THE GOLDEN EAR'S

GUIDE TO AUDIO SYSTEM SETUP AND EVALUATION

Edited by Robert Harley
and the Staff of
the absolute sound

THE GOLDEN EAR'S GUIDE TO AUDIO SYSTEM SETUP AND EVALUATION
the absolute sound

Technical Background: Acoustics

Evaluating Room Acoustics

Norman Varney, *AV RoomService*

Every room sounds unique. It is as individual as the human voice, a fingerprint, or a signature. Evaluating the acoustics of a room entails what we refer to as "characterizing" the room. As sound energy impinges on an object, the object moves. Each object transfers energy uniquely regarding the amount of energy and at what frequencies it absorbs or reflects. Room walls and furnishings interact with the original signal by means of their resonances and their reflective and absorptive properties, which become constructive and destructive to the direct signal. Sound energy, emitted by a source such as a loudspeaker, is absorbed, reflected, and scattered within the room until the energy is eventually converted to heat. Rates of energy dissipation within the room are controlled by the room size, its construction and furnishings, along with the positions of the listener and the speakers, and their amplitude. The room has the final say in the playback quality presented to our ears.

Characterizing the room can be as simple as subjectively judging it as live or dead, fast or slow, etc. From a scientific approach, it means using instrumentation to collect data from different forms of sound energy, and then comparing it to psychoacoustic criteria or a standard. The same size room will have different criteria depending on its function. For example, a conference room will require shorter reverberation times at different frequency bands than a music practice room. Some of the data collected is meaningful and some is not. There is also meaningful data that we cannot collect via today's instrumentation, such as those that define soundstage depth, image size, etc. Our ears are in most cases still our best tools for evaluations. Instrumentation just helps us to categorize, record, and compare without bias.

Psychoacoustics is still in its infancy and there is much to be learned. Quite a lot is known regarding how humans sense sound (i.e., frequency response, loudness, location, types of distortion, etc.), but less is understood concerning how we recognize and process sound cognitively and how it triggers our emotions.

Why be interested in audio? To get closer to the artistic intent for better understanding and enjoyment. In order to achieve this, timbre, soundstage, and dynamics must be accurate. Each of us has our own audio references stored in mind. They can vary in quality from person to person. People differ in their tolerance of certain types of distortion, and in preferences for bass, dynamics, perspective distance, etc. Our personal references are developed, created by motivation and established by learning, experience, and the quality we've been exposed to. The quickest way to developing quality references is by exposing yourself to "unplugged" musical events. Such events are often available at little or no costs in your local neighborhood. You may think your system is pretty great until you get a reality check from a church choir.

When I evaluate a two-channel or home-theater system, I first listen to it with recordings that I am vary familiar with. Some I've continued to use for decades. I listen for distortions to the original recorded signals, what's standing out or what's obscured. I am actually evaluating the entire electro-acoustical system, but for the sake of this article, I am listening to the influences of the room on the recording, rather than say, the influences of the amplifier. Ideally, I won't hear the room; however, there is a room in the system and its attributes are unavoidable. My ears will tell me what needs to be addressed, adjusted, or investigated via instrumentation. It then becomes a matter of prioritizing how audible the distortions are versus how addressable they are, given the many possible constraints such as budget, space, construction, décor, etc.

Audio performance design philosophy should be prioritized in the following order:

1. **Setup.** This has to do with how the listener is oriented to the loudspeakers and how both are positioned within the room. A stereo boom box is not stereo unless your head is oriented correctly in front of it.
2. **Calibration.** The signal must be adjusted to meet electrical standards and to be as pure, linear, symmetrical, and repeatable as possible.
3. **Acoustics.** Noise control and sound-quality issues are maximized.
4. **Electronics.** The remaining budget is spent on the quality of the equipment.

Though most consumers don't follow this approach, it offers the best performance value. I can't tell you the number of times I've encountered the investment of state-of-the-art electronics set up poorly in a bad environment. The system doesn't have a chance to perform to its full potential. A mid-fi system set up right in a decent room will beat it every time.

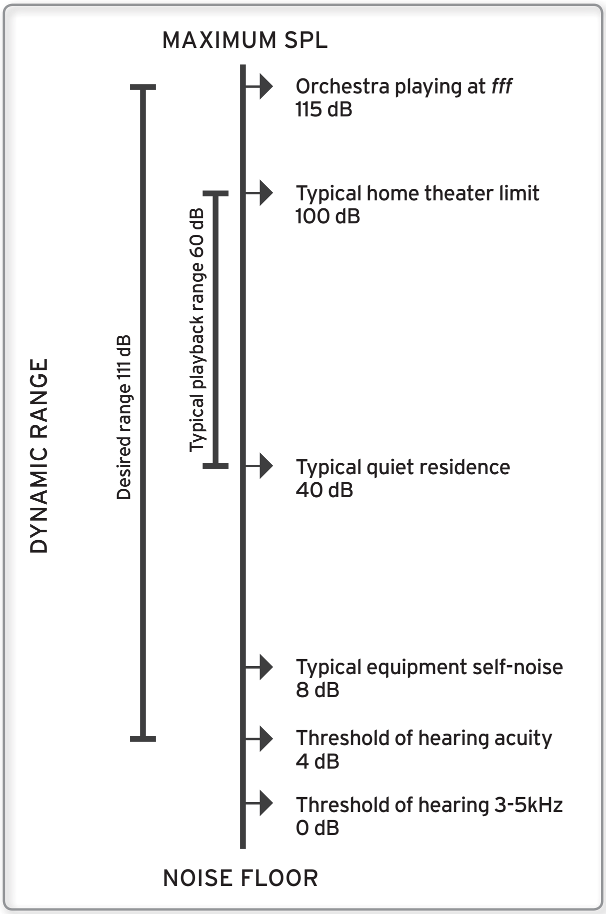


Fig. 1 This chart suggests that the usable dynamic range for listening to live acoustic music can be 111dB, and that typical residential electronic playback is only 60dB.

What to listen for when evaluating a room acoustically?

I'm not listening to the equipment *per se* (driver phase, channel balance, etc.), but rather how the room reacts to the stimulus of the equipment energizing it.

1. **Noise Floor.** Sound quality begins with noise control. Performed with the equipment powered up and the gain at reference level, the noise floor determines the possible dynamic range and the low-level resolution limits. This is about isolation and what comprises the ambient noise floor. Can you detect outside traffic, footfalls, equipment fans, HVAC, transformer hum, electronic signal-to-noise, etc.? Can you hear yourself blink? If not, you may be missing some information. A good noise floor would be a Noise Criteria Curve rating of NC 20 or below. A

typical private residence would be a NC 25–30 *without* the HVAC on. Noise can mask details in timbre, dynamics, soundstage, etc., as well as be a distraction.

2. **First Order Reflections.** Meaning the first reflection of each speaker off of each room surface. The left speaker signal will arrive at the listener's ear first directly, then likely, there will be a reflection arrival from the left wall, then the floor, then the ceiling, then the front wall, then the right wall and finally from the rear wall. That's six reflections arriving at different times behind the original and also colliding with new direct signals. Multiply that by the number of additional speakers in the system. Then add the second-order reflections, the third, etc. You can imagine how obscured and altered the original signal becomes. It turns out that if you absorb or diffuse these first reflections, you gain huge dividends in better sound quality, especially for timbre and spatial cues.

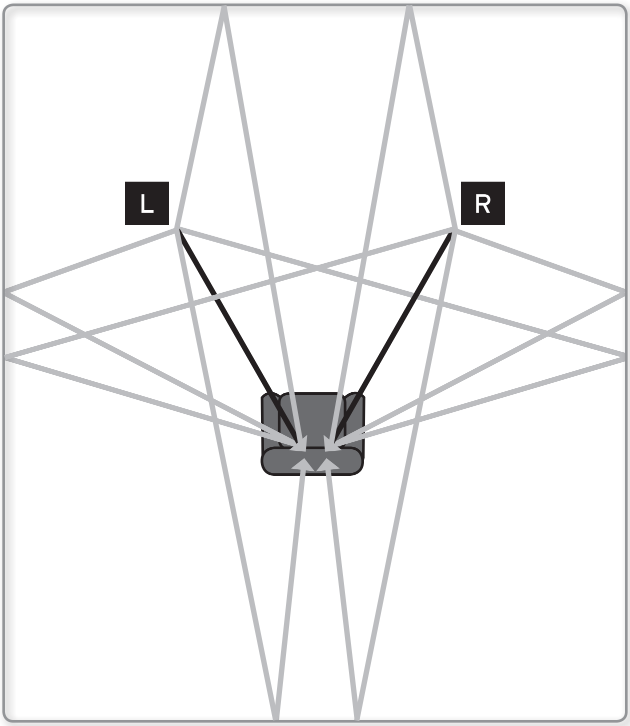


Fig. 2 Shows first order reflections for two speakers, on the walls only.
If we don't treat these 1st reflections, we would have 2nd, 3rd, etc.
Black lines indicate direct sound.

3. **Reverberation.** Probably the most recognizable characteristic of the room. People can easily understand descriptors such as "live" or "dead"—"harsh" for rooms with longer decay times in the upper octaves, and "boomy" for rooms with lingering lower octaves. These general reverberation qualities can be determined by just holding a conversation in the room. However, to really hear the impact on the recording, more bandwidth and changes in amplitude are needed. The source should be the playback system itself, and the audio observer should be located in the listening zone, principally the sweet spot. I prefer reverberation times between 0.25–0.35 seconds from about 125Hz to 20kHz, with a slightly longer decay below 125Hz. *(See Fig. 3)*. I want to hear the recorded ambience, not the room's. Reverberation can distort and mask details in timbre,

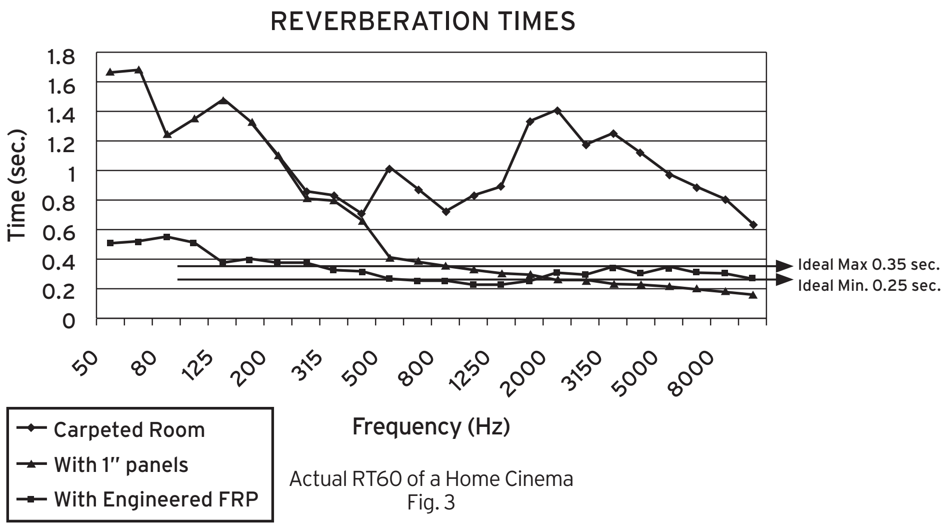
dynamics, soundstage, etc.

a. **Short Reverb.** In general, this is good, but I often run into rooms that have been treated with too much absorption in the mid and high frequency range and none below about 500Hz, leaving them spectrally unbalanced. *(See Fig. 3).* Such rooms sound and feel very unnatural and are uncomfortable to be in. Bass hang makes the low-end muddy, while everything above seems to be sucked away like a vacuum, killing dynamics due to lack of support. More amplifier power is needed, and yet the differences in amplitude are difficult to detect.

b. **Long reverb.** Typical rooms throw a blanket of lingering sound over newly introduced sounds, inhibiting their detection. Too much support also means restricted dynamics, and a room with lots of hard surfaces exacerbates the problem further. Sound builds up faster than it can be absorbed until an equilibrium is achieved between the direct sound field and the reverberant sound field. Each room has a unique density of sound energy at this point of equilibrium, which limits the possible perceived dynamic range. Obviously, reverberation distorts frequency and time, as well as energy, reeking havoc on dynamics, timbre, localization, and intelligibility.

c. **Reverberation from another area.** This is often found in multimedia rooms that have open floor plans. Hearing delayed sound returning from an adjacent room can be very annoying. Clap your hands to hear how much of a problem this may be.

d. **Linear decay.** We want neutral, natural-sounding rooms, not unique ones. We don't want to hear the Beacon Theater along with Joe's listening room. This means that the room's reverberation times must decay at a linear rate across the entire audible bandwidth. Also, it must do so in a controlled window of time. It turns out "psychoacoustically" that we need to detect some presence of room sound because we are conditioned that this is normal. When we don't, as in an anechoic room, we are uncomfortable, even ill, because the brain is alarmed by the fact that what our eyes see, compared to what our ears hear, don't equate. This phenomenon means that we must also allow for a slightly longer decay rate below around 125Hz, about B2 on the piano. *(See Fig. 3).* Note that when we are out in an open field, with no reflective surfaces, sound seems and feels normal.



Actual RT60 of a Home Cinema
Fig. 3

The top line represents the room typically furnished. Note the sonic "signature."

The next lower line represents *optimum* 1" fiberglass panels. Usually we would see more area treated, resulting in yet a steeper dive than shown here. Consequently, this treatment over damps above 500Hz and does not address below 500Hz. The lower line represents application of an engineered, broadband acoustic treatment system.

4. **Flutter echoes.** These are a series of equally spaced echoes of rapid succession, reflected geometrically between the speaker, two parallel room surfaces, and the listener.

FLUTTER ECHO IMPULSE RESPONSE

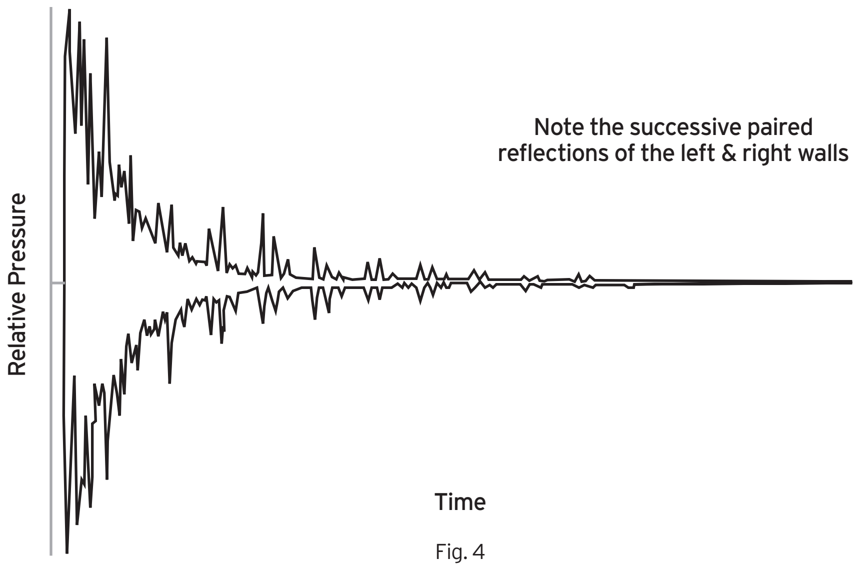


Fig. 4

5. **Room modes.** Low-frequency room resonances dictated by the length, width, and height of the room and creating high and low pressure points resulting in non-linear bass response. All rooms have modes. Rooms that offer even distribution of mode frequencies sound best. Avoid placing speakers and listeners in problematic areas, especially corners.
6. **Cavity resonances.** Just about any object, especially a three- or four-sided container of air, can become excited when its resonant frequency is played into it. Floors, walls, and ceilings are notorious for this. Just hit your fist between the studs of your wall and hear a bass drum boom in the room. Typical construction methods will store sympathetic energy and then release it back into the room later in time. This makes the room sound very slow, muddy, and nonlinear. Large membranes, like windows, can also resonate sympathetically with the soundtrack and play back into the room.
7. **Boundary effects.** These are constructive and destructive interferences which have to do with room surface reflections and resonances, and how far away the speakers and the listeners are from the boundaries of the room. The further away, the less energy transferred to the boundary and then to the listener, and therefore, the less audible the effects. These distortions are the result of energy exciting the room and then returning later in time to interfere with the direct signal

from the speakers. Ideally, the speaker/listener interaction with the room is inaudible. There are two types of boundary distortions: a) *reflections* from sound energy reflecting off of large surfaces back into the room; b) *resonances* from low-frequency standing waves (room modes) that are created by parallel room surfaces and/or from cavities of poor construction methods, or from room furnishings, all of which can produce non-linear low-frequency response. These two types of time, energy, and frequency distortions cause the following anomalies: spatial localization shifts; timbre distortions; loss in articulation; loss in dynamics.

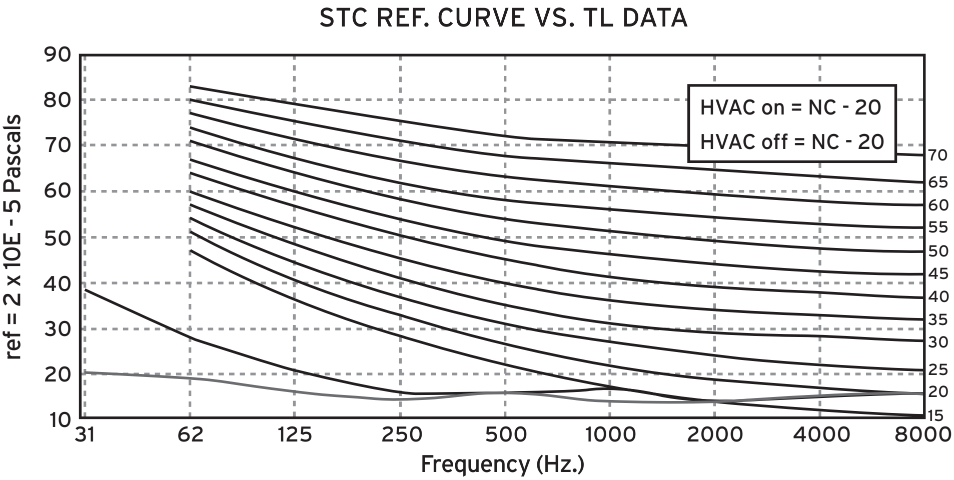
8. **Buzzes and rattles.** Annoying vibrations heard coming from energy-excited room construction or furnishings.
9. **Auditory scene.** A good soundstage presents realistic spatial cues of the width, depth, and even the height of the recording. Sound should appear from beyond the speakers and seem solid and holographic, without voids or meandering voices. Instruments should sound clearly individual in their space and easy to focus on. Instrumental voices should also sound realistic in size and weight. The location of the speakers should not be obvious. A great system set-up can be destroyed by room interferences making it sound flat, congested and muddled. I listen for:
    - **a.** Localization
        - **1.** Direction
        - **2.** Distance
    - **b.** Soundstage
        - **1.** Size
        - **2.** Solidity
    - **c.** Timbre
        - **1.** Naturalness
        - **2.** Image size
10. **Speed.** A room whose construction design has the proper balance of mass and low-frequency absorption and damping is fast and articulate. This means no muddy bass and no masking of low-level details. A room's speed opens up the dynamics of the recording and makes listening much more enjoyable. Macro- and micro-dynamics are especially important in the expression of music. Without it, music is basically melody and rhythm, and sounds very impersonal.
11. **Articulation and intelligibility.** Defined by room reverberation times and resonances, this is about definition and clarity, resolution and dynamics. Are the bass player's notes distinctive, or do they all run together? Can I understand the dialogue, or is it difficult to distinguish between words like bad, bath, bat, etc.?

Though humans are good at auditory adaptation, when the acoustic environment is well controlled, it allows the music to happen freely without all the extraneous distractions that make our brains work harder. To make sense of, or to believe in what we are hearing, the primary auditory cortex must be selective about which audio features to reduce neuron action to and which to enhance for further processing by moving the data on to the secondary audio cortex. This is where relational modeling from our aural memories occurs. When this work of suppression and filling in of data is relieved, we become more easily captivated and immersed in the music, rather than the sound. When the acoustic environment is beautiful, you'll find me in my zone, either tapping my feet or with tears in my eyes.

Using test instrumentation for evaluating room acoustics

Again, not paying attention to equipment setup but only to how the room behaves when sound energy impinges upon it, there are many acoustical tests which can be performed in order to establish the different sound quality characteristics of the room. They can then be used to evaluate performance objectives against standards, as well as to analyze problems for further investigation, and compare or confirm solutions. The test instruments are designed to measure time, frequency, and/or energy portions of sound. At this point in technology there are still many sound attributes that only our ear/brain system is good at deciphering, for example, depth of image, naturalness, etc. With some exceptions, these tests are best performed using the actual playback system as the source for generating the test signals to be recorded. This way the complete interacting playback chain is being weighed regarding performance.

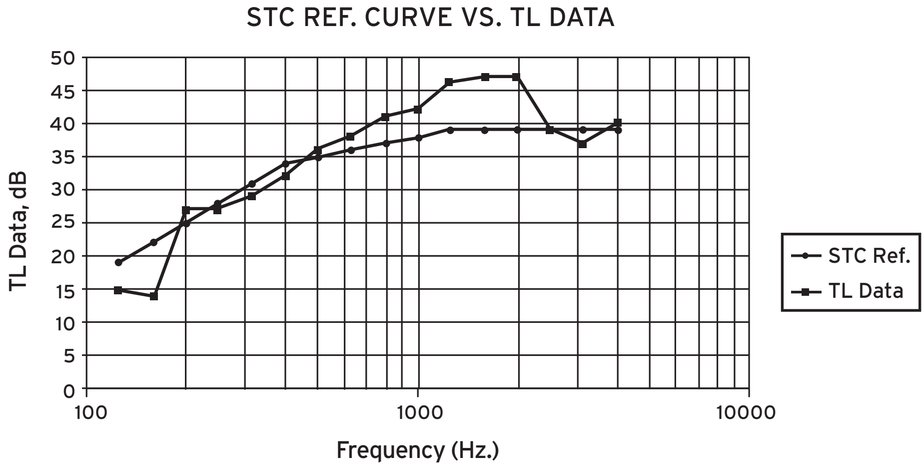
1. **Noise floor.** A method for evaluating and rating the noise floor by measuring and then plotting the results onto tables of Noise Criteria (NC) Curves. A set of spectral curves is used to obtain a single number rating describing the "noisiness" of environments. NC, or RC, is typically used to rate the relative loudness of ventilation systems, but is also handy for characterizing frequency and amplitude noise from fans, external sources, etc. The lower the NC rating, the quieter the noise floor *(See Fig. 5)*.



Recorded NC-20 curve of well designed ambient noise floor.

Fig. 5

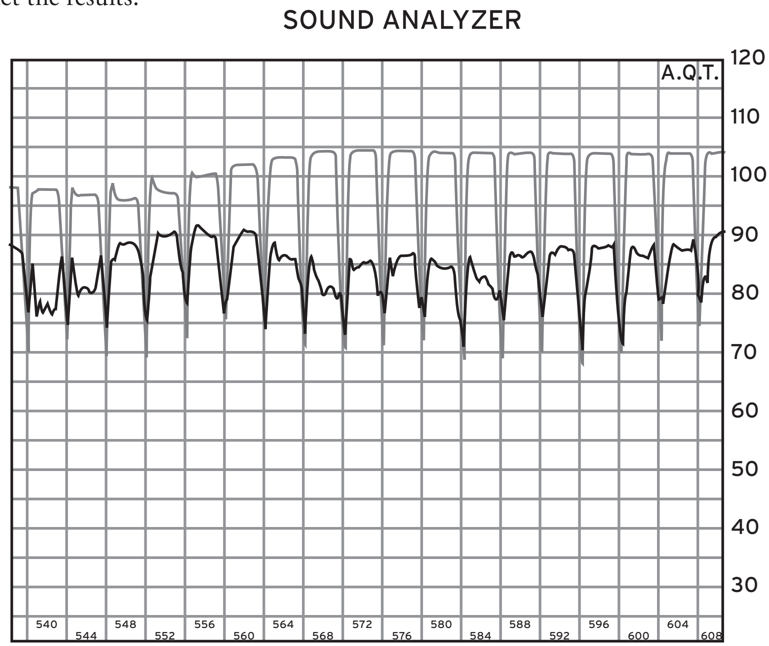
2. **Sound transmission.** A partition (wall, floor, or ceiling) does more than absorb or reflect sound energy, it transmits it as well. Because noise is a two-way street, it may be important to determine the characteristics of an existing partition in order to understand what is needed to mitigate potential problems. The Field STC test is designed to measure sound isolation between two spaces. The higher the STC rating, the better the isolation.



Standard 2 x 4 stud construction w/o insulation = STC 35.

Fig. 6

3. **Room modes/resonances.** A gated sweep from about 16-400Hz plots and reveals slurring frequencies where bass articulation is poor. Location of both the microphone and the loudspeaker impact the results.



Before and after FRP acoustical treatment

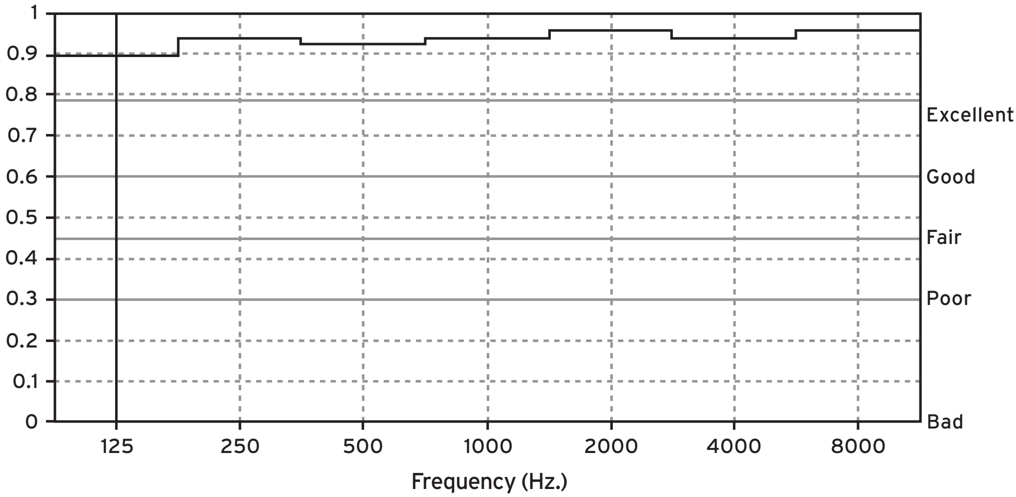
Fig. 7

4. **Pink noise.** By using band-limited pink noise, several meaningful insights into room behavior can be quickly observed. I'll use two different noise-shaped bands with energy 6dB down at the octaves. One centered around 6kHz to help determine reflection anomalies from odd-shaped areas like alcoves, and another centered around 125Hz to distinguish bass reflections.

5. **Reflections/flutter echo.** By controlling how big the time window is, it is possible to plot polar energy-time curves and discover where first-, second-, third- order reflections are located in the room from the source-surface-receiver geometric path.

6. **Reverberation times.** Energy-time curves to plot sound energy decay rates at different octave band centers. Typically, at least three decays from at least three positions in the room are averaged together to represent the room's decay times.

7. **Speech intelligibility.** Speech Transmission Index (STI) and the Rapid Speech Transmission Index (RSTI) are measurements of speech intelligibility or understandability in the listening environment.



Speech Transmission Index after FRP acoustic treatment

Fig. 8

8. **Buzzes, rattles, and vibrations.** A tone generator slowly swept manually from about 20Hz to 5kHz around 100dB SPL can reveal noises from construction and furnishings that need to be addressed.

9. **Room frequency response.** Used mainly to confirm or document what is heard. I'll use this to compare improvements with overlays. It is also helpful in determining how applicable electronic equalization may be for the room. If it is subtle enough, EQ might be the cheapest fix. However, it may reveal problems that are beyond the help of any EQ or DSP.

In conclusion, the acoustic environment dictates the potential sound quality of any playback system to such a high degree that it should be considered before the purchase of the electronic components themselves, as the investment in electronics can only be as good as the environment allows.