

# Harmonic Distortion Measurement using Spectral Warping

Donald Bailey

*Institute of Information Sciences and Technology  
Massey University, Palmerston North, New Zealand  
D.G.Bailey@massey.ac.nz*

## Abstract

*Harmonic distortion may be characterised by the proportion of energy of a sinusoidal signal transferred to the harmonics. Differential time scaling resulting from the spectral warping transform allows the fundamental and harmonics to be separated, and thus measured separately. Two spectral warping transforms for distortion measurement are compared: the standard all-pass mapping, and a piecewise linear mapping. Both are shown to be effective at measuring distortion, although the piecewise linear mapping is computationally less expensive.*

## 1. Introduction

Many electronic systems are modelled as linear time-invariant systems. A system is linear if, when two signals are added on the input, the output is the sum of the corresponding outputs produced by each of the inputs on their own. A time-invariant system is one in which the response is independent of when the input is replied. Delaying the input signal will give a delayed response.

While no practical systems are truly linear or time-invariant, this is a useful approximation, and many analogue electronic circuits (amplifiers, filters) are designed to be both linear and time-invariant over some range of inputs. A key property of such systems is that a sinusoid input will produce a sinusoid output at exactly the same frequency, with only the amplitude and phase affected. Linear systems theory enables any arbitrary signal to be decomposed into the sum of sinusoids, by using a Fourier transform.

Any non-linearity or time variance in a system will result in introduction of harmonics when a single sinusoid is input, and frequency mixing (sum and difference terms) when multiple frequencies are present. One way of measuring the distortion introduced by a 'linear' circuit is therefore to provide a sinusoid to the input and measure the harmonic distortion by measuring the proportion of signal power in the harmonics relative to that in the fundamental. If

$A_i$  is the amplitude of the  $i$ th harmonic, then the total harmonic distortion is defined as

$$D = \frac{\sum_{i>1} A_i^2}{A_1^2} \quad (1)$$

Factors of electronic circuits that may result in distortion include:

- clipping or limiting of high amplitude outputs due to power supply constraints;
- cross-over distortion in class AB or B amplifiers resulting from mismatched output transistors or incorrect biasing;
- nonlinearity of the amplifier gain with signal amplitude;
- slew rate limitations.

Digital signal processing (DSP)-based testing involves employing digital tools and methods to test both digital and analogue components of the device under test. In the basic arrangement of DSP-based testing both signal generation and output measurement are realised by means of pure digital circuitry. For distortion measurement, the test signal is usually a digitised sinusoid and the response is analysed using a discrete Fourier transform (DFT) (normally computed using a fast Fourier transform (FFT)), or filtering. This paper describes how the filtering may be performed by using spectral warping.

### 1.1. Spectral warping

The spectral warping transform (SWT) is a time domain to time domain transformation on a signal that effectively warps the frequency content of the original signal [1-4]. Mathematically, it is the transformation of a discrete time signal such that the samples of its Fourier transform correspond to unevenly spaced samples of the Fourier transform of the original signal.

In the  $z$ -domain, the DFT of the original signal corresponds to sampling its  $z$  transform at equally spaced locations around the unit circle. The DFT of the warped signal corresponds to non-uniformly spaced samples around the unit circle. The transformation therefore is one that maps the unit circle in the  $z$ -

domain onto itself. The simplest such transformation is a first order all-pass mapping:

$$\hat{z} = \frac{z-a}{1-az} \quad -1 < a < 1 \quad (2)$$

where  $a$  is the warping factor. Consider a windowed input signal,  $f$ , with  $N$  samples. After warping, the output signal,  $g$ , has  $M$  samples. The inverse frequency mapping is used to determine the locations of  $M$  non-uniformly spaced points around the unit circle. For the mapping of equation (2), these are located at [5]:

$$\omega_m = \tan^{-1} \left[ \frac{(1-a^2) \sin \frac{2\pi m}{M}}{(1+a^2) \cos \frac{2\pi m}{M} + 2a} \right] \quad (3)$$

The z-transform of the input signal is evaluated at these  $M$  frequencies

$$\begin{bmatrix} G[0] \\ G[1] \\ \vdots \\ G[M-1] \end{bmatrix} = \begin{bmatrix} 1 & e^{-j\omega_0} & \dots & e^{-(N-1)j\omega_0} \\ 1 & e^{-j\omega_1} & \dots & e^{-(N-1)j\omega_1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\omega_{M-1}} & \dots & e^{-(N-1)j\omega_{M-1}} \end{bmatrix} \begin{bmatrix} f[0] \\ f[1] \\ \vdots \\ f[N-1] \end{bmatrix} \quad (4)$$

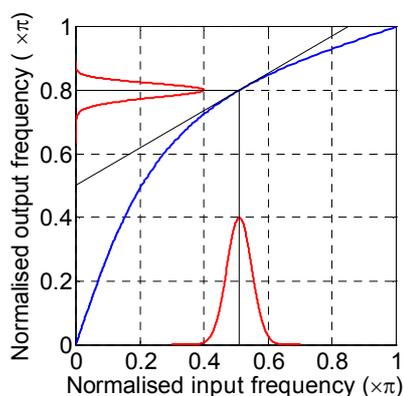
or  $\mathbf{G} = \mathbf{H}_{M,N} \mathbf{f}$ , where  $\mathbf{f}$  are the  $N$  samples arranged as a vector,  $\mathbf{H}_{M,N}$  is the  $M \times N$  z-transform matrix and  $\mathbf{G}$  are the samples in the z-domain at the frequencies indicated by equation (3).

These  $M$  samples are then redistributed uniformly around the unit circle, so the inverse z-transform corresponds to an inverse DFT. Therefore the complete SWT may be represented in matrix form as [5]

$$\mathbf{g} = \mathbf{D}_M^{-1} \mathbf{G} = \mathbf{D}_M^{-1} \mathbf{H}_{M,N} \mathbf{f} \quad (5)$$

where  $\mathbf{D}_M$  is an  $M$ -point DFT matrix.

In the context of distortion measurement, one of the key properties of the SWT is the time distortion that is associated with the frequency transformation [4]. The spectrum of a windowed narrow-band signal will take on the shape of the spectrum of the windowing function. When the signal is warped, the spectrum will be warped according to the slope of the transformation. This process is illustrated in figure 1.



**Figure 1: Frequency warping warps the bandwidth of the output signal depending on the slope. Shown here is  $a = 0.5$ .**

The reciprocal relationship between the time and frequency domains means that if the bandwidth of a signal is reduced, the duration of the envelope of the signal will increase, and visa versa. The duration of the output signal will be scaled by the reciprocal of the frequency transformation. Therefore, for an all-pass transform, if  $\omega$  is the input frequency, then the time scaling is given by:

$$s(\omega) = \frac{1-2a \cos \omega + a^2}{1-a^2} \quad (6)$$

If the signal being transformed is a distorted sinusoid, then each harmonic will undergo a different time scaling. The principle behind using the SWT to perform harmonic distortion measurement described in this paper is to use this differential time scaling to separate the fundamental from the harmonics to enable separate measurement.

Section 2 of this paper looks in more detail at the design considerations for using the SWT for measuring the distortion of the waveform. First the all-pass transform and then a custom piecewise linear transform are considered. Section 3 describes a FIR based implementation of the SWT for distortion measurement, and presents some results from measuring the harmonic distortion of a number of waveforms. Section 4 discusses the effectiveness of the SWT for this application.

## 2. Design

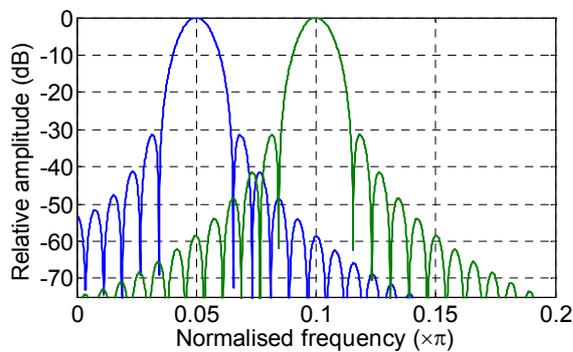
In measuring the distortion in a waveform using digital signal processing, it is important to avoid artefacts introduced by aliasing. The Nyquist sampling criterion requires that the sample frequency be at least twice the maximum frequency of the input signal to avoid aliasing. This poses a problem when measuring distortion because distortion will introduce an unknown number of harmonics. This necessitates that the sample frequency is significantly higher than the input sine wave frequency to allow adequate harmonics to be captured. If the sample frequency is too low, the harmonics around the sample frequency plus and minus the signal frequency will be aliased onto the signal frequency. The consequence is that the energy in those components is indistinguishable from that within the fundamental. This effectively moves those aliased harmonics from the numerator of equation (1) to the denominator, with the result that the total harmonic distortion is underestimated. If necessary, aliasing may be reduced through the use of an appropriate filter prior to sampling, again with the result that the total harmonic distortion is underestimated.

The actual sample frequency will therefore depend on the expected frequency content of the distorted

signal. In the examples provided here, the sample frequency was set to 40 times the sinusoid frequency, giving a normalised input frequency of  $\omega = 0.05\pi$ .

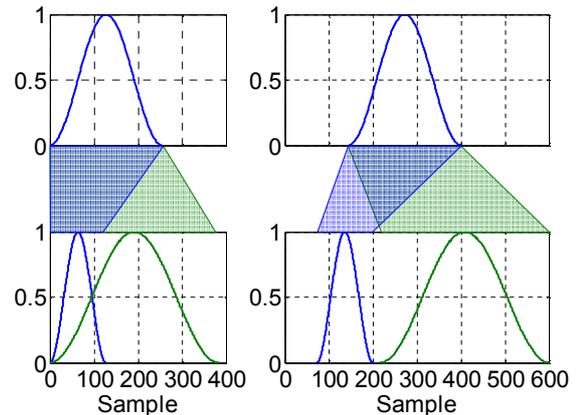
A second consideration is that the SWT requires a finite sequence to be transformed. This may be accomplished by windowing the input signal. A simple rectangular window will result in significant spectral leakage, making it difficult to determine how much energy was in the fundamental and how much in the harmonics. An appropriate window function will reduce the side-lobes of the signal (and the spectral leakage) at the expense of a broader main lobe. The window needs to be sufficiently long to be able to capture the important details of the waveform. If the window is too long, more computation is involved; if too short the main lobe of the fundamental will begin overlapping with that of the second harmonic, making it impossible to distinguish between them.

A Hanning window is used in the examples here because the first side-lobe is -31.5 dB below the central peak, and the side-lobes continue to decrease with increasing distance from the main lobe (see figure 2). A window of width 256 samples has a main lobe half-width of  $0.015\pi$  and is below -40 dB by the time it crosses the sidelobes of the adjacent harmonics. This was considered a reasonable compromise between accuracy and computational expense. A 256 sample window will capture about 6.4 cycles of the distorted waveform



**Figure 2: Relative power within the windowed fundamental and second harmonic.**

It is also necessary to pad the data before processing. The reason for this is shown in figure 3. Without padding, no level of differential scaling will completely separate the fundamental and harmonics. However, if the data are padded at the start with zeroes, these zeroes will also be scaled, allowing space for the fundamental to be separated from the harmonics. The width of padding required will depend on the differential scaling between the fundamental and second harmonic.



**Figure 3: With no padding (left) the fundamental and second harmonic always overlap after the SWT. Pre-padding with zeroes (right) provides space before the second harmonic for the fundamental.**

## 2.1. All-pass mapping

With the all-pass mapping (figure 1), the time scaling is given by equation (6). Let the duration of the signal be  $T$  samples pre-padded with  $T_0$  zeroes. That is

$$N = T + T_0 \quad (7)$$

Separation of two frequencies,  $\omega_1$  and  $\omega_2$  requires

$$(T + T_0)s(\omega_1) < T_0s(\omega_2) \quad (8)$$

Substituting (6) into (8) and rearranging gives

$$\frac{T_0}{T} > \frac{1 + a^2 - 2a \cos \omega_1}{2a(\cos \omega_1 - \cos \omega_2)} \quad (9)$$

To include all of the samples in the output (as required to prevent aliasing during the SWT [5]), the output length must be

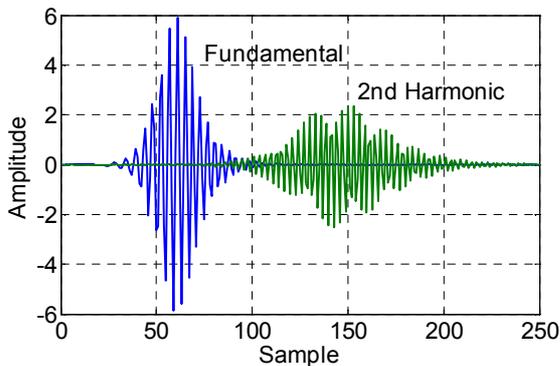
$$M > (T + T_0) \frac{1 - 2a \cos \pi + a^2}{1 - a^2} = (T + T_0) \frac{(1 + a)^2}{1 - a^2} \quad (10)$$

If the objective is to minimise the computation, then the inequality of equations (9) and (10) can be replaced with equalities. Since the SWT involves an inverse DFT, relaxing the constraint on  $M$  to the next largest power of 2 allows an FFT to be used. This relaxes the constraint on  $T_0$  and allows a margin to be included between the warped fundamental and second harmonic.

Substituting  $\omega_1 = 0.05\pi$  for the fundamental and  $\omega_2 = 0.1\pi$  for the second harmonic suggests that  $a = 0.85$ ,  $T_0 = T$  and  $M = 32T$  satisfy the constraints. The result of this combination is shown in figure 4. Two observations may be made.

First, there is still considerable overlap between the envelopes of the fundamental and second harmonic. This results from dispersion of the waveform as a result of windowing. Windowing will broaden the frequency content of the input signal because the

Fourier transform of the window function has finite width. The slope of the frequency mapping varies with frequency, therefore the different frequencies associated with a peak are subject to different time scaling. The lower frequencies are compressed slightly more than the higher frequencies, resulting in a significant broadening of the envelope after the SWT.



**Figure 4: Separation of the fundamental and 2nd harmonic with  $a = 0.85$ ,  $T_0 = T$  and  $M = 32T$ . The input signal consists of windowed sine waves with amplitude 1.**

The second observation is that as a result of time scaling the amplitude is also scaled. As a signal is compressed in time, the amplitude is increased in inverse proportion to the scale factor. This scaling will affect the relative power of the fundamental and harmonic. These results are compared in table 1. The power scaling is not the same as the amplitude scaling because of dispersion. The different frequencies within the spread peak are subject to different scale factors, and the ratio reflects the average scaling.

**Table 1: Amplitude and power scaling of waveforms in figure 4.**

	Fundamental	2 <sup>nd</sup> harmonic
Input power	48.2	48.2
Output power	309.0	126.9
Ratio	6.413	2.634
Amplitude scaling	6.390	2.625

The first dispersion effect may be overcome by allowing for the width of the main frequency lobe as a result of windowing. The tail of the warped fundamental envelope corresponds to the higher frequencies (those with the largest time scale factor). Therefore setting  $\omega_1$  to the position of the first null ( $\omega_1 = 0.065\pi$ ) will account for most of the tail. Similarly the start of the warped second harmonic corresponds to the frequencies with the smallest time scale factor. These are the lower frequencies within the main lobe. Again setting  $\omega_2$  to the first null ( $\omega_2 = 0.085\pi$ ) will account for most of the early tail.

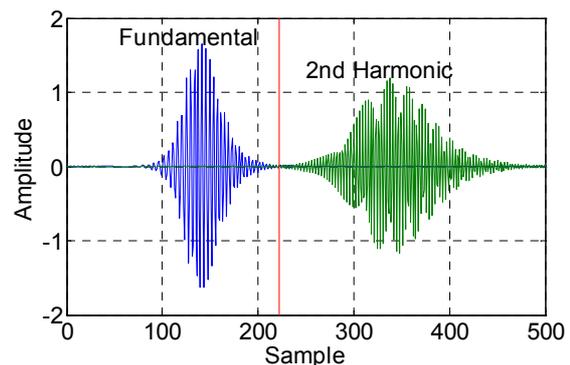
The power scaling may be compensated by weighting each frequency component of  $\mathbf{G}$  before performing the inverse DFT. Since the power scales with the square of the amplitude, the frequency response of the compensating filter needs to be:

$$H_c(\omega) = \sqrt{\frac{1 - 2a \cos \omega + a^2}{1 - a^2}} \quad (11)$$

This filter is able to be combined with the SWT of equation (5) to give a frequency-weighted SWT.

$$\mathbf{g} = \mathbf{D}_M^{-1} \text{diag}(H_c(\omega_m)) \mathbf{H}_{M,N} \mathbf{f} \quad (12)$$

The effect of these two changes is illustrated in figure 5, with the results summarised in table 2. The difference in power between the input and the output reflects rounding errors in performing the computations of the transform. There is still a little overlap which effectively comes from the side-lobes of the windowing function. This is reflected in the leakage measure of table 2, and is at the level expected for the spectral leakage indicated in figure 1. Additional padding with zeroes is unable to reduce this any further. The leakage can only be improved further by increasing the window width,  $T$ .



**Figure 5: Separation of the fundamental and 2nd harmonic with  $a = 0.85$ ,  $T_0 = 3T$  and  $M = 64T$ . The input signal consists of windowed sine waves with amplitude 1.**

**Table 2: Power scaling and leakage of waveforms in figure 5.**

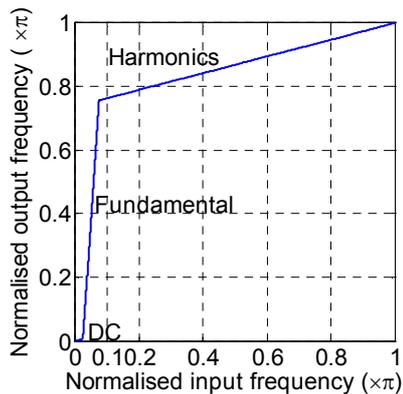
	Fundamental	2 <sup>nd</sup> harmonic
Input power	48.2	48.2
Output power	48.2	48.2
Power difference	-137 dB	-139 dB
Leakage	-40.4 dB	-40.6 dB

## 2.2. Piecewise linear mapping

The major limitation of the all-pass mapping is that the slope changes relatively slowly with frequency. As a consequence, the differential scaling is small,

requiring significant pre-padding of the waveform. This, in turn, increases the length of the output required. A second limitation is the dispersion caused by the differential scaling within the main lobe of the windowed waveform.

Both of these effects may be overcome by designing a piecewise linear mapping. Such a mapping can avoid dispersion from the differential scaling by making the slope constant through the main lobe of each harmonic. The differential scaling can be made arbitrarily large by separately setting the slopes for the fundamental and the harmonics (see figure 6).



**Figure 6: Piecewise linear mapping for separating the fundamental from the harmonics.**

This mapping is represented by:

$$\omega_m = \begin{cases} \frac{7.6m\pi}{M} & 0 \leq m < \frac{1}{304}M \\ \frac{m\pi}{7.5M} + \frac{7\pi}{285} & \frac{1}{304}M < m < \frac{115}{304}M \\ \frac{7.6m\pi}{M} - \frac{14\pi}{5} & \frac{115}{304}M < m < \frac{189}{304}M \\ \frac{m\pi}{7.5M} + \frac{35\pi}{19} & \frac{189}{304}M < m < \frac{303}{304}M \\ \frac{7.6m\pi}{M} - \frac{28\pi}{5} & \frac{303}{304}M < m < M \end{cases} \quad (13)$$

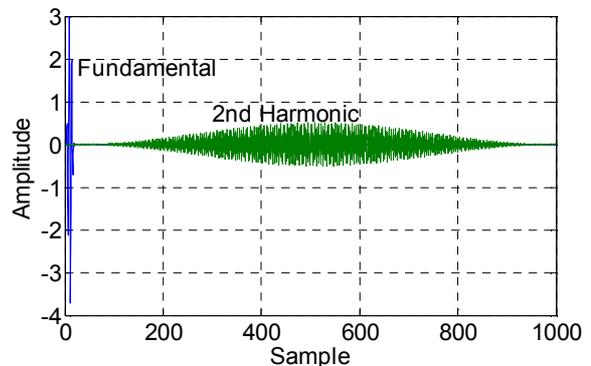
Again, a frequency weighted SWT is used to preserve the power through the transformation. The weights (based on the scaling factors) are:

$$H_c(m) = \begin{cases} \sqrt{3.8} & 0 \leq m < \frac{1}{304}M \\ 1/\sqrt{15} & \frac{1}{304}M < m < \frac{115}{304}M \\ \sqrt{3.8} & \frac{115}{304}M < m < \frac{189}{304}M \\ 1/\sqrt{15} & \frac{189}{304}M < m < \frac{303}{304}M \\ \sqrt{3.8} & \frac{303}{304}M < m < M \end{cases} \quad (14)$$

This is a piecewise constant frequency response, and will not have a particularly nice time response because of the sharp jumps. The results of filtering a fundamental and second harmonic are shown in figure 7, and summarised in table 3.

These results are similar to those for the all-pass mapping, although there is now a smaller difference between the input and output power. The fundamental has been transformed into the first 19 samples, and this

will simplify the calculations in the final implementation.



**Figure 7: Output from piecewise linear SWT for  $T_0 = 7$  and  $M = 2048$ .**

**Table 3: Leakage of waveforms in figure 7.**

	Fundamental	2 <sup>nd</sup> harmonic
Input power	48.2	48.2
Output power	48.2	48.2
Power difference	-57.7 dB	-54.9 dB
Leakage	-41.4 dB	-47.3 dB

### 2.3. Implementation

The matrix form of equation (12) implies a direct FIR implementation [5]. Since a significant fraction of the input sequence is zero as a result of padding (at least for the all-pass mapping), these columns of the  $\mathbf{H}$  matrix do not need to be calculated. After frequency weighting, a DFT may be taken of each of the columns  $\mathbf{H}$  to give the complete transform.

Since the power is conserved in the frequency weighted SWT, it is not necessary to actually compute the power in the harmonics. The power in the harmonics is the difference between the power in the input waveform, and the power in the fundamental component after filtering with the SWT. This means that only the output samples representing the fundamental need to be retained. For the all-pass mapping, the matrix resultant matrix is 162x256, and for the piecewise linear mapping the matrix is 19x256.

The reduction in size of the matrix does not mean that the  $M$  can be reduced. It must be kept at the value described earlier to prevent aliasing in the time domain when performing the inverse DFT [5].

After windowing the input signal, the total power in the signal may be calculated. The input signal is then multiplied by the transform matrix to filter the fundamental component. The power in the fundamental may then be calculated. The harmonic distortion is then given by

$$THD = \frac{\mathbf{f}^T \mathbf{f} - \mathbf{g}^T \mathbf{g}}{\mathbf{g}^T \mathbf{g}} \quad (15)$$

where the superscript  $T$  represents the vector transpose.

### 3. Results

To test the effectiveness of the method, a series of 5 waveforms were input, and the distortion calculated:

- A simple sine wave:

$$f_1[n] = \sin 0.05\pi n$$

- A square wave:

$$f_2[n] = \text{sqr } 0.05\pi n$$

- A sine wave passed through a cubic distortion to simulate crossover distortion:

$$f_3[n] = \sin^3 0.05\pi n$$

- A clipped sine wave:

$$f_4[n] = \begin{cases} \sin 0.05\pi n & |\sin 0.05\pi n| < 0.8 \\ 0.8 & \text{otherwise} \end{cases}$$

- An unbalanced distortion with a quadratic term:

$$f_5[n] = \frac{3}{4} \sin 0.05\pi n + \frac{1}{4} \sin^2 0.05\pi n$$

The results are compared in table 4, along with the distortion measured using a 256 point FFT.

**Table 4: Harmonic distortion measurements for selected waveforms.**

	All-pass	Piecewise linear	Theoretical	FFT
Sine	0.009%	0.007%	0	0.004%
Square	23.13%	23.13%	23.37%	23.12%
Cubic	11.12%	11.12%	11.11%	11.12%
Clipped	0.83%	0.82%	0.81%	0.82%
Unbalanced	2.64%	8.34% 2.86%	2.78%	2.78%

In measuring a sine wave, all systems measure a small distortion. This results from spectral leakage from the side-lobes of the window function, and can only be reduced by using a longer window.

All of the measurement methods also underestimate the distortion contained within the square wave. This is because the harmonics only decay slowly, and there is significant aliasing. This can only be avoided by using a much higher sample frequency, as discussed earlier.

The cubic and clipped waveforms provide a reasonable estimate of the distortion. With the unbalanced waveform, a significant DC term is generated. This should not be included in the calculation of harmonic distortion. With the all-pass mapping, the DC is not completely separated from the fundamental, so the fundamental component is overestimated, underestimating the distortion. With the piecewise linear mapping, the DC component is

removed in the output, but is present in the input. Equation (15) will therefore include the DC with the harmonics, significantly overestimating the distortion. The second figure represents modifying equation (15) to remove the DC term from the numerator to give a true estimate of the harmonic distortion.

### 4. Discussion

The estimates of distortion provided by spectral warping are close to those obtained using conventional methods. Both of the SWT methods described, the all-pass mapping and the piecewise linear mapping, provide similar results. The piecewise linear mapping is slightly more accurate because of the improved differential scaling.

In terms of computational expense, the piecewise linear mapping is considerably more efficient than the all-pass mapping, requiring only 12% of the operations. Both SWT techniques are more efficient than filtering techniques based on conventional FIR filters. The piece-wise linear mapping requires similar computational expense to using an FFT, although a matrix multiplication would be simpler to programme either on a DSP or in a hardware implementation.

In conclusion, the frequency weighted spectral warping transform provides an efficient method of estimating harmonic distortion to moderate accuracy, and in some applications would be a viable alternative to more conventional filtering methods.

### 5. References:

- [1] A. V. Oppenheim, and D. H. Johnson, "Discrete Representation of Signals," *Proc. IEEE*, vol. 60, June 1972, pp. 681-691.
- [2] S. N. Demidenko, and K. Lever, "Analysis of the Frequency Characteristics of a Digital Spectral Warping Network," in *International Workshop on Design Methodologies for Microelectronics and Signal Processing*, Gliwice-Cracow, Poland, 1993, pp. 397-401.
- [3] W. Allen, D. Bailey, S. Demidenko, and V. Piuri, "Analysis and Implementation of a Spectral Warping Network," *Proc. International Symposium on Integrated Circuits*, Singapore 2001, pp. 492-496.
- [4] W. Allen, D. Bailey, S. Demidenko, and V. Piuri, "Analysis and Application of Digital Spectral Warping in Analog and Mixed-Signal Testing," *IEEE Transactions on Reliability*, vol 52, 2003, pp 444-457.
- [5] D.G. Bailey, W. Allen, S.N. Demidenko, "Spectral Warping Revisited", *Proceedings of the IEEE International Workshop on Electronic Design, Test, and Applications*, Perth, Australia, 2004, pp 23-28.