

サウンド・デザイン

福岡女学院大学 2021年度 前期 木曜2限 第11週

講師：松浦知也

teach@matsuuratomoya.com

teach.matsuuratomoya.com

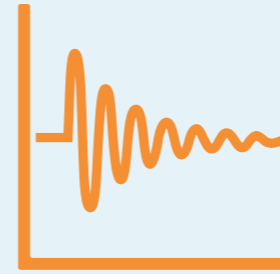


音と知覚、音と本物らしさ
空間と音

今回の内容：

- 音を知覚する時のもう一つの大きな側面、空間知覚について知る
- 空間再現のための主要な技術やフォーマットについて知る
- 未来の音楽のフォーマットについて考えてみよう(音楽体験そのもののデザイン)

今日考えるところ



量的な共有
Quantitative

どうして音の方向を知覚できる？

統計的な共有
Statistical

空間知覚は人により異なる？



質的な共有
Qualitative

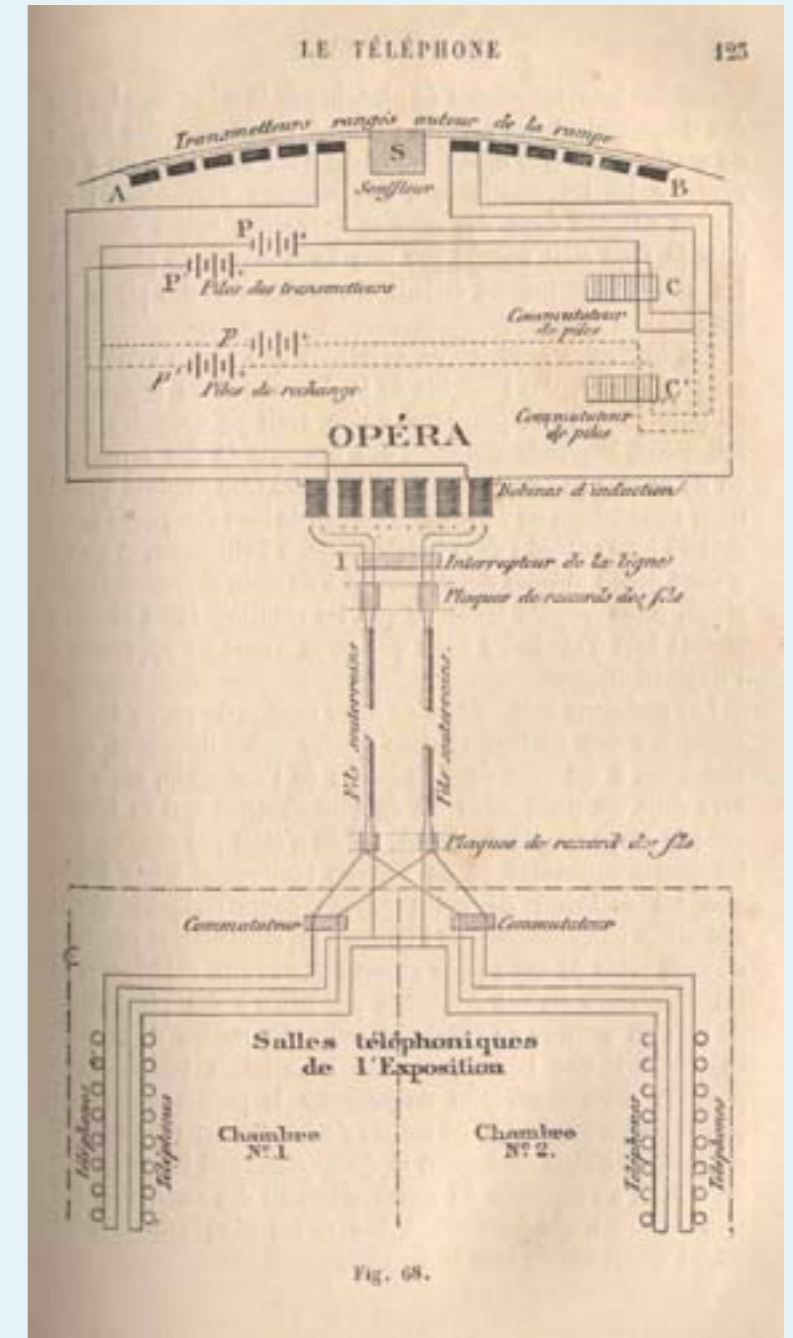


大勢にサラウンドを同時に聴かせられる？

ステレオの誕生

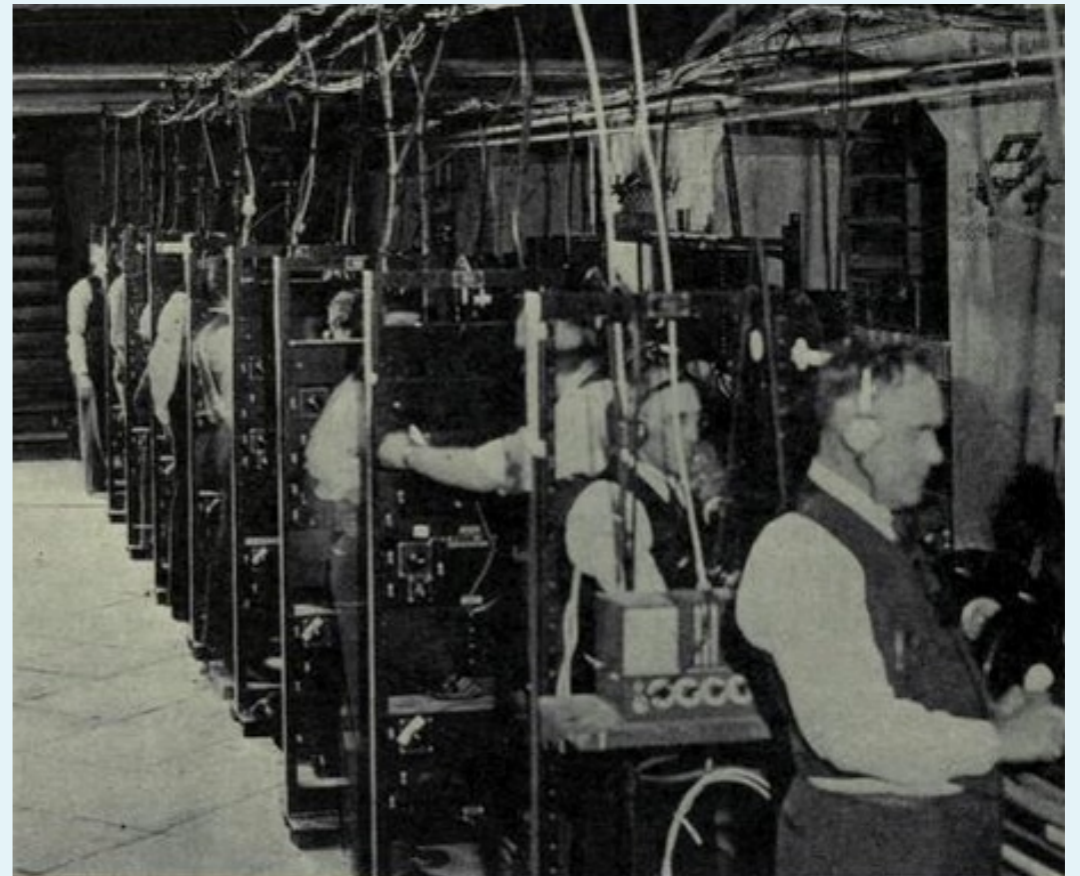
パリ万博(1881)のThéâtrophone

- オペラハウスから万博会場まで2kmほどを電話で繋ぎ、万博会場で客は受話器を当てて聴く
- 両耳それぞれに離れた位置のマイクが対応している
- ある意味、ストリーミングの先駆け？



ディズニー“Fantasia”のFantasound(1940)

- 初めてステレオフォニックサウンドを採用したアニメーション
- 時期によって様々な構成があるが、3chオーディオトラックを8chスピーカーに分配など
- 初期はオペレーターが手動でパンを振っていたが難しすぎたため、パン位置をコントロールするトラックが導入される
- クリックトラックを用いた同時多重録音の最初期の例でもある



大阪万博(1970)

鉄鋼館 スペース・シアター

- スピーカー計1008個
- 武満徹、高橋悠治、Iannis Xenakisらが作曲
- 今は中に入れないがガラス越しには見れる



大阪万博(1970) ペプシ館

- David Tudorによる37つのスピーカーをマイクからのフィードバックを用いて鳴らす”Microphone”などが演奏された



<https://www.expo70-park.jp/cause/expo/pepsi/>

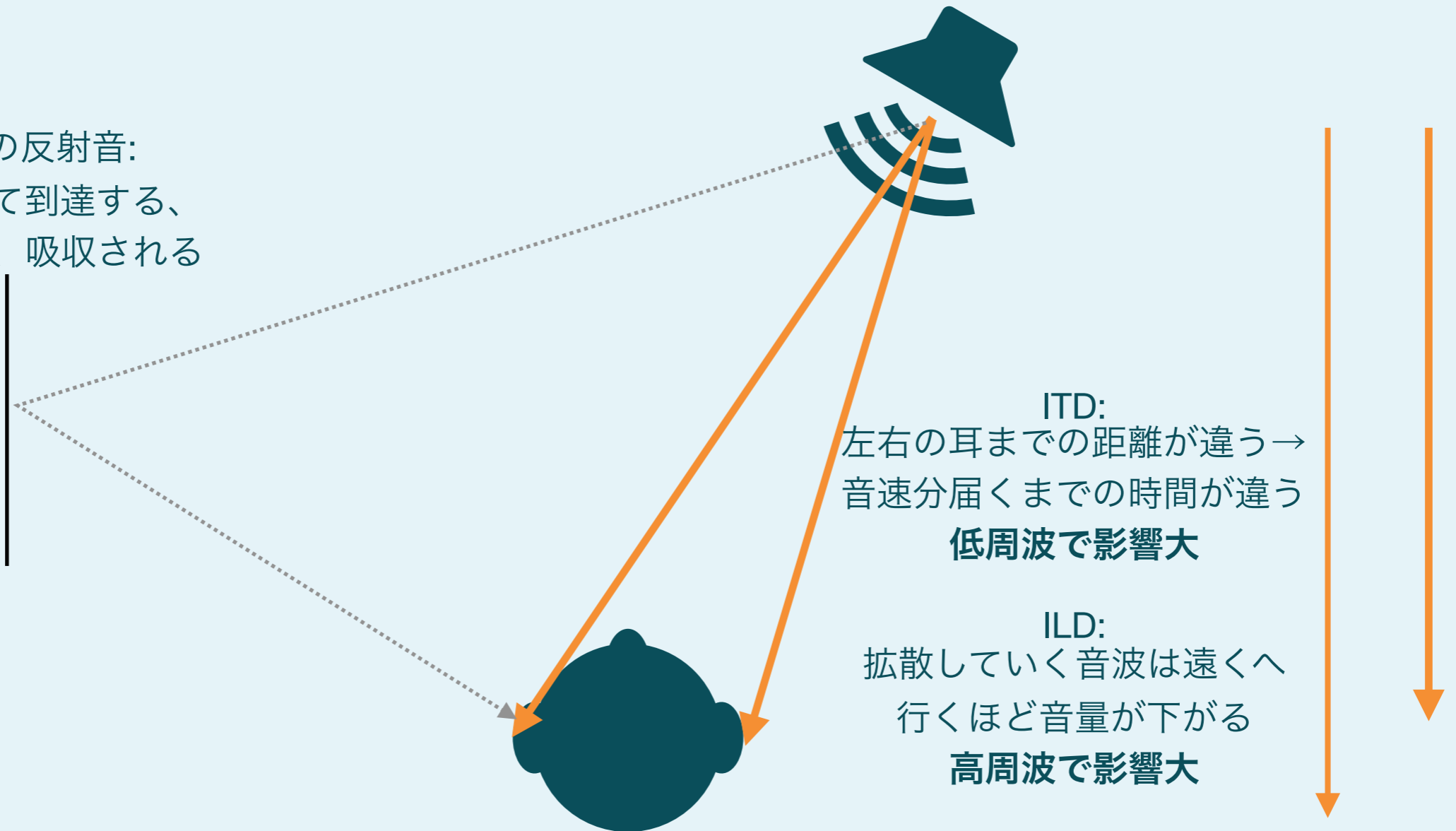
音の方向を知覚する仕組み

- 左右方向：
 - 両耳間レベル差：ILD(Interaural Level Difference)
 - 両耳間時間差：ITD(Interaural Time Difference)
- 上下、前後方向：頭部伝達関数(Head-Related-Transfer-Function)
- その他、残響の変化など

音の方向を知覚する仕組み

左右の知覚

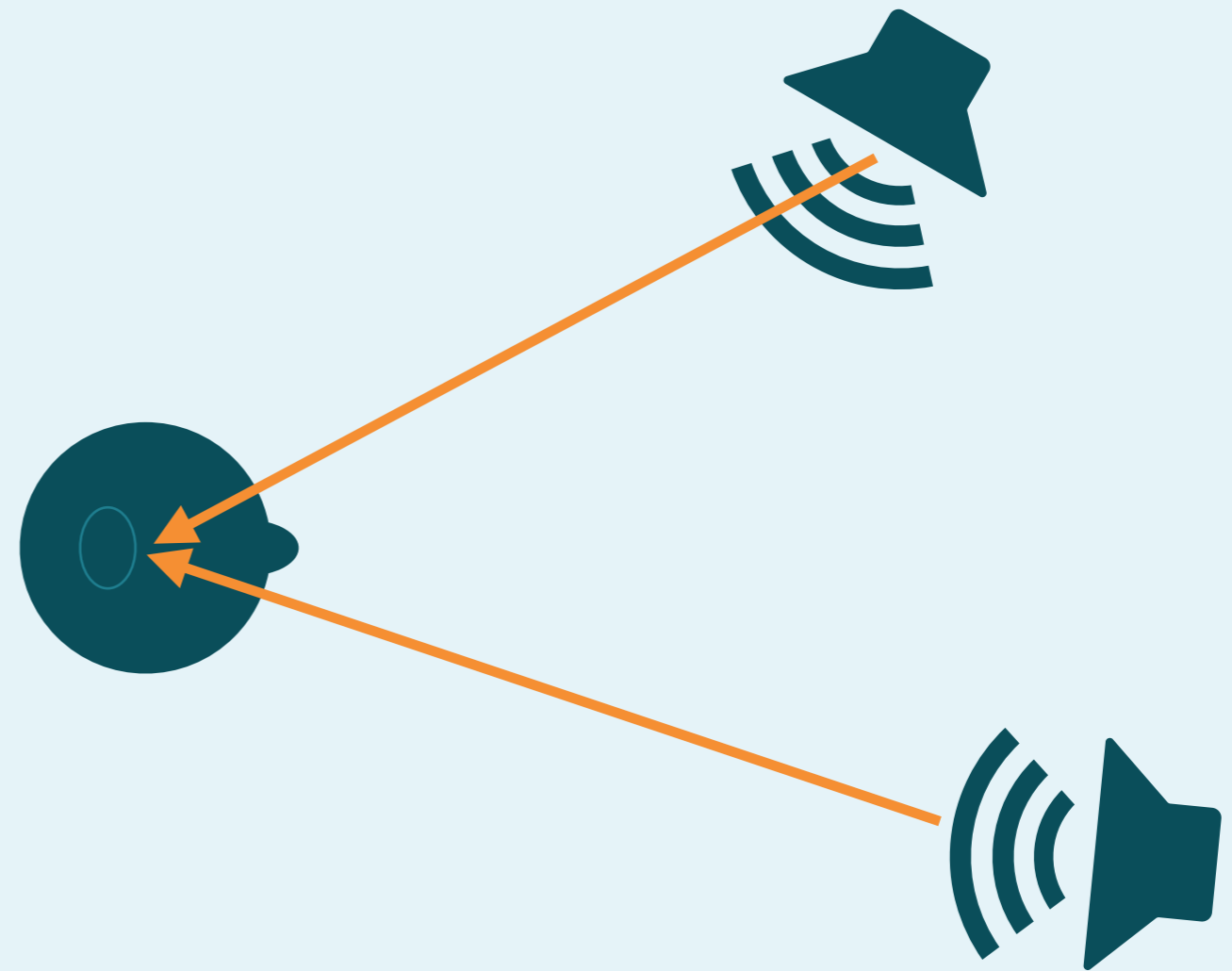
壁からの反射音:
さらに遅れて到達する、
高域が拡散、吸収される



音の方向を知覚する仕組み

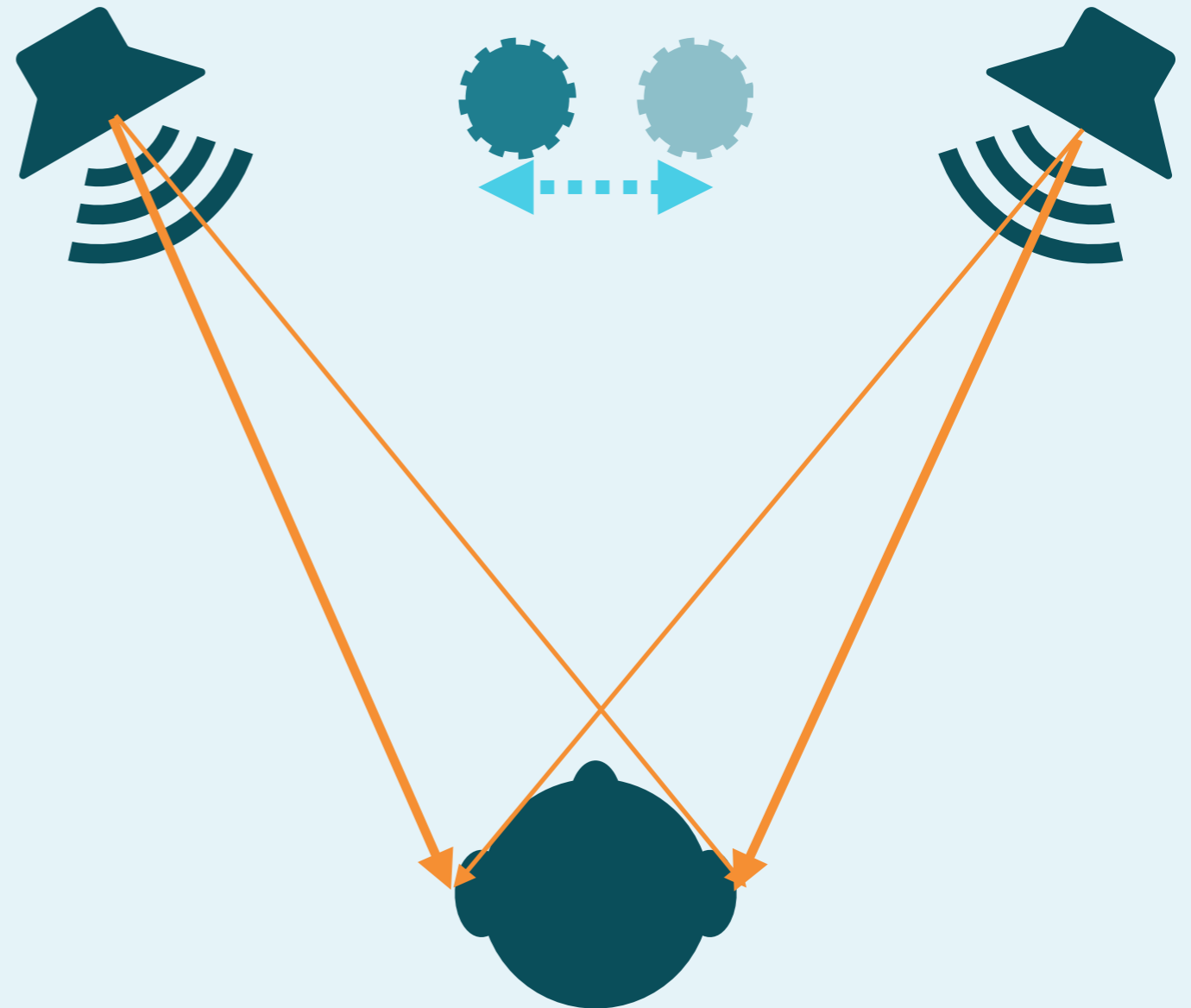
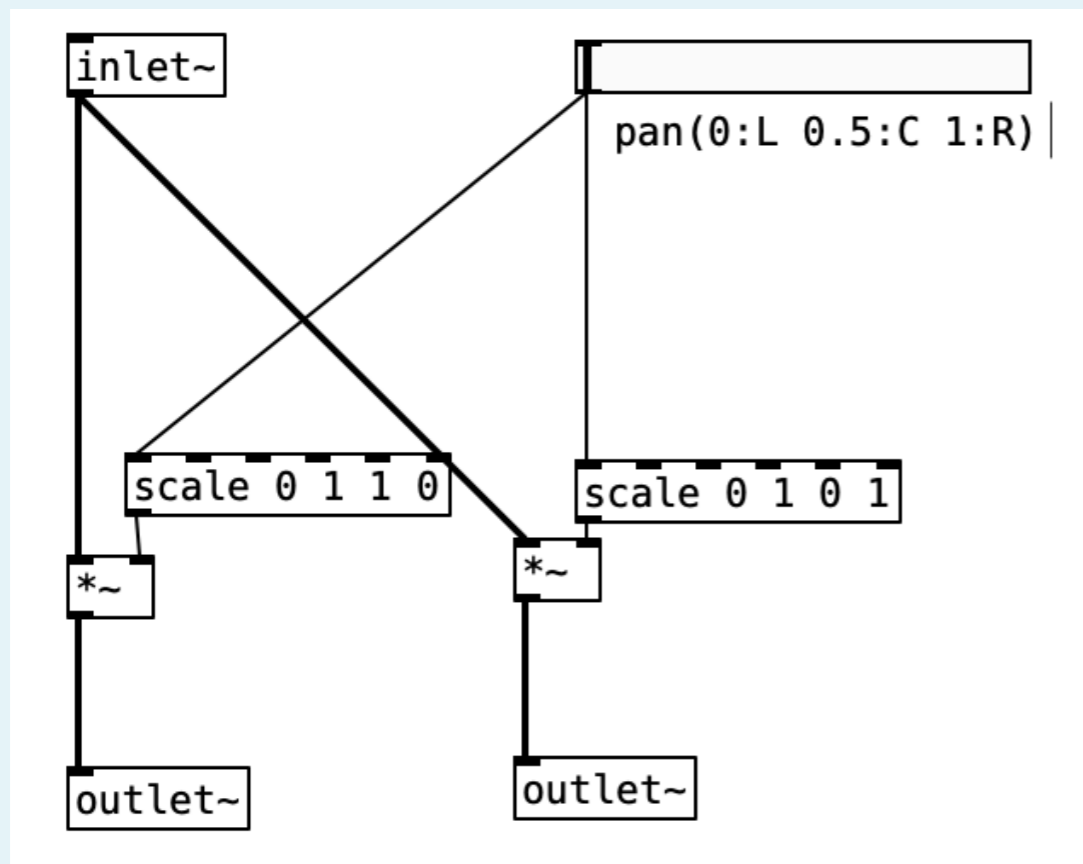
上下、前後の知覚

- 耳介や頭における反射で音色が変化する
- 頭の向きが動くことで起きる音色変化の影響
- 視覚による補強
- 前後対称の位置だと誤認することもある



ステレオ再生

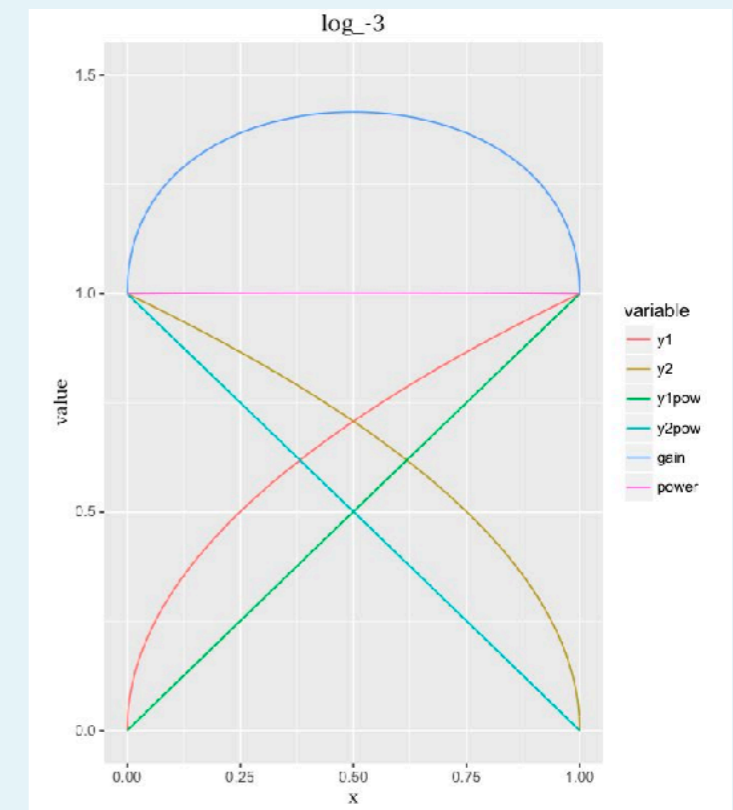
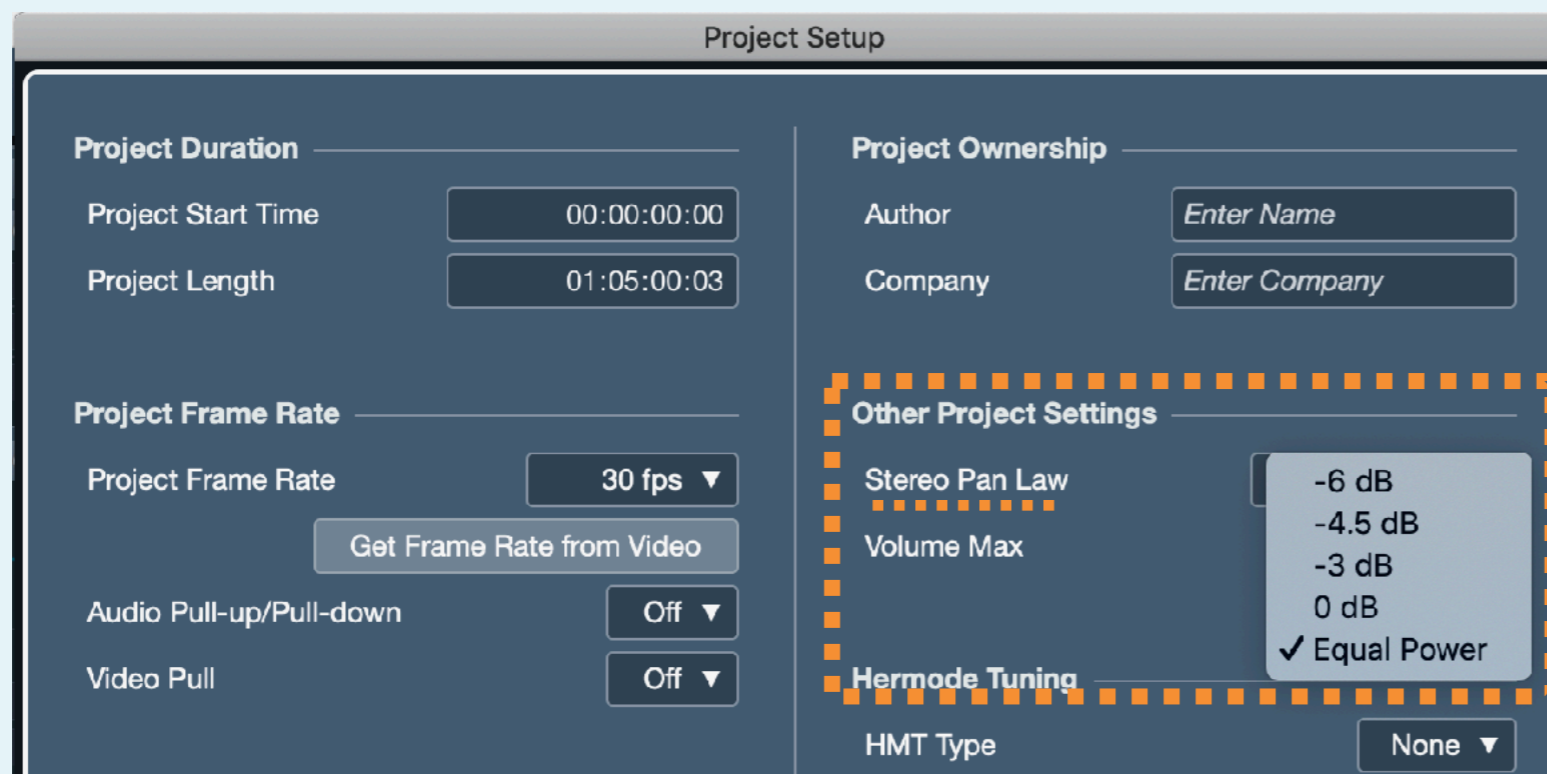
Level (Amplitude) Based Panning



左右の音量を変化させるだけ（世の中の8、9割のパンナーはこれベース）

ステレオ再生

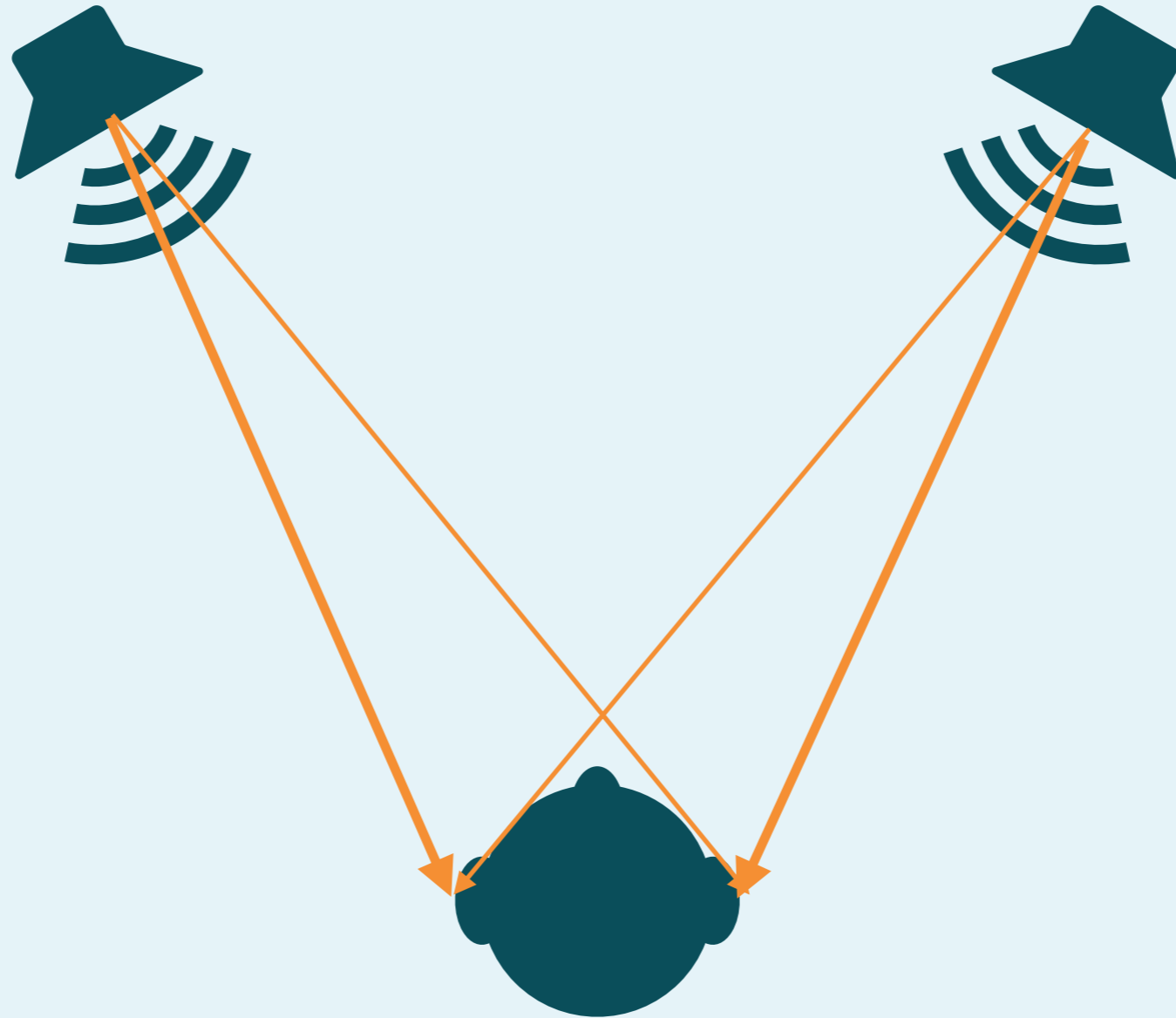
Level (Amplitude) Based Panning



- センターなら左右それぞれ音圧0.5倍.....**ではない!**
- センター0.5(=-6dB)だと、部屋の残響の影響などで音量が小さく聞こえてしまう。対数カーブでセンターを-3dBにするとエネルギー的に等価
- DAWではカーブが選べることが多い。迷ったら-3dB, Equal Power, Constant-Powerと書いてるものを選ぶ(スタジオなど、残響時間の短い部屋では-4.5などを使うこともある)。

2chステレオ再生

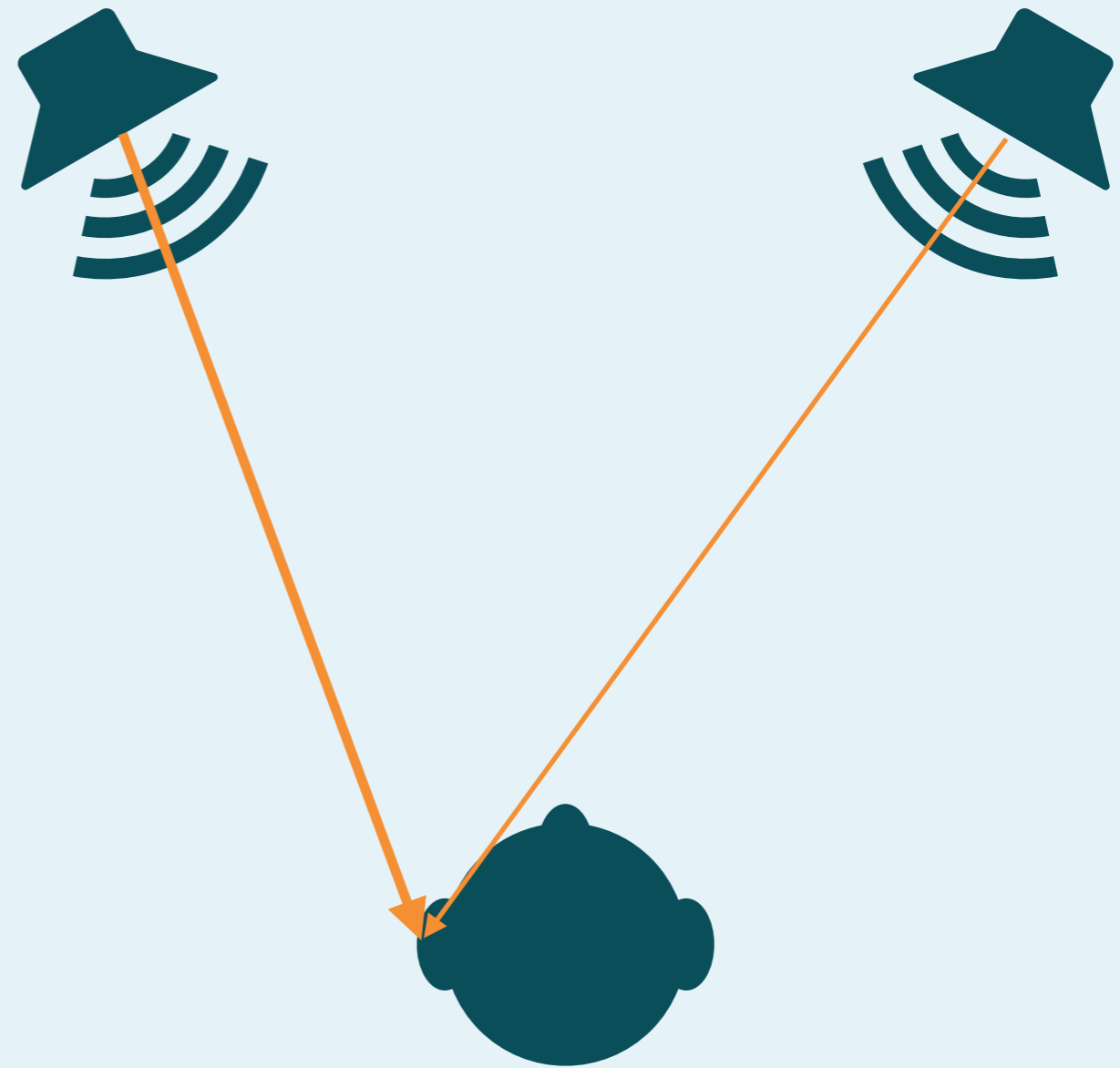
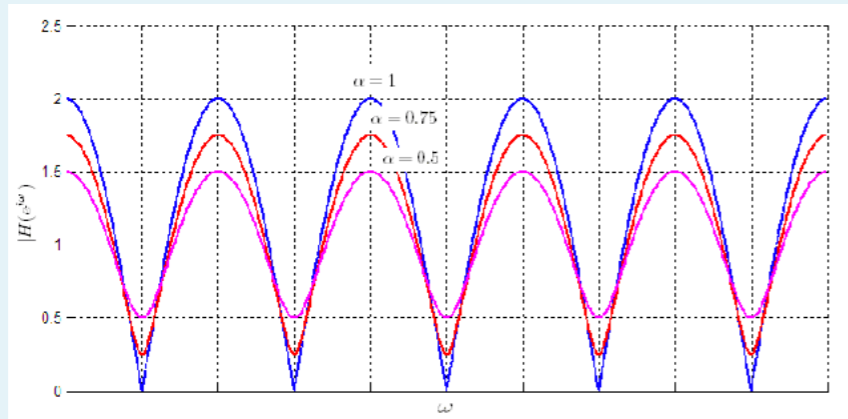
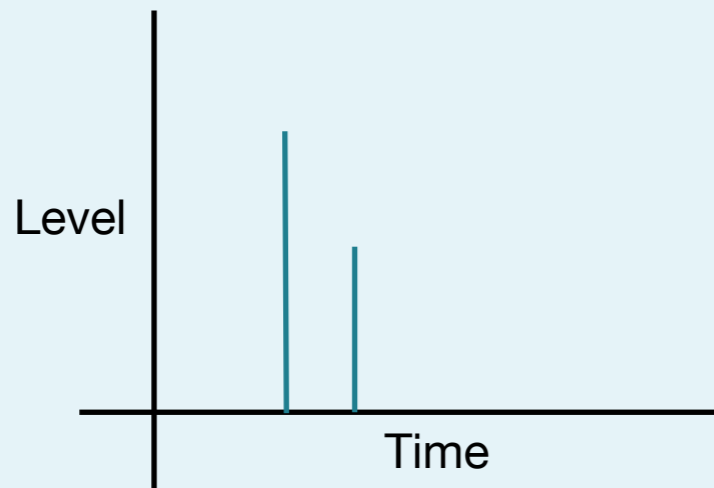
チャンネル間クロストーク



スピーカーでステレオ再生する場合、右チャンネルから出た音も左耳に届いている

2chステレオ再生

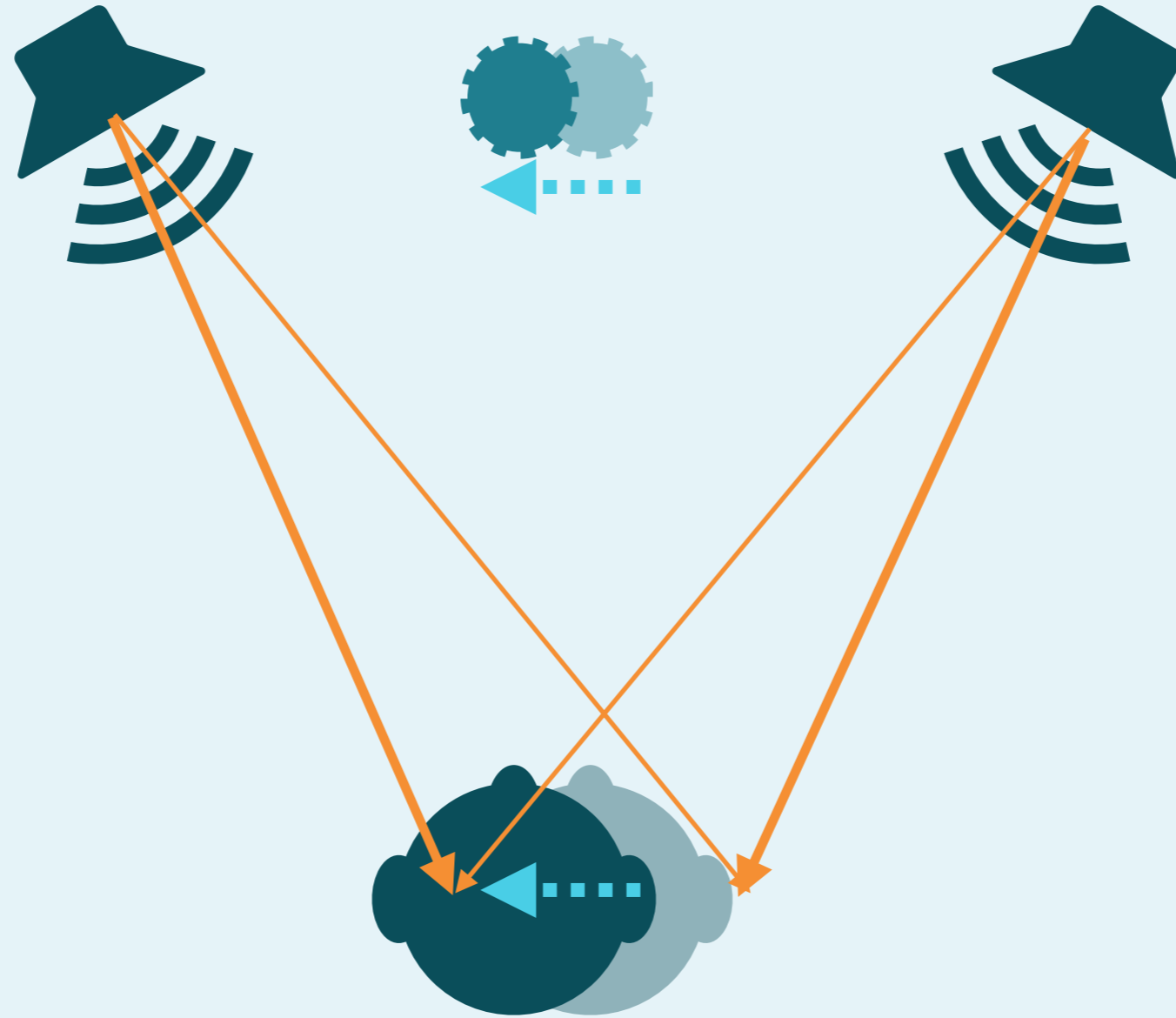
スピーカー再生による音色の変化



左右で同じ音を出力する（センター）と、遅れて届いた音が重なって、コムフィルタがかかる
→ヘッドホン再生とは異なる

2chステレオ再生

音像のズレ



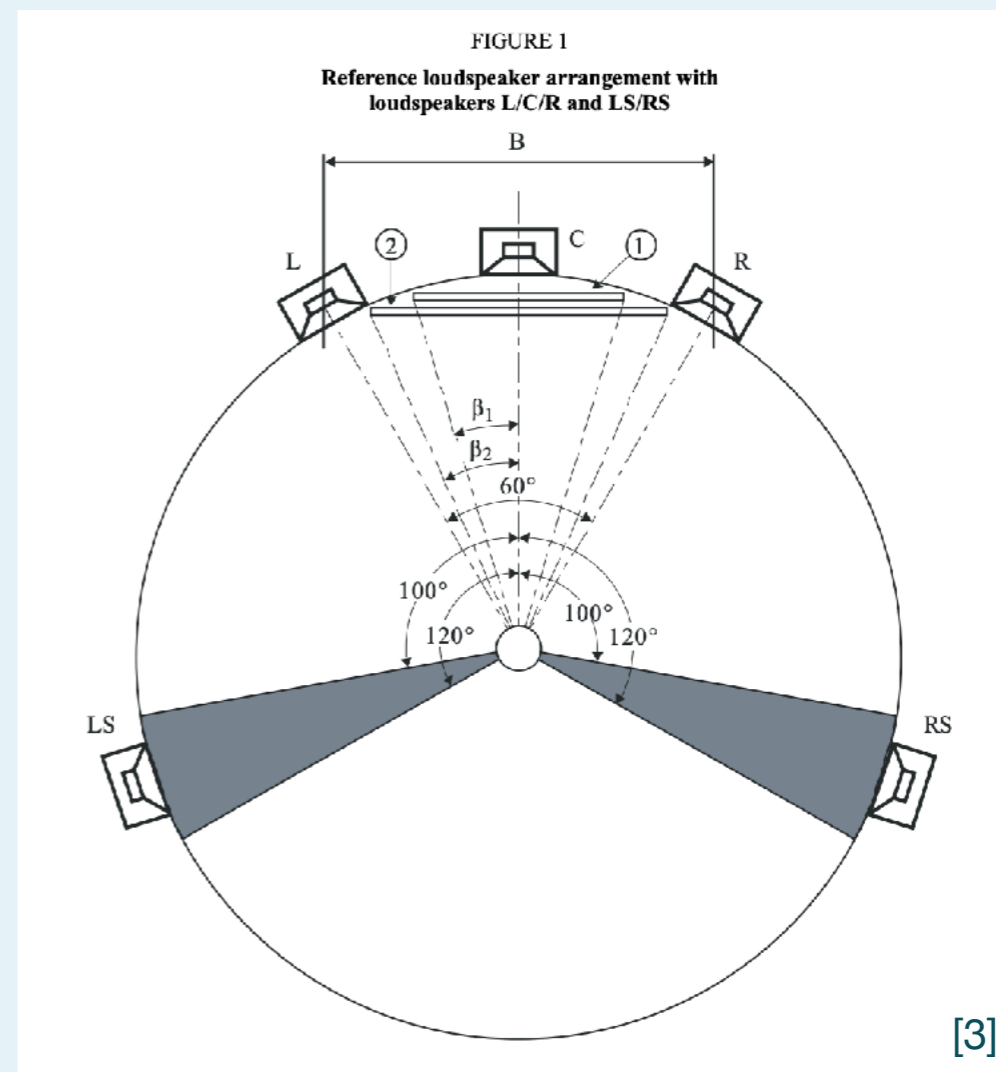
スピーカー同士が等距離でないと音像がずれてしまうし、左右で音質も変わってしまう

先行音効果、ハース効果

- Lチャンネルだけに僅かにディレイを加えることで音像をずらすことができる (ITDの活用)
- うまく使うと、ヘッドホンでもL100%よりさらに左側から鳴っているように聴かせることもできる

チャンネルベース・サラウンド

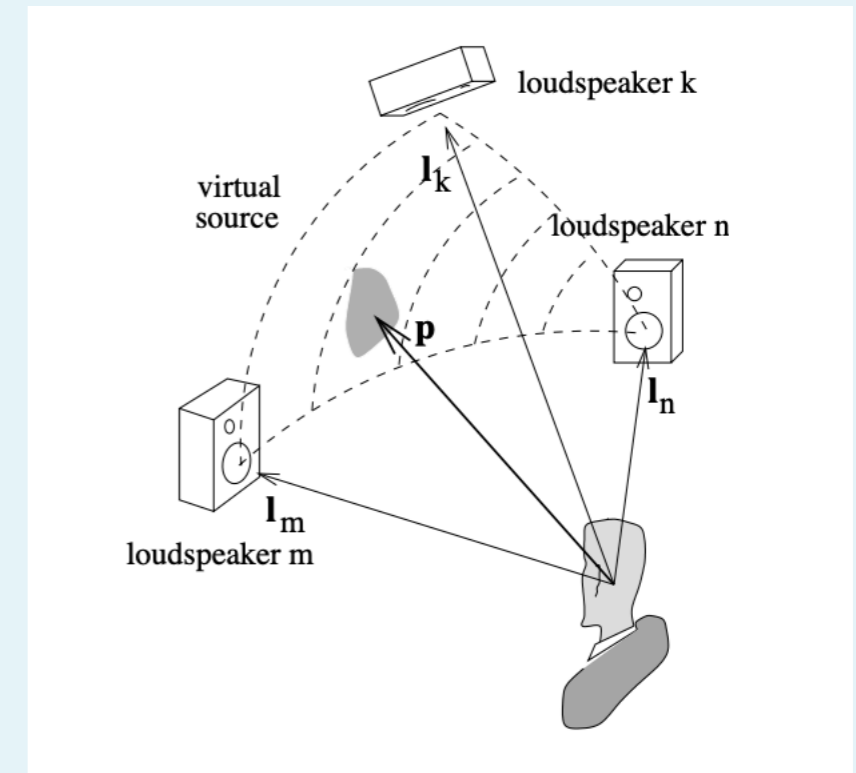
- 5.1ch(.1はLFE：低音専用のチャンネル)
- スイートスポットは中央一か所のみ
- 映画館などではLS,RSを広範囲に広げる配置（ディフューズドサラウンド）にすることもある
- セリフなどは通常Cのみから出すが、L、Rによる中央の定位(ファンタムセンター)を敢えて使うことで音像をぼやけさせることもある



3次元定位

VBAP(Vector Based Amplitude Panning)

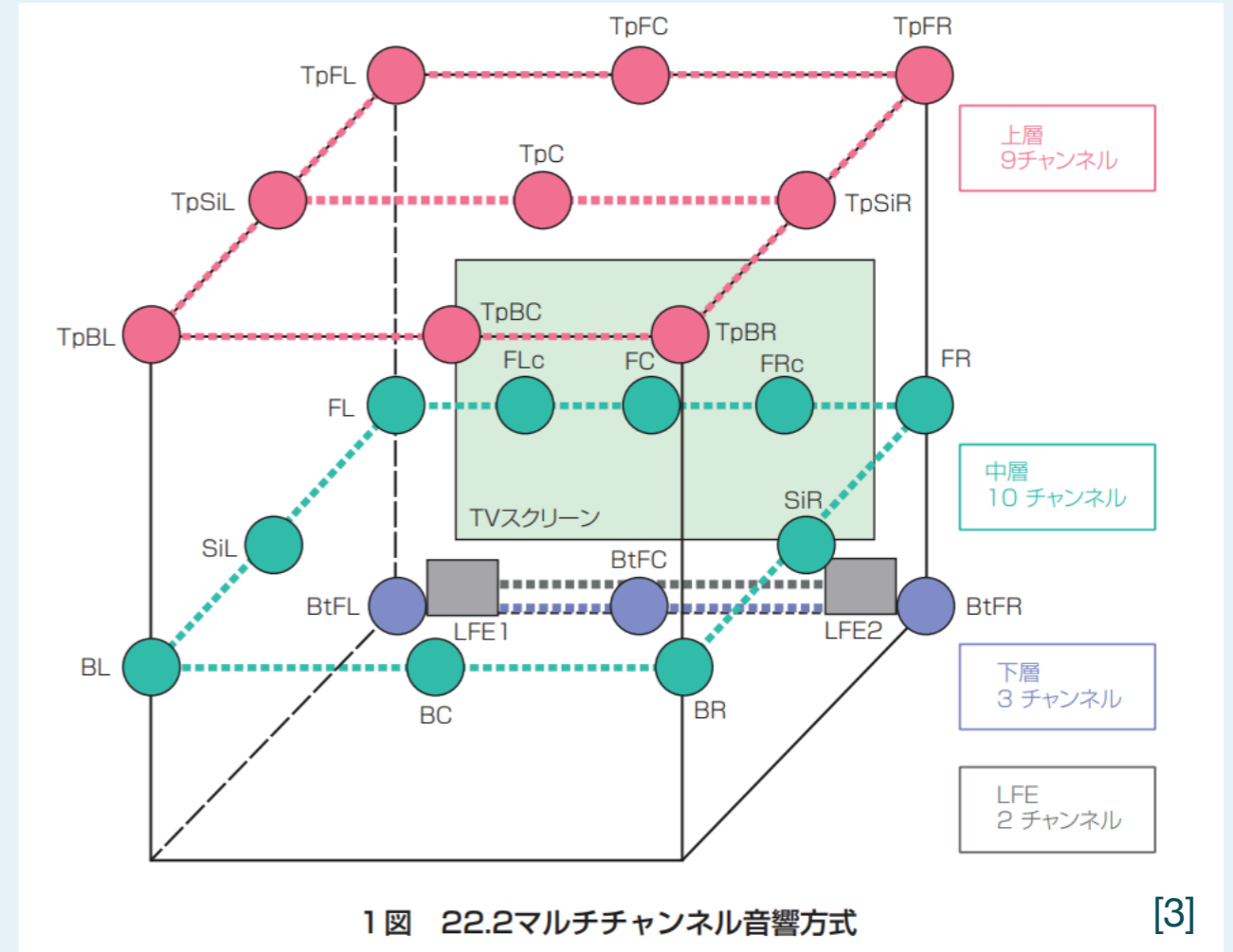
- Level-based Panningを3次元方向に拡張
- 聴取点から等距離（球面上）にスピーカーを配置して三角分割
- 3つのスピーカーへ、それぞれの音量を2乗した合計が1になるように音量を分配



Pulkki, V. (2001). Spatial sound generation and perception by amplitude panning techniques. Helsinki University of Technology.

チャンネルベース・サラウンド

- 22.2 マルチチャンネル
 - 上層9(前後左右斜め:8 頭上1)
 - 中層10(前後左右斜め:8 LとC、RとCの間に計2つ)
 - 下層3(L,C,R)
 - サブウーファー (低音用スピーカー) 2つ
- NHKのスーパーハイビジョン (8K映像)の規格のひとつ



チャンネルベースの問題

- どこまでスピーカーを増やしても基本的にスイートスポットは中心1ヶ所のみ
- 映画館とかだと、リアにスピーカーたくさん配置してあるけど実際はLs、Rsの2chにしか使われてなかったりでもったいない
- 2ch、5.1ch、7.1ch、22.2ch用に全部ミックスを用意するのか？
 - 自動ダウンコンバート(多→小)/アップコンバート(少→多)

オブジェクトベースド・フォーマット

- Dolby Atmosが代表的（最近Apple Musicに入った）
- チャンネルベースのトラック (bedと呼ばれる、7.1.2chなど)に加えて、モノラル音源+位置情報を記録したデータのフォーマット
- 各映画館や、ヘッドホン用などで再生環境に合わせてリアルタイムでレンダリングする
- フォーマットごとに制作しなくてもよいなどのメリット
- 専用のレンダラーが必要になってくる

バイノーラル



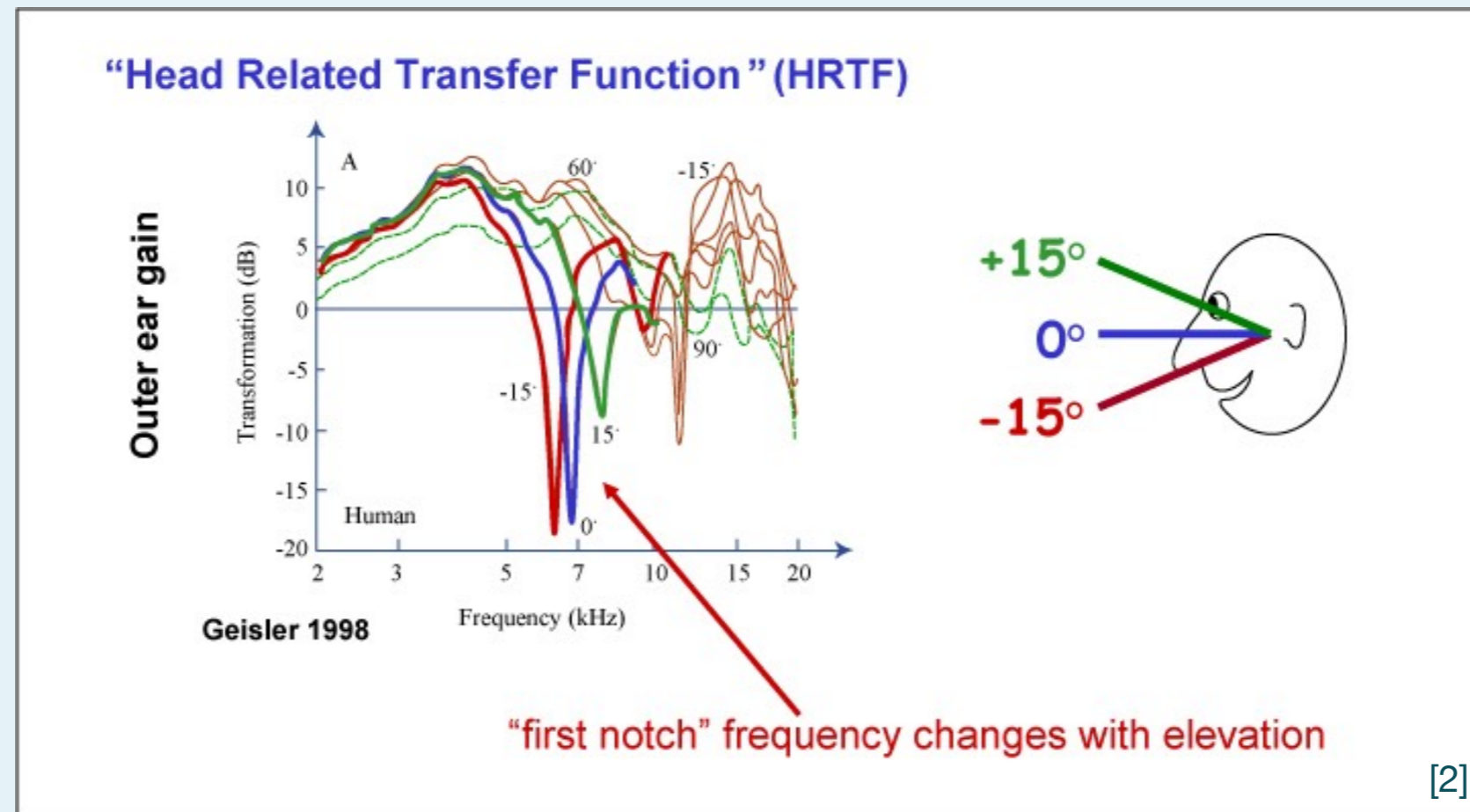
NEUMANN KU100



Roland CS-10EM
<https://www.roland.com/jp/products/cs-10em/>

- 人間が聞いている音を直接録音してしまえばよいのでは？という発想
- バイノーラルマイクロフォン（イヤホンのように耳元に付ける小型マイクや、）
ダミーヘッドマイクロフォン（頭部を模して耳にマイクを埋め込んだもの）を使って録音
- ヘッドフォンで聴いているのに外で音が鳴っているように聞こえる
- 音量変化/ミックス以上の編集が難しい（エフェクトをかけると定位は崩れてしまう）
- 音の定位を動かしたければ実際に音源を動かしながら録音するしかない（後から編集できない）

HRTFによるバーチャルサラウンド



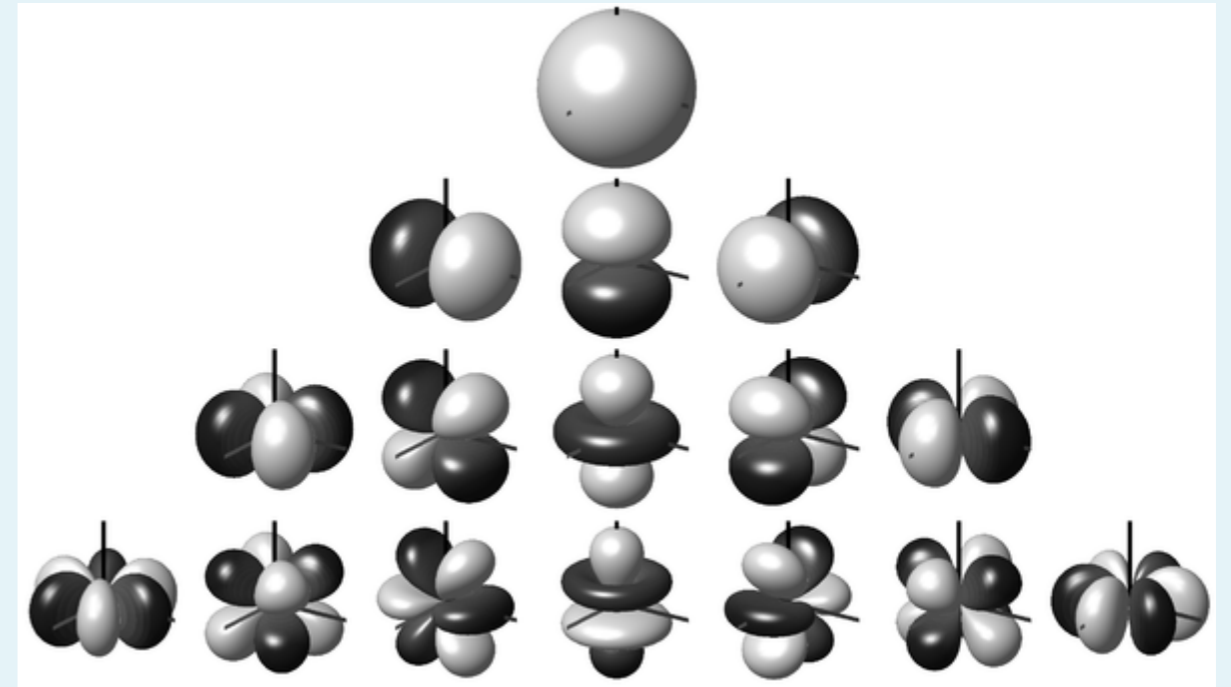
- 音源から耳に届くまでの音色の変化をフィルタとして掛けることでバイノーラルを再現
- このフィルタのことを頭部伝達関数 (Head-Related Transfer Function, HRTF) と呼ぶ
- サンプリングリバーブと同じ要領で、音源位置にスピーカーを置き、測定信号を鳴らしてバイノーラルマイクで録音することでフィルタの特性を記録できる(インパルス応答と呼ばれる)
- 音源位置ごとに測定が必要、動く音源はフィルタを補完しなければならない、HRTFの個人差など問題点は色々

トランスオーラル

- HRTFベースのバーチャルサラウンドをスピーカーで再現したい！
- 実際に使うスピーカーから耳までのHRTFを測定し、特性を打ち消す逆フィルタを作る→スピーカーでヘッドフォンを再現
- さらに、仮想音源→耳のHRTFを適用
- 部屋の残響や、聴取位置の影響などで純粋なバイノーラルほど明確な定位が得られないことも多い
- いわゆるバータイプのサラウンドスピーカーはだいたいこれ

Ambisonic

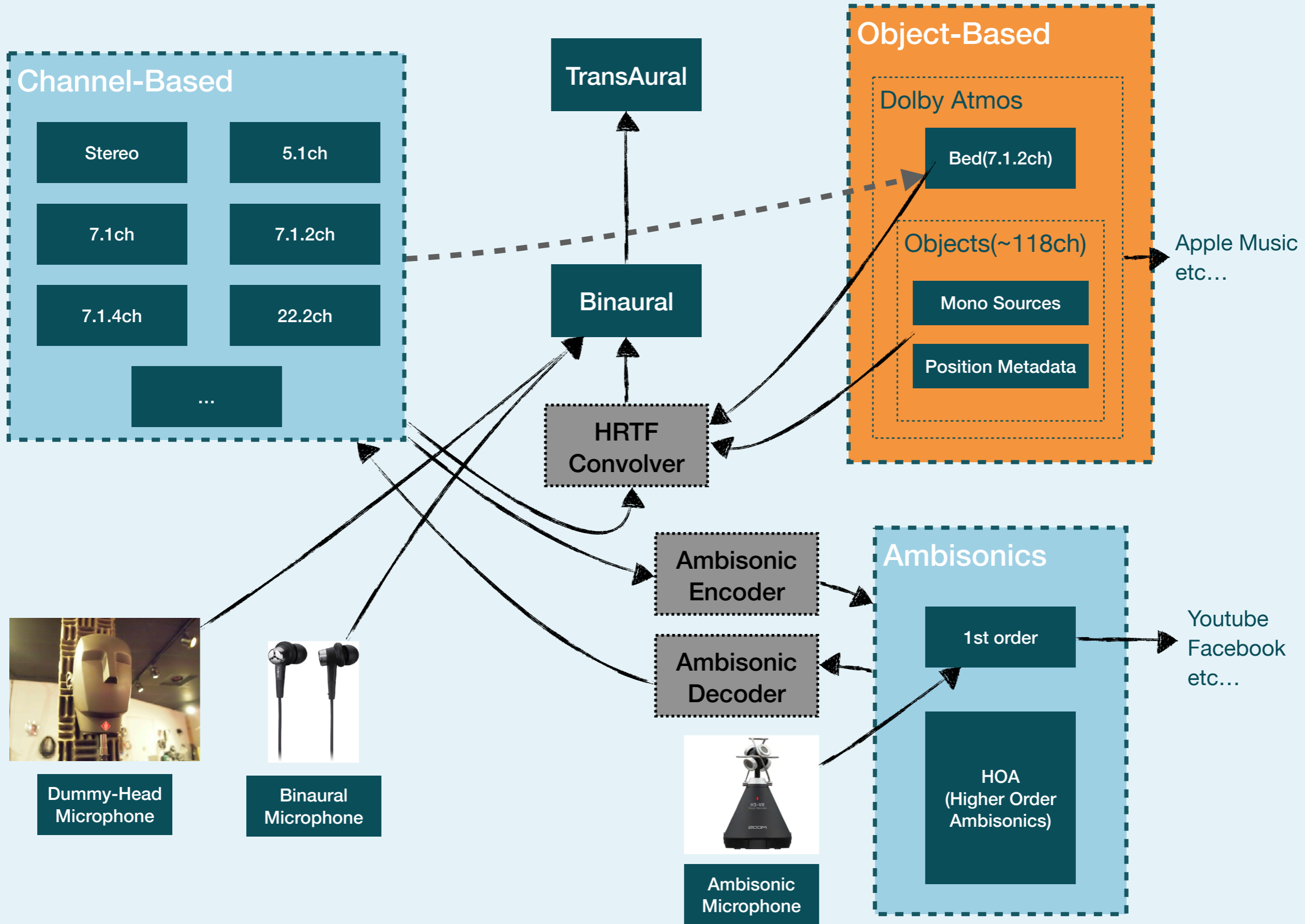
- ある点における球面状の音圧分布を記録するフォーマット
- 次数を挙げていくことで空間解像度が増やせるが、必要なチャンネル数が4,9,16,,,と増えていく
- 1次アンビソニックは専用マイクも普及してきているし、YoutubeやFacebookなどでも対応し始めている
- 異なるスピーカーレイアウト同士を仲介するフォーマットとしても便利



上から0,1,2,3次アンビソニックに必要な音源成分[5]

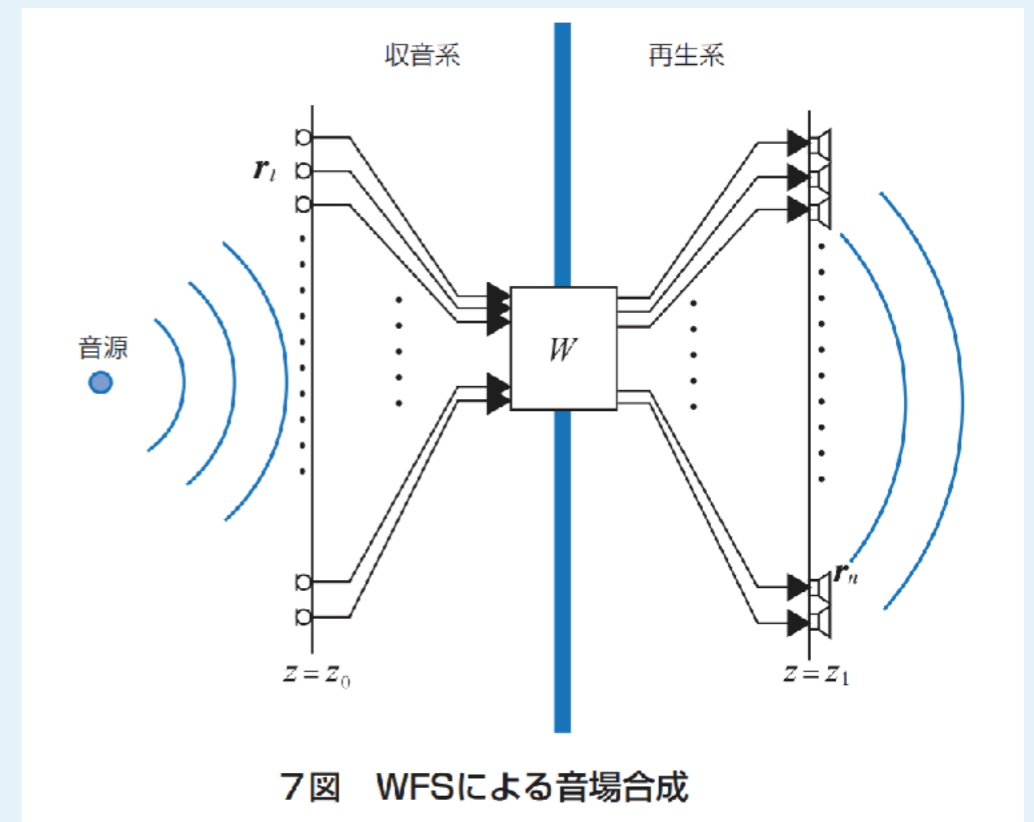


Zoom H3-VR



Wave Field Synthesis

- 空間の物理的な音場をそのまま再現してしまおうという発想
- スピーカーを大量に、一列に敷き詰める
- 聴取者が頭を動かしたり、移動しても大丈夫！
- 演算処理は大変



“物理音響モデルに基づく音響システムの研究動向 - NHK技研 R&D No.126 2011.3” 安藤章男 より

Beyond Immersive Sound?

- チャンネルベースにしても、オブジェクトベースにしても、WFSにしても、空間的な情報量を増やそうと思うとデータ量か演算負荷のどちらかが大きくなる
- 現実空間の完全な再現が理想なのか？
- 全員に同じ再生環境を用意してもらうのも難しい

Image Credits

- [1]By EJ Posselius - Flickr: *_X301277, CC BY-SA 2.0, <https://commons.wikimedia.org/w/index.php?curid=25030807>
- [2]MIT OpenCourseWare-Head Related Transfer Function, CC BY-NC-SA 4.0, <https://www.flickr.com/photos/mitopencourseware/4809307652>
- [3]NHK技研 R&D/No.126 2011.3 濱崎公男 “22.2マルチチャンネル音響方式の標準化動向”
- [4]もんじゃ - Own Work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=50683782>
- [5]By Dr Franz Zotter <zotter@iem.at> - Dr Franz Zotter <zotter@iem.at>, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=30239736>