The AudioMachina XTAC Master Reference System
(Patent Pending)

The White Paper

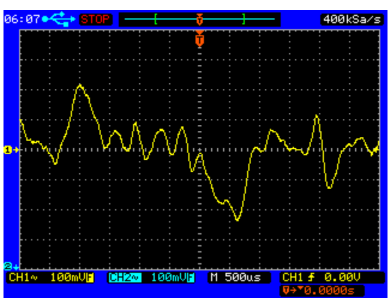
Dr. Karl Schuemann
AudioMachina Inc.
Copyright 2017

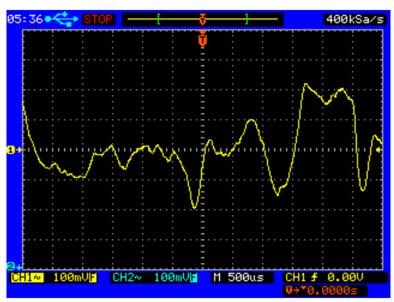
An Introduction to "High Fidelity" Recorded Music, and its Basic Properties

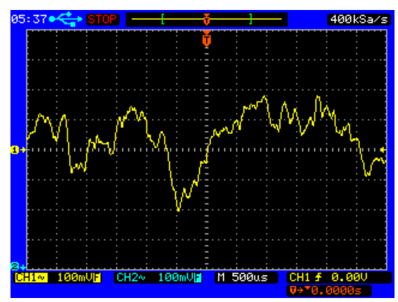
When considering precisely what is meant by the phrase "High Fidelity Loudspeaker," we must first have a very clear understanding of precisely what “Fidelity" means. A good working definition, compiled from numerous dictionary entries, is:

Fidelity: The Degree of Accuracy with which Music is Recorded and Reproduced.
Synonyms for Accuracy: Exactness, Precision, Correctness.

In order to understand how this definition should be applied to optimum loudspeaker design, it is essential to first understand the basic form of recorded music. Below is a sample of images which show a brief moment of a single channel of recorded music in visual form. These are "screenshots" (aka "brief moments in time") taken from a high-performance digital storage oscilloscope being fed a recorded music signal from a high-fidelity preamplifier. The horizontal (X) axis represents Time, and the vertical (Y) axis represents Amplitude. In this case, amplitude is in units of voltage, as that is the conventional basic unit of recording and playback. Note the scale of the screenshots: Time is 500us (500 microseconds) per block, or 5ms (0.005 seconds) for the entire screen. Each tiny division is therefore 100us (0.0001 seconds). Amplitude is 100mV (100 millivolts) per block, or 20mV per tiny division.







The exact names (artists, songs, albums) of these particular images do not matter at all-- these images are intended only to give an understanding of what music actually "looks like" in real time. Within the entire catalog of recorded music known to mankind, there are literally billions upon billions of such unique images. (A single standard CD alone can hold nearly a million of these screenshots.) And these screenshots are, philosophically speaking, exactly like snowflakes-- they all have certain inherent properties which they all share, and yet you can look for the rest of your life and never find two which are exactly identical-- every single one of them is absolutely unique.

So, based on these visual images, what are the inherent, defining properties of music itself (and therefore, high fidelity recorded music)?

1. *It is Continuous*. It never jumps from one value to a completely different value in zero time, but rather, it flows continuously from one value to the next over time.

2. *It is Singular*. At every single moment in time, it has one and only one single, specific amplitude, never more than one nor less than one. In other words: It traces a single line through time.

3. *It is Complex*. It is not reducible to a simple equation, and it is constantly changing shape in unpredictable ways. Another way of saying this is: Music is *always transient* in nature.

4. *It is Unique*. At every single, precise, unique instant of time, it has a single, precise, unique corresponding amplitude. This fact is at the very heart and soul of every piece of music ever played, and every piece of music ever recorded. If you change either the amplitude at a precise moment in time, or the time at which a precise amplitude occurs, the music is no longer itself, and the reproduction can no longer be considered "High Fidelity," because the fundamental *unique shape* of the waveform has been changed. In other words: *Time and Amplitude are absolutely inseparable if the music is to remain as it was originally, or if music is to be considered "High Fidelity" when reproduced.*

An Introduction to the Capabilities of the Human Hearing System

Next, in order to understand what capabilities are absolutely essential to a "High Fidelity" loudspeaker, and more specifically, *how good* each of those capabilities must be, we must first investigate the capabilities (and limitations) of the human hearing system. Any loudspeaker (or other component) which aspires to "High Fidelity" must meet at least a minimum level of performance in *all* of these areas, or else the human hearing system will be able to detect very easily that the "reproduced music" is fundamentally wrong compared with "real music." The following 4 criteria are all different, but *every single one is fundamentally important to High-Fidelity music reproduction:*

1. Frequency Response: The range of human hearing is traditionally stated as 20Hz-20kHz. Music can have a wider range, but most music is within these limits. (Some basic facts: The lowest frequency attained by common instruments is A0 on the standard 88-key piano, at 27.5Hz. The lowest frequency on a standard four-string bass is E1, at 41.2Hz. During music reproduction, most domestic (and mastering) rooms exhibit "room gain" in the deep bass, beginning around 40Hz and increasing at lower frequencies, and thus it is advantageous to have the loudspeaker begin a very gentle rolloff at around 40Hz, to avoid overpressure at extremely low frequencies. Finally, most adults cannot hear much above 16kHz, regardless of what information is above that.) Thus, in the real world, we can say that the loudspeaker system should have relatively flat anechoic response from 40Hz-20kHz, with a very gentle rolloff below that, keeping the in-room response flat from 20Hz-20kHz.

2. Dynamic Range and Signal-to-Noise Ratio: These are two very similar criteria, so are discussed together. The human hearing system has a basic dynamic range of 0dB-120dB SPL, from the quietest detectable sound to the limit of brief exposure before physical pain or hearing damage. Typical *extremely* quiet rooms, with truly heroic acoustic isolation, have a background noise level of 20dB (below which any signal gets buried under the background noise), with typical *very* quiet rooms around 30dB background noise, and typical untreated rooms around 40-50dB background noise. Thus, we can

state that we should strive for a minimum S/N ratio, in any reproduction system, of at least 100dB (120dB minus 20dB), and a minimum usable dynamic range of 100dB also (20dB-120dB SPL). And 120dB for both figures would be welcome. Because most real music has a maximum in spectral energy content in the octaves on either side of 200Hz (i.e., 100Hz-400Hz), this is generally where the highest output is necessary, with slightly lower requirements over the remainder of the audio band. (As an aside, and without wishing to start any arguments, it should be noted that, as recording and playback media, neither vinyl nor 16/44 digital meets the 100dB minimum, not to mention the 120dB that we might desire. However, 24/96 digital (or above) easily achieves 120dB, with room to spare. Fortunately, many recording studios and engineers are now at least recording, editing, and storing original tracks at 24/96 or higher, even if the music is finally issued on poorer-resolution formats.)

3. Amplitude Resolution: Under perfectly ideal laboratory conditions, the human hearing system can resolve an amplitude difference of 0.5dB. In the real world, while playing music, a 1.5dB difference in amplitude is somewhat difficult to resolve, even for expert listeners, while 3dB is rather easy even for untrained listeners. Keep in mind, however, that these numbers represent huge increments in loudness level. A change of 3dB is literally twice the acoustic power (or half the power), meaning a change in signal voltage level by a factor of 1.414 (the square root of 2). Even a 1.5dB change in level represents over a 40% change in acoustic power, or nearly a 20% change in signal voltage. To think about it another way, even if we say that a good listener can distinguish 1.5dB increments at any volume level while listening to music, there are only 80 discrete music volume levels that his/her hearing system can possibly distinguish, from softest to loudest! (120dB divided by 1.5dB.) In other words, the human hearing system is really quite insensitive to changes in signal amplitude. Nonetheless, the traditional standard +/-3dB specification for frequency response in loudspeakers is quite appropriate as a basic requirement for "high fidelity" music reproduction. And +/-1.5dB would be preferable.

4. Time Resolution: Under perfectly ideal laboratory conditions, the human hearing system can resolve time differences of less than 10us (0.00001 seconds, or 10 microseconds). Recent scientific experiments have shown that this is true of both binaural hearing (via sound localization studies) and monaural hearing (meaning that each individual ear has the same inherent 10us time resolution capability, as would logically be predicted). In the real world, while playing music, a 40us time difference is somewhat difficult to resolve, even for expert listeners, while 80us (0.00008 seconds) is rather easy even for untrained listeners. (As an easily understood example, 80us represents an "image shift" in a stereo playback system, from dead-center to 10 degrees off-axis. This image shift will be easily noticed by even casual listeners. More attentive listeners will be able to notice image shifts from center to only 5 degrees to one side (equal to 40us), and many listeners can do even better than this. Similar time-resolution capabilities apply to each ear individually, even if stereo image shift is not used as the test.) Thus, similar to our amplitude data above, we can state that a "high fidelity" playback system should introduce time errors of no more than 80us in the signal, and preferably no more than 40us. This standard should apply throughout the majority of the audible frequency spectrum, but can be relaxed significantly in the low bass and high treble, as the human hearing system becomes quite insensitive to timing at very low and very high frequencies.

Now that we have a basic understanding of human hearing capabilities, let's briefly revisit the above screenshots of music. If we insist on time errors no greater than 40us, and amplitude voltage errors of no more than 20% (both the "preferable" requirements for high fidelity above), we notice that the eyes and the ears do not see (or hear) things the same at all. At the scale of these screenshots, a time error of 40us is only 4/10 of one tiny division! This is extremely difficult for the eye to resolve. On the other hand, with a peak-to-peak voltage of 4 blocks as seen on these screenshots, a 20% change in voltage amplitude is 4 full tiny divisions of error in amplitude, *10 times more* than the allowable visual error in the time scale, and incredibly easy for the eye to resolve. If we reduced the displayed amplitude to where a 20% change in peak-to-peak amplitude represented the same *visual* error as on the time scale, the vertical signal voltage displayed would have an amplitude of only +/- 1 tiny division!! In other words, it would be so shrunken in vertical scale that the eyes would hardly be able to resolve any changes in amplitude in the signal at all. This should give a visual illustration of just how critically important time errors are, relative to amplitude errors. Do not allow your eyes to deceive you about the capabilities of your ears-- they are two entirely different physiological systems, and their relative capabilities are not at all the same. *The human hearing system is vastly more sensitive to Time than it is to Amplitude*.

It should be emphasized once again that the above four criteria must *ALL* be met simultaneously, in order for a music playback system to ever have any hope of presenting reproduced music in a form which the human hearing system will recognize as "like real music." Any system which does not meet *ALL* four criteria simultaneously simply cannot be described as "High Fidelity," because *the human hearing system's innate capabilities will easily be able to recognize that it is not.*

A Brief Summary of Historical Loudspeaker Types, and their Limitations

It is now necessary to investigate the inherent capabilities and limitations of the major types of historical loudspeakers, and then to understand why those limitations fundamentally prevent them from attaining the label "High Fidelity," regardless of cost.

1. Horn Loudspeakers: The earliest form of sound reproduction device, dating to the 1800s and used by Edison in the earliest forms of sound recording and playback. Still used extensively for low-fidelity sound reinforcement applications, where output capability and efficiency are paramount. Problems include: (a) Non-linear air pressure swings during compression vs. rarefaction, resulting in audible distortions, (b) "Horn Colorations" due to suboptimal physical horn geometry, also an audible form of distortion, (c) limited bandwidth of individual horns, necessitating the use of multiple drivers with crossovers, which automatically precludes high fidelity (discussed in more detail below), and (d) Necessity of use either with dynamic woofers (with all the problems discussed below), or with bass horns which, if sized for true 20Hz extension, are the size of entire rooms.

2. Electrodynamic or Dynamic ("direct radiator") Loudspeakers: Also rather old, with the earliest crude forms dating back to the late 1800's. The basic modern form of this type was described by Rice and Kellogg in 1925, nearly 100 years ago, and all modern iterations operate on the same fundamental

physics. The fundamental limitation of the dynamic loudspeaker is that it operates (in physics terms) as *a mass on a spring.* This will be covered in much greater detail below. Briefly put, because it has mass, it has inertia, and because it has inertia, it is always and forever trying (unsuccessfully) to catch up to the input signal. It can't be started moving when it should, and it can't be stopped when it should either. And at every point in between, it is always behind where it should be, in the time domain. Even worse, its time lag is both transient-dependent and frequency-dependent, meaning that its time delays are not consistent across the frequency spectrum—the lower frequency components of the signal are delayed in time worse than the higher frequencies, and therefore these problems *cannot* be fixed by simple physical driver offsets-- it is *mathematically impossible*. Therefore, it cannot possibly meet the basic requirements for "High Fidelity," even as a single driver without the additional problems of crossovers, because it is a complete disaster in the time domain relative to the requirements of "High Fidelity."

3. Multiway Electrodynamic (Dynamic) Loudspeakers: A variation of the above, but with multiple drivers, each of which covers a limited frequency range, usually with crossovers dividing the signal between individual drivers. By far the most popular modern form of the loudspeaker. This type takes the fundamental Achilles' Heel of the electrodynamic driver above (the "mass on a spring" problem), and makes it even worse in the time domain. There are two main reasons for this:

3.1 Woofer diaphragms have 5-10 times the mass of midrange diaphragms, which in turn have 5-10 times the mass of tweeter diaphragms. Yet the drivers all have relatively similar magnetic field strengths. This means, based on basic physics F=ma, that the acceleration of tweeters is vastly faster than midranges, which in turn are vastly faster than woofers. This can be seen very clearly by looking at the impulse response of a multiway loudspeaker, even many which claim to be "time aligned": First to arrive is the tweeter impulse, followed (after a delay of typically 200us) by the midrange impulse, followed (after an even longer delay, typically 1000us) by the woofer impulse. This is the natural consequence of a mass responding to an input force: A lot more mass takes a lot longer to get it moving. And notice the delay times: *all* of them are extremely obvious relative to the known real-world capability of the human hearing system at 40us. Furthermore, we have already established that *all music is transient in nature*. Thus, whenever the musical signal changes direction unpredictably (which, as we already know, is *all the time*), the tweeter's change in response to that signal will arrive at the ears *long* before the midrange's, which in turn will arrive *long* before the woofer's.

3.2 The crossovers typically used in multiway systems contribute even more frequency-dependent non-linear phase shift, and those phase shift errors are added to the innate responses of the drivers. And this problem gets worse as the crossover slope goes higher. It is mathematical fact that no crossover type above first-order can *possibly* sum correctly in time and amplitude under transient conditions (aka *real music*). It is not merely difficult; it is *mathematically impossible*. And since these phase errors are again non-linear with frequency, they contribute non-linear time errors to the system's response. And again, these time errors *cannot possibly* be fixed with physical driver offsets, because they vary with frequency. When combined with the inherent mass-related time delays above, it is normal in multiway dynamic systems to have phase error differences in the range of 720 degrees or more across the frequency spectrum. This is a complete disaster in the time domain.

The practical consequence of this behavior, in *all* conventional dynamic loudspeakers, regardless of type or cost, is that for any instrument which generates fundamentals and overtones (which includes *virtually any instrument one could possibly name*), many overtones will arrive at the ears *long* before the fundamentals. Certainly a single-driver speaker is superior in this regard relative to a non-time-aligned multiway with high-order crossovers, but the fundamental problem remains. Please imagine, for a moment, just how incredibly irritating this is to the human hearing system, to constantly be bombarded by high frequency overtones *long* before the arrival of the lower frequency fundamentals. This, in a nutshell, is the source of "brightness" and "glare" and "listener fatigue" in speakers which otherwise may measure "flat" in frequency response, and also the fundamental reason why dynamic speakers are instantly recognized by the human hearing system as "speakers" and "not real." It is also the reason why many dynamic loudspeakers have a deliberate pronounced "downward slope" in frequency response from bass to treble, often 10dB or more: Their designers are trying to compensate for the irritation caused by the early arrival of the high frequencies, relative to the low frequencies, by progressively boosting the lower frequencies. This is basically a very crude attempt to try to fool the ear into paying more attention to the (late-arriving) lower frequencies, because they are *louder* relative to the (early-arriving) higher frequencies, thus supposedly "balancing out" the perceived sound. But it doesn't work, and can never work: *Two wrongs will never make a right*. It is *utterly impossible* to fix an inherent problem in the time domain by creating an equally egregious problem in the amplitude domain.

In conventional dynamic loudspeakers, given the magnitude of the time delays between various frequency components in the music, even from a single dynamic driver, it is ridiculously obvious to the ears that something is very, very wrong. But because this type of (time arrival) error occurs *nowhere in nature and nowhere in natural sounds*, humans have never adapted to it evolutionarily, and the ear can't recognize what the problem is, although it knows for sure that *something is very wrong*. It knows that there is a very big difference between what it's hearing, and what *real natural music* sounds like.

4. Panel Dipole (Electrostatic or similar) Loudspeakers: First seen 60 years ago in Peter Walker's legendary Quad in 1957. Historically speaking, the last big breakthrough in loudspeaker performance, and the first wide-range transducer in the history of the world to have, at least approximately, correct Time vs. Amplitude characteristics. (And also the reason that it actually *sounds like real music* in the upper half of the human hearing range.) However, the electrostat (or any planar dipole variation) cannot be considered "high fidelity" due to the fact that *it is a dipole*. Because it is a dipole, it creates a full-power inverted-phase acoustical backwave at exactly the same time as the front wave. And at frequencies beginning in the midrange and steadily worsening at lower frequencies, the inverted-phase backwave becomes progressively less directional, and begins to combine with the front wave, but with a large time delay. This results in enormous errors in both time and amplitude, with the result being that dipoles, by definition, cannot be considered "high fidelity" loudspeakers. Furthermore, the limited excursion available in all electrostats creates power-handling problems in the bass which, added to dipolar bass cancellation, seriously compromises amplitude accuracy and dynamic range at lower frequencies. Many speakers have tried to mate dynamic woofers to electrostats with crossovers, but they all suffer from the same (unsolvable) problems in the time domain as multiway dynamics.

5. Bending-Wave Loudspeakers: These fall into both flexible-diaphragm and semi-rigid-diaphragm types, with many variations. However, all of them suffer from the same problems: (a) Presence of flexure and mechanical standing waves on diaphragms, resulting in significant errors in both time and amplitude, and (b) limited bandwidth, typically resulting in the necessity (yet again) of combining them with dynamic woofers and crossovers, again precluding high fidelity.

Imagining the Ideal High-Fidelity Loudspeaker

Now that we have gained a basic understanding of the limitations of historical loudspeaker types, we must ask the question: If we were to attempt to design an "ideal" High-Fidelity loudspeaker, what requirements would we expect of it? Following are 8 criteria, *all of which* must be met simultaneously by our ideal loudspeaker:

1. First, and by far the most important: It must be fundamentally correct in its Time vs. Amplitude acoustical output, across the entire frequency spectrum of human hearing. In the real world, this means that it must have no apparent inertia, i.e., if it has mass, it must include a way of precisely negating the time delays associated with forcing that mass to change its velocity in real time.

2. As a corollary to (1), it must have essentially full-range flat frequency response in-room. Please Note: Any system which has fundamentally correct Time vs. Amplitude response to signals within the normal audio range will *automatically (by mathematical definition)* have flat frequency response within that range. It is critically important to understand this fact, so if necessary, please read that sentence again.

3. It must have a dynamic range and S/N ratio of at least 100dB, and preferably as much as 120dB, preferably throughout the audio range but at least above 100Hz, which is the lower end of the "power range" of most real music.

4. It must not contain multiple drivers of different types, or crossovers of any kind. Such designs are automatically disqualified from the definition of "High Fidelity," as detailed above. This means that all drivers used must be fundamentally capable of full-frequency-range performance, and must be of the same exact type.

5. Any driver diaphragm(s) must not undergo mechanical flexure or standing waves within the normal audio frequency range (20Hz-20kHz), i.e., diaphragms must behave as rigid (pistonic) surfaces throughout the entire audio frequency range.

6. Any electronic circuits which pass audio signal should remain in the analog domain, avoiding the inevitable compromises in fidelity which occur during needless extra A/D, D/A, and DSP stages.

7. It should create a spatially uniform acoustic wavefront in-room without significant lobing, wave

interference effects, phase cancellations, etc. Only two fundamental wave radiation patterns qualify under this requirement: Spherical (point source) or Cylindrical (line source). This essentially precludes the use of spaced multiway systems or large-diameter diaphragms of any type.

8. It should be capable of being installed in normal rooms, and of avoiding strong early reflections from room boundaries, with only the use of standard, easily installed room acoustic treatments. This essentially precludes any system design which could have a "floor bounce" interfere with its acoustic output, as floors are extremely impractical to treat acoustically.

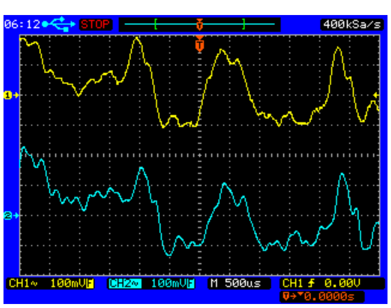
The Design of the AudioMachina XTAC Master Reference System

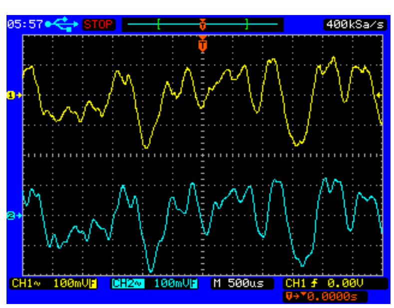
Armed with the above 8 requirements for "High Fidelity" loudspeaker design, we are now finally ready to begin to discuss the design of the AudioMachina XTAC Master Reference System.

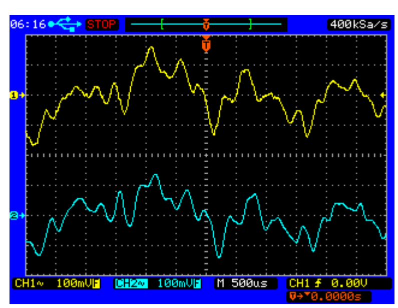
The Time vs. Amplitude Problem

Since this is not only the most important of the 8 criteria above, but also the design aspect most likely to be deemed "impossible to solve" by the scientifically ignorant (or perhaps more to the point, by the marketing departments of all other loudspeaker manufacturers), we will first present *the results* of solving this problem, and then we will proceed to the "how" it was solved.

Below is a new set of screenshots, again taken from the same high-performance digital storage oscilloscope. The musical selections again do not matter, as the same basic results will be obtained no matter which particular piece of music is fed through the system. (That is the whole point of "High Fidelity"!!!) But in these screenshots, instead of a single yellow trace, there are now two traces, a yellow and a blue. As before, the yellow trace is simply the output of a high-fidelity playback preamplifier; in other words, it is a "High Fidelity" form of the original recorded musical signal, taken at the exact same time as the exact same signal is fed to the input of the XTAC System. The blue trace is simply the final *acoustic output* of the AudioMachina XTAC Master Reference System, as picked up by a high-quality condenser (aka monopolar electrostatic) measurement microphone, amplified by a high-quality microphone preamplifier, and then fed directly back to the oscilloscope in real time.







Note that the *acoustic output* of the XTAC Master Reference System bears a *shocking resemblance* to the original musical input signal, in both Time and Amplitude, but most critically, in Time. (If you look carefully, you will see a very slight but very consistent time delay between the yellow and blue traces (approximately one tiny division), which is the result of the very small but still noticeable sound-wave travel time between the driver voice coil and the microphone capsule.) These results should be absolutely eye-popping and jaw-dropping to anyone who understands just how poorly traditional loudspeakers perform on this test, *regardless of price*. This is not only the very essence of "High Fidelity" music reproduction, it is also the *very first time in the entire history of the world* that a loudspeaker has actually achieved this breakthrough in a design which meets *all* of the above 8 criteria.

Physics and Mathematics

Upon seeing these results, the obvious question is: "Since XTAC is clearly using dynamic drivers, how can it possibly behave as if those drivers are *essentially massless*, as it is clearly doing?" To answer this question, it is necessary to discuss some basic physics and mathematics. We will try to keep it as simple as possible, but there is no other way to explain it.

To begin, let's think about the physics of the dynamic driver. Fundamentally, as a mechanical system, it is a "mass on a spring with damping." This is a basic physics problem seen in every college physics (and in mathematics, differential equations) curriculum. And it is the reason why the Time vs. Amplitude response of every conventional dynamic driver is absolutely terrible, regardless of cost. Due to the time delays created by the inertia inherent in the mass, the conventional dynamic driver simply cannot follow the input signal in real time. Even worse, these time delays are not constant with frequency, nor constant under transient input conditions. “High Fidelity” is simply impossible in these circumstances.

Next, it is necessary to introduce the concept of the "transfer function." Simply defined, a "transfer function" is a mathematical equation which describes *the behavior (or output) of a system* based on some input variable(s). So, in the case of a "mass on a spring with damping," there is a specific differential equation (or "transfer function") describing the motion of that system in response to some input. In this case, the "input" is an audio signal in the form of a varying voltage, with the properties discussed above. In a dynamic driver, that varying voltage causes a varying current to flow through the voice coil, which, being immersed in a magnetic field, generates a force proportional to that current. That force, in turn, acts on the "mass on a spring with damping," creating a varying acceleration according to F=ma, which in turn creates a varying velocity of the cone, and thus an acoustic output (sound) by transference of that velocity into the air molecules which are in contact with the cone. The differential equations which describe this system's transfer functions will not be shown here, as they are well beyond most people's math skills. Suffice to say that they are *extremely non-linear functions, in both Time and Amplitude*, which explains why the fidelity of conventional dynamic drivers is so poor. It is simple enough to look up these equations if anyone wishes, but for reasons soon to be explained, their precise mathematical form need not be understood in order to understand the basic operating principles of the XTAC Master Reference System.

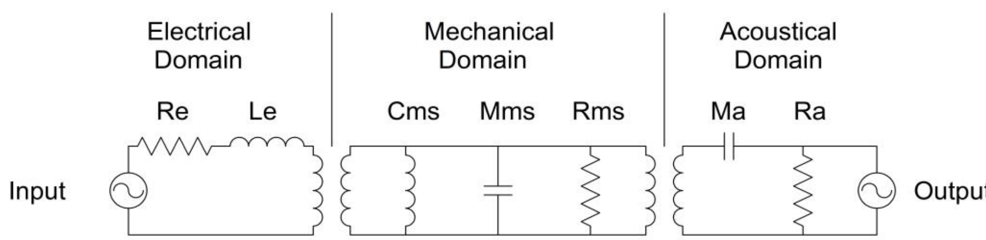
To sum up, the conventional dynamic driver transforms electrical energy (voltage and current) into mechanical energy (alternating kinetic and potential energy in the "mass on a spring" system), resulting in a *delayed and non-linear mechanical response* to the electrical signal, and that mechanical energy is then transformed into acoustical energy, resulting in a “low fidelity” form of the original input signal.

Now, we will make the following statement without proof, as anyone can look it up and confirm it to be true: The mathematical differential equations which describe the transfer function of the "mass on a spring with damping" Mechanical system are *absolutely identical* to the mathematical differential equations which describe the transfer function of an LCR (Inductor, Capacitor, Resistor) Electrical system. Note that we did not say that the equations are *similar*; we said that the equations are *identical.*

Furthermore, we will make the statement of fact that if you multiply a mathematical function (*any* mathematical function) by its inverse, you get unity. Simply put, if you multiply "f" times "1/f", you get the answer "1". (And it obviously doesn't matter what "f" is; you always get the answer "1".) And that means that, for any transfer function multiplied by its inverse, by mathematical definition*, input equals output, in both Time and Amplitude*. Period.

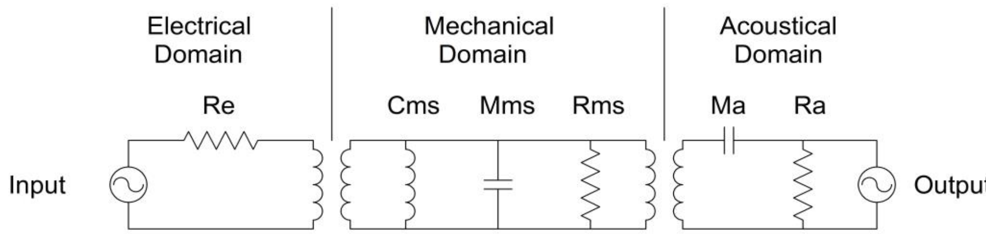
The Basic Model of the Electrodynamic Loudspeaker

We will now present the "full basic model" of the modern electrodynamic loudspeaker driver, as shown in many basic acoustics textbooks:



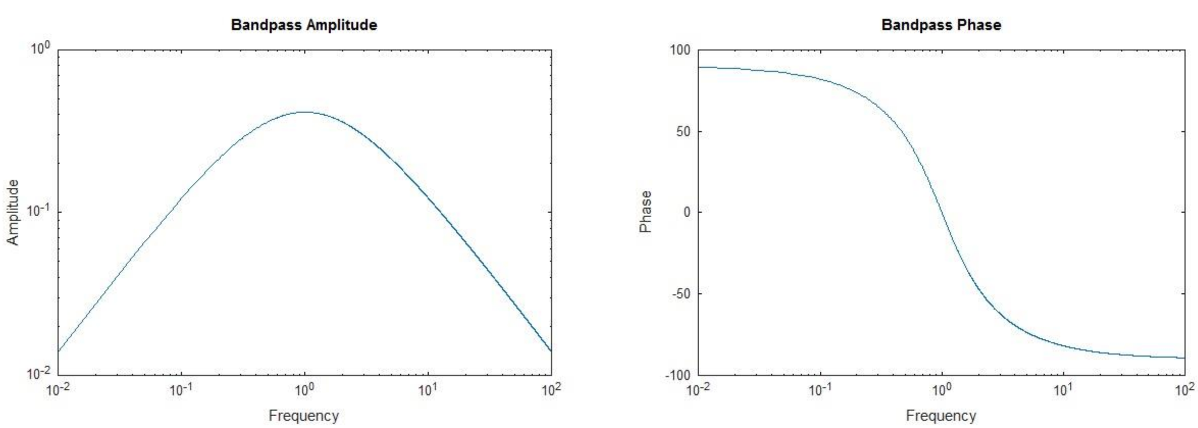
While this may seem complicated, it's not too difficult if we take it one section at a time. First of all, there are 3 main sections (or "domains") seen in the model: On the left, Electrical Domain. In the middle, Mechanical Domain. And on the right, Acoustical Domain. These 3 domains are separated by 2 transformations. It is inaccurate to call them "transformers," although visually they use the electrical symbol of a transformer. The symbols actually represent the "transformation" of energy from one form to another-- first, from electrical to mechanical energy, and second, from mechanical to acoustical energy.

Beginning with the Electrical Domain, we see a signal "input source" denoted by a circle with a wave in it. This would normally be an amplifier in the real world. Next, we see a resistor and an inductor in series, which represents the electrical resistance and inductance of the driver's voice coil (and wiring, etc.). In larger drivers, the inductance of the voice coil is often high enough that it forms an electrical low-pass filter which attenuates high frequencies. However, in small drivers with small voice coils, such as those used in the XTAC System, the effect of the inductance within the audio range is negligible, and can therefore be ignored. The basic model can therefore be simplified to:

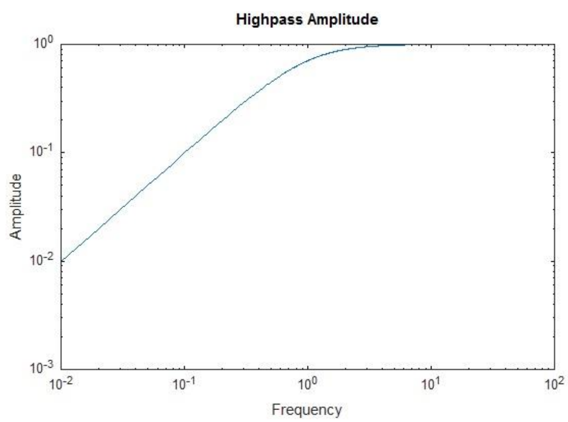


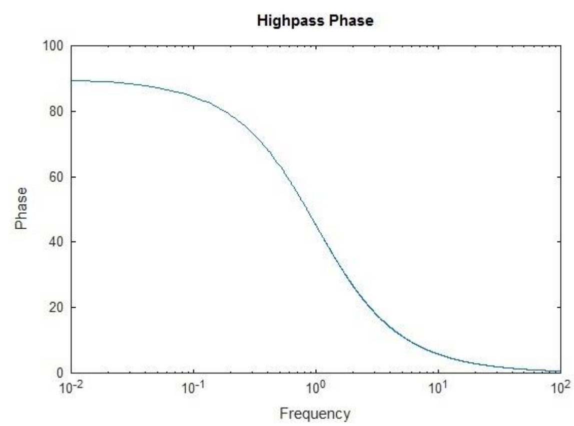
With the Electrical Domain now consisting of only a source and a resistor, its transfer function is extremely simple and completely linear in both time and amplitude: Current is directly proportional to Voltage. This part of the system needs no further attention.

Moving to the Mechanical Domain, we see the LCR representation (or "Analogy") of the "mass on a spring with damping" Mechanical system. To repeat, the LCR form is mathematically identical to the "mass on a spring with damping" Mechanical form, so it is shown visually in LCR Electrical form here. The transfer function of this LCR circuit is commonly known as a "second order band-pass filter," with the center (resonant) frequency of the band-pass filter being the mechanical resonant frequency of the driver system (in the enclosure), and the Q of the band-pass filter equal to the total Q of the driver system (in the enclosure). Note that the term "second order band-pass filter" defines a function with a first-order downward slope on either side of the center (resonant) frequency, thus the name "second order"-- there is no such thing as a "first order band-pass filter." The effective Amplitude vs. Frequency and Phase vs. Frequency plots of the "second order band-pass filter" transfer function are shown below:



Moving to the Acoustical Domain, we see a capacitor in series with a resistor, with the acoustical output taken across the resistor. Again, similar to the Mechanical Domain, this is the *electrical analogy* of the Acoustical transfer function (or more specifically, the air's "acoustic impedance" function)—obviously, air is not actually made from physical capacitors and resistors in the real world. But similar to the Mechanical Domain's analogy, the air's acoustic impedance function can be represented by an Electrical RC circuit here, because the mathematical differential equations are the same. Fundamentally, the air functions as a "first order high-pass filter", wherein its ability to transform mechanical diaphragm motion into acoustical energy remains essentially constant at high frequencies, then as frequencies go below a certain value (approximately where the driver's circumference equals the wavelength), its efficiency in transforming mechanical energy into acoustical energy falls off steadily with decreasing frequency. The effective Amplitude vs. Frequency and Phase vs. Frequency plots of the "first order high-pass filter" transfer function are shown below:



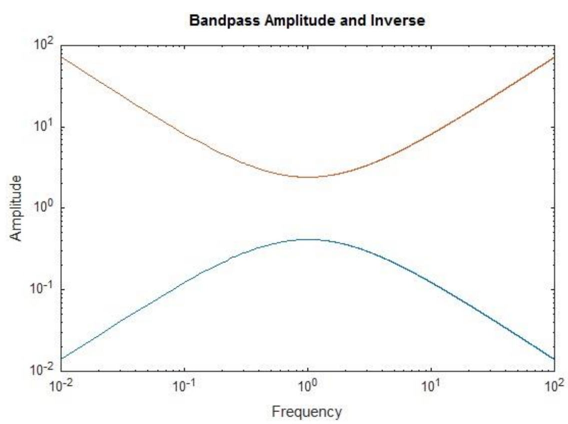


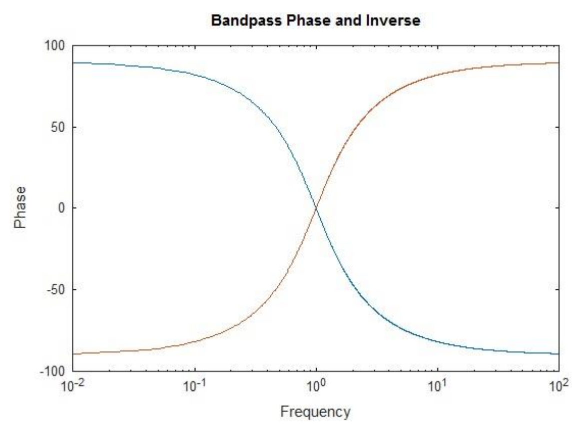
Now, the basic operating principle of the *traditional* dynamic driver is this: The falling slope of the LCR transfer function cancels the (opposing) rising slope of the RC transfer function in the Amplitude vs. Frequency domain, resulting in flat *acoustical power output* between (1) the driver's resonant frequency and (2) the point at which the driver's circumference is approximately equal to the wavelength. This is typically a decade (3 octaves) or so, in terms of frequency response. However, since the two transfer functions (Mechanical LCR and Acoustical RC) are *NOT* mathematical inverses of one another, the *phase response* (and thus, the Time vs. Amplitude performance) of the total system is *badly damaged*. This is the heart of the problem with conventional dynamic drivers, and up to now, it has been considered essentially "impossible to solve," because all drivers have mass and therefore inertia.

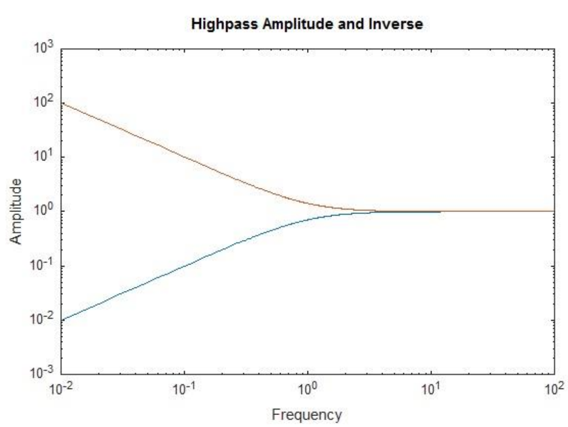
Solving the Dynamic Driver Problem

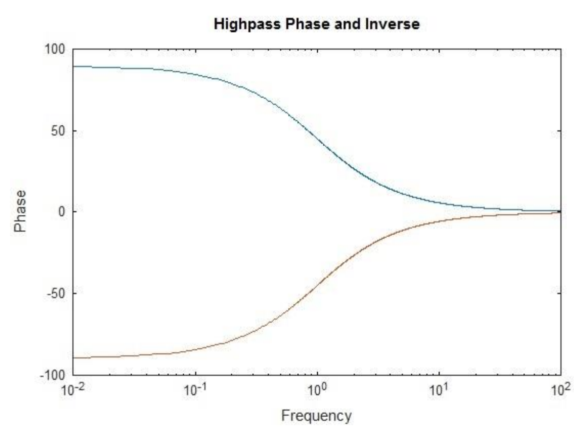
However, there is a way of solving this problem by thinking completely "outside the box" and *solving the problem at its very source*. And it is this: If we apply the exact *inverse* transfer function of the (Mechanical) LCR circuit, in series with the exact *actual* transfer function of the LCR circuit, we instantly convert both its Amplitude and Phase responses to *unity* (as stated above, "f" times "1/f" equals unity, regardless of the definition of "f"). And furthermore, if we apply the exact *inverse* transfer function of the (Acoustical) RC circuit in series with the exact *actual* transfer function of the RC circuit, we instantly convert both its Amplitude and Phase responses to *unity* once again. Both inverse transfer functions can be applied in the (real-world) electrical domain, before the signal ever reaches the loudspeaker (and, being in the electrical domain, will take effect at very nearly the speed of light, vastly faster than is needed to correct problems in the audio frequency range). But because the transfer functions and inverse transfer functions are all in series (in terms of the *combined system*), their effects are all combined (or "cascaded"), with the result that the final *acoustical output* of the AudioMachina XTAC Master Reference System is now virtually identical to the original *electrical input* from the preamplifier, *in real time*. The Time vs. Amplitude problem has been solved in the *purest and cleanest possible way*.

This is shown graphically below, where the blue lines represent the original transfer functions, and the red lines now represent their mathematical inverses. Note that at any frequency, multiplying the two amplitudes results in unity amplitude, and adding the two phase shifts results in zero phase shift. This is, both mathematically and in the real world, unity: There is no significant change in the original, unique form of the music signal, when comparing *electrical input* and *acoustical output* in real time.









By taking this approach, we have eliminated every inherent deviation from pure linearity in the entire "basic model" of the dynamic driver, *in both Time and Amplitude*. If we look at the entire composite (cascaded) transfer function of the complete XTAC system, it becomes essentially a *straight wire with gain*, to use the common phrase. Perhaps even more shocking (at least until you understand the physics and mathematics): The drivers now behave (*in the real world!*) as if they are *essentially massless*. Or, to put it more generally: *Input equals Output, in both Time and Amplitude, with the Acoustical output now being essentially identical to the Electrical input in real time*. And that is the core operating principle of the AudioMachina XTAC Master Reference System, and that is why it is so utterly revolutionary.

Practical Considerations for XTAC in the Real World

While the basic operating principle of XTAC is a revolutionary breakthrough, its real-world form is a carefully balanced optimization of many competing factors, taking many years to achieve. These factors include, among many others:

1. Full frequency range without the use of different driver types or any crossovers.
2. Adequate Dynamic Range and S/N ratio.
3. Absence of diaphragm flexure or breakup in the audible range.
4. Idealized acoustic radiation pattern.
5. Real-world room installation and performance optimization.
6. Reasonable manufacturing and installation difficulty.
7. Reasonable cost.

In the end, there is only one "best answer" to optimize all the competing factors simultaneously, and that is the final form of XTAC. First of all, for several reasons, the system must be made up of a large number of small identical drivers, all operating in acoustical parallel. This is the only way to achieve an idealized radiation pattern while simultaneously achieving sufficient dynamic range and high fidelity in a full-frequency-range system. Once this reality is accepted, then the only two possible idealized physical configurations are spherical (simulated point source) or line-array (simulated line source). And the enormous problem with spherical is that when it is placed in a room, because it is essentially omnidirectional, it has enormous problems with strong early reflections off *all* nearby room surfaces--floors, walls, and ceilings. And when one is interested in true “High Fidelity,” strong early reflections are a very bad thing. Thus, spherical is a really bad solution in the real world. Similarly, a freestanding line array of small drivers, if placed a small distance from walls (and acoustically speaking, a "small distance" is anything under 10 feet, or 3 meters, to any nearby surface), again has enormous problems with strong early wall reflections and standing waves. Thus, a freestanding line array is also a really bad solution in the real world. The only choice left (and by far the best choice in the real world) is to place the line sources at the intersection of room surfaces, thus eliminating the early reflection issue altogether. This approach has proven to have *vastly higher performance and realism, in every way that matters*, than the historical (and now obsolete) "speakers sitting on the floor partway out into the room" approach, because it essentially eliminates *all* strong early reflections from the room acoustics, allowing the original acoustic venue to seemingly transport itself into the listening room.

The full-height corner line array also has several additional benefits:

1. When installed from room boundary to room boundary as designed (normally from floor to ceiling), it launches an essentially ideal cylindrical wavefront into the room, without any significant acoustic interactions with either floor or ceiling, and also without any significant issues from lobing or comb filtering. Thus, to achieve extremely high performance in-room with a standard 2-channel stereo installation, it is *only* necessary to treat the two side walls, two rear vertical corners, and rear wall of the room with basic acoustical treatments (preferably a mixture of standard acoustic absorbers, diffusers, and bass traps). And vertical walls and corners are very easy to treat, relative to floors and ceilings.

2. Because a tall line array can be assembled from multiple small identical line array "modules" without any compromises whatsoever in fidelity, it is possible to achieve a practically ideal full-height floor-to-ceiling line array installation in a room of any height, while still meeting practical concerns in manufacturing, shipping, and installation.

3. In addition, having multiple small identical line array modules makes the requisite total power amplifier output easily "distributable" among multiple smaller (and higher quality) individual amplifier sections, achieving higher fidelity than would otherwise be possible with more powerful amplifiers.

4. Because the boundary-to-boundary corner installation constrains the acoustic wave on 4 sides (left , right, top, and bottom) and forces it to remain essentially purely quarter-cylindrical as it propagates into the room, the acoustic efficiency of the system is greatly enhanced compared to the hemispherical (or, at lower frequencies, essentially omnidirectional) radiation pattern seen in typical historical loudspeaker designs. This greatly improved acoustical room coupling results in enormous gains in both linearity and dynamic range, as driver excursion for a given loudness level is greatly reduced.

5. The ubiquitous "baffle step problem" (the transition from omnidirectional to hemispherical radiation patterns due to the speaker baffle), which causes large (and again unsolvable) disparities between on-axis frequency response and in-room power response in virtually all conventional "box" speakers, again leading to *unnatural sound*, is essentially eliminated outright by the inherent superiority of the full-spectrum uniform cylindrical wavefront of the AudioMachina XTAC Master Reference System.

6. Because of the extremely high uniformity and purity of the in-room cylindrical wavefront, and the almost total lack of destructive interaction with either floor or ceiling, the sound remains the same at any height in the room, from the floor to the ceiling and everywhere in between.

7. Lastly, the almost total lack of early room reflections yields an almost unbelievable increase in the clarity, purity, and intelligibility of both the original music and also the original acoustic venue. The wall between the speakers essentially "disappears" acoustically, leaving behind only the original music and acoustic space. The increase in "naturalness" due to this effect cannot be overstated, and is simply a revelation to those accustomed to traditional loudspeaker designs and traditional room placements.

Conclusion

The AudioMachina XTAC Master Reference System is a revolutionary breakthrough in loudspeaker fidelity, and we believe it to be the first loudspeaker in the history of the world to simultaneously meet, in an optimized form, *ALL* of the fundamental requirements necessary to define a loudspeaker as truly "High Fidelity," and also to simultaneously avoid any obvious faults which can easily be identified as "unnatural" by the human hearing system. We invite you to experience XTAC at your earliest possible opportunity, because if you truly value high fidelity music reproduction, we believe it will change your life forever.

Dr. Karl Schuemann
AudioMachina Inc.
2017