# Measuring the impulse response of microphones using white noise

Dr. Hans R.E. van Maanen (Temporal Coherence B.V., Netherlands)
Rens Tellers (Sonic Designs B.V., Netherlands)

Date of issue: 23 April 2023

## Abstract

The temporal resolution of a recording chain is also determined by the impulse response of the microphones. But these are rarely specified by the manufacturers because these state that these are hard to obtain. An alternative technique to determine the impulse response of microphones is described, which reproduces very well and avoids the pitfall of non-linear responses. The applicability with real loudspeakers is determined in the first step, after which a Monte Carlo simulation of the complete system is performed. This step showed that this approach should be feasible as it reproduced the imposed impulse response correctly, so the third step is to test it with real microphones. Using relatively standard test equipment, the algorithms yielded frequency response curves, which were in close agreement with the specification of the manufacturers, so it can be expected that the impulse responses are determined with a high accuracy. As only short time intervals are required, the use of an anechoic room is not necessary. As this measurement also reveals the phase response of the microphone, it is also possible to apply corrections, both analog and digital, to improve the impulse response of microphones without any mechanical changes. Preliminary tests have revealed that this leads to audible improvements.

## 1. Introduction

The evidence that the temporal response of audio systems is of prime importance to achieve the goal to reproduce sounds in a more natural way, is mounting. The perceptual improvement of High-Resolution Audio recordings is supporting this observation as it has a higher temporal resolution than the CD-format (refs. 1 and 2). But a chain is as strong as its weakest link and as a most recordings are made using microphones, the temporal resolution of these components is crucial in the whole chain and its optimization. So, the temporal resolution of microphones should be known, e.g., by its impulse response, but these are rarely published by manufacturers. The common way to measure the impulse response is by using sparks as these can be regarded as the closest approximations to δ-impulses in sound which can be realized. The commonly cited arguments against this technique are i) that sparks do not reproduce sufficiently well, so a large variance in the results is seen and ii) that the sound pressure level can be so high that the microphone is driven into a nonlinear response during the measurement. These arguments sound reasonable, yet some results have been published (refs. 3 and 4). Still, knowledge of the impulse response is attractive for recording engineers and optimization of the audio chain. The authors have the impression that the temporal resolution of microphones leaves room for improvement as often resonances in the high frequency range are used to increase the response without sacrificing the sensitivity of the device. However, resonances introduce time smear, so these should be avoided as much as possible. As these resonances are hard to determine from the frequency response without the accompanying phase response, measurement of the impulse response is

essential to detect resonances and to select devices for perceptual comparisons. An indication that microphones have a steep roll-off, caused by the use of resonances, is the near absence of information above 20 kHz in many recordings, which is consistent with a fourth order low-pass filtering, a bit similar to a fourth order Butterworth characteristic. But as Boyk (ref. 5) has shown, certainly metal percussion instruments have significant contributions above 20 kHz as is shown in one of his measurements (fig. 1):

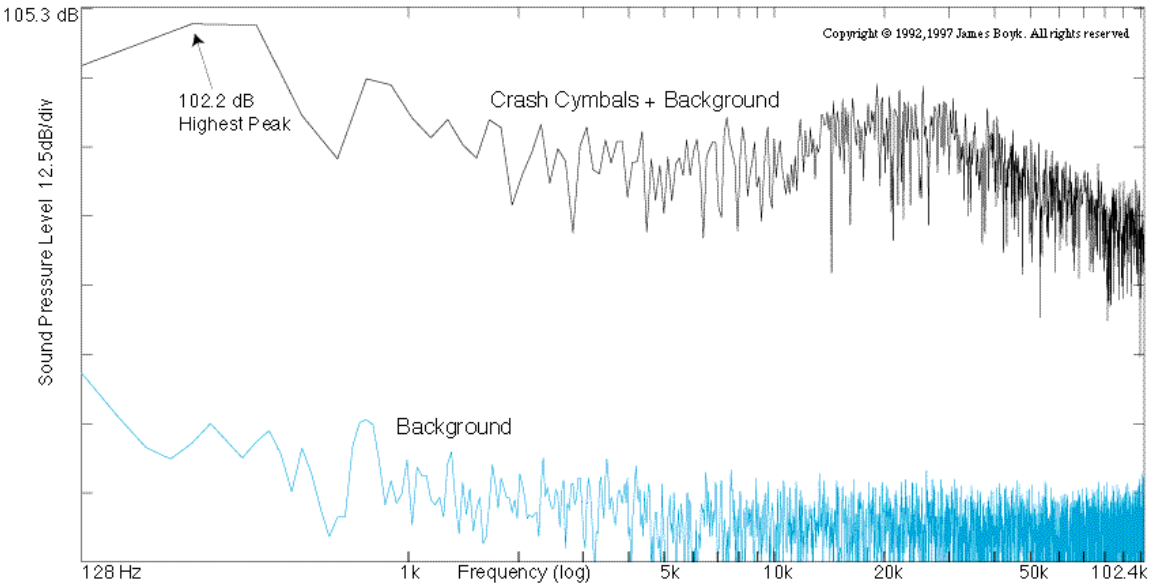


**Figure 13(a)** (Amplitude *vs* frequency). **Crash Cymbals**; 108.3 dB at B&K 4135 microphone with grid off, approximately one and a half feet away. Upper trace: Cymbals + background, corrected to 100 kHz. Lower trace: Background alone.

**Figure 1**: *The spectrum of Crash Cymbals, as measured by Boyk (ref. 5). Note the contributions above 20 kHz.*

In this paper we will describe an alternative way to measure the impulse response of microphones using (white) noise. By this approach, some stumbling blocks, which occur with the more common approach using sparks, can be circumvented, but other hurdles need to be overcome.

In sec. 2, the principles of the technique will be described and in sec. 3, it will be analyzed whether the algorithms can successfully be applied when a real loudspeaker is used. In sec. 4, a Monte Carlo simulation of the measurement with noise is presented, which will be discussed in sec. 5. In sec. 6, the tests of real microphones is described and the results are discussed in sec. 7. In sec. 8, the conclusions of this work will be listed and in Appendix 1 and 2, the procedure to generate Gaussian distributed white noise in the computer is described and its properties analyzed.

## 2. The basic technique and the related mathematics

### *2.1 The problems with the spark technique*

A logical option is to use the sound, generated by a spark discharge, as the sonic equivalent of a δ-pulse. However, a close look at its properties reveals that it is not perfect in frequency domain (ref. 4) and thus neither in time domain. Also, no two sparks are completely identical, so the

deviations and the consequences for the impulse response are unknown. And there is another problem: the spark generates a high sound pressure level, so how can be certified that the microphone and the subsequent electronics are still operating in their linear ranges? This problem could be tackled by using a smaller, less powerful, spark, but that will decrease the signal-to-noise ratio (SNR). The SNR can be improved by 'conditional averaging' (ref. 6), but this requires a complete absence of ambiguity in the sound, generated by all the individual sparks. This is, probably, asking too much, leading to a 'smear' of the averaged impulse response, an undesirable phenomenon, exactly the opposite of what needs to be achieved.

It thus shows that the 'spark' technique has its limitations and stumbling blocks, which hamper its application. Therefore, a different technique to measure the impulse response would be attractive, especially when this could be realized without the use of very specialized equipment.

### *2.2 The use of white noise for the measurement of impulse responses*

From stochastic signal analysis, it is known that there is a one-to-one relation between the transfer function of a system and its response to Gaussian distributed white noise (ref. 6). Without repeating the theory, some of its theorems will be stated without proof. The ones, which are needed in this paper are:

- The autocorrelation function of white noise is a delta function at $\tau = 0$ (albeit with an amplitude of 1 (one)).
- The cross correlation function between the input Gaussian distributed white noise signal and the response of the system equals the impulse response of the system (including its time delay).

From Fourier theory, it is known that multiplying complex valued frequency responses of systems equals convolution of the impulse responses in time domain (refs. 7 and 8).

This basically opens the door to measure the impulse response of a system using Gaussian distributed white noise. This has the advantage that it can be certified that the system is never driven into non-linear responses. Any level of accuracy / uncertainty can be reached by using an averaging time as long as needed.

### *2.3 Basic set-up for the measurement of the impulse response using white noise*

The problem with this approach when applied to microphones is, of course, the generation of the white noise sound. It is no real problem to generate a file of Gaussian distributed noise signal of the required bandwidth and any desired length with computers (ref. 9 and Appendix 1). Amplifiers, which are able to deliver the noise signal with sufficient power over a wide frequency range to a loudspeaker, are no problem either. But to create the related sound field, a loudspeaker, which has both a very wide frequency response and an excellent impulse response itself, is required. This is not possible with the current 'state of the art' of loudspeakers. Actually, even the best tweeters available have an impulse response, which is inferior to the impulse response of many microphones (ref. 2), this seems an insurmountable hurdle. However, there is a way around this.

The loudspeaker can be regarded as a filter and its properties are completely determined by its impulse response. Using an excellent wide-band measurement microphone (which are available, e.g., ref. 10), the impulse response of the loudspeaker, used to generate the sound field, can be measured. It will be labelled a $f_{ls}(t)$. The complex transfer function of the loudspeaker can be determined by Fourier Transformation of the impulse response. It will be denoted as $F_{ls}(\omega)$. The microphone under test (MuT) is also a filter and its complex transfer function will be denoted as $G_{mp}(\omega)$ and its impulse response $g_{mp}(t)$. The overall response of the combination of the loudspeaker and the microphone, denoted as $H_{lm}(\omega)$, is, of course, the product of $F_{ls}(\omega)$ and $G_{mp}(\omega)$. In time domain, this equals the convolution of $f_{ls}(t)$ and $g_{mp}(t)$ and this will be denoted as $h_{lm}(t)$. When the measurement microphone is replaced by the MuT, the measured impulse response equals $h_{lm}(t)$ because the noise is filtered both by the loudspeaker and the MuT. Using the independently measured $f_{ls}(t)$, $g_{mp}(t)$ can be determined by deconvolution of $h_{lm}(t)$ and $f_{ls}(t)$.

Deconvolution is -in general- a bit tricky. So, the first step in this study is to verify the feasibility to obtain the impulse response of a simulated microphone using the measured impulse response of a tweeter. The tweeter impulse response has been measured by Geoff Hill (Hill Acoustics) in the UK and it is shown in fig. 2.

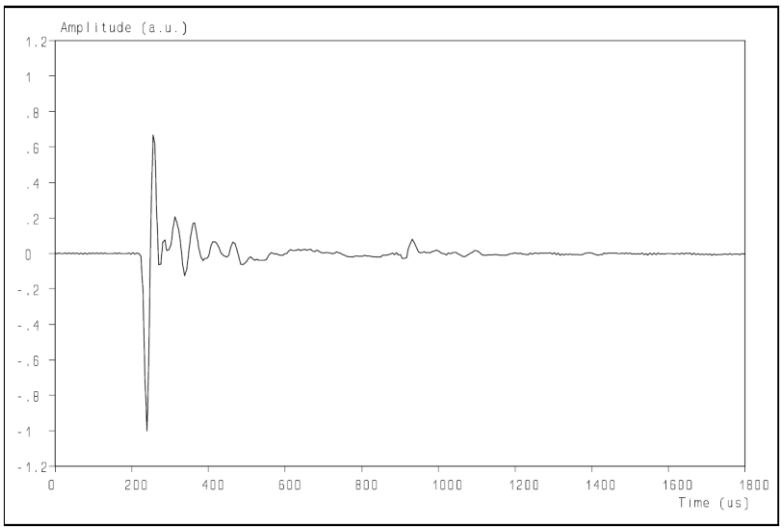


**Figure 2**: *The measured impulse response of the tweeter, used in this study (courtesy of Hill Acoustics). The small peak at approx. 950 µs is probably caused by a damped reflection in the tetrahedral anechoic chamber.*

The properties of the simulated microphone are illustrated in fig. 3 (frequency response) and fig. 4 (impulse response). The properties of the microphone are based on a general approach: the frequency response tends to decrease at higher frequencies in the audio band. To increase the response in the upper part of the audio band, a slight resonance has been added. This resonance can be discerned in the frequency response, but it is far clearer in the impulse response.

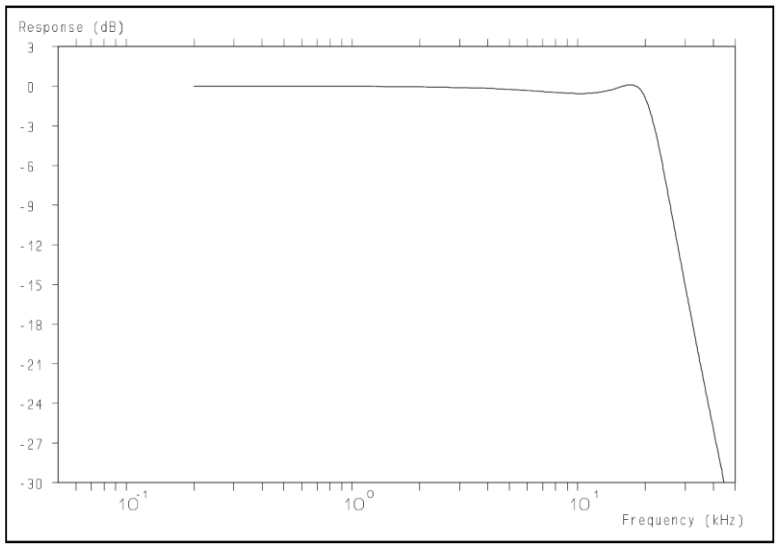


**Figure 3**: *The frequency response of the simulated microphone.*

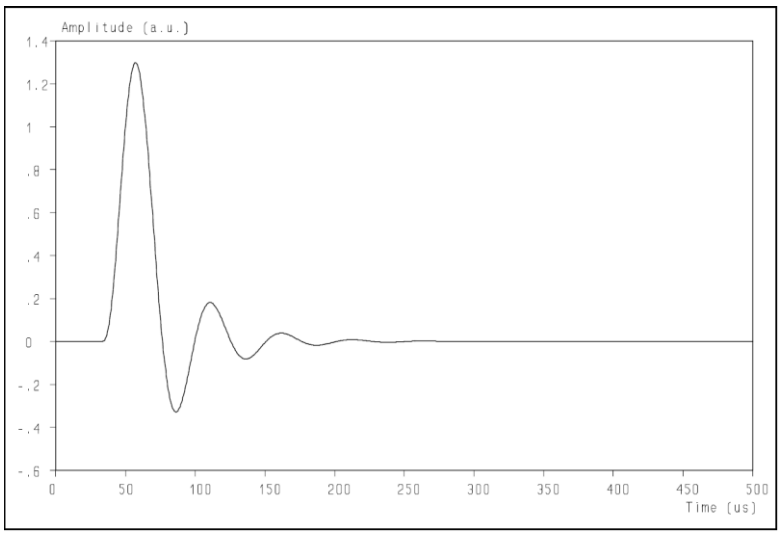


**Figure 4**: *The impulse response of the simulated microphone as derived by Fourier Transformation of its (complex) frequency response.*

The calculation of the convolution of the impulse responses of the loudspeaker and the microphone is straightforward and the result is shown in fig. 5: the signal in time domain as it would come out of the MuT when the impulse would be delivered to the loudspeaker directly (and ignoring non-linearities). The question is now whether it is possible to retrieve the properties of the MuT from the impulse response of the loudspeaker and the signal of fig. 5.

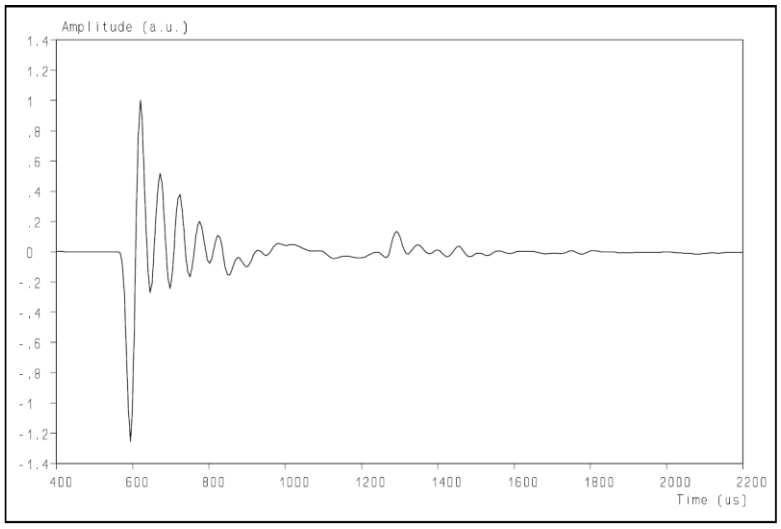


**Figure 5**: *The convolution of the impulse responses of the (measured) tweeter and the simulated microphone.*

## 3. Results and discussion of the first step

The deconvolution process includes the following steps:

- Fourier transformation of $f_{ls}(t)$ to obtain $F_{ls}(\omega)$.
- Fourier transformation of $h_{lm}(t)$, which will yield $H_{lm}(\omega)$, the product of $F_{ls}(\omega)$ and $G_{mp}(\omega)$.
- Dividing $H_{lm}(\omega)$ by $F_{ls}(\omega)$, yielding $G_{mp}(\omega)$.
- Use $G_{mp}(\omega)$ to determine the frequency response of the simulated microphone.
- Inverse Fourier Transformation of $G_{mp}(\omega)$ to determine the impulse response of the simulated microphone.

The result for the frequency response is shown in fig. 6. The frequency response as used for the modelled microphone is shown in fig. 3 and the comparison shows a very good agreement.

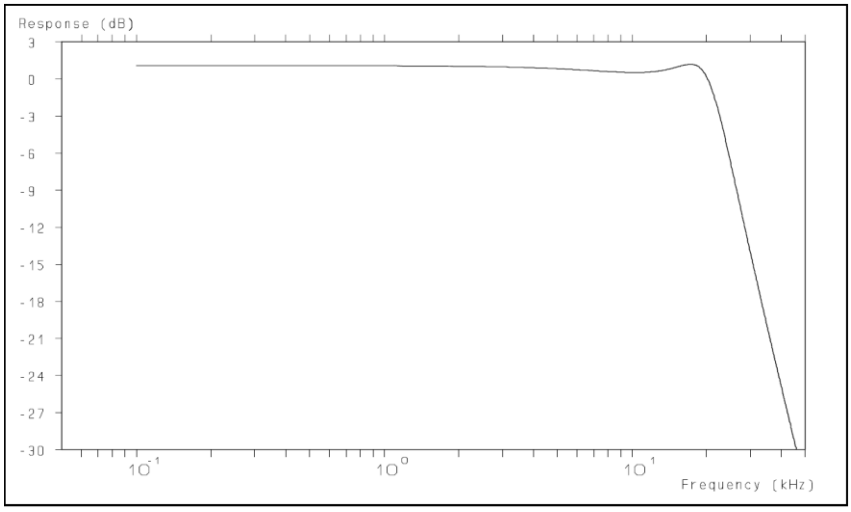


**Figure 6**: *The frequency response of the simulated microphone as determined from the convoluted impulse response of fig. 5. Compare with fig. 3.*

The impulse response of the microphone, using the deconvolution algorithm, is shown in fig. 7. Compared with the impulse response, directly derived from the simulated microphone as shown in fig. 4, there are only minor differences, mainly before the onset of the impulse.

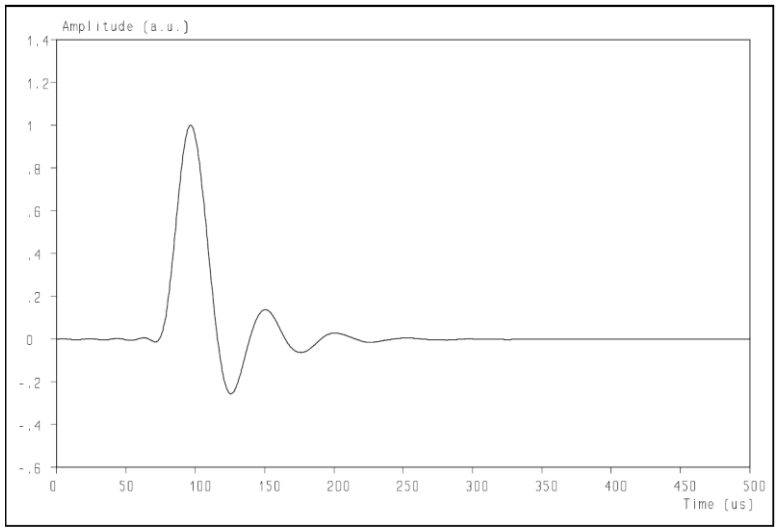


**Figure 7**: *The impulse response of the simulated microphone as retrieved from the convoluted impulse response of fig. 5. Compare with fig. 4.*

The slight oscillation before the onset of the impulse might be caused by aliasing effects and / or truncation errors, but it is not unlikely that a further development of the algorithm can improve the result.

The deconvolution did yield the frequency response curve very accurately. The impulse response also came very close to the actual response. As the impulse response of a real tweeter has been used, this major hurdle can be tackled. Even though the impulse response of the tweeter is inferior to the impulse response of the simulated microphone, the algorithm reproduced it well. Yet, it is obvious that the shorter the impulse response of the loudspeaker is, the better. Which is why it is recommended to use a tweeter which has excellent temporal properties (ref. 2), but it will still be inferior, compared to the white noise signal applied and those of the MuT. It can thus be concluded that that the deconvolution procedure ‘as such’ can be applied to retrieve the impulse response of the microphone from the convoluted responses of tweeter and microphone. Yet, a major step in the whole procedure will now be included in the study: the measurement procedure with Gaussian distributed white noise will be evaluated using a Monte Carlo simulation.

## 4. The use of Gaussian distributed white noise in a Monte Carlo simulation

The technique under study is to excite the loudspeaker with Gaussian distributed white noise and to record the sound from the loudspeaker with the MuT. The signal from the microphone is stored simultaneously with the exciting noise signal so these can be cross correlated. For the limiting case of the measurement time going to infinity, the cross-correlation function is identical to the convolution of the impulse responses of the loudspeaker and the MuT (ref. 6). So, now a Monte Carlo simulation is performed using a computer file, consisting of 5.2 seconds of Gaussian distributed white noise with a sampling frequency of 192 kHz. The same sampling frequency has

been chosen identical as has been used for the loudspeaker impulse measurement and the record entails 1 million samples. The procedure for its generation can be found in ref. 9 and Appendix 1. **(N.B.** It should be noted that this signal extends up to 96 kHz as white noise when directly used as input to a D/A converter, which could introduce further limitations, due to its reconstruction filter. The noise signal has been analyzed to verify whether it fulfills the requirements for this simulation; this analysis is reported in Appendix 2.)

This noise signal is subsequently convoluted (in time domain) with the measured impulse response of the tweeter (see fig. 2), resulting in the loudspeaker output signal in time domain. We will refer to the sound, coming from the loudspeaker, as the 'loudspeaker filtered noise'.

The loudspeaker filtered noise can subsequently be convoluted (in time domain) with the impulse response of the microphone. This results in the temporal output signal of the microphone as it would be recorded when the sound field, reaching the microphone, is the loudspeaker filtered noise. We will refer to this signal as the 'microphone output signal'.

The microphone output signal can now be cross correlated with the Gaussian distributed white noise which has excited the loudspeaker. In the ideal case with an infinite averaging time, the cross-correlation function would be identical to the convolution of the loudspeaker impulse response and the microphone impulse response, shown in fig. 5. The cross-correlation function is shown in fig. 8 and as can be seen, the differences between figs. 5 and 8 are invisible to the unaided eye. But this does not yet mean that the results, derived for the properties of the microphone are (almost) identical when the cross-correlation function is used in the deconvolution process. This is an essential requirement for the practical application of the technique. Yet, as a preliminary conclusion, it can be stated that the averaging time of 5 seconds and 1 million samples is sufficient to approach the theoretical result with a close resemblance.

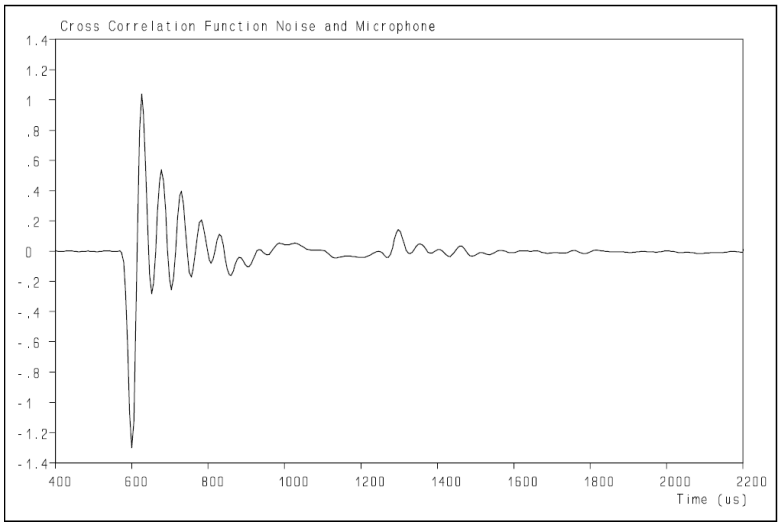


**Figure 8**: *The cross-correlation function between the exciting Gaussian distributed white noise and the microphone output signal, converging to the impulse response of the loudspeaker + simulated microphone. Compare with fig. 5.*

## 5. Results and discussion of the second step

The cross-correlation function of fig. 8 can be used as input for the deconvolution algorithm instead of the impulse response of fig. 2. The output of the deconvolution algorithm consists of the frequency response curve and the impulse response. These are shown in figs. 9 and 10.

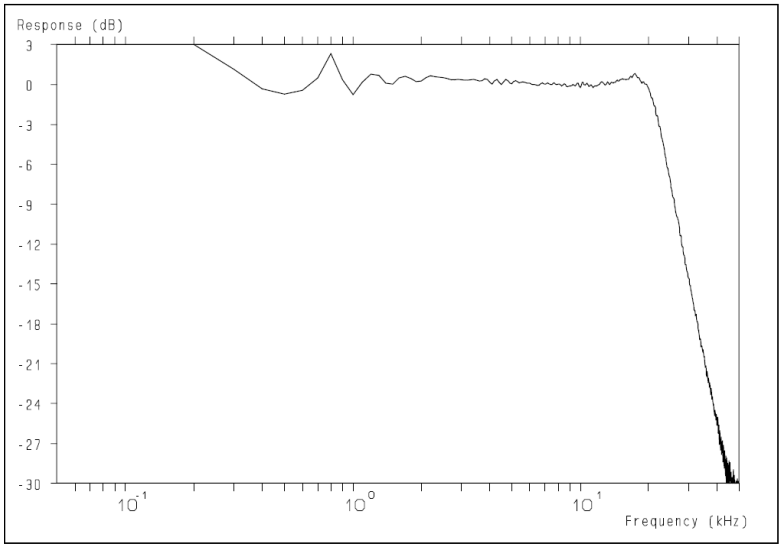


**Figure 9**: *The modulus of the frequency response of the simulated microphone as derived by deconvolution of the cross-correlation function of fig. 8, using the measured impulse response of the loudspeaker as shown in fig. 2. Compare with figs. 3 and 6.*

The frequency response curve is obviously 'wigglier' than the input curve, which is shown in fig. 3 and the retrieved frequency response curve, using the measured tweeter response, shown in fig 6. Yet, the overall properties can still be revealed.

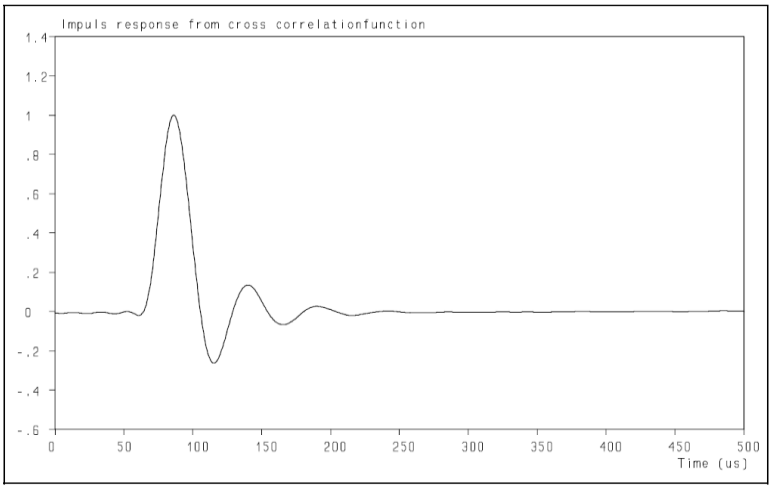


**Figure 10**: *The impulse response of the simulated microphone as derived by deconvolution of the cross-correlation function of fig. 8, using the measured impulse response of the loudspeaker as shown in fig. 2. Compare with figs. 4 and 7.*

The impulse response though, is barely indistinguishable from the impulse responses as shown in figs. 4 and 7. A very detailed comparison shows that the impulse response, derived from the cross-correlation function, has a very small offset just in advance of the impulse and shows a very slight wiggle in the tail, these imperfections are neither present in fig. 7. Overall, the retrieval of the impulse response is excellent and the degradation, compared to the response as shown in fig. 7, can be ignored in most practical applications.

As shown in fig. 9, the variance increases at lower frequencies. This can, qualitatively, be understood by the decreasing statistics for lower frequencies: the number of cycles in the 5 sec. signal is less at lower frequencies. Furthermore, the signal strength decreases with lower frequencies, due to the properties of the tweeter as can be seen from fig. 11. On top of this, the recording is only for short periods of time (2 millisec), compared to the cycle time of lower frequencies (300 Hz corresponds to 3.3 millisec).

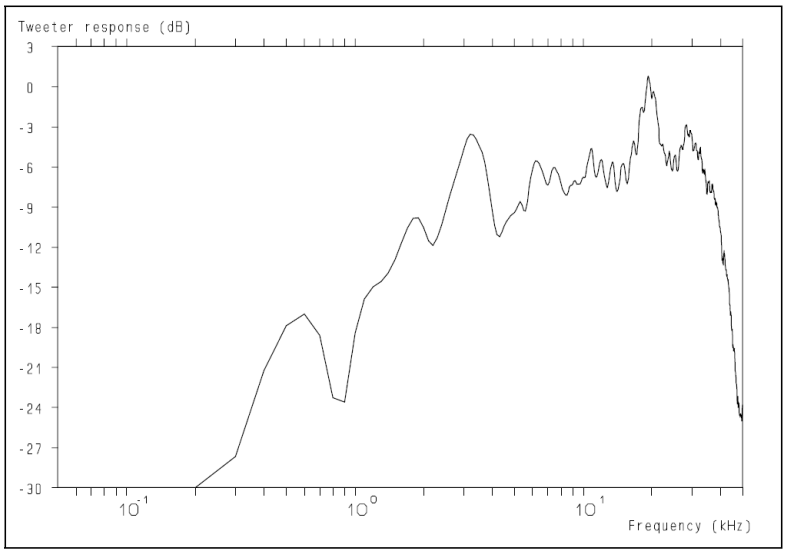


**Figure 11**: *The modulus of the tweeter response as a function of frequency, as derived from the measured impulse response by Fourier Transformation. Note the decrease in response below 300 Hz.*

Fig. 11 shows the modulus of the tweeter response as a function of frequency, derived from the measured loudspeaker impulse response. Note that the technique of this study focusses on the short duration of the impulse response and not on an optimal determination of the frequency response curve. Other techniques are far better suited for that purpose.

A part of the slight imperfections could be attributed to the statistical approach which is inherent to the use of Gaussian distributed white noise for the determination of the impulse response. It is possible to determine the impulse response of the loudspeaker by cross correlation of the loudspeaker filtered noise and the exciting Gaussian distributed white noise. When this is done, it is not unreasonable to assume that the statistical variability in it will be similar to that in the microphone output signal. The use of this loudspeaker impulse response, instead of the independently measured one, could result in less imperfections for the microphone results. To verify this assumption, the loudspeaker filtered noise has been cross correlated with the Gaussian

distributed white noise input signal, as shown in fig. 6. Subsequently, this result has then been used for the deconvolution algorithm and the results are presented in figs. 7 and 8.

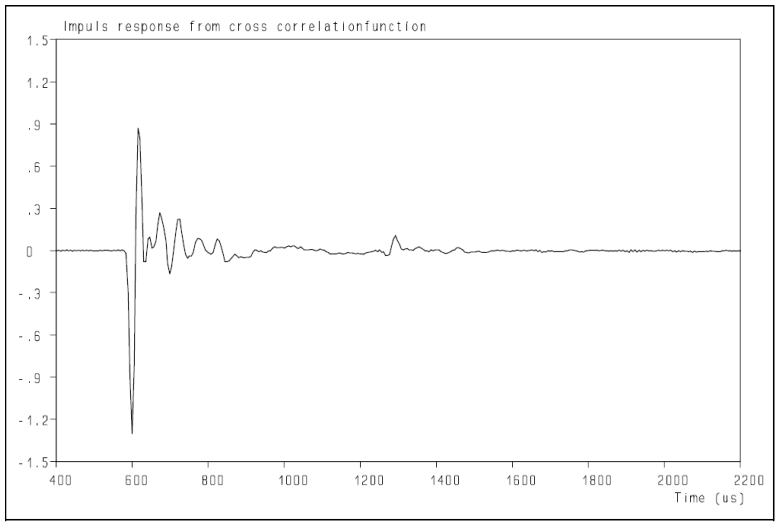


**Figure 12**: *The cross-correlation function between the exciting Gaussian distributed white noise and the loudspeaker filtered noise, converging to the impulse response of the loudspeaker. Compare with fig. 2.*

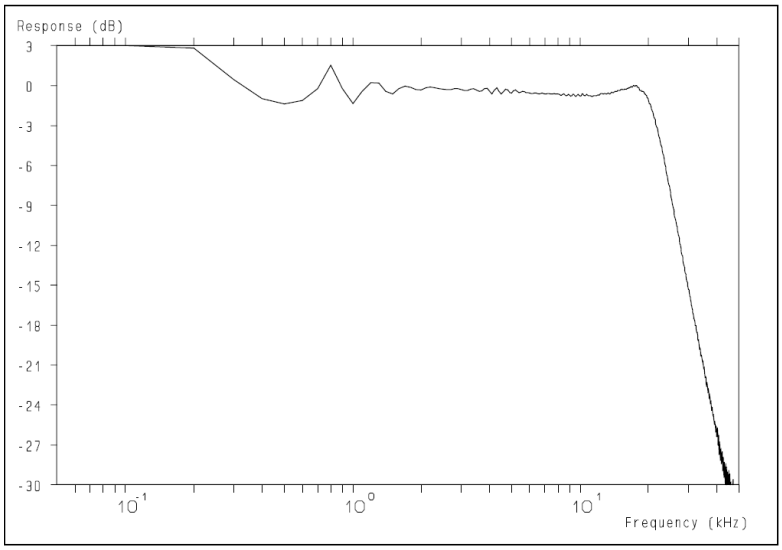


**Figure 13**: *The modulus of the frequency response of the simulated microphone as derived by deconvolution of the cross-correlation function of loudspeaker + microphone (see fig. 8). Instead of the independently measured impulse response (see fig. 2), the cross correlation of the loudspeaker filtered noise and the Gaussian distributed white noise (see fig. 12) has been used. Compare with fig. 9.*

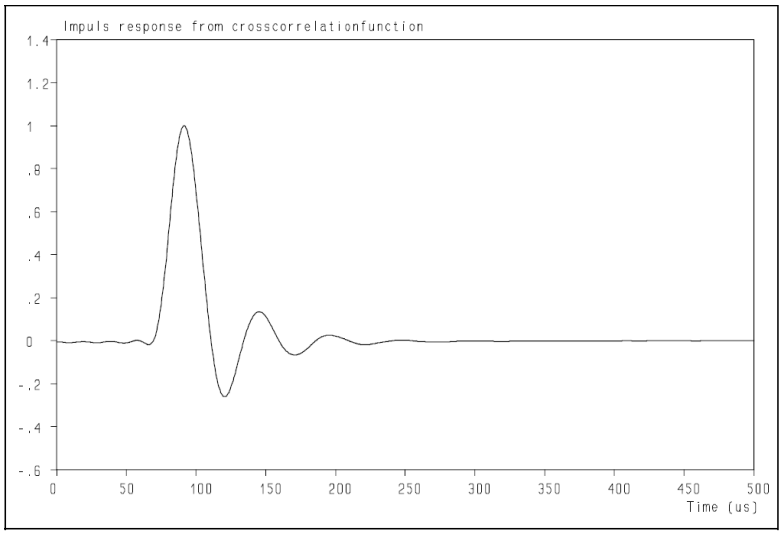


**Figure 14**: *The impulse response of the simulated microphone as derived by deconvolution of the cross-correlation function of loudspeaker + microphone (see fig. 8). Instead of the independently measured impulse response (see fig. 2), the cross-correlation of the loudspeaker filtered noise and the Gaussian distributed white noise (see fig. 12) has been used. Compare with fig. 10.*

These results, presented in figs. 13 and 14, do show a decrease in the variability when compared with figs. 9 and 10. By enlarging the vertical scale by a factor of 25, the differences between the two results become clearer, as is shown in figs. 15 and 16. So at least a part of the variability can be attributed to the statistical nature of the measurement technique. Another advantage is, of course, that the impulse response of the loudspeaker is determined 'on the fly'. So, there is no need to verify whether the loudspeaker properties have changed between the initial measurement and the measurement of the MuT.

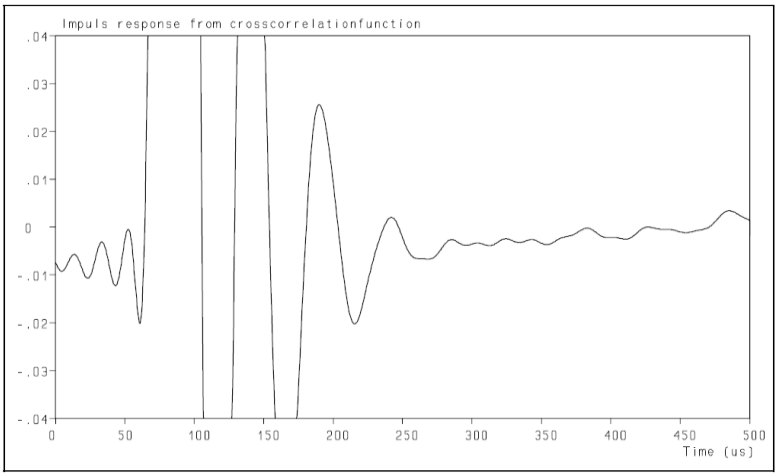


**Figure 15**: *The imperfections of the impulse response of fig. 10 are shown more clearly by extending the vertical scale by a factor of 25. Note the offset before the onset of the impulse response and the variability of the response in the tail.*

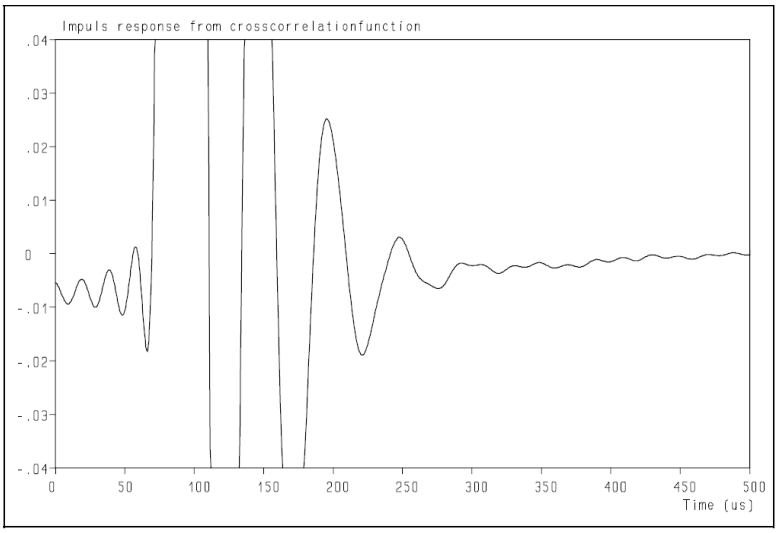


**Figure 16**: *The imperfections of the impulse response of fig. 14 are shown more clearly by extending the vertical scale by a factor of 25. Note that both the offset before the onset of the impulse response and the variability in the tail are reduced, compared to fig. 15.*

The second step of the feasibility study has revealed that the use of Gaussian distributed white noise for the determination of microphone impulse responses is a realistic option, even with noise signals as short as 5 seconds. It showed that the cross-correlation function, obtained is very close to the direct convolution of the loudspeaker and microphone impulse responses.

Based on the results of the Monte Carlo simulation, it is logical to assume that the determination of the loudspeaker impulse response using the same cross-correlation technique shortly before the measurement of the MuT is the best approach as the location of the measurement microphone and the MuT can be as close as possible.

The advantage of Monte Carlo simulation is that the properties of the MuT are known in advance, so the output can be compared with the input (ref. 9). These simulations have shown that the technique provides accurate results for the frequency response and the impulse response of the MuT. The next step is to apply the technique to a real microphone to determine whether practical hurdles can be overcome.

## 6. The impulse response measurement in an experimental setup

The setup to determine the impulse response of a real microphone uses a tweeter, which is known to have one of the best impulse responses (ref. 2), albeit that it is nowhere near to a Dirac pulse and its impulse response is wider than the microphone impulse response itself. It has been mounted in a frame and it is excited by white noise from the generator of Audio Precision APX 525. A Gras measurement microphone Model 46DD was first mounted 150 mm in front of this tweeter. The frequency response of the Gras microphone, as specified by the manufacturer for the one used, is shown in fig. 17.

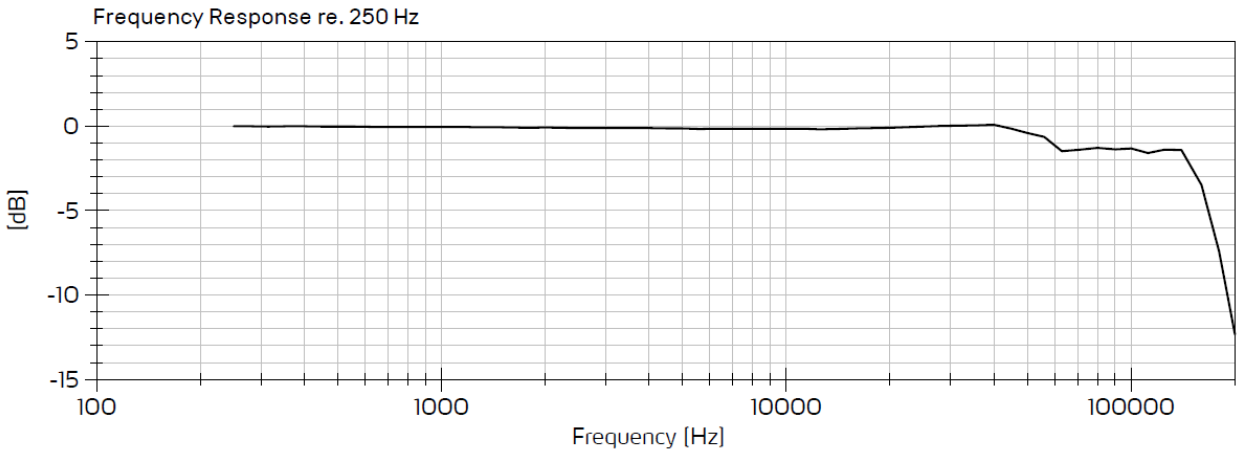


**Figure 17**: *Frequency response of the Gras measurement microphone Model 46DD as used in the tests, described in this paper.*

A picture of the setup is shown in fig. 18. The tweeter is on the left, the Gras microphone to the right in the black microphone stand.

**Figure 18**: *Setup of the measurement. For details: see text.*

The driving signal and the signal from the microphone have been recorded simultaneously for 100 seconds with a 24-bits bitrate and sampling frequency of 624 kHz by the Audio Precision and are edited in a wave-editor (Steinberg WaveLab) to remove the pre-sweep and settling trails, to determine the gain level of measurement signal, to normalize the microphone level to match

measurement peak level and to export text-file output for further processing. From these files, the cross-correlation function has been calculated, which is used to determine the impulse response of the tweeter (ref. 6). The result is shown in fig. 19. The modulus of its Fourier Transform is shown in fig. 20. The frequency response corresponds well with its specifications.

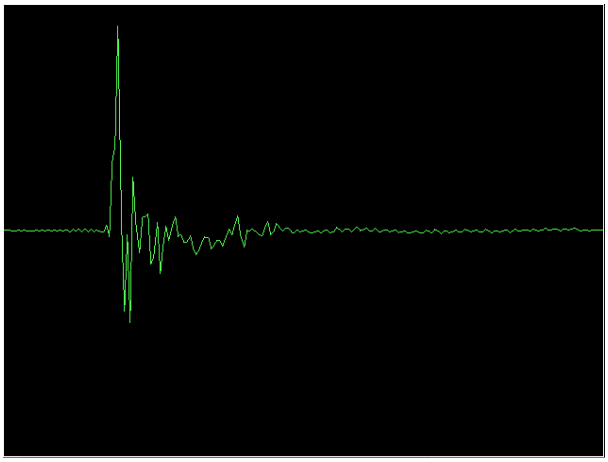

**Figure 19**: *Impulse response of the tweeter. Vertical is the correlation coefficient (dimensionless), horizontal scale 1.04 ms.*

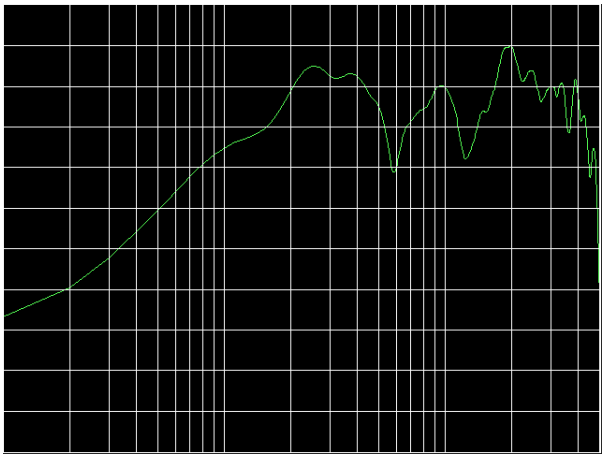

**Figure 20**: *Frequency response of the tweeter as derived from the impulse response of fig. 19. Vertical scale 3 dB/div. Horizontal scale 100 Hz – 50 kHz.*

Subsequently, the Gras measurement microphone has been replaced by two different microphones under test (MuT). As the aim of this study is to study the feasibility of this technique, no brands or type number are reported here to avoid any competitive advantage or disadvantage; brands and models are known to the authors. The position of the membrane of the MuT was located as close to the position of the membrane of the Gras microphone (150 mm in front of the tweeter) as possible under the given conditions. Then the measurement as described above has been repeated. The results for both microphones (being the convolutions of the tweeter impulse response and the microphone impulse response) are presented in figs. 21 and 22. The scales are identical to those of fig. 19. **N.B.** The same scales for similar graphs will be used throughout this paper.

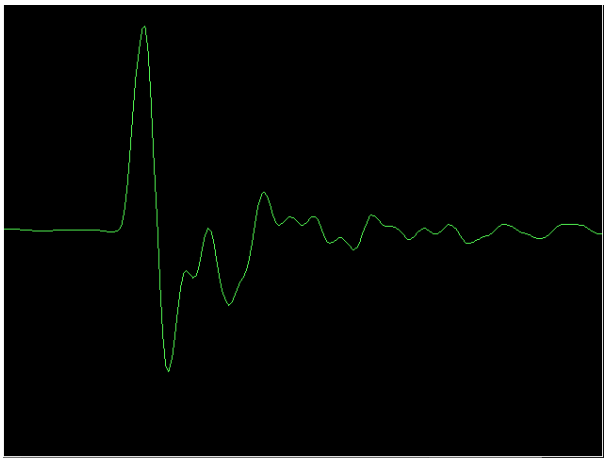

**Figure 21**: *Impulse response Dynamic Handheld Microphone and tweeter before any correction has been applied.*

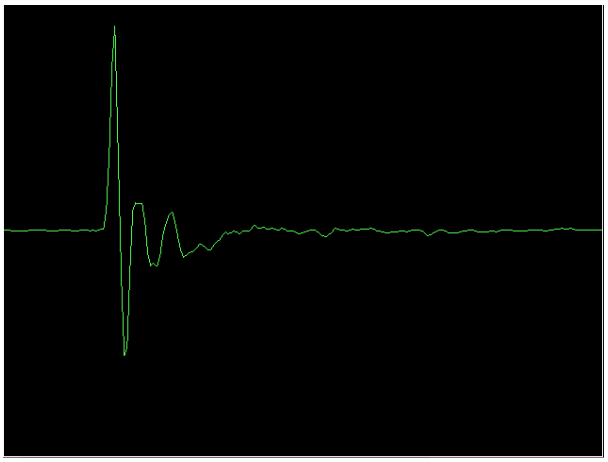

**Figure 22**: *Impulse response of the Condenser Measurement Microphone and tweeter before any correction has been applied.*

The calculation of the deconvolution is done by software, written by the one of the authors, which has been applied to these responses. To see whether the deconvolution works well, the deconvolution can also be used to determine the frequency response of the MuT's. The results are shown in figs. 23 and 24 and these correspond well with the specifications as provided by the manufacturers.

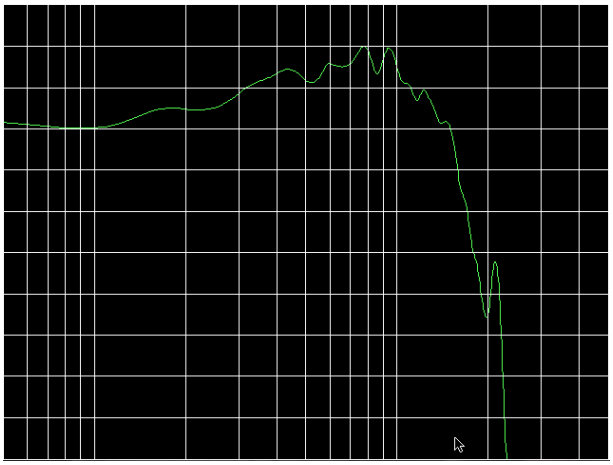

**Figure 23**: *Modulus of the Fourier transform of the impulse response of fig. 21 after deconvolution.*

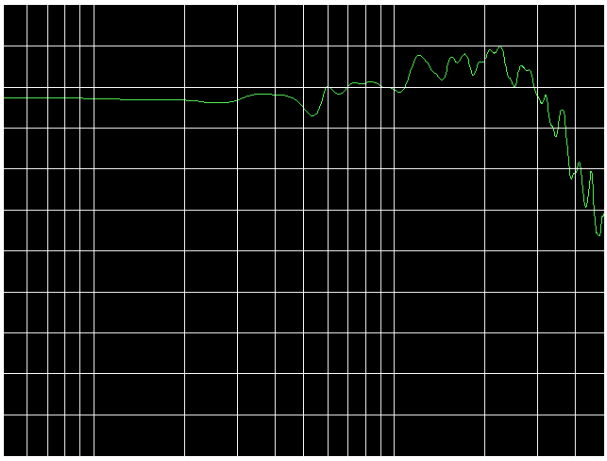

**Figure 24**: *Modulus of the Fourier transform of the impulse response of fig. 22 after deconvolution.*

The final step is to determine the impulse response of the MuT's after deconvolution. To see the effect of the deconvolution more clearly, the impulse responses before and after the deconvolution are presented below.

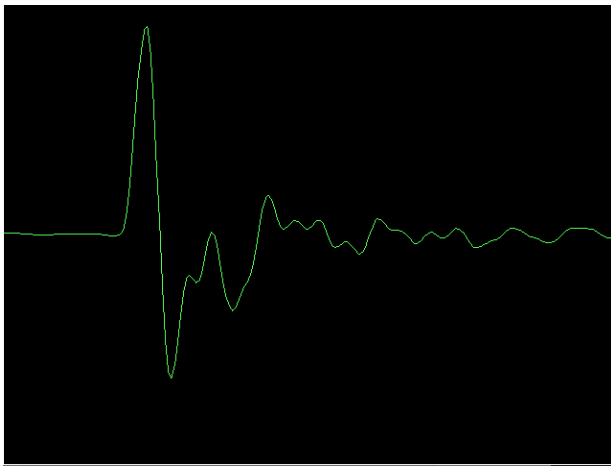

**Figure 21**: *Impulse response of the Dynamic Handheld Microphone and tweeter before any correction has been applied.*

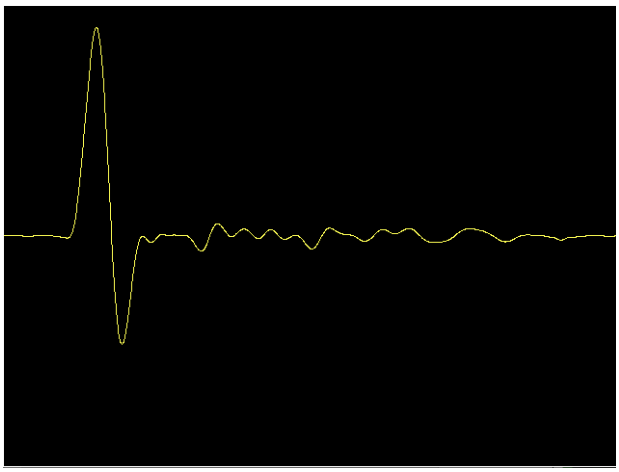

**Figure 25**: *Impulse response of the Dynamic Handheld Microphone and tweeter after the deconvolution has been applied.*

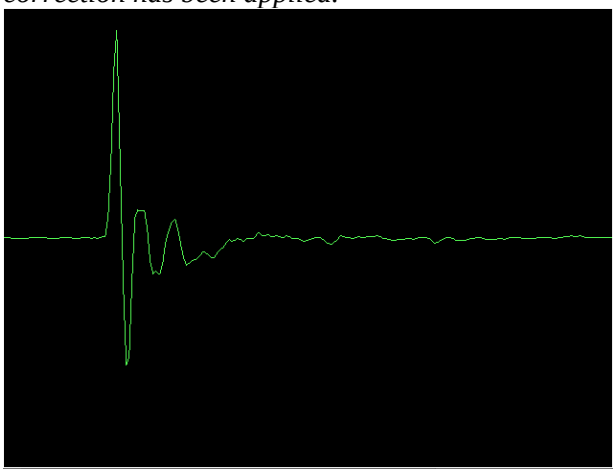

**Figure 22**: *Impulse response of the Condenser Measurement Microphone and tweeter before any correction has been applied.*

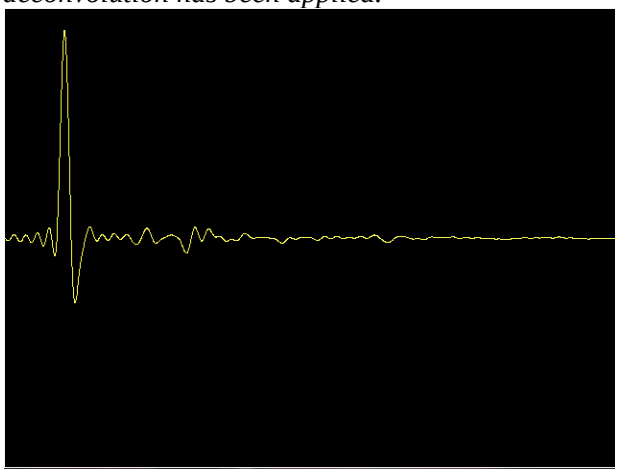

**Figure 26**: *Impulse response of the Condenser Measurement Microphone after the deconvolution has been applied.*

## 7. Discussion

The experiments show that this technique, which feasibility has been demonstrated by Monte Carlo simulations, also works when applied to actual microphones. The frequency responses, which are obtained, correspond well to the manufacturer's specifications and the differences between the 'raw' and the deconvoluted impulse responses is striking. Fig. 20 shows that the high-pass filtering, caused by the tweeter, is clearly visible, but in figs. 23 and 24, it has vanished completely. This further underpins that the deconvolution provides the correct outcome of the MuT properties.

The simplicity of the setup is surprising, but as we are mainly interested in the impulse responses, which show up at short time intervals, the influence of the room reflections can be avoided by using short time windows. Still, if these occur within the time window, these will be identical as long as the setup of the experiment remains the same. It will thus be an artifact in the impulse

response of the tweeter (the actual sound source), but it will cancel during the deconvolution calculation.

That does not mean that using an anechoic room could not result in further improvement of the results, but it is not really necessary to obtain a good and useful result for the impulse response of different microphones, which can be important for the selection of different recordings.

The slight oscillations in fig. 26 are probably caused by truncation errors as the calculation of the frequency response is limited to 50 kHz. Similar effects showed up in the Monte Carlo simulations. Therefore, the measurements have been repeated with sampling frequencies of 192 and 640 kHz. Both results are shown in fig. 27. The differences are minor, albeit that the first flank of the impulse response, as determined by the higher sampling frequency, is indeed a little bit faster and the oscillations have shifted a bit. Yet, the impulse responses are only distinguishable with a direct comparison. However, a higher sampling frequency and a wider noise bandwidth are to be recommended when a microphone with a shorter impulse response is tested.

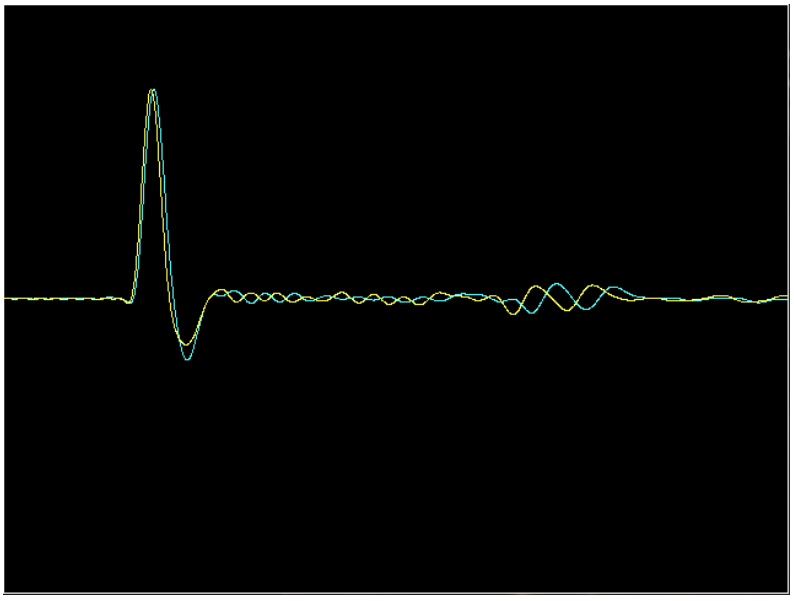

**Figure 27**: *Impulse responses of the of the Condenser Measurement Microphone as determined using a sampling frequency of 192 kHz (blue) and 640 kHz (yellow).*

Also, an extension of the duration of the measurement will, by pure statistical averages, lead to a further improvement of the results. But in our view, these are refinements and optimizations of the technique, just like a different mounting of the tweeter and the microphone. Yet, the determination of the impulse response of microphones does not require expensive equipment and / or an anechoic room. We showed it can be done in a simple laboratory setting with standard test equipment. And this technique neither carries the risk of low reproducibility, like the ‘spark’ technique, nor of overload of the microphone by a high sound pressure level. The authors, present during the tests in the laboratory without any hearing protection, did not experience the sound pressure level as a problem. The only issue was that they should not produce any sound as long as the test was running.

As the technique also provides the complex transfer function of the microphone, it is, theoretically, possible to create an electronic correction to extend the response to higher frequencies and to shorten the overall impulse response. One of the authors has applied this approach in the past, using an analog circuit, to successfully reduce the time smear by the mechanical resonance of moving magnet pick-up cartridges. For microphones, a digital correction is, nowadays, a better approach. By calculating the required inverse impulse response and convolving the digitized microphone signal with it, a significant reduction of the width of the microphone impulse response can be achieved, as shown in fig. 28, where the technique has been used to determine the resulting impulse response of a Condenser Measurement Microphone. The frequency response is roughly an octave higher, the width of the impulse response is halved and the system is now time-correct. So even good microphones can be improved. If the inverse impulse response used is sufficiently long, also 'wiggles' in the microphone characteristic can be reduced.

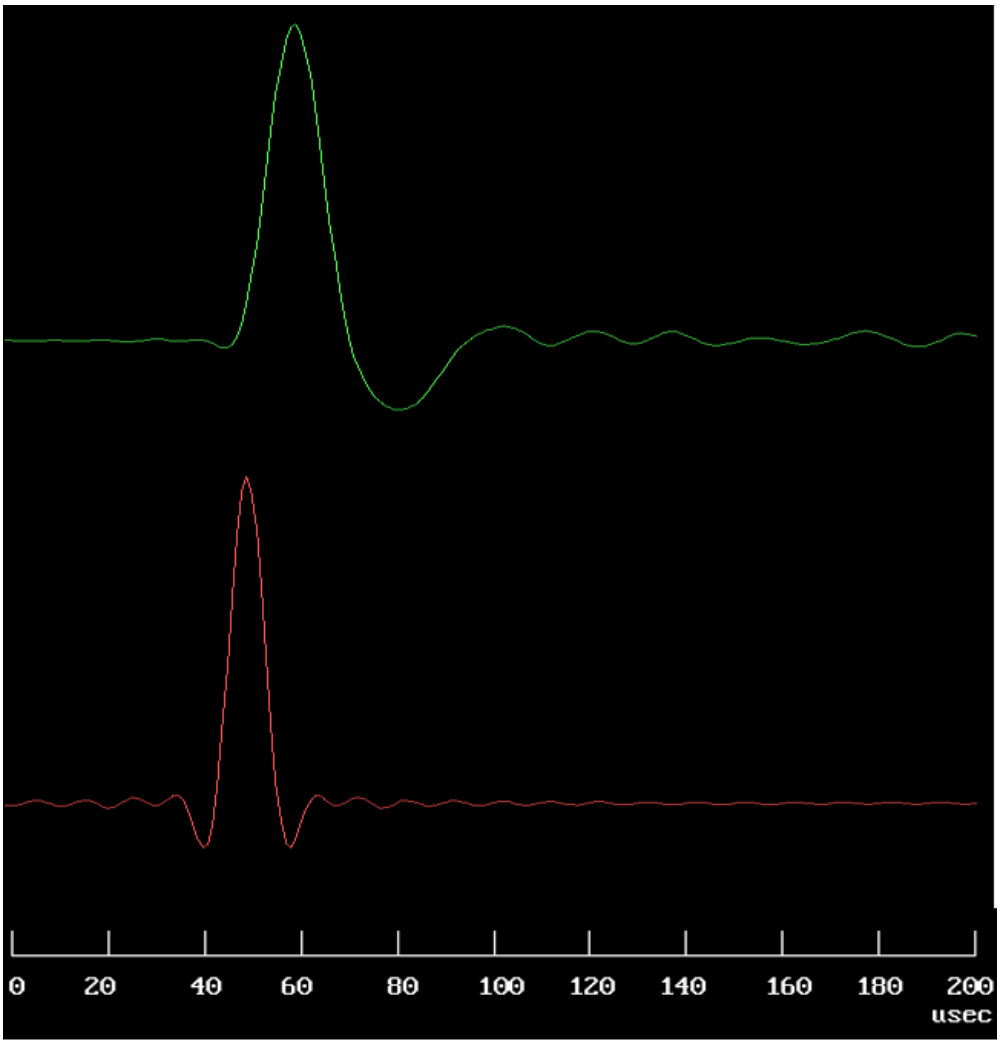


**Figure 28**: *Impulse responses of the of the Condenser Measurement Microphone, upper trace, green, as derived from the 640 kHz data (= yellow curve in fig. 27), 'as is', and after deconvolution with the inverse impulse response, lower trace, red.*

As a final step of these tests, the application of the inverse impulse response to real sound signals has been tested by using a microphone, commonly used for recording of music: a short signal has been perceptually adjudicated 'as is' and after convolving it with the correction impulse response. The original impulse response is shown in fig. 29, green curve, after applying the correction, it should be improved to the red curve. This resulted in a perceptual improvement of the convolved track, which was -not surprisingly- most clear with impulsive sounds, like metal percussion. Although these results are preliminary, it is a clear indication that this approach is at least very promising, and it also further underpins the evidence that the temporal response of systems is of prime importance for natural reproduction of sound.

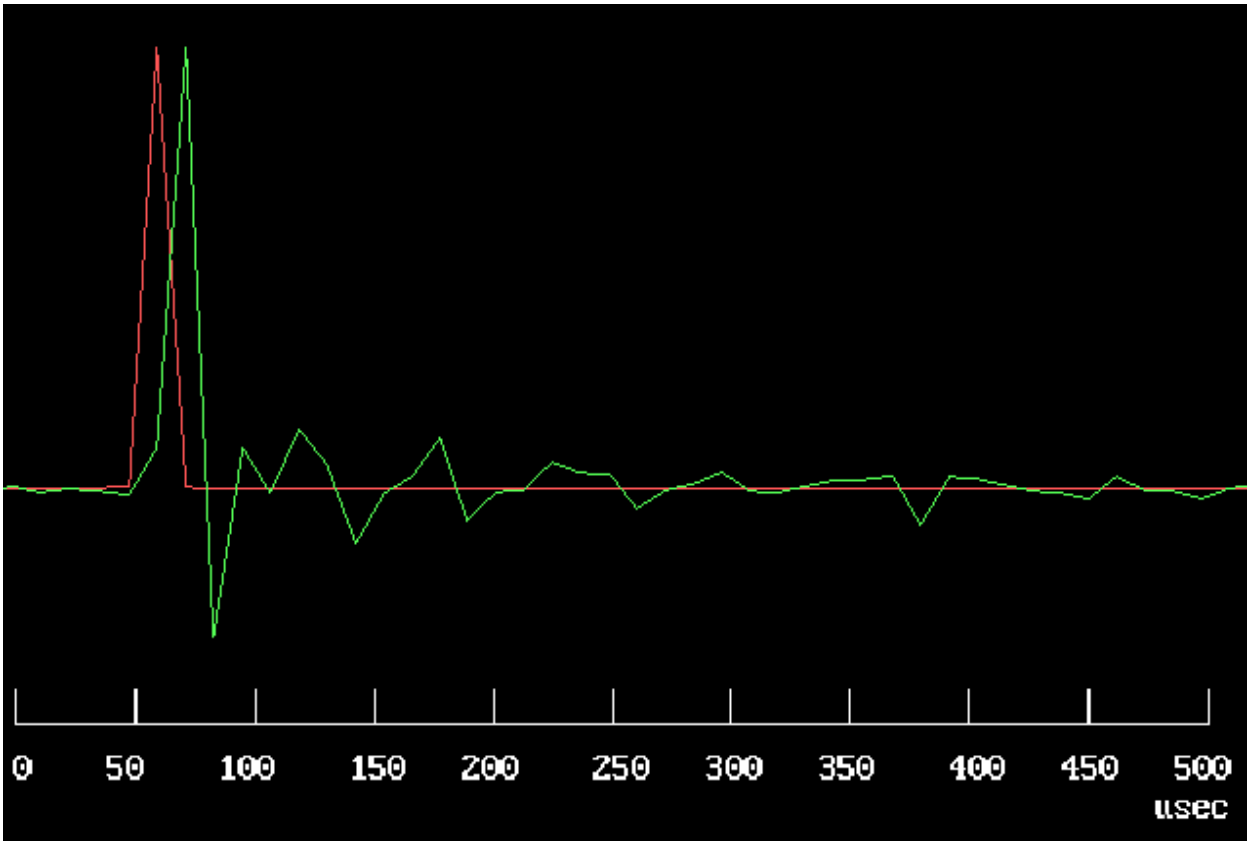


**Figure 29**: *Impulse responses of the of the Recording Microphone, green, 'as is', and after convolution with the inverse impulse response (calculated), red. The sound recorded using the Recording Microphone has been compared 'as is' and after convolution to achieve a better impulse response and less time smear.*

## 8. Conclusions

The impulse response of microphones is an important aspect of the optimization of audio recordings as not only the anti-aliasing filters, the formats and reconstruction filters determine the overall time-smear. However, the number of published microphone impulse responses is small, it is rather the exception than the rule. The technique, described in this paper, enables the determination of the impulse response with relatively simple equipment, independent of the manufacturers willingness to provide these.

The use of a real loudspeaker / tweeter is possible, even though its impulse response is inferior to that of the microphone to be tested. In the first step, it has been demonstrated that the required deconvolution is sufficiently accurate for this purpose.

The determination of the convolution of the impulse responses of the tweeter and the microphone by cross correlation of the driving white noise signal and the microphone output signal has been demonstrated by a Monte Carlo simulation. The determination of the microphone impulse response from this cross-correlation function has been shown to be sufficiently accurate: with a Monte Carlo simulation the properties of both the input and of the output are known and can thus be compared.

The technique can also be applied to actual microphones. Only standard test equipment has been used, so no expensive apparatus is required. The data processing (calculation of the cross-correlation function and deconvolution) has been done by software, written by one of the authors. An anechoic room is not required as long as the transit time of room reflections from the tweeter to the microphone is longer than the time window used.

Deconvolution of the cross-correlation functions yields the impulse responses of the microphones, which are clearly different from the cross-correlation function, in agreement with the Monte Carlo simulations. It also provides microphone frequency response curves, which are close to the specifications of the manufacturers, further underpinning the quality of the results.

Further improvement and optimization can be achieved by using longer time series and / or higher sampling frequencies, a wider bandwidth of the Gaussian distributed white noise and more damping around the setup. But because the impulse response is looking mostly at short time intervals, the effect of reflections is minor to begin with and, if present, will also show up in the determination of the impulse response of the tweeter. So, this is not a real limitation of the technique, but longer time windows are useful to suppress ‘wiggles’ in the frequency response.

This technique is without risk of low repeatability due to reproducibility of the exiting signal (like sparks) and overload of the microphone due to high sound pressure levels. The only requirement is that there are no disturbing sounds during the data-acquisition.

From a perceptional point of view, it would be very interesting to study if there is a relation between the perceived quality of microphones and their impulse responses. As different CD reconstruction filters with different impulse responses give different perceived quality assessments, it is very likely that there will also be differences between microphones, related to their temporal properties. This can be used by recording engineers for optimizing the sound quality under different conditions and for different sound sources.

As this technique also provides the phase response of microphones, and thus its complex transfer function, it opens up the possibility to correct its response to improve the impulse response without any modification of the mechanical system. Such corrections could also simulate microphones with different properties in time domain with the same actual microphone. This can also be helpful with the above-mentioned perceptual comparisons. The first results sound promising, and this should therefore be evaluated further.

**Acknowledgements**

The authors want to thank especially Ir. Menno van der Veen and Prof. Dr. Ir. Ronald van Zolingen for their critical review of the concept and their recommendations for improvements, which certainly helped a lot to raise the quality of this paper.

## APPENDIX 1

### Generation of the simulated Gaussian distributed white noise

---

The file, containing 1 million samples of Gaussian distributed white noise is generated using the Random Number Generator (RND) of the computer. However, the RND generator provides numbers with a rectangular distribution between 0 and 1. This distribution has to be converted into a Gaussian distribution, which is done in the following way:

- Call the RND generator, it returns a value between 0 and 1
- Subtract 0.5 from its value, so it will be in the range between -0.5 and +0.5
- Call again the RND generator, it will return another value between 0 and 1
- Subtract 0.5 from this value too, so it will be a value between -0.5 and +0.5 and add this to the previous value
- Repeat the above procedure until 21 values in the range of -0.5 to +0.5 have been summed

Statistics learn that numbers, obtained in this way, will have a Gaussian distribution, but the related variance will differ from 1 (one). This can easily be solved by normalization. The correct value for normalization is the square root of 1.75 because of the number of summations (21) and the amplitude of the rectangular distribution (0.5).

Each individual sample in the file is generated by the following instructions:

```
Sigma0 = SQR(1.75)          Normalization constant
A = RND - 0.5               Call RND generator and subtract 0.5
FOR L = 1 TO 20             Repeat another 20 times
   A = A + RND - .5         Sum all these values
NEXT L
A = A / Sigma0              Normalize
STORE A                     Store the sample in the file
```

This procedure is repeated 1 million times and the file thus includes 1 million samples of Gaussian distributed white noise with a variance of 1 (one), which can be used for the Monte Carlo simulation. The sampling frequency will be equal to the playback frequency used by the D/A converter. A more detailed description of the generation procedure can be found in ref. 9.

## APPENDIX 2

### Properties of the simulated Gaussian distributed white noise

The results and trustworthiness of the Monte Carlo simulation rely on the correctness of the properties of the Gaussian distributed white noise input, as the theory is based on the assumption that the input noise signal fulfills all the requirements, imposed on it. The three major requirements are:

- The probability distribution of the noise is Gaussian.
- The noise signal should be completely random, it should have no 'memory', in other words, the correlation between any two consecutive samples should be zero.
- The requirement of randomness also means that the correlation between *any* pair of samples should be zero.

To verify whether the data in the noise input file fulfill the above-mentioned requirements, its properties have been verified by common signal analysis techniques. In fig. A-1, the non-normalized probability distribution is shown (wiggly line), together with the theoretical Gaussian distribution curve (drawn line). (**N.B.** Non-normalized means that the integral of the curve is not equal to 1, so it is not a probability *density* distribution, but, for the sake of simplicity, it just has been scaled to 1 (one) at its maximum.)

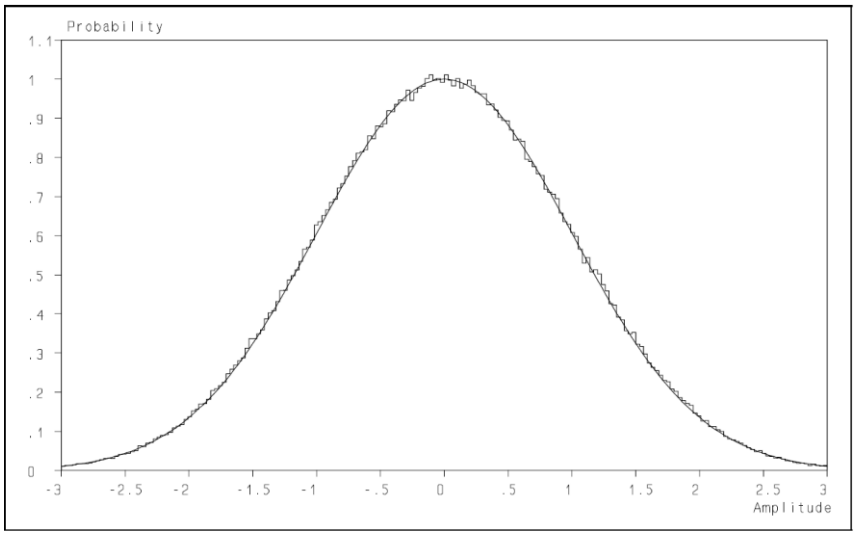


**Figure A-1**: *The non-normalized probability function of the noise, as used in the Monte Carlo simulation (wiggly line) and the theoretical Gaussian distribution function (drawn line).*

**N.B.** The probability function of the noise is wiggly, because the calculation of the distribution uses 'bins' (in this case 200) of a certain width and each sample is assigned as an element to a bin, depending on its value. Therefore, the number of elements in each bin is limited and as a result, this number will show a certain statistical fluctuation, leading to the 'wiggly' character.

It is obvious that the probability distribution follows the theoretical Gaussian distribution very well, in full agreement with the results of ref. 6 and thus fulfilling the first requirement.

The noise should be completely random, so any two samples from the noise file should be completely independent. In other words, there should be no relation between the two, a sample should not be related to *any* other sample. This can be verified by the calculation of the autocorrelation function, which should be 1 (one) for zero time shift (because a sample should fully correlate with itself) and 0 (zero) for all other time shifts (in units of the inverse of the sampling frequency).

The autocorrelation function is shown in fig. A-2 and it behaves accordingly. In order to get a more detailed view of the correlation coefficients with time shifts larger than 0, the value at zero time shift has been put to zero and the vertical scale expanded by a factor of 20. The result is presented in fig. A-3. As can be seen, the correlation coefficients are very close to zero, with some remaining statistical fluctuations, as is to be expected with a limited number of samples (even if this number is a million).

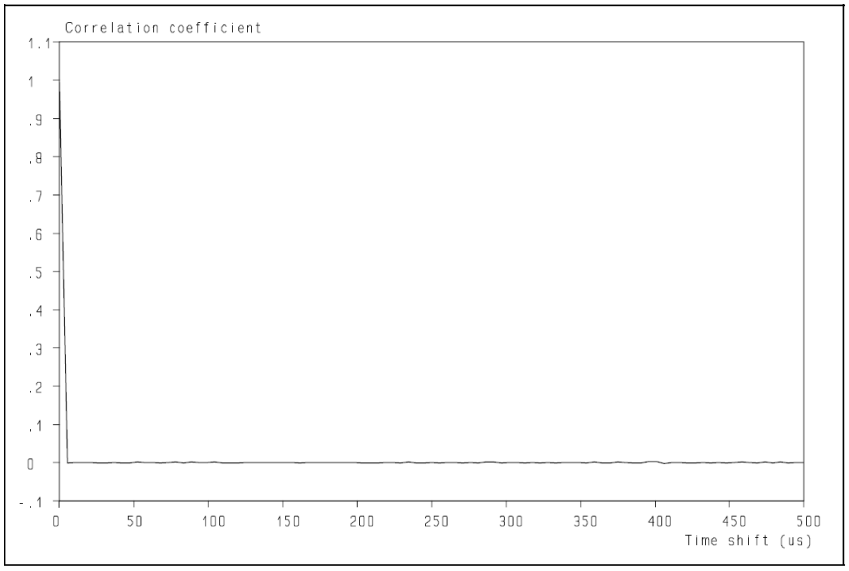


**Figure A-2**: *The autocorrelation function of the noise as used in the Monte Carlo simulation.*

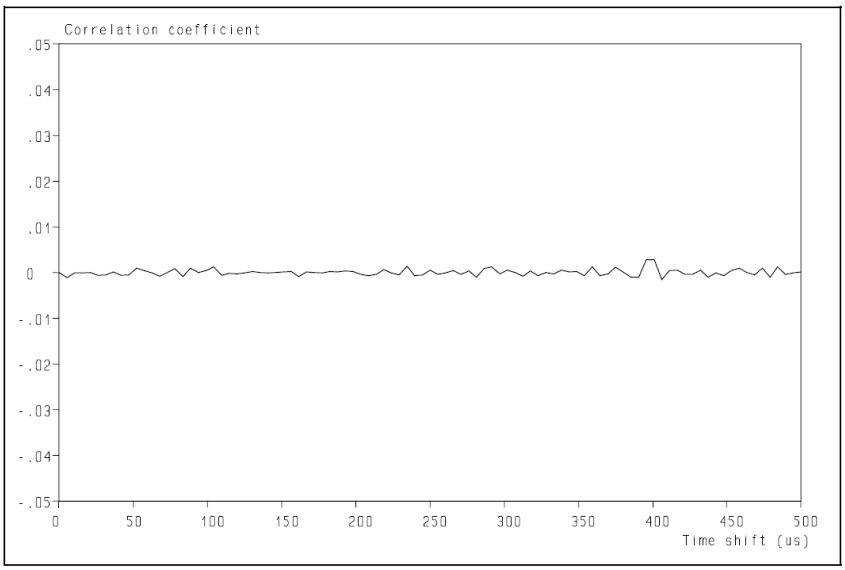


**Figure A-3**: *The autocorrelation function of the noise as used in the Monte Carlo simulation with an extended vertical scale. For details: see text.*

So it can be concluded that the noise in the file, used for the Monte Carlo simulation, also fulfills the second and third requirement. The noise is therefore suited for the simulation and the conclusions from the simulation can be qualified as trustworthy.

**References**


1. Helen M. Jackson, Michael D. Capp, and J. Robert Stuart, "The audibility of typical digital audio filters in a high-fidelity playback system", Audio Engineering Society, 137th Convention, Los Angeles, (2014), Paper 9174
2. "Perception of Temporal Response and Resolution in Time Domain", Workshop #3, AES convention Berlin (Germany), 2017
3. https://earthworksaudio.com/products/microphones/measurement-series/m50/
4. Alex Khenkin, "How Earthworks Measures Microphones", http://recordinghacks.com/pdf/earthworks/how-earthworks-measures-mics.pdf
5. James Boyk , "There's Life Above 20 Kilohertz! A Survey of Musical Instrument Spectra to 102.4 KHz", California Institute of Technology, Music Lab, 0-51 Caltech, Pasadena, CA 91125, USA, Copyright © 1992, 1997 James Boyk.
6. J.S. Bendat and A.G. Piersol, "Random Data, Analysis and Measurement Procedures", John Wiley & Sons, New York (1986)
7. A. Papoulis, "The Fourier Integral and its Applications", McGraw-Hill Book Company, New York (1962)
8. A. Papoulis, "Signal Analysis, McGraw-Hill Book Company, New York (1984)
9. H.R.E. van Maanen, "Retrieval of Turbulence and Turbulence Properties from randomly sampled Laser-Doppler Anemometry data with noise", (chapter 2), Ph.D. Thesis, Delft University of Technology (Delft, Netherlands), 1999
10. https://www.bksv.com/en/transducers/acoustic/microphones/microphone-cartridges/4939