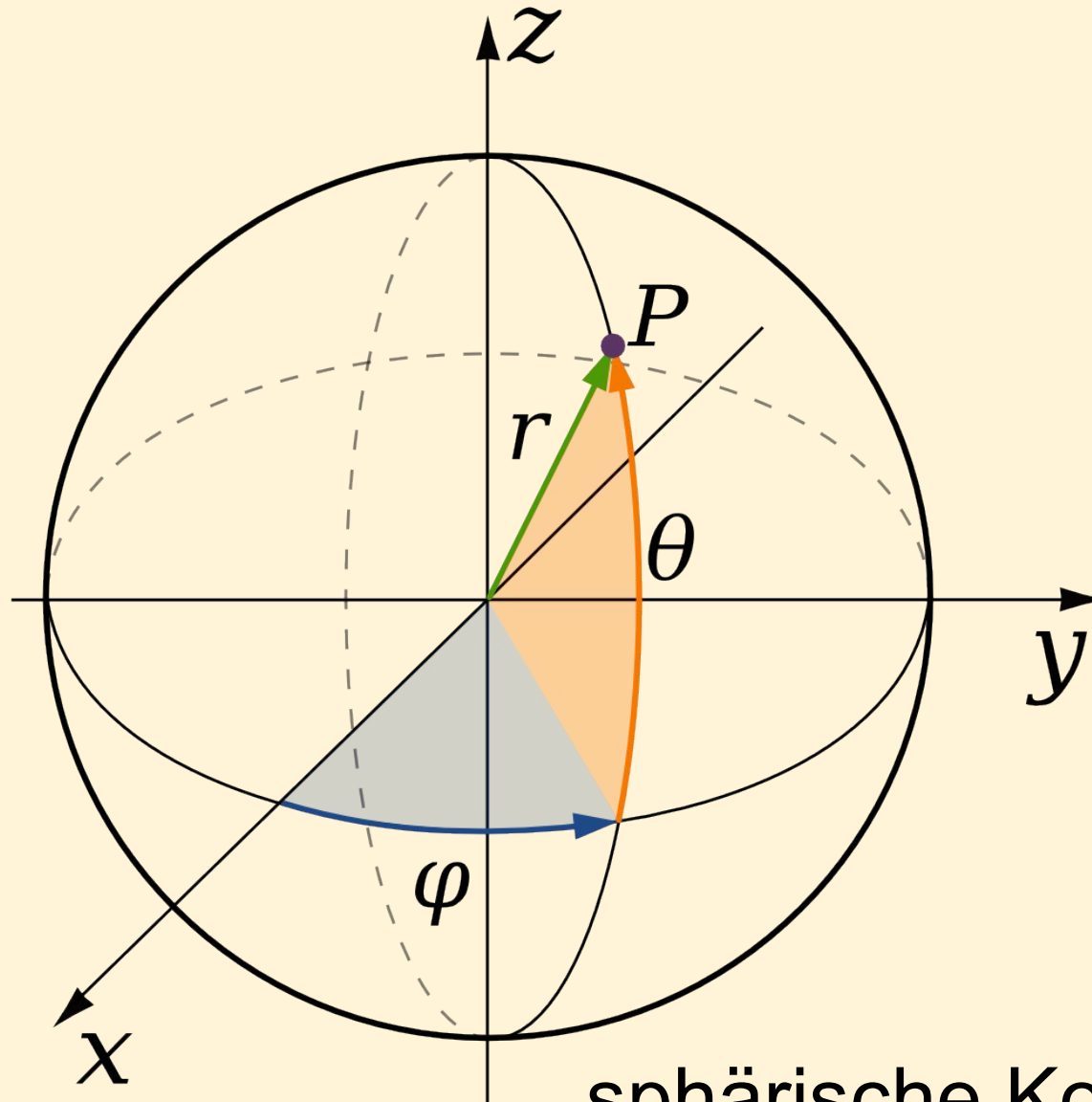
**kartesische Koordinaten**

# Grundbegriffe: **Koordinaten**

Beschreibung des Ortes einer Schallquelle:

- Distanz in x-, y- und z-Richtung:  
**kartesische Koordinaten**, oder
- Azimuthwinkel  $\varphi$  (phi), Elevationswinkel  $\theta$  (theta), Radius  $r$ :  
**sphärische oder Polarkoordinaten**

# Grundbegriffe: Koordinaten



sphärische Koordinaten

# Grundbegriffe: Koordinaten

In sphärischen Koordinaten sind Richtungen in begrenzten Intervallen darstellbar:

$$0^\circ \leq \varphi \leq 360$$

$$-90^\circ \leq \theta \leq 90^\circ$$

Die Winkel sind periodisch.

Nur  $r$  ist einseitig unbegrenzt:

$$0 \leq r \leq \infty$$

Für kartesische Koordinaten dagegen gilt

$$-\infty \leq x, y, z \leq \infty$$

# Grundbegriffe: Koordinaten

In sphärischen Koordinaten sind Richtungen in begrenzten Intervallen darstellbar:

$$0^\circ \leq \varphi \leq 360$$

$$-90^\circ \leq \theta \leq 90^\circ$$

Die Winkel sind periodisch.

praktisch für Panner!

Der Radius ist einseitig unbegrenzt:

$$0 \leq r \leq \infty$$

Für kartesische Koordinaten dagegen gilt

$$-\infty \leq x, y, z \leq \infty$$

# Grundbegriffe: Koordinaten

In sphärischen Koordinaten sind Richtungen in begrenzten Intervallen darstellbar:

$$0^\circ \leq \varphi \leq 360$$

$$-90^\circ \leq \theta \leq 90^\circ$$

Die Distanz ist direkt ablesbar!

Die Winkel sind periodisch.

Nur  $r$  ist einseitig unbegrenzt:

$$0 \leq r < \infty$$

Für kartesische Koordinaten dagegen gilt

$$-\infty \leq x, y, z \leq \infty$$



# Grundbegriffe: Koordinaten

In sphärischen Koordinaten sind die Winkel in begrenzten Bereichen definiert.

Richtungen und Distanzen sind nur über trigonometrische Operationen ermittelbar.

Daher werden kartesische K. in der Akustik nur selten verwendet.

Die Winkel  $\theta$  und  $\phi$  sind

Nur  $r$  ist einseitig begrenzt:

$$0 \leq r < \infty$$

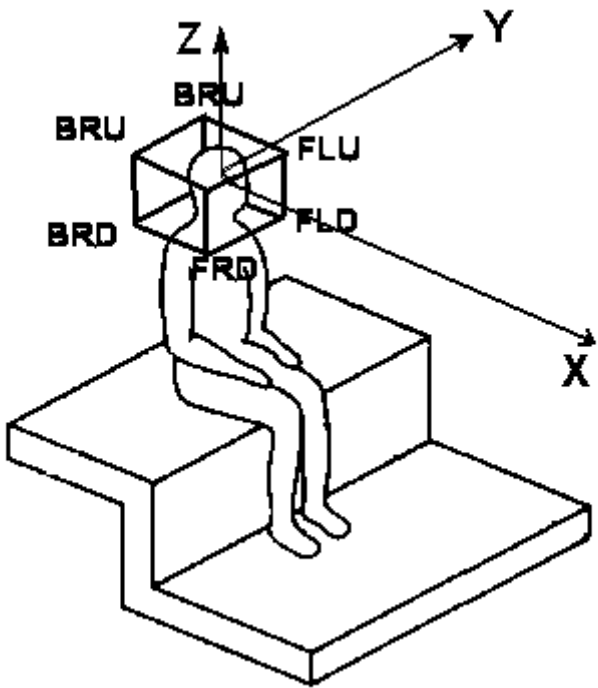
Für kartesische Koordinaten dagegen gilt

$$-\infty \leq x, y, z \leq \infty$$

# Grundbegriffe: **Koordinaten**

Konvention der Achsen:

Rechtssystem, hörerzentriert



X: links positiv  
Y: vorn positiv  
Z: oben positiv

(Illustration: Angelo Farina)

# Grundbegriffe: Dimensionen

Alltagsbegriffe wie „Dimension“, „3D“ etc. sind durch Marketing-Sprache geprägt und irreführend.

Mit welchen (und wie vielen) Dimensionen arbeiten wir eigentlich?

# Grundbegriffe: Dimensionen

Arbeitsdefinition:

Dimensionen = Freiheitsgrade der Translation  
(Translation = Ortsveränderung)

# Grundbegriffe: Dimensionen

Punkt: 0-dimensional

# Grundbegriffe: Dimensionen

Punkt: 0-dimensional

Linie: 1-dimensional (<- unendlich!)

# Grundbegriffe: Dimensionen

Punkt: 0-dimensional

Linie: 1-dimensional (<- unendlich!)

Also: Stereo-Ortung ist eingeschränkt 1-dimensional (nicht unendlich, begrenzt durch den Winkel zwischen den Boxen).

# Grundbegriffe: Dimensionen

Punkt: 0-dimensional

Linie: 1-dimensional (<- unendlich!)

Also: Stereo-Ortung ist eingeschränkt 1-dimensional (nicht unendlich, begrenzt durch den Winkel zwischen den Boxen).

**Tiefensimulation** erfolgt separat durch:

- Pegelabsenkung ( $1/r^2$ -Punktquelle)
- Höhendämpfung (Luftabsorption)
- Anhebung des Hallanteils



# Grundbegriffe: Dimensionen

Fläche: 2-dimensional

# Grundbegriffe: Dimensionen

Fläche: 2-dimensional

Aber: in horizontalem Surround bewegen sich Schallquellen auf einer Kreislinie.

# Grundbegriffe: Dimensionen

Fläche: 2-dimensional

Aber: in horizontalem Surround bewegen sich Schallquellen auf einer Kreislinie.

Also: horizontaler Surround ist ebenfalls nur 1-dimensional, allerdings ohne Einschränkung.

# Grundbegriffe: Dimensionen

Fläche: 2-dimensional

Aber: in horizontalem Surround bewegen sich Schallquellen auf einer Kreislinie.

Also: horizontaler Surround ist ebenfalls nur 1-dimensional, allerdings ohne Einschränkung.

Nur mit **Tiefensimulation** kann man sinnvoll von 2D-Surround sprechen.

# Grundbegriffe: Dimensionen

Raum: 3-dimensional

# Grundbegriffe: Dimensionen

Raum: 3-dimensional

Aber: in sogenannten „3D“-Surroundverfahren bewegen sich die Quellen auf einer Kugelfläche.

# Grundbegriffe: Dimensionen

Raum: 3-dimensional

Aber: in sogenannten „3D“-Surroundverfahren bewegen sich die Quellen auf einer Kugelfläche.

Also: „3D“-Surround ist nur 2-dimensional, mit den *Freiheitsgraden*  $\varphi$  und  $\theta$ .

# Grundbegriffe: Dimensionen

Raum: 3-dimensional

Aber: in sogenannten „3D“-Surroundverfahren bewegen sich die Quellen auf einer Kugelfläche.

Also: „3D“-Surround ist nur 2-dimensional, mit den *Freiheitsgraden*  $\varphi$  und  $\theta$ .

**Tiefensimulation** ergänzt das System mit dem dritten Freiheitsgrad  $r$  zu echter Dreidimensionalität.



# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

- eingeschränkt horizontal

# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

- eingeschränkt horizontal
- eingeschränkt vertikal

# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

- eingeschränkt horizontal
  - eingeschränkt vertikal
- eingeschränkt in der Tiefe

# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

- eingeschränkt horizontal
    - eingeschränkt vertikal
  - eingeschränkt in der Tiefe
- = „Guckkastentheater“, frontal

# Grundbegriffe: Dimensionen

Wie 3D ist eigentlich 3D-Kino?

- eingeschränkt horizontal
    - eingeschränkt vertikal
  - eingeschränkt in der Tiefe
- = „Guckkastentheater“, frontal

Dagegen umfasst 3D-Ton den gesamten erlebbaren Raum um den Hörer herum.

# Grundbegriffe: Dimensionen

Also: 3D-Sound ist eine wichtige Erweiterung des nur eingeschränkt dreidimensionalen visuellen Kino-Erlebnisses.

# Methoden

- diskreter Mehrkanalton
- Wellenfeldsynthese
- Ambisonics

Vergleich der Signal-Repräsentationen



# Methoden: **diskreter Mehrkanalton**

Vertreter:

Quadrophonie, DTS 5.1, 7.1, 10.2, 22.2, ...

Erweiterung der Stereo-Idee:  
diskretes Mischen auf Lautsprecherpaare

# Methoden: **diskreter Mehrkanalton**

## kanalbasierte Repräsentation:

An diskreten Orten wird das Schallfeld abgetastet  
oder aus Mono-Quellen durch paarweises  
Panning ein willkürliches Schallfeld konstruiert.

Jeder Kanal entspricht einem Lautsprecher-  
Signal.

# Methoden: **diskreter Mehrkanalton**

## Vorteile:

- starke Marktpräsenz
- stabile Vorne-Ortung

## Nachteile:

- anisotropische Wiedergabe (bevorzugte Richtungen)
- Phantomschallquellen „kleben“ an LS
- unflexible Lautsprecher-Anordnung
- Formate untereinander inkompatibel

# Methoden: Wellenfeldsynthese



Huygens'sches Prinzip:

Jede Wellenfront kann in eine Folge von  
Elementarwellen  
zerlegt werden.

# Methoden: Wellenfeldsynthese



Huygens'sches Prinzip:

Jede Wellenfront kann in eine Folge von  
Elementarwellen  
zerlegt werden.

Umgekehrt: eng angeordnete Lautsprecher  
können beliebig geformte Wellenfronten  
synthetisieren.

# Methoden: **Wellenfeldsynthese**

objektbasierte Repräsentation:

diskrete (monophonische) Schallereignisse  
plus Metadaten (Position, Bewegung)

Es gibt keine geschlossene Signal-  
Repräsentation.

Das Schallfeld wird aus Einzelsignalen und einer  
parametrisierten Beschreibung (out-of-band!)  
generiert.

# Methoden: Wellenfeldsynthese



## Vorteile:

- sehr stabile Ortung ohne sweet spot
- Möglichkeit, fokussierte Quellen zu erzeugen
- physikalisch korrekt (siehe: Doppler-Effekt!)

## Nachteile:

- Materialschlacht
- Content muss für jedes System neu gerendert werden

# Methoden: **Ambisonics**

Idee:  
räumliches Sampling

auch: Erweiterung von MS-Stereo  
in drei Raumrichtungen

Das Signal (B-Format) ist  
lautsprecherunabhängig!

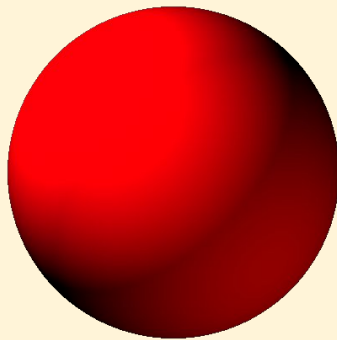


# Methoden: Ambisonics

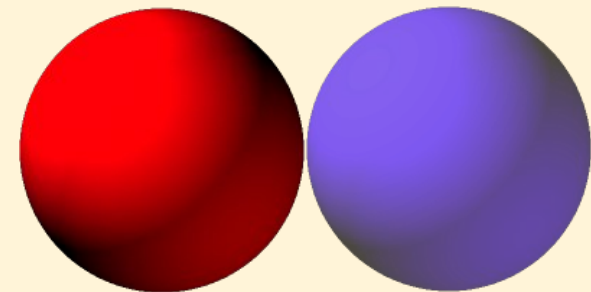
## Exkurs M/S-Stereophonie:

Mittensignal + Seitensignal

M



S



Linker Kanal:  $M + k \cdot S$

Rechter Kanal:  $M - k \cdot S$

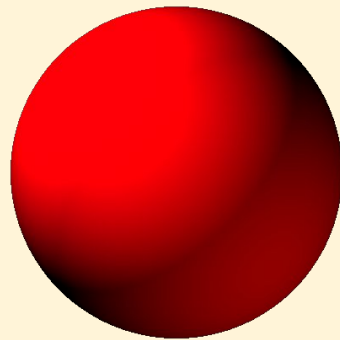
Basisbreite über  $k$  variabel !

# Methoden: Ambisonics

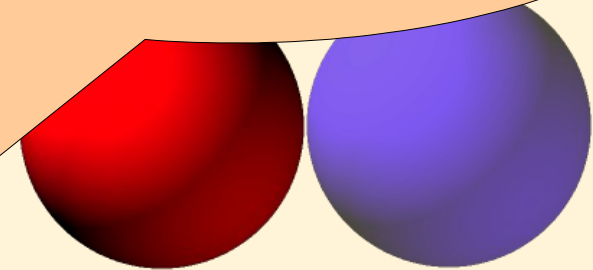
## Exkurs M/S-Stereophonie

Mittensignal

M



MS (und auch Ambisonics)  
muss zur Wiedergabe  
decodiert werden!



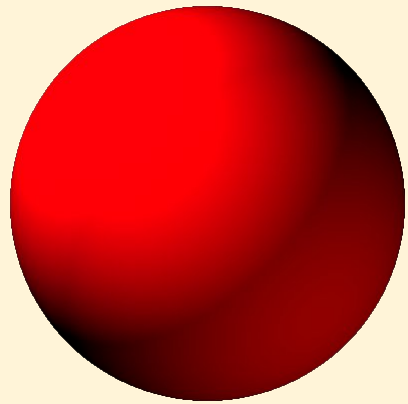
Linker Kanal:  $M + k \cdot S$

Rechter Kanal:  $M - k \cdot S$

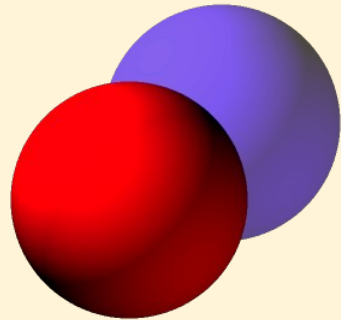
Basisbreite über  $k$  variabel !

# Methoden: Ambisonics

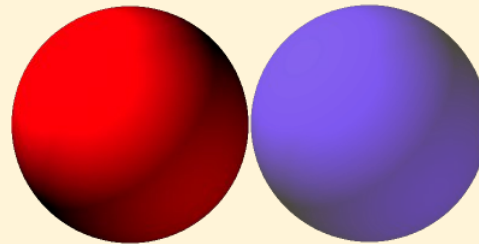
mono + vorn/hinten + links/rechts + oben/unten



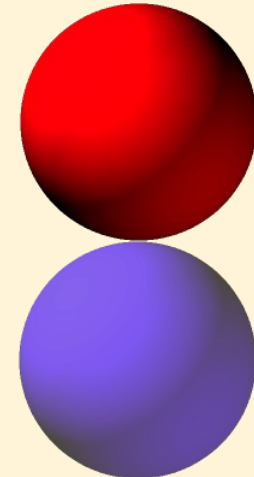
W



X



Y



Z

**Lautsprecher-Quadrat:**



vorn links:  $W+X+Y$

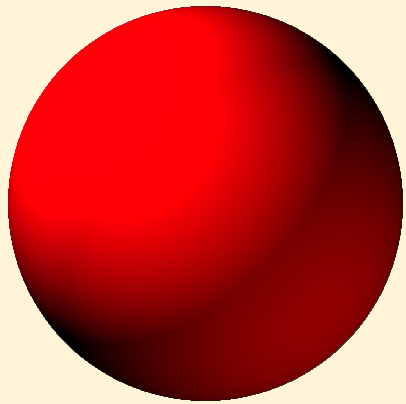
vorn rechts:  $W+X-Y$

hinten links:  $W-X+Y$

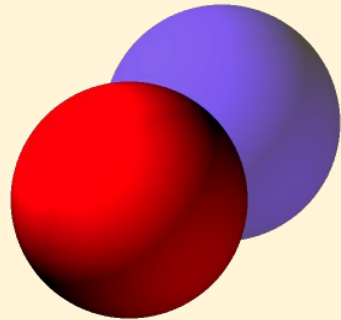
hinten rechts:  $W-X-Y$

# Methoden: **Ambisonics**

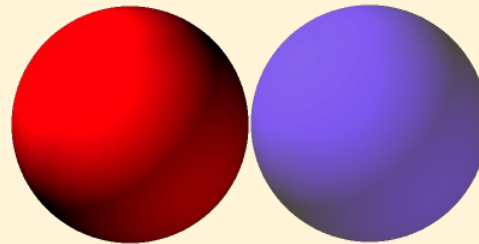
mono + vorn/hinten + links/rechts + oben/unten



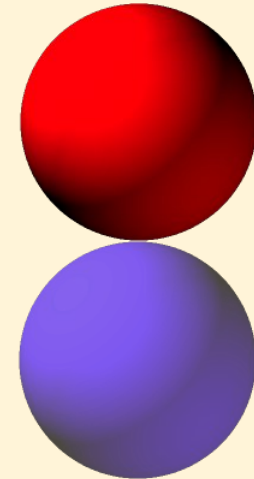
W



X

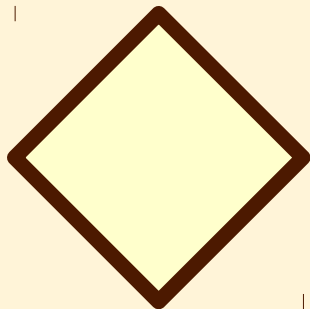


Y



Z

**alternatives Layout:**



vorn:  $W + X$

Seite links:  $W + Y$

Seite rechts:  $W - Y$

hinten:  $W - X$

# Methoden: Ambisonics

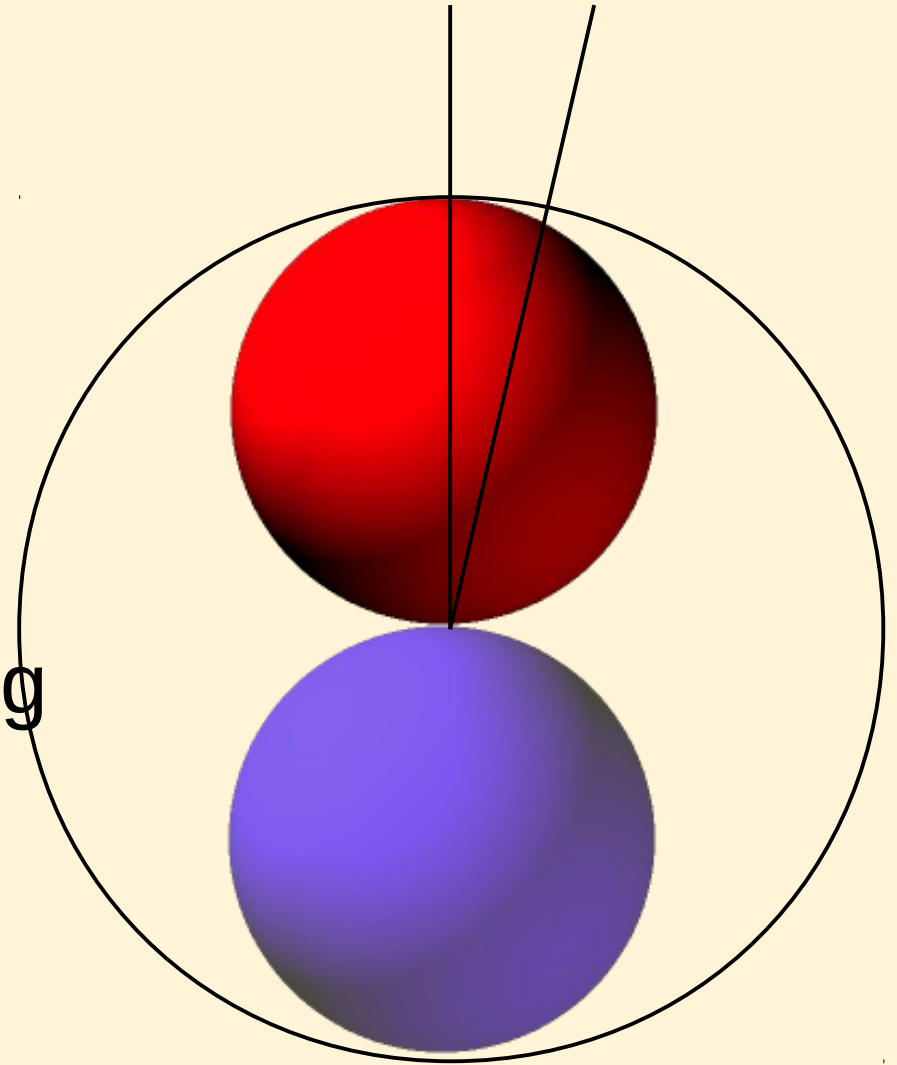
Problem:

geringe Winkelauflösung

gesucht:

höher auflösendes Sampling

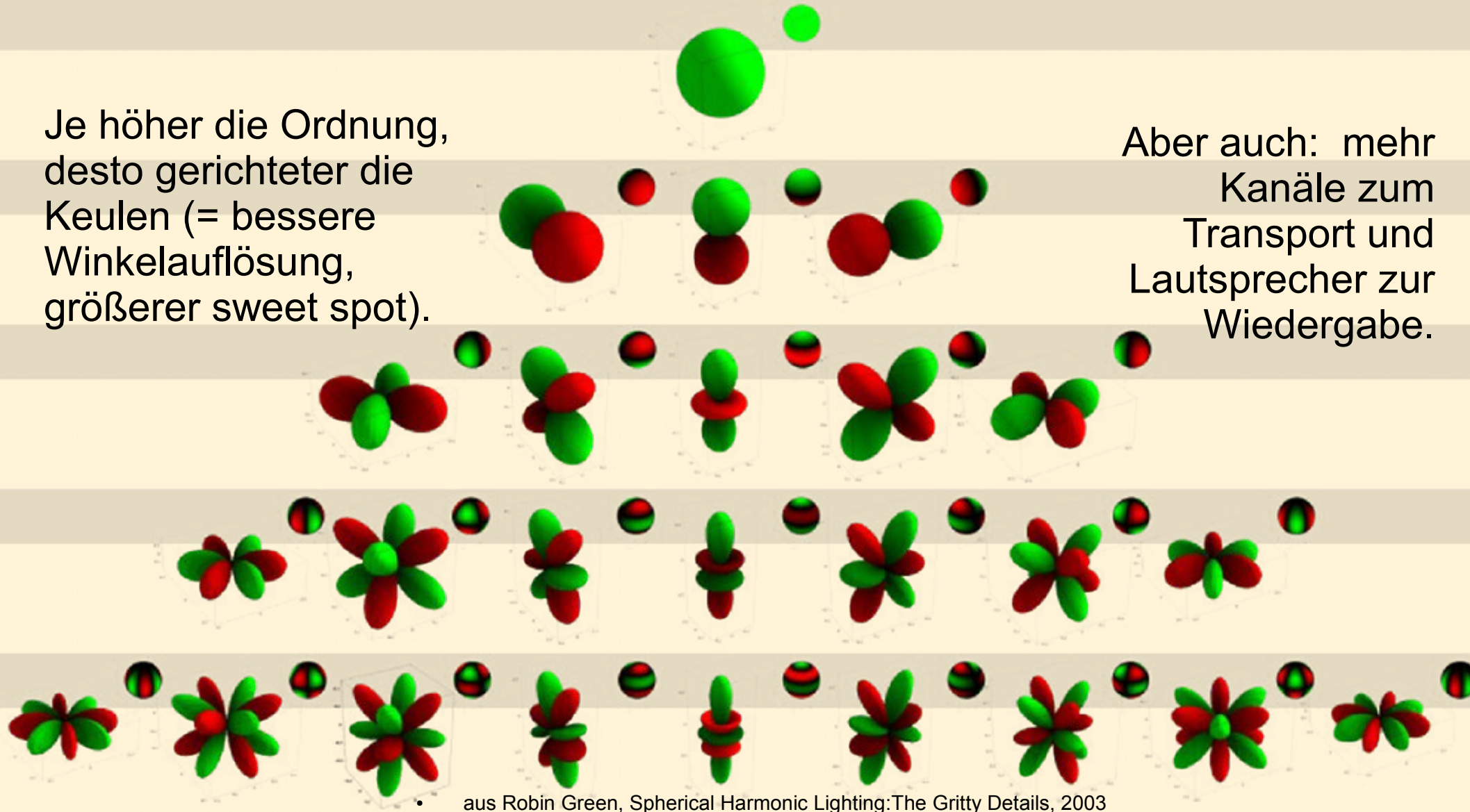
-> Higher-order Ambisonics



# Methoden: Ambisonics

Je höher die Ordnung,  
desto gerichteter die  
Keulen (= bessere  
Winkelauflösung,  
größerer sweet spot).

Aber auch: mehr  
Kanäle zum  
Transport und  
Lautsprecher zur  
Wiedergabe.



• aus Robin Green, Spherical Harmonic Lighting: The Gritty Details, 2003

# Methoden: **Ambisonics**

## schallfeld-basierte Repräsentation:

An einem Ort wird das Schallfeld systematisch räumlich abgetastet (B-Format).

Die Wiedergabe in diesem Punkt ist theoretisch perfekt.

# Methoden: **Ambisonics**

## **Vorteile:**

- skalierbar, optimale Nutzung der vorhandenen LS
- flexibel weil lautsprecherunabhängig
- psychoakustisch optimiert

## **Nachteile:**

- sweet spot
- Signalrepräsentation unintuitiv
- schlecht vermarktet



# Methoden: Vergleich der Signal- Repräsentationen

- a) **diskret** oder **kanal-basiert** (Stereo, 5.1, ..., 22.2, ...)  
lautsprecherabhängig, in-band
  
- b) **objekt-basiert** (WFS)  
lautsprecherunabhängig, out-of-band
  
- c) **schallfeld-basiert** (Ambisonics)  
lautsprecherunabhängig, in-band

# Methoden: Vergleich der Signal-Repräsentationen

## Out-of-band (d.h. Durch Metadaten):

- lautsprecherunabhängig
- sehr zweckmäßig für Komposition und Sound-Design, jederzeit leicht zu überarbeiten
- aber: keine Möglichkeit, natürliche Schallereignisse aufzuzeichnen  
(höchstens: Analyse->Resynthese)
- schwierig in Standard-Audio-Workflows zu integrieren
- Komplexität  $O(N)$  für  $N$  Quellen
- Hohe, konstante Anforderungen an die Wiedergabeseite

# Methoden: Vergleich der Signal-Repräsentationen

## in-band:

- lautsprecherunabhängig
- geeignet für Mikrofonaufzeichnung
- gut in Standard-Audio-Workflows zu integrieren
- aber: unintuitiv, die einzelnen Kanäle haben keine leicht erfassbare Bedeutung, schwer bis unmöglich nachträglich zu manipulieren
- Komplexität  $O(1)$  für beliebig viele Quellen
- skaliert fließend für  $1 - \infty$  Kanäle (auch auf der Wiedergabeseite)

# Aufnahmetechniken

- 1. naiver Ansatz:** Einzelquellen
- 2. systematischer Ansatz:** B-Format, spherical sampling
- 3. hybride Aufnahmetechniken**

# Aufnahmetechniken: **naiver Ansatz**

„Microphone curtain“ (Kirchhoff-Helmholtz!)  
oder: große Anzahl von Spot-Mikrofonen  
+ künstliches Panning  
+ evtl. diffuser Nachhall (Hamasaki etc.)

 Direktsound

 später Hall

frühe Reflexionen 

# Aufnahmetechniken: **systematischer Ansatz**

Soundfield-Mikrofon:

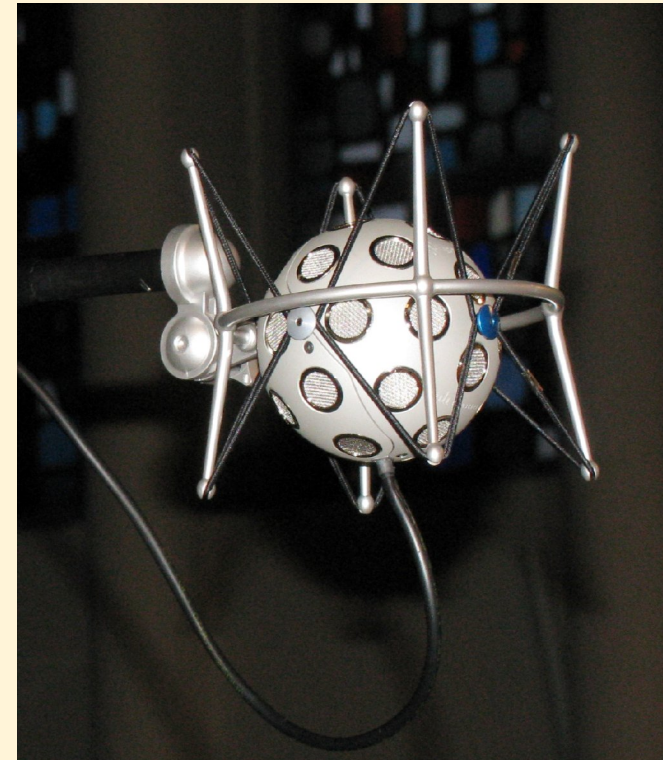


absolut korrekte Wiedergabe  
aller Parameter am Hörort



aber: u.U. zu diffus und instabil

Higher-order-Mikrofone noch  
nicht serienreif.



# Aufnahmetechniken: hybrider Ansatz

Soundfield-Mikrofon  
+ Stützen mit künstlichem Panning in HOA

 absolut korrekte Wiedergabe aller Parameter  
am Hörort

Stabilisierung und Fokussierung des  
Klangbildes durch Stützen.



13 VL L KM-184 14 VL R KM-184 21SF1 22SF2 23SF 24SF4 25TM1 26TM2 27TM3 28TM4 Soundfield TetraMic Master MonA3 MonA2 MonA1 MnUJH

Record Record Rec Rec Rec Rec Rec Rec Rec \*4\* \*4\* \*104\* Mas Mas Mas Mas

Simple delay li Simple delay li SfProc TetraProc UHJnc

Simple delay Simple delay

13 VL L KM-184: AMB order 3,3 panner (by Joeri)

Presets Save Bypass

Controls

Elevation 10.000

Azimuth -24.375

14 VL R KM-184: AMB order 3,3 panner (by Joeri)

Presets Save Bypass

Controls

Elevation 10.000 Manual

Azimuth 24.375 Manual

Connections - JACK Audio Connection Kit

Audio MIDI ALSA

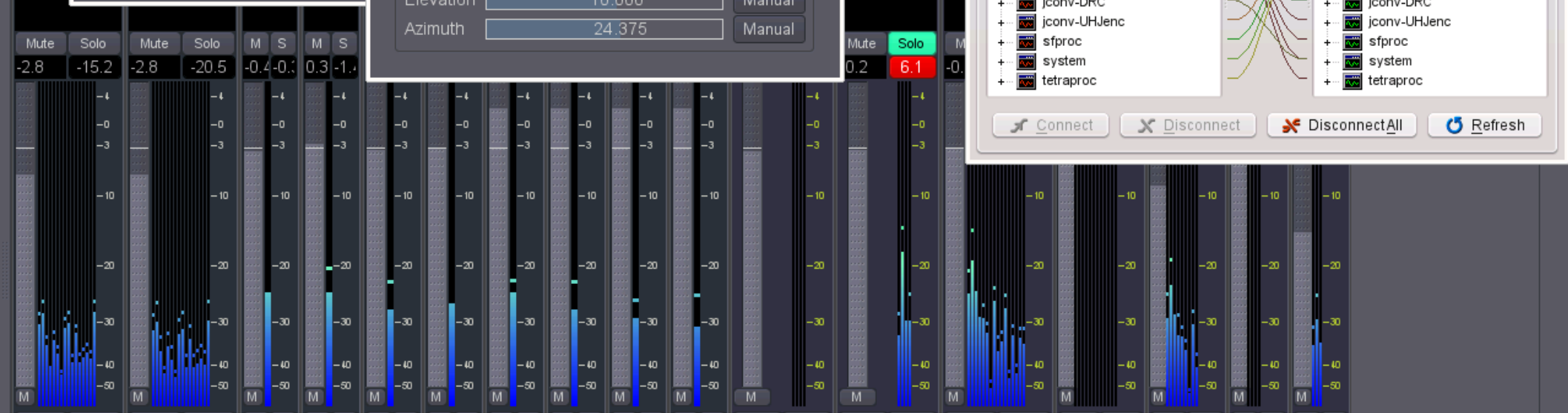
Readable Clients / Output Ports

- ardour
- hex1
- hex2
- jconv-DRC
- jconv-UHJenc
- sfproc
- system
- tetraproc

Writable Clients / Input Ports

- ardour
- hex1
- hex2
- jconv-DRC
- jconv-UHJenc
- sfproc
- system
- tetraproc

Connect Disconnect DisconnectAll Refresh



Master: AMB order 3,3 rotator (by Joern He)

Presets Save Bypass

Controls

Angle 0.000 Manual

AmbDec - 0.4.2 [hex2]

Volume

Config Int 1 3 5 Ext 2 4 6

HPF 10 20 40 80 160

Mute LF RF LB RB

Invert X Y Z EndFire

Config

TETRAPROC - Tetrahedral Microphone Processor - 0.6.2

Core Sound Tetramic ~2050 FuMa

W X Y Z

Volume -50 -40 -30 -20 -10 0 10

Xtalk Mon

Meters Inp Mon Monit Rec Ext

Elev 9

B L Azim R

0 Angle 18

0mni Card Fig





Eine Ambisonics-Produktion unter Linux:  
Ardour (<http://ardour.org>)  
AMB panning plugins,  
AmbDec Lautsprecherdecoder,  
Tetraproc Mikrofonprozessor  
(<http://kokkinizita.net>)

Freie Software, läuft auch unter MacOSX.

# Weiterführende Informationen vom Autor zum Thema Ambisonics-Produktion unter Linux:

[http://cec.concordia.ca/econtact/11\\_3/nettingsmeier\\_ambisonics.html](http://cec.concordia.ca/econtact/11_3/nettingsmeier_ambisonics.html)  
[http://stackingdwarves.net/public\\_stuff/linux\\_audio/ambi\\_at\\_home/](http://stackingdwarves.net/public_stuff/linux_audio/ambi_at_home/)  
[http://stackingdwarves.net/public\\_stuff/linux\\_audio/ambisonic\\_organ\\_remix/](http://stackingdwarves.net/public_stuff/linux_audio/ambisonic_organ_remix/)  
[http://stackingdwarves.net/public\\_stuff/linux\\_audio/lac2010/](http://stackingdwarves.net/public_stuff/linux_audio/lac2010/)

## Ambisonics-Aufnahmen:

<http://ambisonia.com>





**Ich danke für Eure  
Aufmerksamkeit.**

**Noch Fragen?**