The Physics and Psychophysics of Surround Recording

David Griesinger

Lexicon

dgriesinger@lexicon.com

www.world.std.com/~griesngr

Main Message:

- The recording venue is CRITICAL!!!
  - Large (>2000 seat) concert halls can make stunningly beautiful recordings.
    - A wide variety of techniques can achieve satisfactory results.
  - A technique that works well in a large concert hall will probably NOT work well in a hall with 1200 seats.
- It is the job of the engineer to make a stunningly beautiful recordings in the hall that happens to be available.
  - Working to a world-class standard in a small space takes both science and art.

Science and Art

- This talk attempts to demonstrate the science behind the art.
  - We call ourselves engineers. We need to use our scientific skills to perfect what we do.
  - I will show with musical examples which techniques work best in a particular venue, and why.

What is our goal?

- To do justice to the music!
  - Through excellent clarity and balance)
  - To recreate the performance space in the listening room!
  - To make the listening area as large as possible.
- How do we do it?
  - Through Localization – the recreation of the spatial position of musical lines.
  - Through Envelopment – the recreation of the original space throughout the listening room.

The Venue is Critical

- We will look closely at four venues:
  - Boston Symphony Hall
    - 2631 seats, 662,000ft^3, 18700m^3 RT 1.9s
  - Jordan Hall, Boston
    - 1019 seats, ~200,000ft^3, 5,600m^3, RT ~1.5s (occupied)
  - Swedenborg Chapel, Cambridge MA
    - ~200 seats, 50,000ft^3+, 1,450m^3 RT ~1.3s (empty)
  - Sonic Temple Studio, Roslindale MA
    - 44,000ft^3, 1,250m^3, RT 1.4s, with blankets, 0.9s

Boston Symphony Orchestra in Symphony Hall

Boston Cantata Singers in Symphony Hall. March 17, 2002

Jordan Hall, Boston

Boston Cantata Singers in Jordan Hall

Swedenborg Chapel, Cambridge

Oriana Consort in Swedenborg Chapel

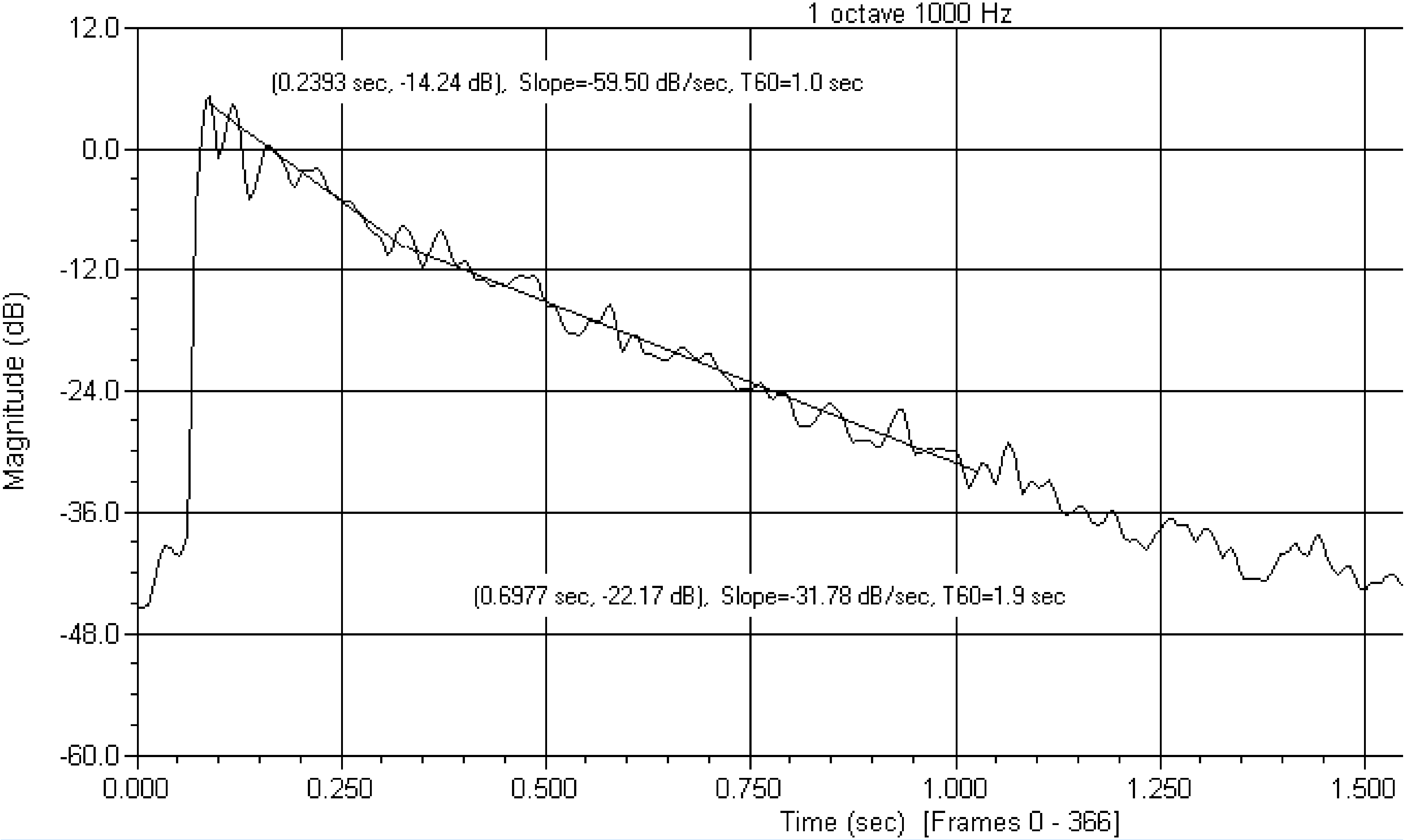
Revels Chorus in the Sonic Temple

Boston Symphony Hall

Boston Symphony Hall

- 2631 seats, 662,000ft^3, 18700m^3, RT 1.9s
  - It's enormous!
  - One of the greatest concert halls in the world – maybe the best.
  - Recording here is almost too easy!
  - Working here is a rare privilege
    - Sufficiently rare I do not do it. (It's a union shop.)
  - The recording in this talk is courtesy of Alan McClellan of WGBH Boston. (Mixed from 16 tracks by the presenter)
  - Reverb Radius is >20' (>6.6m) even on stage.
  - The stage house is enormous. With the orchestra in place, stage house RT ~1 sec

Boston Symphony Hall, occupied, stage to front of balcony, 1000Hz



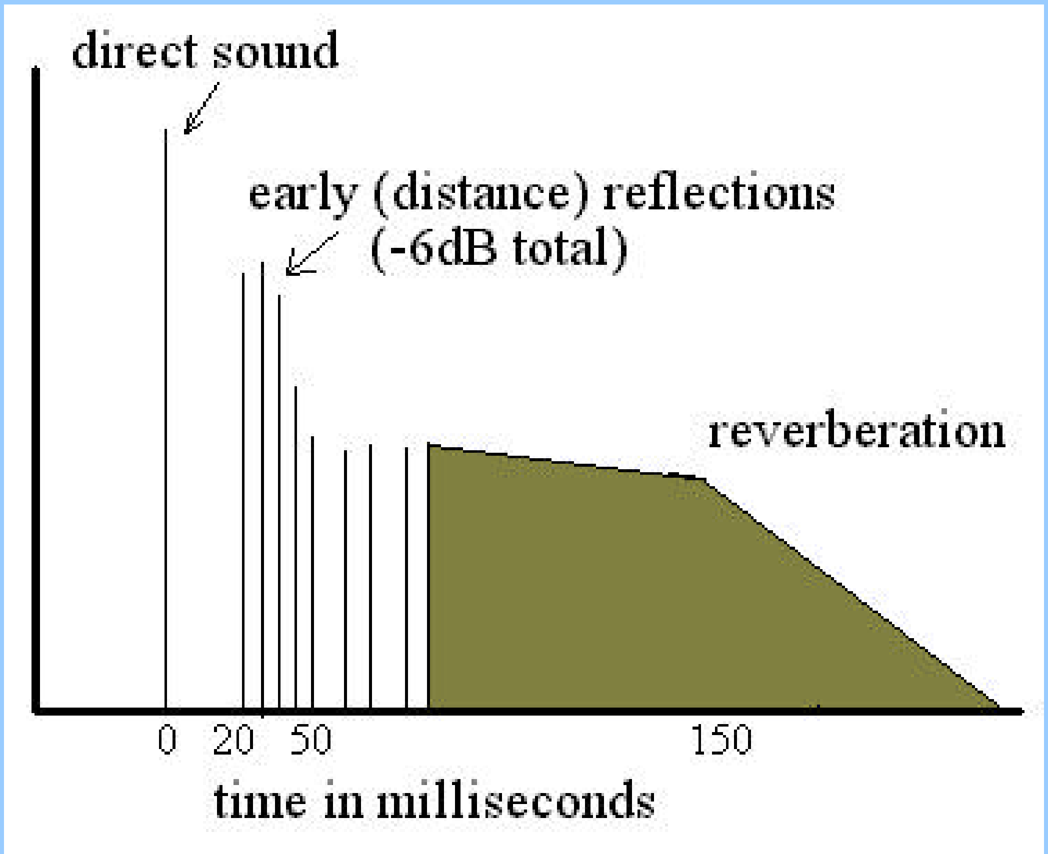
Why is the impulse response relevant?

- Because the early decay (from the stage) is short enough to get out of the way before it muddies he sound.
- And the late decay (from the hall) is long enough to provide envelopment.

The Ideal Reverberation

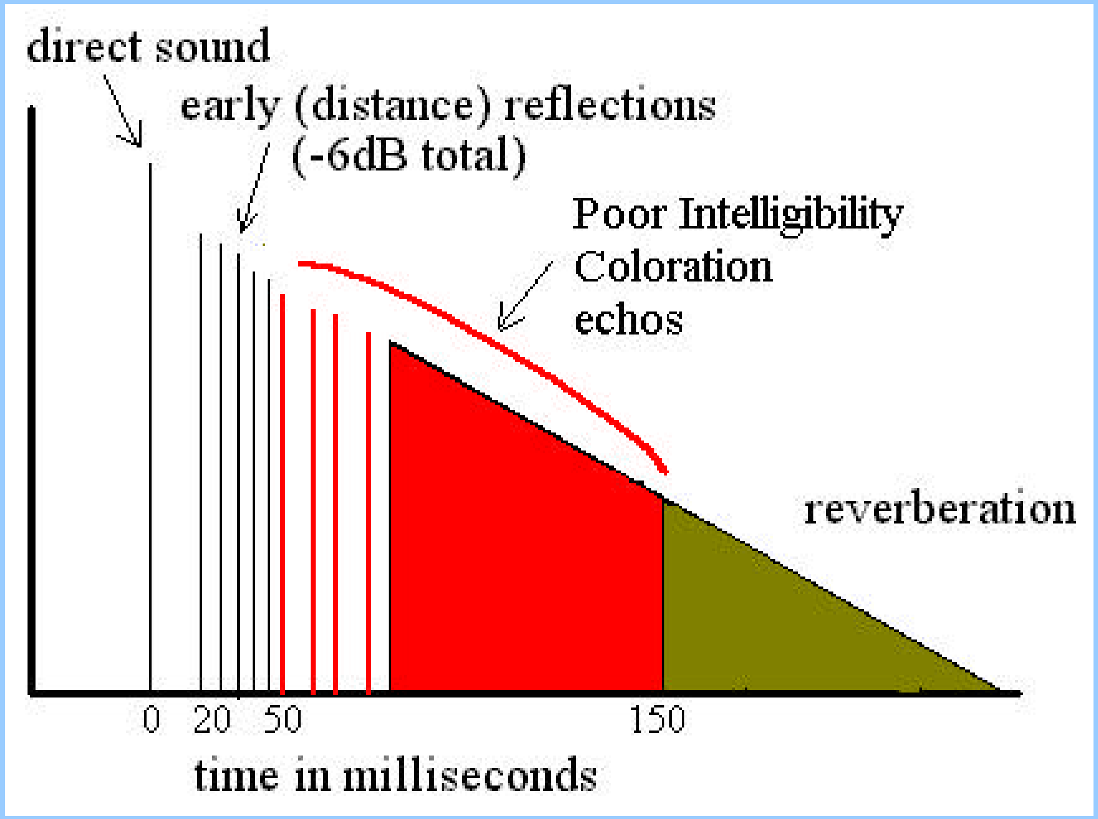
- There is an ideal reverberation profile
  - This profile is required by human perception.
- The ideal profile is NOT provided by most acoustic spaces.
  - Most real rooms have too few very early reflections
  - and too many reflections in the 50 to 150ms time range
- Common microphone technique partially achieves the ideal profile,
    - but only for some instruments in a group.
    - And at the cost of a restricted listening area, reduced intelligibility, and excessive coloration.
- We can often make a more natural recording by using artificial reflections and reverberation.

The Ideal Reverberation



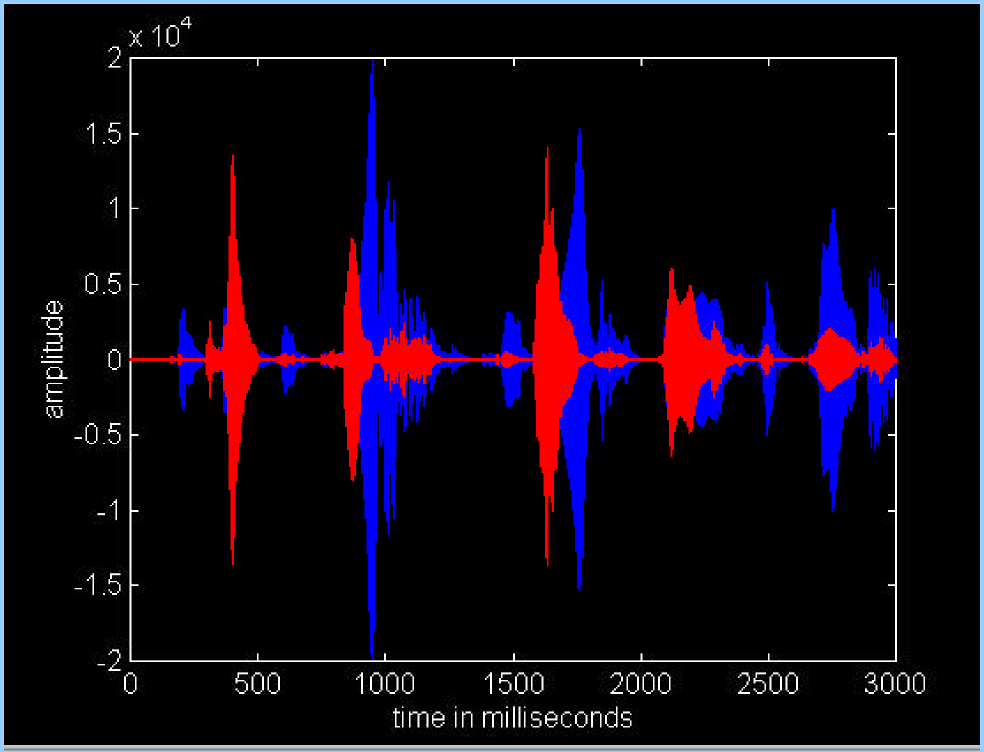
- has 20ms to 50ms reflections with a total energy -4dB to -6dB
- has relatively little energy from 50 to 150ms.

Most real rooms



- Have exponential decay
- to get enough early reflections and reverberation, we get coloration and poor intelligibility.

Why do we want low energy at 50-150ms?

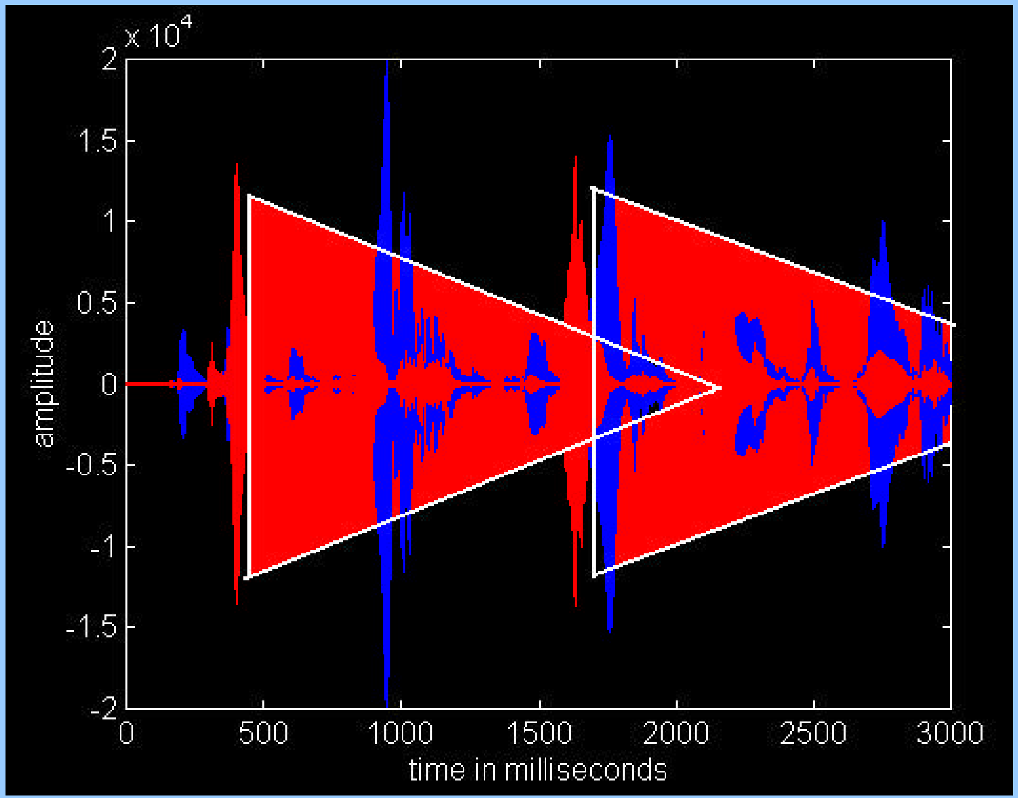


Third-octave filtered speech.

Blue 500Hz.
Red 800Hz

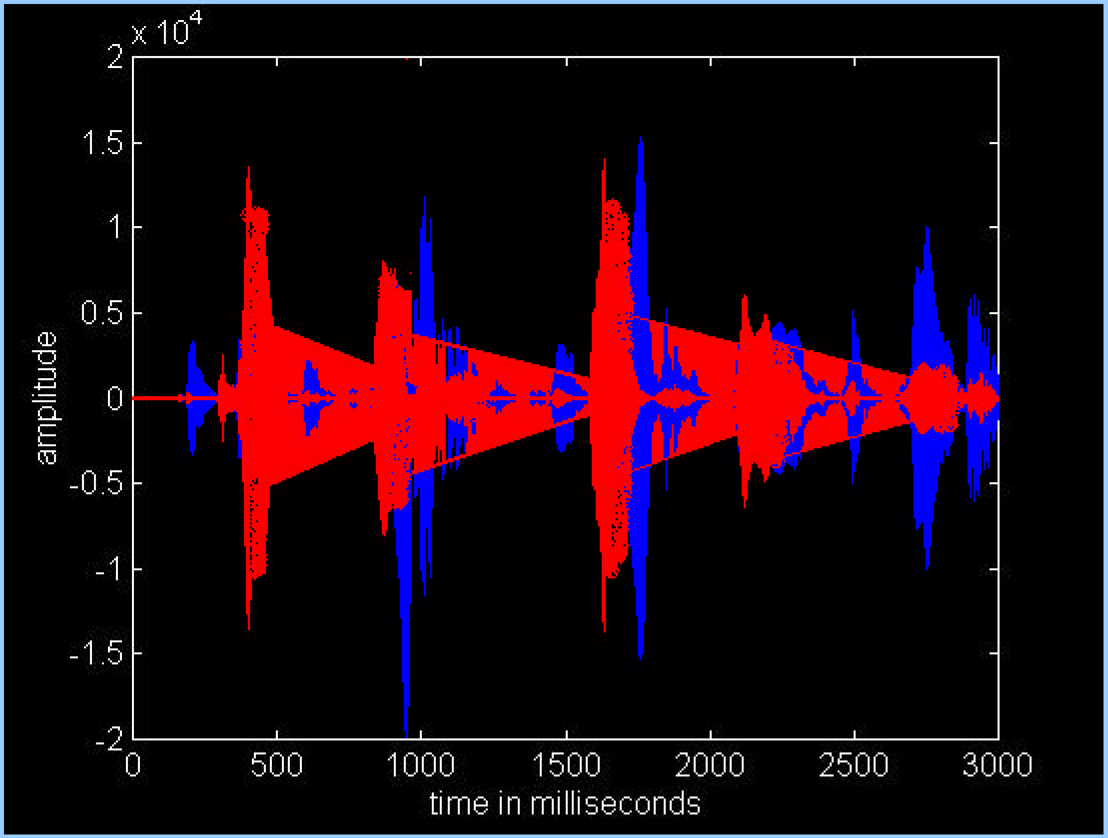
- Because speech (and music) are composed of streams of sound events (notes)
  - with ~ 200ms spaces between each event.
  - a series of such events form a perceptual stream.
  - reflections at 50-150ms make separation of events impossible.

Why exponential decay is problematic.



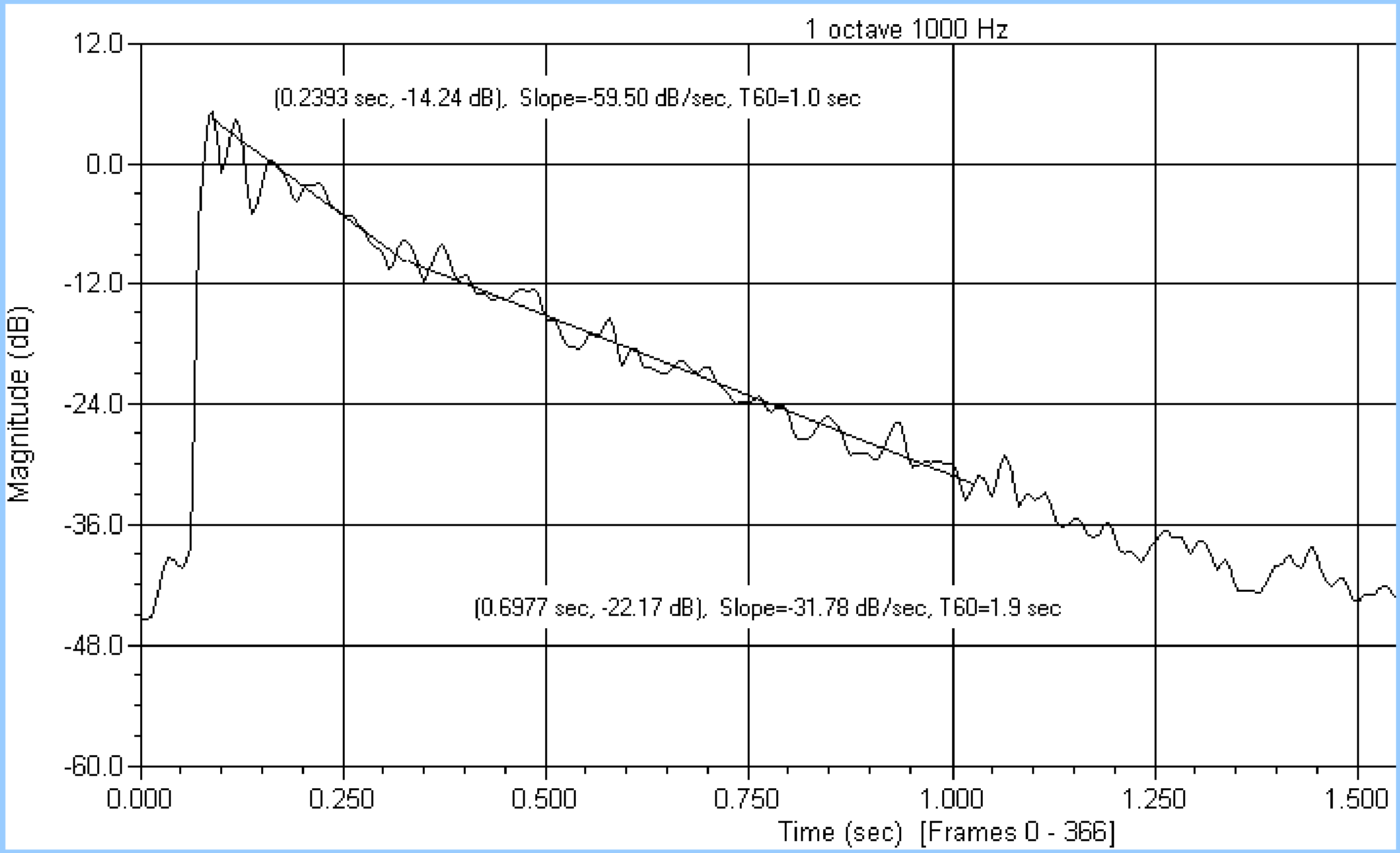
- Adding reverb with exponential decay masks many sound events.
- Comprehension becomes impossible.

Why our "ideal" decay is better



- Strong early reflections combine with the direct sound,
  - But give a sense of blend and space.
  - Intelligibility stays high because the reverberation does not obscure the spaces between sound events.

Back to Boston Symphony Hall:



- Notice the initial decay is ~1s RT, which concentrates the energy in the first 50ms.
- The later, longer decay supplies reverberance without adding mud.
- Smaller spaces usually do NOT have this profile.

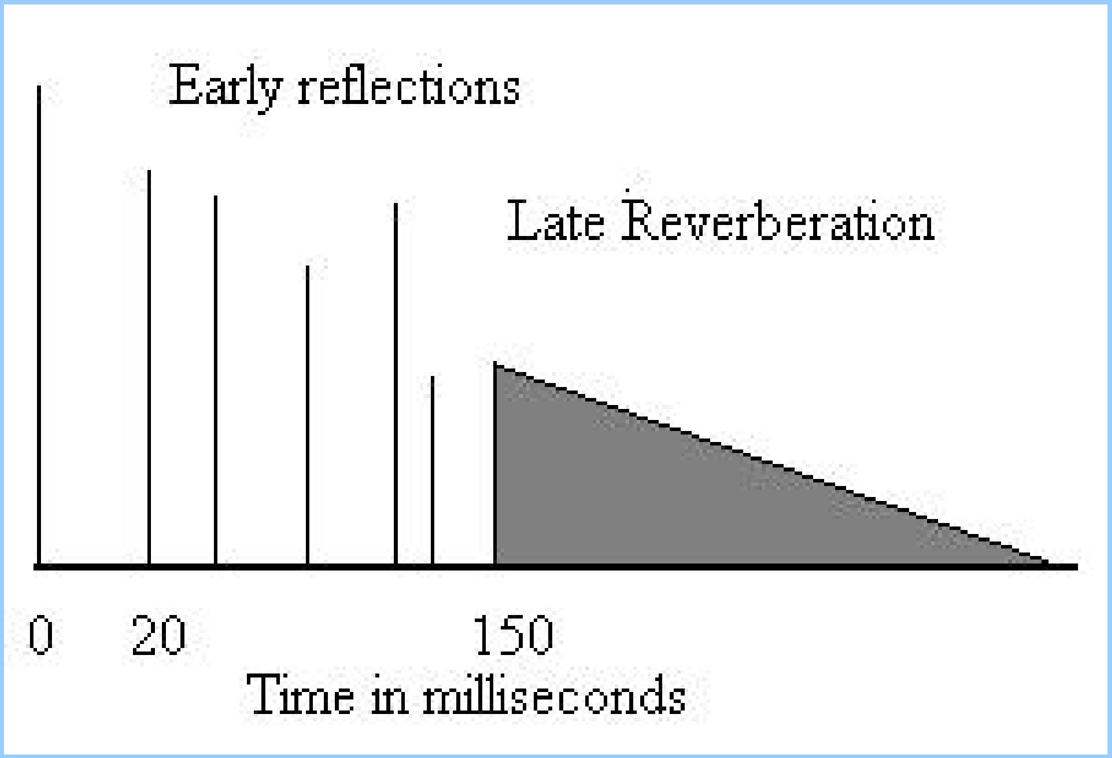
Spatial Hearing

– Horizontal localization and the perception of the spatial properties of a room BOTH rely on the same neural circuitry.
– We detect both through the ITD (Interaural Time Differences) and IID (Interaural Intensity Differences).

– But human hearing (and all human perception) only responds to stimuli that are NOT CONSTANT.

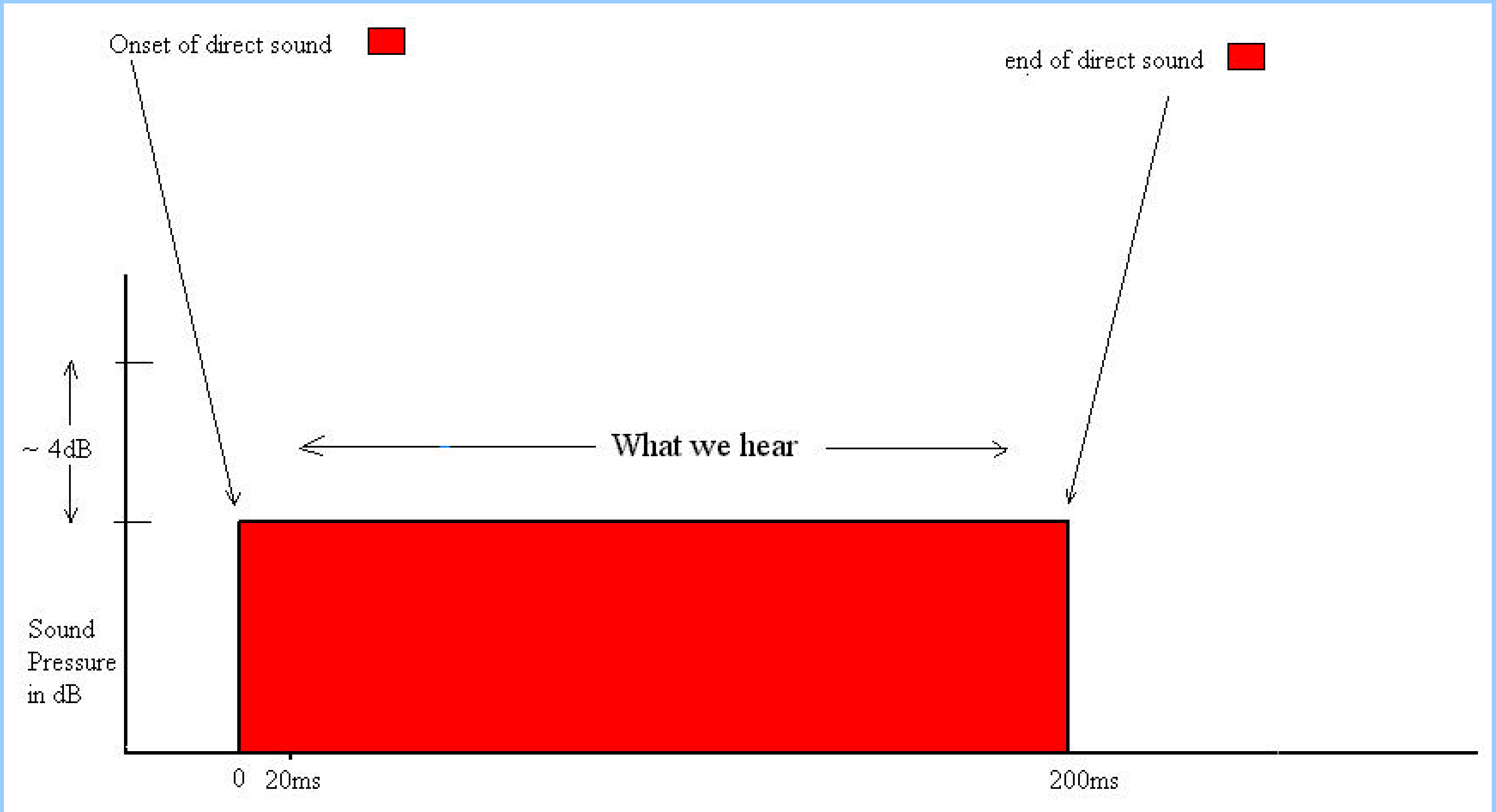
– The START of stimuli are important for localization
– The ENDS of stimuli are important for room perception

Impulse response is not what we hear



- The impulse response is the sound of a pistol, not the sound of music.

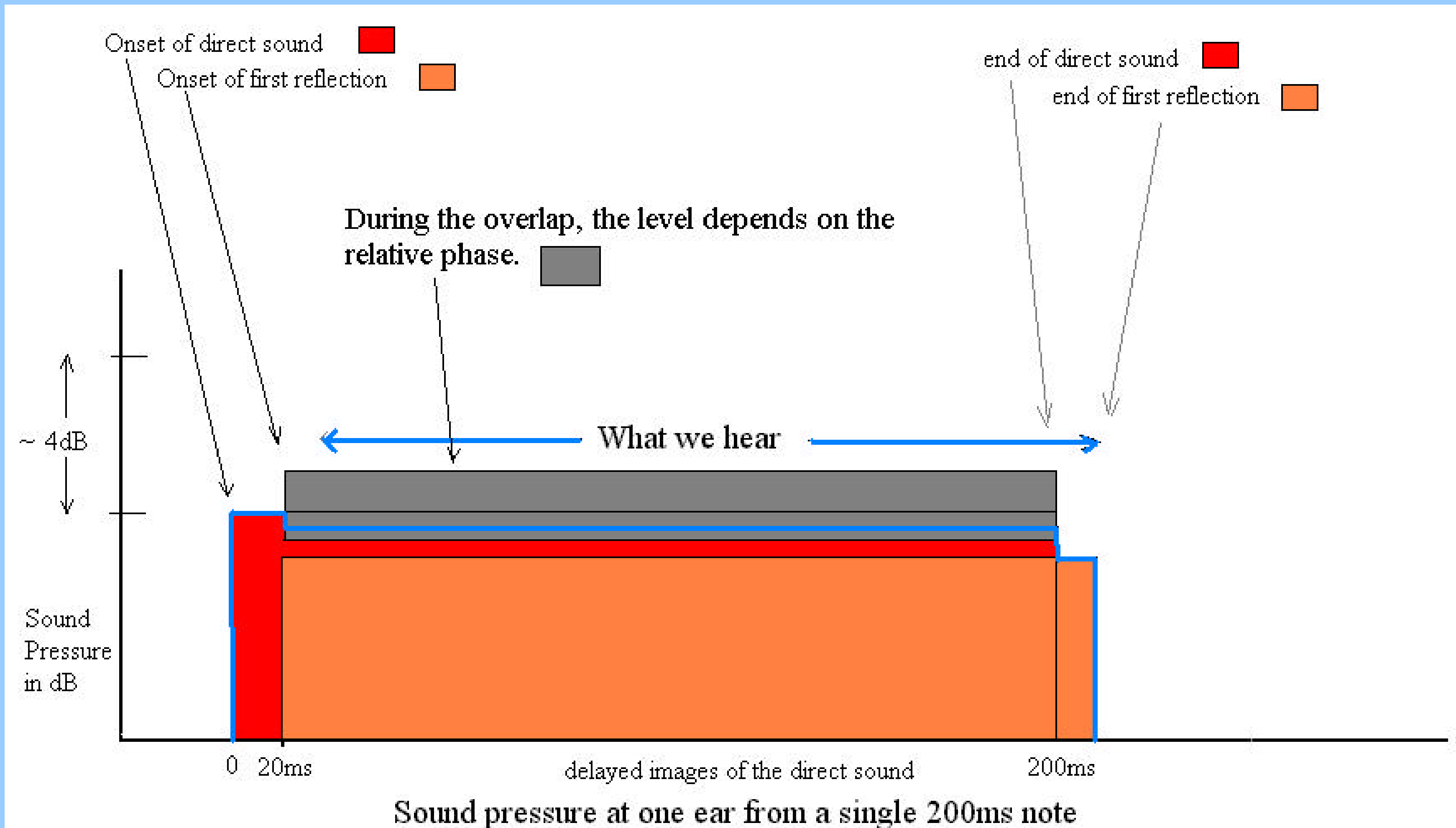
What we hear with notes:



Sound pressure at one ear from a single 200ms note

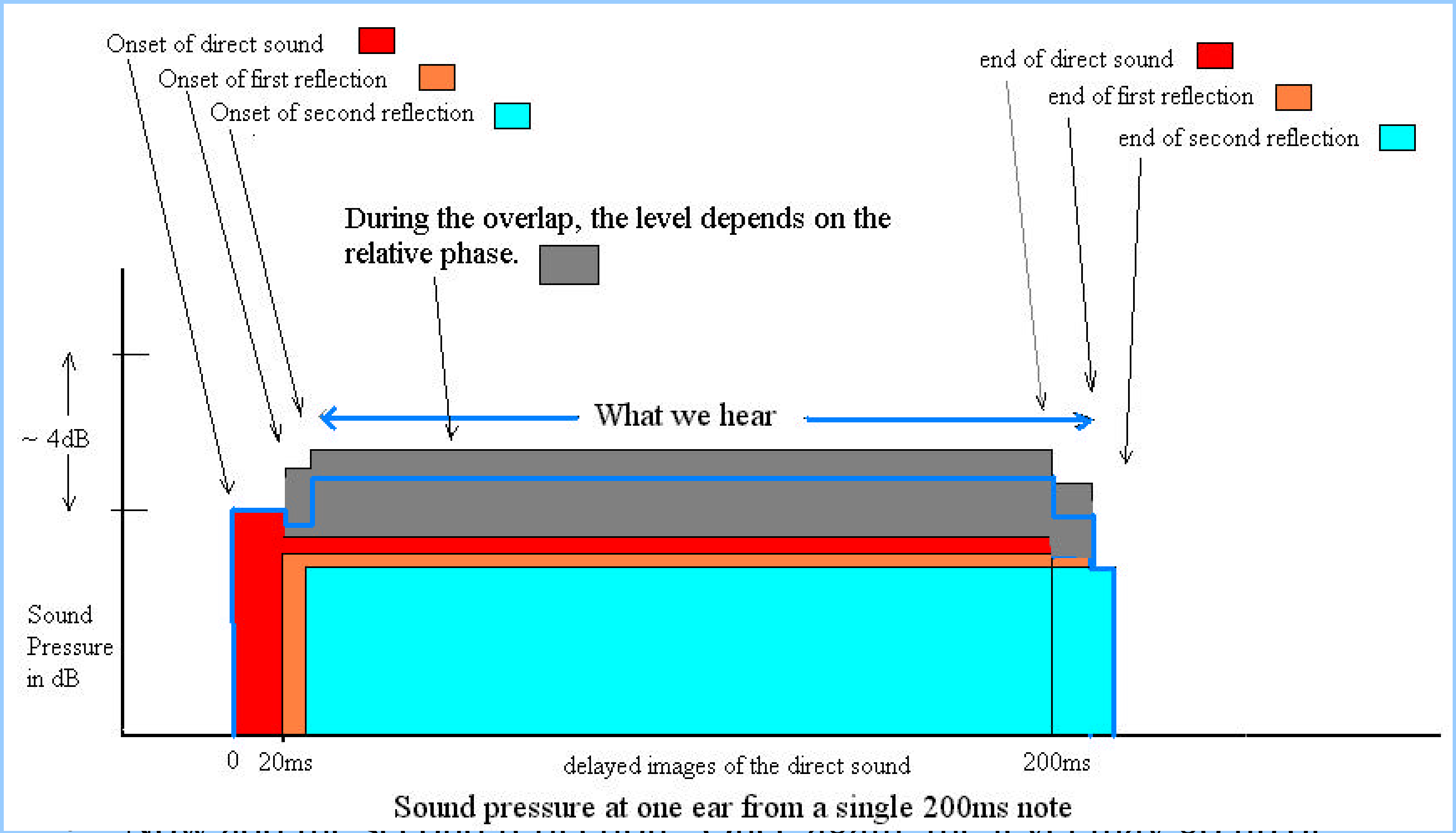
- Notice it has a clear beginning and a clear end.

What we hear with notes:



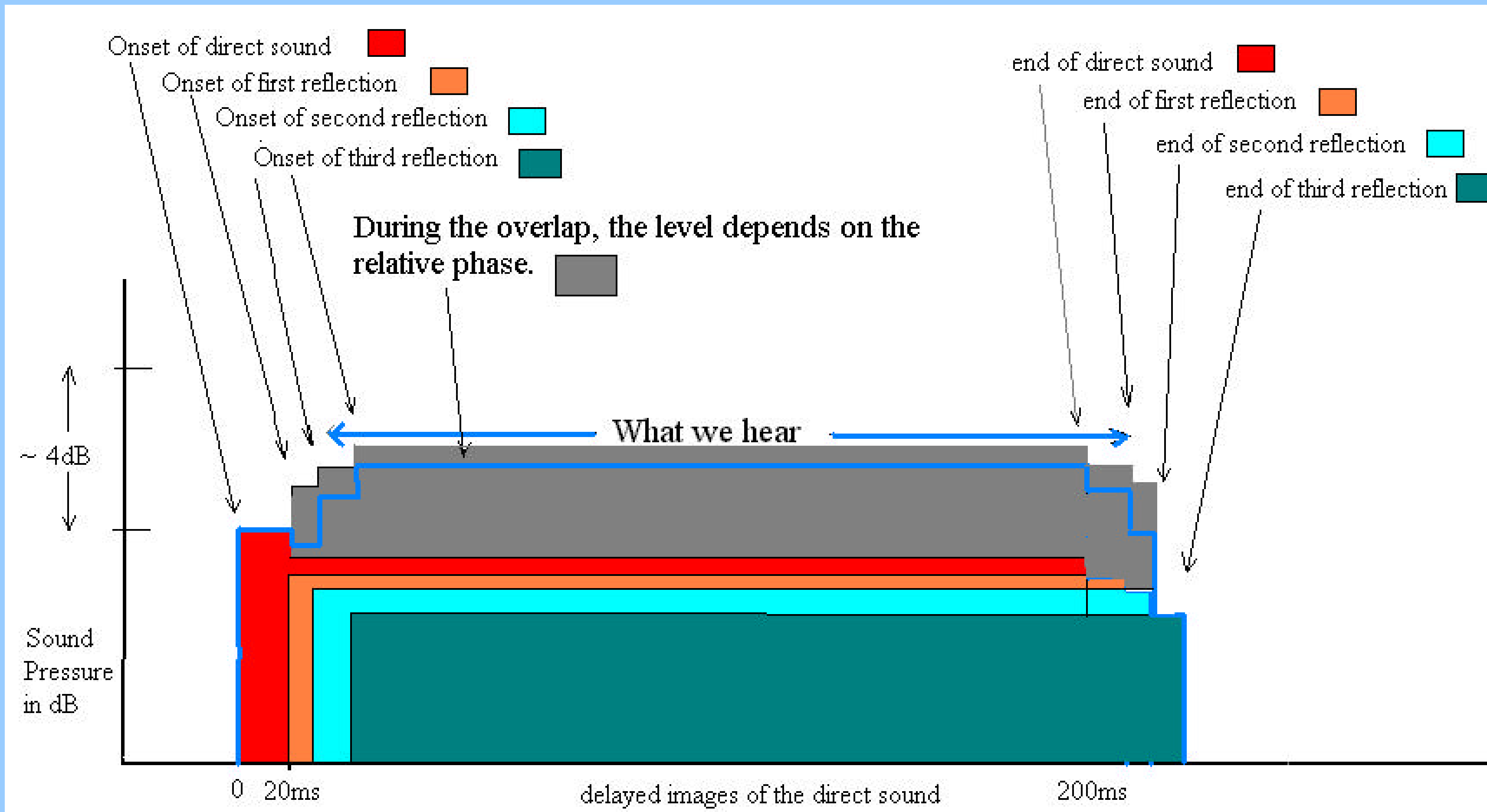
- Now add the first reflection. The level may go up or down, depending on the phase.
- Notice the direct sound is not corrupted for the first 20ms.

What we hear with notes:



- Now add the second reflection. Once again, the level may go up or down, depending on the phase. If the reflection is lateral, the other ear may have a different phase, and a different level.

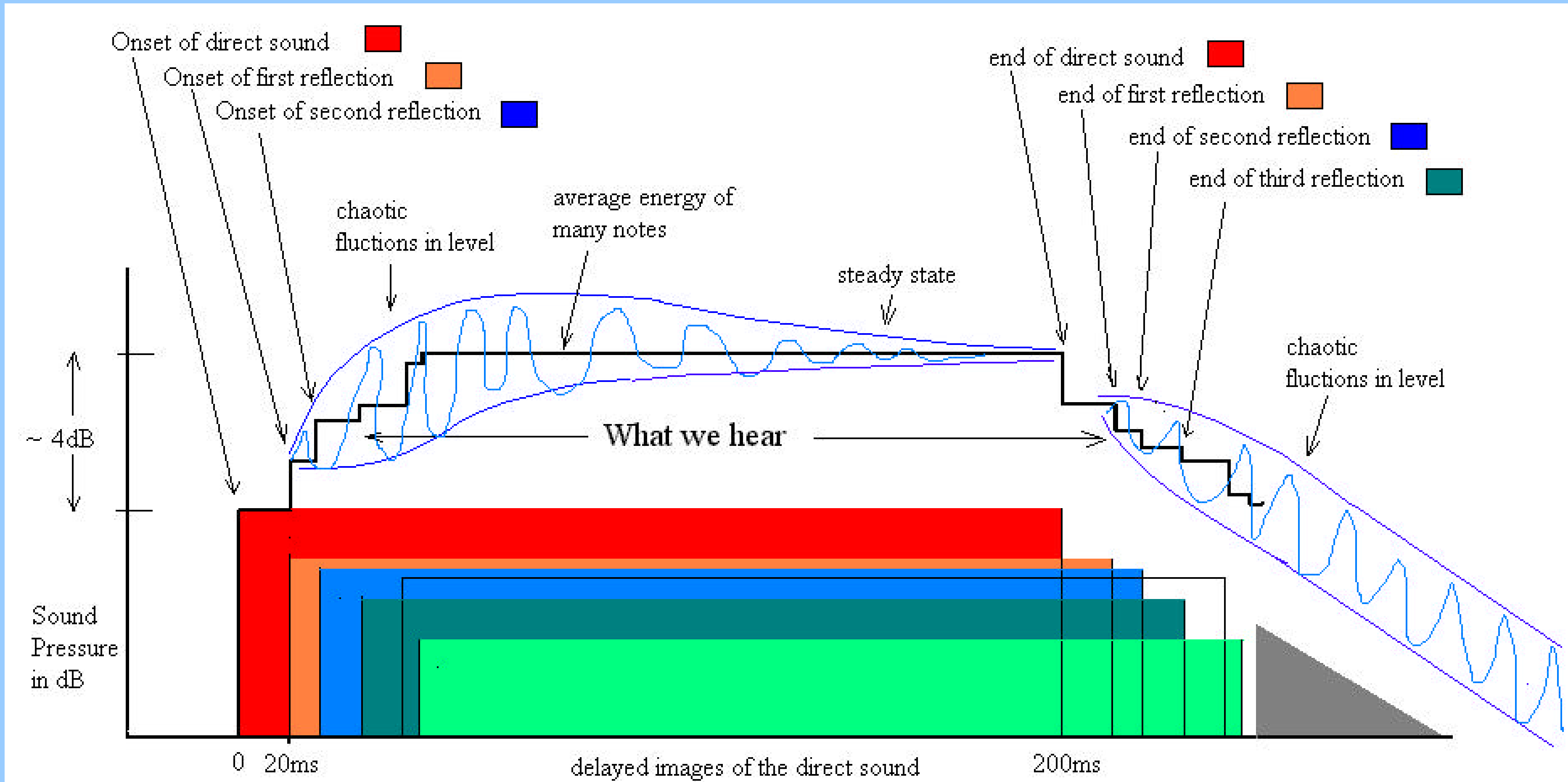
What we hear with notes:



Sound pressure at one ear from a single 200ms note

- Now add a third reflection. Once again the level changes seemingly randomly.

What we hear with notes:

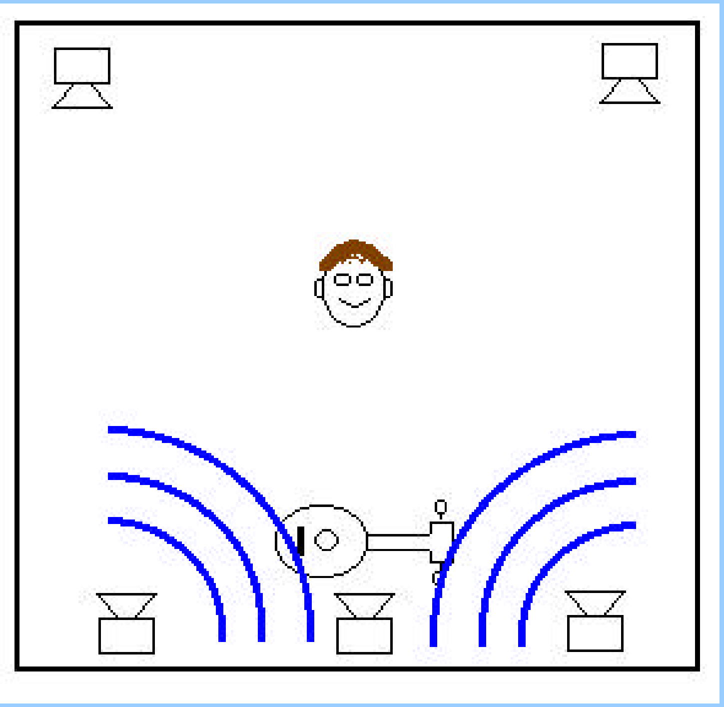


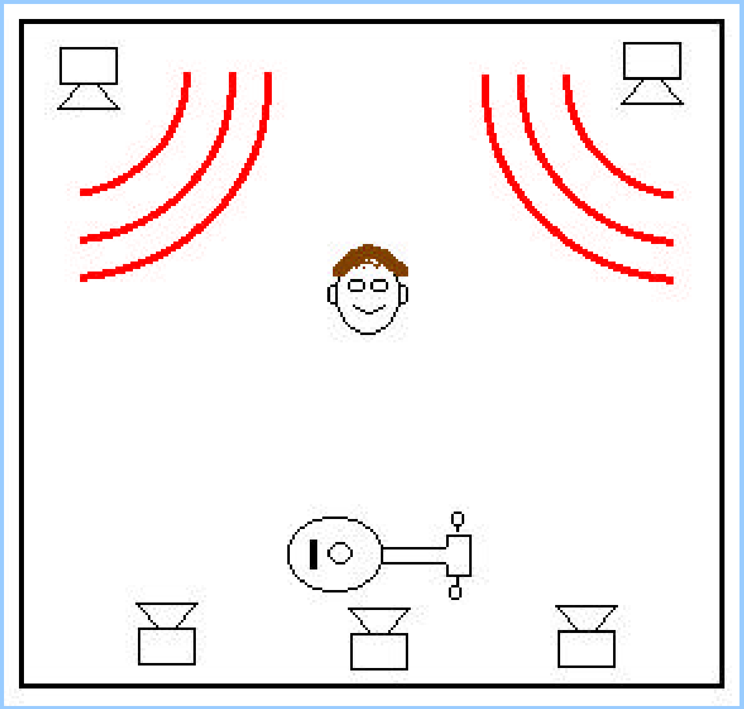
Sound pressure at one ear from a single 200ms note

fluctuates until the room reaches steady-state.

- When the direct sound ends, the fluctuations start again.

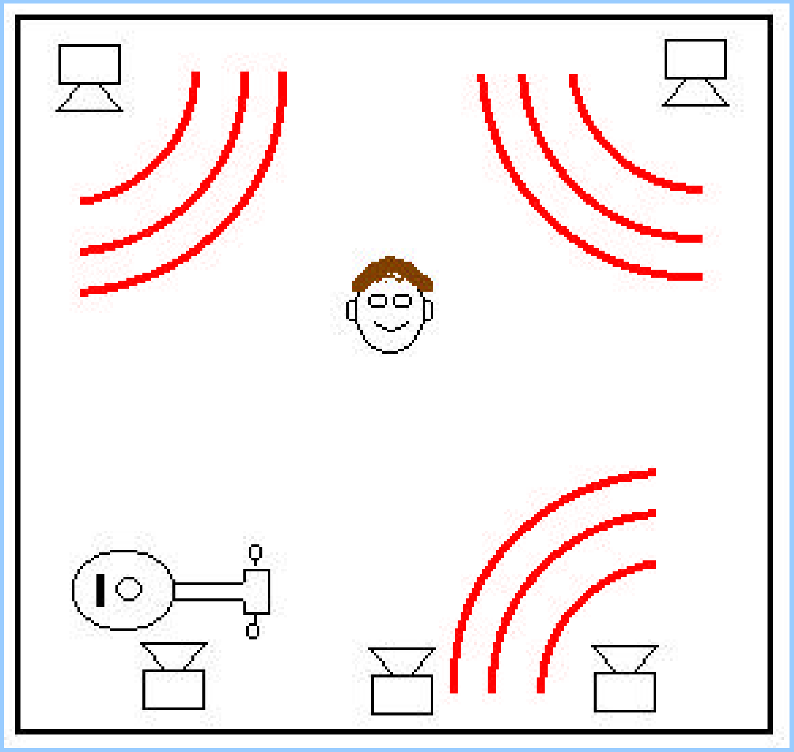
Direction of early reflections





- For reflections that arrive 20ms to 50ms after the end of a sound,
- It is not possible to detect if they come from the front or the rear.
- But it is more natural if they come from both front and rear.

When the sound comes from the right, the early reflections should come from all three other directions.



- Reflections from the same direction as the source are either inaudible or undesirable.

Reflections between 50ms and 150ms

- Reflections between 50ms and 150ms add a sense of distance, but at the cost of reduced intelligibility
- Reflections in this range sound “muddy”.
- They do not create envelopment.

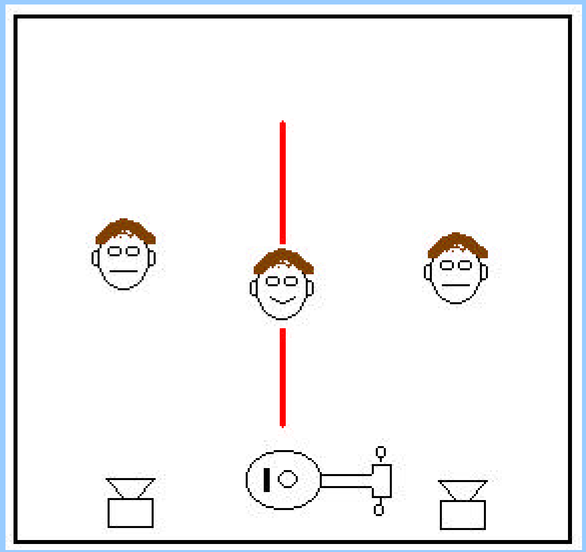
Sound engineers need to control both perceptions separately!

- A recording with to little early lateral reflections sounds too close and artificial
  - There is an optimum level for early reflections
    - -4 to -6dB total energy relative to the direct sound

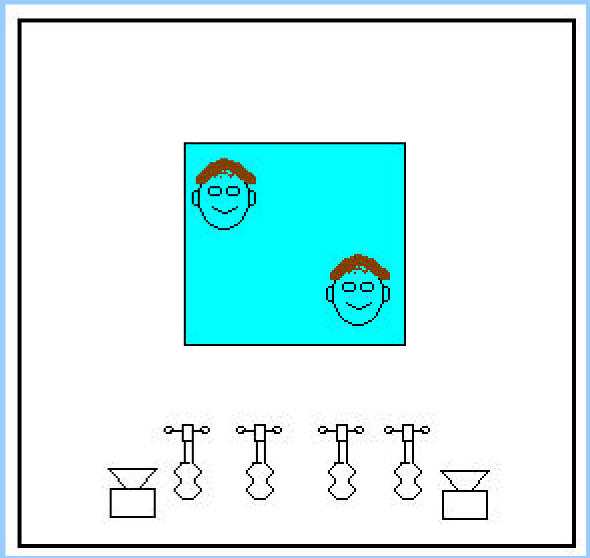
- The level of energy >150ms is critical
  - There is a ~3dB change in audibility for a 1dB change in reverberant level
  - Audibility depends strongly on reverberation time.

We Can Do It!

- The secrets are: high amplitude separation, no time delay panning, and decorrelated reverberation.



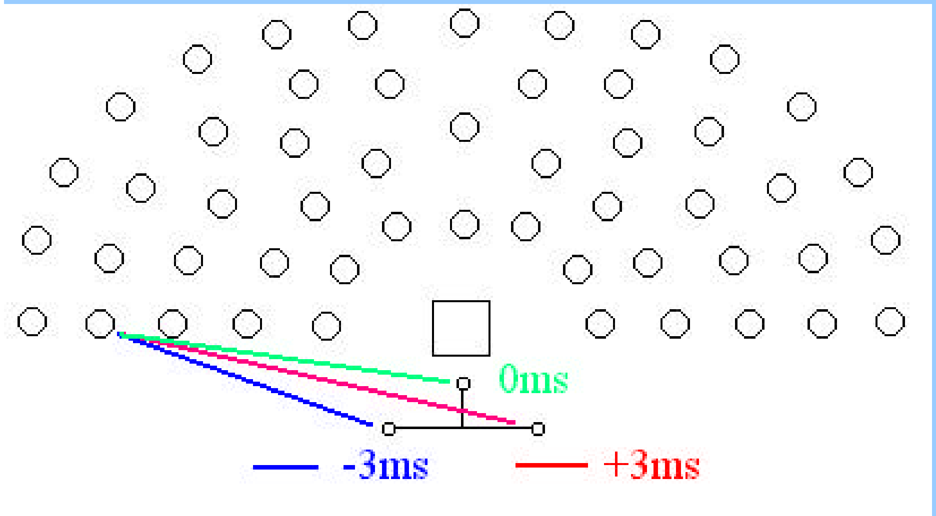
Two channel recordings with a strong center image localize well only on a line between the speakers.

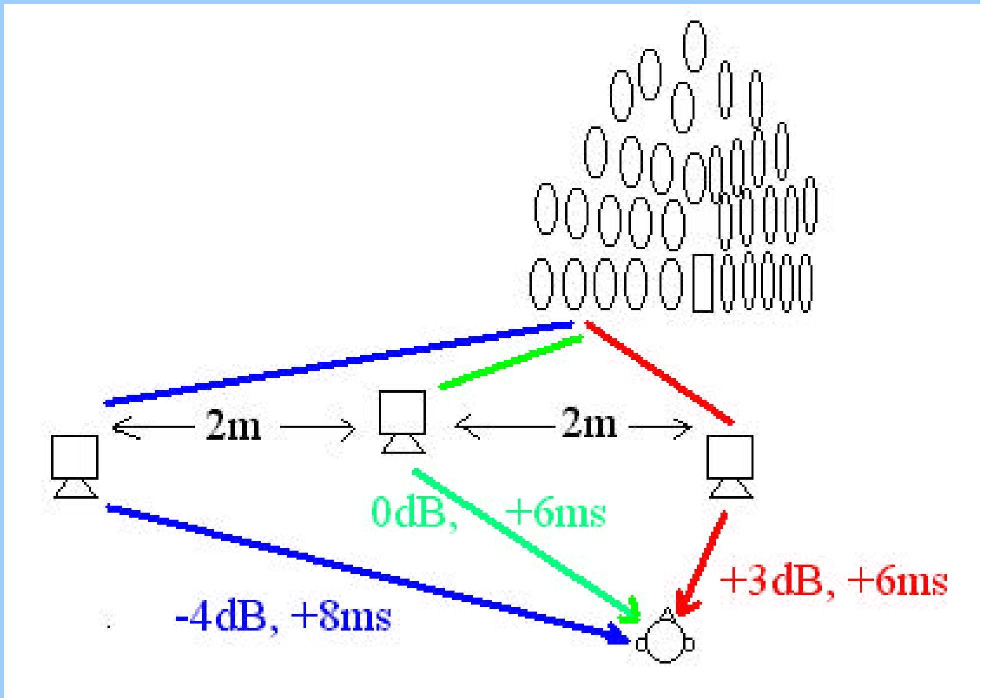


Two channel recordings with a broad source sound good over a wide area if they have high amplitude separation.

Adding a hard center channel improves the listening area for both sources.

Time delay panning outside the sweet spot.

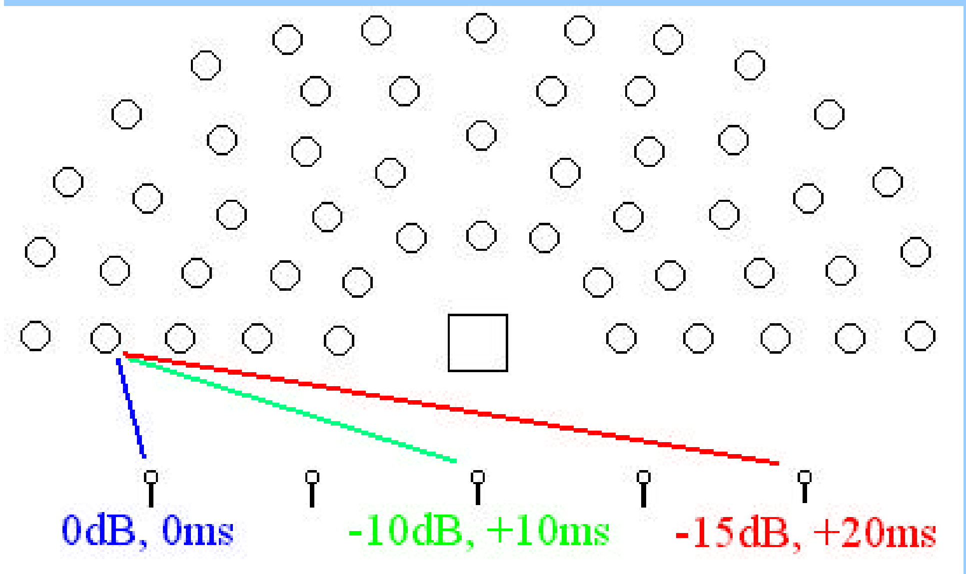


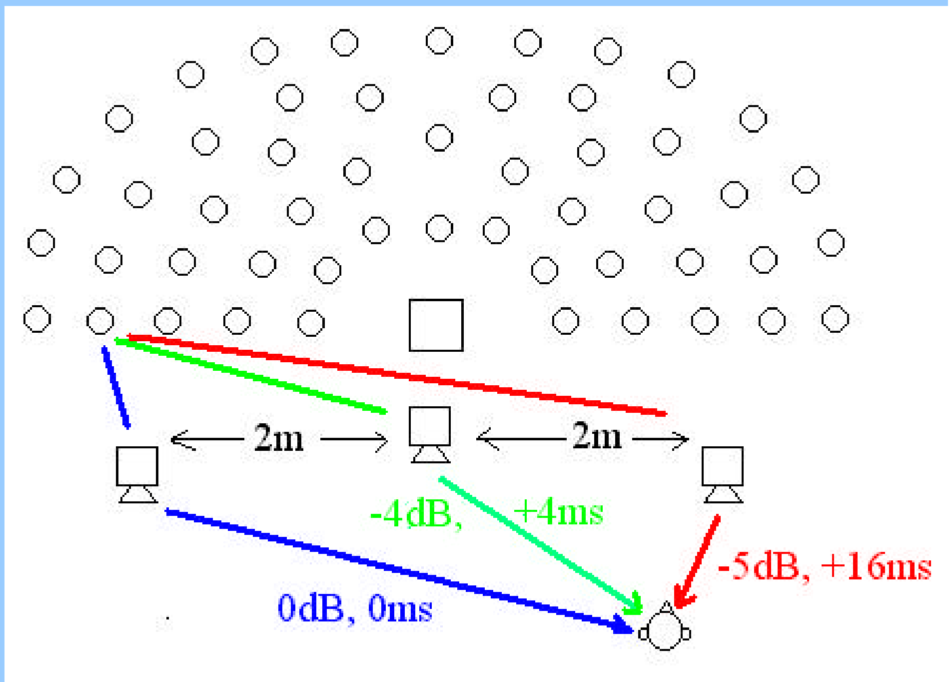


Record the orchestra with a "Decca Tree" - three omni microphones separated by one meter. A source on the left will give three outputs identical in level and differing by time delay.

On playback, a listener on the far right will hear this instrument coming from the right loudspeaker. This listener will hear *every* instrument coming from the right.

Amplitude panning outside the sweet spot.



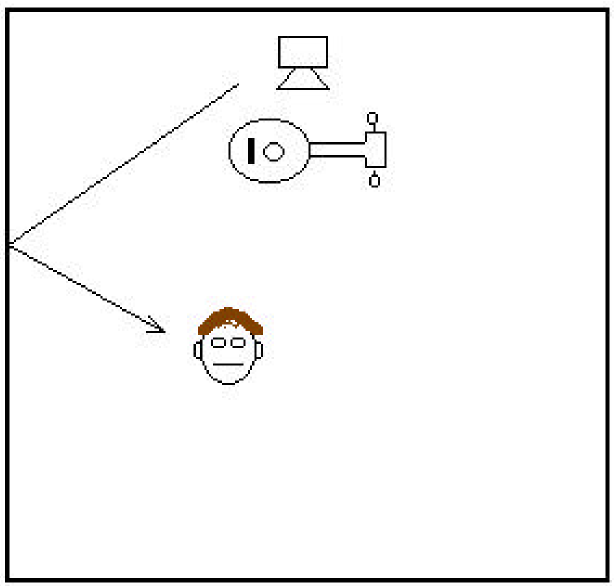


If you record with three widely spaced microphones, an instrument on the left will have high amplitude and time differences in the output signals.

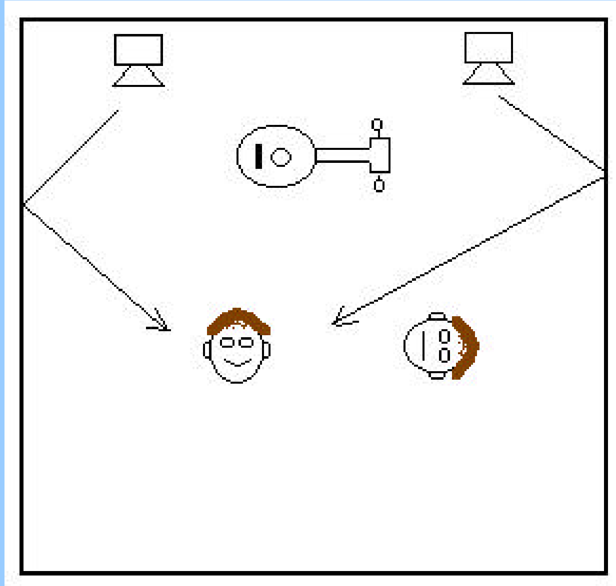
A listener on the far right will hear the instrument on the left. Now the orchestra spreads out across the entire loudspeaker basis.

3/0 versus 3/2

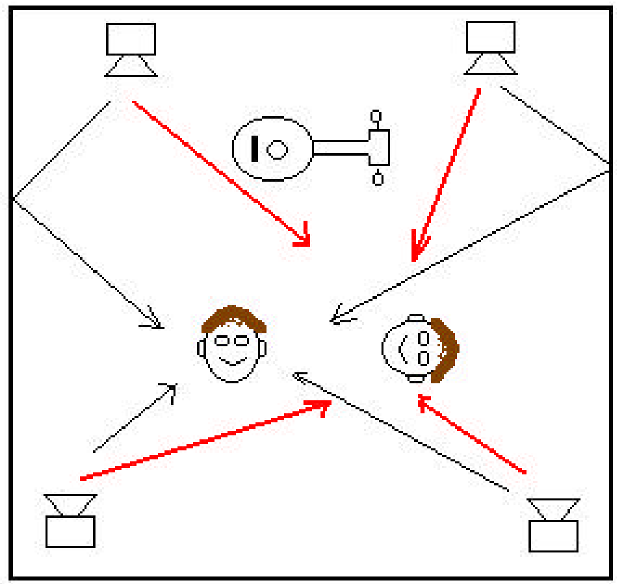
- OK, perhaps we need three speakers in the front, and amplitude panning in the front.
- Why do we need two additional speakers and channels?



Mono sounds poor because it does not reproduce the spatial properties of the original recording space.



With decorrelated reverberation a few spatial properties come through, but only if the listener faces forward.



We need at least four speakers to reproduce a two dimensional spatial sensation.

What is Envelopment?

RT > 7/musical bandwidth

- Envelopment is reverberation that is perceived as coming from all around the listener. Sometimes called “Spaciousness”.
- Envelopment is perceived in a room because the room is causing chaotic fluctuations in the pressure <u>difference</u> at the listener’s ears. (Fluctuations in the ITD and ILD.)
- These fluctuations only occur when the room is NOT in steady state.
- The room is only non-stationary during a time period ~RT/7 at the beginning of a note, and after a note ends.
- In practice spaciousness is not audible unless the bandwidth of the music is GREATER than 7/RT.
- Thus a small room might sound spacious for pink noise, but completely non-spacious for a string bass. (bandwidth ~3Hz.)

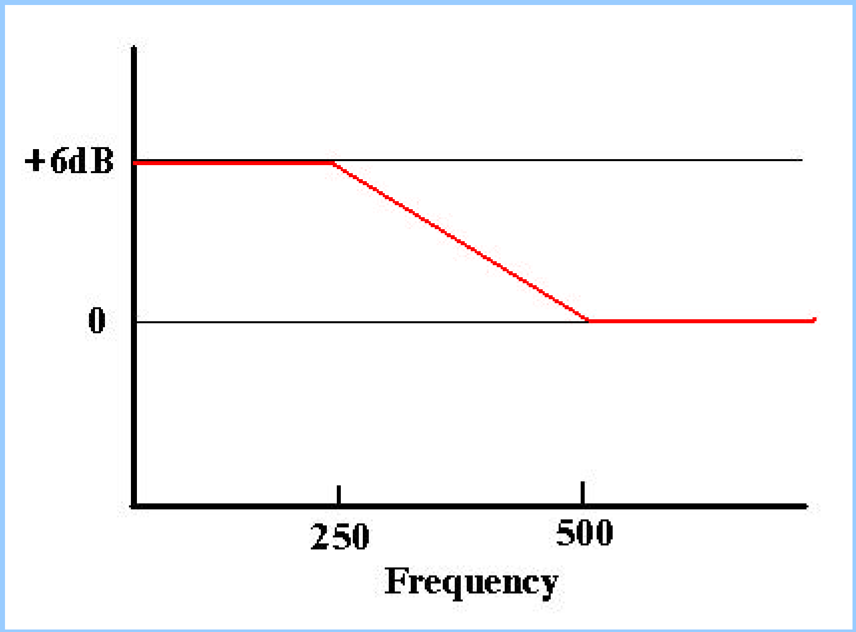
How can we reproduce envelopment?

- The reverberant field of a LARGE room can be reproduced in a SMALL room if:
  - We can excite a fluctuating sound VELOCITY across the listener's head that mimics the fluctuating velocity in the original space.
  - To do this we MUST have at least two LF drivers on opposite sides of the listener.
  - If the listener is allowed to turn the head, we must have at least 3 independent drivers, and four is better!
  - All the LF drivers must be driven by independent (uncorrelated) reverberation signals, derived from a large, non-steady-state room.

Low frequencies are particularly important!

- In our concert hall and opera work it is frequencies below 300Hz where the major benefit is achieved.
  - The result is “inaudible” but highly effective in increasing the emotional power of the music.
- It is commonly believed that because we “cannot localize” low frequencies in a playback room we need only one LF driver
  - We can however easily hear the difference on reverberation.
- It is often the case that using a shelf filter on the rear channels can greatly improve the surround impression.

Shelf filter for rear channels



Applying a shelf filter to the rear channels increases subjective envelopment dramatically without drawing attention to the rear speakers.

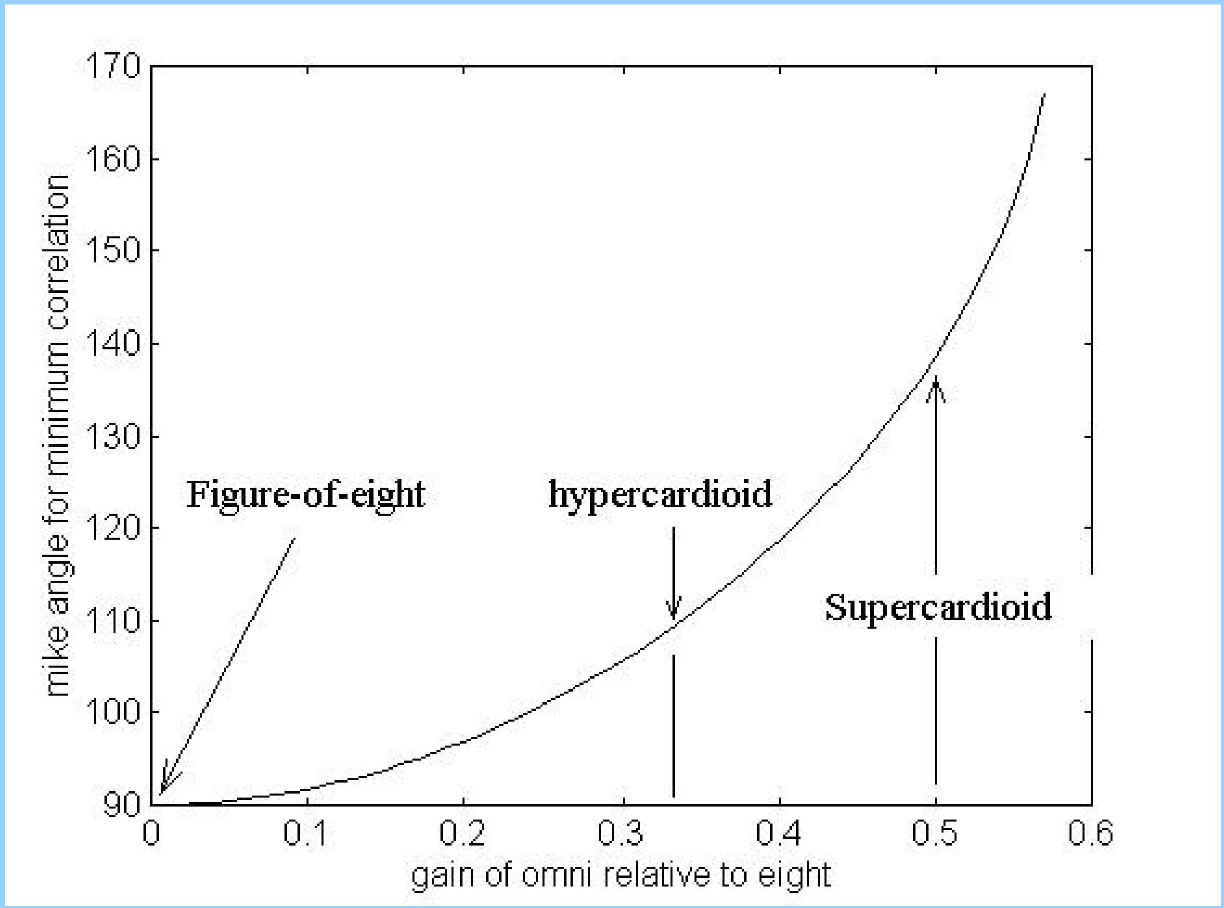
What is Correlation?

- The concept of correlation is difficult to define. We really want some measure of how much chaotic fluctuation can be created by the signal at the position of the listener.
  - Simply delaying one channel by a millisecond or so greatly reduces the mathematical correlation at time zero, but has little effect on the perceived correlation.
  - A measure of correlation MUST be a function of frequency, because we hear the fluctuation in different frequency bands independently.
- Thus a broadband "correlation meter" does not measure the degree of difference that we hear.
  - A broadband meter is insensitive to low frequencies, and may miss high correlation at small time delays.

Decorrelated reverberation with coincident pairs

- It is possible to have a coincident pair with completely decorrelated reverberation, but only certain combinations of pattern and microphone angle will work
  - Blumlein (figure of eight microphones at 90 degrees) works fine because sound velocity due to reverbeartion is completely independent in two orthogonal directions.
  - It is NOT possible to have decorrelated reverberation with a cardioid microphone pair.
  - Or with an MS array with an omnidirectional front microphone.
  - An MS pair with a cardioid front microphone is capable of decorrelated reverberation.

Ideal angle as a function of microphone pattern for decorrelated reverberation in a coincident pair.



- It is NOT possible to achieve decorrelation with cardioid microphones!

Correlation can be reduced in a microphone pair by separating the two microphones in space.

- Mathematical correlation at time zero can be reduced by separating the microphones by more than half a wavelength.
- However correlation at short time delays may still be audible.
- Correlation is always low when the microphones are separated by the reverberation radius.
  - In this case the reverberation seen by each microphone is different.

Formula for decorrelation of an omnidirectional pair as a function of distance and frequency

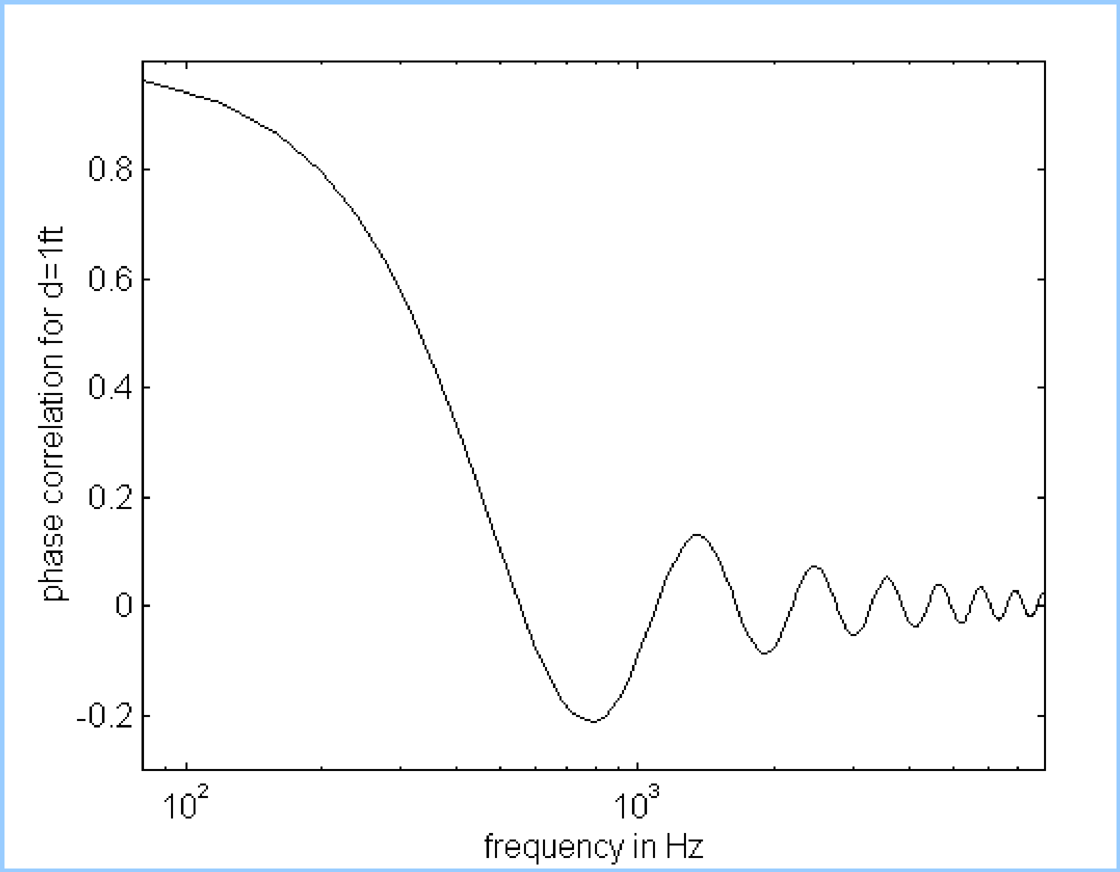
$$\int_{0}^{pi/2} \sin(a)*\cos[2*pi*f*d*\cos(a)/c]\,da$$

- Where f is the frequency, d is the distance between the two microphones, and c is the speed of sound.
- Formula derived by integrating the phase resulting from incident sound as a function of the angle of incidence.
- We assume d is much less than the reverberation radius.

Simplified Formula

- The above formula is equivalent to a synch function of the form sin(x)/x.
- Correlation = sin(2*pi*f*d/c)/(2*pi*f*d/c)

Correlation of two omnidirectional microphones in a reverberant field as a function of microphone separation.

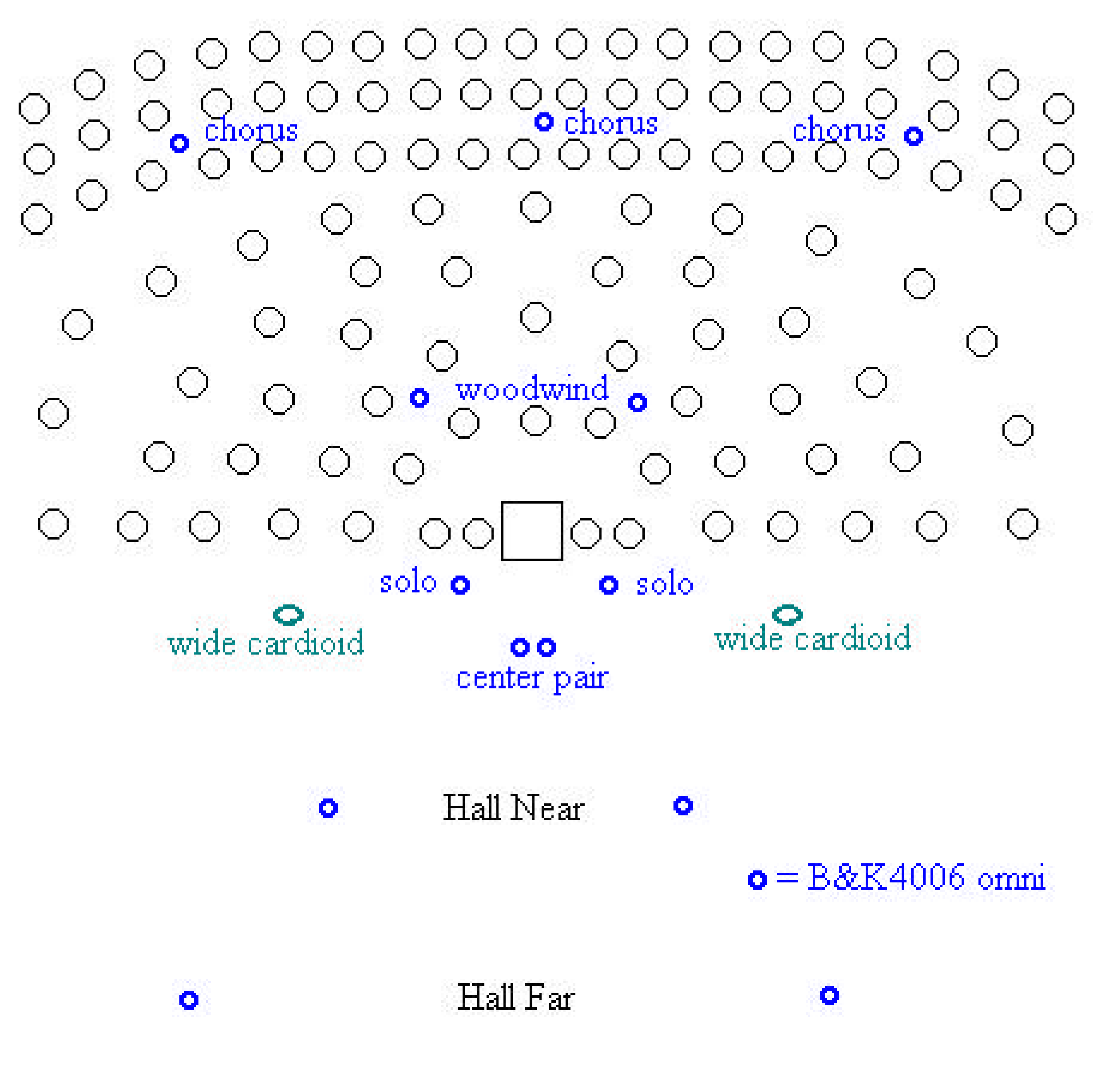


- Notice high correlation below 300Hz, and negative correlation at 800Hz.
- Frequency and distance are inversely proportional.

Recording in Symphony Hall

- With large forces the main problem is balance.
  - We will need a lot of microphones!
- Clarity is almost always good, as there is little reverberant energy in the 50-150ms time range.
- Our major problem will probably be maintaining the sense of distance while correcting balance.

Slavery Documents – as set up by WGBH



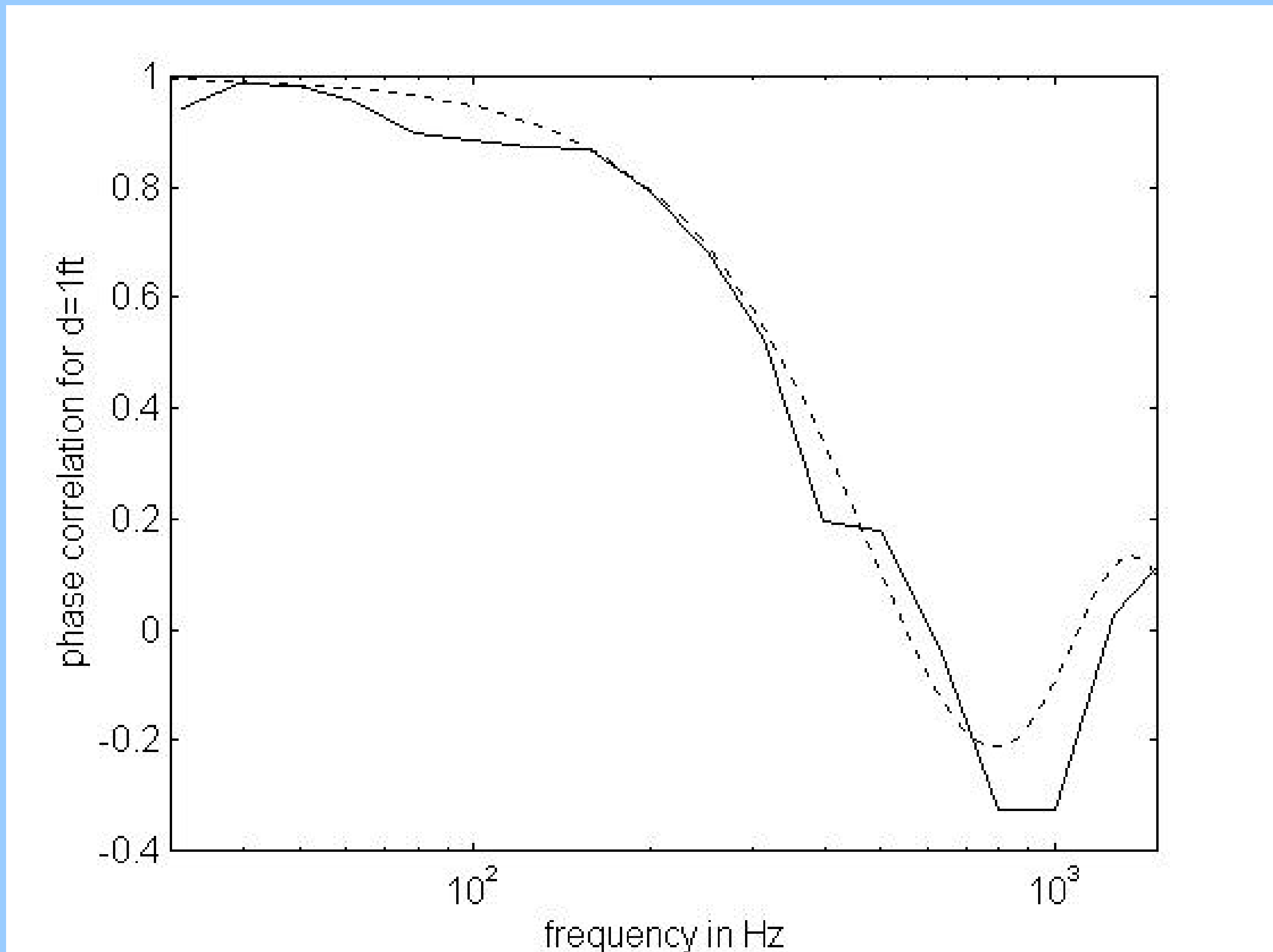
Boston Cantata Singers in Symphony Hall. March 17, 2002

Major Characteristics

- Large reverberation radius – direct sound dominates the pick-ups
- Large distances between microphones –
  - Concept of “main microphone” is meaningless, as there is no position where the balance is remotely acceptable.
  - leakage of distant instruments into other microphones is masked by closer instruments.
    - The early reflections (blend) such leakage often supplies is absent.
  - Thus the sound tends to be up-front and too clear
- Highly decorrelated reverberation with widely spaced omni microphones

Audio Demos

Correlation in the omni “Hauptmicrophone”



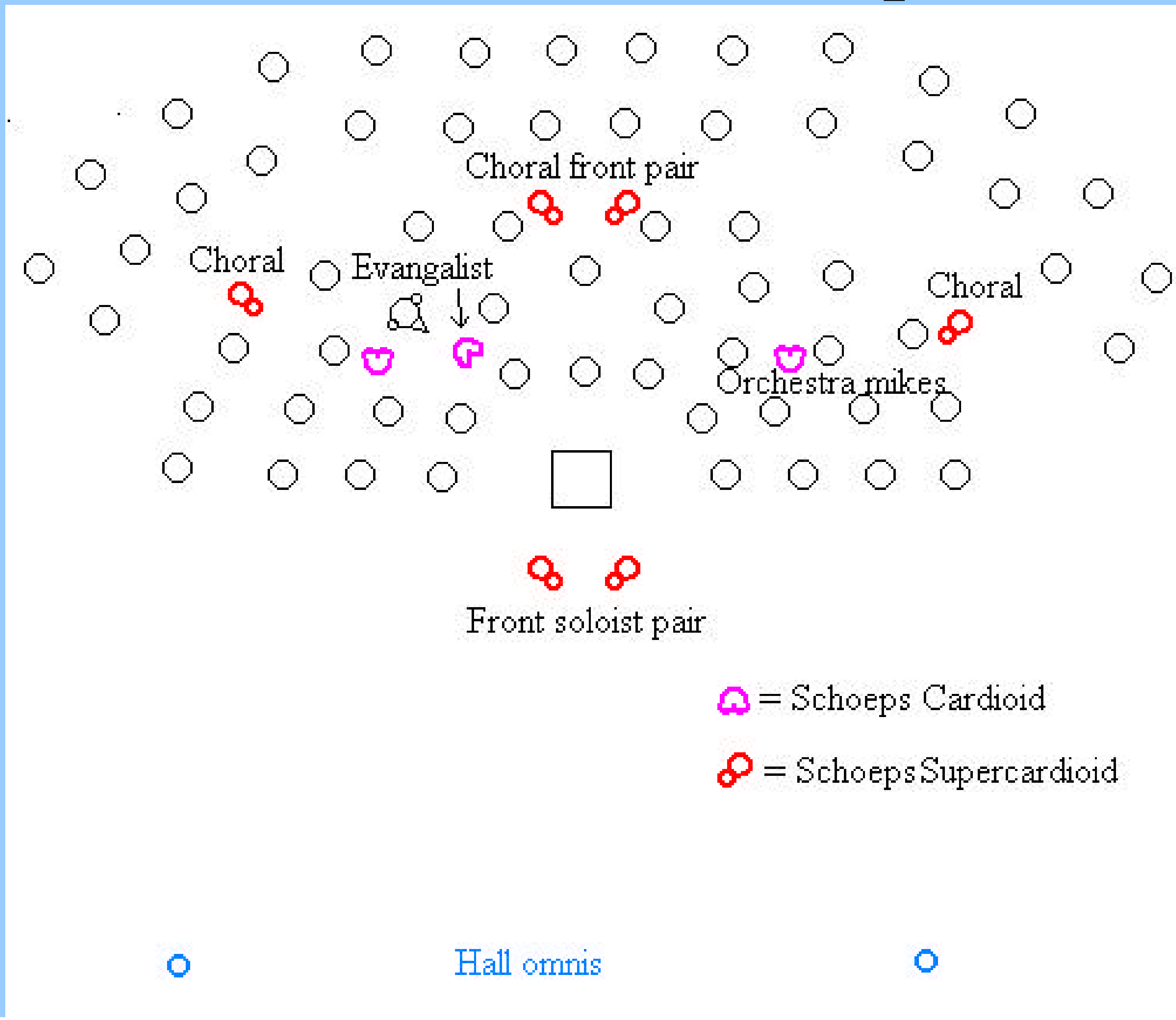
___ = measured correlation; - - - = calculated, assuming d=25cm

Boston Cantata Singers in Jordan Hall

Major Characteristics

- Chorus is deep in an enclosing stage-house with significant reverberation.
- Small distances between microphones results in unwanted leakage.
- Microphones pointed into the stage house increase the amount of undesirable reverberation.
  - Thus the chorus mikes, which must face the chorus, are supercardiod to minimize reverberation pick-up.
  - And the orchestra mikes face the hall, not the stage house.
- Microphones in front do not pick up enough direct sound from the chorus to supply the sense of distance without also getting considerable mud.

Jordan Hall Setup



Solutions

- Add distance to the chorus at the mixing stage with controlled early reflections
- Minimize stage-house pickup wherever possible

Audio Demos

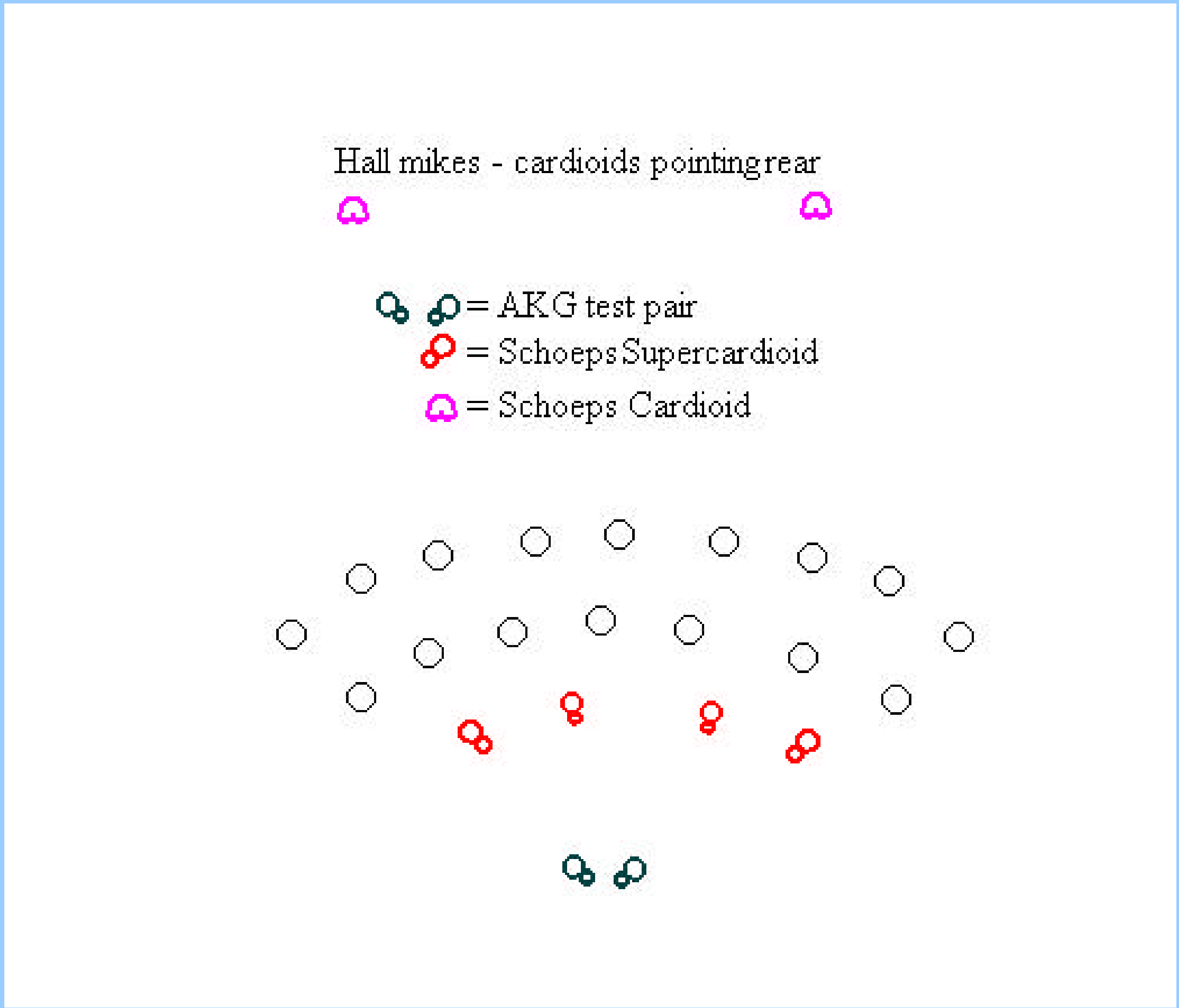
- Early reflections
- Late reverberation

Oriana Consort in Swedenborg Chapel

Major Characteristics

- Hall has relatively low volume of 1450m^3 at the same time as medium RT ~1.5s
  - Low Volume and high RT means the reverb LEVEL will be very high!
    - We will have to keep the microphones close
  - Reverb time is a bit to short for this type of music.
  - With a small group it might be possible to use a microphone pair for a two channel recording.
    - But it might sound better if you did not.

Oriana Setup



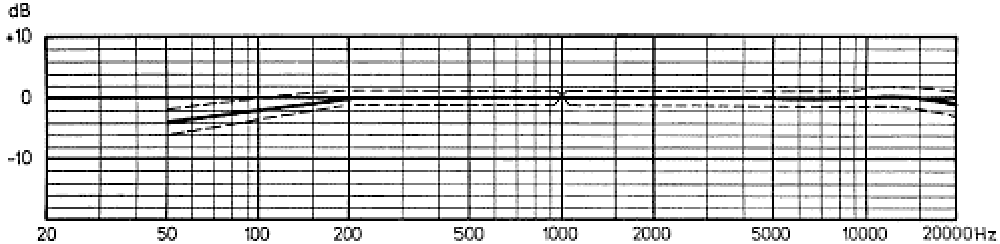
Surround Recording

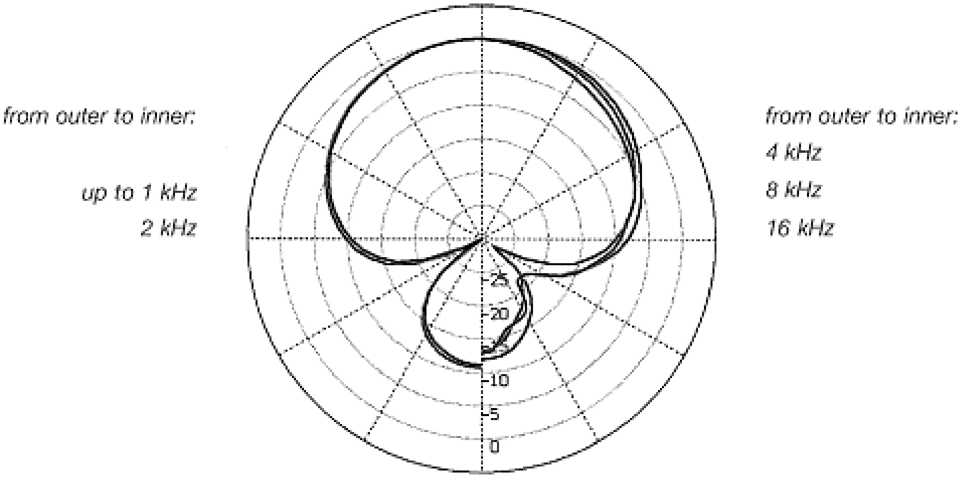
- The recording is created using the multimicrophone front array, (equalized)
  - Augmented with an early reflection pattern from Lexicon in all four outer speakers.
- The surround environment is created using the rear microphones (equalized for the bass roll-off) for the rear channels.
  - And Lexicon late reverberation for the front,
    - And some in the rear also.

The Microphone Pair

A venerable pair of multi-pattern microphones

Another possibility





Pressure Gradient Microphones

- Pressure gradient microphones are a combination of an omni and a figure of eight.
- When the two are mixed with equal on-axis sensitivity, a cardioid results.
  - Reduce the gain of the omni by 6dB and you have a Supercardioid.
  - Reduce the gain of the omni by 10dB and you have a Hypercardioid.

Problem:

- The figure of eight in nearly all available microphones has a bass roll-off, typically at about 120Hz. (Depends on diaphragm size.)
  - When we combine this with the omni – which (may) be inherently flat at LF:
    - The overall sensitivity decreases at LF
      - The mike sounds weak in the bass compared to an omni
    - The pattern may become omni directional at low frequencies
      - This is particularly true for large dual-diaphragm mikes.

Solution

- One solution to this problem is to “correct” it by measuring the microphone at a distance of 1M from the sound source!
  - A spherical sound wave increases the LF velocity of the sound at 6dB/octave when the distance to the sound source approaches ½ the wavelength.
  - A one meter measurement distance exactly compensates for the inherent roll-off of the velocity transducer, and an apparently perfect microphone results.
- A more satisfactory solution would be to equalize the figure of eight pattern electronically before combining it with the omni.
  - The “Soundfield” microphone does this.
- One can also roll off the omni response (electronically or mechanically) to match the figure of eight.
  - Mr. Wuttke (Schoeps) takes this approach.

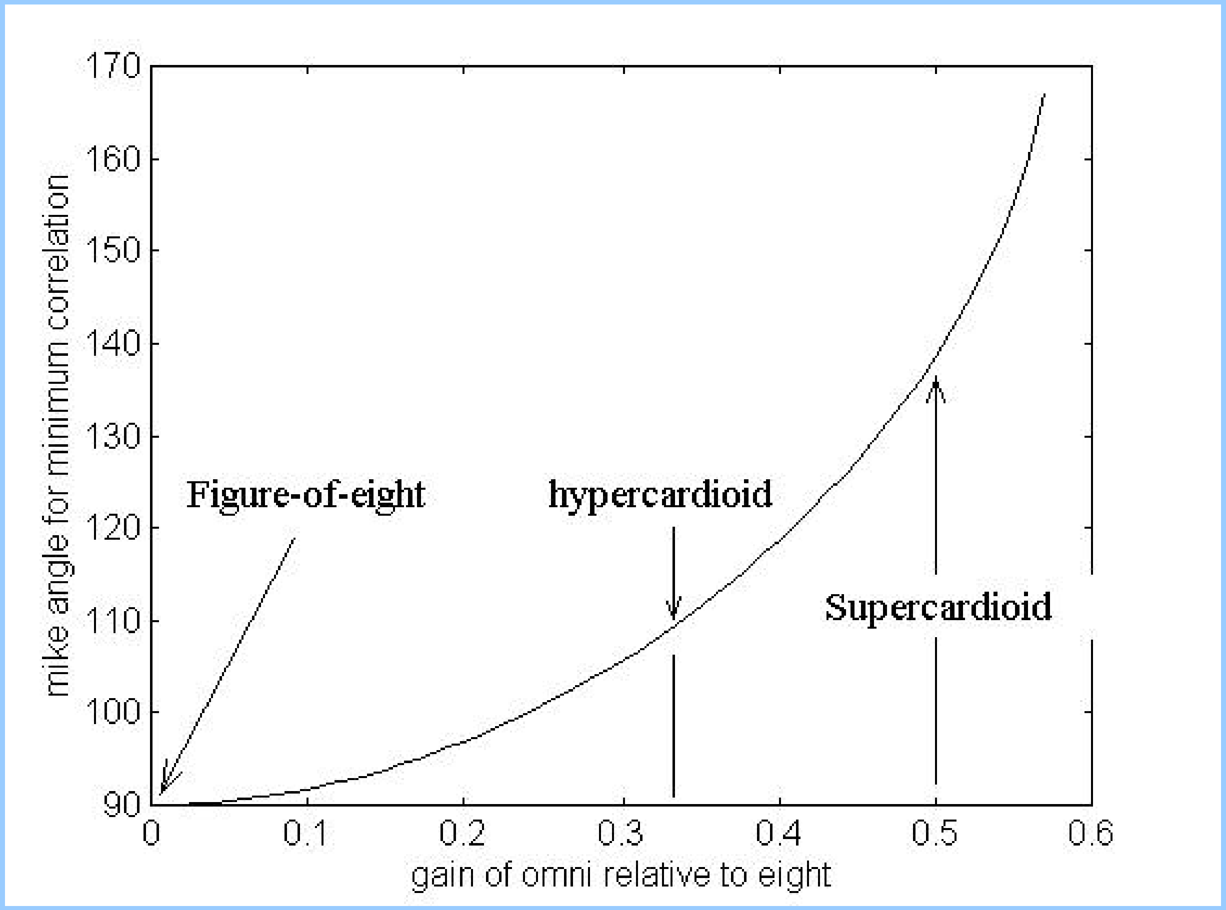
Consequences

- Nearly all available directional microphones either roll off the bass response,
  - Which can be compensated for at the mixing desk
- Or they become omnidirectional at low frequencies,
  - Which usually cannot be compensated.
- Or they do both.
  - The venerable microphones shown earlier do both.
- The consequence for a ORTF – style pair is
  - The low frequencies will be generally weak
    - Which can be compensated.
  - The low frequencies may be monaural
    - Which is more difficult to compensate.
    - But which can be fixed with a Blumlein shuffler circuit
- Be sure to equalize the LF when you use directional microphones!

Correlation of reverberation

- Remember we are (desperately) trying to keep the reverberation decorrelated.
  - We can do this with a coincident pair if we choose the right pattern and the right angle

Ideal angle as a function of microphone pattern for decorrelated reverberation in a coincident pair.

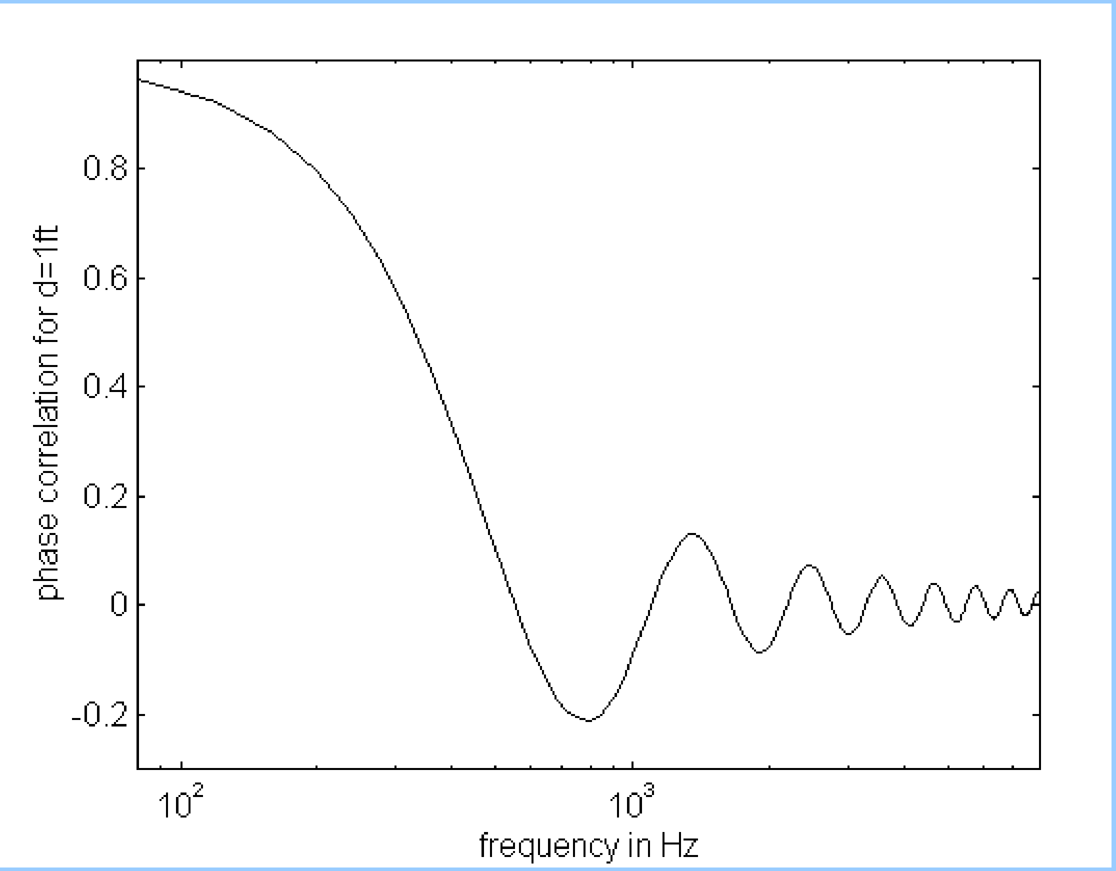


- It is NOT possible to achieve decorrelation with cardioid microphones!

Correlation through distance

- Normal ORTF technique with cardioid microphones reduces the correlation at HF by adding distance.
- But the trick does NOT work at LF,
- And LF correlation is exceedingly audible

Correlation of two omnidirectional microphones in a reverberant field as a function of microphone separation.

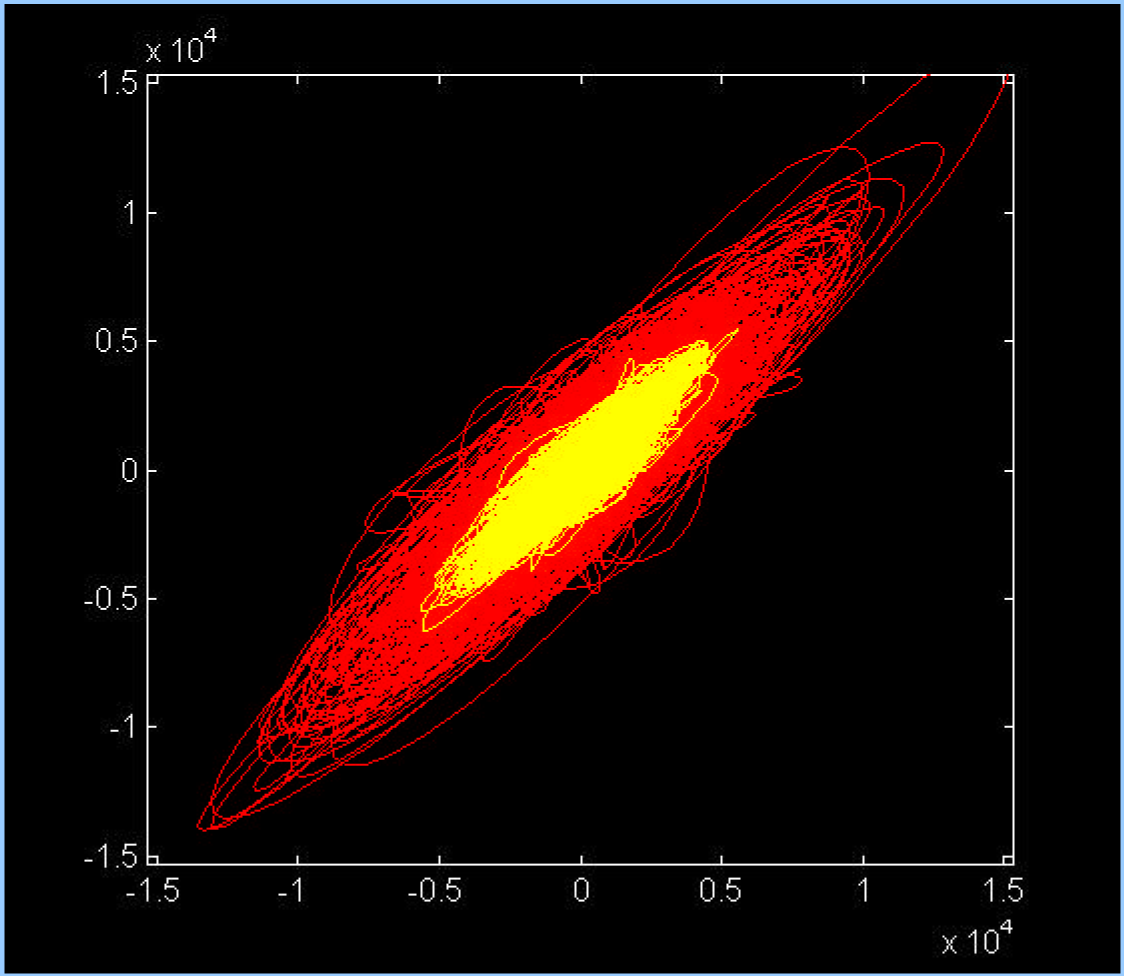


- Notice high correlation below 300Hz, and negative correlation at 800Hz.
- Frequency and distance are inversely proportional.

Audio Demos

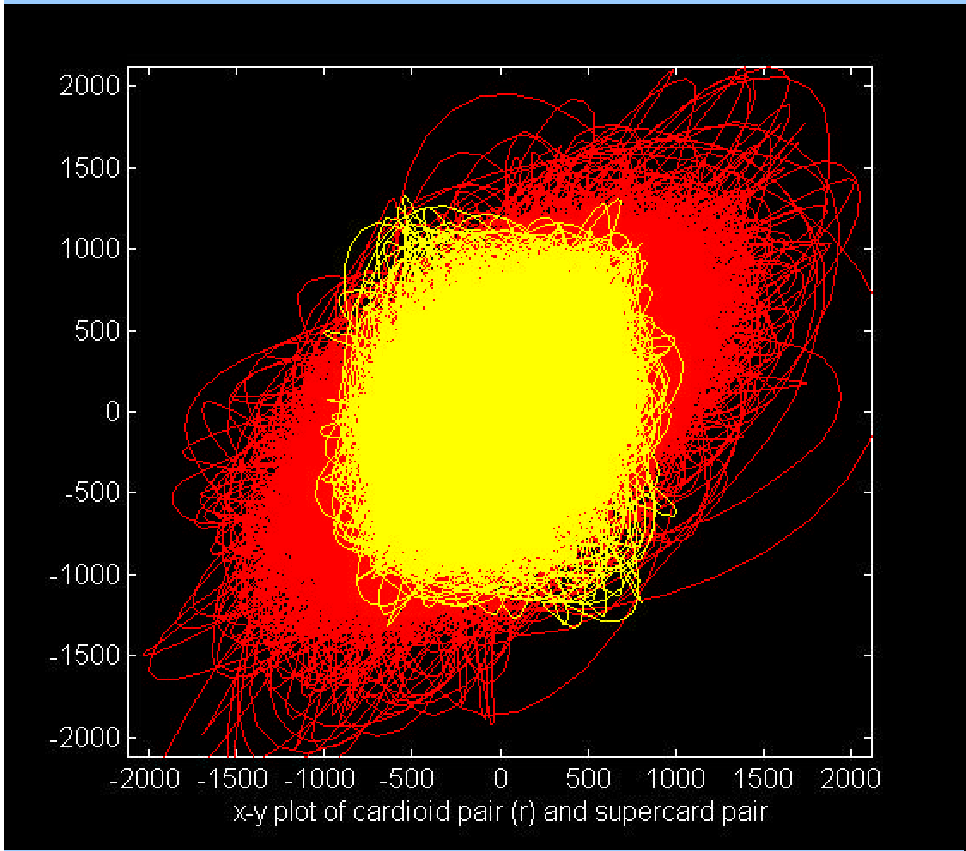
- Omni pair
  - Slavery Documents
- Cardioid Pair
  - AKG large diaphragm mikes with Oriana
- SuperCaridoid pair.
  - AKG large diaphragm mikes with Oriana
- Multimiked front image.
  - Oriana with four Schoeps Supercardioids

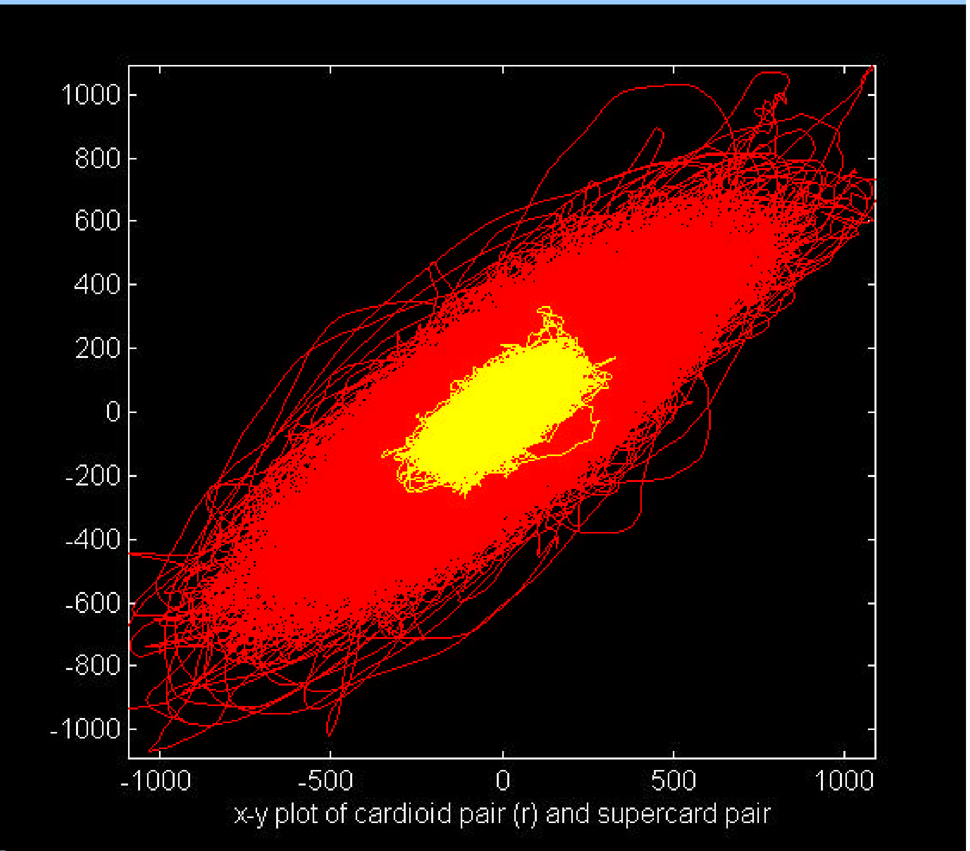
We can use a goniometer



X-Y plot of the omni front pair in Slavery Documents. Red trace is Low pass filtered at 200Hz, Yellow trace LP filtered at 100Hz.

Goniometer with AKG pair





X-Y plot of Oriana front pair with a 200Hz LP filter. Red is Cardioid, Yellow is Supercardioid

The same data, filtered at 100Hz. Note that now the supercardioid is behaving like an omni.

Revels Chorus in the Sonic Temple

Characteristics

- Main problem here was excessive reverberation level.
  - Solution was to add blankets – a LOT of them. 648ft^2
  - Here we list the measured reverberation times

Hz	blankets	empty
8000	0.6	0.9
4000	0.8	1.2
2000	0.9	1.4
1000	0.9	1.4
500	1.0	1.3
250	0.9	1.3
125	1.1	1.4
63	1.0	1.5

  - Reverb radius before the blankets: ~6 feet (2 meters)
  - Reverb radius after the blankets: ~8 feet (2.7 meters)

After the blankets

- Reverberation time drops below 1 second, the magic number for the early decay in Boston Symphony Hall
  - Recording the band is easy, as we can mike them all quite closely.
  - Recording the chorus is hard, as there are >20 singers, and we cannot get the microphones close enough to each.
    - Adding more microphones simply results in picking up more reverberation!
  - With the blankets we can record with adequate clarity using only four supercardioid microphones.
    - Once again we augment the early reflections in all outer channels using the Lexicon.
    - Late reverberation is also created using Lexicon late reverberation.

Reverberation Radius

- The reverberation radius changed from 6’ to 8’ when we added blankets. ~2dB.
  - This is not a large enough change to account for the perceived difference in sound.
- But the change in the total reflected energy in the time range of 50-150ms (the undesirable time range) is much larger: 4.5dB.
  - This is a highly significant and desirable decrease!
- The decrease in the late reverberation (150ms and greater) is 6dB.
  - But we make this back up with the Lexicon.

Conclusions

- Recording is a lot of fun!!!
- It is a great pleasure, and is often useful, to understand some of the science behind the microphones.
- Although simple techniques using microphone pairs or arrays can be seductive, a world-class sound usually requires many microphones, a lot of work, and artificial acoustic augmentation.
  - Time delay panning is undemocratic. Avoid it.
- Make SURE your reverberation is decorrelated, particularly at low frequencies.