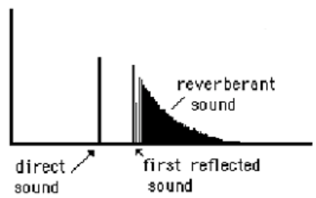
# A GENERAL SOUND REPRODUCTION FRAMEWORK

We have been able to eliminate sweet spots, free up loudspeaker placement restrictions and create reproduced sound fields of placed objects that stay put when you move and turn as in reality. The resulting formats are not overly complex but new improved loudspeakers are required. The new formats are completely compatible with the surround formats and can in fact be used to correct some shortcomings of these, particularly relating to speaker placement issues and sweet spot location – that is if the listener wants sweet spot targeted reproduction, of course.

To achieve this, new rendering equipment is also required. But we must remind ourselves that at this stage we have only considered the direct sound field component from the sources. We now need to consider the remaining sound field components.

### What other sound components are there?

The traditional framework for classification of sound types considers the direct, early reflection and reverberant field components. These are generally considered separate because of their differing behaviours. One aspect not formally treated in the classification is to distinguish low frequencies. We do this because of the lack of human directional perception for low frequency sound means that all the low frequency sound field components, - the direct, early reflections and reverberant sound field parts can be considered as a group from the point of view of creating placed sound objects.




This is not so with the medium to high frequency components. Low frequency sound reproduction will be treated separately in a later paper.

## RECONSTRUCTING THE COMPLETE SOUND FIELD

### Direct sound component

We understand the direct sound field component. This is the first arrival sound and travels directly to the listener from the sources. Given that VWF can place sound sources in space so that the correct listener perspective is provided, the direct sound field component is taken care of.

### Early reflection component

The early reflection sound is usually the second arrival, having been reflected by objects and boundaries in the vicinity of the sources. What we know of the early reflections is that the apparent source of each reflection is not the location of the original sources. Each early reflection will be modified by specular (hard) and diffuse (illuminatory) reflection components. The diffuse reflection component will come from the point of impact on the reflecting surface. The specular reflection will appear to come from an image point behind the surface. Both reflection types will then interact with the directly radiating sound field components to create colourations that characterise the capture environment and each listener location. The characterization has both location and spectral modification parts that relate to the properties of the reflective surfaces.

Given that with VWF we can arbitrarily place sound sources at fixed locations in space, we can now place each and every one of the reflection sources correctly at their real and virtual locations. As we are also correctly placing each of the direct sound sources in space, the resulting interacting sound field will behave correctly for each and every listener location just as was the case in the original environment for each location.

This does not necessarily mean that we have solved the problem of early reflection reproduction. We still need to locate the early reflection sources to behave as if we are listening in the capture environment. Fortunately a number of factors assist here.

First, most capture auditoria have dominant early reflections, and these are predominantly specular in nature. Accurately placing the sources and so reproducing the first, say 10 early reflections may well suffice.

Second, we could spatially characterise the auditorium and so derive its spatial transfer function. This could be done once and prior to any performance as it will largely be a characteristic of the building and the general region that any sources (orchestra etc) will be placed. This could be done by either actual measurement or by acoustic modelling of the structure.

Third, if we have appropriate microphones, these will automatically capture the location of all sources including the early reflections [[1]].

If a suitable microphone is used for the recording, the set of all reflections "seen" by the microphone will be correctly captured from the microphone vantage point and therefore will automatically be correctly placed at reproduction. Unfortunately the result, whilst very good will be less than ideal because the reflections will be those "seen" at the microphone vantage point and will not correctly scale with listener relocation in the listening environment at any other listener location. All listeners bar the one at the position corresponding to the capture location will receive "a view of a view".

Another alternative is to model or measure the early reflections themselves and then use post processing to correctly but artificially place the early reflections at reproduction. By this means the early reflections can be correctly placed or rendered to represent the original set for all listeners regardless of location.

This approach would enable the parameterising of both real and virtual recording environments and thus the ability to recreate these environments at will.

Metadata (data about the venue acoustic behaviour) could be used to distribute the characteristics of any desired venue and the location of the microphones at the same time as the recorded sound signals themselves.

This could be of great value to the gaming and virtual reality markets. You could be exploring a virtual cave in darkness and could sense approaching a wall, then, for a quick change, switch to the intimacy of a phone booth, for example. Other treatments would also be possible. Scene changes in movies and games could be accompanied by appropriate early reflection acoustic behaviour changes.

---

[1] Sound field capture and microphones will be the subject of later papers.

Whilst accurate reproduction of the early reflections in recording venues is now possible with VWF, in practice, we find that for applications other than gaming and virtual environments, all but the worst early reflections are completely overwritten by the reverberant field components.

### Reverberant field component

The medium to high frequency reverberant component of the sound field is different. It is effectively "what is left over" after the direct sound and the early reflections. The reverberant sound field part is related to the sound from all the original sources but contains no directional information. It gives the impression of the capture venue size and nature through decay time constants and spectral acoustic colouring over time. A different set of rules applies to correctly reproducing the reverberant sound field than those for the direct and early arrival sound when listeners are free to move and turn in the reproduction environment.

The reverberant sound field is the result of many reflections from objects and surfaces in the original listening environment. It must be related to the sum of all the sources present in the recording environment, but is by definition characteristically lacking in any discernible acoustic directional information for any listener location.

With a correctly reproduced reverberant sound field, the listener should not sense any directional information when moving around in the listening environment. This means that the source of the reverberant sound field must be hard to find. An implication of this would be that the placement of such a source would have to be non-critical.

The reverberant sound field will be considered next.

Graeme Huon HuonLabs 2008