Audio Engineering Society

Convention Paper

Presented at the 119th Convention
2005 October 7–10 New York, New York USA

Improving Loudspeaker Transient Response with Digital Signal Processing

David W. Gunness

Loud Technologies, Inc., Whitinsville, MA, 01588, USA
David.Gunness@eaw.com

ABSTRACT

The transient response of a loudspeaker represents the combined effect of a multitude of physical behaviors. Some of these behaviors are time-variant, nonlinear, or spatially variable and are not good candidates for digital correction. Others are sufficiently LTI (linear, time-invariant) and sufficiently consistent directionally to be largely correctable with specialized digital filters. In the particular case of high powered, horn-loaded loudspeakers, most of the observed transient misbehavior is the result of stable, correctable phenomena. Consequently, the transient response of such loudspeakers can be significantly improved with signal preconditioning. Measurements demonstrate the improvements that are possible.

1. INTRODUCTION

It is a relatively simple matter to measure the frequency response of a loudspeaker at a particular point in space; and then invert the measured response to generate a complementary digital filter. With commonly available digital signal processing tools, this filter can then be implemented as an FIR (finite impulse response) preconditioning filter. However, in many cases the results are less than satisfactory.

Such a process allows nonlinear, time-variant, and spatially variant phenomena to be incorporated into the preconditioning filter. Consequently, the filter can actually makes the response worse in some directions or at output levels different from the original measurement level. The filter can also cease to be helpful as certain characteristics of the loudspeaker vary over time, with hard use, or from unit to unit.

A more sound approach is to target specific, physical behaviors of the loudspeaker which are known to be LTI (linear and time invariant) and spatially consistent. This paper addresses four, specific, correctable response aberrations which can be addressed with digital filtering. In the case of high powered, horn loaded loudspeakers, these four behaviors account for a large portion of the imperfection of the loudspeakers' transient response.

1.1. Spectrogram Displays of Transient Response

Many of the analyses in this paper will be illustrated with a unique new form of spectrogram. Because this spectrogram's time resolution varies inversely with frequency, it can appropriately be called a "wavelet spectrogram". However, a unique new algorithm allows the spectrogram to be implemented with a non-truncated Gaussian wavelet. This gives an optimum trade-off in frequency and time resolution, and is free from oscillation in either domain.

The unique characteristics of this spectrogram display are particularly well suited to the analysis of loudspeaker transient response phenomena. Temporal effects can be displayed as distinct events in time, but with the affected frequency range clearly defined.

To interpret displays of imperfect response, it is useful to have a reference - the appearance of a perfect impulse. Figure 1.1.1 is a spectrogram of a perfect impulse response. The horizontal axis represents time, and the vertical axis represents frequency on a logarithmic scale. The width of the displayed data at each frequency is indicative of the time resolution at that frequency.

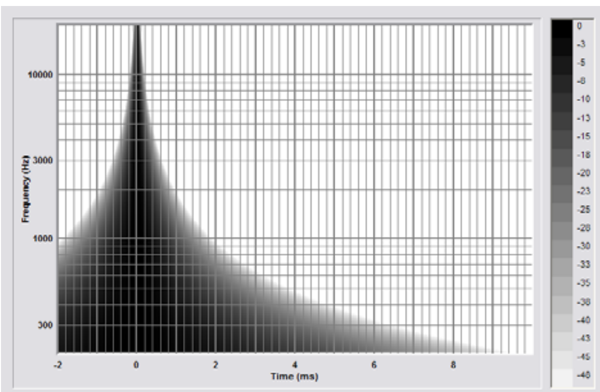

Figure 1.1.1: Spectrogram Display of a Perfect Impulse

The implementation of this new spectrogram display is described in detail in a paper by David Gunness and William Hoy [1].

2. REQUIRED ATTRIBUTES FOR DIGITAL CORRECTION

For the transfer function of a system to be reliably improved by a signal preconditioning filter, the individual subsystems should each be LTI two port systems. Both minimum phase and non-minimum phase systems can be improved.

2.1. Two-Port Systems

A two-port system may be defined as a system with one input and one output. Two port systems can be characterized by a single transfer function, which describes the output of the system as a function of its input. If this transfer function has undesirable characteristics, they can often be eliminated by introducing a signal preconditioning filter before the input to the system, or after the system's output.

Loudspeakers, in general, are not two port systems, because their transfer functions vary with both direction and distance. A preconditioning filter may improve the response in one direction, while making it worse in another.

However, many of the mechanisms that are used to construct a loudspeaker *are* two-port in nature. As a result, a preconditioning filter designed to address a specific subsystem may improve the performance of the loudspeaker in all directions, or at least over the loudspeaker's intended coverage pattern. Four specific loudspeaker mechanisms will be described, all of which satisfy the requirement of being largely correctable, two-port subsystems.

2.2 Linearity

All loudspeakers produce audible, level-dependent artifacts. Those that have a sudden onset, such as voice coil bottoming and cone collapse - can only be avoided, by some form of limiting. Those that change gradually can affect the response in a way that changes over the usable range of a device.

For example, the non-linearity of compliant elements may result in a system which grows effectively stiffer as the excursion increases. The modal behavior or "break-up" of compression driver surrounds and diaphragms may change as the stresses in the structure change; which can cause response peaks and notches to shift in frequency. Voice coils become hot, which results in reduced transduction efficiency and decreased electrical damping.

In the current context, the significance of these and many other level-dependent phenomena is that a preconditioning filter affected by one or more of these phenomena may only be effective over a narrow range of drive levels. At other levels, the filter may well make the response worse.

When developing preconditioning filters, it is important to identify and isolate any level-dependent phenomena. Their effects must either be excluded from the filters, or implemented in a level-dependent manner. Unfortunately, the practical means of accomplishing this are beyond the scope of the current paper. Suffice it to say that the four mechanisms identified in this paper are largely level independent.

2.3 Time Invariance

The characteristics of a loudspeaker can vary over time due to changes in the environment. For instance, a paper cone might increase in mass as it takes on moisture when the humidity is high. Unusually high or low temperatures can affect both the compliance and damping of suspensions, the voice coil resistance, and the strength of permanent magnets.

Characteristics can also vary over time when a loudspeaker is exposed to damaging signals. The stiffness of compliant elements may change due to incidents of high excursion. Cones and diaphragms can become weakened from repeated exposure to high mechanical force or air pressure.

To the extent that such variations are predictable, filters may be optimized for the middle of the range of variation, or for the condition that is expected to be encountered in normal use. However, if a particular mechanism cannot be corrected over a usefully broad range, pains must be taken to eliminate that mechanism's effects from the preconditioning filters.

2.4 Unit to Unit Variability

Some characteristics of loudspeakers can be produced consistently across multiple units built to a given specification. Physical dimensions in particular, such as phase plug slot spacing, can be produced very consistently. Parameters that depend on less easily controlled factors may be much more variable. Magnetic material properties, the stiffness and internal damping of metal foils, and paper cone formulation are just a few examples of sources of potential manufacturing variability.

When developing preconditioning filters for production loudspeakers, it is critical to limit the corrective filters to phenomena that aren't subject to excessive unit-to-unit variability.

2.5 Minimum Phase and Non-Minimum Phase Behaviors

A system is defined as being minimum-phase if both the system transfer function and its inverse are causal and stable. A consequence of this definition is that if a correction or preconditioning filter can be created that corrects the magnitude response of a minimum phase system (meaning the system's response doesn't go to zero at any frequency), it can also correct the phase response - yielding a perfect impulse response with no latency.

Because of this property of minimum phase systems, it has often been stated that non-minimum phase effects cannot be corrected by preconditioning filters. In fact, it is only true that a non-minimum-phase system cannot be corrected *perfectly*. However, the imperfection may simply be latency - which in audio, is an exceptionally benign imperfection.

All loudspeaker systems are subject to significant latency, because of the relatively slow propagation speed of sound in air. Therefore, a small amount of added latency is usually inconsequential. Frequently, a filter can be defined which corrects a non-minimum phase system's magnitude response, and which linearizes its phase response. The net result is a system which is perfect in the sense that its impulse response is a delta function, but in which the impulse is not located at t=0.

In other cases, a filter can be defined which approximately corrects a system's magnitude response and approximately linearizes the system's phase response, while introducing minimal latency.

In short, many non-minimum-phase phenomena are practically correctable in the context of audio applications. Crossover phase linearity is one example. It will be discussed in section 3.4.

3 CORRECTABLE LOUDSPEAKER IMPERFECTIONS

The preceding discussion established that loudspeakers exhibit many imperfect behaviors which are *not* good candidates for corrective filtering. Furthermore, each of these behaviors affects the loudspeaker's measured transfer function. So, a single transfer function measurement is not an optimum reference from which to develop a correction filter.

It is better to develop corrective filters by targeting specific subsystems which are LTI two-port subsystems. Filters to correct each of these subsystems can be developed from their measurable dimensions and parameters, or by painstakingly separating the response of the subsystem from the response of the overall system.

3.1 Compression Driver Phase Plugs

A first subsystem, which is particularly well suited to digital preconditioning, is the time smear produced by compression driver phase plugs.

The openings in a phase plug are arranged in such a way that, from any point on the diaphragm, the path to an opening is relatively short. The designer of the compression driver intends for all of the sound power produced within the driver to leave via the "nearest exit". However, a significant fraction of the sound energy arriving at a phase plug opening will either continue past it or reflect back from it; in either case arriving later at other phase plug slots where the sound is divided again, ad infinitum. Rather than a single acoustical impulse, the response exhibits a decaying sequence of impulses.

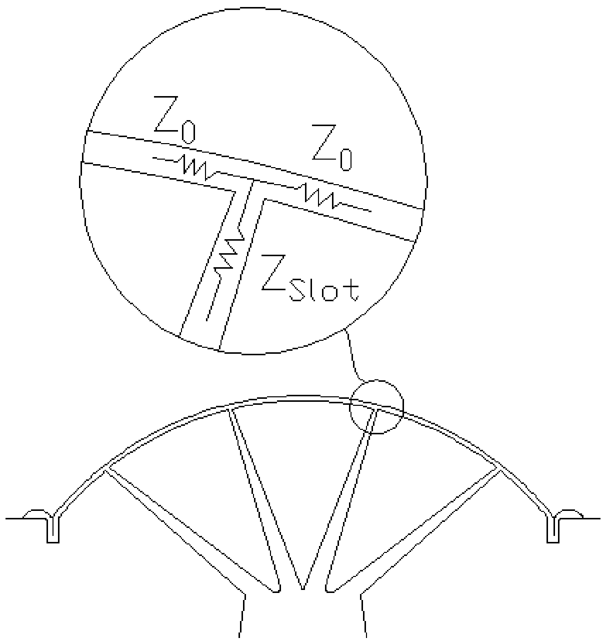

Figure 3.1.1: Schematic Compression Driver Cross-Section, Showing Transmission Line "T"

To understand why this occurs, it is helpful to view the space between the diaphragm and phase plug as a transmission line. At very high frequencies, this is a valid model. Sound propagating through this transmission line will eventually encounter a phase plug slot, which is another transmission line section that can be analyzed as a bulk termination. The characteristic acoustical impedance, Z_A, of each of these sections can be calculated using

$$Z_A = \frac{\rho_0 c}{S} \quad [2],$$

where S is the cross sectional area of the passage.

At the point that a sound wave traveling across the surface of a phase plug encounters a phase plug slot, the characteristic impedance seen by the advancing sound wave is lower, because the cross sectional area, "S", at that point is the combined areas of the phase plug slot and the phase plug to diaphragm spacing.

The reflection coefficient of a transmission line termination is:

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0},$$

and the transmission coefficient is:

$$\tau_0 = \frac{2Z_L}{Z_L + Z_0},$$

where Z_L is the characteristic impedance of the termination, and Z_0 is the characteristic impedance of the transmission line. In the current analysis, the transmission line characterized by Z_0 continues on beyond the termination, and a "tee" branches off from it (the phase plug slot). So our termination impedance is:

$$Z_L = \frac{Z_0 \cdot Z_{Slot}}{Z_0 + Z_{Slot}},$$

which is always smaller than Z_0. Consequently, no matter the slot width, an inverted wave reflects back across the phase plug. Of the power that isn't reflected back, part enters the phase plug slot and part continues on across the phase plug in the same direction.

Interestingly, the maximum power transfer into a transmission line branch occurs when $Z_L = Z_0/2$. In a compression driver, this occurs when the slot width is twice the phase plug spacing. One fourth of the power continues on, and ¼ of the power reflects back with inverted polarity. If the slot is narrower than twice the phase plug spacing, then more of the sound power continues on past. If the slot is wider than twice the phase plug spacing, then more of the power is reflected back. The sound power delivered into the slot never exceeds ½ the power of the wave arriving at the tee.

This analysis demonstrates why all phase plug designs produce significant smearing of the transient response. An example will serve to illustrate both the transient behavior, and the efficacy of preconditioning filters. Figure 3.1.2 is a spectrogram display of the axial response of an example compression driver.

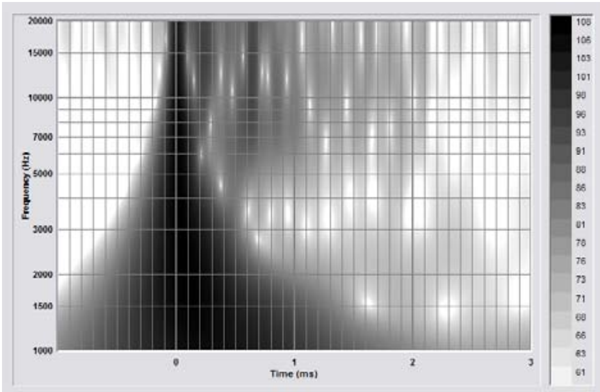

Figure 3.1.2: Spectrogram Displaying Phase Plug Transient Smear - Axial Response, 50 dB Scale

It is a relatively simple matter to invert the transfer function and eliminate the transient smear entirely. However, such a filter would not be effective in other directions. Figures 3.1.3 and 3.1.4 illustrate. With an inverted transfer function as a preconditioning filter, the impulse response can be nearly perfected (figure 3.1.3).

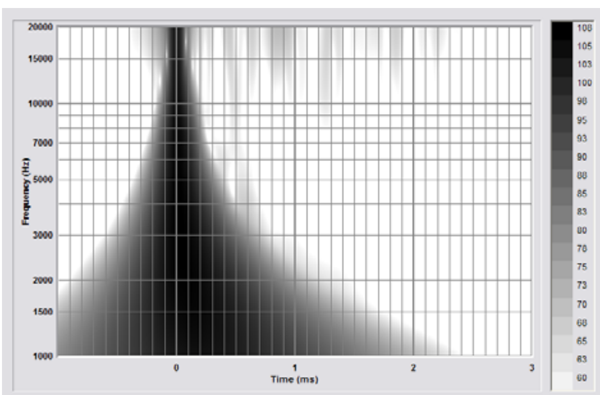

Figure 3.1.3: Axial Response, With Inverted Transfer Function as Preconditioning Filter

However, when this same filter is assessed at 12 degrees off axis, the results are not nearly as satisfying. The spectrogram in figure 3.1.4 displays transient behavior no better than the uncorrected axial response. The reason for the disparity is that the axial measurement includes transient effects that are not two-port in nature. As the direction is changed, the timing and intensity of these effects shifts, such that the axial correction filter is no longer effective.

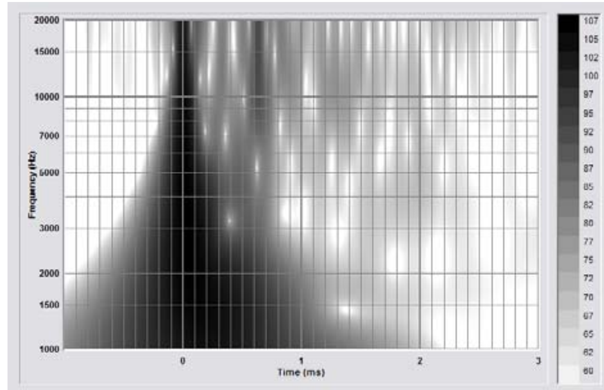

Figure 3.1.4: 12 degrees Off Axis Response, with Inverted Axial Response as Preconditioning Filter

A more carefully constructed filter falls short of perfectly correcting the axial response, but it improves it significantly, while also significantly improving the off axis response. Figures 3.1.5 and 3.1.6 illustrate.

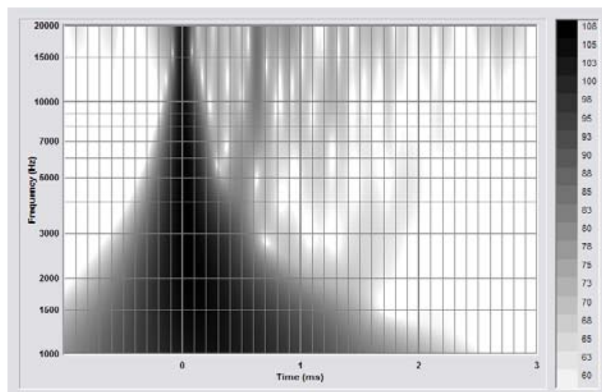

Figure 3.1.5: Axial Response, With Two-Port Correction Only.

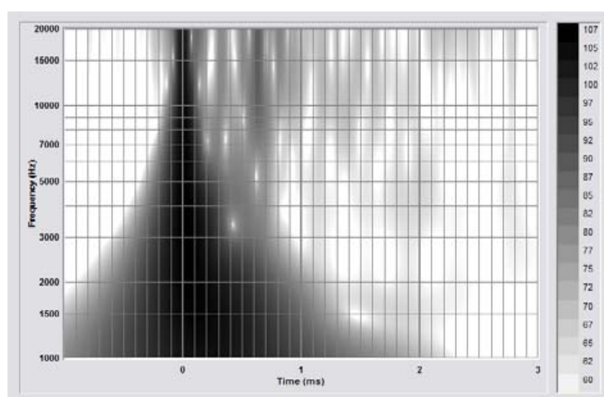

Figure 3.1.6: Off Axis Response, With Two-Port Correction Only

While neither response is as perfect as the simplistically corrected axial response, both show significant improvement. An interesting detail is a somewhat distinct arrival at .6 ms. This particular detail is enclosure edge diffraction - a decidedly non-two-port phenomenon.

If the exit of the phase plug where the various paths converge were a single point, then the system would be a true two-port system. It is not and there is some directional variability in the high frequency response of most compression drivers.

However, carefully constructed phase plug correction filters have been found to improve the transient response everywhere in the coverage pattern of selected horns, with somewhat more improvement in the center of the pattern. With such a preconditioning filter implemented, these loudspeakers exhibit a greatly improved ability to render high frequency detail, an effect that experienced listeners have found easily perceptible.

3.2 Horn Resonance

A second loudspeaker behavior, which yields well to digital preconditioning, is horn resonance. A wavefront progressing down any horn will encounter one or more discontinuities in the area expansion. All horns present a discontinuity at their mouths. Constant directivity horns often employ a diffraction slot to achieve a wide coverage pattern at high frequencies. The exit of this slot represents a severe discontinuity.

A discontinuity in a horn's expansion produces a reflection. A fraction of the sound power reverses course and returns to the compression driver where it is partially absorbed and partially re-emitted, often several milliseconds late. This process is, of course, regenerative, once again producing a decaying series of arrivals. Low frequencies tend to reflect more strongly than high frequencies, so the reflections are most prevalent in the lowest octaves of the horn's usable range. It is this precise phenomenon that produces an audible artifact commonly described as a "honk".

The wavefront does not return to the compression driver in a perfectly coherent fashion, but it appears that the bulk of the reflected energy does converge back at the driver. To the extent that it does, the phenomenon acts as a two-port system, and is correctable via signal preconditioning.

The spectrogram in figure 3.2.1 displays the transient response of a typical high frequency horn. In particular, note the relatively slow decay between 500 Hz and 7 kHz. This characteristic can be greatly reduced with a signal preconditioning filter. The same system with a preconditioning filter applied is shown in figure 3.2.2.

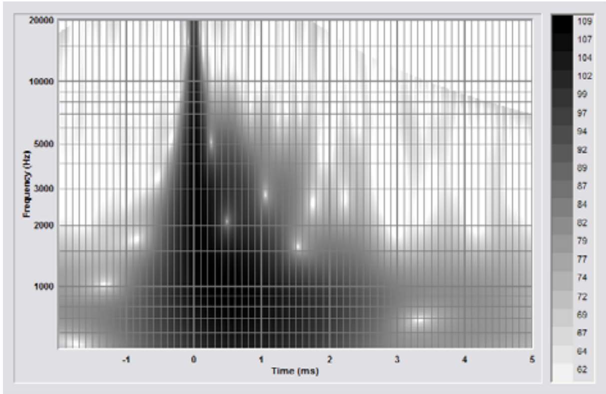

Figure3.2.1: Spectrogram Displaying Horn Resonance, 50 dB Scale

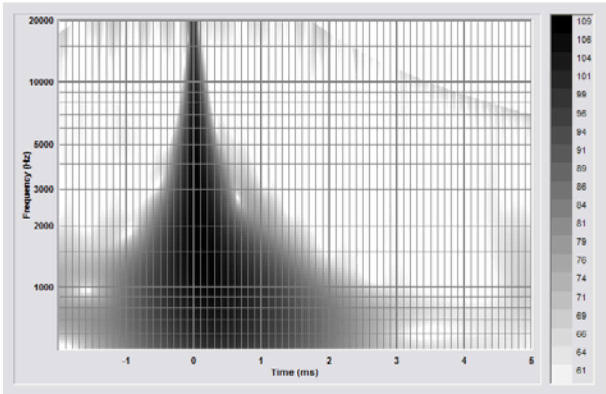

Figure 3.2.2: Spectrogram Displaying Response of Same Horn, with Preconditioning Filter Applied

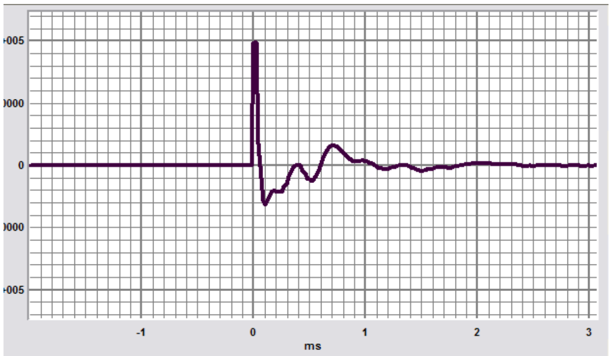

Figure 3.2.3: Impulse Response Displaying Horn Resonance

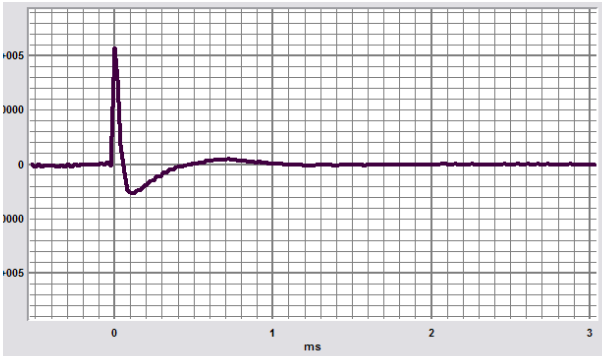

Figure 3.2.4: Impulse Response Displaying Response of Same Horn, with Preconditioning Filter Applied

The improvement in the transient decay is very apparent in the data, and listening tests have verified that the specified filter greatly reduces the colorations listeners describe as "honk".

3.3 Cone Resonance

A third behavior, which also yields to digital preconditioning, is cone resonance. A cone loudspeaker is far from being a rigid piston. In fact, the most successful cone formulations transmit mechanical vibrations at a speed not too dissimilar to the speed of sound in air. A mechanical wave travels from the voice coil to the surround, where it is only partially absorbed. A portion of the energy is reflected back down the cone to the voice coil, where it is, once again, partially absorbed and partially re-emitted. Unlike a horn, the reflections tend to be strongest at the upper end of the transducer's usable range. In many cases the audibility of the mechanical resonance defines the upper frequency limit of usability.

The sound produced by the initial mechanical wavefront combines with sound produced by later, re-emitted mechanical wavefronts. Each of these contributions is produced by the same radiating system, with the same directionality, so the effect of the resonating system is the same in every direction. Therefore, the phenomenon of cone resonance is a two-port system, and is correctable with signal preconditioning.

Figure 3.3.1 is a spectrogram of a 0.38 m (15 in) cone loudspeaker, measured on axis with no filtering applied. Figure 3.3.2 is the same loudspeaker, with a preconditioning filter applied.

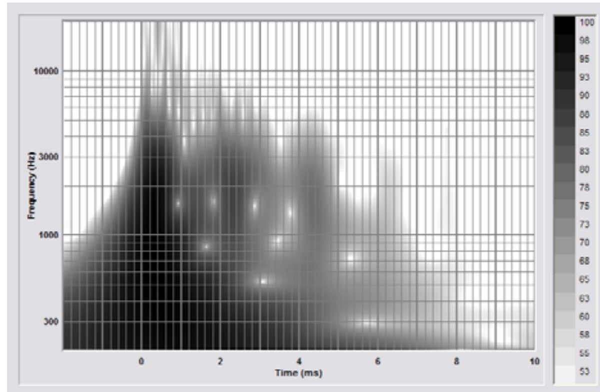

Figure 3.3.1: Cone Speaker, 0.38 m, No Filtering

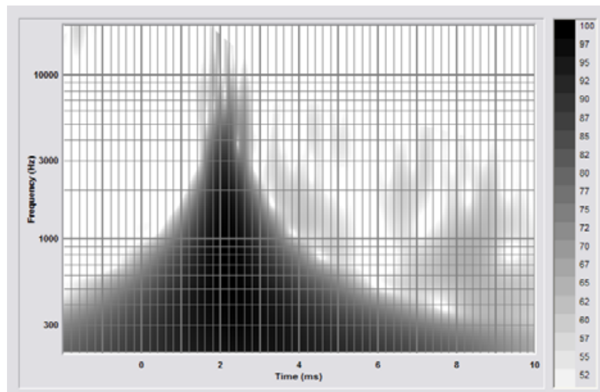

Figure 3.3.2: Cone Speaker, 0.38 m, With Preconditioning Filter

The directionality of a loudspeaker also affects its transient response, but it is not a 2-port characteristic. Hence it cannot be corrected everywhere. However, the transient response can be modified in a way that produces greater consistency throughout the pattern. By improving the transient response in the "worst" direction, at the expense of the transient response in the "best" direction, the sound quality can be made more consistent over the breadth of the coverage pattern. This is exactly what has been implemented in the example shown. The off axis spectrogram, shown in figures 3.3.3 and 3.3.4, shows a degree of improvement similar to that observed on axis.

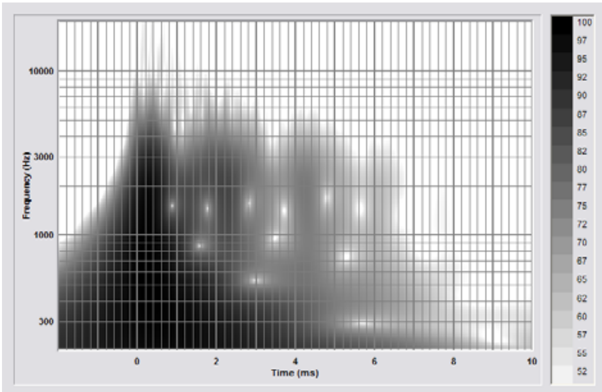

Figure 3.3.3: Cone Speaker, 0.38 m, 15 degrees Off Axis, No Filtering

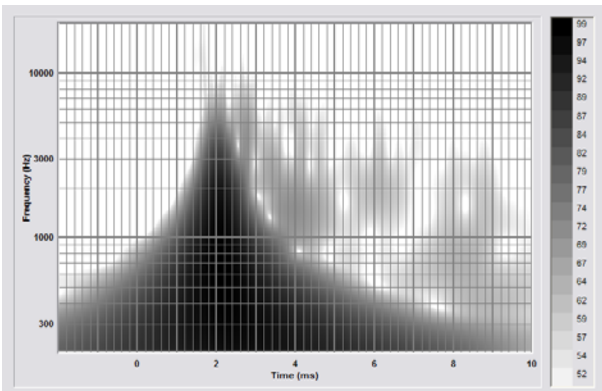

Figure 3.3.4: Cone Speaker, 0.38 m, 15 degrees Off Axis, With Preconditioning Filter

3.4 Crossover Phase Linearity

Multi-way loudspeaker systems typically employ minimum phase high pass and low pass filters. When such filters are summed, a flat magnitude response may be obtained, but the phase response is non-minimum. The resulting system has an all-pass transfer function, with a phase component that is similar to that of the high pass and low pass filters from which it was constructed.

Figures 3.4.1, 3.4.2, and 3.4.3 show the magnitude, phase, and impulse response of a typical set of crossover filters, as well as their summed response.

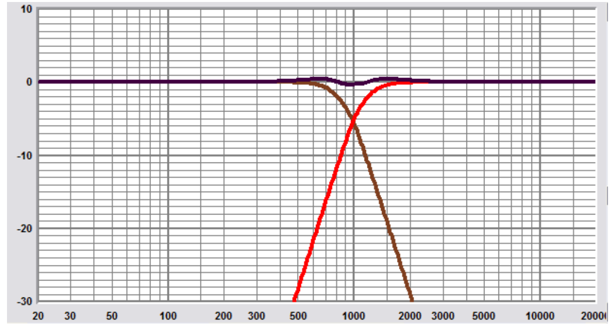

Figure 3.4.1: Crossover Summation - Magnitude Response

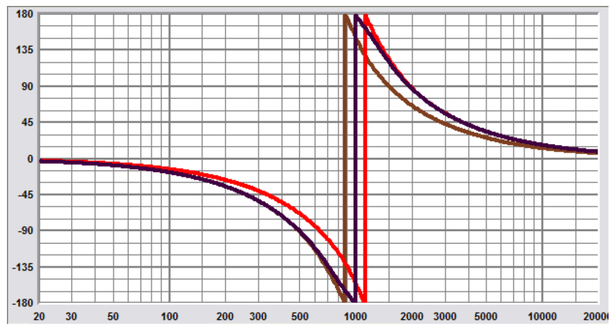

Figure 3.4.2: Crossover Summation - Phase Response

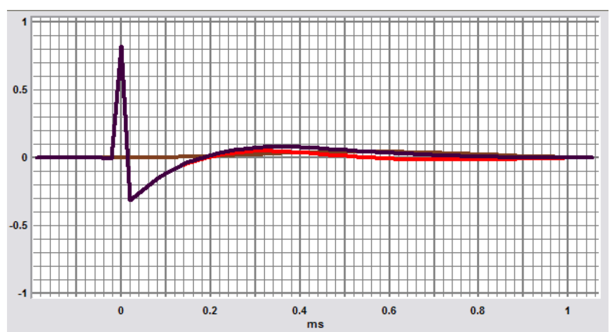

Figure 3.4.3: Crossover Summation - Impulse Response

To linearize the phase response of the summed crossover, we would need a complementary filter, created by inverting the summed response. The result is itself an all-pass filter, but with a phase response that is the mirror image of the summed crossover. In the time domain, this yields an impulse response which is reversed in time. Since the crossover filters themselves have infinite impulse responses, the correction filter is unbounded in negative time. The ideal correction filter is illustrated in figures 3.4.4 & 3.4.5.

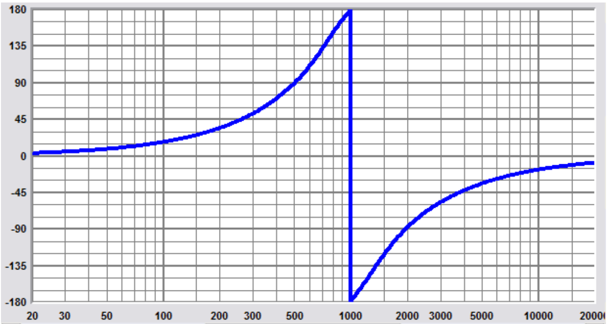

Figure 3.4.4: Ideal Correction Filter (blue) - Phase Response

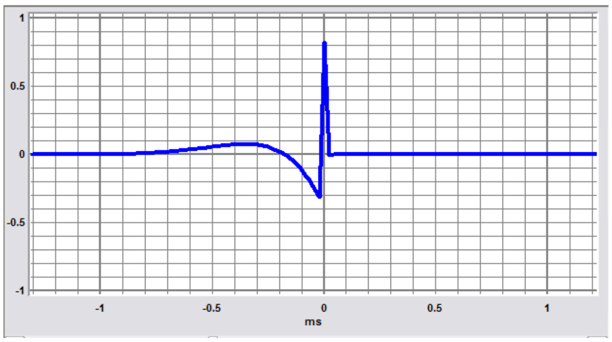

Figure 3.4.5: Ideal Correction Filter - Impulse Response

Strictly speaking, this is an unrealizable filter, even if we accept some amount of latency; but in practical terms, a very useful approximation of this filter is realizable. Through judicial selection of the data window, an FIR filter can be created which gives a sufficiently accurate approximation of the desired response. For this example (4^{th}-order filters at 1 kHz), a practically useful FIR filter can be extracted with a length of .7 ms, or 144 taps at a 48 kHz sample rate. A nearly perfect approximation can be achieved with 1 ms or more.

Figures 3.4.6 & 3.4.7 show the magnitude and impulse response of a .7 ms correction filter. The .7 ms example is used to illustrate the difference from ideal. A 1 ms filter would be indistinguishable from the Ideal filter when viewed in this display format.

The essence of this filter is that the high frequencies are delayed relative to the low frequencies, which counteracts the minimum phase crossover's effect of delaying the low frequencies relative to the highs. From the impulse response display (Figure 3.4.7) it is easy to determine that the added high frequency latency is .7 ms, or the full length of the FIR filter.

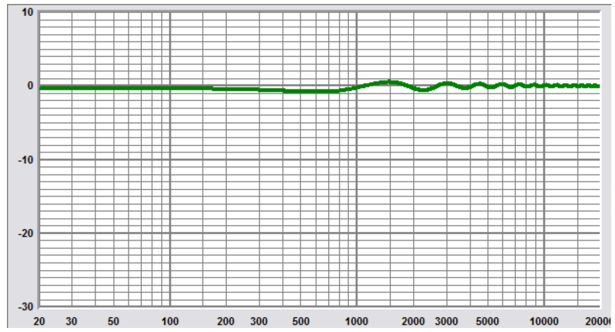

Figure 3.4.6: 0.7 ms FIR Correction Filter - Magnitude Response

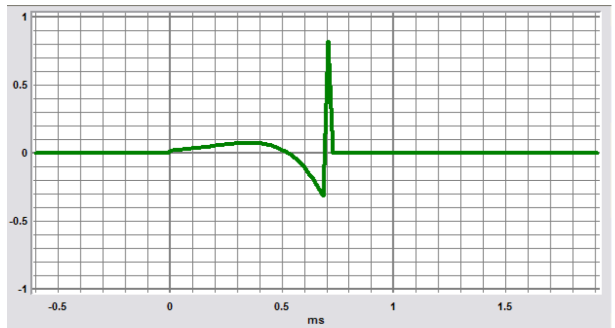

Figure 3.4.7: 0.7ms FIR Correction Filter - Impulse Response

4 CONCLUSIONS AND COMMENTS

A line of horn-loaded two-way loudspeaker systems has been developed, in which the described transient response improvements have been implemented. The clarity and definition these loudspeakers exhibit is more similar to that of studio monitors or high quality consumer loudspeakers - than to other loudspeakers with similar components and horn designs.

A transducer designer who expects his designs to be used in systems that make use of these preconditioning techniques may significantly change his focus in his development work. Rather than optimizing the uncorrected performance of a transducer, he might concentrate on maximizing the loudspeaker's linearity, time invariance, and directional consistency; in other words, the attributes that make it correctable. This change in focus could lead to further performance improvements.

Some of the techniques mentioned in this paper may be covered by one or more patents, applications for which are currently being prepared. Readers are advised to review any of these patents that may issue before implementing similar techniques in commercial products.

5. ACKNOWLEDGEMENTS

The author wishes to acknowledge Bill Hoy, Chuck McGregor, and Nathan Butler - for their contributions to the research presented here and/or assistance in the preparation of this paper.

6. REFERENCES

[1] D. Gunness and W. Hoy, "A Spectrogram for Loudspeaker Transient Response", to be presented at the 119th Convention of the Audio Engineering Society, 2005.

[2] H. F. Olson, *Elements of Acoustical Engineering* (Van Nostrand, New York, 1947).