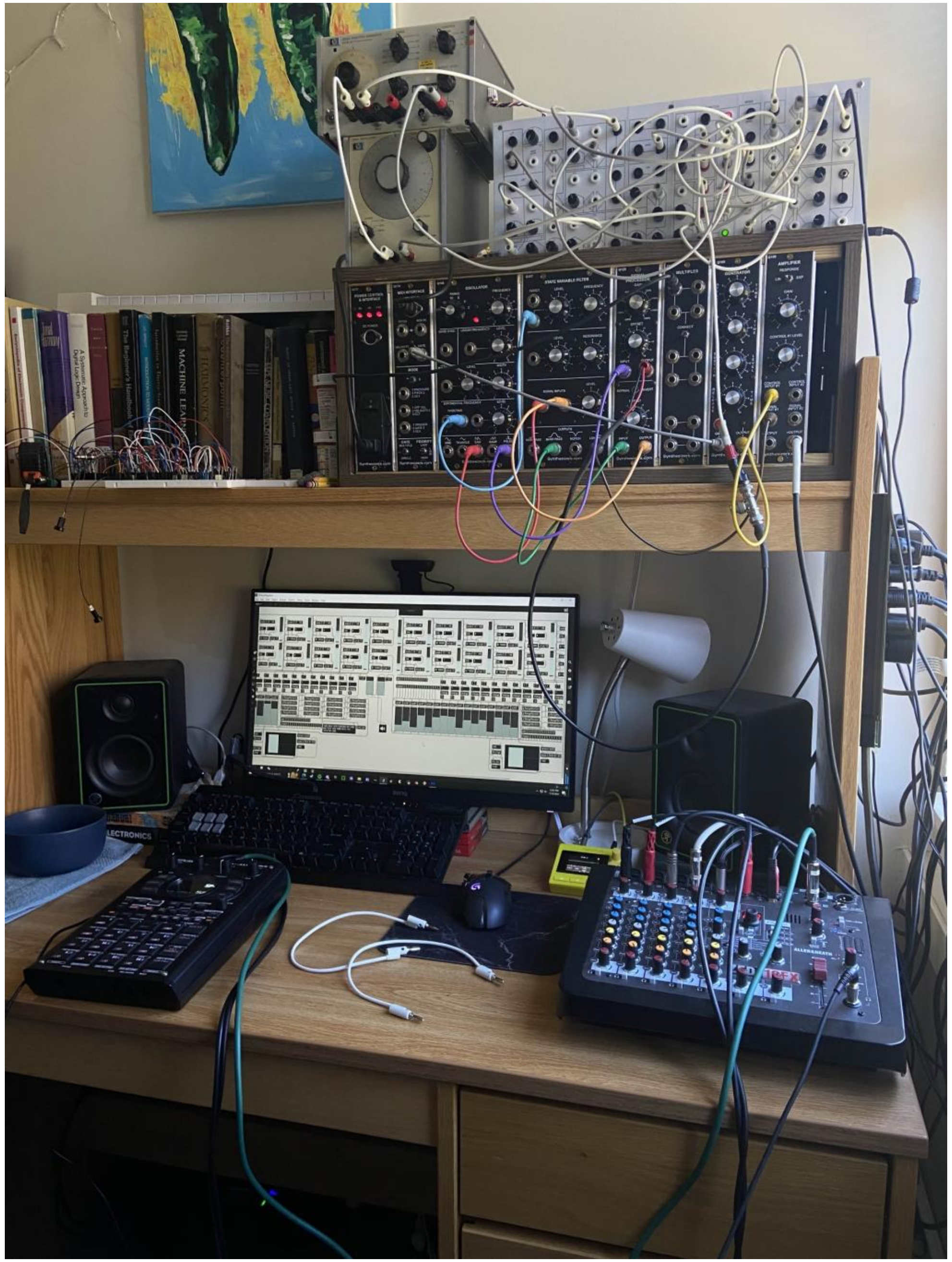
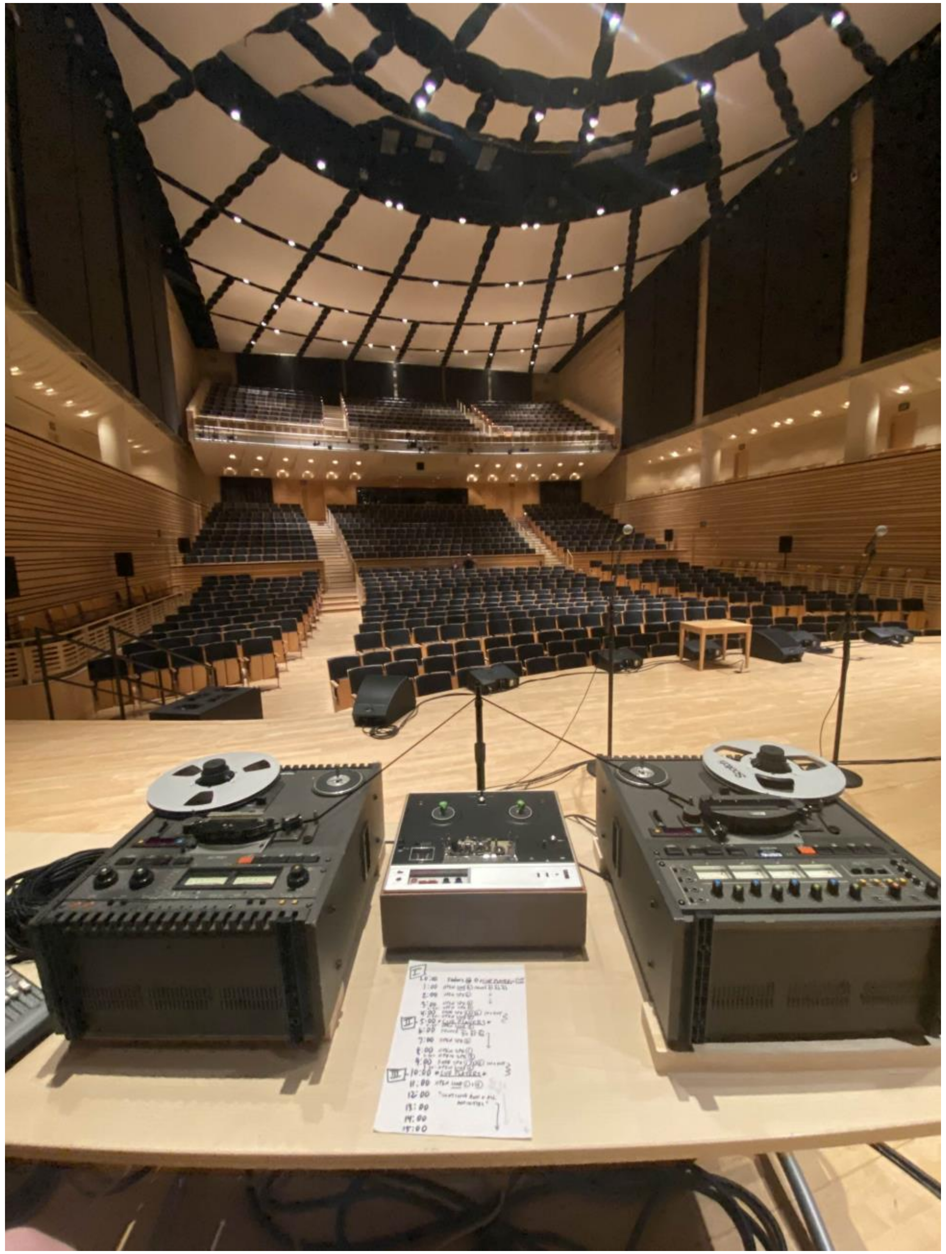
The Psychoacoustics of Spatial Sound


KEYSTEP

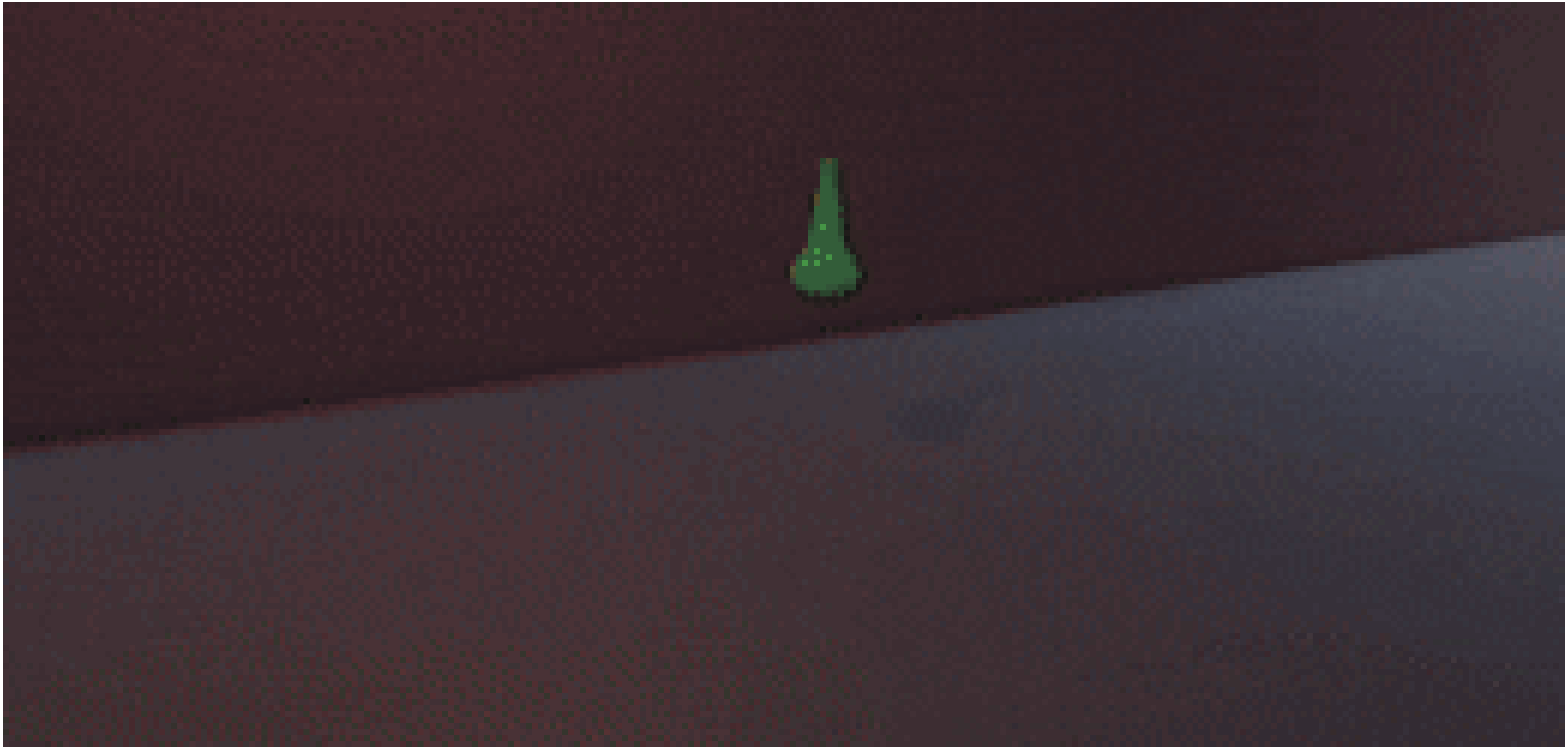

STATE VARIABLE FILTER
AMPLIFIER
ELECTRONICS

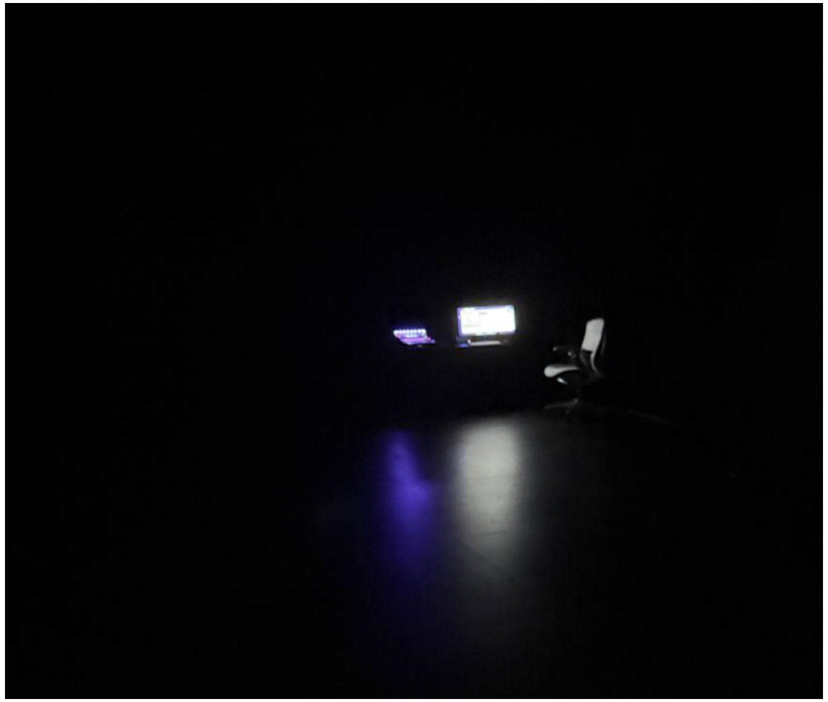


```
@import "../classes/ambisonic_granular_class.ck"
@import "../classes/ambisonic_granular_support.ck"


// instantiation
int nGrans; // number of desired grains (specified in command line)
int mode; // 0 if perpendicular quad setup 1 if angled quad setup
float cursorLocation[2];
string filename; // audio file used as source
string hostname; // address to recieve OSC messages
int port; // port to recieve OSC messages
0 => int device; // where are you getting HID messages
Hid hi; // keyboard
HidMsg msg; // keyboard reader
OscIn mailBox[4]; // recieves OSC messages
OscMsg letterOpener; // OSC reader


int keyArray[8]; // stores the current granulator you want to edit


// check the command line
if( !me.args() || me.args() == 2 ) // take arguments
{
    cherr <= "Input required, format is [audiofile]:[howmany]:[hostname]:[port]" <= IO.newline()
          <= "If no port specified, default to 6449";
    me.exit();
}
else if( me.args() == 3 )
{
    me.arg(0) => filename;
    me.arg(1) => Std.atoi => nGrans;
    me.arg(2) => hostname;
    6449 => port;
}
else if( me.args() == 4 )
{
    me.arg(0) => filename;
    me.arg(1) => Std.atoi => nGrans;
    me.arg(2) => hostname;
    me.arg(3) => Std.atoi => port;
```

```
{
    cherr <= "Input required, format is [audiofile]:[howmany]:[hostnam
          <= "If no port specified, default to 6449";
    me.exit();
}
else if( me.args() == 3 )
{
    me.arg(0) => filename;
    me.arg(1) => Std.atoi => nGrans;
    me.arg(2) => hostname;
    6449 => port;
}
else if( me.args() == 4 )
{
    me.arg(0) => filename;
    me.arg(1) => Std.atoi => nGrans;
    me.arg(2) => hostname;
    me.arg(3) => Std.atoi => port;
```

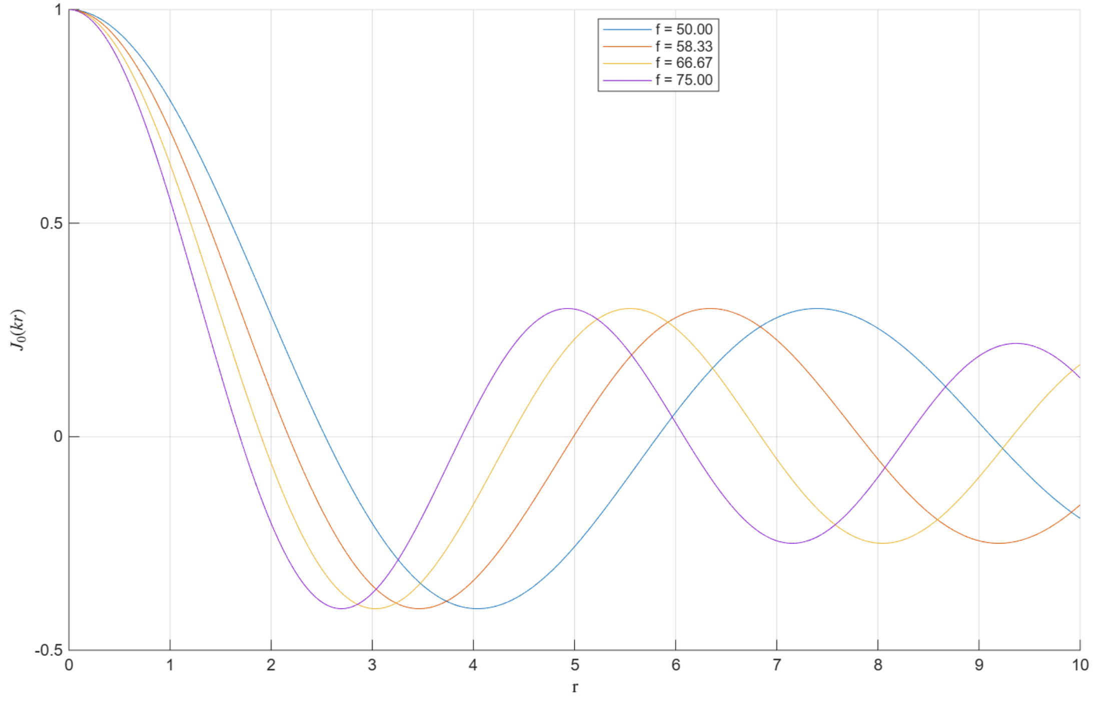
```
}


// state you're address
for( auto x : mailBox ) // set your port for OSC
{
    port => x.port; // set port
}


AmbiGranularSupport assistant;
AmbiGranulator grain(filename)[nGrans];
Encode3 enc[nGrans];
WvOut record[enc[0].channels()]; // record
JCRev reverb[nGrans]; // reverb
BFormat3 bform(1.0/nGrans);
Gain sum(1.0/nGrans)[nGrans];


// all the OSC addresses
mailBox[0].addAddress("/keypresses/down");
mailBox[1].addAddress("/keypresses/up");
mailBox[2].addAddress("/trackpad/x");
mailBox[3].addAddress("/trackpad/y");
```

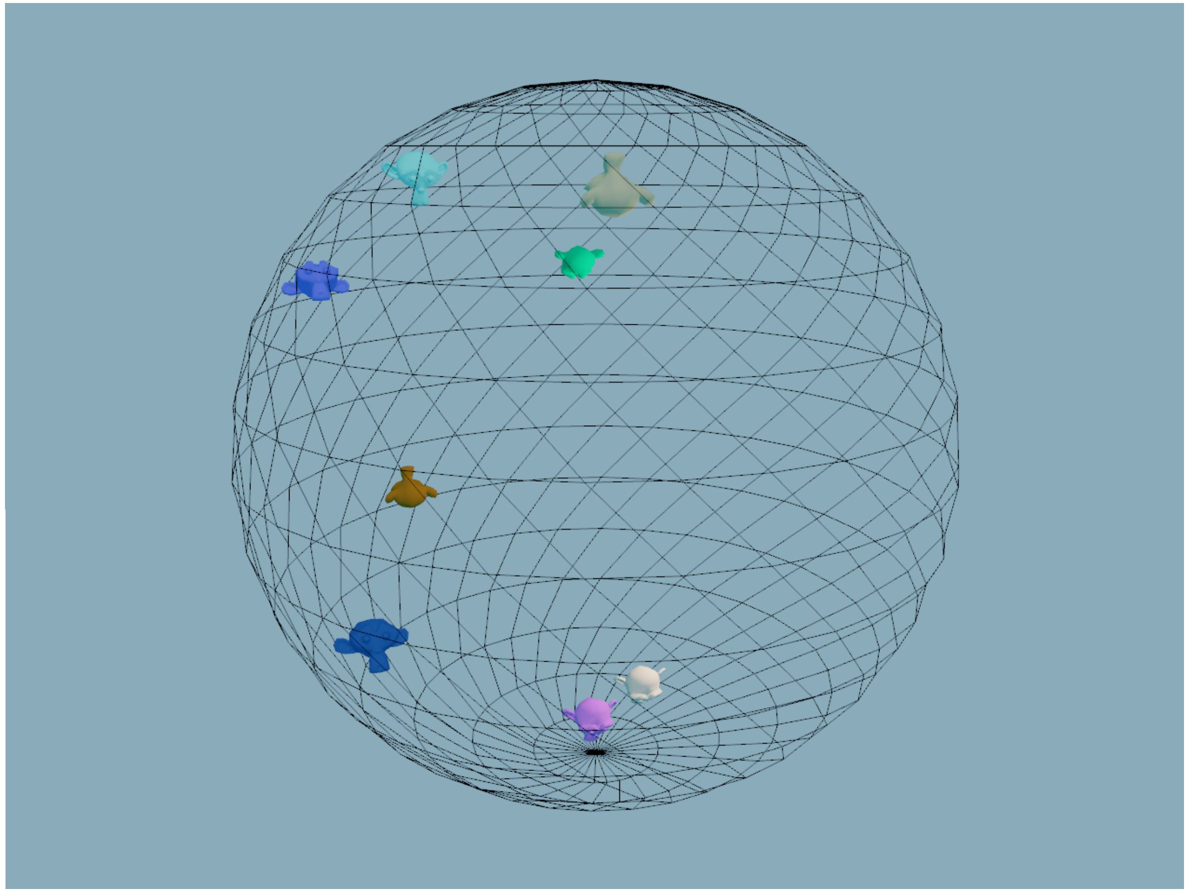




$$p(\vec{r}) = \sum_{n=0}^{\infty} i^n j_n(kr) \sum_{m=-n}^{n} B_n^m Y_n^m(\theta, \phi)$$

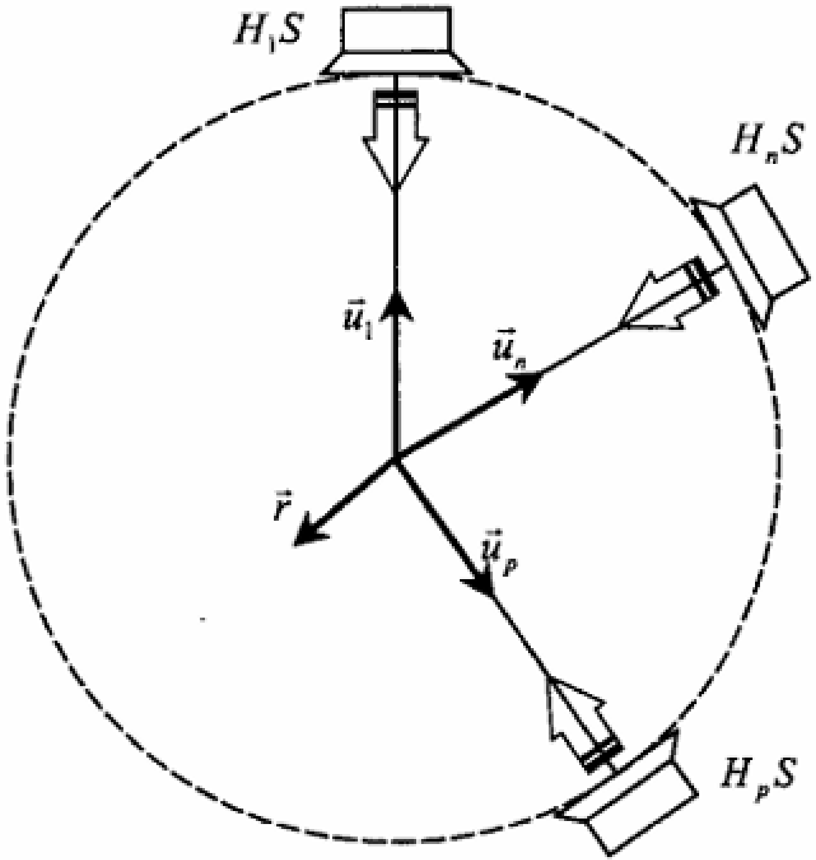
$$\boldsymbol{D}_{\sigma,\delta} = \begin{bmatrix} g_{1,1}(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) & g_{1,2}(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) & \cdots & g_{1,\delta}(\boldsymbol{\theta}_1, \boldsymbol{\phi}_1) \\ g_{2,1}(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2) & g_{2,2}(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2) & \cdots & g_{2,\delta}(\boldsymbol{\theta}_2, \boldsymbol{\phi}_2) \\ \vdots & \vdots & \ddots & \vdots \\ g_{\sigma,1}(\boldsymbol{\theta}_\sigma, \boldsymbol{\phi}_\sigma) & g_{\sigma,2}(\boldsymbol{\theta}_\sigma, \boldsymbol{\phi}_\sigma) & \cdots & g_{\sigma,\delta}(\boldsymbol{\theta}_\sigma, \boldsymbol{\phi}_\sigma) \end{bmatrix}$$





$$D(\theta, \phi) = \sum_{n}^{N} \sum_{m=-n}^{n} Y_n^m(\theta, \phi) B_n^m$$

What is spatial sound?



Reproduced from *ACOUSTIC PROPERTIES AND PERCEPTIVE IMPLICATIONS OF STEREOPHONIC PHENOMENA* by Daniel et al.

A wave…

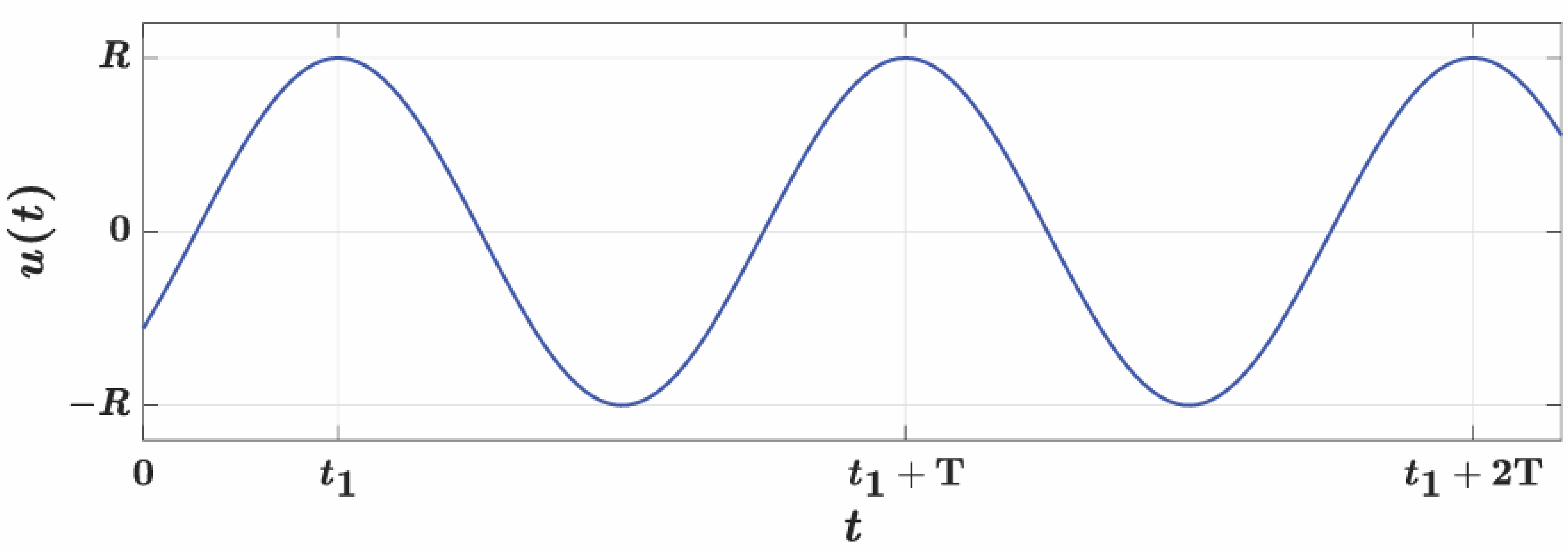
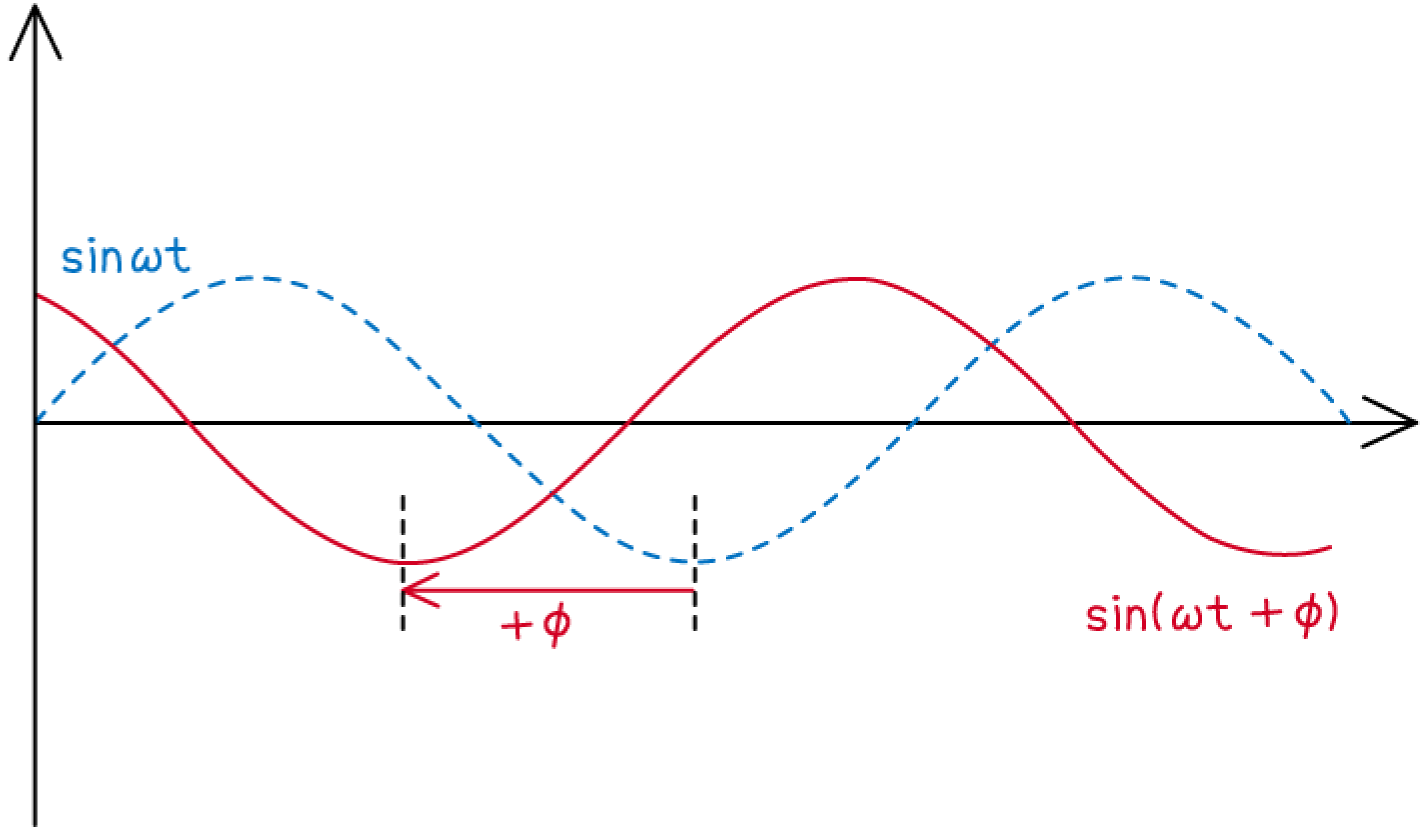


Figure 3.5. *Simple harmonic motion solution for Example 1.*

Reproduced from Mark Holmes' *Introduction to Differential Equations*

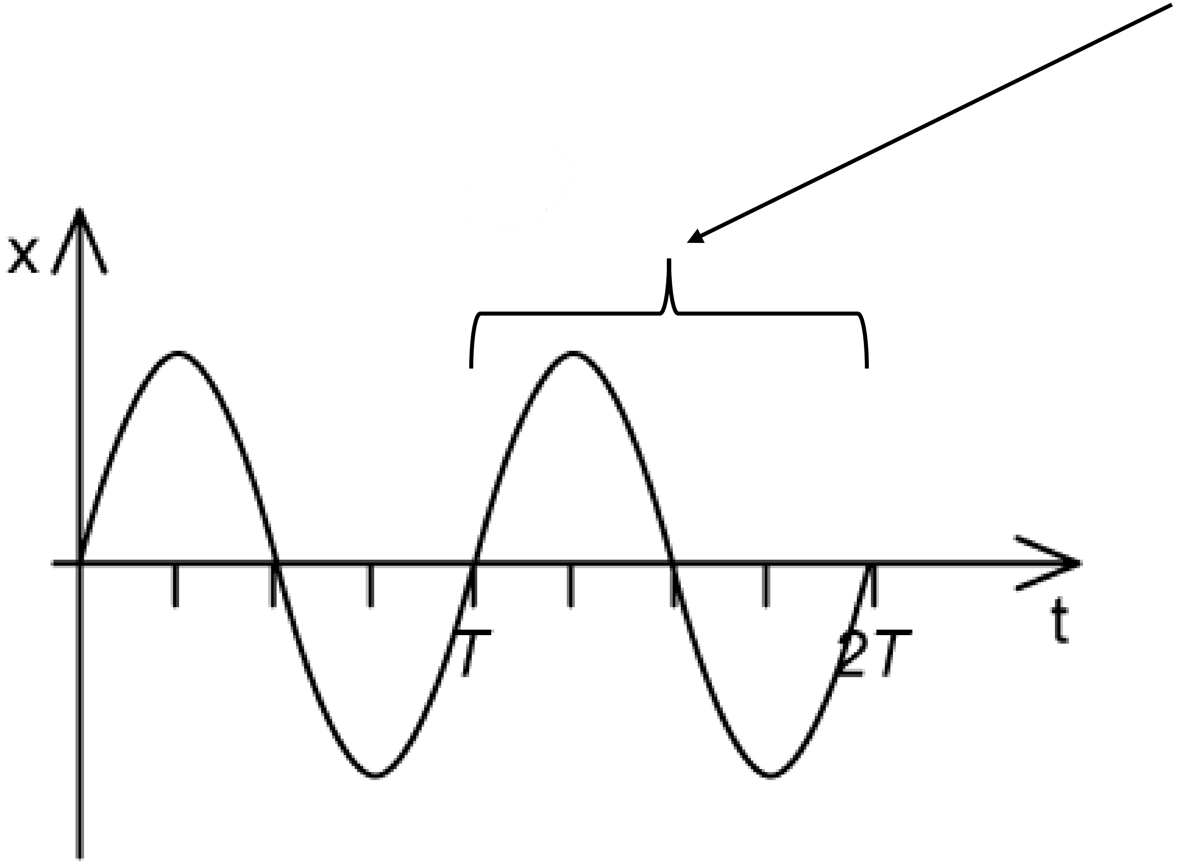
Important components of waves are…



Reproduced from Phase Angles in Simple Harmonic Motion (SHM) | DP IB Physics Revision Notes 2023

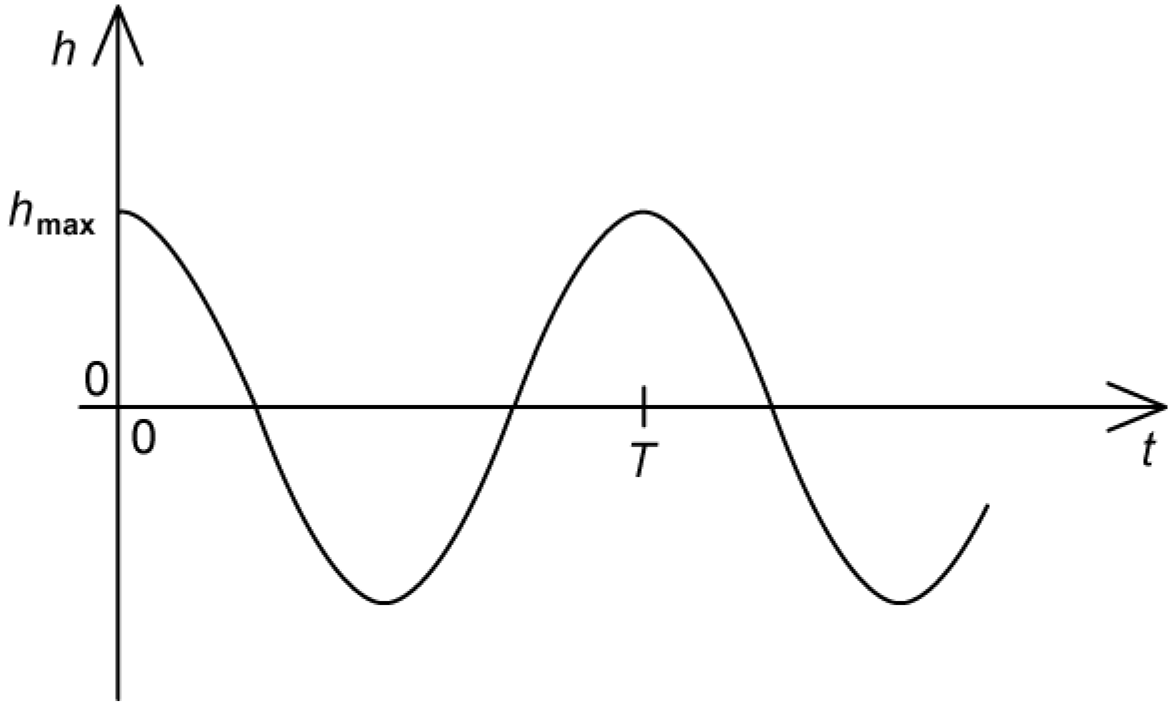
Frequency,
Phase,
Amplitude

But what is this… physically?



Reproduced from Phase Angles in Simple Harmonic Motion (SHM) | DP IB Physics Revision Notes 2023

Wavelength is the *physical length* of wave, in units such as kilometres, meters, centimetres, etc.



Reproduced from Phase Angles in Simple Harmonic Motion (SHM) | DP IB Physics Revision Notes 2023

Wavelength can be calculated by:

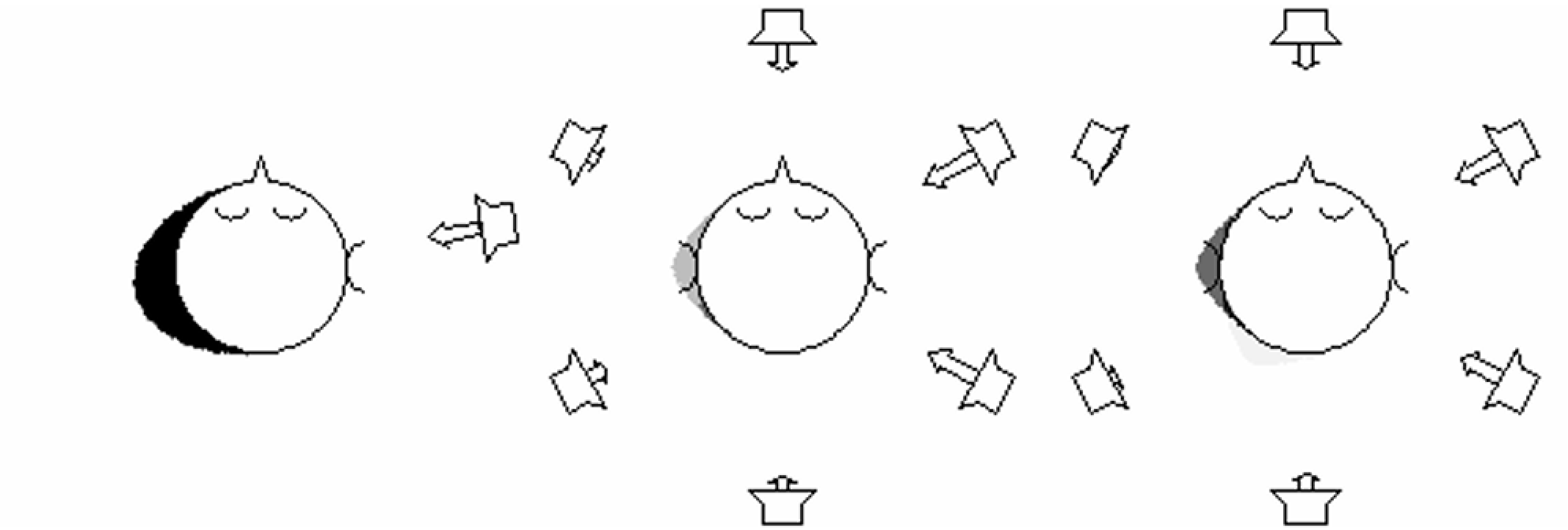
$$\omega = \frac{c}{f}$$

Where c is the speed of sound and f is the frequency of the wave

What if the wavelength of a wave was equal to the distance between the ears?

On average, the ears are roughly 22cm apart.

A frequency of around 1500Hz would be required to create a wavelength equal to the distance between the ears.



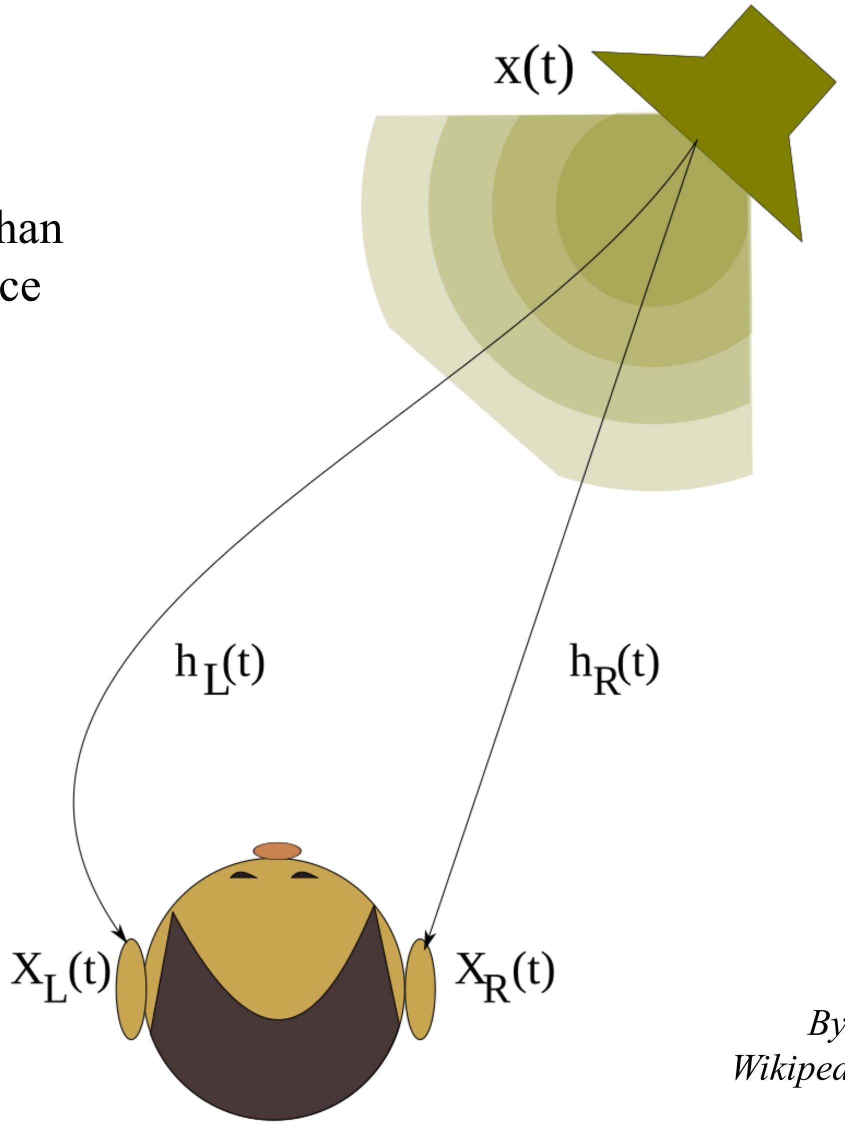
Reproduced from *Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions* by Jérôme Daniel

Surely that’s important right?

Well, if a periodic wave is arriving at the listener, and its wavelength is longer than the distance between the ears. We could probably tell the difference, right?

Such a thing has been proven.

Because when a wavelength is longer than the head, we can *easily* tell the difference from ear to ear.



They’re called…

Interaural Time Differences or ITDs

This is because we perceive the difference between the ears as a *delay*.

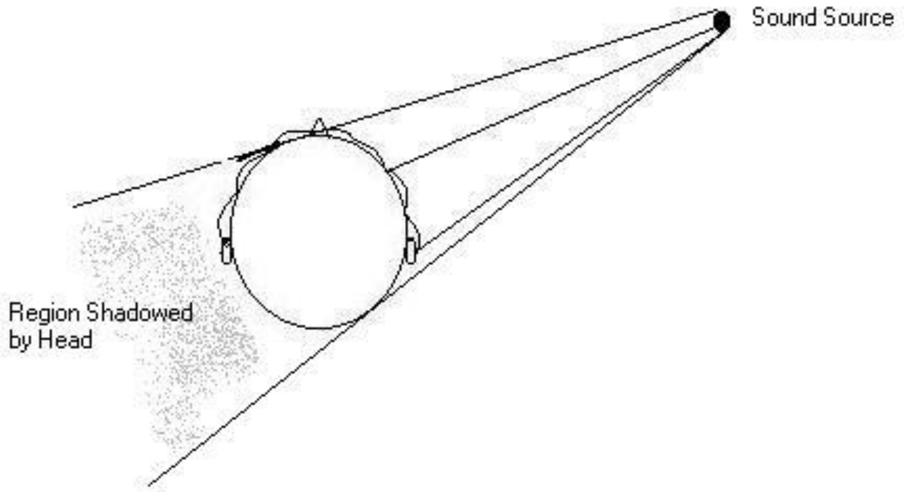
But what about when the wavelength is shorter than the head?

Why can I still tell the location of a high-pitched sound?

There's still a phase difference, but it's hard to make any sense of it.

Why? Because if the wavelength is shorter than the head, the ears may be receiving totally different cycles.

Sound decays over time and space right? You can’t hear a sound from a mile away. This seems to happen especially with higher frequencies.



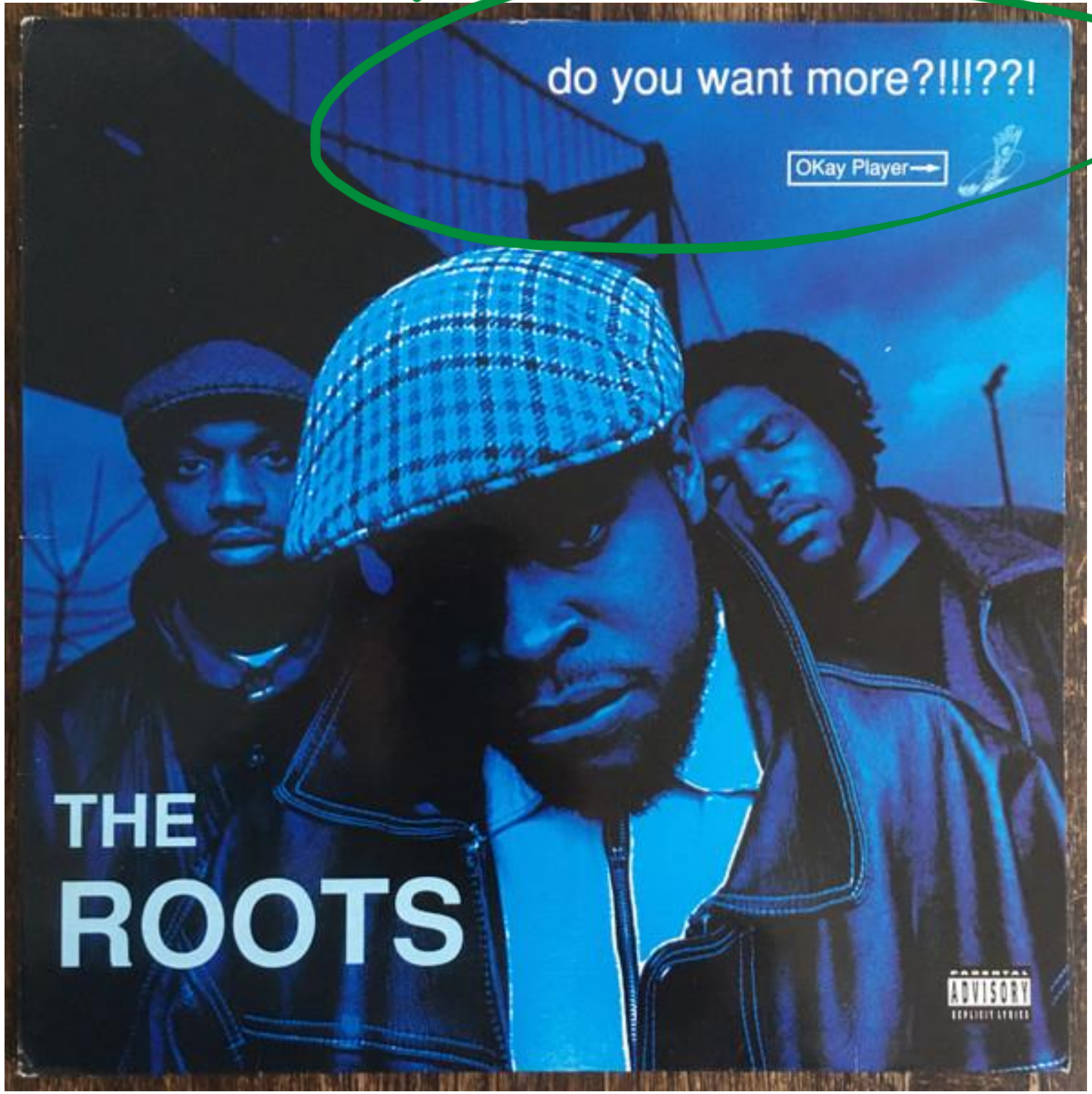
It’s a non-linear decay, meaning it decays increasingly quicker the further the distance.

&

The head will contribute to the decay of a wave as it passes through it.

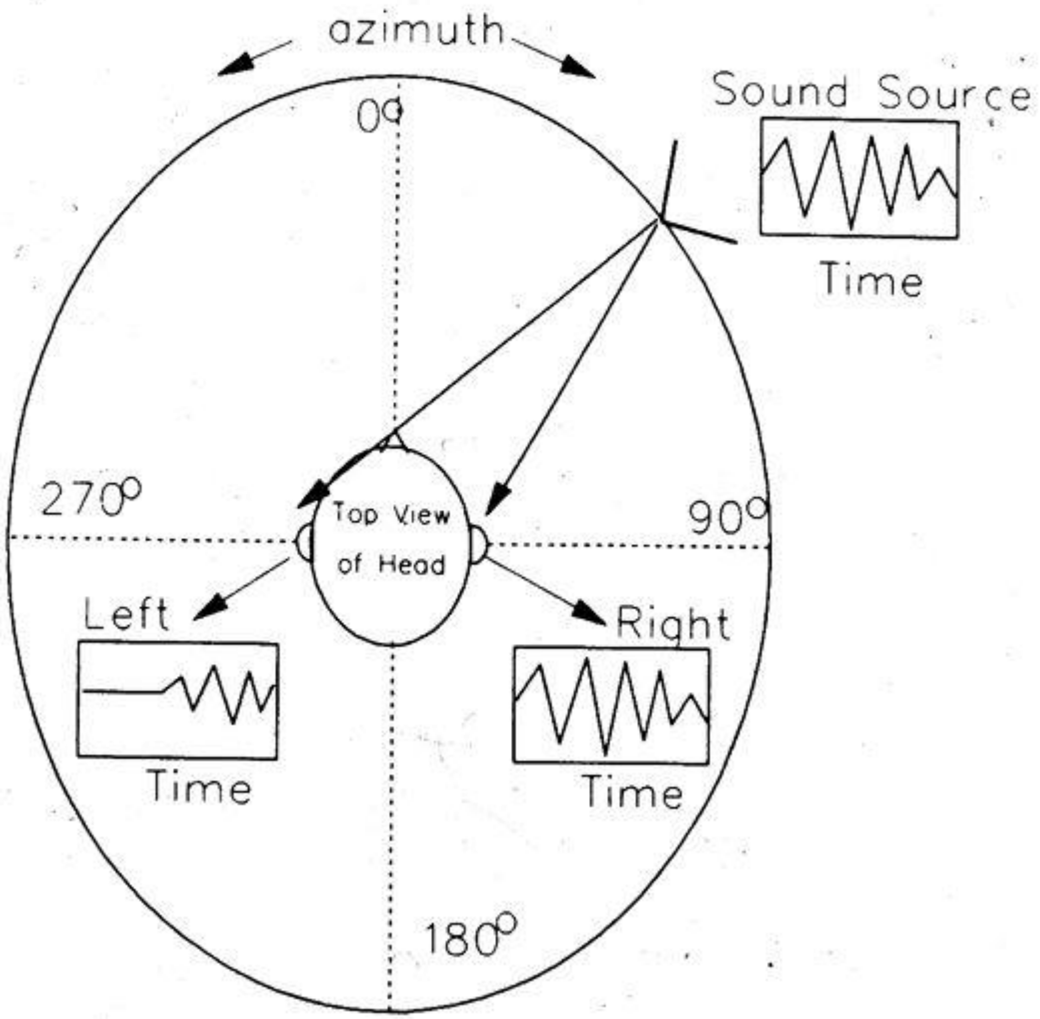
Reproduced from [ANATOMY AND PHYSIOLOGY OF EXTERNAL EAR - Hearing Aid Center in Kolkata]

Two ways of localizing, ITDs and ILDs,
both work well for certain frequency bands.
But is that all?


do you want more?!!!??!
OKay Player
THE
ROOTS
PARENTAL
ADVISORY
EXPLICIT LYRICS

There are many other ways that humans localize sound.

- Visual cues
- Contextual cues
- Scene analysis
- & more

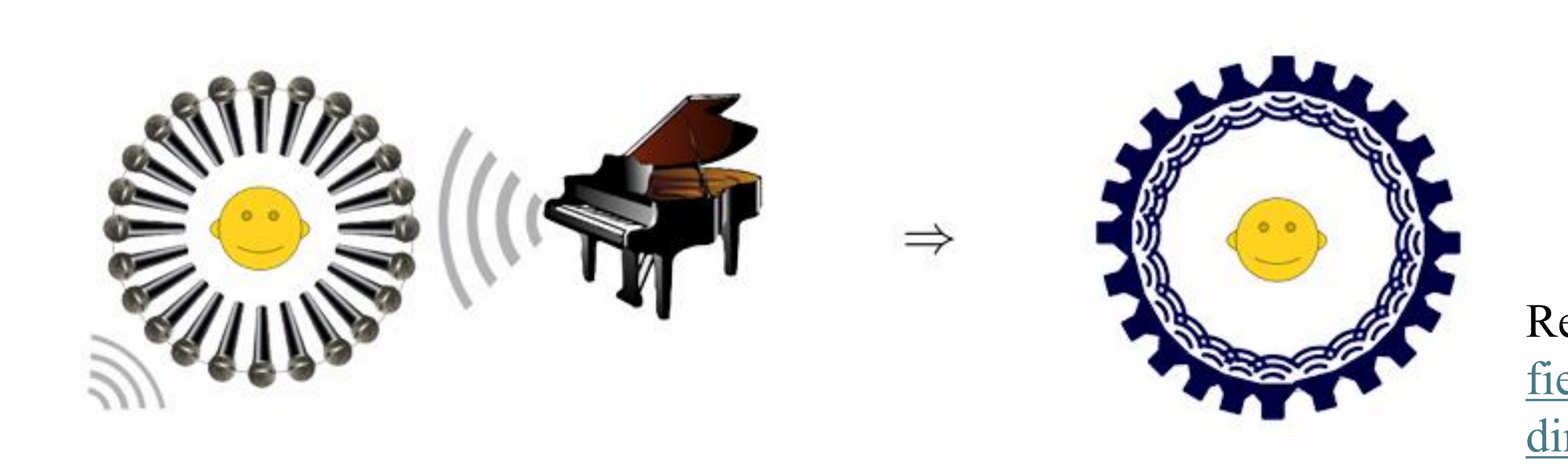


Reproduced from [Perception Lecture Notes: Auditory Pathways and Sound Localization]

Why do we care?

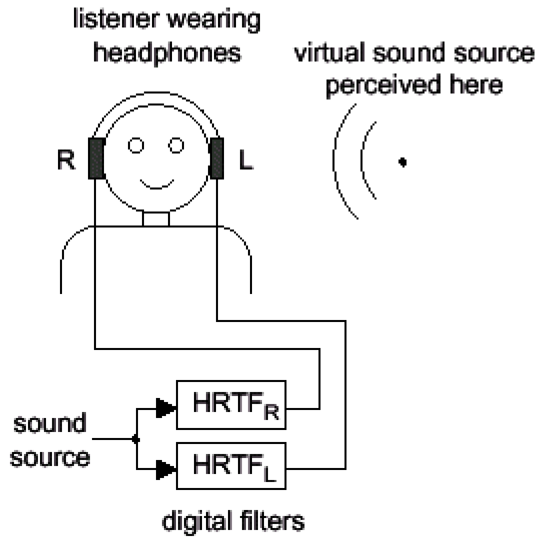
We can exploit this knowledge when creating systems that look to create immersive sound.

Systems like Wavefield Synthesis, Ambisonics and Binaural playback utilize these factors to create immersive experiences.



Reproduced from Wave field synthesis: The next dimension

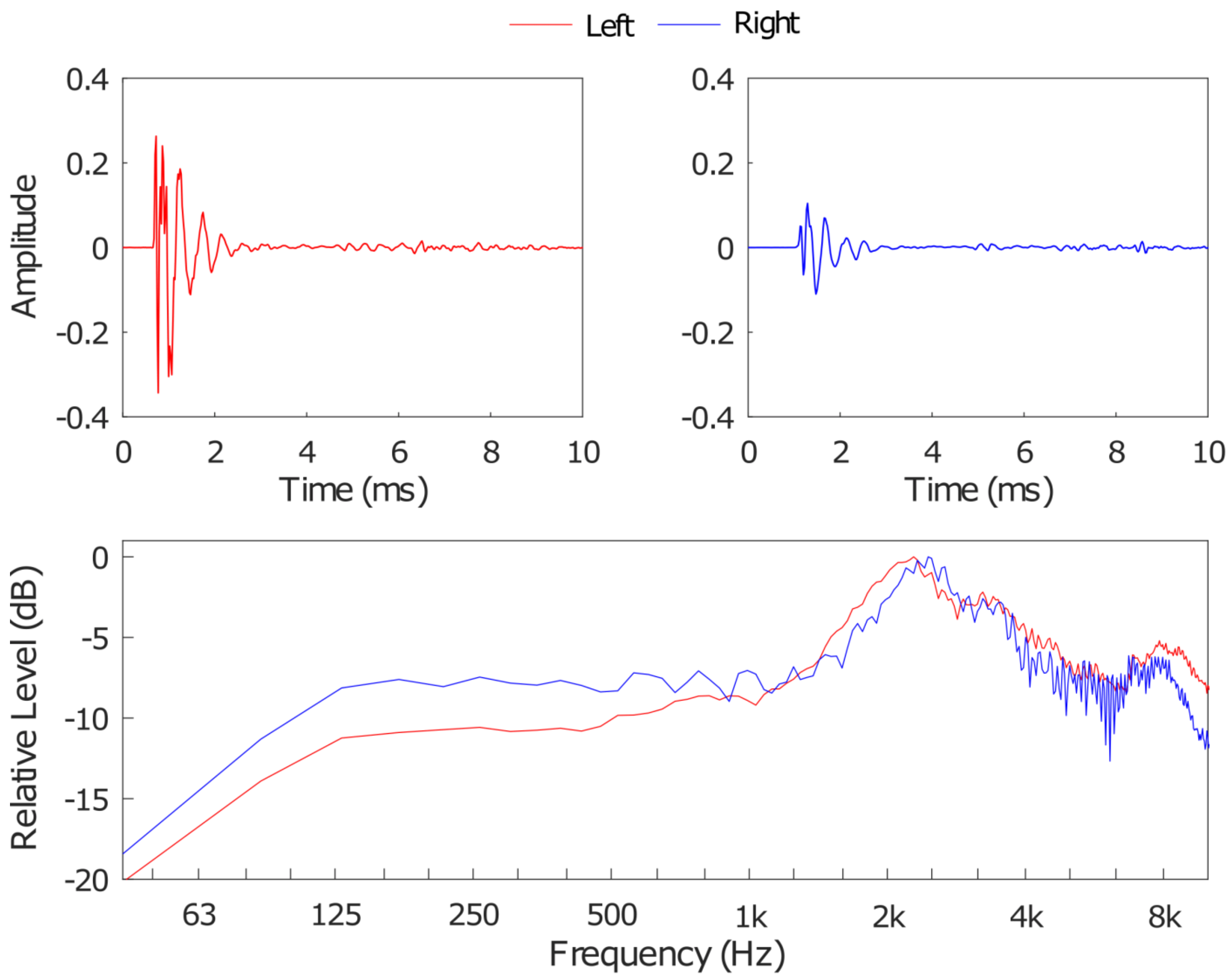
A very common example of this is binaural rendering…



Reproduced from 3D Audio and Acoustic Environment Modeling. – HeadWize Memorial

The Head Related Transfer Function (HRTF)…

We can make a filter that does this!



Reproduced from Head-Related Transfer Functions | VOCAL Technologies

Another application…

Ambisonics uses multiple loudspeakers to achieve spatial sound. An effective method of ambisonic playback is Dual Band Decoding (DBD). Which filters the sound field and optimizes ITD and ILD cues.

Reproduced from [IEM (Institute 17 Electronic Music and Acoustics)]

Lots of a cool stuff, but what are the downsides?

Well, a lot of this is “human specific”, both ILD and ITDs depend on:

- Height of the head
- Width of the head
- Amount of earwax
- Health of the eardrums
- Etc.

And because this is psychological, there are additional factors such as

- Past experiences
- Physical health
- Mental health
- Etc.

To wrap things up…

Some basic methods of sound localization include:

- Interaural Level Differences (ILD)
- Interaural Time Differences (ITD)

Which are more effective based on:

- Frequency range
- Head size
- Distance from original sound

We can use these as a model for making immersive sound, such as:

- Binaural playback
- Multi-loudspeaker playback
    - Ambisonics
    - Atmos
    - Wavefield Synthesis

Q & A