NWC NAFEMS WORLD CONGRESS 2019
incorporating SPDM International Conference Simulation Process & Data Management
A WORLD OF ENGINEERING SIMULATION
17-20 JUNE | QUEBEC CITY | CANADA

Virtual Systems Engineering for Professional Audio Applications

S. W. Hutt
(Mvoid Group, USA)

A. J. Svobodnik
(Mvoid Group, Austria)

T. Nizzoli
(Mvoid Group, Italy)

M.-O. Chauveau
(Mvoid Group, France)

Abstract

This presentation discusses simulations of audio system analysis over the audible frequency spectrum in professional audio applications. A unique splicing technique to merge low to mid frequency FEA with mid to high frequency ray tracing is introduced that results in full frequency band analysis at any listening point in a room. This enables a realistic prediction of the performance of the audio system on a fully virtual basis.

A description of the workflow for virtual product development based on the V-Model will be outlined. It starts with a description on (sub-) component level, i.e. a multiphysical model of an electrodynamic transducer (including compression drivers and nonlinearities in the mechanical, electrical and acoustical domain), that is then coupled to an acoustic enclosure (covering also ported cabinets for woofers for reproduction of low frequencies). The (sub-) system "transducer + cabinet = loudspeaker" is then simulated in a free field radiation condition as a single system (i.e. loudspeaker), and finally assembled to an array of several loudspeakers mounted together, called a line array.

This line array of loudspeakers is then simulated in a listening room (a concert hall for musical presentations). To cover the whole audible frequency response, FEA as well as ray tracing is being coupled, and eventually a room impulse response at arbitrary locations in the listening space is calculated.

Additionally, through auralization a realistic simulation of the audio system playing into the sound field, including also spatial attributes, may be realized. This requires the inclusion of spatial effects from head and torso, as given by means of head related transfer functions based on a semi-analytical approach. Finally, a virtual tuning process is discussed that provides access to modify and enhance the audio system for optimized performance in the listening space in a fully virtual domain without having access to the real domain/system by using advanced signal processing techniques for equalization and delay and gain settings for each active channel.

1. Introduction

The term "Professional Audio" or "Pro Audio" is a term referring to application of amplification of audio signals in pretty much any situation outside of a home. This may be reproduction of recorded sound such as in recording studios, cinema, clubs etc. or reinforcement of a live music, speech, special effects or artistic performance. In all cases the loudspeaker systems need to be tailored to the venue. Virtual Product Development processes have been developed that reduce development time and cost for all aspects of system development. Further, through analysis virtual products may be optimized to refine the designs to meet multiple best practice objectives. Given the multidiscipline characteristics of audio systems this is a particularly valuable tool to achieve optimal objectives.

In the product development concept phase, where no physical hardware exists, engineers will have access to digital prototypes that represent the functional behavior of products in a realistic way, so that product definition and specification can be refined and improved in the virtual domain.

Following progress in the application of advanced numerical analysis techniques for the computer-aided development of products [1], this digitalization trend and the more and more intensive and efficient use of engineering analysis can be seen in all major industries [2], with probably automotive and aerospace industries as pioneers and trendsetters. Also, in terms of applicable physics, there has been much progress. While the focus in engineering analysis was first on structural mechanics for stiffness and stress analysis, nowadays almost all physical domains of mechanical engineering applications are covered, and even strongly coupled multiphysical simulations are on the way becoming a standard to support industrial development tasks.

This success story of digitalization has led to a change in the product development paradigm towards a simulation driven approach. Advanced engineering analysis methods have been continuously moved upfront in the development process to significantly increase engineering efficiency and improve product performance. This enables engineers for a wide variety of products to perform a virtual assessment of product quality and thus opens the door for improvements based on digital prototypes at early development stages.

Multiphysical engineering analysis for the simulation of loudspeakers has been introduced successfully to the audio industry in the last decade [3].A realistic simulation model of an automotive audio system including the listening environment is today available, and used on a regular daily basis in the automotive industry. With the simulation of the passenger compartment of an automobile, having been mastered [4], the method has now been applied to pro audio.

However, defining and assessing product quality of audio systems (i.e. reproduced sound quality) is very challenging. There is no general closed form definition of sound quality, especially not in terms of a metric based on numbers where different systems can be easily compared and judged. Typically, the assessment of sound quality is based on subjective methods by means of listening tests [5]. This is fundamentally supported by a statement out of [6]:

Any sound, noise, music, or in general any signal generated, transmitted, radiated and perceived can more precisely be interpreted and compared by people if it is made audible instead of discussing "levels in frequency bands", "single number quantities" or "dB(A)".

This requirement for the capability of being able to listen to an audio system has a dramatic impact on the simulation models that are required. The generation of audible sound files or sound streams to perform listening tests of digital prototypes is a new requirement for the simulation model. So-called auralization techniques [6] and [7] need to be introduced and coupled to the multiphysical simulation model.

In the following, a complete virtual development environment for automotive audio systems is presented, that ultimately allows listening to a virtual prototype and thus opens the capability of assessing the product quality in a very early development stage where no real prototypes in hardware exist.

2. A Multidisciplinary Virtual Development Environment for Professional Audio Systems

The requirements for listening to virtual prototypes extends a multiphysical approach to a multidisciplinary one, as additional engineering disciplines are required. Figure 1 shows and overview of the building blocks of a multidisciplinary development environment for professional audio systems. The following engineering disciplines need to be added to multiphysical engineering analysis to be able to listen to an automotive audio system:

- Digital signal processing
- Psychoacoustics
- Subjective evaluation
- Sound tuning
- Auralization (binaural audio)

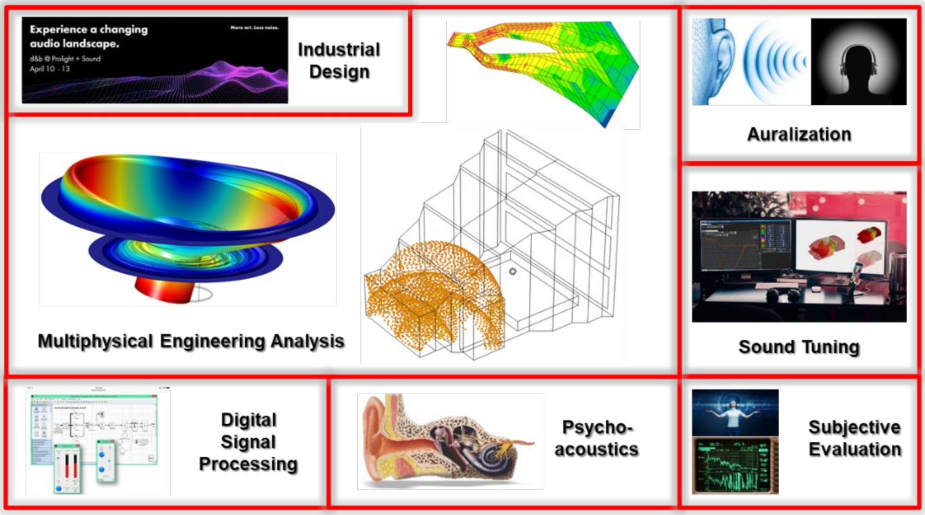

Figure 1: Overview of the building blocks of a multidisciplinary development environment

3. Workflow Process Steps

The multiphysical CAE-based simulation model needs to cover the electroacoustics of a loudspeaker, as well as the acoustical and mechanical interaction with loudspeaker enclosures and the listening space (vibro-acoustics). Thus, we follow a workflow process of the following physical domains, coupled within this multiphysical model:

- Electromagnetics
- Structural vibrations
- Acoustics (propagation of sound waves)

A bottom up approach needs to be followed in a methodical workflow process, starting with the subcomponents of the transducers.

Details of this so-called vibro-electroacoustic model have been published by one of the authors, and we refer to [8]. However, it should be noted, that especially the modeling of the venue is a major challenge because the whole audible frequency range, extending from 20 Hz up to 20 kHz, needs to be covered. Typically, ray tracing methods for the acoustics of venues can only be used down to 100 Hz, so that alternative numerical procedures need to be applied for low frequency simulations. As the auralization is done in the time domain, either we need a simulation in the time domain as well, or a model in the frequency domain that contains magnitude and phase information, enabling a transformation back into the time domain by means of an inverse Fourier transformation. Hence, we use finite element methods for simulating the low frequency spectrum.

It must be noted here, that in general we use a mixture of numerical procedures in the time domain and in the frequency domain to retain the required accuracy for a realistic listening experience but still preserving an efficient solution without too excessive computer power. This needs of course proper tools to transform from the frequency domain to the time domain and vice versa, so that the resulting impulse responses, and frequency responses respectively, contain contributions from both numerical schemes. As we will see later, usually both types of information, frequency or impulse responses, are useful for assessing sound quality. The driving factor here is primarily the capability and limitation of a numerical scheme for calculating an acoustic response.

Multiphysical engineering analyses are used to deliver room impulse responses (RIR), used as a starting point for the auralization process. This RIR needs to cover the vibro-electroacoustic sound path from the speaker terminals, where voltage is applied, to specific measurement points in the seating area of the venue. It is of crucial importance, that this impulse response is accurate enough so that all major psychoacoustically features are covered, and a realistic auralization is possible. The formal definition of this accuracy has been derived by means of listening tests based on A/B comparisons between auralization and real rooms by the authors.

Theoretically, this RIR could be directly used to perform an auralization. However, the raw response of loudspeakers is greatly influenced by interaction with the listening space boundaries. This interaction is very dependent on venue size where for example the acoustics of rooms have significantly more influence on the perceived sound quality than do large rooms or open spaces. Additionally, boundary materials and even varying audience size affects acoustics. As a result, each amplifier channel that drives one or more loudspeakers needs to be enhanced by means of advanced digital signal processing methods to optimize system coherence.

The following topics need to be accounted for to make the workflow comprehensive in terms of product development:

- **Architecture Definition**
    - Transducers
    - Cabinet dimensions
    - Acoustic targets
    - Venue

- **Transducer & Waveguide Engineering**
    - 3D CAD Geometry
    - Clean assembly CAD that setup for meshing & simulation, i.e. no interference fits etc.
    - Physics
    - Materials laws & properties
    - Magnet
    - Soft parts
        - Suspensions
        - Membrane
    - Advanced stage beyond piston analysis (3D models)

- **System Engineering**
    - Cabinet
        - Ports, waveguides, phase plugs, etc.
    - LF/MF/HF tuning
        - Utilization of DSP to optimize/correct performance
    - Validating results to meet product goals
        - Acoustic output & directivity
    - System Integration
        - Cabinet to cabinet coupling

- **Application Engineering**
    - Room interaction & auralization.
    - Virtual tuning
    - Auralization

4. The Complete Simulation Model

While talking about general acoustic principles and their particularities when applied to a room, we are confronted by many degrees of complexity. Its particular geometry, including cavities, edges, small and wide angles between uneven surfaces with very different absorption and reflection properties due to a very complex mix of materials. One of the key attributes to the success of the our approach is the careful consideration of boundary conditions.

Another requirement for room simulations is the need to cover the complete audible frequency range. In fact, the classic Finite Elements Analysis (FEA) fails due to numerical pollution at frequencies above 500Hz to 1kHz. Therefore, we uniquely combine low and mid frequency simulations in the frequency domain based on FEA, and high frequency simulations in the time domain based on geometrical acoustics.

Finally, the degree of complexity and the level of detail considered is another key to the success of our approach. The ability to model a very wide dimensional scale of elements (down to the 0.1mm air gap between the voice coil and magnet of a single loudspeaker, up to the 10 to 40 m length of a venue), complicated mechanical structures and detailed physical phenomena (e.g. loudspeaker nonlinearities) can yield to very complex simulation models.

For this reason, we have developed a "step-by-step" approach. The first step, which is called the concept model, aims at the true reproduction of the room acoustics of a given room. All loudspeaker sources are being considered as perfect pistons, perfectly integrated into a structure as a rigid configuration. Figure 2 depicts a concert hall with line arrays in each corner at left. As a result, the concept model will simply show us the best possible system in a given acoustic environment. Once this is achieved and the locations of all loudspeakers have been chosen, we can consider adding, step by step, further details to the concept model.

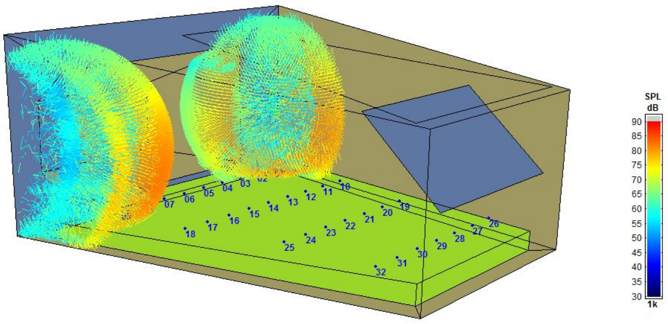

Figure 2: Concert hall with line arrays

The following paragraphs give some information on the different simulation methods used in this approach: their working principle and use cases as well as their limits.

Finite Element Analysis (FEA)

The Finite Element Analysis is one of the best-known deterministic techniques to solve many engineering problems. Vibro-acoustic FEA calculates the behavior of acoustic waves in fluids, as well as the interaction between the fluid and the structure. The acoustic domain is closed, and a coupled structure can be totally or partly connected to the acoustic fluid. The solver predicts the sound wave propagation and the structural vibration. A vibro-acoustic FEA model usually consists of a shell triangular mesh (2D) representing the vibrating structure and of a tetrahedral volume mesh (3D) which discretizes the fluid domain. The use of vibro-acoustic FEA is giving a correct representation of the modal behavior of enclosures and of the reflection and diffraction acoustic phenomena.

But the FEA approach has some limitations: the element dimensions have to be chosen upper frequency dependent. Therefore, when the frequency increases, the wavelength decreases and the number of elements increases, increasing the risk of numerical errors and/or reaching computational (hardware) resources.

In our approach, FEA is used for modeling the room behavior below approx. 100 Hz (depending on the venue size). The material properties as well as the boundary conditions, which take the form of normal impedances for the acoustic domain and displacement constraints for the structural domain, are applied. FEA is also used for modeling the structural behavior of loudspeakers above 1kHz, where the LPM approach fails.

Geometrical acoustics (GA)

Geometrical acoustics is a field of acoustics studying the propagation of sound based on the concept of rays (or cones), which are considered as lines along which the acoustic energy is transported. It is similar to the concept of geometrical optics, which studies light propagation in terms of rays.

This method allows to calculate the path of sound waves through a system with regions of varying propagation velocity, absorption characteristics, and reflecting surfaces. Under these circumstances, wave fronts may bend, change direction, or reflect off surfaces. Ray tracing solves the problem by repeatedly advancing idealized narrow beams called rays through the medium by discrete amounts.

When a ray encounters an object on its path, its energy is divided into three components:

- the absorbed part
- the transmitted part
- the reflected part

Reflections are of two types: the specular reflection, for which the angle between the reflected ray and the normal of the reflecting surface equals the one of the incident rays and the diffuse (or scattering) reflection, where the same normal law applies, but the reflected energy is diffused over multiple path due to the roughness of the surface.

The relative amount of each part of the acoustic energy is depending on both the material properties of the encountered object and the frequency of the considered wave.

However, ray tracing is only valid when the wavelength is small compared to the characteristic dimensions of inhomogeneous inclusions through which the sound propagates as this model does not take into consideration diffraction effects and modal behavior.

In our approach, the ray tracing model is typically used at frequencies above 100 Hz. The loudspeakers are modelled following the LPM approach extended by their directional characteristics, whereas boundary surfaces are defined by an absorption and a scattering coefficient.

5. Virtual Tuning of Professional Audio Systems

The major factors in perceived sound quality from [9], [10] and from our practical experience are:

- Frequency response smoothness on and off-axis
- Perceived bass extension
- Naturalness

Naturalness can be defined as "the feeling of space", i.e., the correct sense of ambience, or the appropriate levels of direct and reflected information.

Misalignment of loudspeakers

There is no single speaker that can reproduce the whole audible frequency range with balanced response. As a result, several loudspeakers of different sizes (subwoofers, woofers, midranges and tweeters) are used, each covering a specific bandwidth. Ideally, they should be placed at the same general location in close proximity to each other and time aligned so that differences in distance to the listener position are minimal, and the arrival is a single wave front arriving and of requisite amplitude of each frequency band. Ultimately, the directivity of a loudspeaker system is derived from the some of the acoustic sources. Figure 3 depicts a generic example of a single enclosure intended to be used in multiples in a line array system.

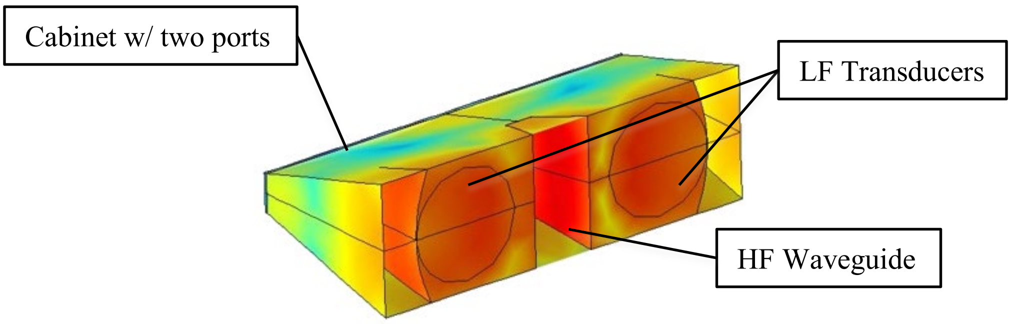

Figure 3: Generic example of a single enclosure

Figure 4 shows the polar axis with 2D and 3D polar dispersion characteristics of the single enclosure.

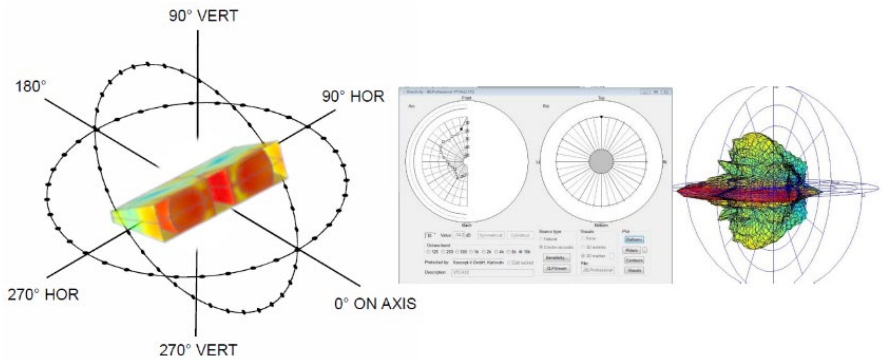

Figure 4: Generic example of a single enclosure

Installation space and transportation requirements of professional loudspeakers face packaging and integration constraints the available installation space for loudspeakers is always kept at a minimum, resulting in small enclosures for woofers for reproducing low frequencies. These constraints have a severe impact on the vibrations of the speaker, and thus deteriorate frequency response smoothness and bandwidth.

By using methods of digital signal processing, these degradations in sound quality can be improved. This process of improving the perceived sound quality of an audio system is called sound tuning and is mandatory for every premium automotive audio system. Frequency response smoothness can be enhanced by equalization based on diverse filters. Equalization here means changing the sound pressure level at frequencies where we have significant peaks and notches by means of electronic filters that can boost or cut specific regions of frequencies.

Filters are also used for crossing over different types of speakers acting in different frequency regions. As said before, one single loudspeaker cannot cover the whole audible frequency range. Audio systems usually have up to three(four)-way arrangements ((subwoofer), woofer, midrange and tweeter). It is therefore necessary to delimit the field of action of each contributor by properly setting high-pass and low-pass filters. Doing so, one will have to take specific care of the phase alignment of all speakers in order to achieve a good overlapping.

The other two main features used in the tuning process are:

- Changes in channel gain
  That is done to compensate for different sound pressure levels of channels, e.g. caused by misalignment of loudspeakers.

- Adding channel delay
  The signal is delayed compensating for differences in the arrival time of sound waves from different loudspeakers.

The third major factor "naturalness" can be classified as a spatial attribute. There is almost no indication of spatial attributes in a frequency response. And also in the time domain, by looking at the RIR, it is almost impossible to judge spatial attributes of an audio system.

While the evaluation of the spectral performance of an audio system purely based on visual data, i.e. frequency responses visualized by X/Y graphs, is to some extent possible, there is no alternative to a listening test to evaluate spatial attributes. However, as in the concept phase no hardware exists, in early development phases there is only a possibility for a listening test in the virtual domain by means of auralizations based on the simulated RIRs.

Line Arrays

A line array loudspeaker system may be comprised of many devices stacked vertically so that the wavefronts of the individual loudspeaker drivers or waveguides interact with each other to form an acoustic beam with narrow vertical dispersion. The vertical directivity is dependent on the height of the array vs. frequency.

Figure 5 shows a photo of a professional line array suspended in the air. The photo shows a high frequency waveguide in the center of each enclosure with low frequency woofers at the left and right of the enclosures. In this example eight enclosures are coupled together to form a somewhat cylindrical wavefront.

Figure 5: Professional line array suspended in the air

6. Auralization of Audio Systems

For a realistic virtual listening experience, all major psychoacoustic features need to be included, and thus we cannot directly use the RIRs at ear locations. Spatial effects and localization of sound is based on binaural hearing, and such effects are not included in RIRs. At our two ears, sound events arrive with differences in time and level, caused by reflection and diffraction due to our head and torso. These differences, which are called interaural time difference (ITD) and interaural level difference (ILD), make significant changes to the incident sound waves, and are dependent on direction.

Binaural cues, because of diffraction of incident sound waves on head and torso, are added to the ear sound pressure. The amount of cues (referenced to a plane wave) is described by the so-called head-related transfer function (HRTF) 15. HRTFs describe the relationship of sound pressure at the eardrum or ear canal entrance to an incident plane wave without reflection and diffraction effects, i.e. without head and torso.

These HRTFs can either be used by means of measurements from individual persons, which might then not fit well for listeners with different physiology, artificial heads (so-called dummy heads), or, as we propose here, based on an analytical model with simplified geometry 15. Figure 6 **Fehler! Textmarke nicht definiert.**shows a set-up, where a rigid sphere is being used as a head model. Due to the simplified geometry, it is possible to derive closed form solutions for the approximation of left ear’s and right ear’s HRTFs due to time and level differences of incident sound waves. It must be noted, that the model is a full 3D model and thus HRTFs vary with azimuth (i.e. horizontally) as well as elevation (vertically) for arriving sound waves.

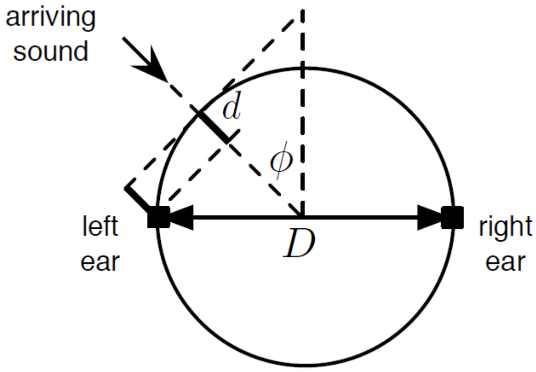

Figure 6: HRTF model set-up

Now, as the HRTFs are known, it is possible to calculate the BRIR for left and right ear by applying the HRTFs to the simulated RIRs coming from our multiphysical simulation model. Despite typical binaural audio applications, where usually only one microphone located at the center of the head is used to derive a single RIR, we use two RIRs from the simulation model corresponding to the ear canal entrance location of left and right ear. Thus, ITD is directly based on the simulation model, resulting in an improved accuracy and a more natural listening experience.

An additional advantage of our analytical HRTF model is that the sound path through the ear canal is not included. At the time during playback through headphones, this part of the sound path is added by the listener's own ear canal physiology. Our HRTF model actually calculates the sound at the ear canal entrance, which is very close to the transducer used in a headphone, and thus gives superior results. If a measured HRTF, either from individual persons or artificial heads, were being used, the ear canal sound path would be twice included during playback and would need to be excluded once.

The HRTF filters are implemented in the discrete time domain, using first order IIR filters and fourth order Lagrange fractional delays, to enable real-time processing. Real-time processing is of crucial importance to realize a virtual reality (VR) like environment for tuning an audio system.

The BRIRs are finally convolved with acoustical test files, sound files containing music, speech or noise, and ultimately a binaural listening experience is created by sending the final signal to headphones.

Figure 7 shows the signal flow of our auralization environment. First, the test signals are sent to a routing matrix, distributing the two-channel (stereo) signal to the individual channels of the sound system. Then each channel's tuning parameters are applied, and the resulting signal is sent through the simulation data, actually the RIRs, and finally binaurally rendered by means of HRTFs, and ultimately send to headphones for listening.

Figure 7: Signal flow for a VR-like tuning environment

All building blocks are designed for real-time processing, so that by changing any tuning parameter, we immediately get a change in the visual response, i.e. the graphics showing the frequency response, as well as a change in the audible response via the headphones. Thus, we have an acoustically VR-like development environment generated for listening to audio systems based on purely computer-generated models.

7. Summary and Conclusions

Professional audio systems and acoustics in rooms has been reviewed and basic elements of a VR-like audio environment for real-time virtual sound tunings and listening tests of professional audio systems has been described. Additional engineering disciplines were added to multiphysical engineering analysis to enable the audibility of audio systems already at concept level. The resulting room impulse responses from advanced numerical analyses based on CAE methods were enriched by binaural effects. Especially for professional applications, where tuning of the audio system is important to improve the perceived sound quality mainly because of limitations due to small room acoustics and constraints in designed space, signal processing and findings from psychoacoustics as well as from subjective evaluation were added as well.

With such a virtual listening system, it is possible to assess the product quality of audio systems already in the concept phase, where no porotypes in hardware exist. Today, listening tests are the only real possibility to rate audio quality – there exists no alternative based on a system that uses a quality metric defined by numbers. Such a virtual listening system therefore opens the possibility of significant improvements at lower engineering and product costs especially compared to design changes at a later stage in the development process, closer to the start of production.

This fully virtual multidisciplinary development environment for professional audio systems has already been applied to first real-life industrial applications, and its maturity could be validated. This approach is an efficient tool for substantial improvements in engineering efficiency resulting in the realization of significant business benefits.

8. References

[1] Turner, M. J., Clough, Ray W., Martin, H. C., and Topp, L. J. (1956). *Stiffness and deflection analysis of complex structures. In: Journal of the Aeronautical Sciences*. Volume 23, Number 9. p. 805–823

[2] Newton, P. (2016). *The NAFEMS Simulation Capability Survey 2015*, NAFEMS

[3] Svobodnik, A. J. (2005). *Acoustic Matrix Methods for Woofers, Tweeters, Horns and Small Transducers*. Audio Engineering Society

[4] Svobodnik, A. J. (2011). *CAE as Key Technology for the Multidisciplinary Virtual Product Development of Automotive Audio Systems*. NAFEMS

[5] Shively, R. (1998). *Subjective Evaluation of Reproduced Sounds in Automotive Spaces*. Audio Engineering Society

[6] Vorländer, M. (2008), *Auralization*, Springer-Verlag

[7] Kleiner, M., Dalenbäck, B.-I., Svensson, P. (1991), *Auralization – An Overview*. Audio Engineering Society

[8] Svobodnik, A. J. (2013). Virtual Development of Audio Systems – Application of CAE Methods

[9] Olive, S. (2004). *A Multiple Regression Model for Predicting Loudspeaker Preference Using Objective Measurements*. Audio Engineering Society

[10] Klippel, W. (1990). *Assessing the Subjectively Perceived Loudspeaker Quality on the Basis of Objective Parameters*. Audio Engineering Society

[11] Shaw, E. A. G. (1982). *External Ear Response and Sound Localization*. Amphora Press

[12] Faller, C., Menzer, F., Tournery, C. (2015). *Binaural Audio with Relative and Pseudo Head Tracking*. Audio Engineering Society

