Head Related Transfer Functions & Multiple Exponential Sweep Method

Markus Murschitz
(Institute of Electronic Music and Acoustics
KUG 2008-1-24
mamut@sbox.tugraz.at)

Graz, Winterterm 07/08

This work is about the Multiple exponential sweep method and its specific advantages at the measurement of Head Related Transfer Functions. It will explain both terms in a short way and will describe the way from the general problem of system identification to the Multiple exponential sweep method. The robustness of this method against weak non-linearities will be explained.

Therefor the signal properties of the exponential sweeps in conjunction of system modeling will be discussed.

Contents

1 Problem and Motivation

To state the problem it is necessary to give a quick overview of the Head Related Transfer Function and its measurement.

1.1 Head Related Transfer Function

A head-related transfer function (HRTF) describes the sound transmission from the free field to a place in the ear in terms of a linear time invariant system.[Majdak 1]
The according impulse response is called Head Related Impulse Response HRIR.

The basic procedure of measuring can be described as follows:

1. Placing speakers in various points of free space specified by radius r, azimuth angle Θ and elevation angle ϕ.
2. Playing back a known test signal for each of these positions
3. Recording the signal y[n] inside the ear of a test person, or an dummy head.
4. Computing the head related impulse response h[n].

If on the one hand a complete set of HRTFs is to be determined several measurements at several positions in space heve to be accomplished for each (r,Θ,ϕ). On the other hand if high quality measurements should be provided it is necessary to have a certain measurement duration as described later.
These two wishes all in all cause a long measurement duration (might be some hours),

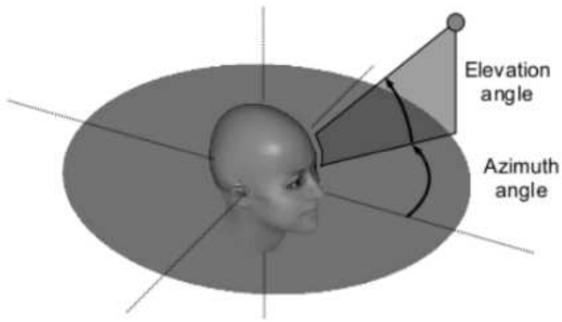

Figure 1: The geometrical properties for HRTF measurement

which is neither very efficient nor very comfortable if a human persons head is used.

The main aim of the Multiple exponential sweep method (MESM) is to reduce this times in a very efficient way.
An other important question is, how many speakers are used to perform this measurement. The answer was found by [Minnaar 1]. Due interpolation of the speakers signal responses it is possible to reduce the number of speakers to 1130. Which is still a quite high number.

Here a little geometric trick is used. There are only some speakers necessary, if they are arranged in an horizontal oriented arc around the test persons head. Then the test person is rotated.

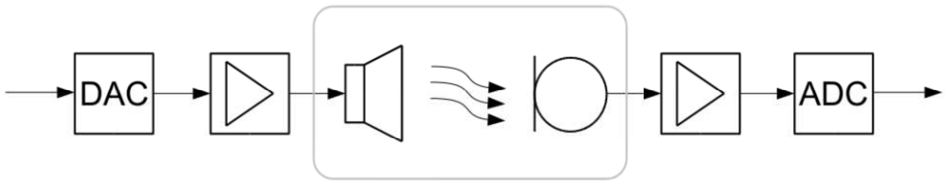

Figure 2: A general Problem description for a one channel system.

2 System Identification

All in all to get one HRTF is a problem of system identification. Signal processing assumes the figure 3 as the model for system identification.

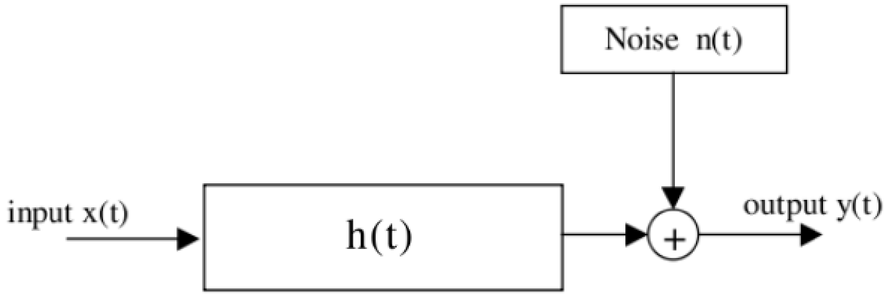

Figure 3: The general model for System identification.

The process to get the h[n] is called deconvolution.

$$H = \frac{\hat{Y}}{X} \tag{1}$$

The $\hat{Y}$ describes the $FFT\{\hat{y}\}$, where $\hat{y}$ is an averaged value of several distinct measurements of y as a result of x.
Which means, for uncorrelated noise and infinite measurements the noise will have no effect on $\hat{y}$. For repeated measurements we have to consider that repetition means that we have to deal with the risk of time aliasing. If the reverberation time of the room is longer than the period of repetition of the input signal x we will get time aliasing in our calculated h because of the overlapping of the consecutive response signals. Which might be a quite silly mistake. But it will be shown that this is the most importand limitation and that we can reach shorter measurement times if we use this knowledge.

$$h = IFFT\{H\} \tag{2}$$

It can also be described by a convolution with the inverse input signal x'. Which means that x fulfills $x' * x = \sigma$. Which explains the name deconvolution.

$$h = y * x' \text{ where } x' = IFFT\left\{\frac{X(-\omega)}{|X(\omega|)}\right\} \tag{3}$$

3 The Input Signal

It will be shown that for the deconvolution process it is important to chose a suitable input signal. Some standard input signals as the Maximum Length Sequences and linear sweeps will be described to show their weaknesses in comparison to the Exponential Sweeps. Therefore the empiric results of the measurements are presented without exact mathematical explanation. Some of this mathematical explanation might be found in 4.1.

3.1 Maximum Length Sequences

A Maximum Length Sequence (MLS) is a pseudo random binary sequence. Which can also be described as nearly white noise. Which means that the noise is uncorrelated and that its spectrum is so called white.
It is normaly generated by a K bit maximal linear feedback shift register. Which can produce an uncorrelated sequence of the length $L = 2^K$. If we assume the model shown in figure 3 we have to consider the measurement noise. So the first thing one has to worry about is that L is bigger than the reverberation time of the room which is the order N of the unknown system. Otherwise time aliasing might occur.
So far so good but if small non-linearities are considered in the system h one can run into deep troubles. The non-linearities are causing a distorted output signal y and therefor a distorted deconvolution result h as shown in 4.

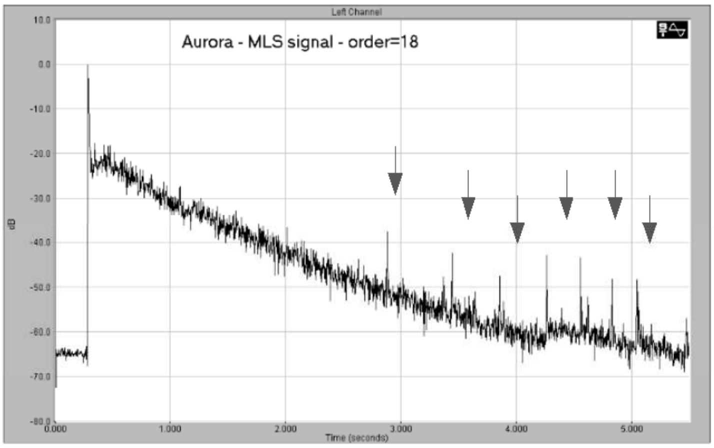

Figure 4: The effect of nonlinearity on the deconvolution result h. The arrows show the scaled down copies of the HRIR overlaid with the original one. The graphic shows a zoomed part of $h[n]$ in dB

3.2 Linear Sinusoidal Sweeps

The Linear Sinusoidal Sweeps or short linear sweeps are of the form shown in 4.

$$x(t) = sin(\phi(t)) \text{ where } \omega(t) = \frac{d\phi(t)}{dt} \text{ ... is linear} \tag{4}$$

The result for weak non-linearities is better than in the MLS case we are getting a $h[n]$ which is free of this scaled down copies of its own. But on the other hand we can see a raised noise floor in comparison to an undistorted measurement result.
All the afforts to get rid of this by averaging as before are not worth trying because this noise is inherent for this specific system signal combination. The idea one might have is to push all the noise together into a region of h which is not holding information. And the way to accomplish that is to change the input signal once more. It will be shown that the exponential sweeps explained in 4.1 are fulfilling this wish.

4 Exponential Sweeps

An exponential sweep is also a sinusoidal sweep the only difference is that its $\omega(t)$ is proportional to an exponential function. It was invented by Angelo Farina, described in [Farina 1]. As an appropriate ansatz we use a formula of the form:

$$x(t) = sin\left(\phi(t)\right) \text{ with } \phi(t) = K_1\left(e^{\frac{t}{K_2}} - 1\right)$$

If the start and the end frequency are ω_1 and ω_2 at time $t = 0$ and $t = T$ we can observe the missing parameters K_1 and K_2 by fulfilling:

$$\omega_1 = \left.\frac{d(\phi(t))}{dt}\right|_{t=0} \qquad\qquad \omega_2 = \left.\frac{d(\phi(t))}{dt}\right|_{t=T}$$

A exponential sweep is now defined by 5.

$$x(t) = sin(\phi(t)) = sin\left(\frac{\omega_1 T}{c}\left(e^{\frac{tc}{T}} - 1\right)\right) \qquad \text{where } c = ln\left(\frac{\omega_2}{\omega_1}\right) \tag{5}$$

4.1 Exponential Sweeps Properties

Some of the mathematical properties which are signal inherent, have to be discussed here.

4.1.1 Harmonics

Analyzing harmonic means in a mathematical way to search for frequencies which fulfill:

$$N\omega(t) = \omega(t) \qquad \text{which means } N\frac{d(\phi(t))}{dt} = \frac{d(\phi(t))}{dt}$$

Applying this formula on 5 leads to a really interesting property of the exponential sweeps.
We obtain:

$$\Delta t = T\frac{ln(N)}{c} = T\frac{ln(N)}{ln\left(\frac{\omega_1}{\omega_2}\right)} \tag{6}$$

Which means that there is always a constant time difference Δt between a harmonic and its fundamental frequency.

4.1.2 Periodogram

If a periodogram of the signal, which is an estimation of its frequency spectrum is computed we can obtain the pink spectrum of the exponential sweep. In order to calculate the periodogram an approximation of the signal has to be used as described in [Zotter 1]. Figure 5 shows the periodogram. One can see the Gibbs phenomenon in time which is caused by windowing the infinite sweep to get a specific sweep from ω_1 to ω_2.

5 A Non-linear System Model

If we are assuming a weakly non-linear System we will have to model it as well as an linear time invariant one. It is quite important to know about the limits caused by the assumed model. The model we are assuming is a Hammerstein Model as shown in figure 6.

It shows the combination of a nonlinear memoryless and a linear time invariant system. If we are examining the effect only of the nonlinear system which is memoryless, on the signal, we will find out, that it can be described by a sum over the Chebyshev Polynomials. These polynomials are described as:

$$y(t) = \sum_{n=1}^{N} = k_n * T_n(s(t))$$

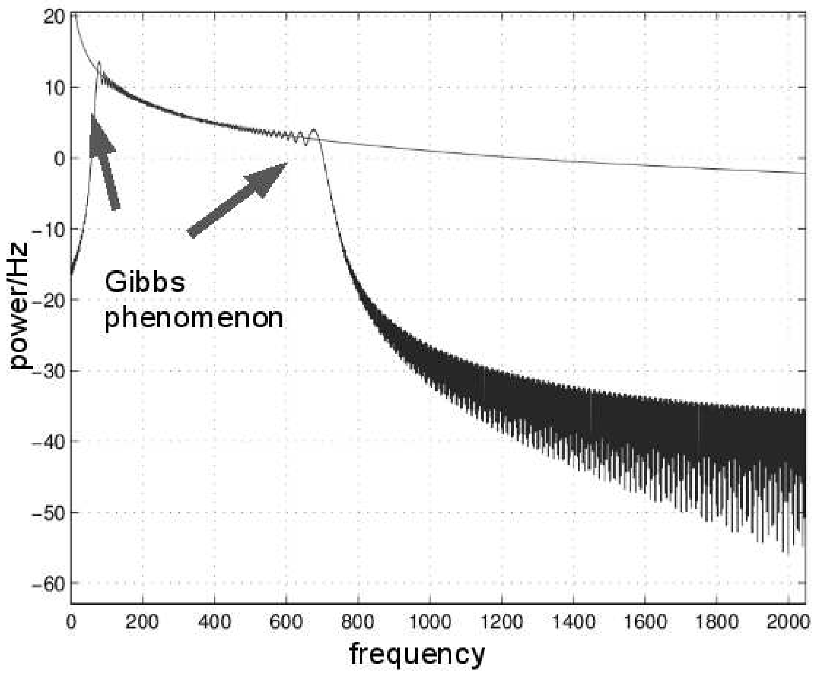

Figure 5: The Periodogram of an infinite and a windowed sweep.

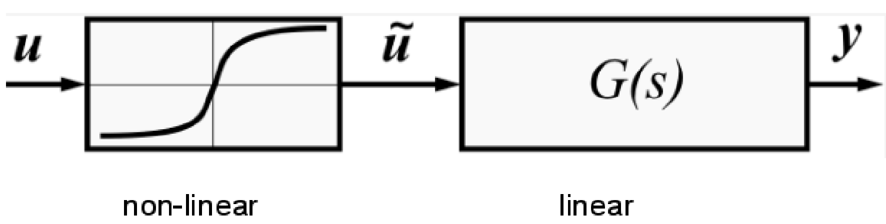

Figure 6: A Hammerstein Model is the combination of a memoryless system and a linear time invariant one.

where

$$T_n(s) = \begin{cases} cos(N * arccos(s)) & \text{if } s \in [-1, 1] \\ cosh(N * arccosh(s)) & \text{if } s \geq 1 \\ (-1)^n cosh(N * arccosh(-s)) & \text{if } s \leq -1 \end{cases}$$

If the first case is assumed and $s(t) = x(t) = cos(\phi(t))$ we get out that:

$$T_n(s) = cos(N * \phi(t))$$

which is exactly as we analyzed in 4.1.1.

6 Exponential Sweeps and Non-linear Systems

All in all we can see that the effect of the nonlinear system is only to produce scaled harmonics of the fundamental frequency (iff $x \in [-1, 1]$).
Furthermore we already know that these harmonics of Exponetial Sweeps are always causing a constant time shift Δt in respect to their fundamental frequency in y.

So if a weak nonlinear system which can be modeled as a Hammerstein System is excited with an exponential sweep the result is the systems response and time shifted versions of it for each harmonic (see fig. 7).

If $x[n]$ stays the same and $y[n]$ gets these additional components $h[n]$ will also contain such shifted components. So $h[n]$ which we get out of the deconvolution shows this repetition in time for each harmonic shown in figure 8. Here we got scaled down repeated versions of the linear impulse response. By windowing the rightmost part out of $h[n]$ we can separate the non-linear effects of a system from the linear ones.

In fact, the exponential sweep method allows to slightly overdrive systems that would otherwise behave linearly on purpose. The amount of overdrive has to be chosen properly, in order to allow for a separation of the nonlinear system responses. The motivation for this rather brave approach lies in the better SNR we get by increasing the playback volume of the test signal.

7 Multiple Exponential Sweep Method

The original problem of determination of several HRIRs as described in chapter 1.1 can be solved now. The straight forward way would be to measure all of the HRIRs, system

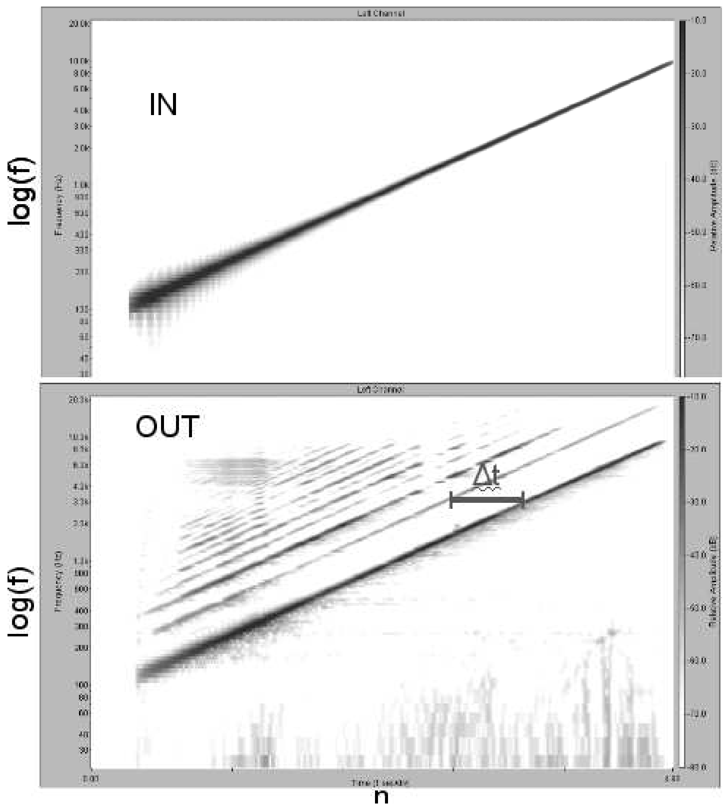

Figure 7: The response of a nonlinear system caused by excitation with an exponential sweep (in logarithmic scale)

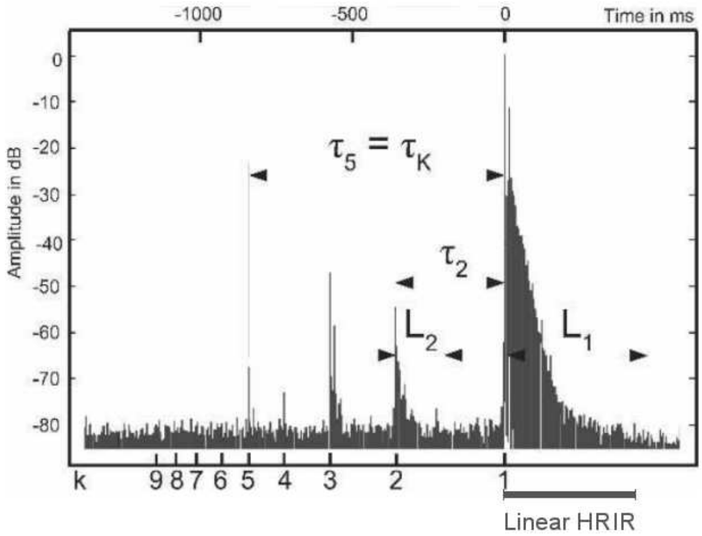

Figure 8: Head Related Impulse Response and the nonlinear scaled down repetitions of it (zoomed in and shown in dB)

by system. But if we consider all the theory behind the exponential sweeps, it is possible to do this measurement much more efficient.
Two strategies are explained in [Majdak 1] and [Farina 1]:

- Interleaving: Measurements of several systems overlapping in time.
- Overlapping: Multiple measurements of one system overlapping in time.

7.1 Interleaving

Interleaving describes the idea to measure one system while the previous systems measurement is not finished yet. This will result in a h where the HIRs of the two systems are interleaved in time. Therefor it is necessary to place the HIR of the second system between the linear and the second order non-linear part of the first system (see figure 9).

As you can see in this image it will not fit in there. The parameter which has to be changed in order to get a bigger gap is the sweep duration. A quick look at 6 shows that changing the total duration will cause a longer gaps between all orders and their fundamental frequency.
The result can be seen in figure 10.

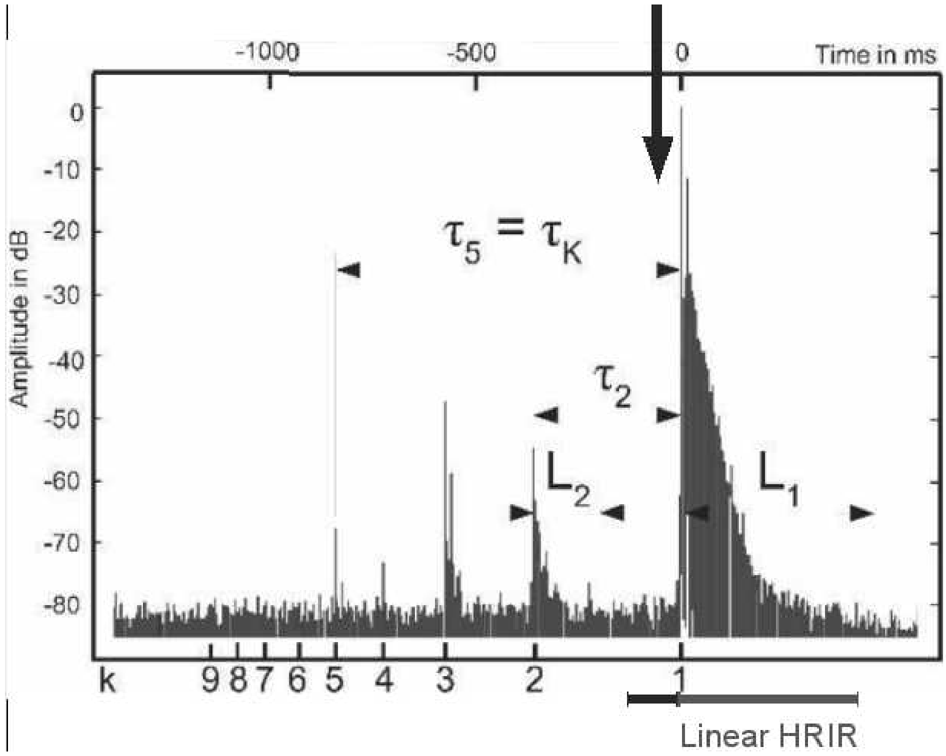

Figure 9: The arrow sows where to place the second systems impulse response.

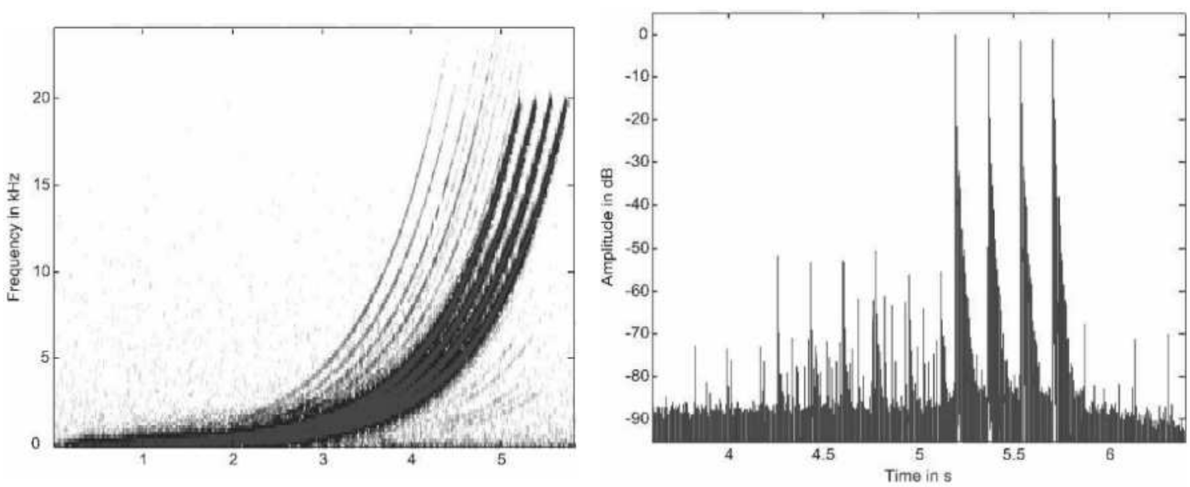

Figure 10: left: Spectrogram of the interleaving method y; right: according $h[n]$ in dB

A very positive side effect is that using a longer sweep will also increase the SNR. Each duplication of the sweep duration means +3dB for the SNR, which might reduce the number of repetitions in order to get the same SNR and could decrease the overall measurement duration.

7.2 Overlapping

If a better SNR is required it is still necessary to average over several measurements. Which means to use a periodic signal. The idea of overlapping is to do one of these measurements while the previous one done on the same system is not finished yet. This is only possible with systems which have a low number of harmonics. This is the case for weakly nonlinear systems. It works as long as the highest harmonic of the system is not interfering with the measurement of the previous system.

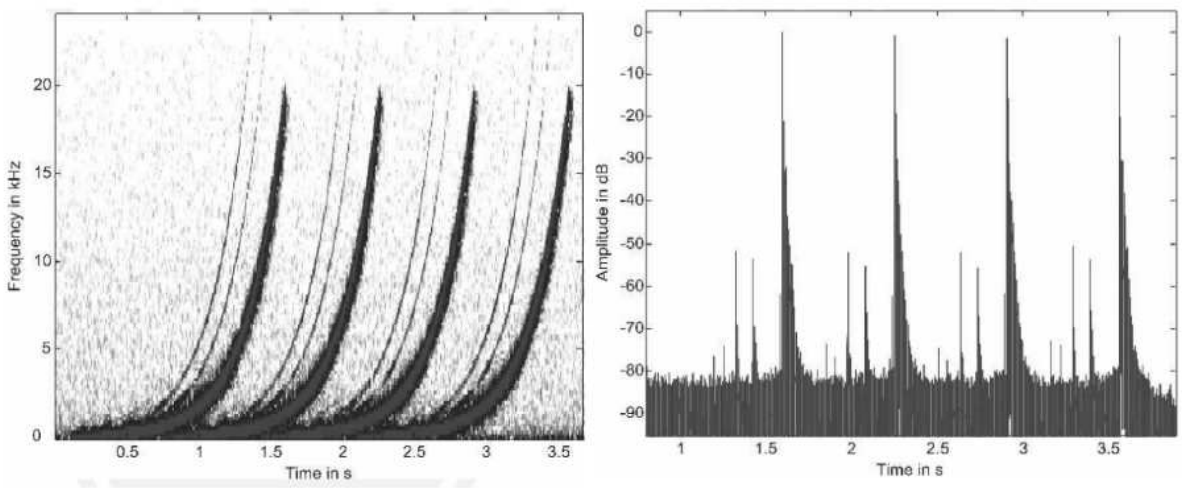

Figure 11: left: Spectrogram of the overlapping methods y; right: according $h[n]$ in dB

8 Conclusions

A short duration of the HRTF measurement impedes the accuracy of the measurement results in terms of SNR on the one hand. However, on the other hand long measurement durations are tedious and unacceptable for the subjects who must endure the whole HRTF measurement procedure in a still position. As a possible consequence to the long duration, subjects might move during the HRTF measurement, and the HRTFs will suffer from poor accuracy as well.

It seems to be a good tactic to use long sweeps which causes a good SNR on there own (without the benefit of averaging). And to pack them as dense as possible with the interleaving and the overlapping methods. Like shown in figure 12.

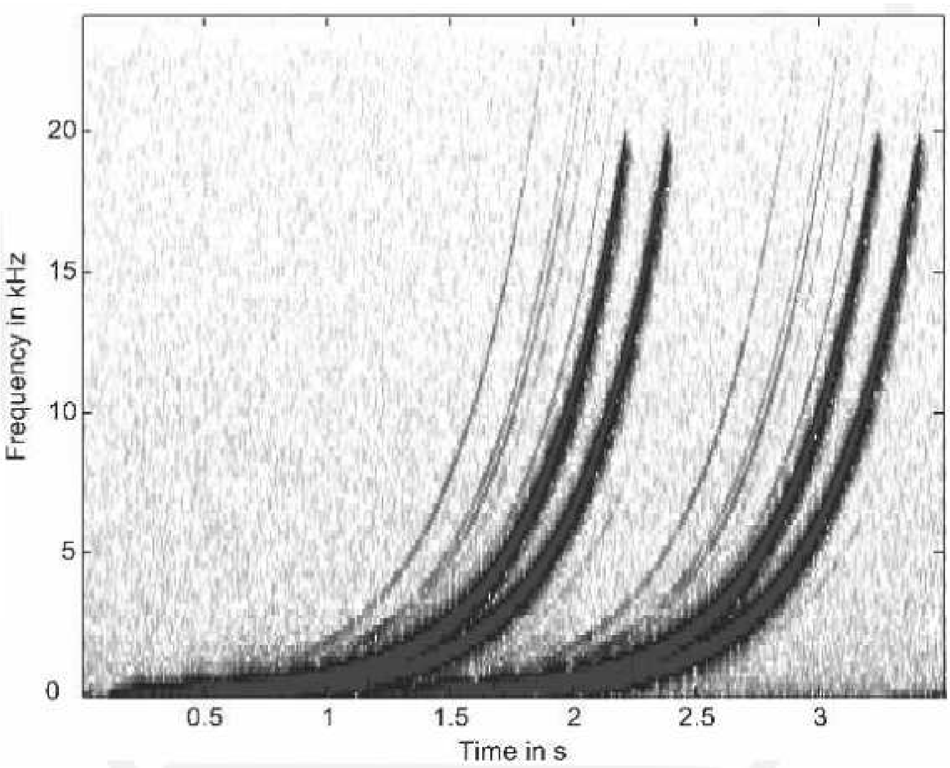

Figure 12: Output signal y of the combination of overlapping and interleaving

References

[Majdak 1] Piotr Majdak et al.,Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions, J. Audio Eng. Soc., Vol. 55, No. 7/8, 2007 July/August.

[Farina 1] Angelo Farina, Simultaneous measurement of impulse response and distortion with a swept-sine technique. Presented at the 108th Convention 2000 February 19-22 Paris, France.

[Minnaar 1] P. Minnaar, J. Plogsties, and F. Christensen, Directional Resolution of Head-Related Transfer Functions Required in Binaural Synthesis, (J. Audio Eng. Soc., vol.53, pp. 919 (2005 Oct.))

[Majdak 2] , Piotr et al.; Multiple Exponential Sweep Method for Fast Measurement of Head Related Transfer Functions; Audio Engineering Society Convention Paper Presented at the 122nd Convention 2007 May Vienna, Austria

[Zotter 1] Franz Zotter; Log Sweep Spectrum; febuary 2007, not published yet.