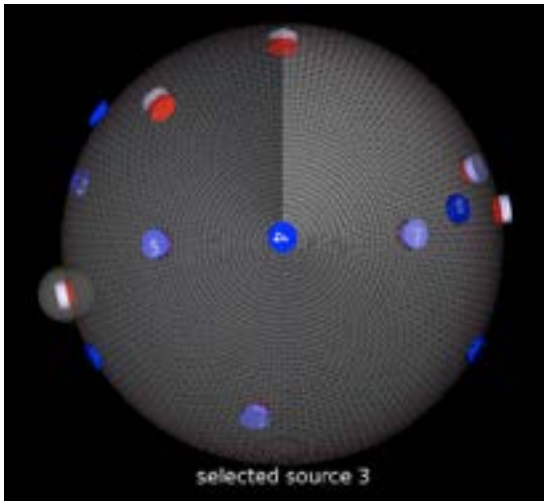
Räumliche Transformationen zur Veränderung von ambisonischen Aufnahmen

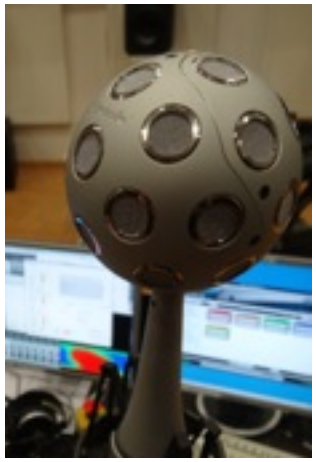
Matthias Kronlachner

Motivation

Warum in der Ambisonics-Domäne?

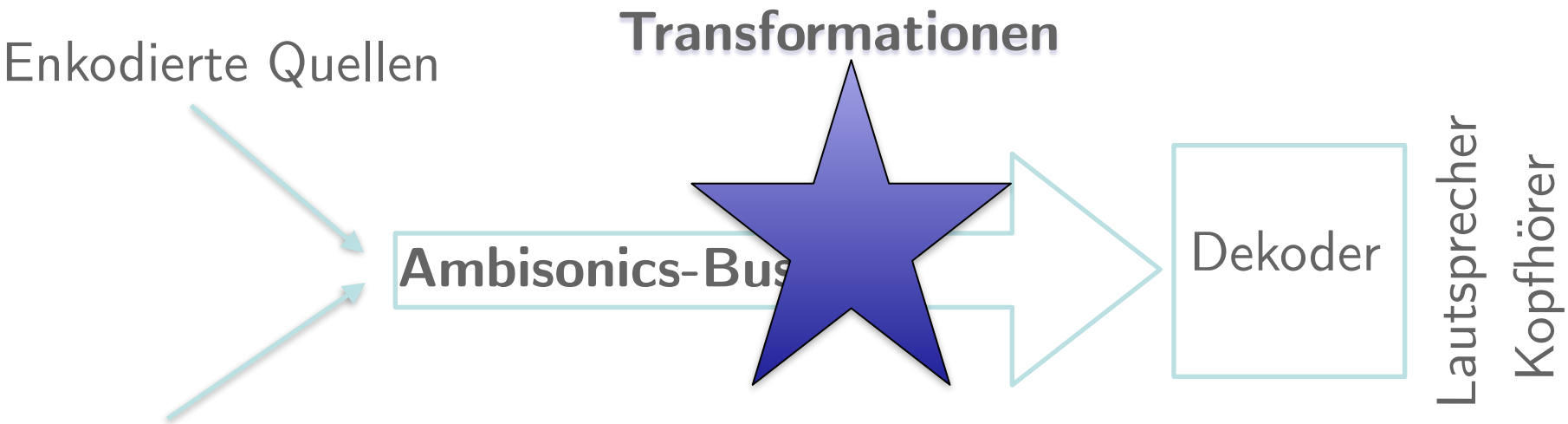


- Objektbasierte Formate
  einfach: Metadaten verändern

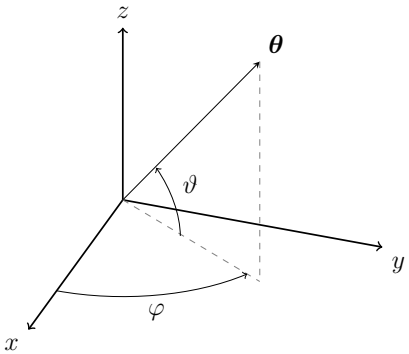


- Aber: nicht immer alle Objekte einzeln verfügbar / Trennung mit Artefakten?

- **Allgemeine Transformation**
  - **Rotation**
  - **Warping**
  - **Richtungsabhängige Lautstärke**



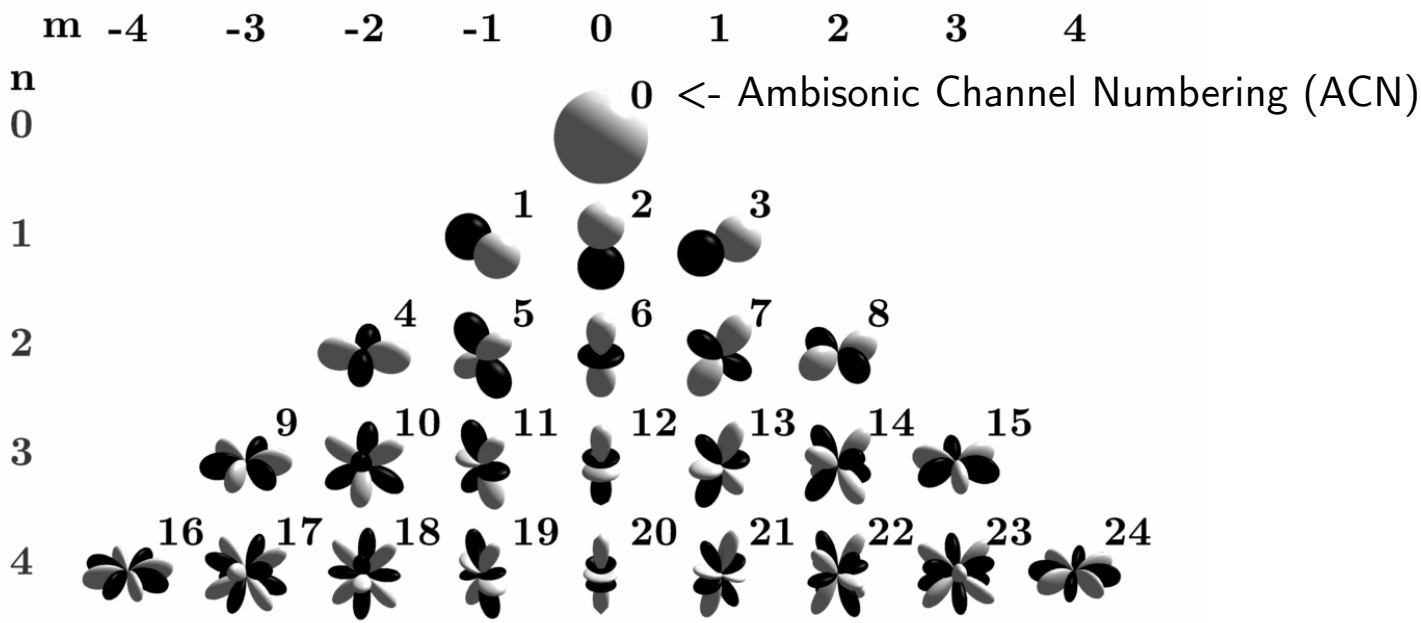
- **Punktanordnungen: *t*-designs**
- **Plug-ins**

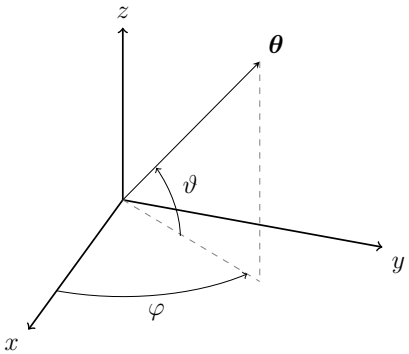


Kugelflächenfunktionen Y_n^m

$$f(\varphi, \vartheta, t) = \sum_{n=0}^{\mathrm{N}} \sum_{m=-n}^{n} Y_n^m(\varphi, \vartheta)\, \phi_{nm}(t)$$

$$Y_n^m(\varphi, \vartheta) = N_n^{|m|} P_n^{|m|}(\sin(\vartheta)) \begin{cases} \sin|m|\varphi, & \text{for } m < 0 \\ \cos|m|\varphi, & \text{for } m \geq 0. \end{cases}$$



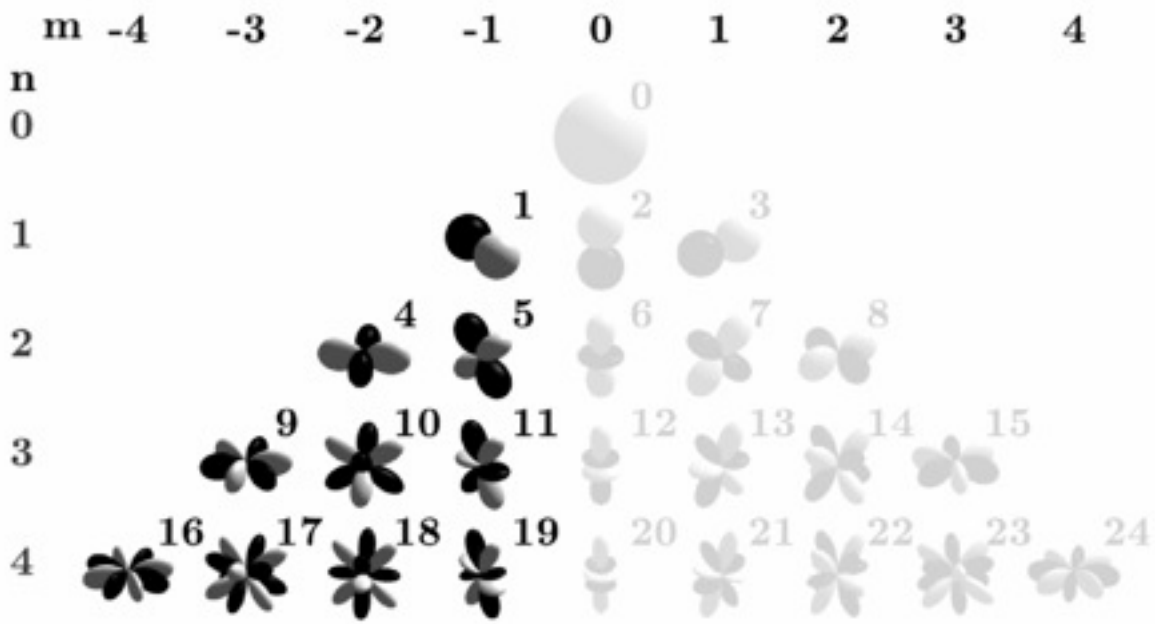


Transformation durch Symmetrie

(Erinnerung an MS Aufnahmetechnik..)

Vorzeichennegation für Signale mit m<0

resultiert in Spiegelung um die y-Achse



Vektornotation

$$f(\varphi, \vartheta, t) = \sum_{n=0}^{\mathrm{N}} \sum_{m=-n}^{n} Y_n^m(\varphi, \vartheta)\ \phi_{nm}(t)$$

$$f(\boldsymbol{\theta}, t) = \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\ \boldsymbol{\phi}_{\mathrm{N}}(t)$$

$$\boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}) = \begin{pmatrix} Y_0(\boldsymbol{\theta}) \\ Y_1(\boldsymbol{\theta}) \\ Y_2(\boldsymbol{\theta}) \\ Y_3(\boldsymbol{\theta}) \\ Y_4(\boldsymbol{\theta}) \\ \vdots \\ Y_{(\mathrm{N}+1)^2-1}(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} Y_0^0(\boldsymbol{\theta}) \\ Y_1^{-1}(\boldsymbol{\theta}) \\ Y_1^0(\boldsymbol{\theta}) \\ Y_1^1(\boldsymbol{\theta}) \\ Y_2^{-2}(\boldsymbol{\theta}) \\ \vdots \\ Y_{\mathrm{N}}^{\mathrm{M}}(\boldsymbol{\theta}) \end{pmatrix} \qquad \boldsymbol{\phi}_{\mathrm{N}}(t) = \begin{pmatrix} \phi_0(t) \\ \phi_1(t) \\ \phi_2(t) \\ \vdots \\ \phi_{(\mathrm{N}+1)^2-1}(t) \end{pmatrix}$$

Ambisonic Channel Numbering (ACN)

Spherical Harmonic Transform

$$\mathcal{SHT}\{f(\boldsymbol{\theta})\} = \boldsymbol{\phi}_{\mathrm{N}} = \int_{\mathbb{S}^2} \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}.$$

Geeignete Wahl von Abtastpunkten

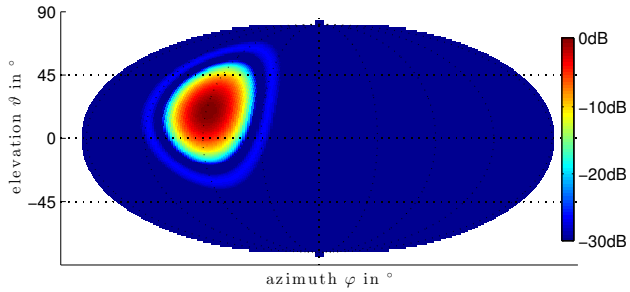
$$\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{\mathrm{L}}]^T.$$

$$\mathcal{DSHT}\{\boldsymbol{f}(\boldsymbol{\Theta})\} = \boldsymbol{\phi}_{\mathrm{N}} = \boldsymbol{Y}_{\mathrm{N}}^{\dagger}(\boldsymbol{\Theta})\, \boldsymbol{f}(\boldsymbol{\Theta}),$$

$$\boldsymbol{Y}_{\mathrm{N}}(\boldsymbol{\Theta}) = \begin{pmatrix} \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta}_1) \\ \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta}_2) \\ \vdots \\ \\ \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta}_{\mathrm{L}}) \end{pmatrix} = \begin{pmatrix} Y_0(\boldsymbol{\theta}_1) & Y_1(\boldsymbol{\theta}_1) & Y_2(\boldsymbol{\theta}_1) & \cdots & Y_{(\mathrm{N}+1)^2-1}(\boldsymbol{\theta}_1) \\ Y_0(\boldsymbol{\theta}_2) & Y_1(\boldsymbol{\theta}_2) & Y_2(\boldsymbol{\theta}_2) & \cdots & Y_{(\mathrm{N}+1)^2-1}(\boldsymbol{\theta}_2) \\ & & \cdots & & \\ \\ Y_0(\boldsymbol{\theta}_{\mathrm{L}}) & Y_1(\boldsymbol{\theta}_{\mathrm{L}}) & Y_2(\boldsymbol{\theta}_{\mathrm{L}}) & \cdots & Y_{(\mathrm{N}+1)^2-1}(\boldsymbol{\theta}_{\mathrm{L}}). \end{pmatrix}$$

Allgemeine Transformationsvorschrift

$f(\boldsymbol{\theta}, t)$



1) Richtungen unterschiedlich gewichten $g(\boldsymbol{\theta})$

2) Richtungen neu zuweisen $\tilde{\boldsymbol{\theta}} = \mathcal{T}\{\boldsymbol{\theta}\}$

$$\tilde{f}(\mathcal{T}\{\boldsymbol{\theta}\}, t) = g(\boldsymbol{\theta})\, f(\boldsymbol{\theta}, t)$$

inverse Abbildung $\tilde{f}(\boldsymbol{\theta}, t) = g(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, f(\mathcal{T}^{-1}\{\boldsymbol{\theta}\}, t)$

Allgemeine Transformationsvorschrift

Transformiertes Signal $\tilde{f}(\boldsymbol{\theta}, t) = g(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, f(\mathcal{T}^{-1}\{\boldsymbol{\theta}\}, t)$

$$f(\boldsymbol{\theta}, t) = \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\, \boldsymbol{\phi}_{\mathrm{N}}(t)$$

$$\boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\, \tilde{\boldsymbol{\phi}}_{\mathrm{N}}(t) = g(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, \boldsymbol{\phi}_{\mathrm{N}}(t)$$

Orthogonalität

$$\int_{\mathbb{S}^2} \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta})\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} = \boldsymbol{I}$$

$$\tilde{\boldsymbol{\phi}}_{\mathrm{N}}(t) = \underbrace{\int_{\mathbb{S}^2} \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta})\, g(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, \mathrm{d}\boldsymbol{\theta}}_{:=\boldsymbol{T}}\, \boldsymbol{\phi}_{\mathrm{N}}(t)$$

$$\tilde{\boldsymbol{\phi}}_{\mathrm{N}}(t) = \boldsymbol{T}\, \boldsymbol{\phi}_{\mathrm{N}}(t)$$

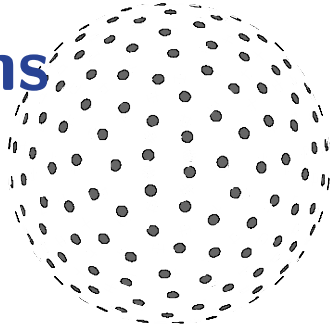
-> Erkennen SHT $\quad \boldsymbol{T} = \mathcal{SHT}\{g(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\mathcal{T}^{-1}\{\boldsymbol{\theta}\})\}$

und verwenden DSHT

$$\boldsymbol{T} = \mathcal{DSHT}\left\{\mathrm{diag}\{\boldsymbol{g}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\}\, \boldsymbol{Y}_{\mathrm{N}}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\right\}$$

$$= \boldsymbol{Y}_{\tilde{\mathrm{N}}}^{\dagger}(\boldsymbol{\Theta})\, \mathrm{diag}\{\boldsymbol{g}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\}\, \boldsymbol{Y}_{\mathrm{N}}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})$$

Transformationsvorschrift mit *t*-designs



$$\boldsymbol{T} = \mathcal{DSHT}\left\{\mathrm{diag}\{\boldsymbol{g}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\}\,\boldsymbol{Y}_{\mathrm{N}}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\right\}$$

$$= \boldsymbol{Y}_{\tilde{\mathrm{N}}}^{\dagger}(\boldsymbol{\Theta})\,\mathrm{diag}\{\boldsymbol{g}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})\}\,\boldsymbol{Y}_{\mathrm{N}}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}\})$$

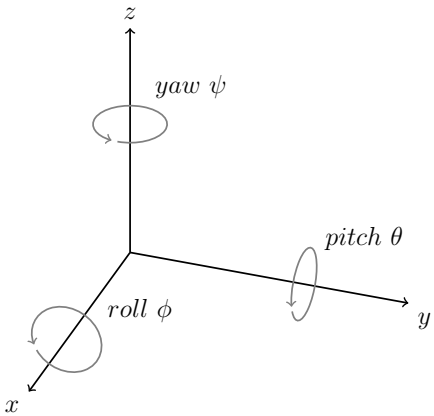
mit *t*-design keine Pseudoinversion notwendig $t \geq 2\mathrm{N}$

$$\boldsymbol{T} = \mathrm{diag}\{\frac{4\pi}{\mathrm{L}}\}\,\boldsymbol{Y}_{\tilde{\mathrm{N}}}^{\mathrm{T}}(\boldsymbol{\Theta}_t)\,\mathrm{diag}\{\boldsymbol{g}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}_t\})\}\,\boldsymbol{Y}_{\mathrm{N}}(\mathcal{T}^{-1}\{\boldsymbol{\Theta}_t\})$$

ohne inverse Abbildung müssten wir für jede Parameteränderung die Pseudoinverse berechnen...

Transformationen erhöhen möglicherweise N -> Ñ

Anwendungen der allgemeinen Transformationsvorschrift

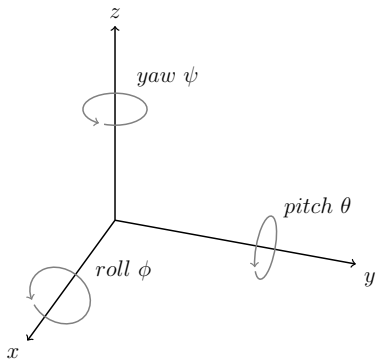


$$\tilde{\boldsymbol{\theta}} = \mathcal{T}\{\boldsymbol{\theta}\} = \mathbf{R}(\phi, \theta, \psi)\ \boldsymbol{\theta}$$

$$\mathbf{R}(\phi,\theta,\psi) \quad = \quad \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}}_{\mathbf{x-axis-rotation(roll)}} \cdot \underbrace{\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}}_{\mathbf{y-axis-rotation(pitch)}} \cdot \underbrace{\begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{z-axis-rotation(yaw)}}$$

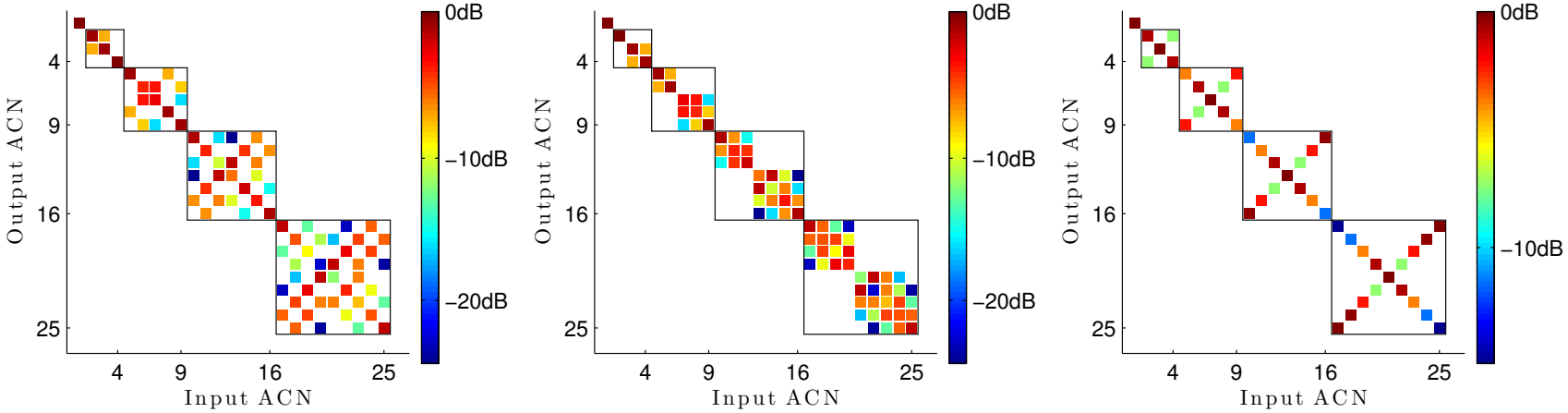
$$\boldsymbol{T}_r^{xyz} = \mathrm{diag}\{\frac{4\pi}{\mathrm{L}}\}\ \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\Theta}_t)\ \boldsymbol{Y}_{\mathrm{N}}(\boldsymbol{R}^{\mathrm{T}}(\phi,\theta,\psi)\ \boldsymbol{\Theta}_t)$$

Rotation



$$\tilde{\boldsymbol{\theta}} = \mathcal{T}\{\boldsymbol{\theta}\} = \mathbf{R}(\phi, \theta, \psi)\ \boldsymbol{\theta}$$

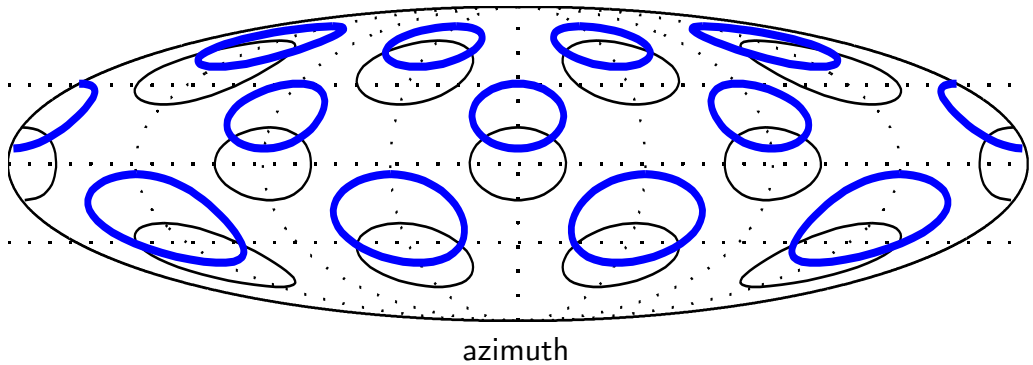
$$g(\boldsymbol{\theta}) = 1$$

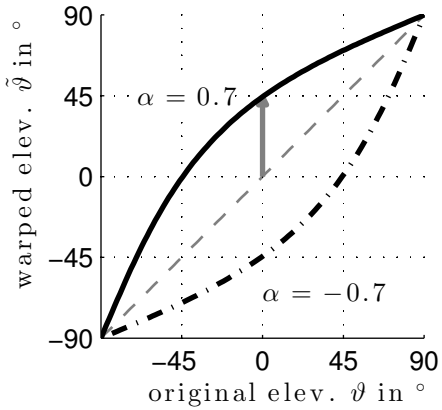


(a) x-axis $\phi = 25°$. (b) y-axis $\theta = 25°$. (c) z-axis $\psi = 25°$.

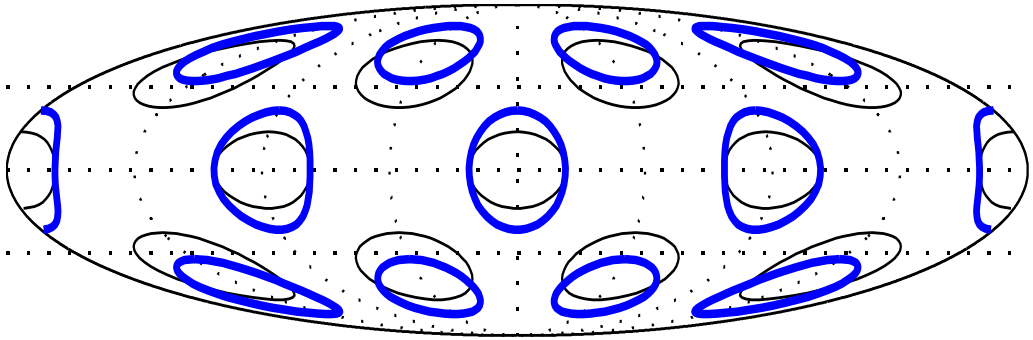
Besetzung der Transformationsmatrix,
Komponenten jeder SH Ordnung n werden gemischt,
verlassen diese aber nicht!

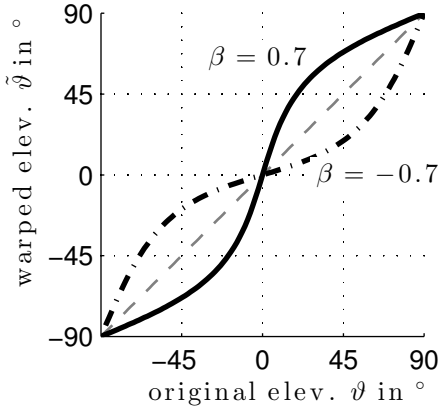
Warping





Zum Pol



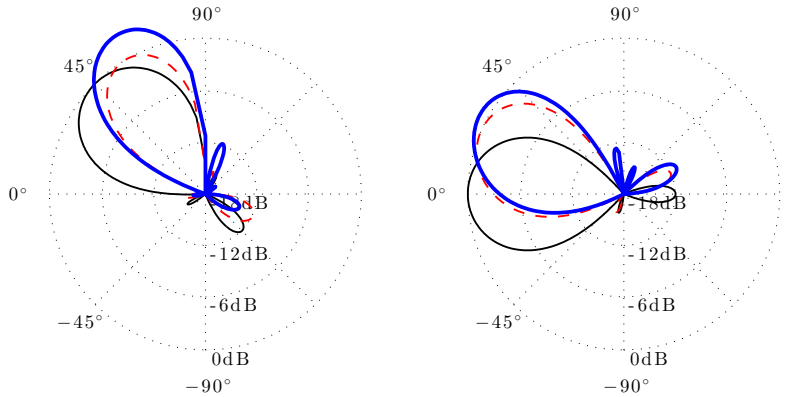


Zum Äquator

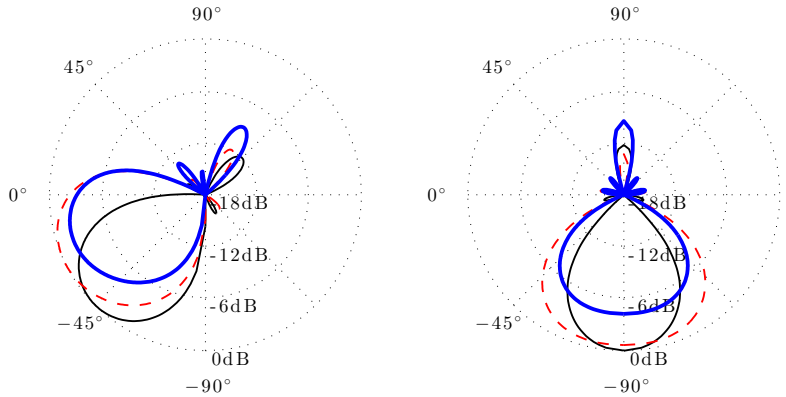
- Lautstärkenkompensation notwendig

Warping

Richtung Pol $\alpha = 0.4 \quad \mathrm{N} = 3$



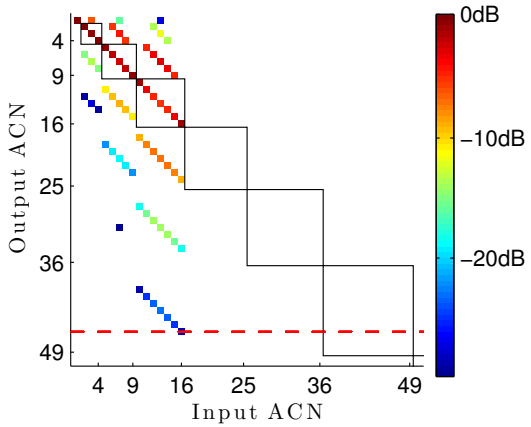
(a) $\vartheta = 45°$. (b) $\vartheta = 0°$.



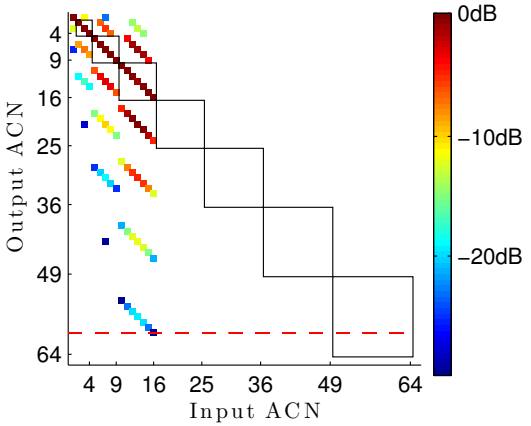
(c) $\vartheta = -45°$. (d) $\vartheta = -90°$.

schwarz: Originalsignal, rot: ohne Lautstärkenkompensation

blau: mit Lautstärkenkompensation



(a) Without loudness compensation.



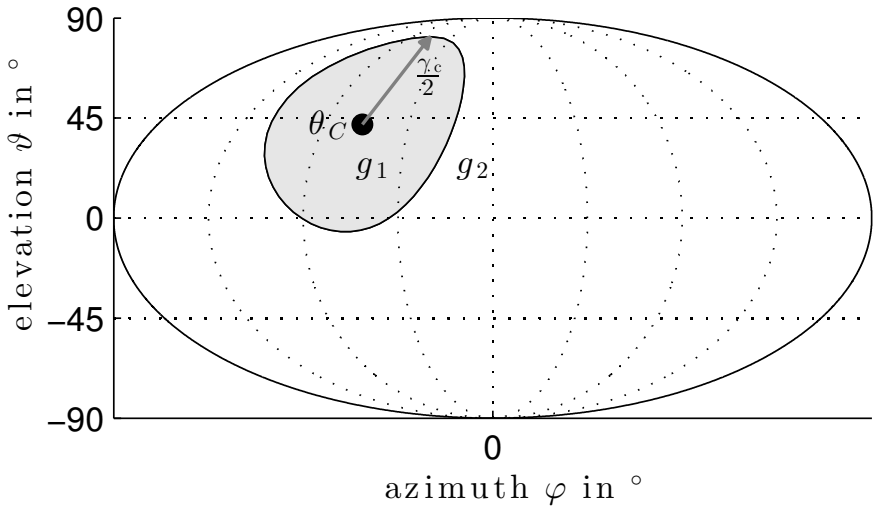
(b) With loudness compensation.

Richtungsabhängige Lautstärkenanpassung

- Kugelkappenfunktion mit Zentrum $\boldsymbol{\theta}_\mathrm{c}$, Größe $\frac{\gamma_\mathrm{c}}{2}$

- Lautstärkefaktor g_1 für Punkte innerhalb der Kappe,

  g_2 außerhalb, neutrale Winkelabbildung $\mathcal{T}\{\boldsymbol{\theta}\} = \boldsymbol{\theta}$



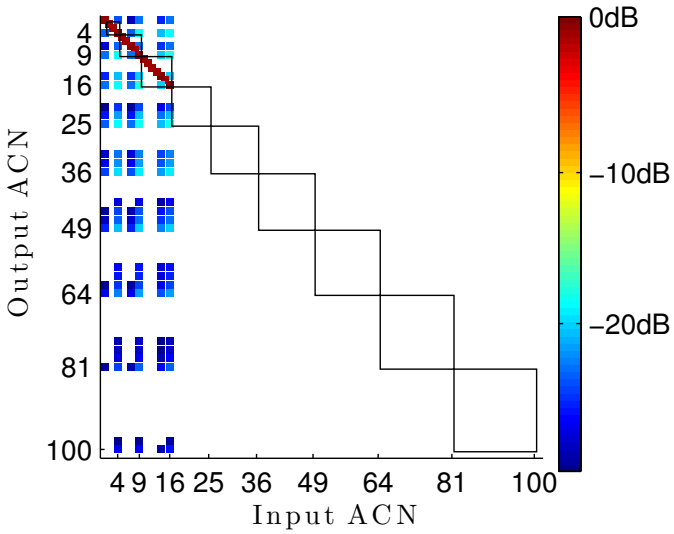
Richtungsabhängige Lautstärkenanpassung

$$\mathrm{N} = 3$$

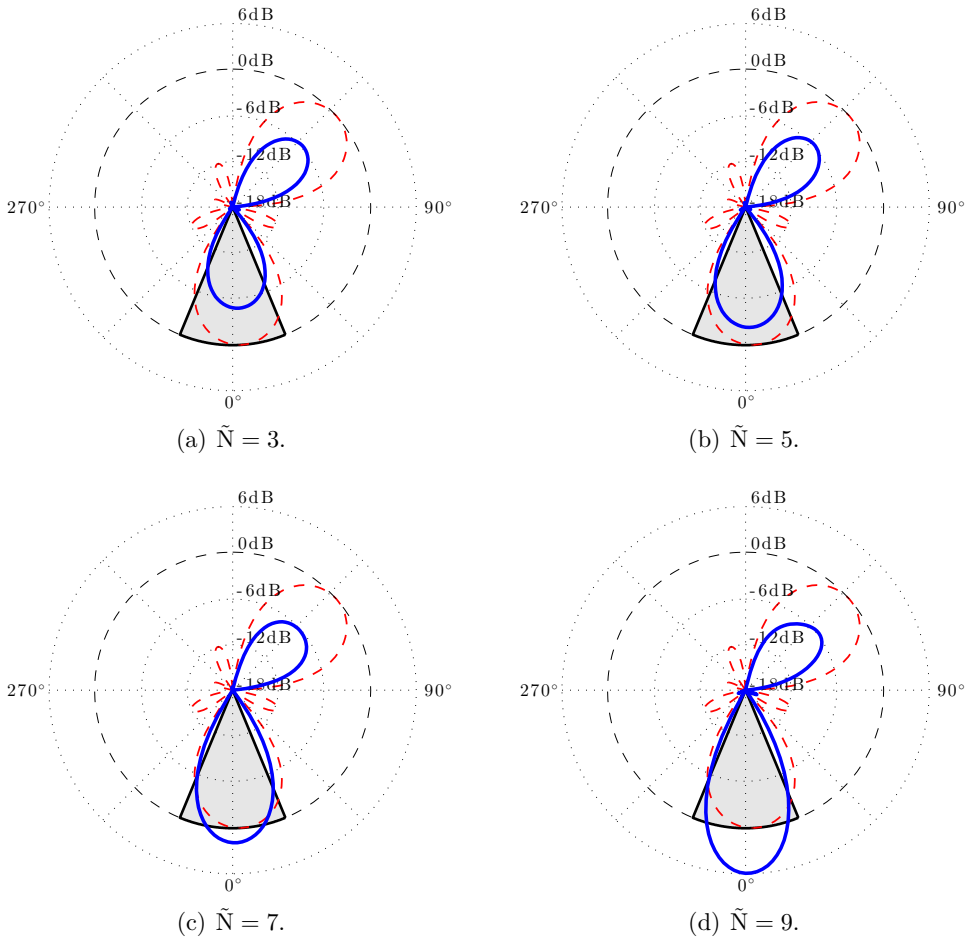
$$\boldsymbol{\theta}_C = (0°, 0°)$$

$$\gamma = 40°$$

$$g_1 = 6\mathrm{dB},\ g_2 = -6\mathrm{dB}$$



Transformationsmatrix



(a) $\tilde{\mathrm{N}} = 3$. (b) $\tilde{\mathrm{N}} = 5$. (c) $\tilde{\mathrm{N}} = 7$. (d) $\tilde{\mathrm{N}} = 9$.

Nachteil: benötigt höhere Ordnungen!

Besser: Spherical Slepian Functions

Teilmenge der Kugelfläche $S^2 \subset \mathbb{S}^2$

Orthogonalität geht verloren $\int_{S^2} \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta})\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta} = \boldsymbol{G}$

g(θ) - Kappenfunktion $\boldsymbol{G} = \int_{\mathbb{S}^2} \boldsymbol{y}_{\tilde{\mathrm{N}}}(\boldsymbol{\theta})\, \mathrm{diag}\{\boldsymbol{g}(\boldsymbol{\theta})\}\, \boldsymbol{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}$

Singularwertzerlegung $\boldsymbol{G} = \boldsymbol{U}\, \mathrm{diag}\{[\sigma_i]_{1\ldots(\mathrm{N}+1)^2}\}\, \boldsymbol{V}^{\mathrm{T}}$

Ersetzen der Singularwerte $\boldsymbol{T} = \boldsymbol{U}\, \mathrm{diag}\{[\varsigma_i]_{1\ldots(\mathrm{N}+1)^2}\}\, \boldsymbol{V}^{\mathrm{T}}$

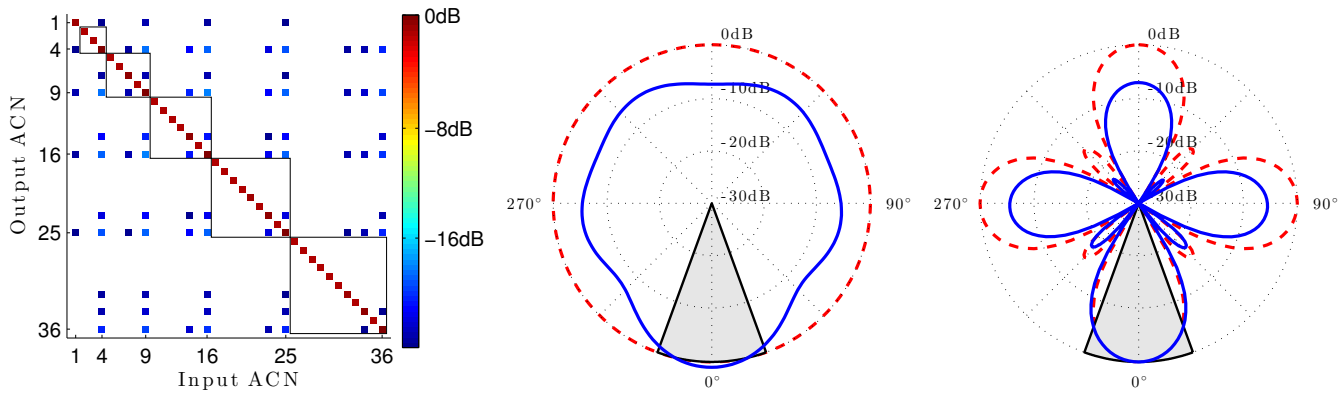
$$\varsigma_i = g_1\, u(\sigma_i - \alpha\, \sigma_1) + g_2\, u(\alpha\, \sigma_1 - \sigma_i) \qquad 0 < \alpha < 1$$

Vorteil: Ordnungen strikt begrenzt!

Besser: Spherical Slepian Functions

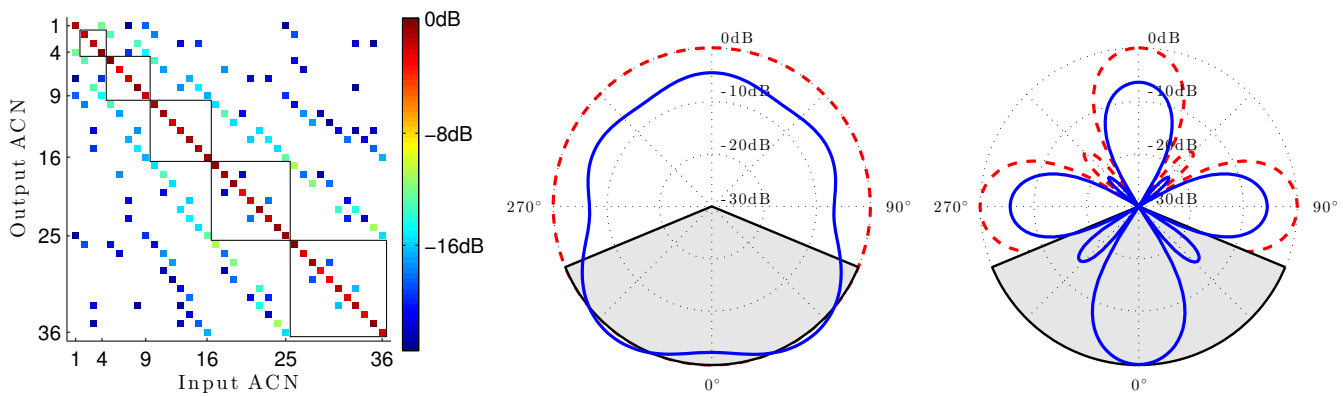
$\mathrm{N} = \tilde{\mathrm{N}} = 5 \qquad \alpha = 0.5 \qquad \boldsymbol{\theta}_C = (0°, 0°) \qquad g_1 = 0\mathrm{dB}, g_2 = -6\mathrm{dB}$

$\gamma = 40°$



(a) (b) (c)

$\gamma = 135°$



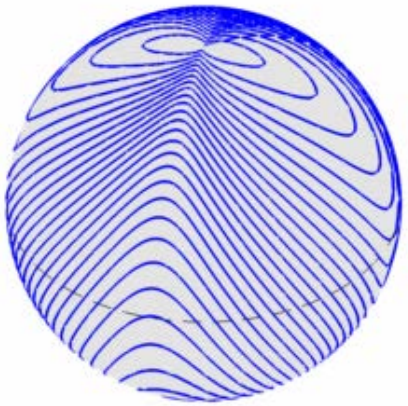
(d) (e) (f)

Andere Transformationskurven

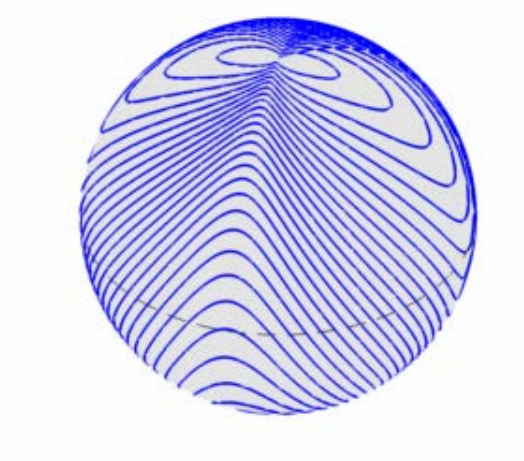
Philips Pavilion, Le Corbusier und Iannis Xenakis,
Weltausstellung 1958 in Brüssel
[wikimedia commons/Wouter Hagens]

Azimuth abhängiges Warping

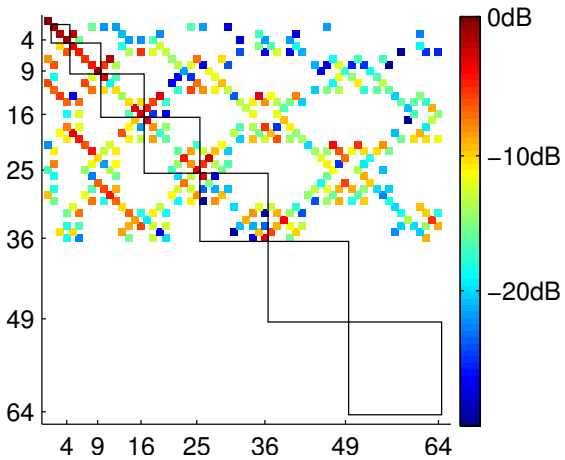
$$\mu = \cos\vartheta,$$
$$\alpha = 0.8\sin 2\phi,$$
$$\tilde{\mu} = \frac{\alpha + \mu}{1 + \mu\alpha},$$
$$\tilde{\vartheta} = \arccos\tilde{\mu}.$$



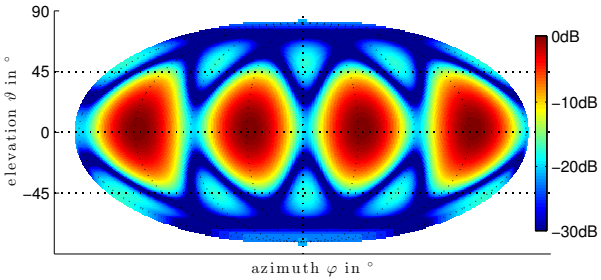
Andere Transformationskurven



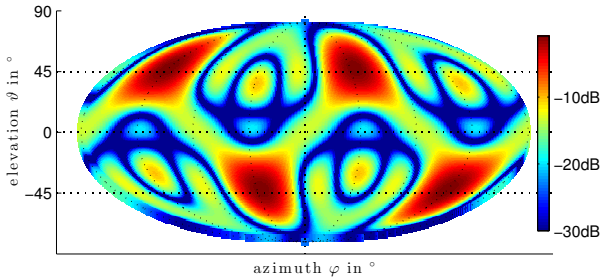
(a) Distortion scheme, lines indicate original elevation levels.



(b) Sparsity of $\boldsymbol{T}$, truncated at $\tilde{\mathrm{N}} = 7$. .



(c) Original surround signal.



(d) Modified surround signal.

t-designs durch nichtlineare Optimierung finden

Mögliche Kostenfunktion für *t*-designs

1. Kondition von **Y**

$$\epsilon_C(\boldsymbol{\Theta}) = \kappa(\,\boldsymbol{Y}_{\mathrm{N}}(\boldsymbol{\Theta})\,) - 1$$

2. Frobenius-Norm

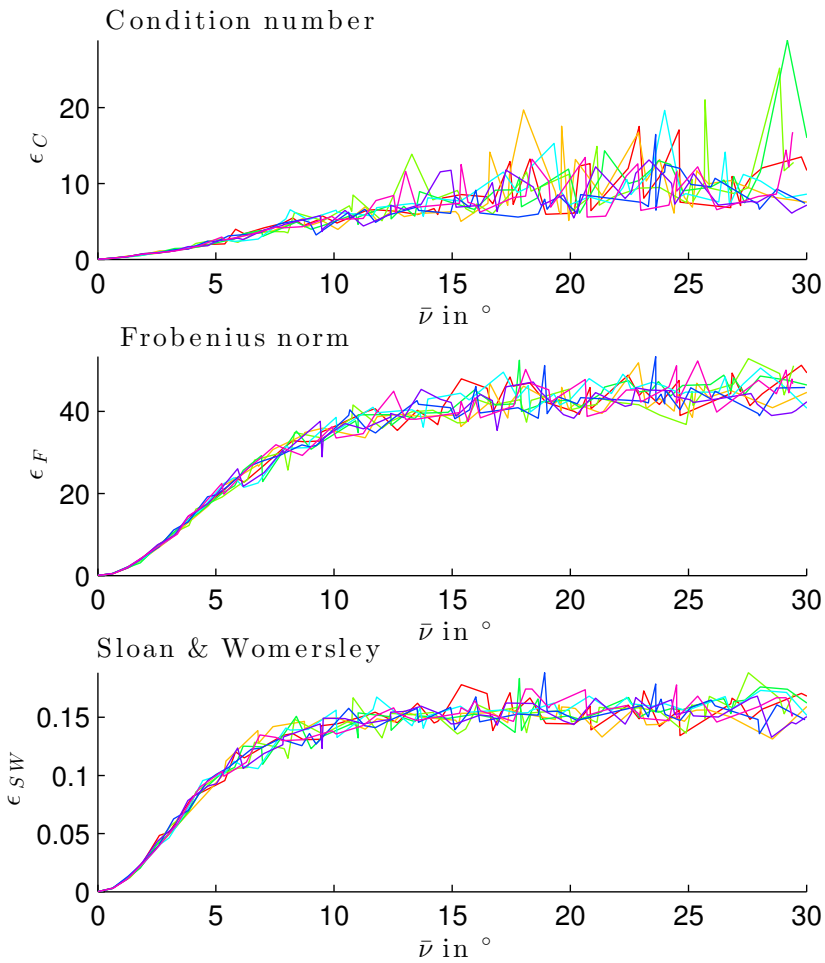
$$\epsilon_F(\boldsymbol{\Theta}) = ||\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\Theta})\,\boldsymbol{Y}_{\mathrm{N}}(\boldsymbol{\Theta})||_F - ||\mathrm{diag}\{\frac{4\pi}{\mathrm{L}}\}||_F$$

$$||\boldsymbol{A}||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

3. Sloan & Womersley

$$\epsilon_{SW}(\boldsymbol{\Theta}) = \frac{1}{\mathrm{L}^2}\sum_{n=1}^{t}\sum_{m=-n}^{n}\left|\sum_{l=1}^{\mathrm{L}} Y_n^m(\boldsymbol{\theta}_l)\right|^2$$

Mögliche Kostenfunktion für *t*-designs



- Starte mit bekanntem *t*-design $\boldsymbol{\Theta}_t$ [1]
- Hinzufügen von Rauschen v

$$\bar{\nu} = \frac{1}{\mathrm{L}} \sum_{l=1}^{\mathrm{L}} \arccos \boldsymbol{\theta}_t^{\mathrm{T}} \boldsymbol{\theta}_\nu$$

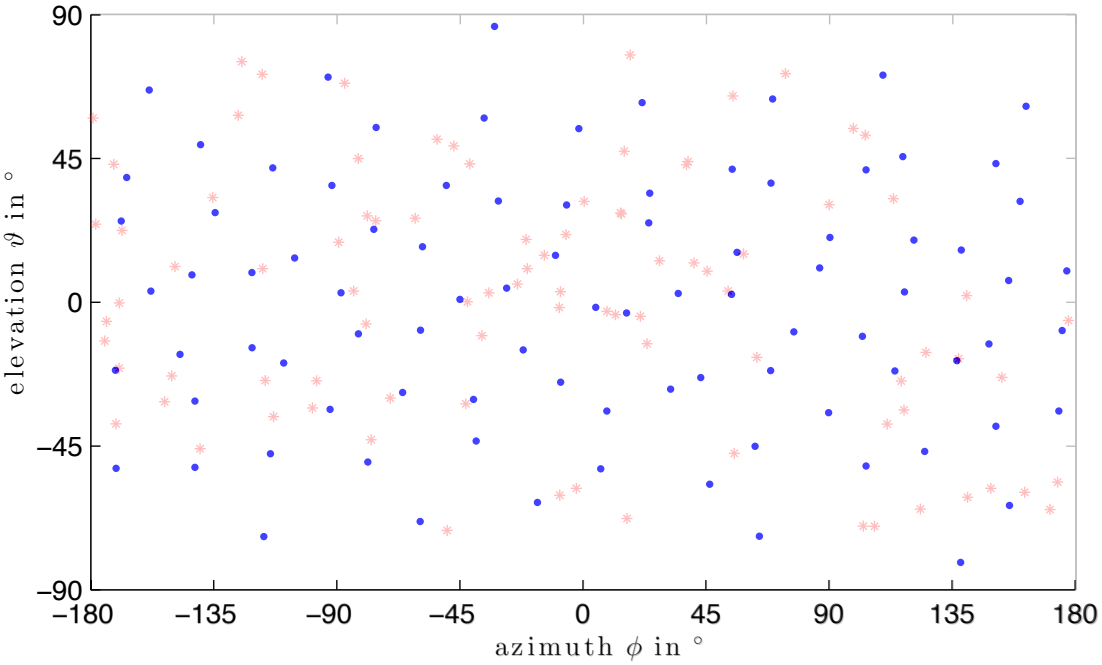
[1] R. Hardin and N. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions," in Discrete Computational Geometry, vol. 15, pp. 429-441, 1996.

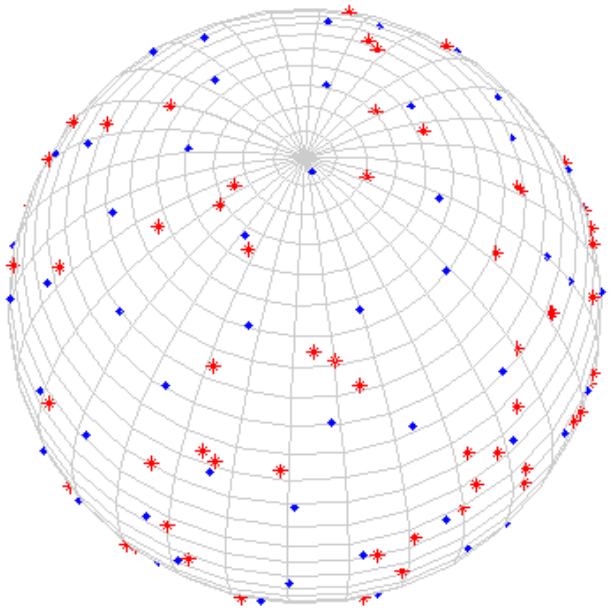
So finde ich *t*-designs...

- Zufällige Startkonfiguration t=10, L=90
- Nach 24m36s...

$$|\epsilon_{CW}| < 10^{-11}$$

$$\kappa(\boldsymbol{Y}(\boldsymbol{\Theta})) = 1.000008$$

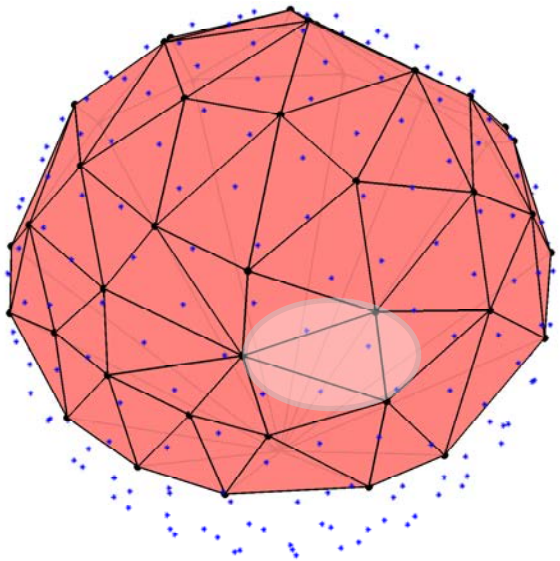




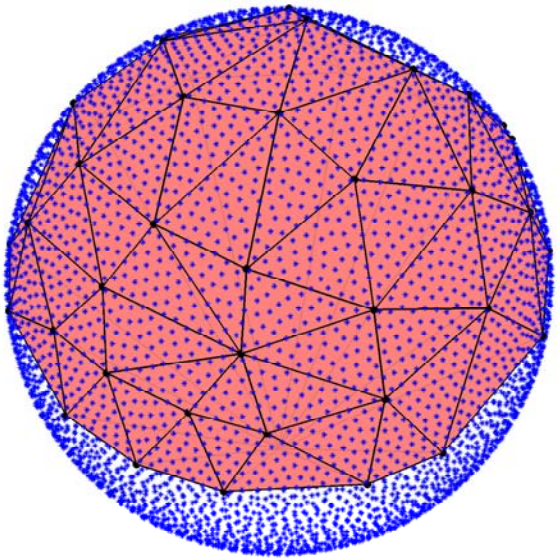
t-designs mit vielen Knoten?

All-Round Ambisonic Decoding

für den ZKM-Kubus mit 43 Lautsprecher



21-design mit L = 240



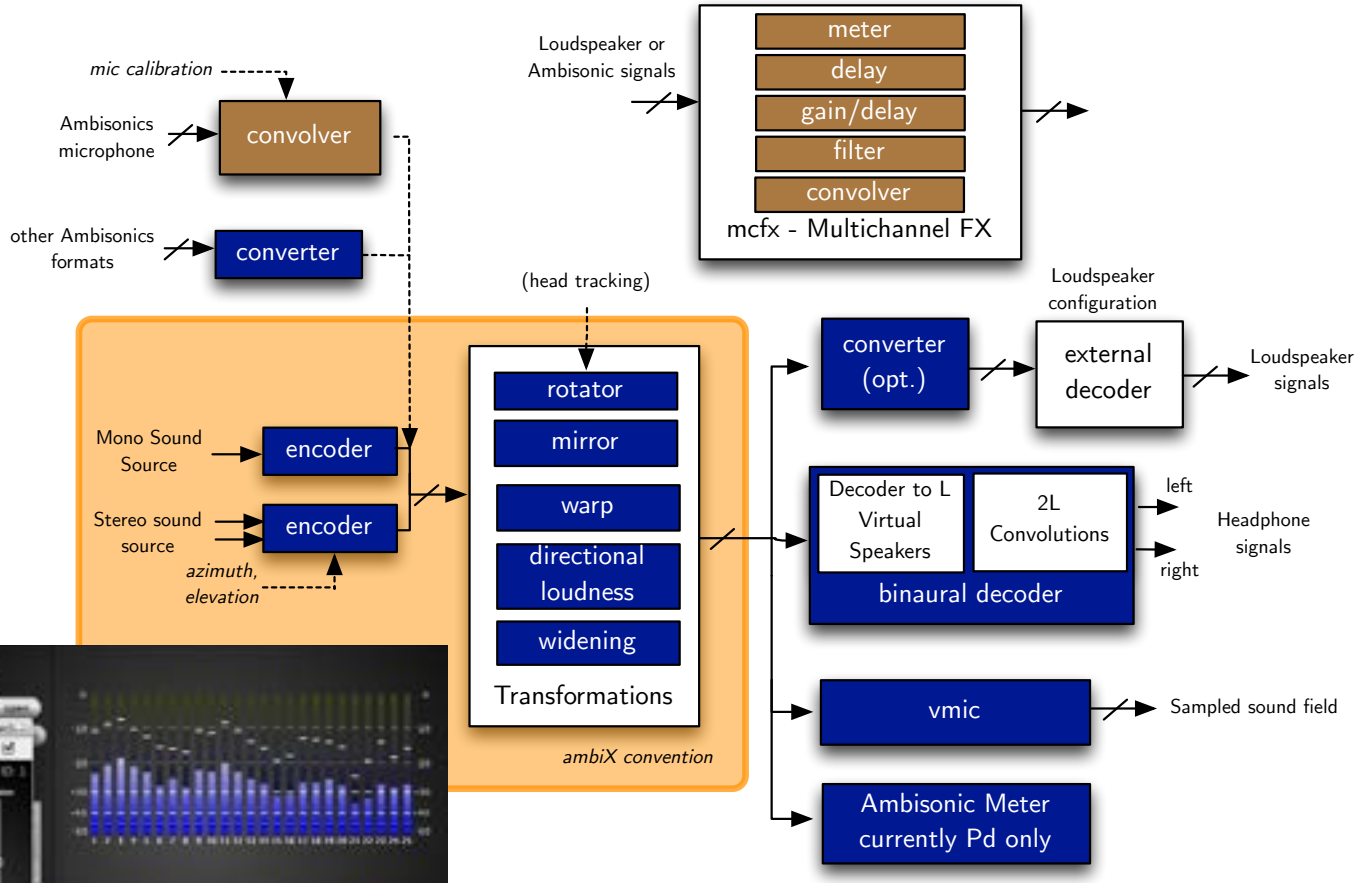
100-design mit L = 5200 [1]

[1] M. Gräf and D. Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms," in Numerische Mathematik Vol. 119 No. 4, p. 699-724, 2011.

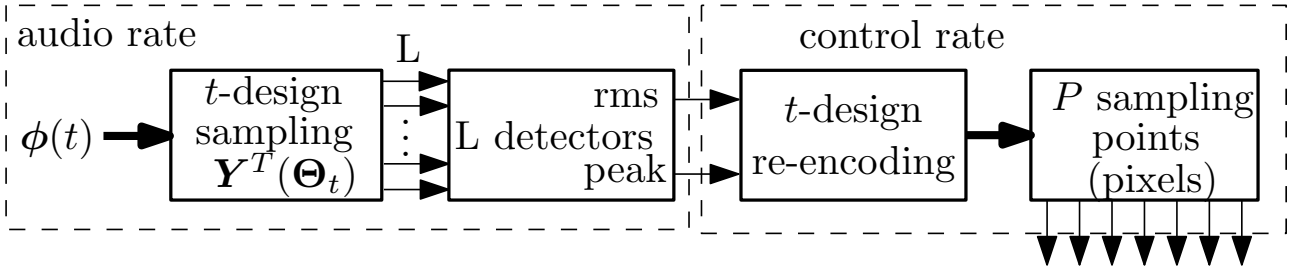
Implementierungen als Plug-ins

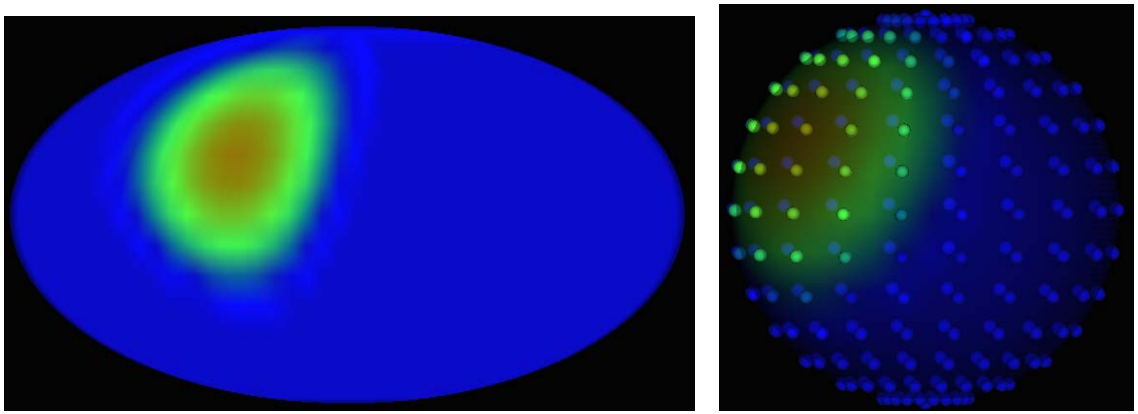
- *JUCE* C++ Bibliothek - Plug-in Infrastruktur, GUI
- *Eigen* C++ Bibliothek für lineare Algebra
- andere Bibliotheken für Samplerate-Conversion, FFT, OSC
- Open Source, getestet unter Windows, MacOS, (Linux)
- Anforderung an Host: flexible Busstruktur -> Reaper/Ardour

Implementierungen als Plug-ins



Visualisierung von Ambisonics-Signalen

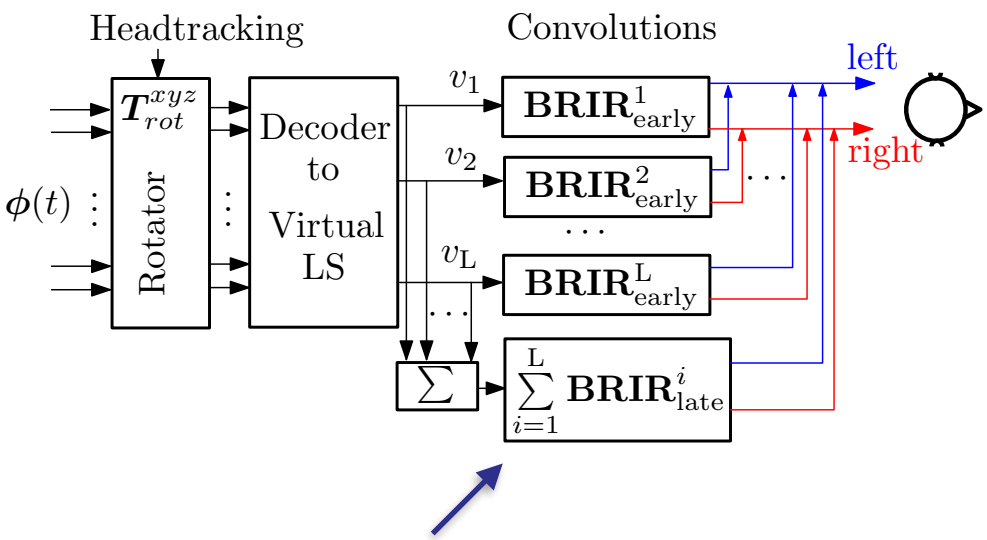




Visualisierung der Richtungslautstärke: Pure Data Prototyp

Dekodierung für Kopfhörerwiedergabe

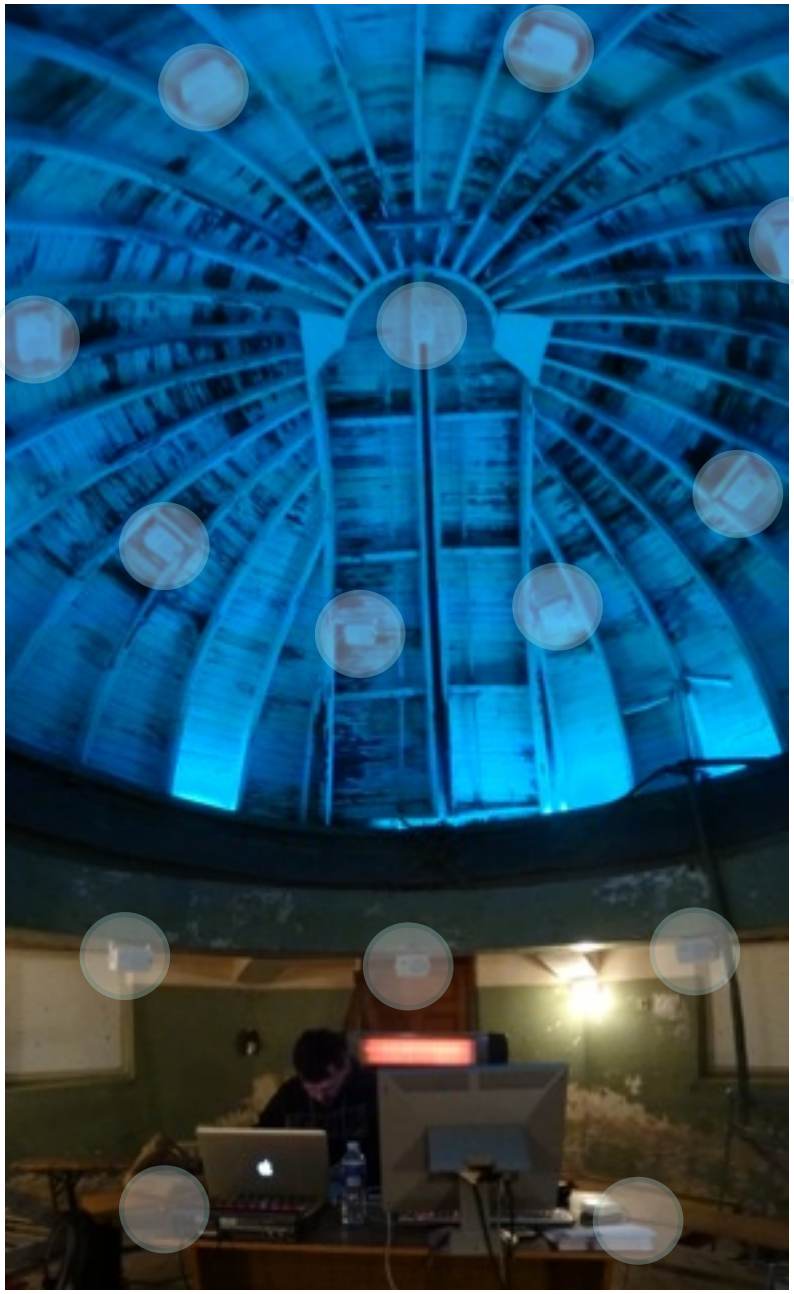
Dekodierung auf virtuelle Lautsprecher,
Faltung der Lautsprechersignale mit
Binauralen Raumimpulsantworten (BRIRs)

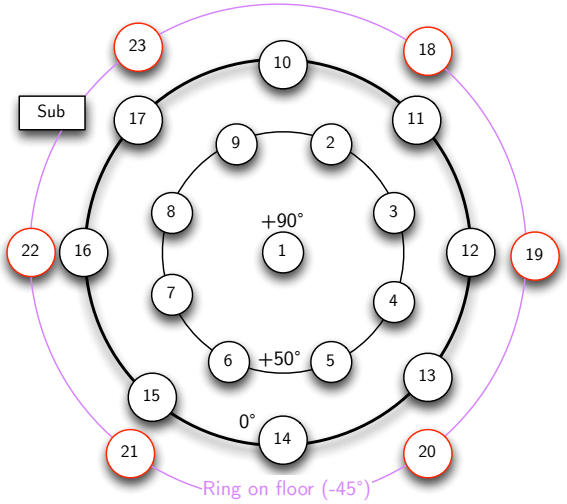


Optimierung durch kürzere Faltungen

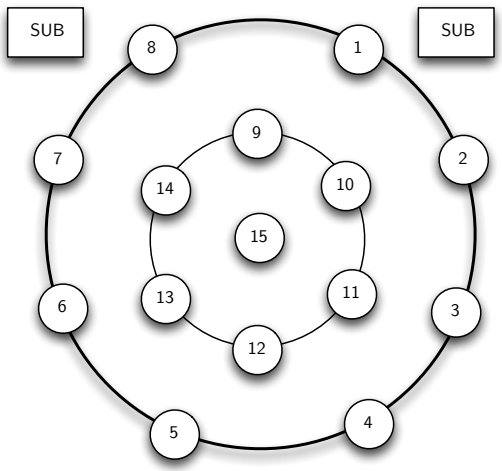
Headtracking mit Arduino [1] und 3-Achsen Beschleunigungssensor, Gyrometer und Magnetometer zur Drift-Kompensation

[1] D. Frie, "open-headtracker," http://code.google.com/p/open-headtracker, 2012.

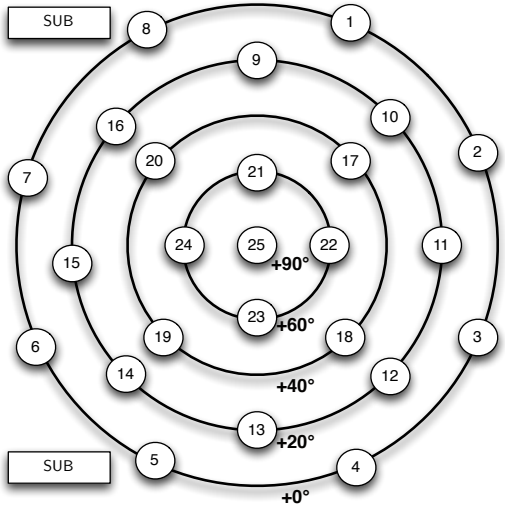




Astronomisches Observatorium der Vilnius Universität, 2013



EU-ICT Messe Vilnius, 2013.



Mobile IEM Ambisonics Kuppel, EAA Symposium Berlin, 2014.

Studienzentrum für Musikinnovation,

Litauische Musik und Theater Akademie, Vilnius, 2014

Zusammenfassung

- Allgemeine Transformationsvorschrift für Ambisonics
- Anwendung für Rotation, richtungsabhängige Lautstärkeanpassungen, Warping, Kreatives...
- Vorteile mit Slepian Functions für Lautstärkeanpassungen
- Suche nach t-designs durch nichtlineare Optimierung
- Plug-in Implementierungen

Fragen?

Danke.

Matthias Kronlachner