The loudspeaker as musical instrument: an examination of the issues surrounding loudspeaker performance of music in typical rooms

by David Moulton
Sausalito Audio Works, 61C Galli Dr., Novato, CA,
dave@sawonline.com)

Abstract

The loudspeaker is the most important and one of the most variable elements in the electroacoustic music performance process. Nonetheless, its performance is subject to a "willing suspension of disbelief" by listeners. In addition, its behavior and variability are usually not accounted for in assessments of the quality of music reproduction or music instrument synthesis, especially as they occur in small rooms. This paper will examine the aesthetic assumptions underlying loudspeaker usage, the general timbral qualities and sonic characteristics of loudspeakers, as well as some of the issues and problems inherent in loudspeakers' interactions with small rooms and listeners.

Humans' Willing Suspension of Disbelief in regard to loudspeakers

Before we begin our consideration of loudspeakers per se, we need to consider how they are regarded by humans. Here we have a very special case: humans tend not to perceive loudspeakers as sound sources. Instead, we are strongly inclined to perceive the virtual source signals in the program material presented by the loudspeaker. We assign cognitive identification to those virtual sources, including their imagined position(s) and environs. This tendency is pervasive, and it tends to make loudspeakers sonically imperceptible as such. We call this substitution of illusion for reality a "willing suspension of disbelief."

The cognitive act of substitution of imagined signal sources for the physical reality of the loudspeaker generally blinds us to the real qualities and characteristics of the loudspeaker and its behaviors. Fortunately, such an act of cognitive substitution also greatly eases the task faced by recording engineers and producers in terms of creating an illusion of a "previous sonic event." Listeners willingly accept, even seek out, such an illusion and consciously regard the virtual sources as perceived reality in place of the actual physical source.[1] The action is voluntary, but generally not acknowledged by the listener.

Functional Perspectives from which to regard Loudspeakers

There are numerous ways to think of a loudspeaker, based on specific applications and functions. The following perspectives are not mutually exclusive, but rather overlap significantly. However, they emphasize different characteristics of the loudspeaker and illuminate qualities that may not be recognized from other perspectives. These perspectives include: Reinforcer, Mimic, Reproducer, Unique Source, and Musical Instrument.

The Loudspeaker As a Reinforcer.

One view of the loudspeaker is as a "reinforcer." In such a worldview, the loudspeaker "makes other instruments sound louder". Such an approach implies that loudspeakers "change" other instruments, instead of replacing them (the change being to make them louder). While the loudspeaker may actually in fact acoustically replace the source instrument(s) by masking it with a louder iteration of the source sound, it MUST NOT APPEAR to replace it. Such an illusion is central to the use of loudspeakers to support live instruments and voices in sound reinforcement for theater, concert halls and any other live performance.

The Loudspeaker As a Mimic.

The loudspeaker can also be thought of as a "mimic," which is generally the viewpoint held by synthesists and music producers. In such a world-view, the loudspeaker "mimics" the sounds of other sound sources: a flute, a snare drum, a rock band, an orchestra, etc. Questions of accuracy and realism exist. Does the loudspeaker sound "exactly" like a guitar? In this perspective, we are concerned with the loudspeaker's sonic resemblance to other sound sources. We also encounter distinctions between "natural" sound sources (i.e. a flute) and "artificial" sound sources (i.e. the loudspeaker).

The Loudspeaker As a Reproducer.

The most traditional way to think of a loudspeaker is as a "reproducer." As a reproducer, the loudspeaker and its drive mechanism "reproduce" a previous sonic event, as observed by a microphone(s) and stored as an electronic signal trace(s). Instead of being concerned with the loudspeaker's sonic resemblance to other sources, we are concerned with it's accuracy of reproduction of a prior event.

When thinking of a loudspeaker in these terms, questions of "accuracy" and "realism" arise, and the metaphor of loudspeaker as lab instrument is applicable. For instance, how "closely" does the reproduction resemble the original? Is the reproduction indistinguishable from the original? If not, how different is it, and what are the causes of those differences? Such questions, of course, imply the existence of a previous "sonic event" and the answers are also dependent on the recording practices used. In current commercial recording practice, there is often no such unique coherent event.

The Loudspeaker As a Unique Sound Source.

From a fourth viewpoint, we can think of a loudspeaker as a Unique "Sound Source With(out) Its Own Sound Quality." This is an unconventional view that arises from the willing suspension of disbelief, but nonetheless a quite useful perspective for studying the behavior of loudspeakers.

It is important to note that we can simply and unequivocally imagine or remember the sound of a tuba, without making mnemonic reference to any other sound. This is generally impossible to do with loudspeakers. We have no way to "imagine" what a loudspeaker sounds like, except in terms of how it "sounds" relative to its abilities to reproduce, mimic and reinforce other sounds, or in terms of its imperfections.

What is the characteristic sound, or timbre, of a loudspeaker? It has none, or rather, it has an essentially infinite range of such timbres. Therefore, the ideal loudspeaker apparently has NO particular distinctive timbral characteristics. We can distinguish a bassoon from a loudspeaker, for instance, simply because the bassoon has a constrained set of "characteristic" sounds and a specific timbral character and range, while the loudspeaker does not.

This leads us to a description of an idealized loudspeaker, as a device that can create:

• any reasonable sonic spectrum, with

• any reasonable temporal envelope, at

• any reasonable loudness, in

• a reproducible manner.

As such, the loudspeaker can be regarded as a "Universal Sound Generator."

To date, there have been two significant constraints for such a loudspeaker, constraints that both limit its performance and tend to cause it to have both a distinct timbral behavior in spite of the above criteria, as well as significant variability in reproduction.

These constraints are that a typical loudspeaker can only obtain its optimum performance in:

• one direction, and

• in a free field (or very large volume space, where the Schroeder frequency is near or below 20 Hz.).

These constraints are significant because essentially all applications of loudspeakers (for listening to music) occur in small reverberant spaces. Timbre is rendered variable as a function of degraded off-axis spectrum and energy folded back into the room via early and reverberant reflections, at both high and low frequencies (short and long wavelengths, relative to room size).

There is an additional feature of loudspeakers that needs to be noted in this regard. Loudspeakers, when combined in phase-locked arrays, can:

• reproduce or mimic any reasonable acoustic space within another smaller space,

• reasonably represent the phantom position of any reproduced sound source within that virtual space, and

• reasonably envelop the listener within that virtual space.

Therefore, a phase-locked array of loudspeakers can be thought of as a "Universal Sound Environment Generator." The explanations regarding why this is so are beyond the scope of this paper.

It is these qualities of "Universal Sound Generator" and "Universal Sound Environment Generator" that make the loudspeaker so hard to sonically characterize, as well as extremely useful to us. If the loudspeaker can make "any" sound, all sounds would appear to be equally "characteristic" of a loudspeaker. Another way to describe this is to say that the most distinguishing sonic feature of the ideal loudspeaker is its sonic featurelessness!

A requirement for such idealized loudspeakers is, of course, that the differences between them should be inaudible. This condition is necessary if reproducibility of common signals is to be consistent and reliable.

Meanwhile, there is a commonly held view that "accuracy" (which is actually defined as "the absence of error") of reproduction is a desirable characteristic of loudspeaker performance, and that a more accurate reproduction is a better one. From this viewpoint, it follows that there must be a strong positive correlation between "accurate" and "beautiful." If this is true, then we should find that as loudspeaker playback of music becomes more accurate, it should also become more beautiful.

However, current recording practice does not maintain rigorous standards of accuracy. There are no standards of absolute accuracy at all (absolute accuracy referring to conformance to absolute values, such as dB SPL, time-of-day, geophysical location or direction, etc.), and only a few standards of relative accuracy. No records are kept, for instance, of absolute magnitude, direction, or time. Further, little attention is paid to the relative accuracy of these attributes, nor is there any effort expended in maintaining them. They are routinely set aside, in practice, for the purpose of improving the "quality of the production," at, of course, the expense of such accuracy. Similar changes are routinely made to spectra as well.

It has been my experience that while there is a significant overlap between the qualities of "accuracy" and "beauty," particularly at the "poor quality" ends of their particular ranges, the qualities of accuracy and beauty tend to diverge as resolution increases, or at least as resolution reaches a point of diminishing enhanced audibility. In our recording practice AND in our loudspeaker design practice, when faced with this divergence, we generally choose to pursue beauty at the expense of accuracy, for very good artistic, aesthetic and professional reasons.

This is in conflict, of course, with the reproducer modality. Precise, accurate reproduction can be thought of as an activity suitable in the laboratory – the loudspeaker as a laboratory instrument. Such a function is not necessarily musical and can be inherently unmusical. That our current recording practice subverts the reproducer modality through the routine introduction of massive errors as a matter of course indicates how important this conflict is.

This leads us to a fifth viewpoint, the point of view that is the basis of this paper: **the loudspeaker as musical instrument**.

When we think of the loudspeaker from this perspective, we can think about it "playing music," in fulfillment of all of the artistic premises and aesthetic principles that are present in musical performance in general. This turns out to be extremely useful and also quite revealing.

We are able to consider the timbral behavior of the loudspeaker, vis-à-vis other instruments and vis-à-vis the environment. We are further able to examine the aesthetic qualities of "loudspeaker music," and to examine the differences between that particular instrumental genre and other instrumental genres.

In the light of conventional definitions, a loudspeaker certainly qualifies as a musical instrument. A loudspeaker "plays music" subject to the same acoustical laws and principles as any other musical instrument, and with the added characteristic that unlike any other musical instrument, the loudspeaker also undertakes to create "acoustic musical environments." There are, of course, some interesting and unique features of the loudspeaker in this regard, the most important of which is that the loudspeaker is operated by remote mechanical control, in parallel with hundreds or thousands of other loudspeakers in other environments at other times, by an ensemble of performers (the artist, producer and engineer) that does not get to hear each individual loudspeaker or its environment, but can only estimate "what will probably sound good" in the broad range of cases.

At the same time, the loudspeaker also continues to function as Reinforcer, Mimic, Reproducer and/or Universal Sound Generator, depending on the specific application at any given moment.

Traditional Music Performance

Traditional music performance generally takes place in reverberant halls, usually with a raised stage and sufficient volume and floor area to accommodate an economically viable audience in order to support the performance. There are significant differences between the acoustics of and expectations for venues for "acoustic" performance (i.e. "classical" music, et al) and venues for amplified performance (i.e. "rock," "pop" music, et al).

During such events, an audience is present, and the music performance is generally highly ritualized, with extremely clear and well established cultural expectations on the part of the audience and the performers. There is a clear psychological interaction between the performers and the audience.

Venue Sizes and Acoustics

A reasonable minimum dimension for any live music venue is equivalent to a single low-frequency wavelength, or approximately 50 feet. Indeed, few venues do not have a longest dimension that is significantly greater than this distance. Venue sizes rise to >>1,000 square meters and many tens of thousands of cubic meters volume (for cathedrals). Not all venues are suitable for all genres of music. The general acoustic behaviors needed for various types of music venues are comparatively well-known and understood.

Sound Pressure Levels

The noise floor for a venue with audience present is seldom less than 40 dBA SPL (approximately equivalent to NC 35). Individual direct sound pressure levels from acoustic instruments range from 3 dB above whatever noise floor is present to about 110 dB SPL (at 10 feet from the source). Electric instruments often obtain levels 10-15 dB above that level (also at 10 feet from the loudspeaker). Sustained acoustic levels in the reverberant field of a concert hall seldom exceed 110 dB SPL, while peak levels may reach 120 dB SPL. Again, electric amplification yields levels 10-15 dB above that. Therefore, we can characterize the dynamic range of performed music as approximately 60 dB (50-110 dB SPL). It is reasonable to speculate that for electrically amplified performance the noise floor goes up as the level goes up and that a similar, or even lesser, dynamic range exists for "electric" genres, even though a greater dynamic range could be available for such genres.

Instrumental and Ensemble interactions with Venue

For acoustic musical performance, considerable attention is usually paid to the acoustics of the venue. Concerns pertaining to the arrival time of early reflections in the hall, lateral reflected energy, on-stage acoustics (for the benefit of the performers), and the duration and spectra of the reverberant decay all are addressed. In venues used for amplified performance, such concerns are a bit simpler, consisting primarily of attempting to reduce the amount of reflected and reverberant information as much as possible, and leaving such elements under the control of the amplified sound reinforcement system.

Loudspeakers Playing Back Music

Loudspeakers playing music mostly exist in a different set of environments (except when they are being used to reinforce acoustic instruments in live performance) than live performance venues. It is useful to consider the physical conditions and constraints of loudspeakers in such environments.

Loudspeakers Play Music In Small Rooms, usually simulating larger concert venues of all types and sizes

Loudspeakers generally play back music in small rooms, i.e. rooms whose largest dimension is significantly less than 50 feet. This means that, in general, the longest wavelengths will be folded back into the acoustic space to create interference patterns before they have fully propagated, for an inherently degraded reproduction.

An important feature of these rooms, vis-à-vis live performance venues, is their variability. While live performance venues may have some commonalities and consideration given to their suitability for music performance, the great majority of small rooms in which loudspeakers are placed have no acoustical attention paid to them at all, other than possible sound isolation standards. Further, the variability in shape, volume, surface treatment, noise floor and other relevant characteristics are significantly greater than for live performance venues. In fact, the only general commonalities that can be described for such small rooms are a volume of less than 300 cubic meters, a ceiling height of less than 3 meters and a predisposition (in residential cases) for rectangular enclosures.

In arrays, loudspeakers are generally called upon to generate reproduced sound environments of spaces with significantly greater volumes than the volume of the playback space, and usually those reproduced environments have much longer reverberant times than the playback spaces (2000-5000 vs. 100-400 milliseconds). The spatial cues for such large virtual spaces are transmitted to the listener by both the direct sound and early reflections in the playback room, and the integrity of those cues is dependent on both high and low frequency spectral content.

Directivity

In general, loudspeakers generate sound omnidirectionally at long wavelengths (>5 feet/<200 Hz.) and increasingly directionally at short wavelengths (<1 foot/>1 kHz.). A traditional design goal has been to maintain flat on-axis amplitude response. Therefore, power response has necessarily rolled off as a function of frequency. This significantly affects the interaction of the traditional loudspeaker with the room in which it is placed. Both early reflections and reverberant energy derived from loudspeaker emission have significantly degraded high frequency content.

Sound Pressure Levels, Sensitivity and Power

Loudspeakers are typically rated for sensitivity by measuring their output at 1 meter on axis with a signal input of either 1 Watt (power) or 2.83 Volts (the energy needed to generate 1 Watt with an 8 ohm load), in a free field. Typical sensitivities for domestic loudspeakers range from approximately 80 to 92 dB SPL. A typical stereophonic array of domestic loudspeakers will yield an output 4-5 dB above that when each speaker is driven by 1 Watt, and program is correlated at low frequencies (which is typical). Maximum output is generally limited by amplifier power, and by the limitations imposed by the crest factor of the program (maximum sustained average levels are usually 10 db below peak level, except for "hyper-compressed" recordings[2]). Assuming 100 Watts of power per channel and a sensitivity of 88 dB SPL (both typical), a single loudspeaker will generate a peak output of 108 dB SPL at 1 meter and a sustained level of 98 dB SPL. Two loudspeakers will yield approximately 102 dB SPL sustained level. At 3 meters (a typical listening distance) the <u>direct</u> output from the two loudspeakers will attenuate by approximately 9 dB, to 93 dB SPL. Depending on the Critical Distance in the room, the direct–plus-reverberant level will be approximately 3 dB above that, or 96 dB SPL. This means that in a typical reasonably high-quality domestic listening situation the maximum sustained output will probably be no greater than 96 dB SPL at the listening position, with peaks reaching 106 dB SPL. This is approximately 14 dB below the maximum outputs of live performances of acoustic ensembles and 24-30 dB below the performance levels of "electric" ensembles.

Linearity and Distortion

Compared to electronic circuitry, loudspeakers are comparatively non-linear in their behavior. Transducers exhibit non-linear behaviors as a function of their motional limits, spectral contents at or beyond the limits of the pass-band, thermal changes in the transducer, and asymmetries and irregularities in the motion of the transducer surface. A "good" loudspeaker exhibits harmonic and intermodulation distortion artifacts of somewhat less than 1% of total amplitude through much of its dynamic range and distortion onset levels are usually determined by a 3% distortion threshold (in contrast, a power amplifier with distortion artifacts greater than .01% would be considered poor). As the physical limits of the various transducers are approached, non-linearity increases significantly for the various reasons noted above.

Accuracy

The term "accuracy," as applied to loudspeakers, is generally taken to mean "spectral resemblance" to various source instruments. This usage is technically a misnomer (as noted above), and systematic studies and evaluations of accuracy, through rigorous blind comparisons of a loudspeaker with a variety of source instruments and ensembles, are simply not undertaken, due to their difficulty, expense and perceived irrelevance.

Loudspeakers can readily be distinguished from source instruments in most cases, and in direct comparisons are found to sound distinctly different from other instruments.[3,4] This is probably due to (a) differences in directivity between the instruments and (b) differences in the nature, size and location(s) of the instruments' vibrating surfaces.

There are two further issues with accuracy, as it pertains to loudness. First, single loudspeakers (and small arrays) cannot reasonably approach the sound pressure levels obtained by large or amplified ensembles. Such a failing introduces large errors in magnitude, of course, but also spectral errors dues to the changes in frequency response of the human auditory system as a function of loudness. Second, "accurate" sound pressure levels may not be appropriate (or even legal) for much playback. Therefore, it is axiomatic that music playback via loudspeaker may be attenuated by as much as 60 dB (relative to a live performance) for a variety of playback environments. Such attenuation and its sensory and spectral effects are generally ignored in considerations of accuracy. Those effects are nonetheless quite real, and have a significant impact on the "quality of reproduction" of music.

Stereophony and Surround Sound

Stereophony and surround sound involve the use of an array of loudspeakers operating in phase-locked synchrony. Such synchrony is unique to loudspeaker arrays among arrays of acoustic sound sources, and it enables loudspeakers to generate complex sound artifact volleys which in turn permit humans to perceive two particular sensations unique to loudspeakers: phantom images and phantom environments/envelopment. The effect of these images and environments is profound on human listeners, and constitutes one of the most powerful qualities of loudspeaker music.

The Timbral Character of the Loudspeaker

Given the above, the following informal observations can be made about the timbral behavior of traditional loudspeakers:

Divided Spectrum

No single transducer can reliably reproduce a bandwidth of greater than 3 octaves (8:1) without electronic intervention. Therefore, common practice is to distribute the ten-octave audible spectrum among multiple transducers. Timbre is audibly affected by the choice of transducers, their assigned bandwidth, the crossover circuitry used to distribute spectra to the transducers, and their relative positions in an assembled loudspeaker system.

Formants

All acoustical instruments have distinctive sets of formants, or resonant structures that are a consequence of the size, shape and materials of the instrument. Such formant sets are, in general, independent of the pitch or frequency content of a given note. The spectral character of the formant set is probably more important to the perceived timbre of a given instrument than is the specific spectrum of any given note played on that instrument. In fact, the individual note spectra (and waveform) change dramatically from note to note.

The loudspeaker, on the other hand, seeks to have no formants or resonant structures, except as induced by ported low frequency enclosure topologies and as a natural, if unintended, consequence of crossover behaviors (i.e. the effects related to two drivers at different points in space generating the same frequency, a behavior that is unique to loudspeakers among musical instruments). This speaks to the Universal Sound Generator quality of loudspeakers. Unlike all other instruments, the acoustically observed spectrum and shape of any given waveform played by the loudspeaker should remain constant, regardless of frequency.

Directivity at low, mid and high frequencies

The output of loudspeakers tends to be omnidirectional at low frequencies and unidirectional at high frequencies, yielding a power response that declines with increasing frequency. Significant interference effects exist at crossover frequencies. These directional behaviors may be the single most characteristic timbral element of loudspeakers, vis-à-vis other instruments.

Crossover Design

Crossover design and implementation has a pervasive effect on timbre. Numerous schools of thought and practice exist. These all introduce variability into the timbral behavior of loudspeakers.

Low Frequency Performance

Low frequency performance of loudspeakers is generally compromised by the desire to reduce the cost and size of loudspeakers. Numerous schools of thought and practice exist regarding low-frequency transducer/enclosure topology, again having a significant effect on timbre.

Interaction with rooms at low frequencies.

A loudspeaker's performance varies as a function of the environment it is operating within, its position within that environment, and the individual frequency content at any given moment. There is no perfect solution to the problem of this variability.

A Brief comparison with conventional acoustic instruments.

Below is a table showing informal estimations of the frequency ranges, approximate directivity and sound pressure level of a variety of acoustical instruments vis-à-vis both a typical good loudspeaker and an ideal one.

Instrument	Low freq. fundamental	High freq. fundamental	Overtones to:	Low freq. radiation directivity	High freq. radiation directivity	Min dB SPL (3m)	Max dB SPL (3m)
Flute	250	2K	10k	omni, transverse	forward horizontal	50	90
Clarinet	150	1.2K	5K	omni, downward	lateral horizontal	50	90
Oboe	250	1.2K	15K	weak omni	omni	60	85
Bassoon	75	400	15K	weak omni	omni	50	80
Saxophone, tenor	125	600	15K	omni, forward	lateral horizontal	55	90
Trumpet	180	1.2K	20k+	forward	lateral, vertical	60	100
French Horn	100	800	5K	rearward	rear, lateral	55	95
Trombone	40	500	20K	forward	lateral, vertical	60	105
Tuba	30	250	8K	forward	omni	60	90
Violin	200	2K	20K+	weak omni	Vertical	45	85
Viola	150	1.5K	20K+	weak omni	Vertical	45	80
Cello	60	1K	20K+	forward omni	Forward	45	85
Doublebass	40	500	20K+	forward omni	Forward	45	80
Snare Drum	200	4K	20K+	vertical, horizontal dipole	omni	55	105
Cymbal	500	10K	20K+	vertical, horizontal dipole	omni	50	95
Tympani	40	250	1K	omni	weak omni	40	95
Glockenspiel	250	1K	20K+	omni	omni	60	95
Marimba	250	1K	2K	omni	weak omni	50	80
Xylophone	250	1K	5K	omni	omni	50	85
Triangle	1K	2K	20K+	none	omni	50	90
Bells	200	800	15K	omni, downward	horizontal plane	50	80
Bass Drum	20	100	2K	horizontal, vertical dipole	weak omni	40	105
Piano	30	2K	15K	omni, vertical, horizontal dipole	omni, forward	40	100
Pipe Organ	20	2K	12K	omni	omni	50	105
Loudspeaker (typical traditional)	**50**	**20K**	**-**	**omni**	**narrow forward**	**0 dB**	**99**
Loudspeaker (ideal)	**20**	**20K**	**-**	**omni**	**horizontal**	**0 dB**	**120**

Table 1: Various acoustical instruments vis-à-vis a typical “good” loudspeaker and an “ideal” loudspeaker, showing the range limits of their various fundamental frequencies and informal estimations of the upper limits of their significant harmonics, and their patterns of polar radiation at low and high frequencies, plus estimations of the minimum and maximum Sound Pressure Levels such instruments can obtain as observed at 3 meters in a free field.

This table illuminates several interesting things.

First, all traditional instruments generate complex waveforms with extensive overtone structures. As a general rule, the frequency range for the fundamental frequencies of musical notes spans four octaves, from 65 to 1040 Hz. (low C to high C), while overtones extend up to five octaves above that range. Loudspeakers, on the other hand, do not generate overtones as such, except when driven into distortion.

Second, the directivity of acoustical instruments varies widely at both high and low frequencies. There is no uniformity, and no particular trend. In a collective sense, we can generalize that high frequencies are radiated laterally and upward, occasionally forward, while low frequency radiation tends to be either bidirectional or somewhat directional as a function of horn coupling with free space. It needs also to be noted that the timbral results of such directivity is also affected by the comparatively large size of performance venues – the strident harshness of a violin in a small practice room, with extremely strong high frequency early reflections off a low ceiling is in marked contrast to the warmth of the same instrument in a concert hall, with the attenuated vertical high frequency radiation lending a silky reverberant patina to the sound.

In contrast, traditional loudspeakers have a comparatively distinctive pattern of directivity, as noted above, a pattern generally unlike almost all acoustical instruments. While low and mid frequencies are radiated approximately omnidirectionally, high frequencies are beamed narrowly, so narrowly that for the most part a listener cannot sit more than 15° off-axis and still perceive the spectrum the designer intended. The result of this is a particular interaction with rooms that sounds quite unlike other instruments.

It is therefore quite important to note that no other instrument requires such a precise orientation by the listener as does the loudspeaker, simply in order to take in the spectral range of the loudspeaker. Also, as noted earlier, the volley of early reflections are spectrally deficient vis-à-vis the on-axis radiation. This has led to compensations, in many instances, during the act of recording and production. By tradition, many recordings are overly bright to fill in for this spectral deficiency and to compensate for off-axis listening.

Finally, when we compare the maximum sound pressure levels that can be obtained by loudspeakers vs. acoustical instruments, we see that loudspeakers have a much larger dynamic range and appear to be equivalent or greater in range in almost all regards. However, comparisons here are extremely tricky when we start accounting for variables such as concert hall vs. playback room, listeners’ distance from the source, etc. Further, instruments aggregate into ensembles that may include as many as 250 members (orchestra and chorus). Finally, the subjective perceived loudness from many sources is not easily matched by the phantom images generated by a loudspeaker pair.

At best, loudspeakers can generate sound pressure levels in a small listening room that effectively mimic the subjective quality of loudness obtained by a full orchestra, but it is safe to say that they cannot replicate the amplitude at bandwidth that an orchestra can achieve. Further, the mimicry of a live rock or heavy metal band is always going to be reduced in subjective loudness (and this is probably as it should be).

We can further generalize that the amplifier power available, when coupled with the given sensitivity of most domestic loudspeakers, is usually insufficient to achieve such levels, introducing a further constraint on loudspeaker playback.

The Aesthetics of Loudspeaker Music

Loudspeaker music can be thought of as a genre, with its own characteristics, style and aesthetics. Because of the willing suspension of disbelief that is present with loudspeakers, the aesthetic qualities of loudspeaker music are not obvious. Similarly, the idiomatic tendencies and gestures that distinguish loudspeaker music from other instrumental genres tend to be obscured as well. Nonetheless, they are present and quite distinctive.

We can begin a brief examination of the aesthetics of the genre by considering the aesthetics implicit in the relevant functional applications we described earlier.

When we listen to music from loudspeakers while regarding the loudspeaker as a reproducer, we generally desire to experience the equivalent of a "musical event" such as a concert or club performance. We wish to experience the sensation of "being there" (or that "they are here"). In short, we have expectations that are related to our expectations for experiencing "live" music. We would like very much to be "transported" into that virtual concert milieu by the illusion yielded by the loudspeakers.

When we listen to music from loudspeakers while regarding the loudspeakers as mimics, we desire to experience "exactly" the sensation we would have listening to the "source" instrument(s). We would like the loudspeaker(s) and our room to "disappear," and to have only the sensation of "musical performers" and "their space." This is very closely related, of course, to the aesthetic desire inherent in the reproducer modality. I speculate that this perspective is most appropriate when listening to recordings of solo instruments and small ensembles.

Loudspeaker Music As a Musical Genre

When listening to popular music, which exists primarily as loudspeaker music, we generally expect and desire to experience "the song," and we generally accept without any qualms that what we are listening to comes from loudspeakers only and is NOT a live performance. We enjoy the "entertainment effect" of music produced for loudspeaker playback.

When listening to recorded classical music, we generally expect and desire to experience music as a private entertainment, a la the Esterhazys (Haydn's wealthy patrons), but without the expense and complication of resident composer and performers. We enjoy the "perceived quality" of the recorded performance, the virtual performance acoustics and sonic artifacts, as well as the perceived quality of our own physical surroundings.

In this case, we regard the recording as a transcription from another medium, with the willing suspension of disbelief firmly in place and little engagement with the new medium. To "transcribe" means (a) "to adapt or arrange a [musical] composition for a voice or an instrument other than the original," and (b) "to record for [playback] at a later date" (American Heritage Dictionary). Both definitions are relevant and applicable to music played back by loudspeakers, but we tend to ignore the first definition, that process of adaptation and change in medium. Nonetheless, it is reasonable and correct to view recordings as "an arrangement of music for an instrument other than the original."

As a general rule when transcribing, we need to consider both the nature of the original AND of the transcribed instrument. The best transcriptions illuminate both qualities of the original that might not have been obvious and qualities that were not available in the original, while maintaining the essential musical identity and "gestalt" of the original. Such considerations are perfectly true for recordings, as well as for more traditional transcriptions. We are "adapting" music to loudspeakers from its original form. From this view, "loudspeaker as reproducer" is not an appropriate perspective.

This raises some questions about loudspeaker music. What sorts of musical qualities are most idiomatic for loudspeakers (and the recording process)? What sorts of musical qualities are least idiomatic (i.e. don't sound very good) for loudspeakers? What is the social and aesthetic experience of "listening to loudspeakers," in comparison to "listening to live players?" What qualities of loudspeakers enhance the "loudspeaker listening" experience? It is in the consideration of these questions that we can find some insight into the nature of "loudspeaker music," as a distinct instrumental genre of music, as distinct, clearly, as "piano music," "orchestral music," or "choral music," and also distinct from issues of musical style.

Let us compare live performance vs. loudspeaker performances from this standpoint for a moment.

• Live music is public and usually occurs in crowded venues.

• Live music is highly social and ritualized.

• Live music has a strong emotional interaction between listeners and performers.

• Live music is mostly limited by human capabilities for performance (except when sound reinforcement is used).

• Live music is a one-time event not under the listener's control.

At the same time:

• Loudspeaker music is generally performed in private.

• Loudspeaker music is casual, ubiquitous and often extremely intimate.

• Loudspeaker music has no interaction between listeners and performers.

• Loudspeaker music is not constrained by human performance limitations.

• Loudspeaker music is under the direct control of the listener, who can vary it's spectrum and level at will. It can be played on demand, stopped, restarted and repeated exactly, ad infinitum.

• Loudspeaker music easily and often becomes internalized by the listener.

We do the bulk of our listening to loudspeakers in private spaces. We do NOT, as a rule, gather in public social groups to listen to loudspeakers (excluding dance clubs and aerobics classes!). Loudspeaker music is casual, ubiquitous and often extremely intimate. We regularly use loudspeaker music as a very strong "mood enhancer" for love-making, dining, meditation and other private ceremonies and activities, including "serious listening." Further, we often use loudspeaker music in our homes and automobiles as a kind of "sonic perfume" (not to mention its use in stores, elevators, restaurants, telephones and AV presentations – sonic pollution as sonic perfume, or vice versa?).

Due to the machine nature of the loudspeaker drive mechanism, the performance does not vary as a function of our listening response. Further, in production, we polish performances to idealize them into best possible performances that lack live blemishes.

Loudspeaker music is not constrained by human performance limitations. Loudspeakers can play indefinitely, at any reasonable level (subject to system design limits). Loudspeakers can play higher, lower, louder, softer, faster and slower than any other instruments.

Because loudspeakers can be played on demand and repeated ad infinitum, recorded performances often become internalized, and given "recordings" obtain mythic status. Listeners often "memorize" recordings (something that almost never happens with live performance). Listeners can and do regularly "program" listening events, or sequences of recordings.

Finally, and most importantly, we must consider the remarkable emergence of loudspeaker music over the past century. In 1900, loudspeaker music (i.e. phonograph performances of music) constituted a very small portion of all music performed, surely less than 1%. A century later, it is reasonable to estimate conservatively that 99.9% of all music experienced is loudspeaker music, and that such music has become ubiquitous. This means that over the past 100 years loudspeaker music has almost entirely displaced "non-loudspeaker music."

Due to the willing suspension of disbelief, this displacement has happened essentially without notice – our willing suspension of disbelief coupled with our preoccupation with "loudspeaker as reproducer" has masked this change in instrumental usage. We now "listen to loudspeaker music" as our primary, often only, musical activity. Further, we even have the ironical situation now in place where much live music is produced, promoted and presented for the express purpose of encouraging the loudspeaker music equivalents of that live music (i.e. "touring to sell the record").

This watershed evolution is, to my mind, the single most important change to occur in the history of musical instruments and musical media. That it has happened without notice simply makes it all the more remarkable!

What Kinds of Musical Gestures Work Well For Loudspeakers, and What Kinds of Gestures Don't?

A.

What Is Idiomatic for single loudspeakers (mono)?	What Is NOT Idiomatic for single loudspeakers (mono)?
Simple sound sources	Complex sources (i.e. many sound sources in a space)
Simple acoustics	Complex acoustics
Simple doublings	Complex textures, doublings, polyphony
Single miking techniques (i.e. "acoustic mixing")	Complex mic arrays in a common space, especially with cross-talk
Consonant music	Dissonant music
Music with clear upper spectrum definition	Music lacking upper spectrum, or with rich mid and upper-bass harmonic content.
Music with moderate crest factors.	Music with extreme crest factors (very high or very low).
"Dry" sounds	"Wet" sounds

B.

What Is Idiomatic for Stereo and Multichannel	What Is NOT Idiomatic for Stereo and Multichannel
Simple arrays of instruments	Complex arrays of instruments in blend
Use of time-domain localization	Use of amplitude cues for precise localization
"Phantom LF" vs. "Discrete HF"	Complex stereophony/blends of multiple stereo and mono recordings from a single space
Spectral Management	Complex ambiences and reverberance, particularly with excessive mid and LF content.
Spatial Distribution	
Simple stereo/mc Ambience	
Simple stereo/mc Reverberance	
"Acoustic mixing" via minimalist stereo and multichannel miking techniques	

Tables 2A and 2B. Musical Gestures and recording techniques that are comparatively idiomatic (column 1) and not idiomatic (column 2) for single loudspeakers (A) and loudspeaker arrays (B).

Many desirable techniques for stereophony are quite undesirable for mono. The problem areas that arise when summing stereophonic signals into monaural ones can be characterized as:

- Changes in level

• Changes in timbre

• Changes in reverberance.

These changes occur due to destructive interference in summation to mono. Time and amplitude difference cues become timbral cues and interference patterns when summed in the drive mechanism for loudspeakers. Therefore, such practices are often suppressed, even though they are idiomatic in stereophony. This is simply in anticipation of the multiple possible modes of playback.

"Musical Instrument Quality" or "Museum-Quality" Loudspeakers:

Music instrument quality loudspeakers are loudspeakers that are equivalent to fine musical instruments in design, function, build quality and appearance. They are appropriate for the playback of music in small concert and recital halls, museums, galleries, etc, as well as private homes.

Such loudspeakers must have

• excellent sonic performance, which can be characterized as:

  • excellent spectral and timbral range, envelope and dynamic flexibility (i.e. "tonal quality"),
  • excellent linearity over their dynamic range, and
  • ability to reasonably "evoke," if not replicate, the sound pressure levels obtained by source ensembles;

• excellent acoustical interaction with the playback room, which can be characterized as:

  • integration into the playback space for management of low frequency response issues, and
  • excellent frontal and horizontal power response at high frequencies, to maintain the spectral quality of phantom images and environments, and to enhance the quality and stability of position of such images and sense of phantom environment;

• "fine quality" appearance. The loudspeaker that would be a fine musical instrument must have the same attention paid to its appearance and build quality as any fine acoustical instrument.

When embracing this worldview, we need to recognize that the loudspeaker is NOT a generic black box, but a singular and distinctive musical instrument (think of "loudspeaker as concert grand piano"), and that the sound quality we perceive comes PRIMARILY from the loudspeaker and via the room, not from the quality of its supporting electronics. Selection of an appropriate loudspeaker becomes the primary musical quality decision for enjoying loudspeaker music.

---

1 Moulton, David - Numerous informal classroom experiments with students in Critical Listening and other audio classes. When asked to identify sound sources, no student has ever identified a loudspeaker, even though all examples have been via loudspeakers.

[2] Katz, Bob, "Mastering Audio, The Art and Science," pp. 128-132, Focal Press, Oxford, 2002.

[3] Moulton, David – "Eight Vocal Microphones Tested and Compared," Home and Studio Recording Magazine, June, 1994.

[4] Moulton, David. "Speaker, Speaker, On The Wall, Who Sounds Coolest Of Them All?," Home and Studio Recording Magazine, June, 1994.