7.3 Theory

This section describes the background to the IR synthesis performed by the PP module with some references to literature.

Impulse response scale and calibration

Created Impulse Responses (IRs) are normalized by not including the acoustical input spectrum (`Lp1m_a` in the source file) since the it is replaced by the spectrum of the anechoic file at convolution. By removing the acoustical source spectrum it is also reasonable to switch anechoic material as long as the implied source has a directivity pattern similar to the one given in the source-file.

The finally heard processed music/speech has been scaled in several steps, see Fig. 7.2.

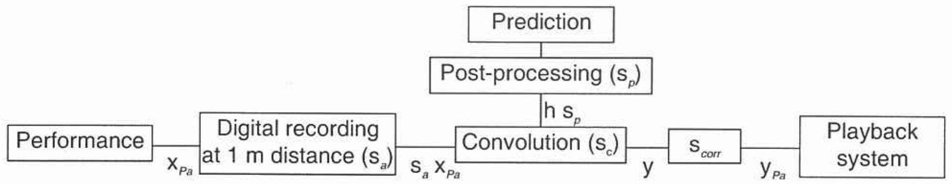

Fig. 7.2 principal scale factors from anechoic recording to audible auralization.

$$y = (x_{Pa}\, s_a)(h\, s_p)\, s_c \qquad (7.3\text{-}1)$$

Where

y final processed music/speech 16-bit integer format

x_{Pa} true pressure signal of anechoic music/speech material at 1 m in front of the real source [Pa]. x_{Pa} includes an extra, typically unknown, scale-factor if the recording was made at another distance.

s_a scale-factor when x_{Pa} was digitally recorded to 16-bit integer format [Pa^{-1}]. For the majority of available anechoic material this factor is unknown.

h impulse response as described above

s_p scale-factor applied at the post-processing in the PP module. The reason for this factor is the 24-bit integer SIM-file format used by the *Lake DSP* convolvers. The factor is saved in the SIM-format output-files header created by the PP module. The SIM-file file information utility displays the value and the calibration and scaling utilities updates it.

s_c scale-factor applied at convolution to create 16-bit integer format. When using *Lake DSP* hardware this factor is unity. When software convolution is used this value is stored in the SCL-file with the same base-name as the processed material. The WAV-file information utility displays this value.

To arrive at absolute calibration the total scaling is thus:

$$y_{Pa} = y s_{corr} \;\; ; \;\; s_{corr} = \left(s_a s_p s_c\right)^{-1} \qquad (7.3\text{-}2)$$

Where

y_{Pa} is the true pressure signal at the receiver position used in the particular hall modeled [Pa]. The playback system must then create this signal at the ears of a listener.

s_{corr} is the correction scale-factor [Pa]

As can be seen from the various scale-factors above, absolute calibration may be very difficult to achieve and requires calibrated amplifiers and well known anechoic source material. Presently, the head-phone equalization filters applied for a binaural receiver are also not calibrated in level.

Relative calibration for a particular anechoic source signal is, however, achievable. The relative level variations within a hall (or across halls) can be preserved by observing the two factors s_p and s_c.

Since the software convolution auto-scales for maximum dynamic range the calibration for software convolution must be done after convolution by observing $s_p s_c$. The WAV-file calibration and addition utilities do this automatically and, by using the corresponding SCL-files, allows only for calibration of processed files that have exactly the same anechoic source-file.

When using *Lake DSP* hardware the calibration must be done of the IRs before convolution so that the scaling of all responses are equal.

See the post-processing flow illustration in Fig. 7.1 for an overview of both desktop/software and *Lake DSP* auralization including relative calibration.

Reflection path (RP) transfer function

The prediction module calculates early part octave-band echograms (125 Hz to 16 kHz). These echograms are saved in the ECH-file together with air absorption, reverberation time in octave-bands and other information.

Reflection path transfer function magnitude

Depending on reflection order, two different methods for reflection path magnitude synthesis methods are employed.

For the direct sound (if present) and first-order specular reflections

Frequency lines in-between the octave-band center-frequencies are interpolated using a cubic spline. To assure a smooth interpolation it is made for log(magnitude) vs. log(frequency). The interpolated function is then returned to lin(magnitude) vs. lin(frequency) and sampled at the requested number of points (half of reflection transform-size).

The part below the lower limit of the 125 Hz octave-band (88 Hz) can be treated in two ways: either suppression by a raised cosine bell or a flat response down to 0 Hz, see Fig. 7.3.