This note provides a description of the design and implementation of my audio system and provides the results of a few measurements. When listening to my system I tend to use several different target curves to make allowances as necessary for the characteristics of recordings. However, in my particular room, I tend to use a flat target curve most of the time, and measurement results in this note are for a nominally flat target curve (unless otherwise noted). In general I find that "optimum" target curves are very room dependent. All DSP room/speaker corrections are made using the design features of Acourate (www.acourate.com).

System Components

My primary system comprises the following components:

Speakers:

2 Quad ESL 2905 Front Left and Right speakers,
1 QUAD ESL 989 Front Center speaker,
2 QUAD ESL 989 Rear Left and Right speakers,
3 home designed & made front subwoofers, one for each of the front audio channels,

Electronics:

1 Pioneer DV-868 AVi player fitted with a dvdupgrades 6 channel SPDIF card
(https://www.dvdupgrades.ch/product/Modification/SPDIF/Output/Six_channel_S_P_DIF_output_board/24308.html)

1 Apogee BigBen
1 RME ADI-192
1 RME HDSP-9652 sound card, mounted in a dedicated audio convolution computer
4 TacT 2150 digital amplifiers

PreSonus Firebox
Matched pair of Earthworks QTC50 microphones

1 Equi=Tech 2Q transformer isolated balanced power system
2 Electronics Specialists Inc. ISO-21

Cables:

Power cables: mostly shielded computer power cables
Dedicated supply from the main electrical input panel, via Romex 8/2 copper cable.
Speaker Cables are Synergistic in-wall
Signal cables mostly Apogee Wide-Eye AES/EBU, but also some home made two conductor, twisted pair "AES/EBU" cables where I want to avoid direct intra-component ground connections.

Listening Room:

My listening room, a key system component, is a dedicated room that I designed and constructed. It is not possible to totally eliminate small room effects in any room by design features; however, very significant improvements can be implemented. A discussion of my room design rationale, goals and implementation follows at the end of this note. A final room configuration is dependent upon many compromises. You may have different goals and compromises! Most of the electronic components are located in an equipment annex adjoining the listening room.

System Configuration.

The dvdupgrades 6 channel SPDIF card in the Pioneer player provides 6 channels of digital output, using three AES/EBU 2-channel outputs (Front Left and Right channels; Front Centre and the .1 channels; Rear Left and Right channels).

The digital signal is picked-off within the player and all outputs, including the multichannel outputs, are full resolution digital (some level shifting for SACD and DVD-A). The DSD signal is converted to 88.2/24 or 176.4/24 PCM on the dvdupgrades card. The 88.2 kHz sample rate output is low-pass filtered at a nominal 38 kHz and the 176.4 kHz at a nominal 64 kHz. This conversion process is relatively straightforward since the PCM sample rates are sub-multiples of the DSD clock frequency.

Figure 1 shows an overall system block diagram.

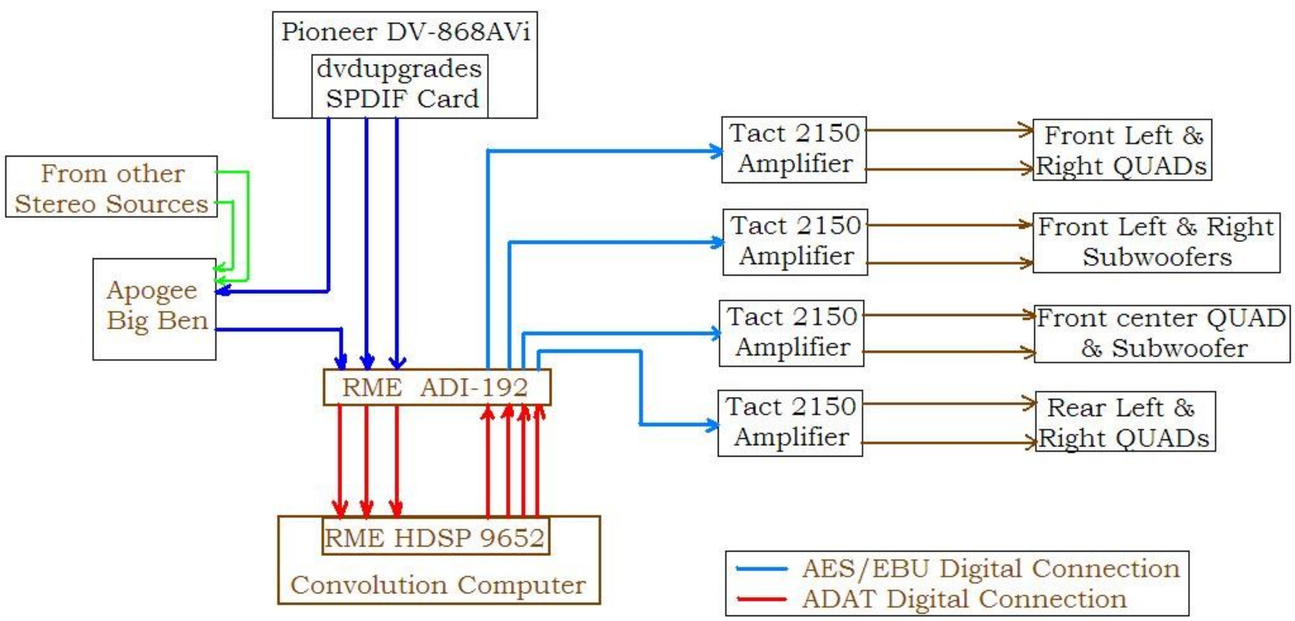

Figure 1 . System block diagram.

The Front Left and Front Right channel signals are passed to an Apogee BigBen, and then on to an RME ADI-192. The Front Centre, .1 , Rear Left and Rear Right signals are passed directly from the Pioneer to the RME ADI-192. The BigBen is used primarily as an input selector for alternative stereo inputs (such as from computers, Squeezebox, LP system – which has it's own A/D conversion, etc). The BigBen also improves the clock quality of the digital output for all inputs that go through it.

The RME ADI-192 (as I am currently using it) re-samples all input sample rates to a 96 kHz sample rate output using the RME internal clock, converts from AES/EBU to ADAT format, and passes the signals to an RME HDSP-9652 sound card in my convolution computer.

The convolution computer performs both crossover and DSP room corrections. All filter coefficients are generated using Acourate. Acourate filters provide better quality sound than anything else I have tried. The “excess phase” correction, not implemented in any other software of which I am aware, makes a substantial improvement in detail, imaging, and “realism”. The convolution computer also adds the appropriate delays to the outputs to assure intra-speaker phase alignment.

The processed ADAT outputs from the RME HDSP-9652 (note no direct ground connection between the computer and the rest of the audio system) go back to the RME ADI-192 where format conversion to AES/EBU takes place and 8 channels of output (front left QUAD, front left subwoofer, front centre QUAD, front centre subwoofer, front right QUAD, front right subwoofer, rear left QUAD, and rear right QUAD) are sent to the TacT amplifiers. The RME HDSP-9652 inputs and outputs are synchronized by the RME ADI -192 internal clock.

It is worth noting that the TacT amplifiers are a good match to the QUAD speakers. The amplifier 'volume control' adjusts the voltage level of the outputs, consequently I am able to set an upper limit to the output voltage, for a full scale digital input signal. This assures that I can not overdrive the QUAD speakers - an important consideration for me as I live at an altitude where the QUADs arc before the self-protection circuitry becomes active.

The system set-up maintains an all digital signal path (except for the passive L-C low pass filter in the output of the amplifiers). Unfortunately I am unable to slave all digital devices to a single master clock.

Measurements:

The measurement set-up comprises the Earthworks microphone(s) feeding into the Presonus Firebox converter (operated in full-duplex mode) with Firewire connection to a Gateway laptop. The laptop provides the stimulus signal output (connected via SPDIF to the system front L&R input switching BigBen) and records the microphone output. I use Acourate, Audacity, and Audition software for all system set-up , measurement and analysis activities. I use Uli's “Memory Stick” implementation of Acourate with my convolution computer. I have operating scripts that allow me to send the stimulus signals to any pair of outputs, consequently the multichannel correction measurements are fairly easy to implement.

http://www.acourate.com/

http://audacity.sourceforge.net/

http://www.adobe.com/products/audition/

All measurements were made at either 96kHz/32-bit or 48kHz/32-bit as deemed appropriate. Calibration files for the individual microphones were supplied by Earthworks. I modeled the low frequency roll-off of the microphones after discussions with Earthworks. I measured the Presonus Firebox microphone input response at both 96kHz sampling rate and 48 kHz sampling rate. The resulting convolved responses are shown in Figure 2, for one of the QTC50 microphones through the

Presonus Firebox. The slight 'droop' above about 10 kHz comes largely from the Firebox microphone preamplifiers. I should note that the Presonus Firebox is the limiting factor in the measurement arrangement, and that some measurements are limited by the electronic noise levels of the microphone preamplifiers and their sensitivity to external electrical noise– I need to get something rather better.

Figure 2. Convolved calibration for Presonus Firebox and Earthworks QTC 50 Microphone

A word of caution is necessary concerning the use of measurements when making DSP corrections. The measured in-room response to an impulse has temporal variations that can lead to results that are not optimum. For one thing it is a matter of how the impulse is gated when computing the FFT. For example, with one impulse it is sometimes possible to gate in two different ways to show either a huge suck-out in one case or a flat response in the other – this is one of the issues that makes a fully automated system difficult to implement with success for all set-ups. Care is required to attempt to make sure that the measurement processing selected is appropriate. The DSP room correction corrects to the “measurement”, and if the in-room measurement is improperly selected the result may not be satisfactory.

I use variable width windowing of the impulse response during analyses, typically varying from about 1.3 seconds at the lowest frequency to about 1ms at the highest frequency. All measurements presented in this note were made using swept-frequency techniques, as implemented in Acourate. Full-range measurements have used a log-sweep frequency sweep from 2 Hz to 24 kHz (or 48 kHz in some cases) in 30 seconds, at ½ full scale digital input. Subwoofer only measurements were swept from 2 Hz to 1000 Hz. Each measurement is convolved with the inverse of the log sweep stimulus to derive the impulse response. This approach provides both primary impulse and harmonic distortion data (see: http://www.acoustics.net/objects/pdf/article_farina02.pdf). The nature of my room allows a relatively slow sweep to be used.

FFT processing, for frequency and phase response analyses, mainly used either 131072 samples

(~0.732 Hz resolution at 96 kHz sampling rate) or 262144 samples (~0.366 Hz resolution at 96 kHz sampling rate) but, for particular purposes with the subwoofers, 524288 samples per impulse were analyzed (~0.09 Hz resolution at 48 kHz sampling rate).

I tend to experiment with a large range of options, as provided within Acourate. The results that follow are mostly from a recent experimental set-up.

In general all measurements were made centred at the listening position with a 96 kHz sample rate. Typical variations from a central measurement are shown in my document addressing the potential capabilities of a QUAD for reproducing bass frequencies - also to be found in the files section of this forum.

Figure 3 shows the characteristics of the nominal crossovers. For this particular set-up I chose Bessel filters as they provide impulses without ringing. The filters sum to a flat response throughout the crossover region.

After room correction the resulting measured impulse responses are shown in Figure 4 for the QUADs and Figure 5 for the subwoofers.

The system response is shown in Figures 6 (raw FFT output) and in Figure 7 (1/6 octave average). Figure 7 also shows the THD measurement.

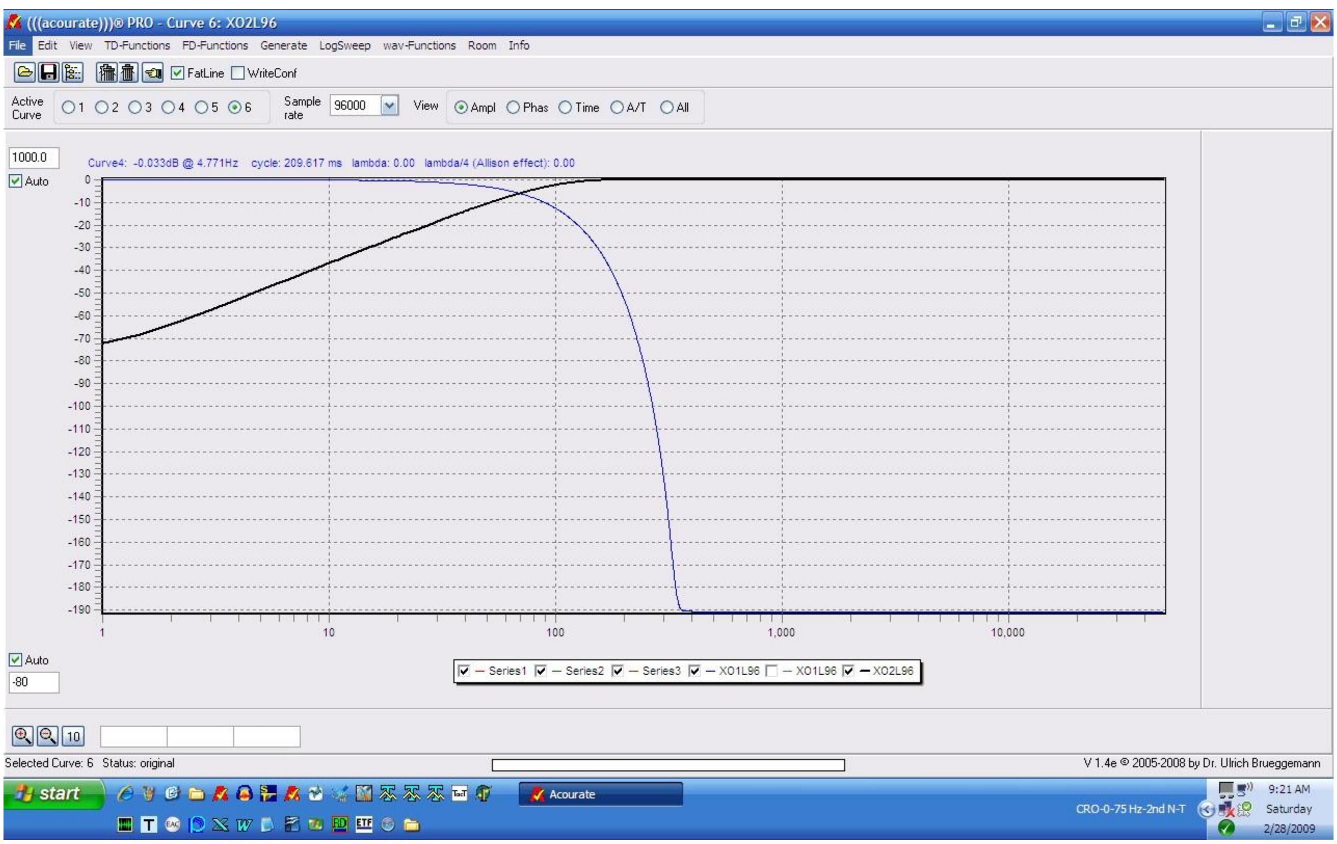

Figure 3. Crossover filter characteristics.

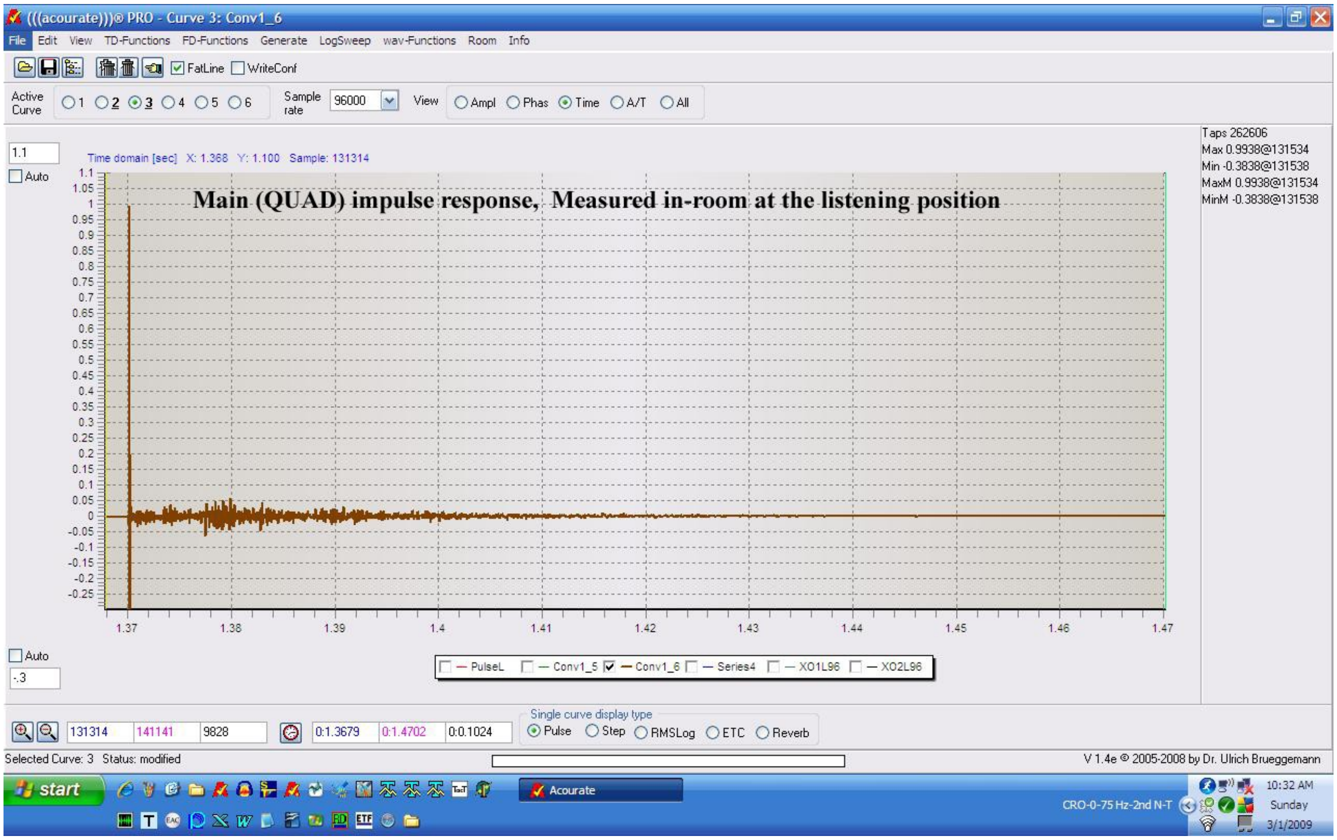

Figure 4. QUAD impulse response (first 100 ms).

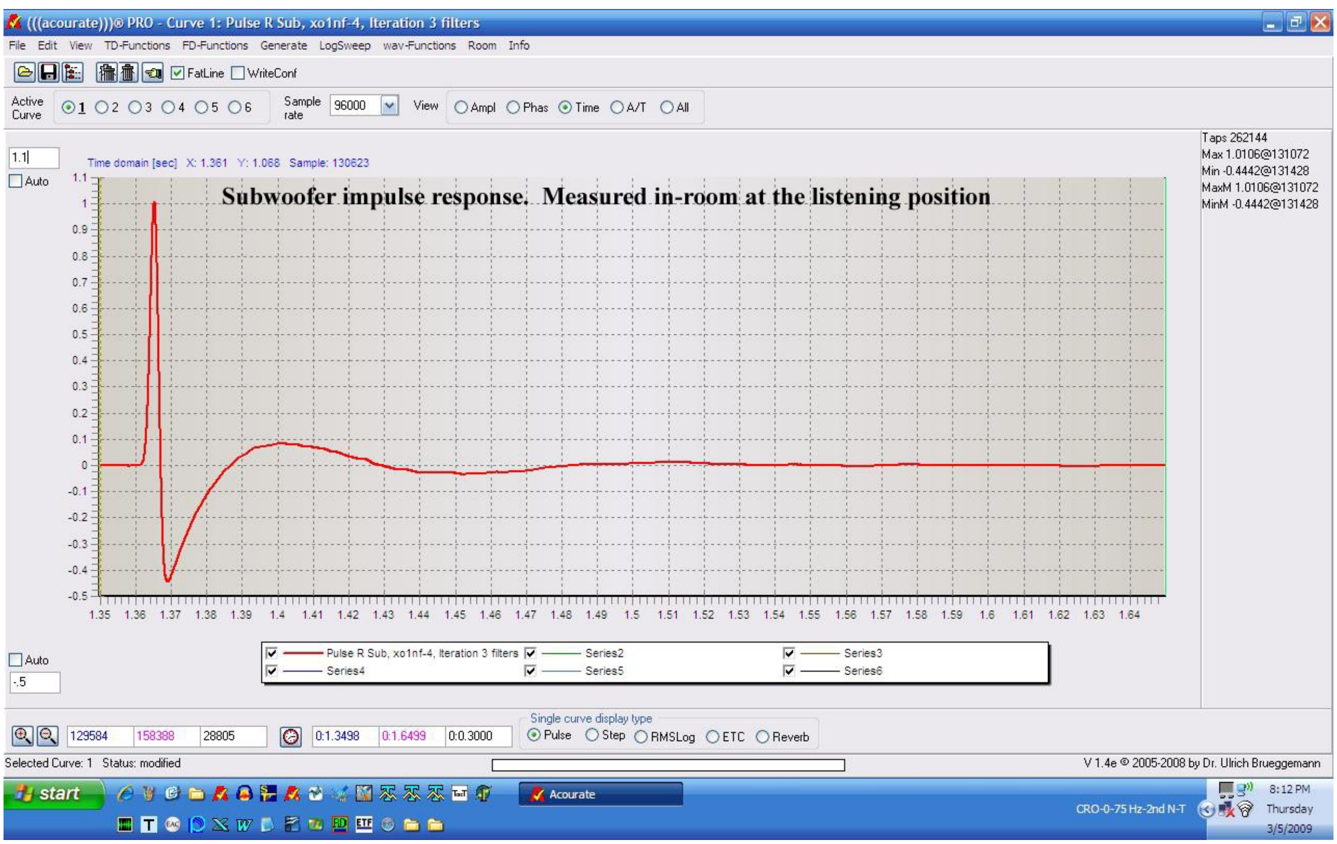

Figure 5. Subwoofer Impulse Response (first 300 ms).

Figure 6 – Typical system raw FFT output. Data points every 0.73 Hz.

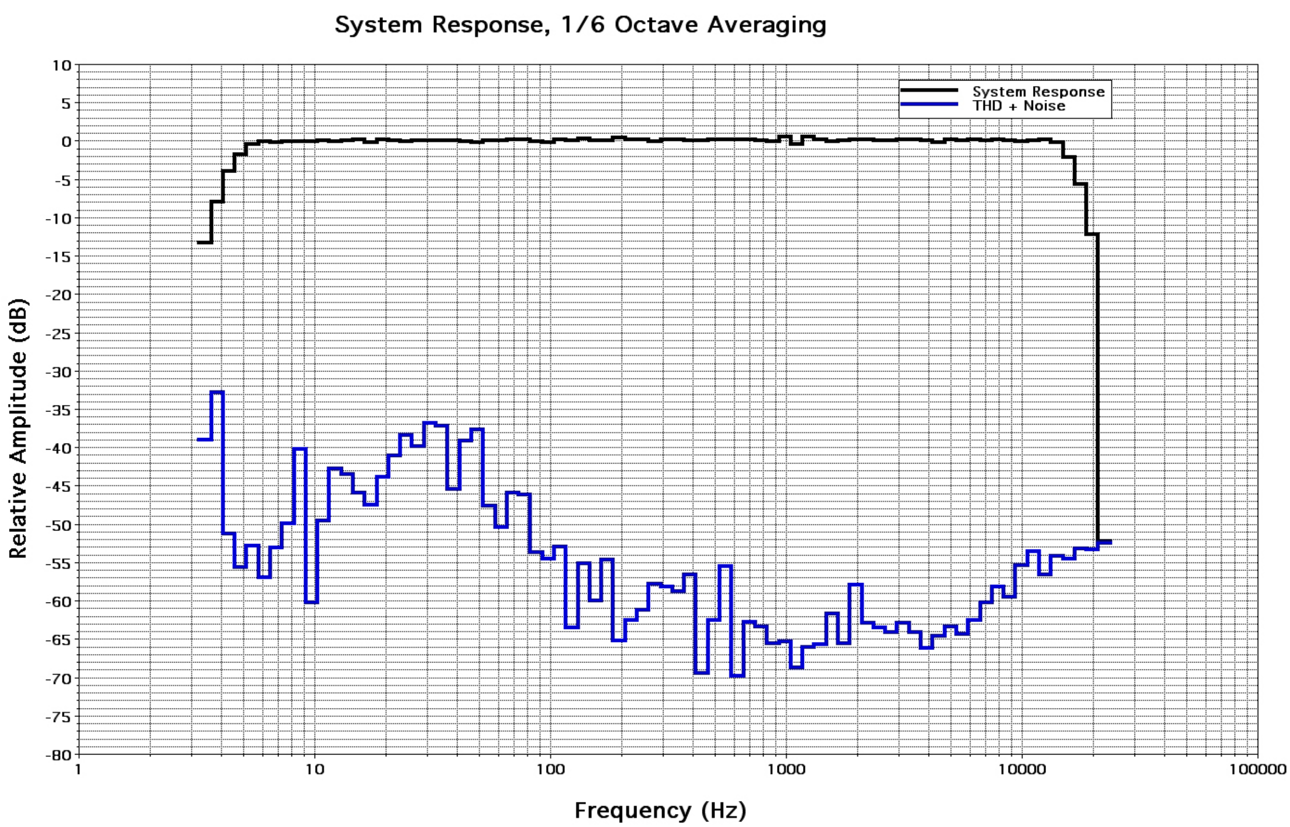

Figure 7. Typical system response and THD at the listening position – 1/6 octave averaged.

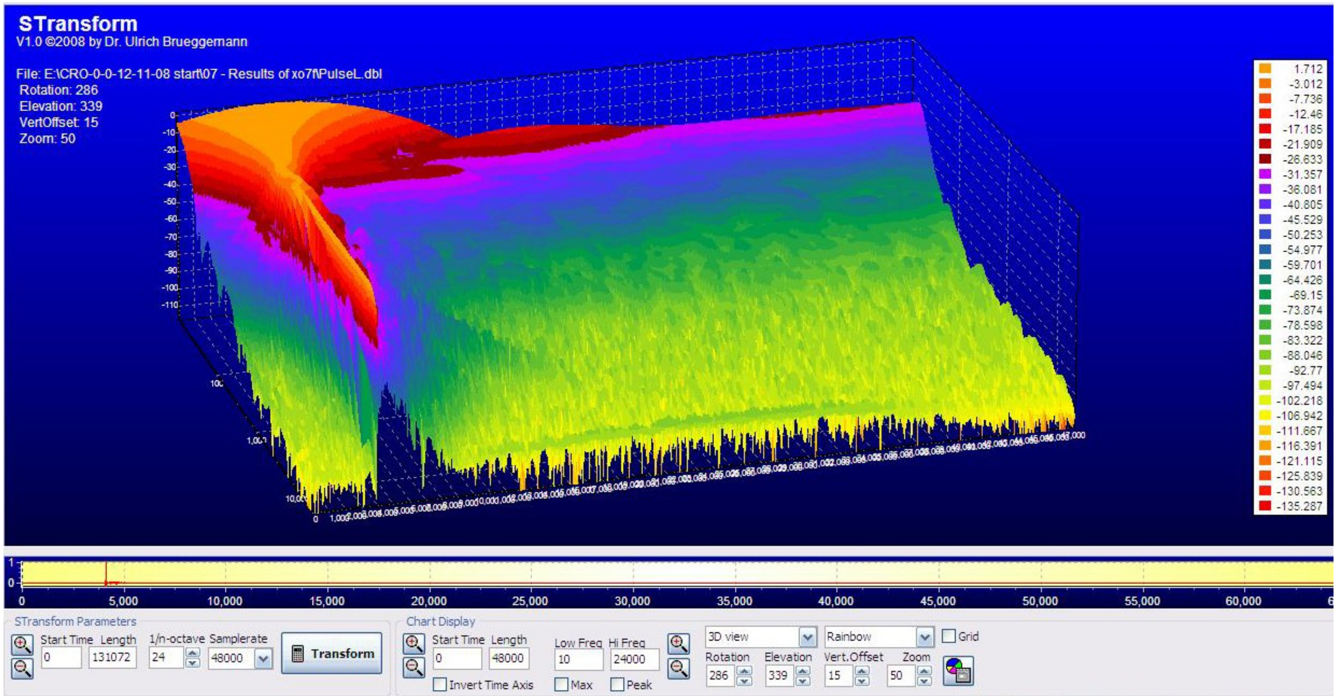

Figure 8. S Transform of system impulse response. Shows low level of time varying room effects. Time axis 0 to 1 second, frequency axis 10 to 24000 Hz, Amplitude axis 0 to -110 dB.

Figure 8, above, somewhat distorts the impulse response but shows that there are no really serious room effects to degrade the sound quality.

Listening Room Discussion:

I have been attending live concerts over a period of more than 50 years and have always wanted to achieve a sound from my audio systems that more closely approaches what I hear in a good concert hall. One often sees the statement made that it is not possible for one's audio system to sound like the real thing. Some of us try hard to get as close as we can. Others accept it as a fact, and concentrate on getting a sound that they like. This is probably why so many 'flavors' exist in 'high-end' audio.

Several years ago, with the impending release of high resolution multi-channel recordings, I decided to design and build a listening room, with the primary goal of trying to achieve a better approximation to what I hear in a concert hall, for both two channel and multi-channel sources. Put in very simple terms, my goal was to transport the listener (me) into the recording space with the orchestra, rather than to bring the orchestra into my listening room. This might initially appear to be a comparatively innocuous difference but it actually provides a fundamental change to the quality of the sound one hears in one's room.

Have you ever pondered the question: "why is it so difficult to get close to the sound of a real musical event in one's home"? There are of course many contributing reasons: interference by the recording engineers and producers, in going from the sound of the orchestra to what actually gets encoded on the discs for one; the very limited number of channels of information for another; and so on. Most parameters are out of our control. However I think that our own listening room is a significant obstacle – and **that** we can do something about!

Consider for a moment the effect of the different sizes of the concert hall and our listening room. In a good concert hall the first hall reflections arrive no sooner than about 20 milli-seconds (ms) after the direct sound. However, in a typical listening room, the room reflections start arriving as soon as about 1.5 ms after the direct sound, and a very significant quantity of reflected energy arrives from all sorts of directions within a 20 ms window. The actual acoustic environment, encoded in the recording from the hall reflections, has little chance of coming through without being swamped by the superimposition of the claustrophobic and obscuring effects of the listening room reflections. I consider that these room reflections, arriving so soon after the direct sound from the speakers, are just one of the major obstacles associated with reproducing true concert hall spaciousness in our listening rooms.

I had three very specific design goals regarding acoustic performance:

1) I wanted to attenuate all room reflections with arrival times at the listening position of less than 20ms after the direct sound. Absorbing just the first side wall reflections, which many people do, only partially meets this requirement. Absorbing some of these first reflections improves imaging, but does not significantly address the issue of too much energy arriving at the listening position well before the arrival of hall reflections in the direct sound from the speakers. It should be noted that, in attenuating all the less than 20 ms arrivals, many of the later room reflections are also attenuated to acceptable levels, so that the hall reflections on a recording can indeed be heard.

2) All enclosed spaces, irrespective of shape, have resonant modes. The conventional approach is to choose room dimensions, which spread the distribution of modes reasonably uniformly and/or position the speakers to minimize resonant mode excitation. This approach smooths the summed resonances as much as possible at the listening position. However, this does not deal with reverberation time and bass smearing issues, and it becomes an even more difficult issue when poor room dimensions are predetermined. I have attempted to mitigate room mode problems by using bass absorbers/traps, tuned to the frequencies of the room resonant modes. These comprise as many Helmholtz slot absorbers and one-quarter wave bass traps as could be conveniently accommodated. Preferential damping of room modes decreases reverberation time and amplitude anomalies for these modes – although it can't eliminate room mode effects altogether in a small room.

3) I wanted, as much as possible, to keep absorption balanced throughout the frequency range. This requires particular attention to the lower frequencies. It is so very easy to absorb in just the higher frequencies, and end up with a room with a poor frequency balance and a 'dry' sound. Consequently I needed to make all absorbers as thick as possible (within practical limits) to extend the broadband absorption as low down in frequency as practical. I used high-density fiberglass boards for absorption. These provide the best broadband absorption for the thickness, and at the lowest cost.

I constructed my audio room in the only space available to me. This was in my basement, and consequently the room dimensions were largely predetermined. The resulting finished dimensions are very far from what could be considered optimum, being a little under 25 ft long, 12 ft wide, and the maximum height is about 7ft 5 in. There is a major structural support beam and also a heating duct running across the full width of the basement ceiling, which results in the finished height for one section being only about 6ft 5in. The constraining walls in the front, right side, and back, and the floor are poured concrete. The left wall comprises separated stud construction and has a stairway to the floor above
. The initially open ceiling comprises 2"x10" joists supporting the floor above.

The first steps, after determining the basic room dimensions, were to identify the positions for absorbers. This was not an easy task to try to do by hand because of the complexity of trying to compute multiple reflection paths, consequently I wrote a computer program to trace reflections in a rectangular room. It turned out that, for the nominal configuration of my room, there are 119 different reflection paths, having a delay time of less than 20 ms. In the more critical range of less than 10 ms delay, where significant spatial imaging problems can occur, there were 25 different reflection paths. In a furnished room many of these reflections will be somewhat attenuated, but none should be ignored in the basic design. It should be noted that the number of reflection paths depends upon the room dimensions and the system configuration within the room. In general, the larger the room the fewer the number of reflection paths meeting the 20 ms criterion. The concept of reflections is rather simplistic in a small room where audio wavelengths can exceed the actual room dimensions. However, it is of use since it is reasonably applicable over much of the frequency spectrum associated with imaging specificity.

The areas to be covered by absorber were computed for a broad range of speaker and listener positions (to allow positioning changes if desired) so that all of the less than 20ms reflections struck at least one absorber surface before reaching the listener. This was a fairly straightforward iterative procedure, starting with computations without any absorbers, to determine all surface positions contributing towards the reflections. Absorbers could then be placed in the most obvious/convenient positions and the calculations re-run to allow selection of the next positions. This process was repeated until all reflections (except for the first floor reflection) struck at least one absorbing surface before reaching the listener. In general, with the use of reasonable absorbers, a single absorber reflection provides enough attenuation to drive the reflection to a sufficiently low amplitude over the frequency range of the absorber effectiveness. In the computations I also added some random forward scattering, at each reflection. Although I use QUAD ESLs, which have a directional radiation pattern, I performed the absorber placement calculations assuming omni-directional radiation characteristics.

In my particular case, about 90% of the front wall comprises 8 in thick Johns-Mansville 6 pound/cu ft “Spin-glas” fiberglass board, covered with Guilford FR701 2100 acoustical fabric. Of the remainder, about 7% is devoted to sub-woofer front baffles, and 3% to the front defining edges of Helmholtz slot absorbers, which are tuned to the room axial fundamental and second harmonic frequencies. The effective area for these Helmholtz absorbers comprises essentially the whole area of the front wall, but the actual slot width does impose some limitations.

The side absorbers are about 4 ft high x 8 ft long, mostly about 6”thick, and use 3 pound/cu ft “Spin-glas”. Absorbers stretch the full width of the ceiling more or less directly above the side absorbers (there is an offset but it is not important in a general description like this) to complete the absorbers required to deal with the reflections. Surfaces towards the front end of the room, not covered by absorbers, are considered as ‘don’t care’.

My room was kept reasonably long by the positional requirements for the rear speakers for the multi-channel set-up. Consequently the listening position is such that reflections from the back of the room at the listening position already more or less meet the 20 ms criterion and do not have to be considered. However, in situations where this is not the case, rear reflections are probably best treated with diffusion rather than absorption.

In a concert hall the sound coming from behind the listener is largely diffusive/reflective in nature, and so it is appropriate to use diffusion for the rear of the listening room. I have a large bank of ‘mid-range’ diffusers on the back wall (good diffusion characteristics over about the 250 Hz to 3.5 kHz frequency range). Also floor-to-ceiling ‘high range’ diffusers (good diffusion characteristics over about the 2.5 k Hz to 14.5 kHz frequency range) on the side walls, and also some of the ‘high range’ diffusers on the back wall and ceiling. These diffusers kill slap echo effects and contribute to the sense of spaciousness in the room. All diffusers are of the quadratic residue type. The ‘high range’ diffusers (prime=11) were cut from nominal 2”x6” planks, and the ‘mid range’ diffusers (prime =31) were built-up one well at a time with a maximum depth of a little over 15 inches.

I have already mentioned the tuned Helmholtz slot absorbers in the front wall. All of the ‘high range’ diffusers are also configured to provide tuned absorption, in addition to the diffusion. Each ‘high range’ plank is separated from it’s neighbor by a gap which, when combined with the cavity behind the plank, provides Helmholtz slot absorbers, tuned to appropriate frequencies, to preferentially damp other room resonant modes.

The view of the front of the finished room is shown in Figure 9, (there is a significant perspective anomaly resulting from the use of a wide-angle lens. The color accuracy is not that good either). The front speakers are about 5-6 feet in front of the wall behind them, and the distance from the speakers to the listening position is about 9 feet.

Figure 9. View looking towards the front of the room.

The sub-woofers are mounted in the front wall and utilize a low volume enclosure design; and are operated at frequencies below the driver/enclosure resonance. They use Lambda Acoustics 15” drivers.

The ‘high range’ diffusers can be seen on the side walls starting where the side wall absorbers end. The CD and LP racks are positioned in the ‘don’t care’ areas.

The three segments of “benches” on the floor are “Spin-glas” absorber blocks, about 12in thick and 18in high, and damp the first floor reflections. The absorption efficiency is not as high as one might think because of the wavelengths associated with the floor reflection cancellations.

The microphone stand is at the nominal listening position (the furniture has been moved to make all pictures a little clearer) and positions the calibrated measurement microphone.

The left and right QUAD ESL989 main speakers in Figure 9 have now been replaced with ESL 2905s.. They are as close to the side walls as possible because of the narrowness of the room. QUADs work better than most loudspeakers when close to a wall because of the directional characteristic of their radiation pattern. However, the proximity of the wall does modify the speaker frequency response somewhat – particularly in the lower frequencies.

The dark vertical strips in the front wall are the fronts of slot type Helmholtz absorbers, tuned to the longitudinal fundamental room resonant frequency (about 23 Hz) and it’s second harmonic.

Figure 10 shows a detail not resolved in the complete view.

Figure 10. Detail view of front wall Helmholtz absorber slots.

Considering the ceiling. All of the space between the joists has been filled with R30 fiberglass insulation and contributes to absorption in one way or another. The special emphasis in the ceiling is for bass absorption. I have added many partitions between joists to provide tuned one-quarter wave bass absorption traps, closed off on the bottom by the ceiling panels and with “Spin-glas” ceiling panels, placed appropriately for the particular tuned frequency at the ‘open’ ends. The region around the periphery of the side and rear walls with the ceiling constitutes broad-band absorption. There are reflective, absorptive and diffusive surfaces on the actual exposed ceiling.

The back end of the room is shown in Figure 11.

The back wall has high frequency diffusers, some absorption, but more importantly mid-range diffusers. The mid-range diffusers, which are fabric covered, diffuse over the frequency range of about 250Hz to 3.4kHz. Figure 12 shows details of part of the mid-range diffuser assembly.

The position of the stairs leading up from the basement is shown a little more clearly in Figure 13. One wall of the stairway is covered with a 2 inch thick layer of "Spin-glas".

Figure 11. View looking towards the back wall.

Figure 12. View of part of the mid-range diffusers, prior to fabric covering

Figure 13. Right rear corner.

RT60 reverberation times are not really meaningful in a small listening room, because a fully reverberant soundfield is never established, and the results can vary with position in the room. Results also become more suspect in the lower frequencies. However they are often quoted and I do so here as general information. The measured room properties at the listening position give decay rates that are shown in Figure 14, as equivalent RT60 values. In general the room shows well-controlled reverberation across the full frequency range.

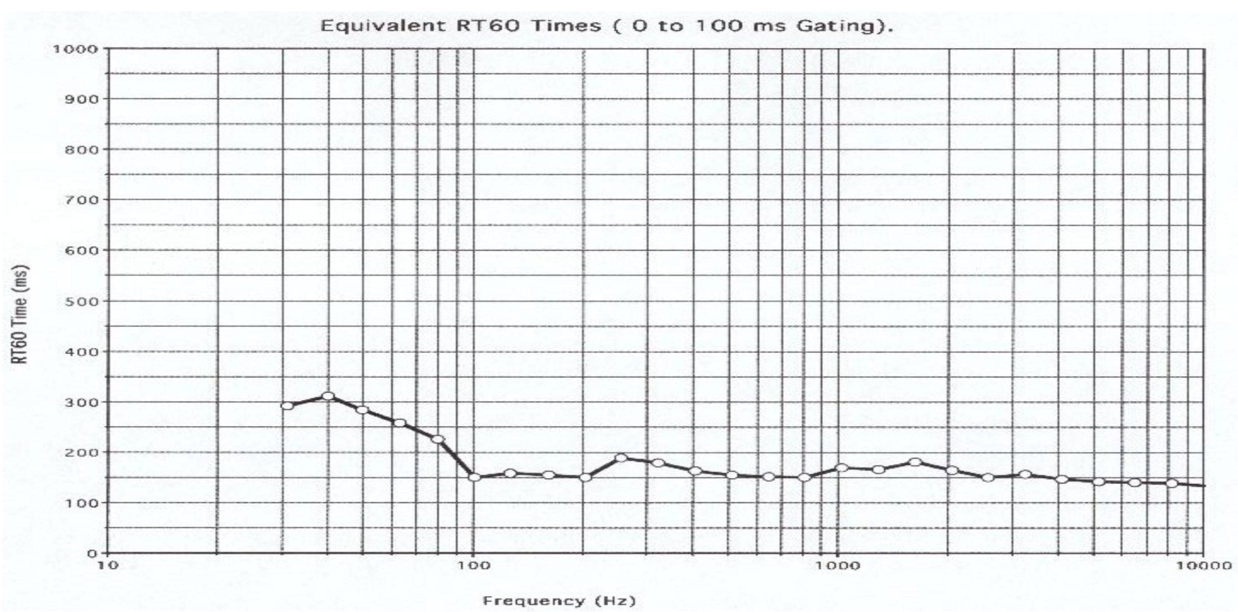

Figure 14. Equivalent RT60 measurements (ETF 5 measurements).

There is a widely held belief that low reverberation times equate to a ‘dry’ sound. However, having completed the room, I can assert that the belief is not true in rooms such as mine. It should be remembered that room reverberation adds (fortunately not linearly) to the reverberation times encoded in recordings. Consequently low reverberation times contribute towards getting closer to the recording space decay characteristics.

The intent of the room design was to try to avoid the dominating nature of room reflections, and to expose the natural open ambience of the recording space. In simple terms this was to take the listener into the concert hall rather than to bring the orchestra into the listening room. The room succeeds admirably in doing this, and in providing a satisfactory acoustical environment in which to listen. The sound is clear, detailed, sweet and dynamic at all listening levels and instruments sound more like the real things. Moving my equipment into this room has provided an improvement in sound quality far exceeding anything I have ever obtained by any change in equipment.

The listening room is usually the most neglected component of any audio system, and it really is one of the most important. My experience indicates that the return on investment, in terms of improved sound quality for the dollar, is greatest if you spend the money on your room.