7 - 12 SEPTEMBER 2014, KRAKÓW
FORUM ACUSTICUM

Avoiding spectral leakage in measurements of distortion-product otoacoustic emissions

Anders T. Christensen, Rodrigo Ordoñez and Dorte Hammershøi
Section for Acoustics, Department of Electronic Systems, Aalborg University, Denmark.

Summary
The typical measurement of distortion-product otoacoustic emissions (DPOAEs) relies on the discrete Fourier transformation (DFT) to extract amplitude and phase of the response at the $2f_1 - f_2$ DPOAE frequency, f_1 and f_2 being the frequencies of the stimulus. The DFT measures the DPOAE response accurately, not when the DPOAE frequency is exactly at a DFT bin, but when the DPOAE frequency and both stimulus frequencies are at DFT bins. The stimulus frequencies are typically varied so that the ratio f_2/f_1 is constant. This strategy is not readily compatible with the fixed spacing between frequency bins of the DFT. This study shows that as the distortion frequency $2f_1 - f_2$ decreases the error at $2f_1 - f_2$ due to spectral leakage from f_1 and f_2 increases to more than 10 dB depending on relative levels. Controlling for this error by selecting 'good' ratios therefore allows DPOAE measurements at low frequencies. Eighteen normal-hearing human subjects had DPOAE levels measured as a function of f_2/f_1. $2f_1 - f_2$ was kept around 246 and 1231 Hz. The ratio giving the maximum DPOAE level increases as the distortion-frequency decreases. Contrary to the expectation one might get from the literature, the 246-Hz DPOAEs are not smaller than the 1231-Hz DPOAEs.

PACS no. 43.64.Jb, 43.64.Yp

1. Introduction

The discrete Fourier transformation (DFT) [13] can be used as an alternative to analog or digital filtering to extract the distortion-product otoacoustic emission (DPOAE) from ear-canal responses to two stimulus tones [16]. The transformation is performed on a windowed time signal which must be band-limited, to avoid artifacts of anti-aliasing, and stationary for the resulting frequency bins to unambiguously reflect the calibrated frequency content. A third source of error is introduced when the analyzed time signal is not entirely periodic. Tonal content between the available, evenly-spaced frequency bins is represented by the DFT with the bins closest to the off-bin content. The error introduced at all those closest neighbor bins is 'spectral leakage', also known as Gibb's phenomenon.

In the context of DPOAEs, spectral leakage might contaminate the amplitude and phase data at the distortion-frequency bins. To avoid it, both stimulus frequencies, f_1 and f_2, and the response frequency of interest, here $2f_1 - f_2$, must be exactly at frequency bins available in the DFT, or equivalently, tones of those three frequencies must complete exactly an integer number of cycles within the analysis window.

If the integer requirement is not met, the resulting spectral leakage can be reduced by windowing of the recorded time signal [4]. Contrary however to windows which have lower sidelobes, the rectangular window (effectively no window, applied in this paper) has the unique property of being able to provide an unbiased estimate of the energy at a given DFT bin.

DFT-based DPOAE measurement systems typically round the specified stimulus frequencies to the nearest DFT bins and adjust the distortion frequencies accordingly [8, 15, 17]. The IEC60645-6 standard [6] for instrumentation of OAE measurements allows the stimulus frequencies to deviate ±1% from their nominal values. This in turn allows slight variation in the ratio f_2/f_1 across frequencies even when specified as fixed. The symmetry of the ratio error depends on the type of rounding applied. The error is larger at lower frequencies and for lower DFT resolutions. The influence of a given error is individually determined by the response gradient around the measurement point.

The present paper outlines consequences of spectral leakage on simulated DPOAEs and describes the selection of ratios without rounding of stimulus frequencies, similar to the selection briefly indicated in [11] and [12]. These ratios are readily compatible with the DFT resolution and can be perfectly fixed across frequency. Compatible ratios were applied to measurements of DPOAE ratio-magnitude functions at low frequencies in 18 normal-hearing human subjects.

---

2. Methods

When the DPOAE is measured across frequencies to probe the state of hearing at corresponding places in the cochlea, the "primary ratio" f_2/f_1 between the stimulus frequencies is kept fixed. This paradigm is known as the fixed-ratio paradigm. Alternatively, several ratios are measured at each frequency, usually by fixing f_2. This paradigm is known as the fixed-f_2 paradigm. Below we describe, for both paradigms, how to select ratios which can measured across frequency without contamination from spectral leakage.

Common to both paradigms, an N-point DFT has the center frequencies of its bins defined by the sampling frequency f_s. The DFT resolution is defined as the frequency spacing f_s/N between adjacent DFT-bins. It is "high" when the frequency spacing is small. The N bins are indexed by the integer variable k, so that the kth bin has its center frequency at $k \cdot f_s/N$.

To make sure that the distortion frequency and both stimulus frequencies of the DPOAE measurement end up at DFT bins, we specify those, as well as the ratio between the stimulus frequencies, in terms of the index variable k. The stimulus-frequencies are denoted k_1 and k_2, $k_1 < k_2$, and the distortion frequency of interest $k_{dp} = 2k_1 - k_2$.

2.1. Fixed-ratio measurements

The ratio between stimulus frequencies, the primary ratio, is specified as the fraction k_2/k_1. The reduced version of that fraction yields for the given ratio the smallest k_{dp} at which a DPOAE measurement can be made without spectral leakage. That smallest k_{dp} yields in turn the highest DPOAE resolution possible. The next k_{dp}'s, fulfilling the requirements to avoid spectral leakage, are multiples of the smallest.

A typical primary ratio is 1.22, which can readily be expressed as $k_2/k_1 = 122/100$, or reduced, as 61/50. The smallest k_1 and k_2 are 50 and 61, and the smallest k_{dp} then 39. This means that measurements can be made without spectral leakage at every 39 bin in the DFT used. The resulting DPOAE resolution when e.g. using a 4096-point DFT and a sampling frequency of 48 kHz is therefore $39 \cdot 48000/4096 = 457$ Hz, although the DFT resolution is 11.7 Hz.

A primary ratio of 1.222 (truncated) can be expressed as 11/9, leading to a DPOAE resolution of 7 instead of 39 DFT bins. Ratios of 6/5 (1.20) and 5/4 (1.25) give DPOAE resolutions of just 4 and 3 DFT bins, respectively. Only for a primary ratio of 3/2 (1.50) is there a one-to-one correspondence between DPOAE resolution and DFT resolution.

2.2. Fixed-f_2 measurements

For the fixed-f_2 paradigm, ratios which have k_2 in common need to be identified. The smallest DFT bin which two smallest k_2's have in common is the product of the two.

Table I. Smallest k's for 16 'good' ratios (truncated).

k_2/k_1	k_{dp}	k_1	k_2
1.100	9	10	11
1.111	8	9	10
1.125	7	8	9
1.143	6	7	8
1.167	5	6	7
1.200	4	5	6
1.222	7	9	11
1.250	3	4	5
1.286	5	7	9
1.300	7	10	13
1.333	2	3	4
1.375	5	8	11
1.400	3	5	7
1.429	4	7	10
1.444	5	9	13
1.500	1	2	3

For instance, the ratios 6/5 and 5/4 have smallest k_2's of 6 and 5. The smallest k_2 they have in common is 30. Fixed-f_2 measurements can then be made with those two ratios without spectral leakage by fixing k_2 at the 30th DFT bin. In that case, the distortion product for the 6/5-measurement is at the 20th DFT bin and at the 18th DFT bin for the 5/4-measurement. Without further ado, 3/2-measurements could also be made at the 10th DFT bin because 30 is readily divisible by its k_2 of 3.

Now, this series of measurements at three ratios can be made with k_2's fixed at every 30 DFT bin. Even though it takes on a slightly different meaning than in the fixed-ratio paradigm, we may say that the DPOAE resolution of these fixed-f_2 measurements is 30 DFT bins. Inclusion of 11/9-measurements dramatically gives a resolution of 330 DFT bins.

This guideline for fixed-f_2 measurements is principally the same for fixed-f_1, fixed-f_{dp} and fixed-$\sqrt{f_1 f_2}$ measurements.

2.3. Selection of DPOAE ratios in general

Table I contains k's for 16 ratios between 1.100 and 1.500. They are 'good' in the sense that they allow for relatively high-resolution DPOAE measurements with both the fixed-ratio and the fixed-f_2 paradigms.

Several ratios can be measured if the smallest k's to be fixed in each are divisible. For instance, fixed-k_{dp} can be measured at the 8th DFT bin, that is, with a relatively high DPOAE resolution for all ratios having smallest k_{dp}'s at 1, 2, 4 and 8. Similarly, at the 9th for those at 1, 3 and 9. For the 7th, only that at 1 and the two at 7 can be picked. Otherwise the highest possible DPOAE resolution decreases dramatically, because k_{dp}'s need to be multiplied to yield the smallest DFT bin they have in common.

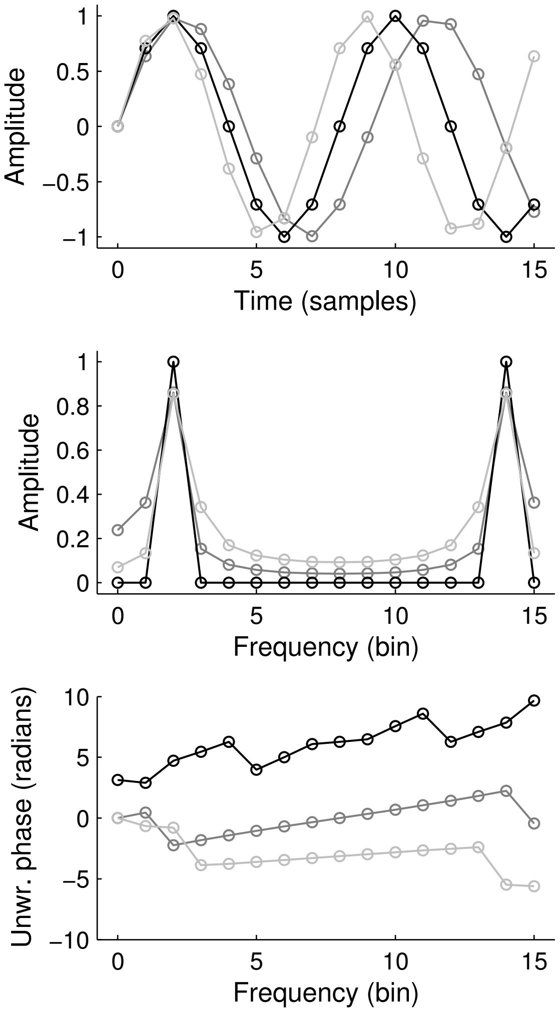

Figure 1. Time and frequency domain of three tones completing 2 (black), 1.75 and 2.25 (gray) cycles within the analysis window of 16 samples.

2.4. Simulation of spectral leakage

Figure 1 demonstrates the spectral consequences of tones which do not complete an integer number of cycles within the analysis window.

Spectral leakage is apparent for the two tones which complete close to, but not exactly, two cycles within the analysis window. In addition to the tone at the second bin, the DFT uses those at neighboring bins to make up the off-bin content. Even though the amplitude is actually 1 for all tones plotted, that amplitude is not readily reflected in any of the bins, except when the tones being analyzed complete an integer number of cycles within the analysis window. In the simulation, the phase of the spectrum for the tone completing exactly two cycles actually accumulates no phase across the spectrum, but numerical inaccuracy for values close to zero makes it look like it does. The off-bin tones do accumulate a bias phase across the spectrum.

A windowed DPOAE recording contains multiple tones and, as already described, the energy of one may leak into the bins of the others, if they are not all placed exactly at a DFT bin. Figures 2, 3 and 4 demonstrate such leakage for the fixed-ratio and fixed-f_2 paradigms, respectively. The relative levels are similar to those measured in the ear.

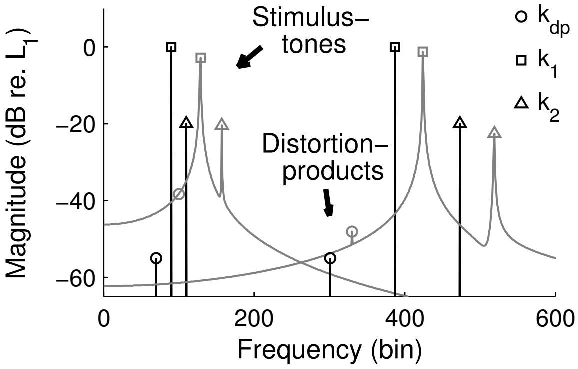

Figure 2. Spectrum of four simulated fixed-ratio DPOAE recordings ($N = 4096$), where the k_{dp}-, k_1- and k_2-levels are -55, 0 and -20 dB re. full-scale. The DPOAE ratio is fixed at 11/9 (1.222) and the smallest k_{dp} is then 7. Black curves are recordings where k_{dp} is a multiple of 7. Gray curves do not have k_{dp} at a multiple of 7.

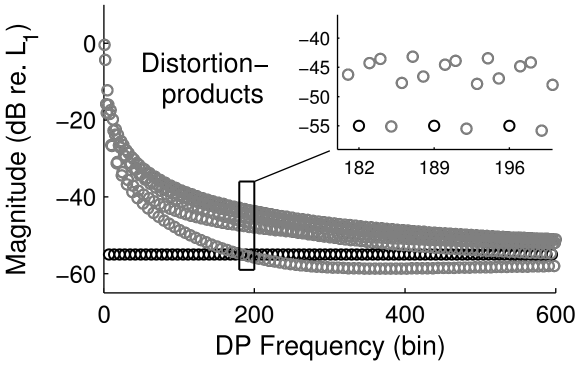

Figure 3. Same simulation as in Figure 2 for stimulus-frequency pairs giving a distortion-product frequency at each DFT bin. Only every 7th DFT bin (black) reflects the correct distortion level at -55 dB re. full-scale.

The higher the amplitude of the leaking tone, the greater the effect on neighboring bins. Tones that are closer in frequency affect each other more, because spectral leakage rolls off with bin distance to the leaking tone. Thus, DPOAE measurements with shorter analysis windows, lower sampling frequencies, smaller primary ratios and larger stimulus levels relative to the response levels are more prone to spectral leakage. Fixed-ratio DPOAE measurements towards lower frequencies are also particularly prone, because the stimulus-frequency bins are closer to the response-frequency bins.

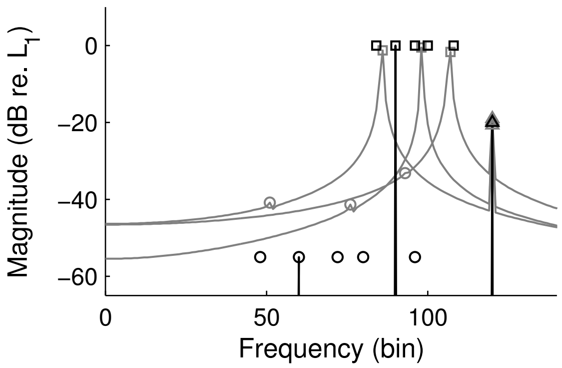

Figure 4. Spectrum of eight simulated fixed-k_2 DPOAE recordings ($N = 4096$), where the k_{dp}-, k_1- and k_2-levels are -55, 0 and -20 dB re. full-scale. k_2 is fixed at the 120th DFT-bin. Black curves are recordings with DPOAE ratios 10/9, 6/5, 5/4, 4/3 and 10/7 (full curve only shown for 4/3). For all these ratio 120 is divisible by the smallest k_2. For gray curves the DPOAE-ratios are 11/9, 9/8 and 7/5. These smallest k_2's do not divide into 120 without a remainder.

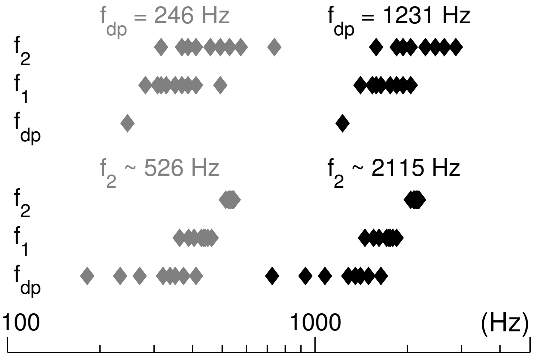

Figure 5. Overview of the four conditions where measurements were made.

2.5. Experiment

DPOAE ratio- and level-dependence have only been established for mid- and high-frequency DPOAEs, e.g. [5, 3, 2, 7]. 'Good' ratios selected as described above were applied to measure ratio- and level-dependencies of low- and mid-frequency DPOAEs in 18 clinically normal-hearing human subjects.

With sampling frequency 48 kHz and N 32768, DPOAEs were measured with f_{dp} fixed at 246 Hz, f_{dp} fixed at 1231 Hz, f_2 fixed at 513-545 Hz and f_2 fixed at 2051-2120 Hz. Figure 5 gives an overview of the frequencies in these configurations.

In the fixed-f_{dp} conditions the following ratios were included: 1.125 (9/8), 1.200 (6/5), 1.222 (11/9), 1.250 (5/4), 13/10 (1.300), 1.333 (4/3), 1.364 (15/11), 1.400 (7/5) and 1.500 (3/2). These have smallest k_{dp}'s 7, 4, 7, 3, 7, 2, 7, 3 and 1, respectively. The first DFT bin they have in common is $1 \cdot 2 \cdot 3 \cdot 4 \cdot 7 = 168$ (246 Hz) which thus constitutes a low-frequency fixed-f_{dp} condition. The fifth multiple of this common k_{dp} (1231 Hz) constitutes a mid-frequency fixed-f_{dp} condition.

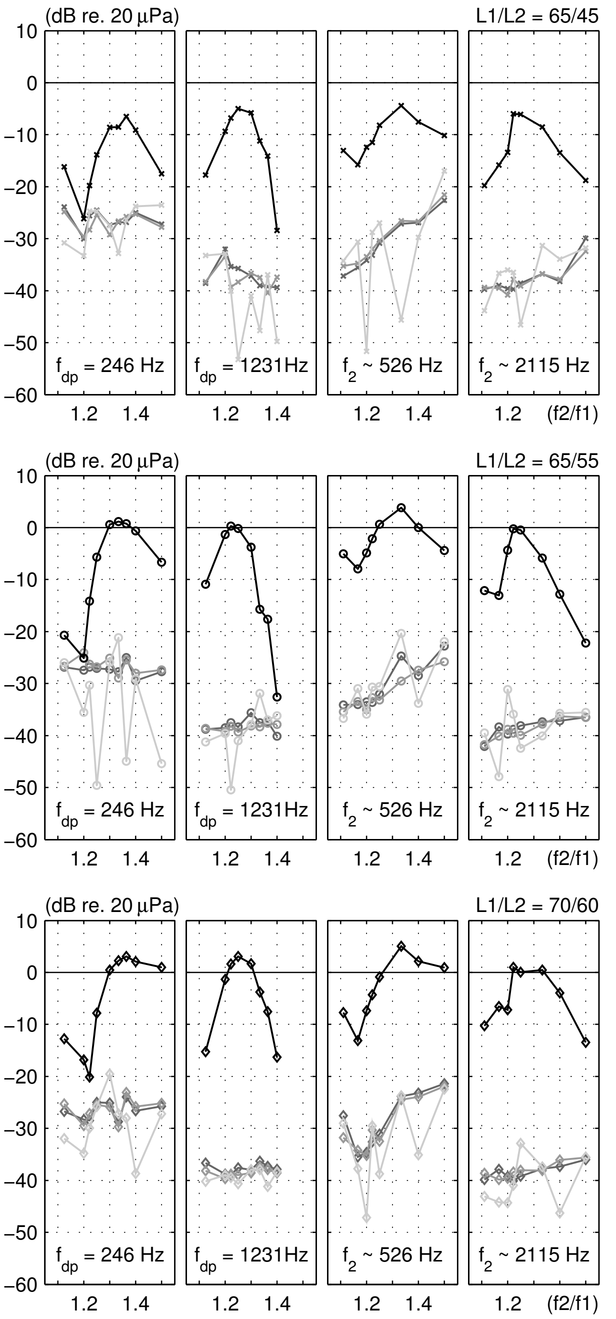

Figure 6. DPOAE ratio-magnitude measurements from subject 2. Stimulus levels L_1 and L_2 are specified in dB SPL. Gray curves are noise floors calculated in three different ways. Dark gray as the average amplitude of DFT bins 5 to 16 to each side of the $2f_1 - f_2$ bin. Medium gray the same but from the average difference between repeated measurements where the response has thus been cancelled out. Light gray is the level at the $2f_1 - f_2$ bin in the difference signal.

In the fixed-f_2 conditions the following ratios were included: 1.111 (10/9), 1.667 (7/6), 1.200 (6/5), 1.222 (11/9), 1.250 (5/4), 1.333 (4/3), 1.400 (7/5) and 1.500 (3/2). To allow the measurement of all at a low frequency, k_2 changes slightly between 513 and 545 Hz. In the mid-frequencies between 2051 and 2120 Hz.

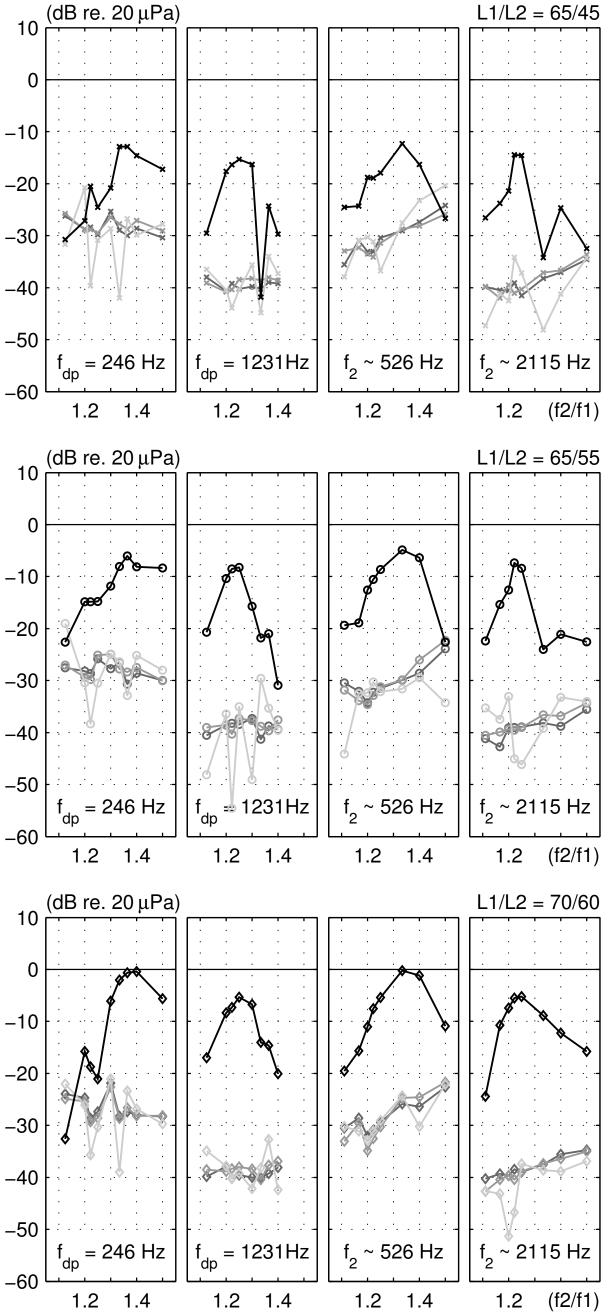

Figure 7. DPOAE ratio-magnitude measurements from subject 10. See caption of Figure 6.

3. Results

Exemplary ratio-magnitude results from two different subjects are shown in Figures 6 and 7.

The mid-frequency DPOAE level is maximum between primary ratios of 1.20 and 1.25. The ratio-magnitude responses increase and broaden as the stimulus levels increase. These two results are in agreement with existing literature.

As has been demonstrated for mid- and high-frequency DPOAEs, e.g. [5], mid- and low-frequency DPOAEs verify that the primary ratio, giving the highest DPOAE level, increases as the DPOAE-frequency decreases. Low-frequency DPOAEs have maxima at ratios above 1.30.

Low-frequency DPOAEs had the same prevalence as mid-frequency DPOAEs. The noise floor increases as the DPOAE frequency decreases but maximum response levels were the same for low- and mid-frequency results.

4. Discussion

DPOAEs were measurable at low frequencies, i.e., f_{dp} fixed and varied around 246 Hz. This was enabled by an approximate doubling of the averaging times for low-frequency with respect to mid-frequency measurements, and by the selection of 'good' ratios as described in the method section.

If the DPOAE resolution measured in DFT bins is not fine enough, the number of points in the DFT can be increased. The time signal being analyzed must however remain stationary within the analysis window. Alternatively, the sampling frequency can be changed to meet certain requirements for the DPOAE resolution.

5. CONCLUSIONS

The DFT can be used to extract the magnitude and phase of the DPOAE from ear-canal responses to two stimulus tones. Results at the DPOAE bin of the DFT are, however, affected by spectral leakage from the stimulus tones, if the distortion frequency and both stimulus frequencies are not exactly at bins available in the DFT. Expressing the DPOAE ratio as a reduced fraction, instead of a decimal number, allows the identification of what referred to as smallest k's. These k's reflect in terms of DFT bins the highest resolution with which DPOAE measurements can be made without spectral leakage.

References

[1] S. Dhar, C. L. Talmadge, G. R. Long, A. Tubis (2002). Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. J Acoust Soc Am, 112(6), 2882-2897.

[2] L. E. Dreisbach, J. H. Siegel (2001). Distortion-product otoacoustic emissions measured at high frequencies in humans. J Acoust Soc Am, 110(5), 2456-2469.

[3] S. A. Gaskill, A. M. Brown (1990). The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity. J Acoust Soc Am, 88(2), 821-839.

[4] F. J. Harris (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. Proceedings of the IEEE, 66(1), 51-83.

[5] F. P. Harris, B. L. Lonsbury-Martin, B. B. Stagner, A. C. Coats, G. K. Martin (1989). Acoustic distortion products in humans: Systematic changes in amplitude as a function of f2/f1 ratio. J Acoust Soc Am, 85(1), 220-229.

[6] IEC60645-6. Electroacoustics – Audiometric equipment – part 6: Instruments for the measurement of otoacoustic emissions, 2010. International Electrotechnical Commission.

[7] T. A. Johnson, S. T. Neely, C. A. Garner, M. P. Gorga (2006). Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions. J Acoust Soc Am, 119(1), 418-428.

[8] M. C. Liberman, S. Puria, J. J. Guinan Jr (1996). The ipsilaterally evoked olivocochlear reflex causes rapid adaptation of the $2f_1f_2$ distortion product otoacoustic emission. J Acoust Soc Am, 99(6), 3572-3584.

[9] G. R. Long, C. L. Talmadge (1997). Spontaneous otoacoustic emission frequency is modulated by heartbeat. J Acoust Soc Am, 102(5), 2831-2848.

[10] G. R. Long, C. L. Talmadge, J. Lee (2008). Measuring distortion product otoacoustic emissions using continuously sweeping primaries. J Acoust Soc Am, 124(3), 1613-1626.

[11] M. Mauermann, B. Kollmeier (2004). Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. The J Acoust Soc Am, 116(4), 2199-2212.

[12] S. W. Meenderink, P. M. Narins, P. van Dijk. (2005). Detailed f_1, f_2 area study of distortion product otoacoustic emissions in the frog. J Assoc Res Oto, 6(1), 37-47.

[13] A. V. Oppenheim, R. W. Schafer: Digital Signal Processing. Prentice Hall, 2nd edition, 1999.

[14] Robinette, M. S., Glattke, T. J. (Eds.): Otoacoustic emissions: clinical application. Thieme, 2007.

[15] C. A. Shera, C. L. Talmadge, A. Tubis (2000). Interrelations among distortion-product phase-gradient delays: their connection to scaling symmetry and its breaking. J Acoust Soc Am, 108(6), 2933-2948.

[16] M. L. Whitehead, B. B. Stagner, B. L. Lonsbury-Martin, G. K. Martin (1994). Measurement of otoacoustic emissions for hearing assessment. Engineering in Medicine and Biology Magazine, IEEE, 13(2), 210-226.

[17] R. H. Withnell, L. A. Shaffer, C. L. Talmadge (2003). Generation of DPOAEs in the guinea pig. Hearing research, 178(1), 106-117.