DRC

(Digital Room Correction)

Matthew Wrobel

Introduction

Digital room correction is an effort to employ digital filters to compensate for undesirable effects of speaker placement, room resonance, reflections, or any other coloration or "unfaithful" effect on music reproduction caused by both the listening environment and the audio equipment. The principle is simple in concept, usually involving finite impulse response (FIR) filters. Finite impulse response filters are simply a set of phase-delayed superpositions of the original signal. Acausal filters allow this delay to be negative—a sort of a "pre-echo." The assumption here is obvious: that room (and equipment) effects can be modeled simply as varied-amplitude reflections (delays). The assumption is of course dubious, given that nonlinear distortion can be present in the equipment as well as building materials that sympathetically oscillate, but not necessarily with a linear spring coefficient. At a low enough listening volume, however, one might be comfortable assuming the entire system is acting essentially linearly. Given that the room effects are considered can be described as a finite impulse response, the (exact) inverse is trivial to compute—it is just the opposite sign of the coefficients.

To ascertain just what the room response is, for a single listening position, a signal must be auditioned, and the recording compared to the original. A high quality microphone is essential. The microphone will *not* be part of the final listening system, and should therefore not factor in to the correction. (There can be a provision to supply a microphone correction curve a priori, if necessary). Each channel of playback must be auditioned separately.

Now one might think that the best signal to audition in order to establish an impulse response would indeed be an impulse (Dirac delta). But a Dirac delta is not resolved well by both the reproduction equipment and the microphone. Indeed such an impulse might drive the nonlinear modes of the equipment[1]. Instead a gentler approach yields better results. The package used here (the one found at http://drc-fir.sourceforge.net/) advocates the use of a log-sweep. The package also generates an inverse filter, which when applied to the sweep, results in an impulse.

Procedure

Picking up from where we left off above, the first step is to generate a log-sweep. The package above provides a utility *glsweep* which can generate a log sweep of variable length; varied beginning and end frequency; varied amplitude; variable lead-in/trailing silence, and variable windowing. Here a 45 second sweep from 10Hz to 21kHz with 15 seconds of leading and trailing silence, and fractional .05 and .005 lead-in and fade-out windowing, respectively. The next step is to audition the sweep through a single channel while simultaneously recording

---

[1] F. Alton Everest (2000). *Master Handbook of Acoustics* (via Wikipedia)

it. Audacity was used in this case, though any sound recording/editing software with that capability should suffice. There are packages available that attempt to bundle and automate the entire DRC-generation process, but they are of suspect quality and are not recommended here.

When glsweep was run, it not only created the sweep, but an inverse filter as well. This "inverse" is a filter that will turn the original sweep into a Dirac delta impulse. Again, the supposition that the room effects can be modeled as an FIR filter allows the algebraic equivalence of applying the "inverse" after the room effects, as to applying it beforehand (resulting in an impulse) and playing that impulse through the system (and recording it) thus obtaining the room effects. So, theoretically, an impulse with the room effects applied to it—the impulse response—can be obtained without ever playing an impulse. The utility *lsconv* will convolve the "inverse" filter with the recorded sweep to obtain the impulse response.

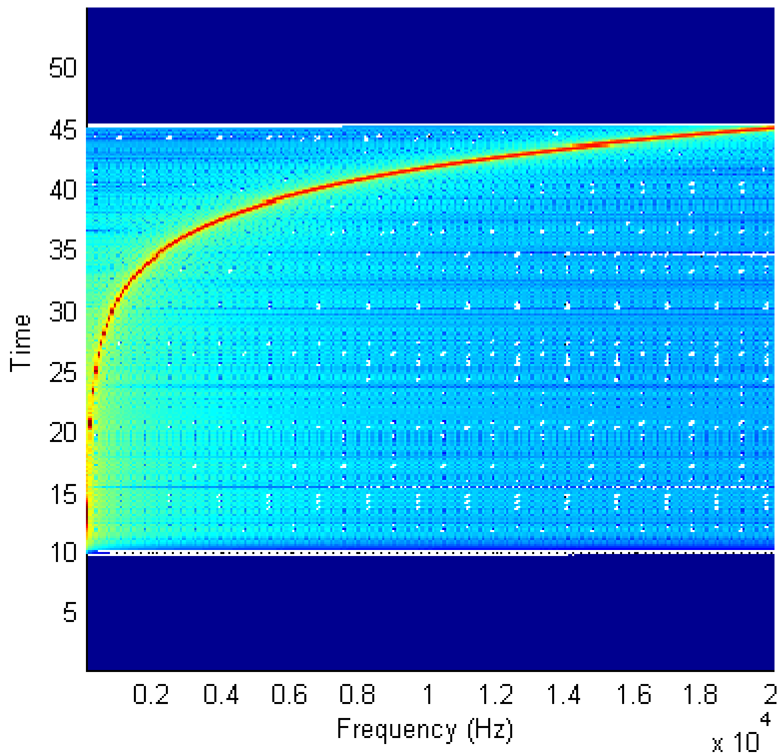
Theoretically it would be conceptually simple to generate the "inverse" filter for the room response—the one that would turn the recording back into a clean Dirac delta. Solving a large linear system (Ax=b) consisting of a matrix whose column-vectors are zero-padded phase-delayed copies of the recording, and where b is a vector corresponding to a centered delta function, could give an exact inverse. The exact filter would be undesirable as it is extremely sensitive to listening position, i.e. down to the centimeter.

The next utility, *drc*, will compute the desired corrective room response—a sort of anti-room. Rather than do an exact inverse, it will use some softening parameters and psychoacoustic effects that are not well-discussed. The aggressiveness of the correction can be tuned by editing the .drc profiles supplied. The result is a more reasonable listening sweet spot. Drc does not simply ouput the necessary filter, but applies it to an impulse, resulting in, as stated above, the anti-response of the room.

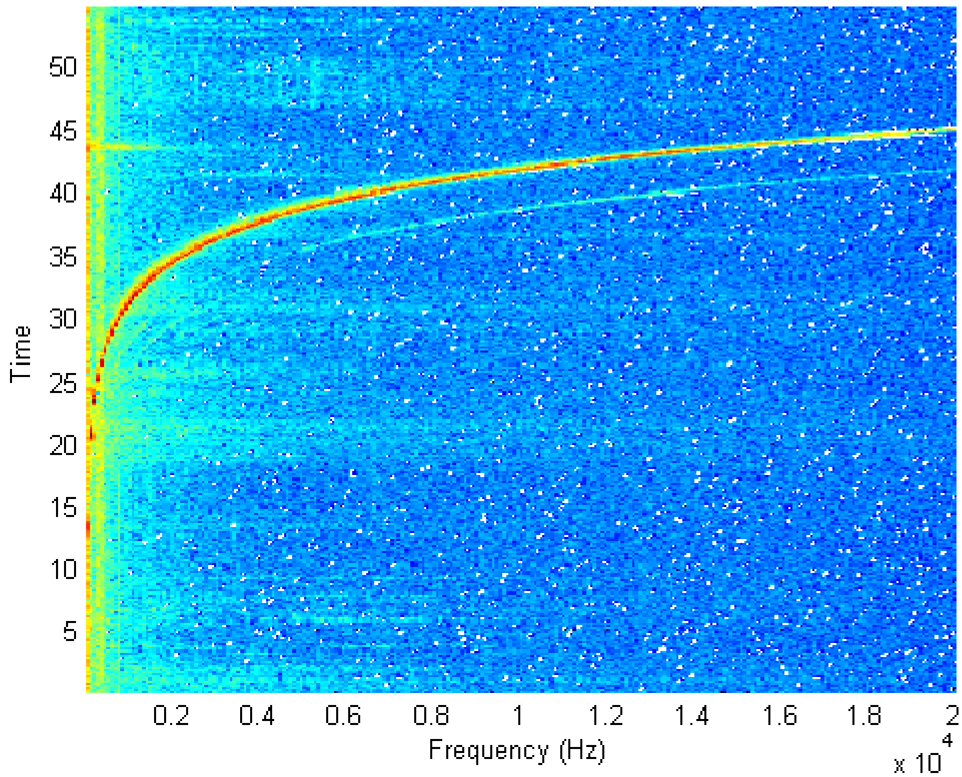
Finally a utility is necessary to apply the filter to the audio stream. This is not supplied by the above package. There are various plugins for mp3 players such as Winamp and Foobar 2000, and plugins for Windows Media, as well as VST plugins that can be used with various sound editing utilities and some sound card drivers (to apply it to the entire system-wide audio stream). In this case, Convolver (http://convolver.sourceforge.net/) was used. It supplies both a DirectShow (Windows Media Player et. al.) and a VST plugin. The VST plugin was used with Audacity to analyze and re-audition (and record) corrected signals.

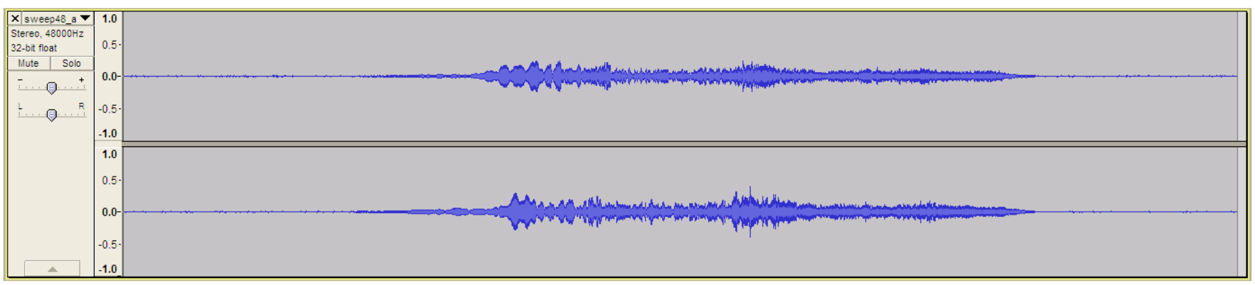
Results

The original sweep is displayed below, both in waveform and spectrogram. The spectrogram is obtained from MATLAB using the *spectrogram()* command. It will do the sequence of short-FFTs in time and generate the plot. In this case, the Goertzel algorithm was used, with 250 frequency bins, as the necessary sample-length at the 48kHz sampling rate to support 10Hz sound produces too many frequency bins for MATLAB (it crashes the plotting routine). As can be seen below, the plotting routine still has issues.



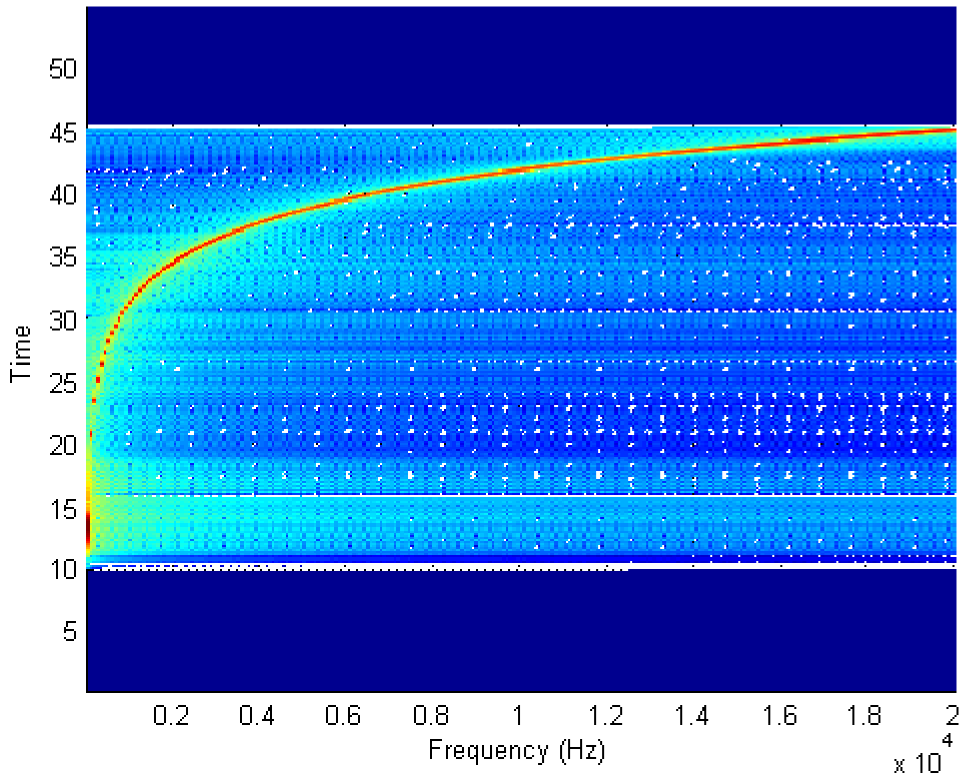
The waveform shows smooth lead-in windowing, and a consistent amplitude. The spectrogram shows very little errant frequencies as the signal sweep smoothly changes frequency. And now, the recorded audition of the above file.

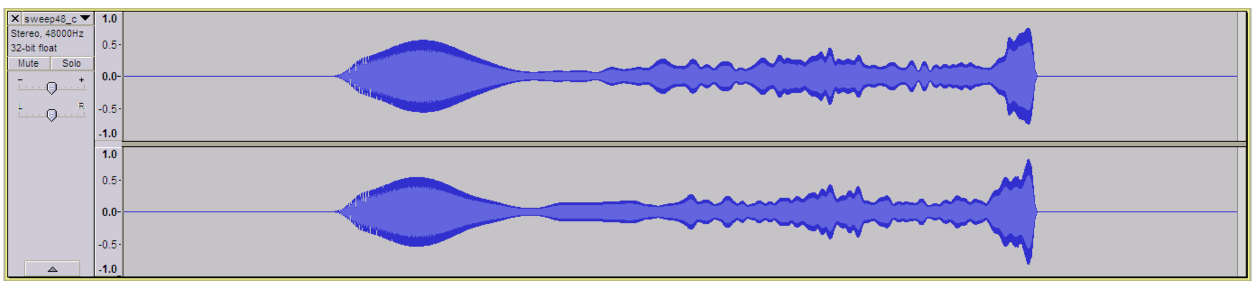




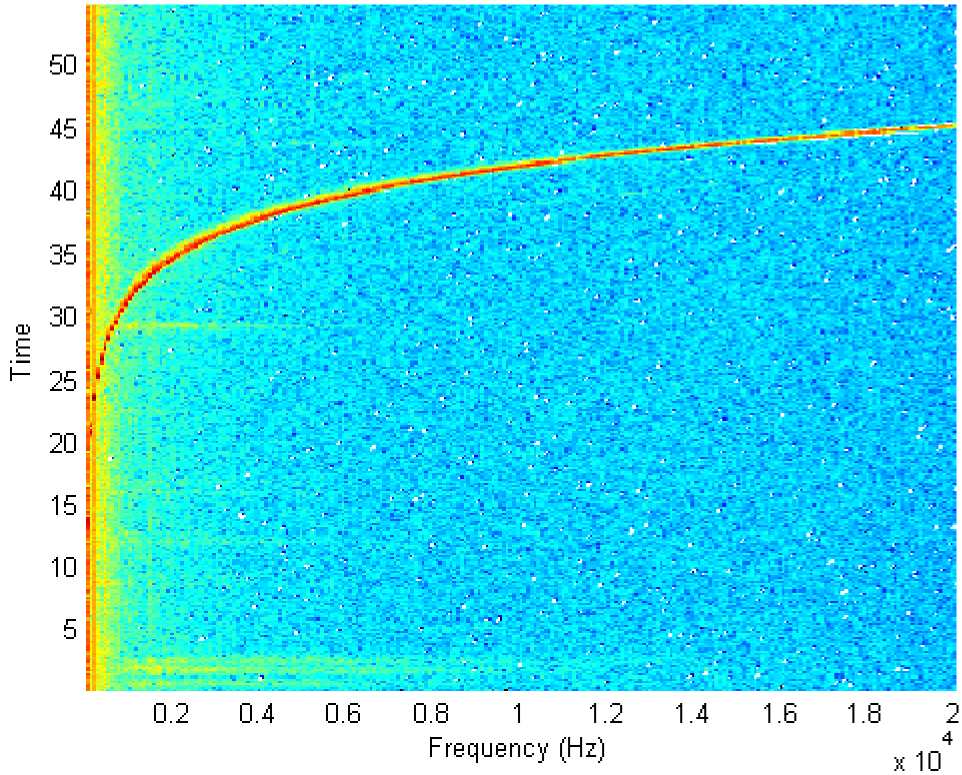
The listening environment was the square-loop hallway outside of the POM lab. The speakers were placed against the outer corner nearest to the doorway to the lab. The listening position was equidistant from both speakers, against the inner wall. First and foremost, the response is obviously not smooth as the original sweep is. There are resonance peaks. Also striking, is the difference in response from both channels. One speaker was very near to a metal cabinet, clearly and audibly capable of sympathetic vibration. Also the center axis of the entire speaker array was not symmetric with the square loop of the hallway, and was canted slightly toward north. What is also clearly present in the spectrogram (of the left channel—right is not appreciably different) is the presence of higher harmonics. The use of a logarithmic sweep makes this appear indistinguishable from a physically-impossible multiple-second pre-echo. But however they were introduced—likely mechanically from the Electro-Voice PAs—they are there. There is also constant low-frequency background noise present; perhaps it is a 60Hz bleed-through or ground-loop.

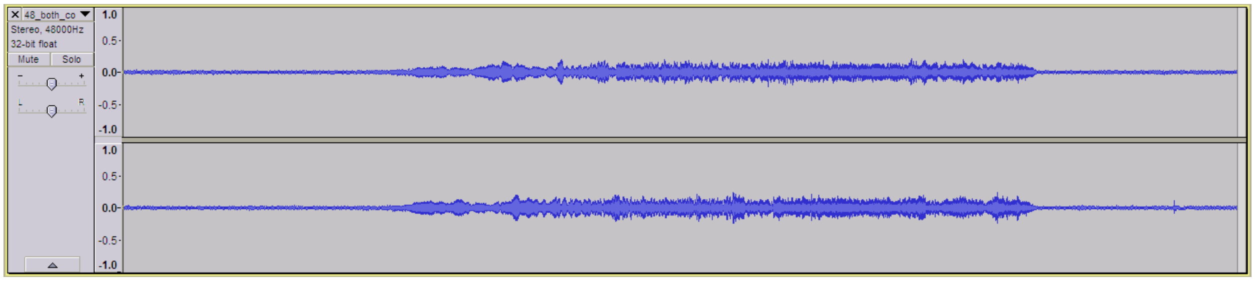
Now the correction can be applied to the sweep. The following spectrogram and waveform will be for the corrected sweep *before* auditioning.





The filtered sweep seems to be an attempt to bolster the very low (10-20Hz) frequencies which the system failed to produce and to quiet some of the resonances. There is no curve corresponding to the higher harmonics seen on the recording (there would be an out-of-phase component of similar amplitude, which would show up roughly the same since this visualization of the spectrogram does not show phase). This makes sense since FIR filters should not be able to produce a compensation for that effect. This filter also seems to act smoothly, and not sharply attack the individual resonance peaks, especially at low frequency. The next spectrogram and waveform will be the recording of the auditioned corrected sweep.





The correction has achieved better reproduction of some of the earlier low frequency content. It also shows some reduction in the room resonances. It must be pointed out, however, that the shown waveform and spectrogram are from digitally-amplified data, because the filter reduces much of the overall volume. The filter would actually need to increase the dynamic range of the recording in its attempt to achieve uniform frequency response. To avoid (digital) clipping then, for most of the content the gain is actually reduced. To compensate for this, the physical Marantz amplifier could have been turned up, but it was already at about half-volume. Aggressive DRC can damage either the woofers (especially in these ported cabinets) or tweeters at inaudible frequencies, so no attempt was made to ameliorate this effect through increased amplification. An important consequence, then, is that the overall sound generated was *quieter* and may have been less-capable of driving sympathetic vibrations or driving nonlinear vibrations. Nonetheless—although it is difficult to see in the plot—the higher harmonic is still present (for some reason the plots lost quality when they were exported.)

Qualitatively, using the streaming version of the filter on actual musical content, it is clear that the correction has an effect beyond simply turning down the volume. The best way to describe the effect is that it sounds similar to listening on modestly-priced earbuds. It produces the same flat, “clinical” sound as expensive circumaural reference headphones without obtaining the benefit of lively dynamic range and deep frequency reproduction.

Indeed, the effect may not be attractive to many listeners. Many listeners may be expecting to hear the room when they are listening to speakers, in a room. The speakers may sound more capable when allowed to ring up the room and fill it with sound. To be sure, all sense that the music was being played on two 15-inch PA cabinets was lost. This should not be seen as completely disparaging. Many speakers would have difficulty competing even with a pair of $9 Sony earbuds in terms of faithful reproduction. And to say that it falls short of reference headphones is essentially tantamount to saying that the sound is still being reproduced by speakers. Furthermore, the listening environment was terrible. The hallway has very little absorptive materials and all sorts of thin metallic surfaces to buzz and vibrate. The result is a very muddled listening experience, with fairly long reverb decay time. Under the circumstances, transforming that listening environment into something that sounds “clinical” is impressive. But still, to the primitive listener in all of us, the resultant sound may seem emasculated. Why not listen to headphones if it will sound like headphones? With a long enough cord, the headphones might afford more freedom of motion since the correction is only valid for one small spot in the room. Addressing the room acoustics directly (i.e. physically, not digitally) may be more rewarding, as would simply picking a more appropriate listening environment. Room acoustics are not necessarily disagreeable.