

This is a short description of what happens in the listening room that affects our perception of the recorded music.

It is important to separate between acoustics for concert halls and listening rooms. You can also separate between acoustics for rock or choir and everything in between. The graphs below will show a very general situation.

The field of study of the ear/brain and its perception of sound is called psycho-acoustics. There are mainly three clues that tell us where a sound is coming from: Time, level differences between left and right ear together with differences in frequency content. These differences occur because of the distance difference between the ears and the fact that the head is masking the direct sound.

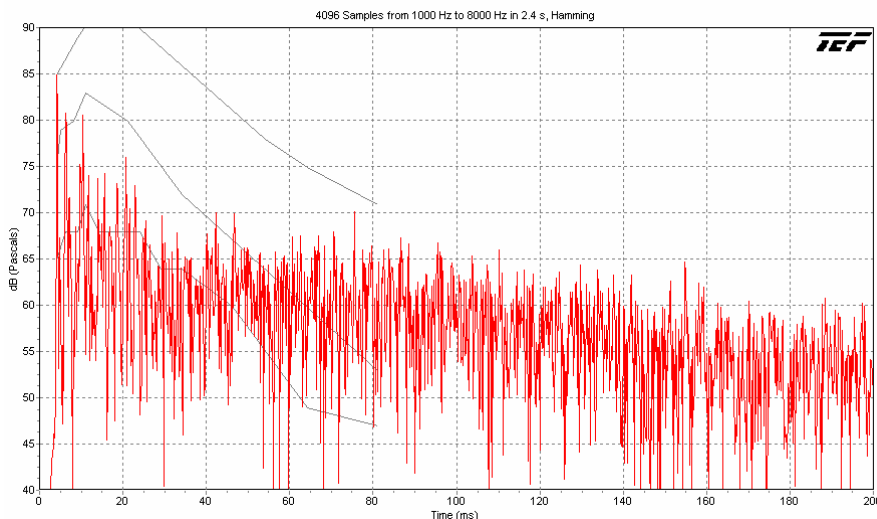
Unfortunately most recordings today should be labeled two-channel mono instead of stereo because there is very little actual stereo information recorded. There are a bunch of true stereo microphone techniques but these are rarely used in pop/rock recordings and even if you do see the use of two microphones in modern recording it doesn't mean that it is actually stereo. This limits the realism of the reproduced sound because it is the direct sound together with the reflections in the room that define to us where the sound source is placed. The most important reflections are those in the first 20 to 50 ms after the direct sound. If the clues are recorded properly the ear/brain will be tricked into believing that we do have a real room in front of us and we will be able to tell the distance, size and level the instrument in playing.

As we will see it's not only the reflections that enter the microphones that are important. The reflections in our listening room are at least as important.

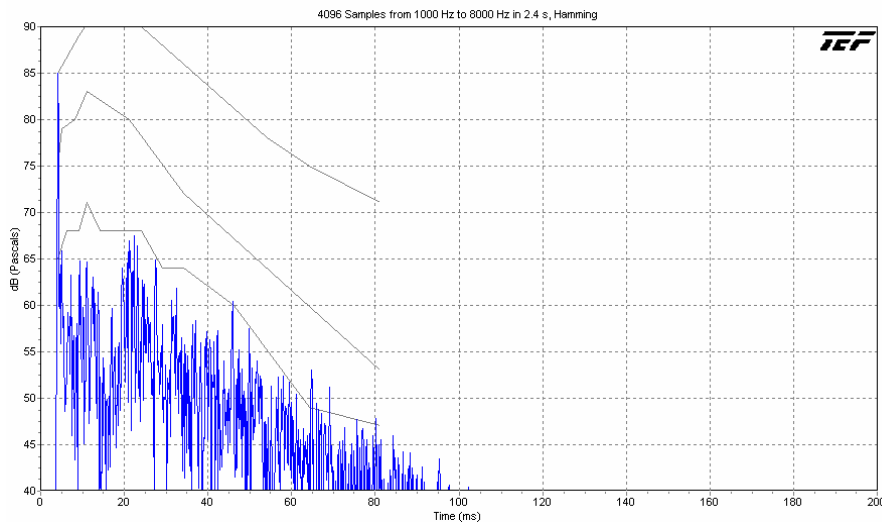
So what happens when we listen back to a recording in our room? We know that the sound contains all the important clues to trick our brain when it is reproduced by our loudspeakers but what happens next?

The graphs I'm using to illustrate this are called ETC measurements (Energy time Curve). It displays energy over time and shows us how the sound decays as it bounces around your room. The horizontal axis is Time and starts at 0 ms and ends at 200 ms on the right hand side. The vertical axis indicates dB SPL.

Graph 1 is a measurement of a relatively large recording room. This is an example of what kind of room information that could be present in a recording.



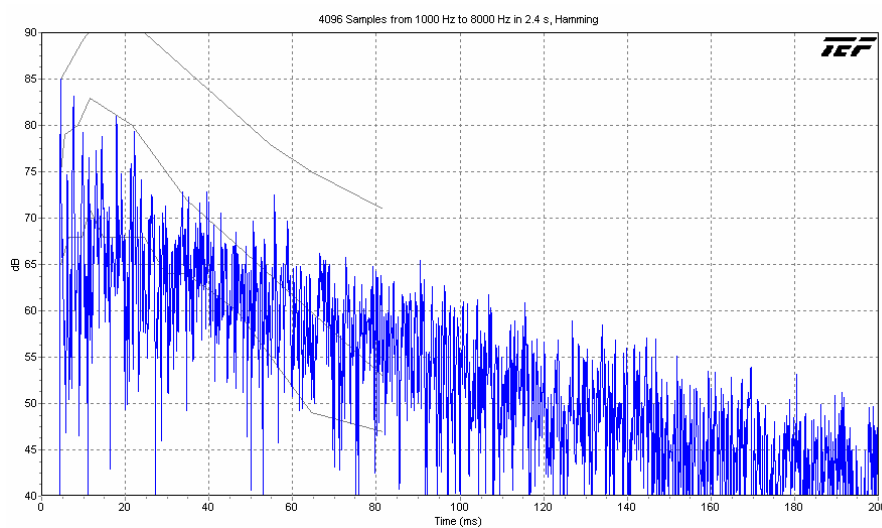
Graph 2 shows a properly treated listening room.



To illustrate the effect the graphs also contains the contours of the so-called Haas curves. If we had reflections arriving after the direct sound (the first peak on the left) these reflections would be perceived as an echo. If the reflections arrives our ears with a level over the middle gray line we would find it difficult to localize the sound source and it could also seem like it changes position depending on frequency content. This is an effect that is used more or less unconsciously in hi-fi to create a “larger than life” listening experience together with reflections with levels above the bottom gray line where we will find “ambience” and “room feeling”. Reflections in this area are very useful on a stage for acoustic performance of music. Reflections here will give a “full” sound with an added sense of “spaciousness” although it might not be very noticeable.

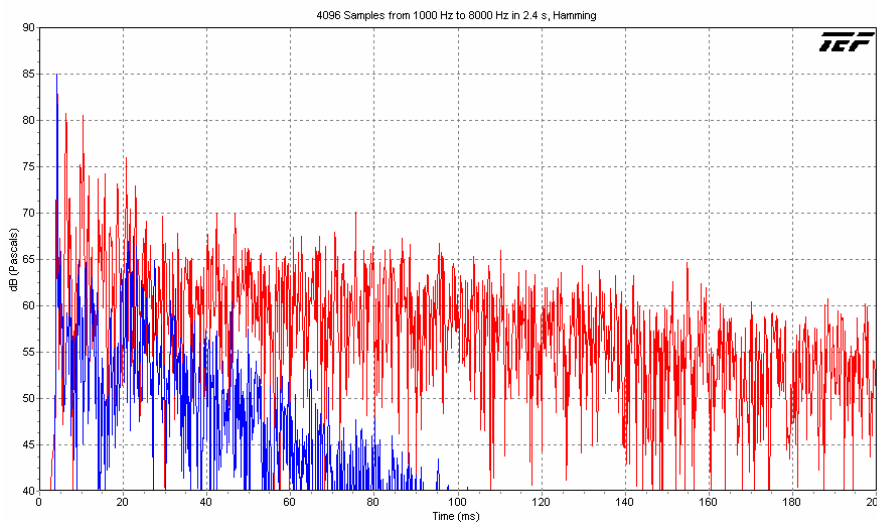
Reflections below the bottom gray line will have very little influence in our perception. Books like ”Audio Encyclopedia” and ”Master Handbook of Acoustics” are good reading with lots of info. Reflections will be inaudible if the level is -28dB in relation to the direct sound.

Graph 3 is a measurement of an untreated listening room. The room had chairs, sofas, curtains and a normal living room feel.



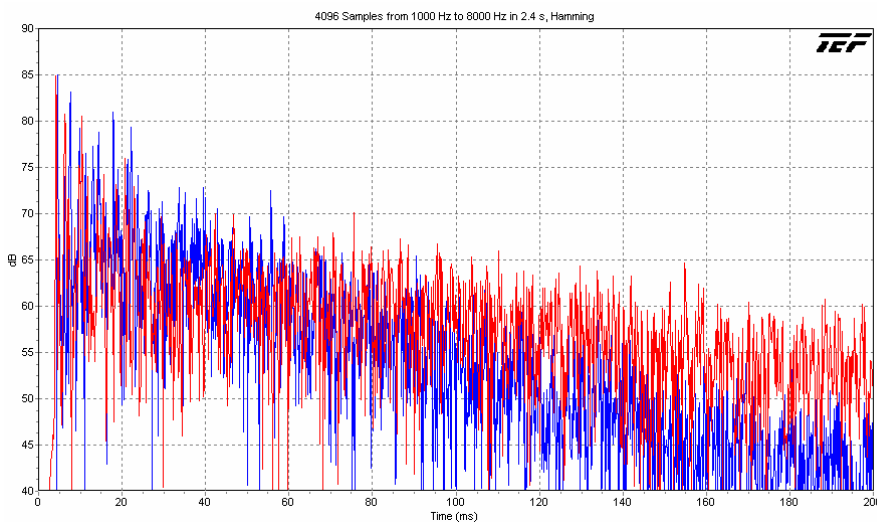
In the graphs 4 and 5 we see the situation like it actually is when we listen. We are listening to two rooms at the same time.

Graph 4 shows the treated listening room and the ETC of the recording overlaid. Even visually we don't have much problem telling what is what. The ear/brain will have little problems separating the two rooms and a clear image of the space information in the recording is what we are tricked to believe.



When we look at the graph of the untreated room and the recording we can see that they have a lot of similar information. The question is: What will our ear/brain make of it? The information is almost impossible to separate and our perception is completely confused and all the spatial clues are distorted and this makes it very difficult to tell what is what. In nature we don't find any situation where we would be listening to two rooms like this.

I should add that this might sound "nice" enough but it is not accurate and since hi-fi is all about accuracy in every other component of the audio chain why shouldn't this also apply to our room?



Usually it doesn't take that much to make significant improvement to a listening room like the one shown above.

In addition to these issues we also have frequency issues in our listening position caused by reflections. This is called comb-filter. If we don't treat the reflections correctly we will see that the manufacturer's anechoic room frequency response with no more than ± 2 dB is more like ± 10 dB (or worse) in our listening position.

I will not go further into this because if the reflections are treated properly it will not be an issue. With a lot of new DSP power available it is easy to fall into the trap of equalizing the frequency response but this will not at all help us with the issues this paper have been all about.

I also want to say that even if you listen to music with a limited amount of actual stereo information it is not wasted to treat the room correctly. Even though there are no digital reverb box capable of truly tricking the brain to believe it is listening to a real 3D space a good engineer will use his or her tools so to create a new artificial space for the music they are mixing and all this information is in there for a reason and an untreated room will mask this.

You can easily detect if there are stereo information in the recording if you can move quite a bit from left to right without the soundstage collapsing. What I call two-channel mono will only sum correctly in one place and that is directly equidistant from the loudspeakers. If you have a mono button on your pre-amp, try it and then listen to some of your CD's, you will hear that left/right positioning of instruments are gone and the reverbs with uncorrelated information will all seem to come from the phantom image directly between the speakers. If you have a "wide-sweet spot" also in mono mode you will probably find there is an excess of reflections or the loudspeaker it self might have timing issues.

© Lars Tofastrud