



**POLITECHNIKA
GDAŃSKA**



Ambisonia

Mgr inż. Bartłomiej Mróz
bartlomiej.mroz@pg.edu.pl

Problemy tradycyjnych systemów (stereo, 5.1)

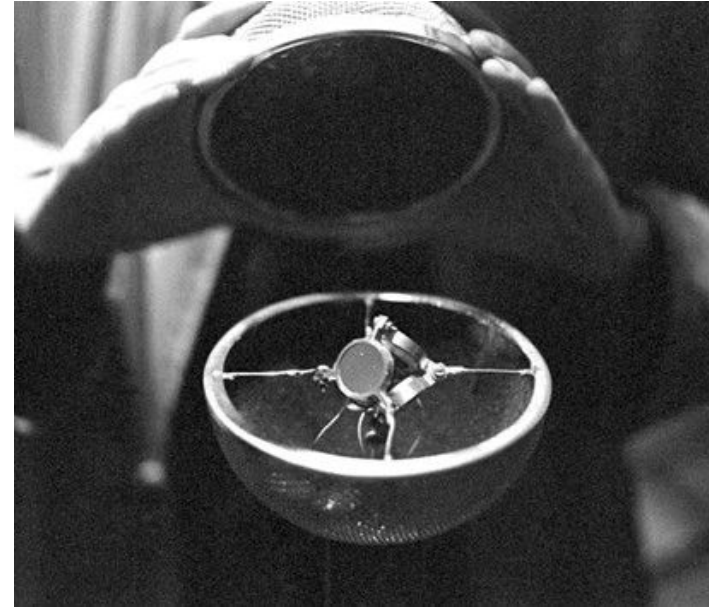
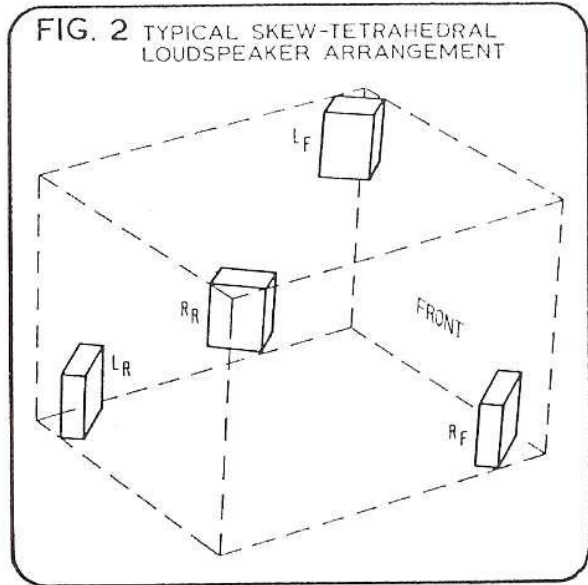
- Dobra lokalizacja tylko z kierunku frontального
- Niewielki obszar “*sweet spot*”
- Myślenie o sygnale w kategorii pozycji głośników

Czym jest ambisonia – definicja

W przeciwieństwie do innych wielokanałowych formatów surround, ambisoniczne kanały transmisji nie zawierają sygnałów głośnikowych. Zamiast tego zawierają one niezależną od głośnika reprezentację pola dźwiękowego (tzw. *B-format*), który jest następnie dekodowany do konfiguracji głośników słuchacza. Ten dodatkowy krok pozwala producentowi myśleć w kategoriach kierunków źródłowych, a nie na głośnikach i zapewnia słuchaczowi znaczną elastyczność w układzie i liczbie głośników używanych do odtwarzania.

Ambisonia – rys historyczny

- Opracowana przez zespół brytyjskich naukowców w latach siedemdziesiątych
- Wywodzi się z techniki o nazwie “peryfonia” (ang. *periphony*) – technika odtwarzania dźwięku zarówno w pionie, jak i w poziomie, wokół słuchacza
- Również w tamtym czasie: kwadrofonia (ang. *quadrophonics*)
- W szczególności należy podkreślić prace Michaela Gerzona ([link](#))
krótki – 8min – filmik dokumentalny o M. Gerzonie ([link](#))

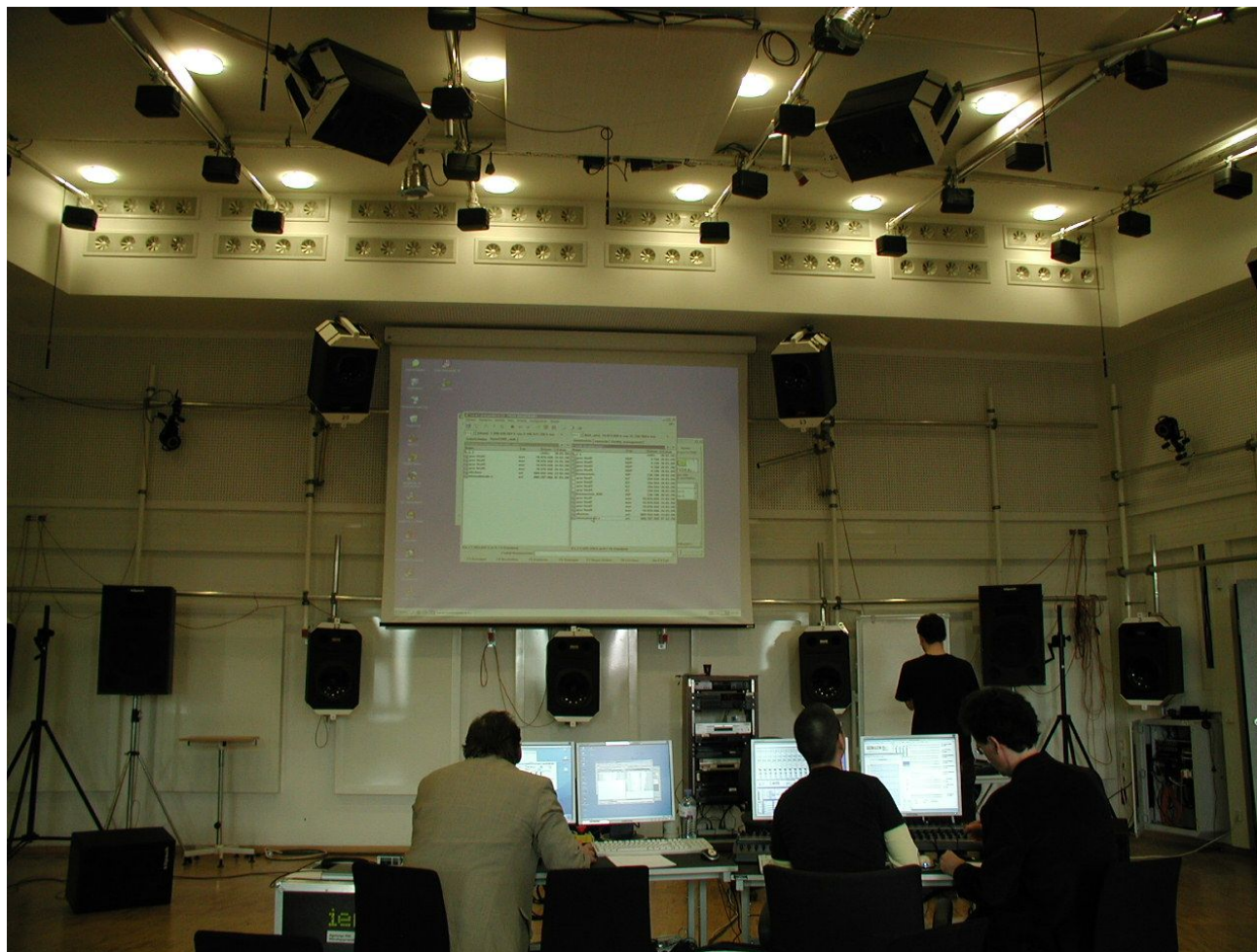


Przykładowe układy do odsłuchu oraz nagrywania przestrzennego,
opracowane przez Michaela Gerzona (1971)



Współczesne mikrofony ambisoniczne

Po lewej: Soundfield SPS200;
na środku: mh acoustics em32 Eigenmike®;
po prawej u góry: Sennheiser Ambeo® VR Mic;
Po prawej na dole: mikrofon Zylia ZM-1 oraz rekorder ZR-1



Kunstuniversität Graz, the IEM-Cube (Austria) – pół-sfera złożona z 24 głośników

Zyia 6-Degrees-of-Freedom system

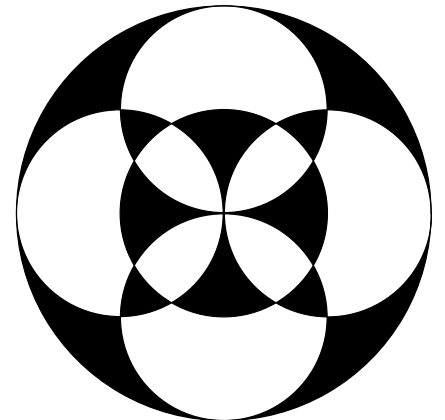
- <https://www.zyia.co/zyia-6dof.html>

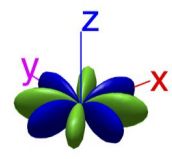
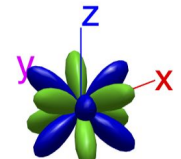
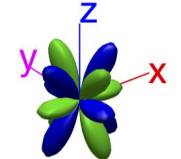
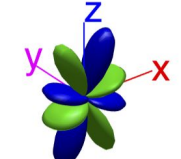
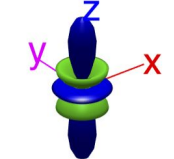
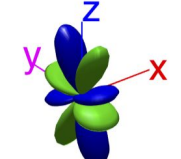
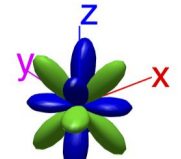
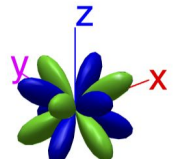
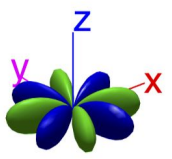
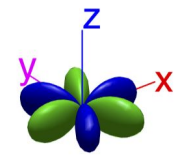
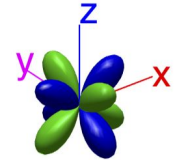
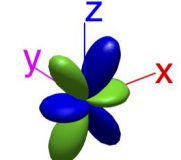
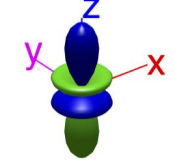
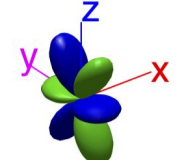
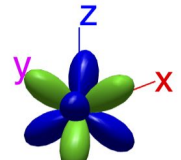
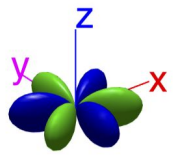
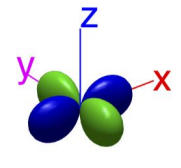
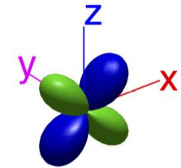
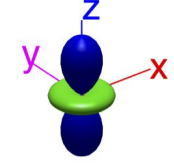
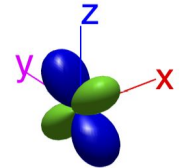
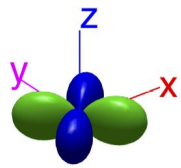
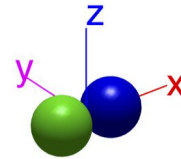
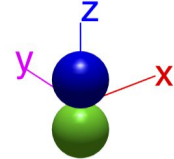
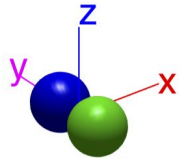
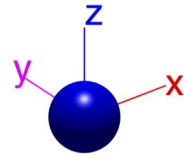




Podstawy ambisonii

- Przełożenie koincydentalnej pary M-S Blumleina do trzech wymiarów
- *A-format*: zależny od modelu mikrofonu
(Pierwszy kanał przenosi informację o amplitudzie sygnału, podczas gdy pozostałe kanały określają kierunkowość poprzez zależności fazowe między sobą)
- *B-format* (kanały WXYZ):
 - W = omni (front+back+left+right+up+down)
 - X = front – back
 - Y = left – right
 - Z = up – down





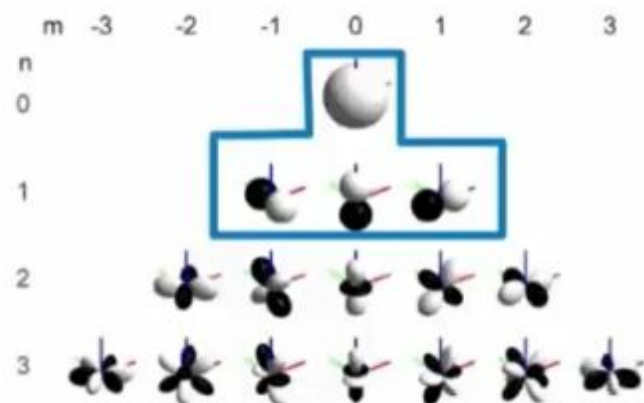
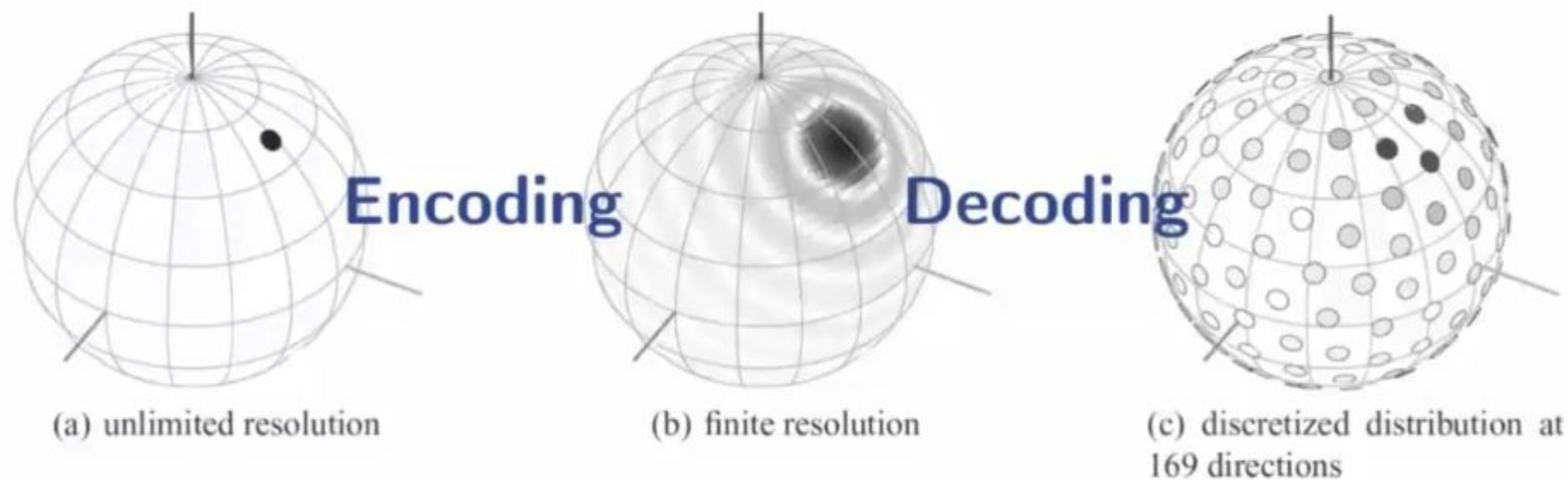


Fig. 1: First 16 spherical harmonics with order n and degree m .



*Main mic./close-mic,
channel/obj-based + FX,
Ambi Mics.*

Ambi production world

Effects

(mirror, rotate, directional
loudness, width, reverb, etc.,
directional compression, ...)

Scene-based transmission
(YouTube, MPEG-H, ITU-adm,
HOAST hoast.iem.at)

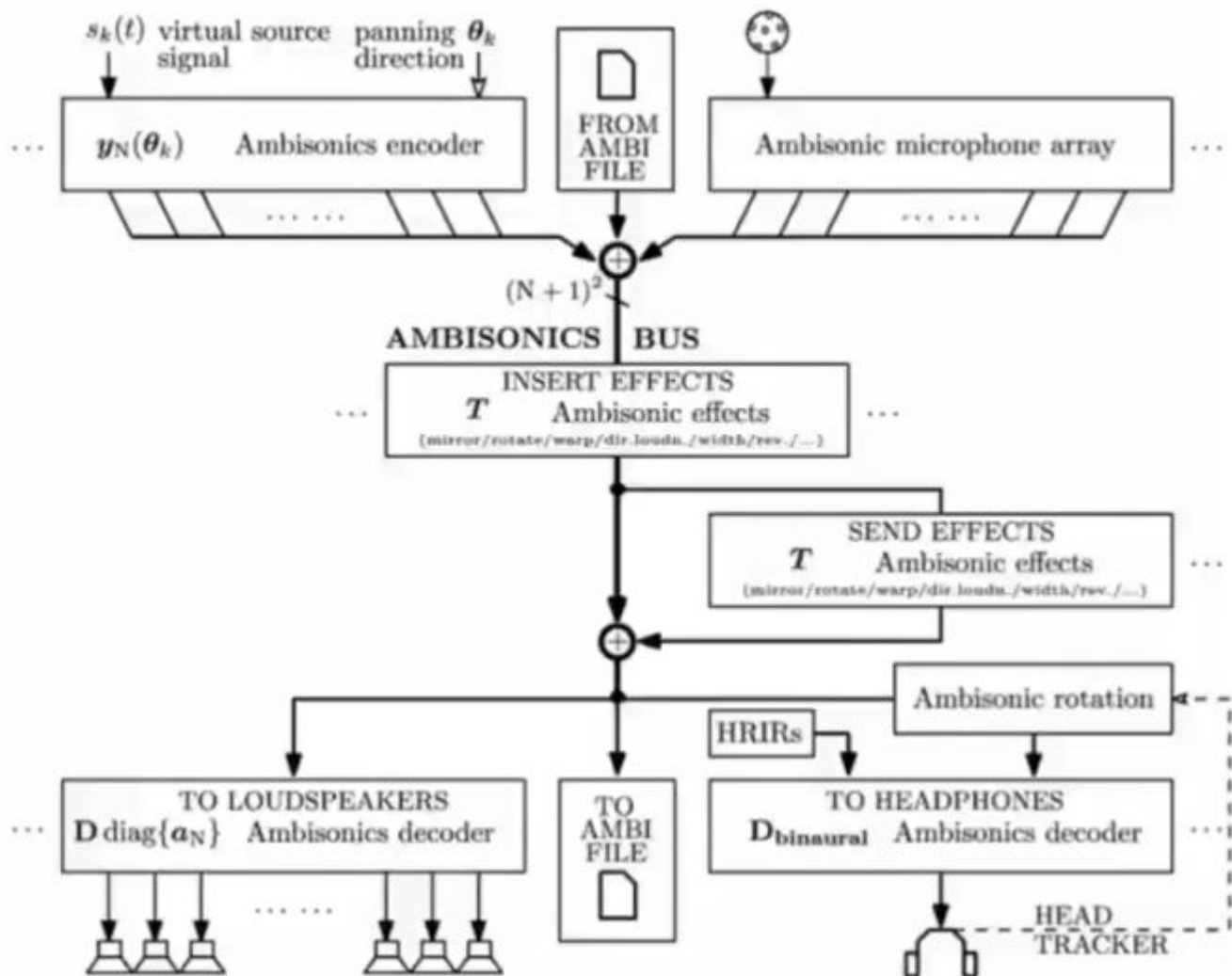
*Standard formats
(5.1, 7.1.4, ITU-R2051, ...)*

Custom loudspeaker layouts
Head-tracked **headphones**

INPUTS

AMBISONIC BUS EFFECTS

OUTPUTS



Podstawy ambisonii

- Ilość kanałów w ambisonii 2D: $2N + 1$ (tylko płaszczyzna horyzontalna)
- Ilość kanałów w ambisonii **3D**: $(N + 1)^2$ (pełna sfera)
- Numeracja kanałów: FuMa (*Furse-Malham*), ACN (*Ambisonic Channel Number*)
- Normalizacja kanałów: głównie SN3D, czasem też N3D, MaxN

Numeracja kanałów

Ambisonic Order

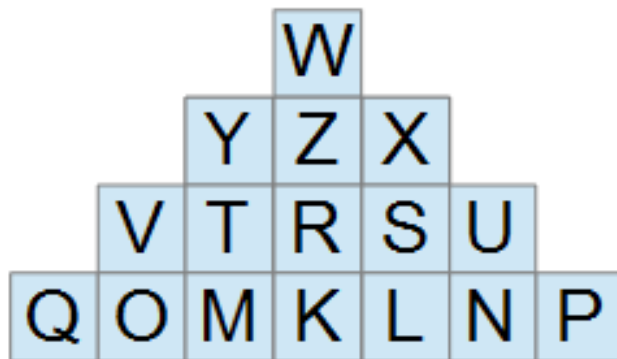
Mono

1st-order full sphere

2nd-order full sphere

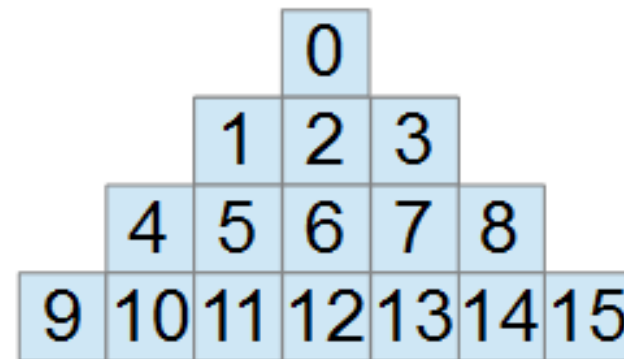
3rd-order full sphere

FuMa



Read alphabetically in groupings of three orders (W,X,Y,Z)(R,S,T,U,V)(K,L,M...)

ACN



Read numerically (0,1,2,3,4,5,6...)

Normalizacja kanałów

- MaxN - normalizuje każdy pojedynczy komponent, aby nigdy nie przekraczał wzmocnienia 1,0 dla spanoramowanego źródła monofonicznego – używany w FuMa
- N3D – podobna do SN3D – ortonormalna podstawa dekompozycji 3D. Zapewnia równą moc kodowanych komponentów w przypadku idealnie rozproszonego pola 3D.
- SN3D – (w kolejności kanałów ACN) jest szeroko stosowana. W przeciwieństwie do N3D, żaden komponent nigdy nie przekroczy wartości szczytowej komponentu 0 rzędu dla źródeł jednopunktowych. Ten schemat został przyjęty przez proponowany format AmbiX

Jak to wygląda w (matematycznej) praktyce?

ACN	Order	Angle/Elevation Representation	Cartesian Representation
0	0	1	1
1	1	$\sin(a)\cos(e)$	y
2	1	$\sin(e)$	z
3	1	$\cos(a)\cos(e)$	x
4	2	$\sqrt{3/4}\sin(2a)\cos(e)\cos(e)$	$\sqrt{3}xy$
5	2	$\sqrt{3/4}\sin(a)\sin(2e)$	$\sqrt{3}yz$
6	2	$(1/2)(3\sin(e)\sin(e)-1)$	$(1/2)(3zz-1)$
7	2	$\sqrt{3/4}\cos(a)\sin(2e)$	$\sqrt{3}xz$
8	2	$\sqrt{3/4}\cos(2a)\cos(e)\cos(e)$	$\sqrt{3/4}(xx-yy)$
9	3	$\sqrt{5/8}\sin(3a)\cos(e)\cos(e)\cos(e)$	$\sqrt{5/8}y(3xx-yy)$
10	3	$\sqrt{15/4}\sin(2a)\sin(e)\cos(e)\cos(e)$	$\sqrt{15}xyz$
11	3	$\sqrt{3/8}\sin(a)\cos(e)(5\sin(e)\sin(e)-1)$	$\sqrt{3/8}y(5zz-1)$
12	3	$(1/2)\sin(e)(5\sin(e)\sin(e)-3)$	$(1/2)z(5zz-3)$
13	3	$\sqrt{3/8}\cos(a)\cos(e)(5\sin(e)\sin(e)-1)$	$\sqrt{3/8}x(5zz-1)$
14	3	$\sqrt{15/4}\cos(2a)\sin(e)\cos(e)\cos(e)$	$\sqrt{15/4}z(xx-yy)$
15	3	$\sqrt{5/8}\cos(3a)\cos(e)\cos(e)\cos(e)$	$\sqrt{5/8}x(xx-3yy)$

Niniejsza tabela zawiera dwa różne sposoby panoramowania formatu B, w zależności od tego, czy położenie dźwięku jest wyrażone za pomocą kąta i wysokości, czy reprezentacji kartezjańskiej. W pierwszym przypadku, kąt jest mierzony w kierunku przeciwnym do ruchu wskazówek zegara i wzniesienia w górę. Jeśli użyjesz tego drugiego, współrzędne x, y, z muszą być składowymi wektora jednostkowego w kierunku dźwięku (tj. $xx + yy + zz$ musi wynosić 1).

Te równania są tutaj w celach informacyjnych. Nie trzeba mieć o nich wiedzy, aby korzystać z technologii i tworzyć nagrania ambisoniczne!

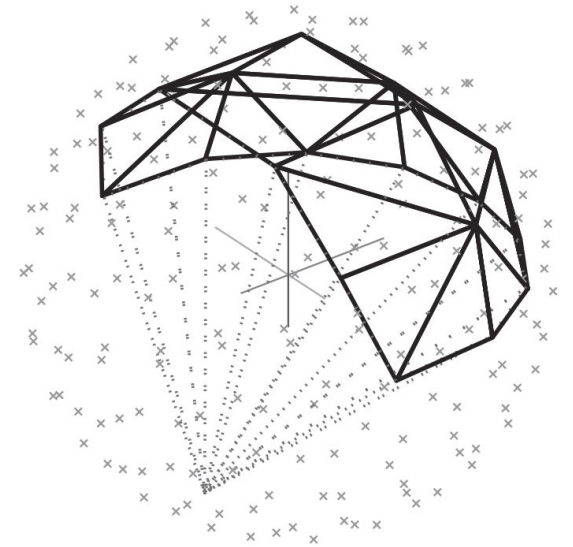
Formaty kodowania

- UHJ – historyczny
- AMB – Microsoft Wave Format Extensible (WAVE-EX). Zapis B-Formatu z współczynnikami wagowymi FuMa. Główną wadą jest 4GB limit pliku, ze względu na Microsoftowy format WAV jako kontener.
- AmbiX – *Ambisonic Exchange*. Apple'owski .caf (Core Audio Format) jako kontener. Normalizacja SN3D, numerowanie kanałów ACN – przyjęte przez Google jako podstawa formatu YouTube 360.

Dekodowanie

- VBAP
- MBAP
- AllRAD

All-Round Ambisonic Panning (AllRAP) jest algorytmem do arbitralnego ustawiania głośników, mającym na celu tworzenie fantomowych źródeł o stabilnej głośności i regulowanej szerokości. Odpowiednik All-Round Ambisonic Decoding (AllRAD) pasuje do koncepcji formatu ambisonicznego. Konwencjonalne dekodowanie ambisoniczne jest proste tylko z optymalnymi ustawieniami głośników, dla których osiąga niezależny od kierunku rozkład energii, szacowaną głośność i szerokość źródła fantomu. AllRAP / AllRAD jest nadal prosty, ale bardziej wszechstronny i wykorzystuje połączenie wirtualnego optymalnego ustawienia głośników z Vector-Base Amplitude Panning.



Źródła do dalszych poszukiwań

- IEM Plugin suite: <https://plugins.iem.at>
- Facebook Spatial Audio Workstation:
<https://facebook360.fb.com/spatial-workstation/>
- Zoom H2n conversion plugin:
<https://endabates.wordpress.com/2016/10/27/zoom-h2n-conversion-plugin/>
- YouTube spatial audio:
<https://support.google.com/youtube/answer/6395969>



POLITECHNIKA
GDAŃSKA

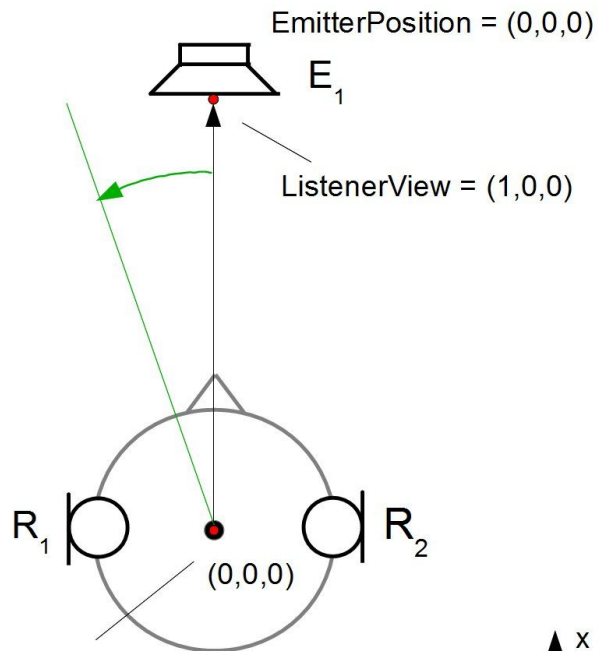


Dźwięk binauralny

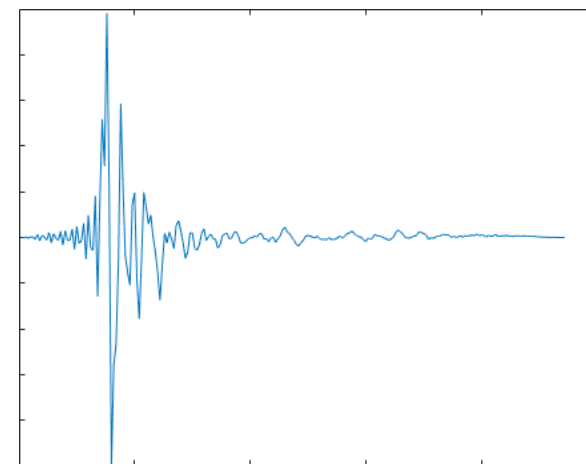
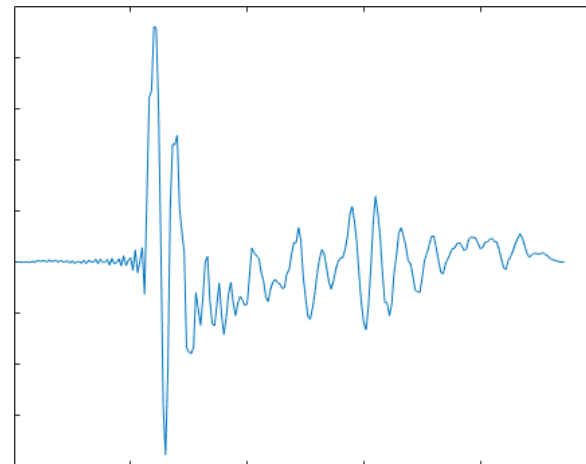
Mgr inż. Bartłomiej Mróz
bartlomiej.mroz@pg.edu.pl

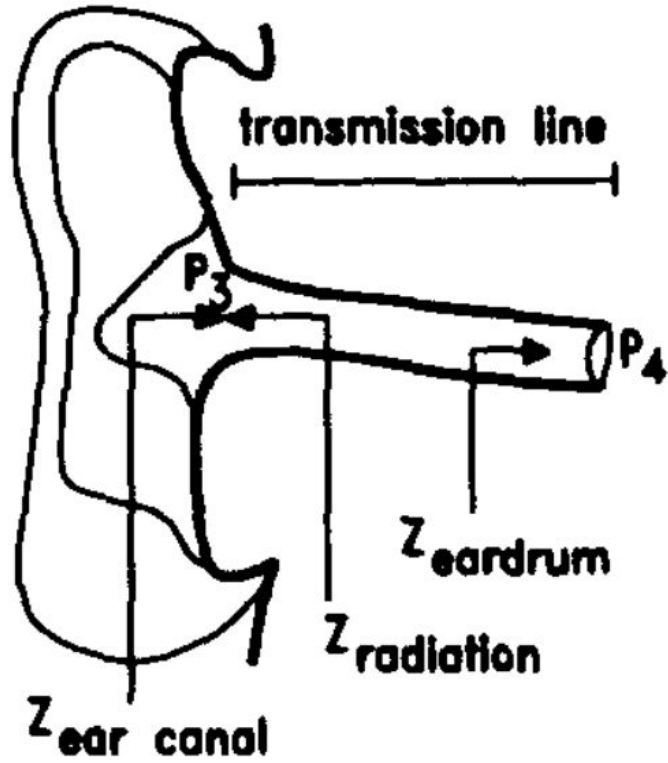
Funkcje HRTF

SourcePosition = (azimuth, elevation, radius)

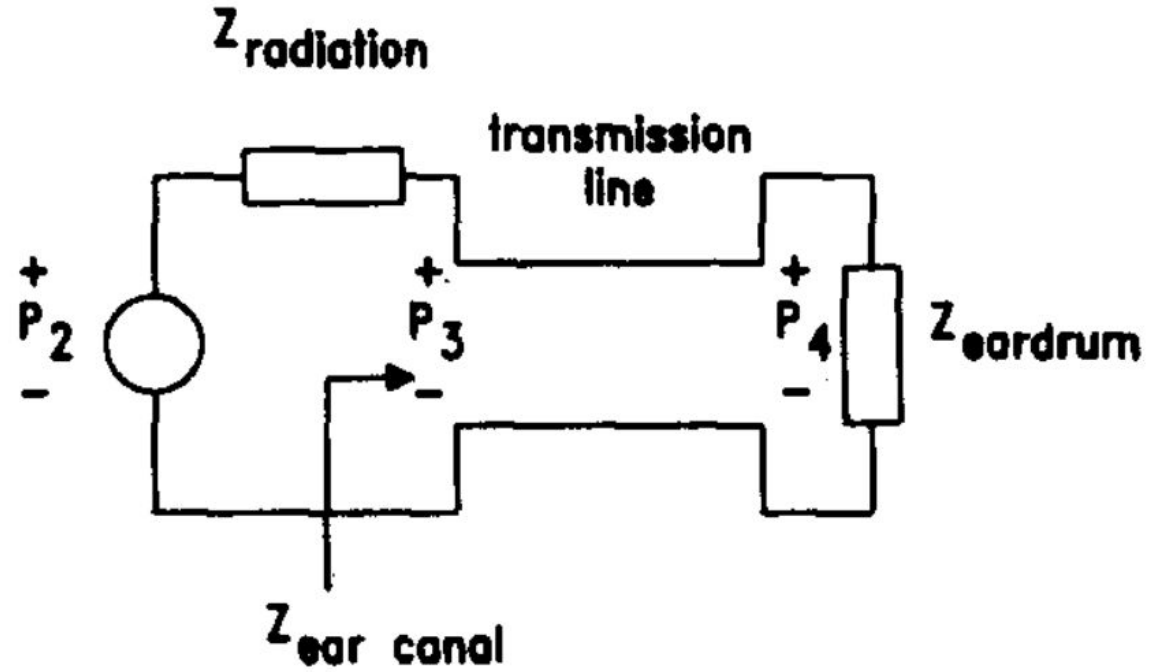


ListenerPosition = $(0,0,0)$

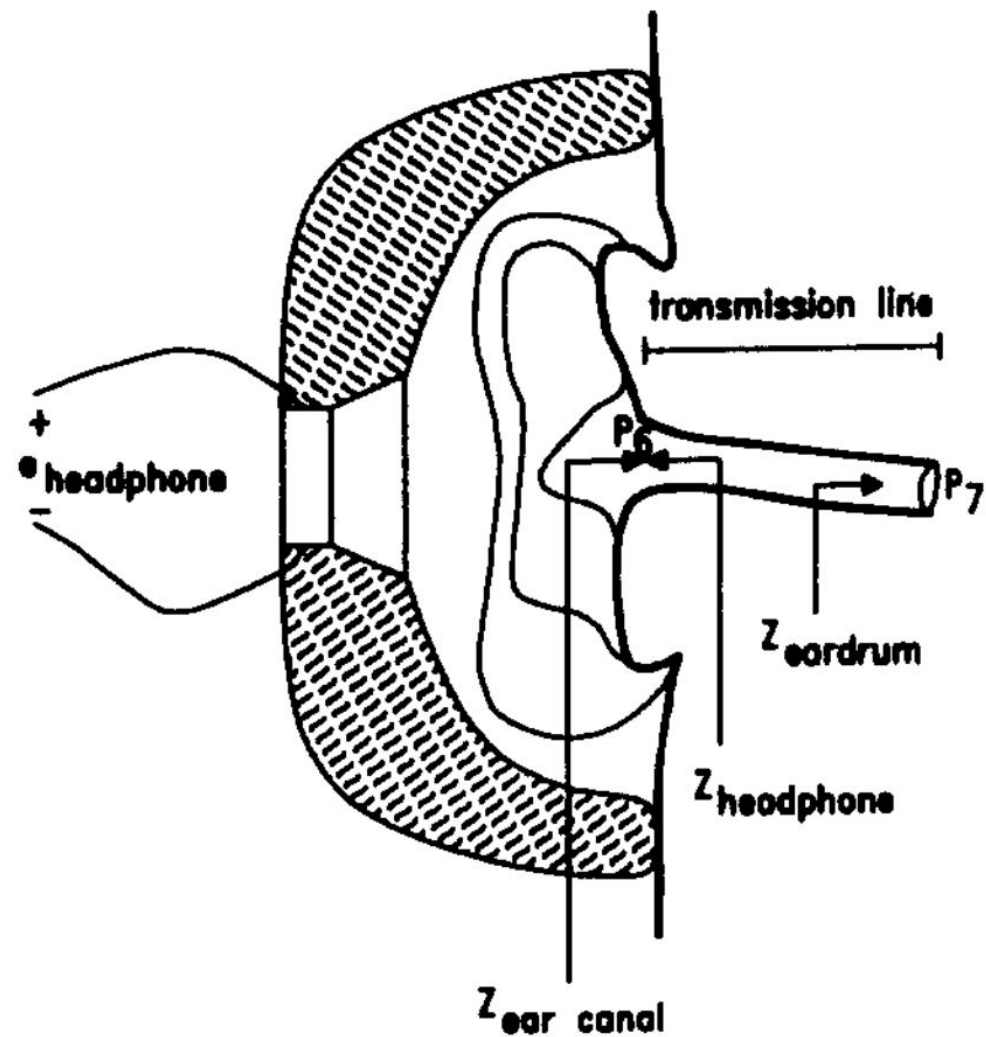




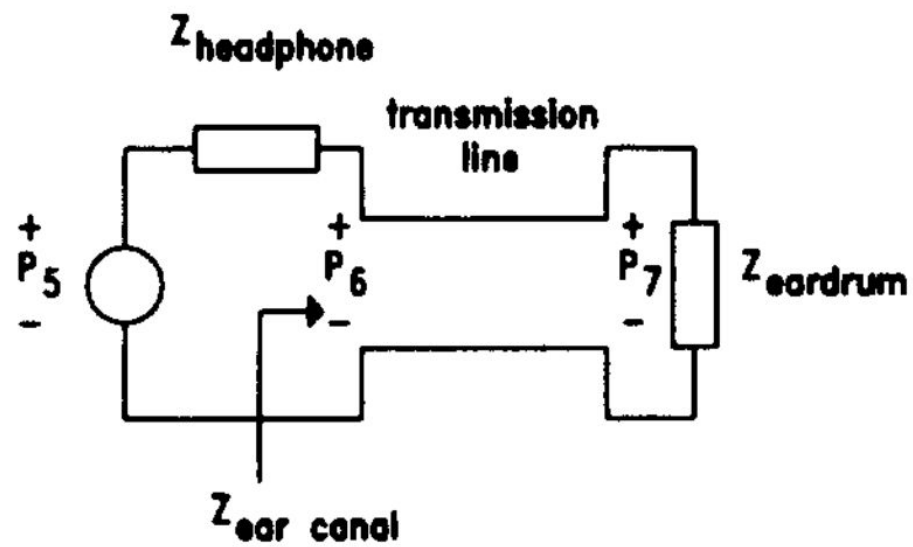
(a)



(b)



(a)



(b)

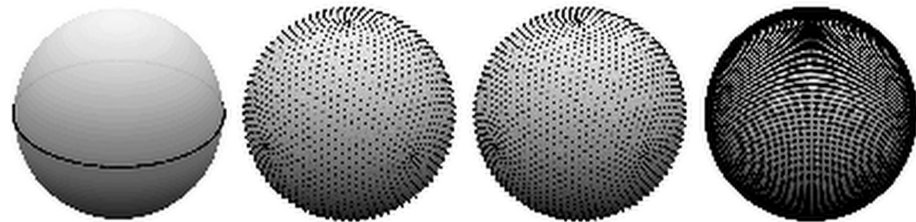
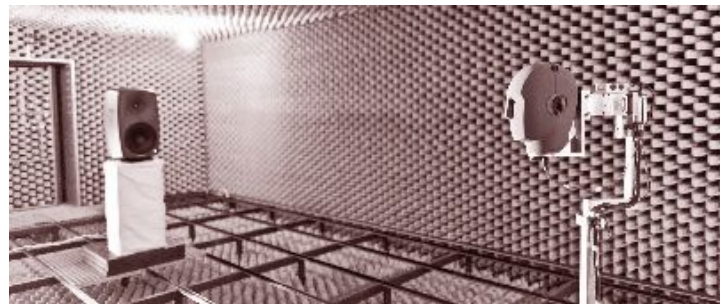
SOFA (Spatially Oriented Format for Acoustics)

SOFA to format pliku do przechowywania przestrzennie zorientowanych danych akustycznych, takich jak związane z głową funkcje przenoszenia (HRTF) oraz binauralne lub kierunkowe odpowiedzi impulsowe w pomieszczeniu (BRIR, DRIR). SOFA została znormalizowana przez Audio Engineering Society (AES) jako **AES69-2015**.

Otrzymywanie funkcji HRTF

Manekin

- KEMAR
- Neumann KU100



<https://sound.media.mit.edu/resources/KEMAR.html>

<https://audiogroup.web.th-koeln.de/ku100hrir.html>

Otrzymywanie funkcji HRTF

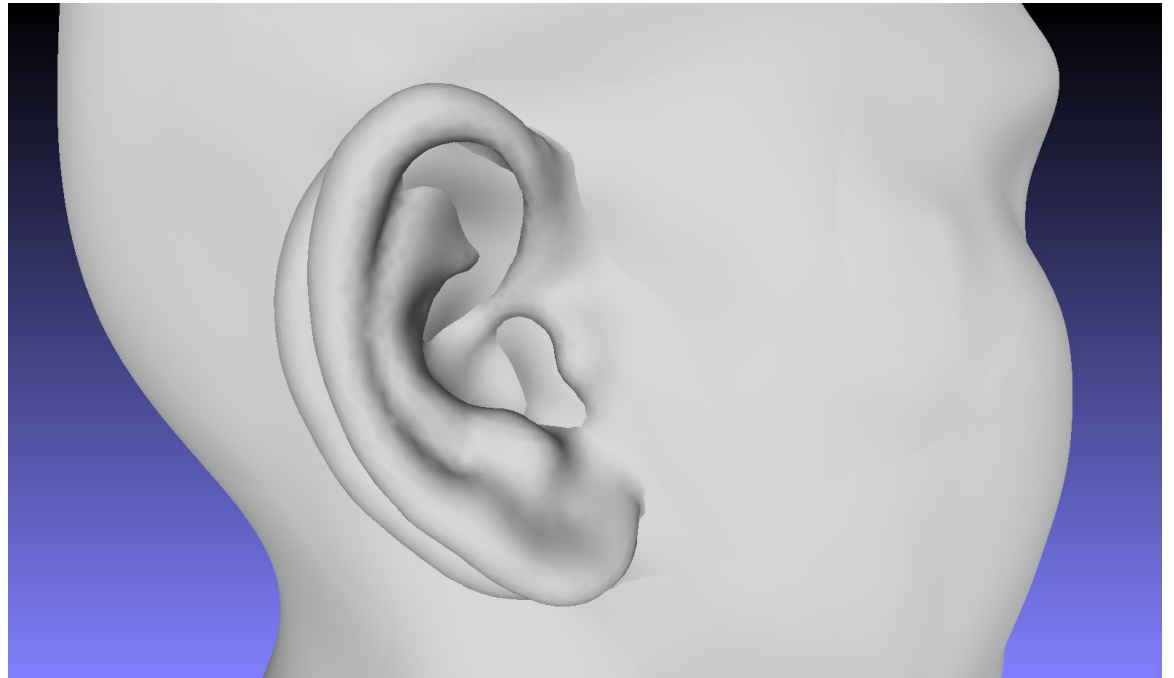
Pomiar indywidualny



Otrzymywanie funkcji HRTF

Symulacje numeryczne

- zdjęcia
- skan 3D



Dopasowywanie funkcji HRTF

- Za pomocą danych antropomorficznych
- Za pomocą informacji zwrotnej

- ♦ Adaptacja
- ♦ Selekcja

Eksternalizacja

Wrażenie słyszenia dźwięków *spoza głowy*

- Indywidualne / dopasowane funkcje HRTF
- Equalizacja słuchawek
- Pomiar HRTF przy błonie bębenkowej
- Efekt zgodności pomieszczenia
- Dodanie mechanizmu śledzącego ruchy głowy, aby móc je kompensować i dynamicznie przełączać funkcje HRTF zgodnie z pozycją głowy