Audio Engineering Society
Convention Paper
Presented at the 126th Convention
2009 May 7–10 Munich, Germany

Generalized multiple sweep measurement

Stefan Weinzierl[1], Andre Giese[1], Alexander Lindau[1]

[1] Audio Communication Group, TU Berlin
Sekr. EN-08, Einsteinufer 17c, 10587 Berlin, Germany

stefan.weinzierl@tu-berlin.de, andre@fosburyflop.de, alexander.lindau@tu-berlin.de

ABSTRACT

A system identification by impulse response measurements with multiple sound source configurations can benefit greatly from time-efficient measurement procedures. An optimized method by interleaving and overlapping of multiple exponential sweeps (MESM) used as excitation signals was presented by Majdak et al. (2007). For single system identifications, however, much higher signal-to-noise ratios (SNR) can be reached with sweeps whose magnitude spectra are adapted to the background noise spectrum of the acoustical environment, as proposed by Müller & Massarani (2001). We investigated on which conditions and to what extent the efficiency of multiple sweep measurements can be increased by using arbitrary, spectrally adapted sweeps. An extension of the MESM approach towards generalized sweep spectra is presented, along with a recommended measurement procedure and a prediction of the efficiency of multiple sweep measurements depending on typical measurement conditions.

1. INTRODUCTION

Fast methods for measuring impulse responses of multichannel sound systems are increasingly important. This holds true for the measurement of loudspeaker arrays, e.g. individualized HRTF measurements using circular rigs of multiple loudspeakers [7], multichannel sound reinforcement systems or compound sources used to generate arbitrary source directivities [4][5]. All require measurements with multiple input sources. Dynamic binaural simulations require a high number of binaural room impulse response (BRIR) measurements for different head positions. The measurement duration, which is already considerable, multiplies with the number of source positions, if measurements are done step by step [2][3]. The measurement of a complete loudspeaker array for wavefield synthesis can include several hundreds of sound sources.

To increase the efficiency of multichannel impulse response measurement, Majdak et al. have presented an optimized method by interleaving and overlapping of logarithmic sine sweeps used as excitation signals. Due to specific properties of the sweep measurement, this procedure, called Multiple Exponential Sweep Method (MESM) by the authors, was only investigated for logarithmic sweeps. It can be used either to reduce the total measurement duration without loosing signal-to-noise ratio, or to increase the signal-to-noise ratio while keeping the measurement duration constant.

However, the signal-to-noise ratio of sweep measurements also depends on the spectral shape of the sweep. Especially if the spectral shape can be adapted to the spectral distribution of the background noise and to the sensitivity of the loudspeaker used for excitation, as proposed by Müller & Massarani [6], a much higher signal-to-noise ratio can be attained, depending on the equipment used and the measurement conditions available.

For this reason, the use of multiple, temporally interlaced sweeps has been reconsidered with respect to arbitrarily shaped sweeps, in order to investigate, to what extent multiple sweeps can be used to increase the efficiency of multichannel system identification through a combination of overlapping and spectral adaption.

2. SWEEP MEASUREMENT AND DISTORTION

When measuring impulse responses of room-acoustical or electro-acoustical systems such as auditoria or sound reinforcement systems, the system is usually regarded as largely linear and time-invariant. Thus, the system "loudspeaker – room - microphone" can be excited by a determined stimulus, and the complex transfer function can be derived by dividing the response spectrum by the stimulus spectrum (see figure 1). These FFT based measurement techniques can be used with arbitrary stimuli as long as they do contain sufficient energy in all frequency bands of interest.

Sine sweeps turned out to be particularly suited for acoustical measurements, because a) they exhibit a small crest factor of 3 dB thus providing a high signal-to-noise ratio that can be raised by increasing the sweep duration, b) their magnitude spectrum can be shaped in the frequency domain by adjusting the group delay according to the desired spectral energy distribution, c) they are much less sensitive to time-invariance than random or pseudo-random excitation signals such as MLS and d) the product of non-linearities in the signal chain can be discarded by windowing of the deconvolved impulse response. The properties of two popular types of sweeps are shown in figure 2.

For system identification, the measured system output $Y(e^{j\omega})$ is divided by the reference spectrum $X(e^{j\omega})$ to yield the complex transfer function $H(e^{j\omega})$. This corresponds to a multiplication with the inverse signal $X^{-1}(e^{j\omega})$ with

$$X^{-1}(e^{j\omega}) = \frac{1}{X(e^{j\omega})} = \frac{X(e^{-j\omega})}{\left|X(e^{j\omega})^2\right|} \qquad (1)$$

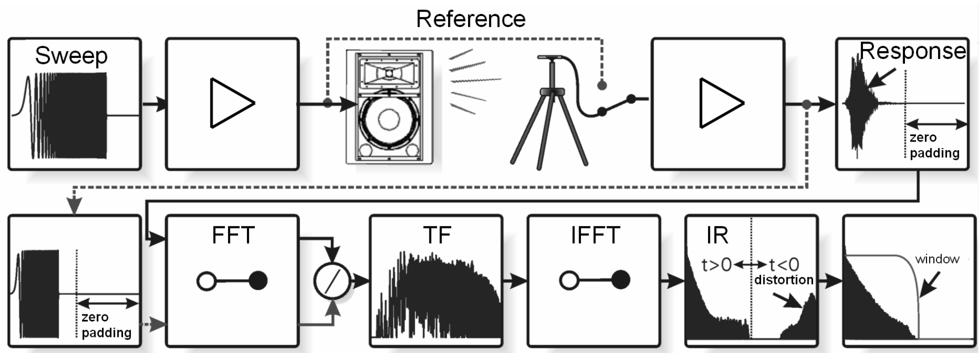

Figure 1: FFT-based measurement of acoustical transfer functions (TF) and impulse responses (IR) with subsequent removal of non-linear products of the signal chain by windowing (adapted from Müller 2008[1]). Aliasing due to the cyclic deconvolution is prevented by zero padding.

In the following, the reader is free to interpret ω as angular frequency in s^{-1} and time parameters (T, τ) in s or ω as normalized frequency in rad and time parameters (T, τ) in samples.

If we assume that an inversion of $X(e^{j\omega})$ is possible, i.e. $X(e^{j\omega}) \neq 0$, the inverse signal $X^{-1}(e^{j\omega})$ has negative logarithmic amplitude, phase and group delay compared to the original signal $X(e^{j\omega})$. Hence, by multiplying the system response $Y(e^{j\omega})$ with $X^{-1}(e^{j\omega})$, the spectrum $|Y(e^{j\omega})|$ will be shaped with the reciprocal magnitude spectrum of the excitation signal, and the group delays of $Y(e^{j\omega})$ and $X^{-1}(e^{j\omega})$ will be added.

Since harmonic distortion products with order k and frequency ω are generated by the fundamental frequency ω/k with group delay $\tau_g(\omega/k)$ but are folded back with the (higher) group delay $\tau_g(\omega)$ of their own frequency ω, they will appear in the deconvolved system response $h(t)$ at negative times given by

$$\tau_k(\omega) = \tau_g(\omega/k) - \tau_g(\omega) \tag{2}$$

$\tau_k(\omega)$ yields the resulting group delay of all deconvolved distortion products. Since $\tau_k(\omega)$ depends only on the group delay of the applied excitation signal, it shall be inspected for some typical types of sweeps.

We will consider sweeps with constant temporal envelope but individual magnitude spectrum. They are favourable, because they can be adapted to the spectrum of the underlying noise floor, thus yielding a largely frequency-independent signal-to-noise ratio while preserving the low crest factor of 3 dB for a sinusoid with constant peak amplitude. Their magnitude spectrum can be constructed in the frequency domain by setting the group delay $\tau_g(\omega)$ to

$$\tau_g(\omega) = \tau_g(\omega_1) + C \int_{\nu=\omega_1}^{\omega} \left|X(e^{j\omega})\right|^2 \mathrm{d}\nu \tag{3}$$

normalized by a constant C with

$$C = \frac{T}{\int_{\omega=\omega_1}^{\omega_2} \left|X(e^{j\omega})\right|^2 \mathrm{d}\omega}, \tag{4}$$

given by the sweep duration

$$T = \tau_g(\omega_2) - \tau_g(\omega_1), \tag{5}$$

the total sweep energy

$$\int_{\omega_1}^{\omega_2} \left|X(e^{j\omega})\right|^2 \mathrm{d}\omega$$

and start and stop frequencies ω_1 and ω_2.

Müller and Massarani (2001) have given a discrete formulation of (3) for direct implementation in code by setting $\tau_g(\omega_1)$ to an arbitrary value and defining $\tau_g(\omega_i)$ through iteration with

$$\tau_g(\omega_i) = \tau_g(\omega_{i-1}) + C' \cdot \left|X(e^{j\omega_i})\right|^2 \tag{6}$$

and

$$C' = \frac{T}{\sum_{\omega=\omega_1}^{\omega_2} \left|X(e^{j\omega})\right|^2} \tag{7}$$

For simplicity we will stay in the continuous domain for the following considerations. However, a discrete, algorithmic solution can easily be obtained by exchanging integrals for sums.

In the following, the deconvolved distortion products for different types of sweeps shall be inspected a little closer, because their form and extension are crucial for further considerations with respect to the overlapping of subsequent impulse response measurements.

2.1. Linear sweeps

Linear sweeps (figure 2, left column) have a linearly increasing instantaneous frequency and a white magnitude spectrum with a power density spectrum given by

$$\left|X(e^{j\omega})\right|^2 = D_{LinS} \tag{8}$$

related to the total sweep energy by

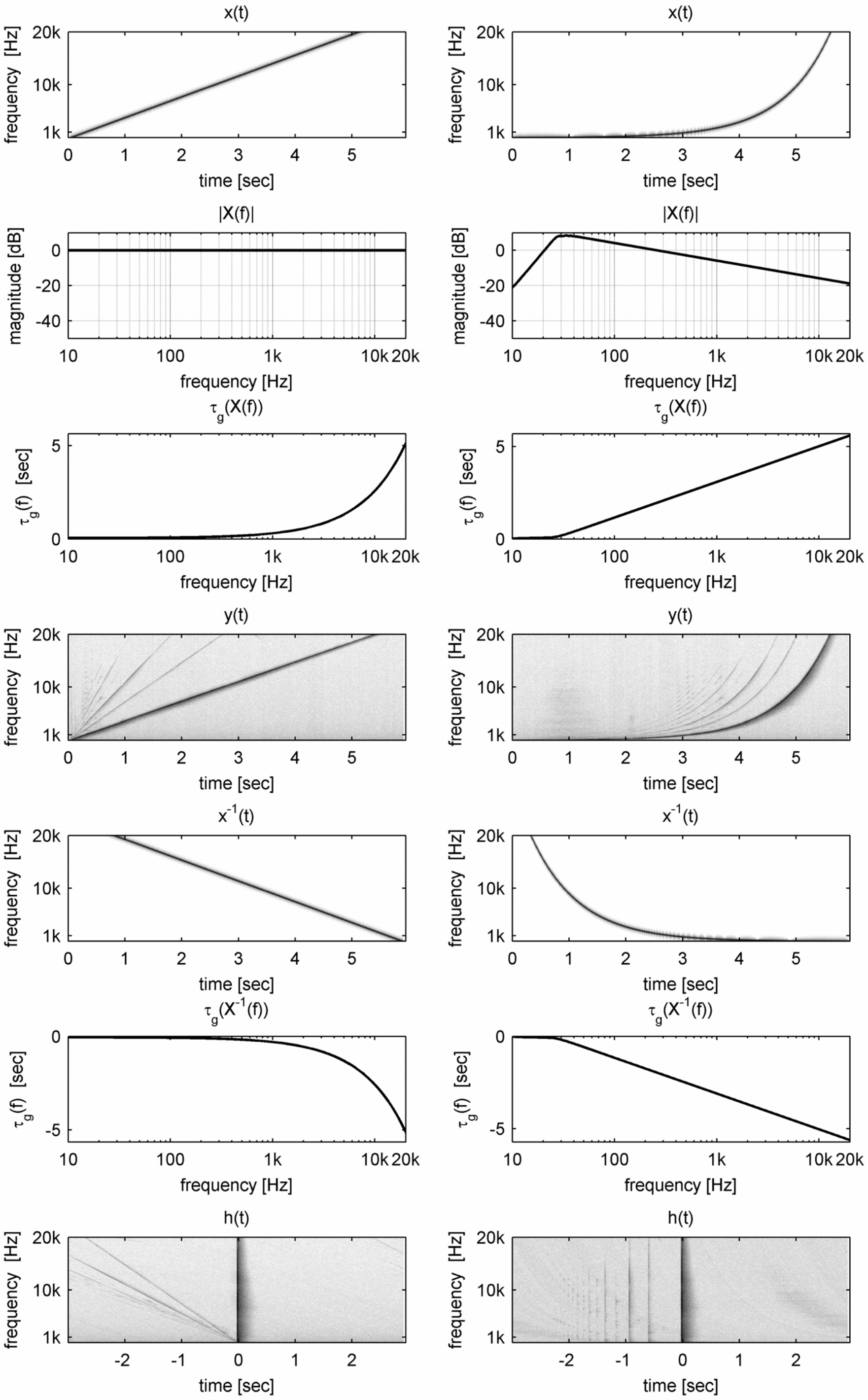

Figure 2: FFT-based measurement with linear (left) and logarithmic, high-pass filtered (right) sweeps. For a typical measurement with moderate distortion the spectrogram, the magnitude spectrum *X(f)*, and the group delay *τ(f)* are shown for the excitation signal *x(t)*, the system response *y(t)*, the inverse signal $x^{-1}(t)$, and the resulting impulse response *h(t)*.

$$D_{LinS} = \frac{\int_{\omega_1}^{\omega_2} \left| X(e^{j\omega}) \right|^2 \mathrm{d}\omega}{\omega_2 - \omega_1} \tag{9}$$

with start and stop frequencies ω_1 and ω_2.

Combining (3), (8), and (9) yields a group delay of

$$\tau_g(\omega) = \tau_g(\omega_1) + \frac{T}{\omega_2 - \omega_1}(\omega - \omega_1) \tag{10}$$

with $\omega \in [\omega_1; \omega_2]$.

Filtering the system response with the inverse signal corresponds to a convolution with the time-reversed sweep $X^{-1}(e^{j\omega})$, i.e. a sweep with decreasing frequency in time (deconvolution). Whereas the linear system response is folded back to $t = 0$, the products of harmonic distortion are folded back to negative times.

The resulting group delay of the *k*th harmonic after deconvolution according to (2) will then be

$$\tau_k(\omega) = \omega \cdot \frac{T}{\omega_2 - \omega_1}\left(\frac{1}{k} - 1\right) \tag{11}$$

yielding a linear down-sweep in time. The slope of the group delay increases with increasing order k while the instantaneous frequency decreases with increasing k by $k/(1-k)$ (see fig. 3, top). The resulting harmonic down-sweeps will extend into negative times from a minimum of

$$\tau_{k\min} = -\frac{\omega_2}{\omega_2 - \omega_1} \cdot T \cdot \left(1 - \frac{1}{k}\right) \tag{12}$$

reached for ω_2 to a maximum of

$$\tau_{k\max} = -\frac{\omega_1}{\omega_2 - \omega_1} \cdot \frac{T}{2} \tag{13}$$

reached for the start frequency ω_1 and $k = 2$. For high values of k and broadband measurements with $\omega_2/\omega_1 > 100$, the down-sweeps will extend into negative times for a time segment almost equal to the duration of the original sweep. In practice they can be masked by the noise floor before, thus $\tau_{k\min}(\omega)$ will have to be assessed by a reference measurement.

2.2. Logarithmic sweeps

Logarithmic sweeps (figure 2, right column) yield an exponentially increasing instantaneous frequency. They are usually called *log sweeps*, because their instantaneous frequency increases linearly on a logarithmic frequency scale corresponding to a logarithmically increasing group delay.

They exhibit a pink magnitude spectrum decreasing with 3 dB per octave and a power density spectrum given by

$$\left| X(e^{i\omega}) \right|^2 = D_{LogS} \cdot \frac{1}{\omega} \tag{14}$$

and normalized to the total sweep energy by

$$D_{LogS} = \frac{\int_{\omega_1}^{\omega_2} \left| X(e^{j\omega}) \right|^2 \mathrm{d}\omega}{\int_{\omega_1}^{\omega_2} \frac{1}{\omega} \mathrm{d}\omega} \tag{15}$$

Combining (3), (14), and (15) yields a group delay of

$$\tau_g(\omega) = \tau_g(\omega_1) + \frac{T}{\ln(\omega_2/\omega_1)}(\ln\omega - \ln\omega_1) \tag{16}$$

The resulting group delay of the *k*th harmonic after deconvolution according to (2) will then be

$$\tau_k(\omega) = -\frac{T}{\ln(\omega_2/\omega_1)}\ln k \tag{17}$$

As already shown by Farina [8], log sweeps exhibit the particular property that after deconvolution with the inverse filter, i.e. an exponential sweep-down, not only the linear part of the system response but also the nonlinear harmonic distortion products appear with a constant group delay (independent of ω) before the linear impulse response. Thus, the deconvolution not only yields a linear impulse response, but a series of harmonic impulse responses (HIRs) for each harmonic order k (figure 3, second from top). They appear at negative times before the linear impulse response and can be extracted by windowing in order to analyze the power and the spectral distribution of each harmonic component.

For increasing k the HIRs will be aligned with decreasing distance, depending on k, the sweep duration T and the ratio of stop and start frequencies ω_2/ω_1. In practice, the magnitude of distortion typically decreases with increasing k and will be masked by the noise floor above a certain value $k_{\max}$ that can be assessed by a reference measurement. Thus, distortion in the deconvolved system response is extending from

$$\tau_{k\,\min} = -\frac{\ln k_{\max}}{\ln(\omega_2/\omega_1)} \cdot T \tag{18}$$

to

$$\tau_{k\,\max} = -\frac{\ln 2}{\ln(\omega_2/\omega_1)} \cdot T \tag{19}$$

2.3. Red-coloured and arbitrarily shaped sweeps

A sweep with red-coloured spectral shape has a magnitude spectrum decreasing with 6 dB per octave and a power density spectrum given by

$$\left|X(e^{i\omega})\right|^2 = D_{\mathrm{RedS}} \cdot \frac{1}{\omega^2} \tag{20}$$

and normalized to the total sweep energy by

$$D_{\mathrm{RedS}} = \frac{\int_{\omega_1}^{\omega_2} \left|X(e^{j\omega})\right|^2 \mathrm{d}\omega}{\int_{\omega_1}^{\omega_2} \frac{1}{\omega^2} \mathrm{d}\omega} \tag{21}$$

Combining (3), (20), and (21) yields a group delay of

$$\tau_g(\omega) = \tau_g(\omega_1) + \frac{T}{(\frac{1}{\omega_1} - \frac{1}{\omega_2})}(\frac{1}{\omega_1} - \frac{1}{\omega}) \tag{22}$$

The resulting group delay of the kth harmonic after deconvolution is then given by

$$\tau_k(\omega) = \frac{T}{(\frac{1}{\omega_1} - \frac{1}{\omega_2})} \cdot \frac{1}{\omega} \cdot (1-k) \tag{23}$$

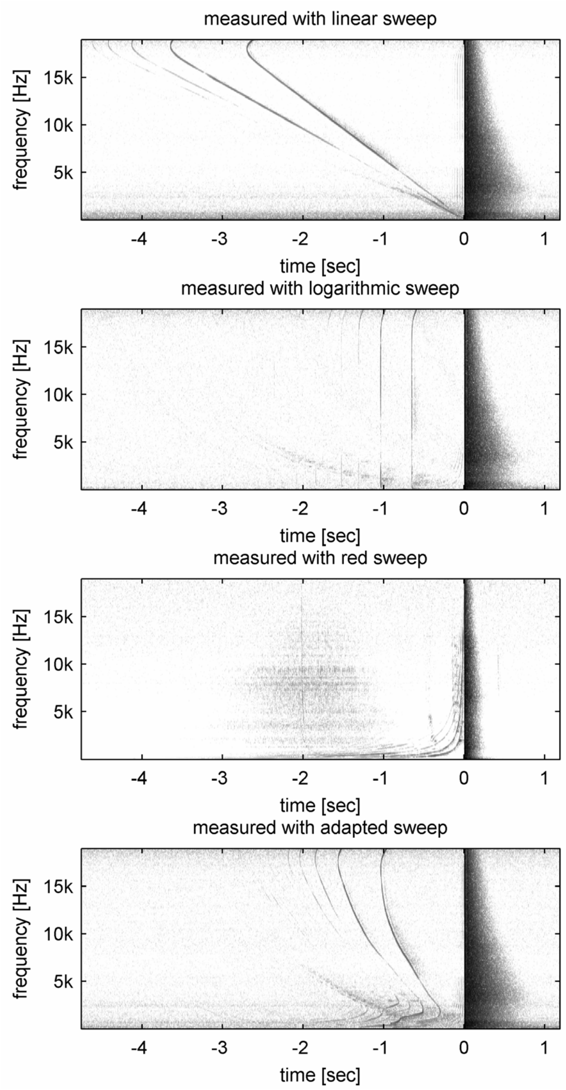

Figure 3: FFT-based measurements with linear sweep (white), logarithmic sweep (pink) and "red sweep", i.e. synthesized with a pink spectral shape. The last plot shows a measurement with a sweep adapted to the spectrum of the noise floor

For different k, (23) yields a series of rising sweeps with increasing instantaneous frequency in time, as can be seen in figure 3 (third from top). Distortion in the deconvolved system response is extending from a minimum group delay of

$$\tau_{k\,\min} = -\frac{(k-1)}{(1-\frac{\omega_1}{\omega_2})} \cdot T\,, \qquad (24)$$

reached for low frequencies, to

$$\tau_{k\,\max} = -\frac{T}{(\frac{\omega_2}{\omega_1}-1)} \qquad (25)$$

reached for high frequencies and $k = 2$.

For arbitrarily coloured sweeps the instantaneous frequency of the deconvolved distortion products in time can take any shape depending on $|H(e^{j\omega})|$. It can be calculated by applying (3) to (2). Figure 3 (bottom) shows the shape of distortion products for a sweep spectrally adapted to the underlying background noise (see fig. 6). Its spectral envelope is between red and pink for 100 - 2000 Hz and between pink and white above 2000 Hz. The same characteristics can be recognized in the temporal behaviour of the distortion products in the deconvolved impulse response.

3. MULTIPLE EXPONENTIAL SWEEP MEASUREMENT

Majdak et al. [7] have proposed a method to optimize impulse response measurements for system configurations with multiple sources by interleaving and overlapping subsequent logarithmic sweeps (called exponential sweeps by the authors). This is possible by starting the excitation at well-defined times, so that after deconvolution the linear impulse responses of a group of subsequent sweeps are located between the linear and the first harmonic impulse response of the last sweep of the group (interleaving). Moreover, by starting the first sweep of subsequent groups before the last sweep of the preceding group has reached its stop frequency, the procedure can be accelerated even more (overlapping, see fig. 4).

The efficiency of the interleaving strategy depends on the duration of the sweep T and the reverberation times L_1 and L_2 of linear and first harmonic impulse response. Only if the sweep is long enough and L_1 and L_2 are short enough, multiple sweeps will fit into the gap between the devonvolved linear and the first harmonic impulse response, given by (19) (see Fig. 3, second from top).

The efficiency of the overlapping strategy depends on the maximum order of distortion $k_{\max}$ visible, i.e. above the general noise floor of the measurement. Only if $k_{\max}$ is low enough subsequent sweeps can be pushed together and the linear impulse response and the distortion of the next sweep will still be separated by windowing.

A quantitative analysis of the two strategies reveals that the duration of the sweep has to be changed from its orginal, non-interleaved duration T to

$$T' = \left[(\eta-1)L_1 + L_2\right]\frac{c}{\ln 2} \qquad (26)$$

so that a group of η - *1* excitation sweeps with their reverberation time L_1 can be interleaved and inserted between the impulse response of the last sweep and the closest distortion product for τ_2 according to (19) and using T' for T, with $c = \ln(\omega_2/\omega_1)$. For small values of η, T' can get shorter than T yielding a loss of SNR. If that is to be avoided, T' has to be set to T in that case.

The measurement duration for η interleaved systems is thus

$$T_{INT} = T' + \eta L_1 \qquad (27)$$

If sweeps are overlapped, they have to be separated by a minimum time segment given by

$$\Delta t_{OV} = \tau_{k_{\max}} + L_1 \qquad (28)$$

with $k_{\max}$ being the maximum number of harmonics visible in the deconvolved impulse response and $\tau_{k\max}$ the corresponding time location according to (19).

If interleaving and overlapping is combined, one has to keep in mind that beginning of the kth harmonic impulse response will be shifted to

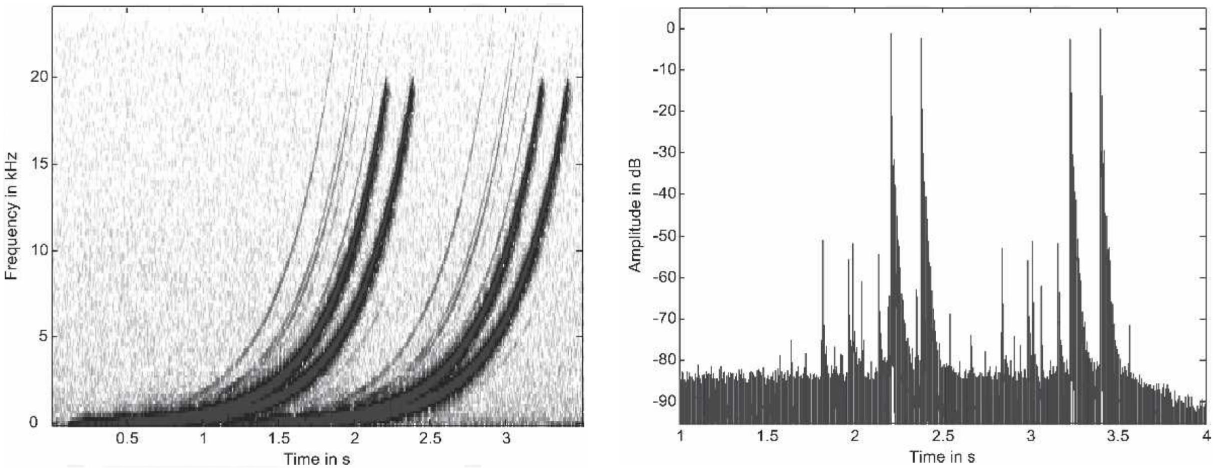

Figure 4: Multiple sweep measurement, compressing subsequent sweeps in time by interleaving and overlapping (from [7], figures 9 and 10). Left: System response after excitation with two overlapped groups of twofold interleaved sweeps. Right: Series of linear and harmonic impulse responses after deconvolution.

$$\tau_k{}' = \frac{T'}{c} \ln k \tag{29}$$

according to the extension of the sweep necessary for interleaving.

The total measurement duration for N sources is then given by the interleaving delay for $\eta = N$ sources according to (27) plus an additional delay for the overlapped groups of τ'_{kmax}, since the reverberation time L_1 in (28) is already accounted for in (27). The resulting total measurement time will then be

$$T_{MESM} = T' + \tau_k{}' \left\lceil \frac{N}{\eta} - 1 \right\rceil + NL_1 , \tag{30}$$

where $\lceil x \rceil$ returns the next higher integer of x.[1]

It can be compared to a conventional, step-by-step measurement with a total duration of

$$T_{con} = (T + L_1)N \tag{31}$$

The number of interleaved sweeps η can be optimized two reach a maximal reduction of measurement time with the same SNR per measurement or a maximal increase of SNR within the the same total measurement time for all systems.

For the optimization towards shortest measurement duration, one can see by inserting (26) and (29) into (30) that the measurement time T_{MESM} increases with increasing η. Hence, the optimal number of interleaved sweeps $\eta_{opt,T}$ will be given be the smallest η that still keeps $T' \geq T$, because otherwise a loss of SNR will occur. Thus the optimal η will be given by (26), using T for T', solved for η and rounded to the next higher integer:[2]

$$\eta_{opt,T} = \left\lceil \frac{T \ln 2}{cL_1} + \frac{L_1 - L_2}{L_1} \right\rceil \tag{32}$$

When calculating T_{MESM} using $\eta_{opt,T}$, i.e. a shortest measurement duration without loosing SNR, one has to keep in mind, that for small values of η a non-

[1] Formula (13) of Majdak et al. [7], giving (N/η - 1), is correct if N/η is an integer, i.e. if N can be partitioned in groups of equal size. Otherwise, (N/η -1) has to be rounded to the next higher integer.

[2] Formula (15) of Majdak et al [7], giving ($L_1 + L_2$), has been corrected for (L_1 - L_2)

interleaved measurement with $\eta = 1$ and a sweep duration T according to (30) might yield a shorter measurement durations T_{MESM} than values from (32), (26) and (29) inserted in (30). The resulting efficiency with respect to an optimal measurement duration can be expressed as a time compression factor

$$v_{tot} = \frac{T_{con}}{T_{MESM}} \tag{33}$$

with values shown in figure 5.

When optimizing the MESM procedure towards an optimal SNR without increasing the total measurement duration, one has to find the highest η so that $T_{\mathrm{MESM}} \leq T_{\mathrm{con}}$, according to (30) and (31), since T' increases with η. The SNR gain due to the application of multiple sweeps with extended duration T' is then given by

$$g = 10\log\frac{T'}{T} \tag{34}$$

with T' according to (26).

For the MESM procedure, both v_{tot} and g primarily depend on the number of systems N, but also on the measurement conditions, i.e. on the parameters L_1, L_2, $k_{\max}$, and T, as illustrated in figure 5 for different values N and L_1 and fixed values for the other parameters, set to typical values of $L_2 = 0.2 \cdot L_1$, $T = 3$ s, $k_{\max} = 5$, and $\omega_2/\omega_1 = 400$.

It can be seen that the efficiency of the MESM procedure strongly depends on N and L_1. Only for short reverberation times enough sweeps can be interleaved (for L_1 = 0.1 s <u>all</u> N sweeps are interleaved in fig. 5, bottom), in order to increase the duration of the single sweep as well as the resulting SNR gain. To accelerate the total measurement duration, lower values of η are favourable (for L_1 = 0.1 s no more than 5 sweeps are interleaved in fig. 5, top), but also the efficiency of overlapping procedure benefits from shorter reverberation times. The erratic behaviour of the curves for certain parameter constellations is due to the fact that the efficiency of MESM is always maximal when the total number of sweeps can be partitioned in groups of equal size for interleaving.

Majdak et al. have considered only exponential (logarithmic) sweeps, which are popular due to their emphasis of low frequencies similar to the typical noise floor of acoustical environments, and where the effect of deconvolution on harmonic distortion products is easy to oversee (see sec. 2.2). However, while log sweeps are favourable in the low frequency region, their 1/ω - decrease of spectral energy in the high-frequency region is usually not very well adapted to characteristics of typical acoustical noise floors.

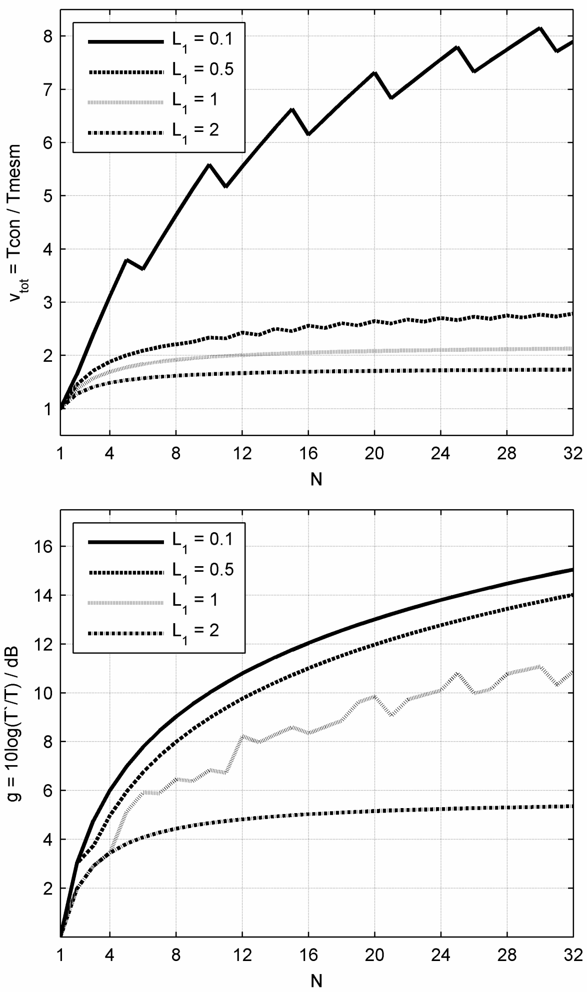

Figure 5: Measurement acceleration v_{tot} (top) and SNR gain g (bottom) attainable by the MESM procedure. Values are shown depending on the number of systems N, different values of L_1 (reverberation time of the linear impulse response) and fixed values of $L_2 = 0.1L_1$, $k_{\max} = 5$, $\omega_2/\omega_1 = 400$. v_{tot} and g are relative to a conventional measurement without interleaving and overlapping and sweep duration $T = 3$ s.

That is why linear sweeps with a low-frequency shelving gain or, ideally, sweeps spectrally adapted to the measured noise floor, have found to provide a higher signal-to-noise ratio than logarithmic sweeps.

4. MEASUREMENT WITH MULTIPLE NON-EXPONENTIAL SWEEPS

The benefit of multiple sweeps results from interleaving and overlapping subsequent excitation signals. So these two procedures have to considered separately.

First, it is easy to see that interleaving will only work effectively for exponential sweeps. For non-exponential sweeps it depends on the ratio of start and stop frequencies ω_2/ω_1 and on the duration T of the original sweep, how large the time gap is between the linear impulse response and harmonic distortion products in the deconvolved system response. For broadband measurements with $\omega_2/\omega_1 \approx 500$, τ_{max} will be about 1 ‰ of the original sweep duration T for the linear sweep and 2 ‰ for the red sweep according to (13) and (25). τ_{max} can be determined for arbitrarily shaped sweeps by applying their group delay to (2) and (3), see fig. 3 (bottom) for the adapted sweep. Since it did not turn out to be effective for the types of sweeps considered, the interleaving strategy was not considered further.

However, the efficiency of the overlapping strategy of subsequent sweeps does not depend on the group delay and thus on the spectral shape of the applied sweep. It will only depend on the time segment, by which nonlinearities extend into negative times before the start of the linear impulse in the deconvolved system response. In the case of exponential sweeps, this time segment is easy to detect in the energy time curve, because the harmonics are realigned as harmonic impulse responses with a clear onset time given by (17) through the deconvolution process. For non-exponential sweeps, however, the distortion products spread out in time, superposed among different orders and gently approaching the noise floor towards the left border. Much clearer than in the energy time curve they can be recognized in the spectrogram where they are visible with their instantaneous frequency as rising or declining sinusoids.

To find a criterion according to which the presence of distortion is acceptable in the deconvolved system response, it might be instructive to consider the consequences if distortion is not sufficiently removed by windowing because it overlaps into the preceding linear impulse response. If the measured impulse responses will be used for convolution with anechoic sound material, as it is usually the case with spatial or head-related impulse responses, the result will depend on the source signal to be convolved with. When triggered with noisy audio content the convolution with a sweep will assume noise characteristics, too. The worst case will occur for transient audio signals triggering exactly the rising or declining sweep that can be seen in the spectrogram. Hence, in order to determine whether the sweep will be masked by the underlying noise floor, the latter should be evaluated with a narrowband, auditory filter to make sure that a simultaneous sinusoid of equal power will be masked. This should be done for all filterbands, because according to the complex behaviour of distortion products of different order in time and different energy it cannot be easily predicted which frequency will have furthest extension into negative times.

The maximum extension $\Delta\tau_{max}$ resulting from an inspection of all filterbands can then be used to calculate the minimum delay Δt_{OV} necessary to prevent distortion from overlapping into the linear impulse response with extension L_1 of the preceding measurement. Like in the case of exponential sweeps it is given by

$$\Delta t_{OV} = \Delta\tau_{max} + L_1 \tag{35}$$

resulting in a total measurement duration T_{OV} for N systems, where $N - 1$ sweeps of length T are overlapped:[3]

$$T_{OV} = T + (N-1)\cdot\Delta t_{OV} + L_1 \tag{36}$$

The total measurement duration T_{OV} will be shorter than for a conventional measurement T_{con} without overlapping according to (28). The reduction of measurement duration can again be (mathematically as well as practically) expressed as an acceleration factor v_{tot} or a gain in SNR, assuming that the duration T of a single sweep is extended by a factor v_{sw} so that the overlapped measurement will have the same measurement duration as the conventional measurement.

[3] Formula (11) in Majdak et al. [7] was augmented with a missing term "+ L_1"

v_{tot} is given as the ratio of non-overlapped, conventional measurement duration according to (31) and overlapped measurement duration according to (36). It depends on L_1 and on the maximum extension $\Delta\tau_{max}$ of distortion, which was set to $r = 35\%$ of the sweep duration for the calculation in figure 6 (top).

In order to calculate a gain in SNR due to the extension v_{sw} of a single sweep, it has to be considered that not only T but also $\Delta\tau_{max}$ will increase with the factor v_{sw}. If the total measurement duration is equal to the conventional measurement duration according to (31), we get

$$(T + L_1)N = v_{sw}T + L_1 + (N-1)(v_{sw}\Delta\tau_{max} + L_1)$$

yielding

$$v_{sw} = \frac{NT}{T + (N-1)\Delta\tau_{max}} \tag{37}$$

Since $\Delta\tau_{max}$ is always proportional to the sweep duration T according to (12), (18), and (24), it can be expressed as

$$\Delta\tau_{max} = r \cdot T \tag{38}$$

yielding

$$v_{sw} = \frac{N}{1+(N-1)r} \tag{39}$$

and corresponding to a SNR gain of

$$g = 10\log v_{sw} \tag{40}$$

The gain in SNR through the overlapping of subsequent sweeps thus only depends on the number of systems N and the maximum extension of distortion into negative times $\Delta\tau_{max}$ relative to the total sweep duration T. For $N \rightarrow \infty$, v approaches $1/r$ asymptotically, as can be seen in figure 6, showing the gain in SNR over N for four different values of r.

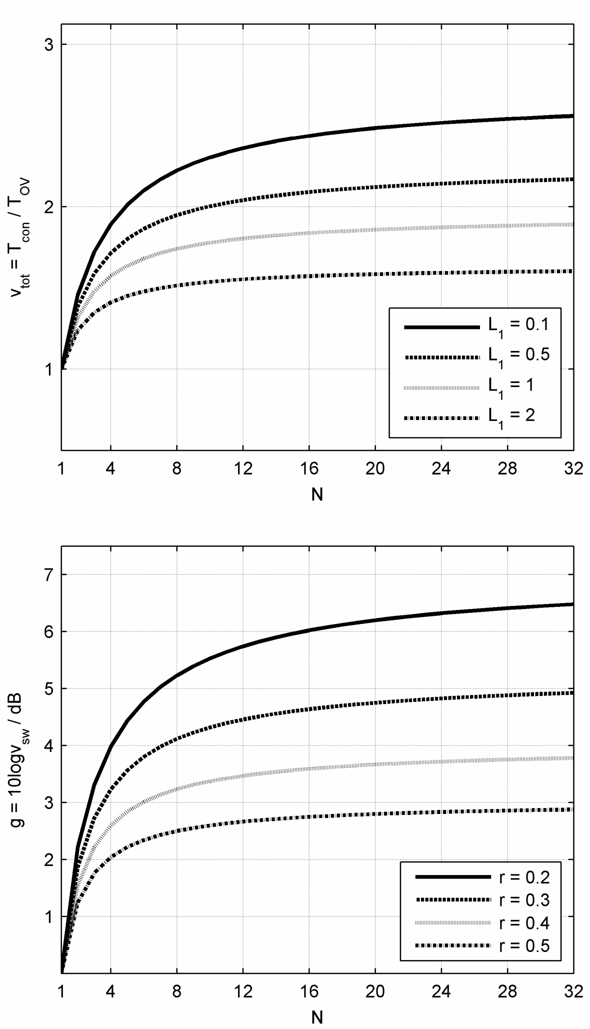

Figure 6: Measurement acceleration v_{tot} (top) and SNR gain reached by overlapping and extension of subsequent sweeps yielding the same total measurement duration (bottom) relative to the non-overlapped measurement. Values for v_{tot} depend on N, L_1 and r (set to 0.35), whereas values for g depend on N and r only.

Both the measurement acceleration and the SNR gain through overlapping alone will be lower than for overlapping and interleaving, as it results from the original MESM procedure, but it will add to the SNR gain ΔL resulting due to the spectral shape of the optimized sweep compared to the pink spectrum of an exponential sweep. The amount of SNR gain through spectral shaping will depend on the spectral noise power and can only be determined for a particular environment by a reference measurement.

5. REFERENCE MEASUREMENT

The application of spectrally shaped sweeps will have two consequences for a measurement with multiple, overlapped excitation signals: It will determine the extension τ_{min} of distortion products into negative times due to the relative group delay of harmonics versus fundamental frequencies in the sweep signal, and it will yield a defined SNR gain compared to a non-adapted excitation spectrum.

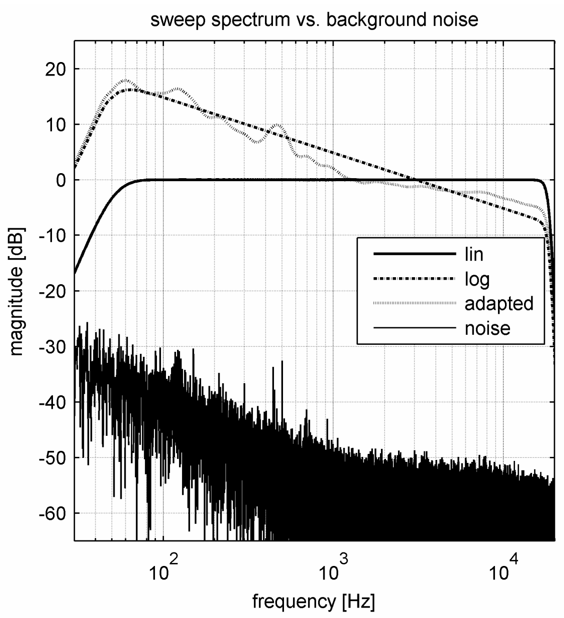

Figure 7: Magnitude spectra for three sweeps with constant temporal envelope and different group delay. Sweeps were high-pass filtered below 50 Hz to avoid loudspeaker distortion and normalized to the same energy.

Both parameters can be determined in a reference measurement, as it was done by the authors in a typical (quiet, but not sound-insulated) laboratory environment, using a loudspeaker with moderate distortion. Measurements were performed using a lin sweep, a log sweep (exponential sweep) and a sweep spectrally adapted to the underlying noise floor. Figure 7 shows the magnitude spectrum for the three types of sweeps along with the spectral power of the noise floor.

The adaption of sweep spectrum to the noise floor was performed by 1/3 octave smoothing of $|N(e^{j\omega})|$ and transfering the result to a sweep synthesis algorithm according to (6). Alternatively, a spectral envelope was formed based on an amplitude peak detection of $|N(e^{j\omega})|$ within semitone ($2^{1/12}$) intervals and subsequent 1/3 octave smoothing of the peaks, in order to have a more precise representation of the spectral noise envelope including potential narrowband noise due to hum or ventilation noise (fig. 8). The result was again resynthesized as magnitude spectrum of a spectrally shaped sweep.

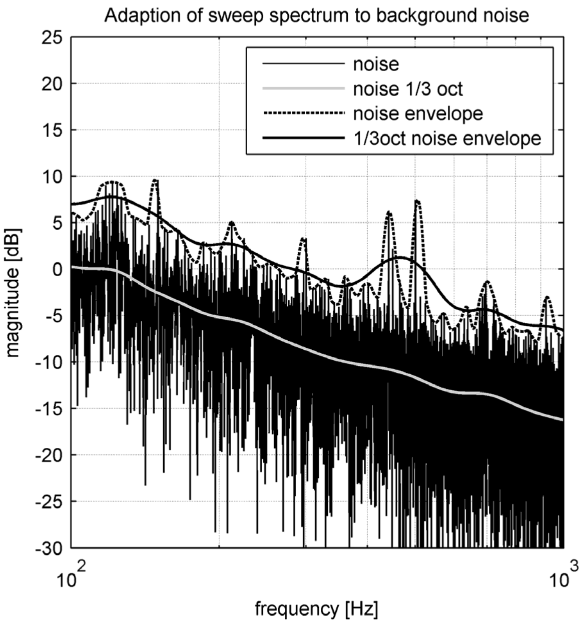

Figure 8: Two approaches to adapt the sweep spectrum to the spectral distribution of the noise floor. First: 1/3 octave band smoothing ("noise 1/3 oct") of the original noise magnitude $|N(e^{j\omega})|$ ("noise"). Second: Spectral envelope based on an amplitude peak detection of $|N(e^{j\omega})|$ within semitone intervals ("noise envelope") and subsequent 1/3 octave smoothing of the peaks ("1/3 oct noise envelope")

After impulse response measurement with these four types of sweeps, the signal-to-noise ratio was determined by measuring noise power in the deconvolved system response within a time interval unaffected by reverberation or distortion (at the left border of fig. 3) relative to the (normalized) signal power of the sweep. In table 1, the measured SNR values of around 75 dB are given relative to the SNR obtained for the linear sweep.

	SNR / dB_{rel}
lin sweep	0 dB
log sweep	3.4 dB
1/3 oct adapted sweep	6.8 dB
envelope adapted sweep	9.0 dB

Table 1: SNR obtained for an impulse response measurement with different types of sweeps, relative to the SNR obtained for a linear (white) sweep

An important parameter for the efficiency of a multiple sweep measurement is the extension of distortion products into negative times in the devonvolved system response. Whereas for the log sweep, four upper harmonics (k = 2 ... k = 5) are clearly visible in the spectrogram (fig. 3) as well as in the broadband energy time curve (fig. 9), the distortion products of the adapted sweep appear smeared in time in the energy time representation. In the spectrogram, at least 5 upper harmonics are visible, extending for about 2 s into negative times relative to the start of the linear impulse response. The same extension is yielded by the the broadband energy time curve (fig. 9, second from top). Analysis with an auditory filterbank (24 bands gammatone filterbank, implemented by [9] according to [10]), however, indicates distortion artefacts up to 2.7 s into negative times. Only with very close inspection of fig. 3 they can be seen as a weak shadow at around 13 kHz, corresponding to the peaks visible in the 12-14 kHz auditory band.

Hence, analysis of the reference measurement yields an extension of distortion products of $\tau_{kmax} = 1.6$ s for the log sweep (corresponding to $k_{max} = 5$) and $\tau_{max} = 2.7$ s according to a very strict criterion and $\tau_{max} = 2.0$ s according to broadband analysis for the adapted sweep, corresponding to a ratio of $r = 34$ % resp. $r = 46$ % relative to the sweep duration T of 5.9 s for this measurement. The broadband reverberation time L_1 (which is not the classical reverberation time RT60, because the whole reverberation tail is taken into account, not only a decrease of 60 dB) was about 1.2 s according to the RMS values shown in fig. 8. If we apply these values to the evaluations shown in figures 5 and 6 and the values from table 1, we can conclude that under these measurement conditions and for N = 8 systems the use of overlapped and spectrally adapted sweeps will yield a higher SNR gain of 8.5 dB compared to 6.5 dB for the use of logarithmic, interleaved sweeps. Only for N > 18 the use of interleaved sweeps will outperform the measurement with spectrally adapted sweeps.

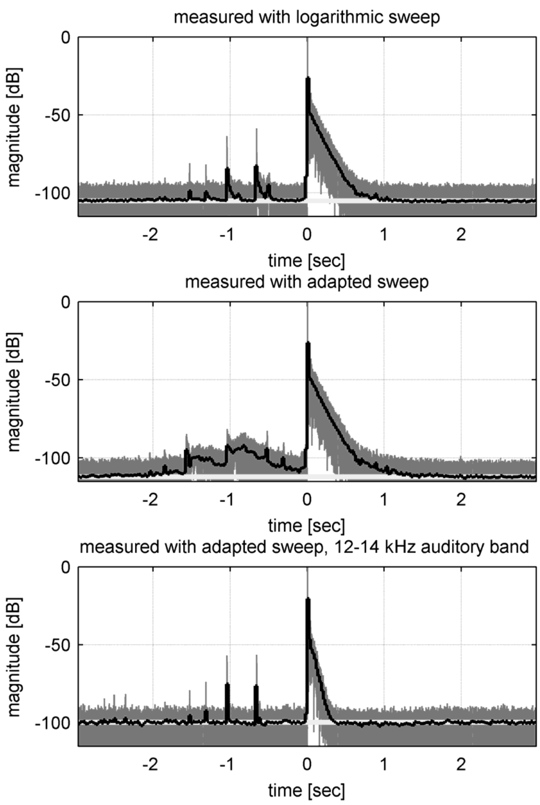

Figure 9: Analysis of the reference measurement with log sweep and spectrally adapted sweep (see Fig. 3, second from top and bottom) showing magnitude (grey) and RMS value determined in a window of 2^{10} samples (top). The adapted sweep measurement has additionally been analysed with a 24-band gammatone filterbank. The result of band No. 18 (12-14 kHz), revealing the furthest extension of harmonic distortion into negative times, is shown (bottom)

6. DISCUSSION

An analysis of the efficiency of multiple sweep measurements has shown that both a procedure with overlapped and interleaved logarithmic sweeps and a procedure with overlapped and spectrally adapted sweeps can yield a considerable benefit in terms of reduced total measurement duration or increased signal-to-noise ratio. If we assume the degree of distortion due to non-linearities of the loudspeaker and determined in a reference measurement as typical ($k_{max} = 5$ for the log sweep, $r = 0.5$ for a spectrally adapted sweep, yielding a gain in SNR of 5-6 dB compared to the log sweep), the trade-off will primarily depend on the number of sound sources to be used and the reverberation time of the acoustical environment. For a short reverberation time of $L_1 \approx 0.1$ s the original MESM approach will outperform the use of overlapped and spectrally adapted sweeps already for $N \geq 7$, while for a longer reverberation time of $L_1 \approx 2$ s, the latter will always outperform log sweeps, because efficient interleaving is no longer possible.

7. ACKNOWLEDGMENT

Alexander Lindau has been supported by a grant from the Deutsche Telekom Laboratories.

8. REFERENCES

[1] Müller, S. (2008): Messtechnik. In Weinzierl, S. (ed.): *Handbuch der Audiotechnik*. Springer Verlag, Berlin.

[2] Lindau, A., Hohn, T., Weinzierl, S. (2007): Binaural resynthesis for comparative studies of acoustical environments. AES 122nd Convention, Vienna, Preprint no. 7032

[3] Lindau, A., Maempel, H.-J., Weinzierl, S. (2008): Minimum BRIR grid resolution for dynamic binaural synthesis. In: Proc. of the Acoustics '08., Paris, pp. 3851-3856

[4] Behler, G., Lundschek, R. (2007): Dodekaeder-Lautsprecher mit variabler Richtcharakteristik. Fortschritte der Akustik, DAGA, Stuttgart

[5] Zotter, F., Sontacchi, A., Höldrich, R. (2007): Modeling a Spherical Loudspeaker System as Multipole Source. Fortschritte der Akustik, DAGA, Stuttgart

[6] Müller, S. and Massarani, P. (2001): Transfer-Function Measurement with Sweeps. J. Audio Eng. Soc., vol. 49, pp. 443-471

[7] Majdak, P., Balasz, P., Laback, B. (2007): Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions. J. Audio Eng. Soc., vol. 55, pp. 623-637

[8] Farina, A. (2000): Simultaneous measurement of impulse response and distortion with a swept-sine technique. AES 108th Convention, Paris, Preprint no. 5093

[9] Slaney, M. (1993): An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank. Apple Computer Technical Report #35, Apple Computer, Inc.

[10] Glasberg, B.R. and Moore, C.J. (1990): Derivation of auditory filter shapes from notched-noise data. Hearing research 40, pp. 103-138