Audio Engineering Society
Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY, USA

Evaluation of Time-Frequency Analysis Methods and their Practical Applications

Pascal Brunet [1], Zachary Rimkunas [2], and Steve Temme [3]

[1] Listen, Inc., Boston, MA, 02118, USA
pbrunet@listeninc.com

[2] Listen, Inc., Boston, MA, 02118, USA
zrimkunas@listeninc.com

[3] Listen, Inc., Boston, MA, 02118, USA
stemme@listeninc.com

ABSTRACT

Time-Frequency analysis has been in use for more than 20 years and many different Time-Frequency distributions have been developed. Four in particular, Short Time Fourier Transform, Cumulative Spectral Decay, Wavelet and Wigner-Ville have gained popularity and firmly established themselves as useful measurement tools. This paper compares these four popular transforms, explains their trade-offs and discusses how to apply them to analyzing audio devices. Practical examples of loudspeaker impulse responses, loose particles, and Rub & Buzz defects are given as well as a demonstration of their application to common problems with digital/analog audio devices such as Bluetooth headsets, MP3 players and VoIP telephones.

1. INTRODUCTION

Time-frequency analysis is a very natural thing: we do it every day. When we listen to music for example we clearly perceive that frequency content changes over time. A melody played by an instrument is a signal that changes frequency (pitch) over time, and sheet-music is a popular example of time-frequency representation.

While the notion of time-frequency analysis is intuitive, its mathematical definition is far from obvious. The very notion of an instantaneous spectrum contradicts the theory of Fourier analysis.

As J. Ville puts it in his seminal paper “Théorie et Applications de la Notion de Signal Analytique” [1], let’s consider the case of a piece of music where a specific note, e.g A440, appears only once for a limited time. The Fourier spectrum of the complete piece will

exhibit a magnitude and a phase at 440 Hz for a steady sine that neither begins nor ends. To account for the fact that the frequency only appears for a specific portion of time, the Fourier analysis introduces adjacent frequencies with such amplitude and phase that they cancel out the main frequency when it is not there and reinforce it when it is there. As Ville explained, it is a great mathematical tour de force but it is not the reality: when we don't hear the A440, it is simply because it is not played.

After this practical definition of the problem, J. Ville defines the instantaneous spectrum as:

$$W_x(t,\omega) = \frac{1}{2\pi}\int x\left(t+\frac{\tau}{2}\right)\cdot x^*\left(t-\frac{\tau}{2}\right)\cdot e^{-j\omega\tau}d\tau \quad (1)$$

where x(t) is the signal and $x^*(t)$ is the complex conjugate.

In fact this equation was a rediscovery of an expression that Eugene Wigner introduced in 1932 in the context of quantum mechanics, and nowadays it is called the Wigner-Ville Distribution (WVD). It is interesting to note that time-frequency analysis draws heavily from the field of quantum mechanics and shares the same interest in joint distribution of conjugate variable pairs (time-frequency, position-momentum, etc.).

The first practical implementations of time-frequency analysis used analog filter banks to make sonograms of speech, then later FFTs were used to make spectrograms, waterfalls, and order analyses. These are just different representations of the same result.

The FFT based method is now by far the most popular technique as it is simple and efficient. An analysis is obtained by making an FFT based on a short window (short compared to the total length of the analyzed signal), storing the spectrum, shifting the window in time, making a new FFT and repeating the operation until the signal is covered. The resulting stack of spectra is the spectrogram. By using a symmetrical window which is maximal in its center, the FFT gives an estimate of the instantaneous spectrum at the center of the window. To distinguish the result from this process, this method is now called Short Time Fourier Transform (STFT). The properties and limitations of the STFT will be detailed later in this paper.

For a long time the FFT waterfall was the only method of time-frequency analysis, but in the 1980's, important papers [2-5] renewed the interest in the Wigner-Ville Distribution. Around the same time, the Wavelet transform was introduced by J. Morlet and A. Grossmann [10]. In fact many different time-frequency distributions were created around that time and thorough reviews of the field can be found in [6, 7], and their bibliographies. Complete books have also been written on the subject [8-10]. Since the 1980's the concept of time-frequency analysis has become more popular, the engineer's PC has become more powerful and, as a result, computationally intensive transforms like Wigner-Ville and Wavelet are now used along with traditional methods like STFT and Cumulative Spectral Decay.

The paper begins by explaining the general theory of time-frequency analysis, presenting the two Cohen's classes of distributions: time-frequency and time-scale. The central role of the WVD with regards to the other distributions is explained. The way in which the Cohen's Classes relate to the main attributes of a signal is discussed, including energy-time curve, energy spectrum, group-delay, and instantaneous frequency.

The paper then focuses on Heisenberg's uncertainty principle applied to signal analysis and its impact on time and frequency accuracy limits.

The following four selected transforms are presented:

- Short Time Fourier Transform (STFT)
- Cumulative Spectral Decay (CSD)
- Smoothed Pseudo Wigner-Ville Distribution (SPWVD)
- Continuous Wavelet Transform (CWT)

For each transform, the exact mathematical definitions as well as details of its implementation are given.

We will show the underlying relationships between the four transforms by expressing each of them in terms of Cohen's Classes. Their general properties will be summarized and compared in a table.

Applications of time-frequency analysis to audio systems are then demonstrated. First the selected transforms are applied to loudspeaker impulse responses

to show how they contribute to the measurement and study of ringing, reflections, time-alignment, linear and non-linear distortion.

We also show how time-frequency analysis, is effective at isolating loose particle defects from harmonically related Rub & Buzz defects.

Finally, we demonstrate how time-frequency analysis is a valuable tool for the observation of transient behaviors in digital audio devices. These characteristics are very difficult to detect in either the time or the frequency domain alone.

A comparison of the four time-frequency distributions is presented with their trade-offs and preferred application types, and general guidelines are proposed in conclusion.

2. THEORY OF TIME-FREQUENCY ANALYSIS

2.1. Cohen's Classes of Distribution

The general purpose of a time-frequency distribution is to show the energy map of a signal, with respect to time and frequency. In theory, an infinite number of solutions exist and in literature you will find a great number of different distributions.

Following some work done in quantum theory, L. Cohen [6] has proven that any energy time-frequency distribution of a signal x(t) can be expressed as:

$$\rho_x(t,\omega) = \frac{1}{4\pi^2} \iiint x\left(u+\frac{\tau}{2}\right) x^*\left(u-\frac{\tau}{2}\right) \cdot \phi(\theta,\tau)\, e^{j\theta u - j\theta t - j\tau\omega}\, du\, d\theta\, d\tau \quad (2)$$

where $\phi(\theta,\tau)$ is an arbitrary two-dimensional function that determines the distribution and its properties. This expression can be conveniently rewritten as a convolution in the time-frequency plane:

$$\rho_x(t,\omega) = \iint \varphi(u-t, \zeta-\omega)\, W_x(u,\zeta)\, du\, d\zeta \quad (3)$$

by setting:

$$\varphi(t,\omega) = \iint \phi(\theta,\tau)\, e^{j\theta t + j\omega\tau}\, d\theta\, d\tau \quad (4)$$

W_x is the WVD of the signal as defined in (1). The function φ is called the kernel function in the literature [4, 6], and defines the distribution and its properties. All the distributions defined by (3) or (3) constitute the time-frequency Cohen's Class.

It is easy to see that the WVD is a member of Cohen's class. We will show later in the paper how the STFT can be expressed in terms of Cohen's distribution.

Because Cohen's distribution describes how the energy of the signal is distributed both in time and frequency domains, the following attributes of the signal can be derived from it.

2.1.1. Energy

The total energy of the signal is given by:

$$E_x = \iint \rho_x(t,\omega)\, dt\, d\omega \quad (5)$$

2.1.2. Instantaneous Power

The instantaneous power or energy-time curve is given by:

$$|x(t)|^2 = \int \rho_x(t,\omega)\, d\omega \quad (6)$$

2.1.3. Spectrum

The energy density spectrum is given by:

$$|X(\omega)|^2 = \int \rho_x(t,\omega)\, dt \quad (7)$$

2.1.4. Group Delay

The group delay which is the energy time of arrival at each frequency is given by the average time for each particular frequency:

$$T_g(\omega) = \frac{1}{|X(\omega)|^2} \int \rho_x(t,\omega)\, t\, dt \quad (8)$$

2.1.5.Instantaneous Frequency

The Instantaneous frequency which is the energy frequency location at each time is given by the average frequency for each particular time:

$$\Omega_i(t) = \frac{1}{\left|x(t)\right|^2} \int \rho_x(t,\omega)\, \omega\, d\omega \tag{9}$$

It can be noticed that all of these signal characteristic functions are based on summation. Consequently they cannot express fully the complexity of the signal. For example, as pointed out in the introduction, the energy spectrum doesn't tell anything about the temporal distribution of the signal.

Another important point to remember is that these derived functions are approximated in different ways depending on the kernel function used.

If we replace frequency by scale, which is a time compression factor, we can obtain an n^{th} octave based analysis. A general class of energy time-scale distributions has also been defined by L. Cohen [9, 11] and is related to the WVD by:

$$\sigma_x(t,\alpha;\varphi) = \int\int \varphi\left(\frac{u-t}{\alpha}, \alpha\zeta\right) W_x(u,\zeta)\, du\, d\zeta \tag{10}$$

where φ is a kernel function. For each specific time t and scale α the density s is given by a sum of the WVD weighted by the kernel function time shifted to t and time compressed by the scale factor α. We will see later, that the Continuous Wavelet Transform is part of the class defined in (10).

2.2. Uncertainty Principle

In Quantum Mechanics, the measurements of position and momentum of an atomic particle yield random results and their standard deviations are interdependent. Increasing the measurement precision of the position decreases the precision of the momentum, and vice versa. The Heisenberg Uncertainty Principle simply gives a lower boundary to the product of the standard deviations of each quantity.

In Signal Theory, it is well known that duration and bandwidth are coupled: the bandwidth of a signal is roughly equal to inverse of the duration. This is often referred to as the Uncertainty Principle. However, as explained by L. Cohen in [6, 9], although the Uncertainty Principle makes perfect sense in Quantum Mechanics where the quantities are random, it is a misleading name in the context of Signal Theory, as there is no such uncertainty when dealing with signals.

The Uncertainty Principle applied to a signal simply states that the product of the duration of a signal by its bandwidth has a lower boundary. The value of the boundary does not depend on the signal.

However this boundary value depends on how duration and bandwidth are measured. A convenient definition of duration and bandwidth is the energy standard deviation:

$$T = \sqrt{\frac{1}{E_x} \int (t - \langle t \rangle)^2 \left|x(t)\right|^2 dt} \tag{11}$$

$$B = \sqrt{\frac{1}{E_x} \int (\omega - \langle \omega \rangle)^2 \left|X(\omega)\right|^2 d\omega} \tag{12}$$

where ⟨t⟩ and ⟨ω⟩ denote respectively the mean time and the mean frequency of the squared signal. T and B are referred as effective or rms duration and bandwidth.

Using these definitions, for every signal we have the following relationship:

$$BT \geq \frac{1}{2} \tag{13}$$

with T expressed in seconds (s) and B expressed in rad/s. If B is expressed in Hz the limit becomes 1/(4π).

The Uncertainty Principle would therefore be better described as the Heisenberg Limit in signal processing applications. Furthermore, the limit applies only to a signal and has nothing to do with measurement accuracy. In theory, both frequency and time can be measured simultaneously with arbitrary precision. However the accuracy limit usually arises due to the measurement process. For example, to calculate a STFT, first we extract a portion of the signal with duration T. In doing this, we create a new signal which has an rms bandwidth (in Hz) greater than 1/(4πT). Due to truncation, the effective frequency resolution or accuracy is therefore inversely proportional to the duration of the time window. It is not to be confused

with the numerical precision of the FFT which is given by: df= 1/T. As we will see later, the Heisenberg Limit applies in a similar fashion to the Wavelet Transform because of the finite duration of each wavelet. On the other hand, the WVD is not subject to such a limit because the instantaneous spectrum is calculated using the whole history of the signal, without any time truncation. For example, for a linear chirp of frequency αt defined by:

$$x(t) = e^{j\frac{\alpha}{2}t^2} \tag{14}$$

the WVD is:

$$W_x(t,\omega) = \frac{1}{2\pi}\int e^{j\alpha t\tau}\cdot e^{-j\omega\tau}d\tau = \delta(\omega-\alpha t) \tag{15}$$

where δ is the Dirac distribution. The WVD is totally concentrated along the instantaneous frequency of the chirp. This is due to a property of the bilinear product $p(t, \tau)= x(t+\tau/2). x^*(t-\tau/2)$ that appears under the integral sign of the WVD (1).

If we denote the instantaneous frequency of x(t) by $\varpi(t)$, then the instantaneous frequency of the product $p(t, \tau)$ is obtained by:

$$\begin{aligned}\frac{d}{d\tau}\arg\big(p(t,\tau)\big) &= \frac{d}{d\tau}\arg\left(x(t+\frac{\tau}{2})\cdot x^*(t-\frac{\tau}{2})\right)\\ &= \frac{1}{2}\varpi(t+\frac{\tau}{2})+\frac{1}{2}\varpi(t-\frac{\tau}{2})\end{aligned} \tag{16}$$

As shown in Figure 1, when $\varpi(t)$ is linear, the two instantaneous frequencies in (16) have opposite slopes with regard to the delay τ and they compensate each other. $p(t, \tau)$ then has a steady frequency in regard to τ and its Fourier transform is therefore a Dirac distribution. The value of the steady frequency is $\varpi(t)$, the instantaneous frequency at time t. It is clear that in this case the Heisenberg limit does not apply: the exact value of instantaneous frequency is known at any time.

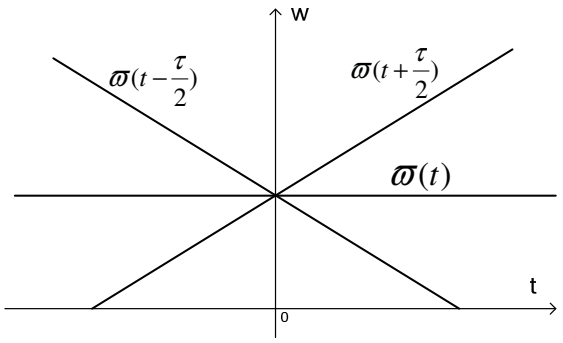

Figure 1 Frequency laws of bilinear product

As explained in [8,12] the assumption of stationary frequency used in Fourier analysis is replaced in Wigner-Ville by a less stringent assumption of linear frequency progression.

2.3. Short Time Fourier Transform

The continuous time definition of the STFT is:

$$S_x(t,\omega) = \left|\frac{1}{\sqrt{2\pi}}\int x(\tau)\cdot h(\tau-t)\cdot e^{-j\omega\tau}d\tau\right|^2 \tag{17}$$

The function h has a limited duration and is usually symmetric with a maximum at its center as to emphasize the signal at the middle point. The function h is shifted to the time t and acts as a window to get the local spectrum S_x around t.

The window used in this case is a time limited Gaussian window:

$$h(t) = \frac{1}{\sqrt{\alpha T\sqrt{2\pi}}}e^{-\left(\frac{t}{2\alpha T}\right)^2}1_{[-\frac{T}{2},\frac{T}{2}]}(t) \tag{18}$$

with a total duration of T, and a unit energy. Its effective duration (11) is αT. The coefficient α is chosen to minimize the side-lobes due to truncation. The Gaussian function has been chosen because it is the most concentrated signal both in time and frequency. It is the only signal that reaches the Heisenberg limit.

For implementation we used the following definition in the digital domain:

$$S[n,m] = \left| \sum_{p=-P}^{N-P-1} h[p] \cdot x[n+p] \cdot e^{-j2\pi mp/N} \right|^2 \quad (19)$$

where n is incremented from a start sample n_0 in steps of Δn: $n = n_0 + k\,\Delta n$. When the increment Δn is smaller than the window length N, there is overlap between two successive windows.

As we will show later, the STFT distribution is smoothed version of the WVD. The effective time resolution is therefore limited to the effective duration of the window used. Similarly, the frequency resolution is limited by the bandwidth of the window. The practical consequence is the well-known fact that a short duration window yields a good time resolution and a poor frequency resolution, and inversely a long duration window yields a poor time resolution and an improved frequency resolution.

When a time-frequency distribution is smoothed by a Gaussian function of effective duration ΔT, we can downsample that smoothed distribution along the time axis, i.e. calculate spectra separated by several samples. The sub-sampling rate must remain higher than twice the bandwidth of the smoothed function to prevent any loss of information. For example, some energy peaks may be ignored if the spectra in the STFT are too far apart. For a Gaussian window, the effective bandwidth (in Hz) is $\Delta f = 1/(4\pi\,\Delta T)$. For the maximum frequency we can use: $6\,\Delta f \approx 1/(2\,\Delta T)$, at which you get 78 dB attenuation of the spectrum. Hence the smoothed time-frequency distribution can be sampled in time with time steps dt smaller than the effective duration of smoothing ΔT without any loss of information. In other words the STFT is complete when: $dt \le \Delta T$. The completeness of the distribution is necessary to get a correct energy spectrum. To obtain completeness of the distribution it is not necessary to calculate a spectrum for each sample as long as the time increment is small enough. The practical consequence is that you can still obtain a complete energy distribution of a signal with a reduced number of spectra as long as they are separated by a time increment equal to the effective window duration.

In order to demonstrate time-frequency analysis, we used SoundMap™ software from Listen, Inc. In figure 2, a STFT is applied to a filter's impulse response. The filter is a band-pass Chebyshev with 3dB ripples. The lower left time-waveform graph shows the impulse response, the upper right graph shows the global spectrum or frequency response, and the upper left intensity map, shows both the time versus frequency plot of the band-pass filter. Note that the intensity map shows both the resonant frequency and decay time (ringing) for all three resonant frequencies at 200, 6.5k, and 10 kHz.

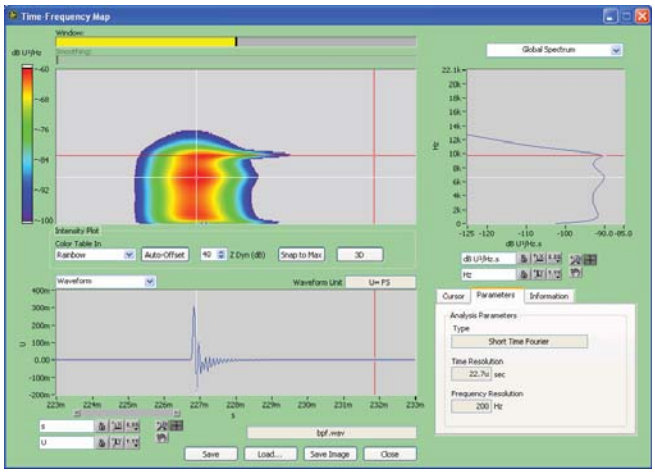

Figure 2 Impulse Response of a Band-Pass Filter (STFT)

2.4. Cumulative Spectral Decay

This transform was introduced in the late 1970's [13], for the study of loudspeakers, and has been popular in the loudspeaker industry ever since. Its purpose is to reveal and analyze the resonance frequencies of an acoustical transducer.

The continuous time definition of the CSD is:

$$C(t,\omega) = \left| \int_t^{t+T} x(\tau)\, e^{-j\omega\tau} d\tau \right|^2 \quad (20)$$

where x(t) is the impulse response of the device under test (DUT) and T is usually longer than the duration of x(t).

The CSD is a STFT with a rectangular right-sided window. It can also be viewed as the squared magnitude of the convolution of x by the stopping tone burst $1_{[-T,0]}(t)\,\mathrm{Exp}(j\omega t)$ of duration T.

C(t, ω) gives the temporal decay of a stopping tone of frequency ω applied to the DUT. Applied to a loudspeaker impulse response it clearly shows the resonance frequencies and ringing. Since CSD uses a

rectangular window, the integration time T must be longer than the impulse response duration, otherwise severe ringing artifacts will occur.

For implementation we used the following definition in the digital domain:

$$C_x(n,k) = \left| \sum_{u \geq n}^{Min(Q,N)-1} w(u) \cdot x(u) \cdot e^{-j2\pi \frac{uk}{N}} \right|^2 \qquad (21)$$

with Q: end of time block, N: FFT size.

A Tukey window w(u) (rectangular window with cosine tapers) is applied to the time block before performing the FFT to attenuate the wide transient spectra that would otherwise pollute the map.

Because the analyzed data shortens as the integration window slides, there is a low frequency limit for each spectrum under which the values are meaningless. The limit is equal to the inverse of the effective duration of the integration. Under that limit, the levels are usually zeroed or masked.

Another type of CSD exists with the integration limits reversed:

$$B(t,\omega) = \left| \int_{t-T}^{t} x(\tau)\, e^{-j\omega\tau} d\tau \right|^2 \qquad (22)$$

It is called Cumulative Spectral Attack and it gives the response of a system to a starting tone. It is the same STFT shifted in time by T. This version is rarely used because the transition state of the attack is superimposed on the steady state and as a result is more difficult to visualize.

In figure 3, the CSD is applied to the impulse response of the same band-pass filter as in Figure 2. The constant energy zone before the impulse begins at about 227ms is where the analysis window encompasses the entire signal. Therefore, the instantaneous spectra are all equal to the frequency response of the DUT. The wide frequency spread that occurs above the constant energy zone when the impulse response begins is an artifact due to the signal truncation by the quasi-rectangular window. Compared to the STFT analysis of the filter, it is more difficult to see the three distinct resonances.

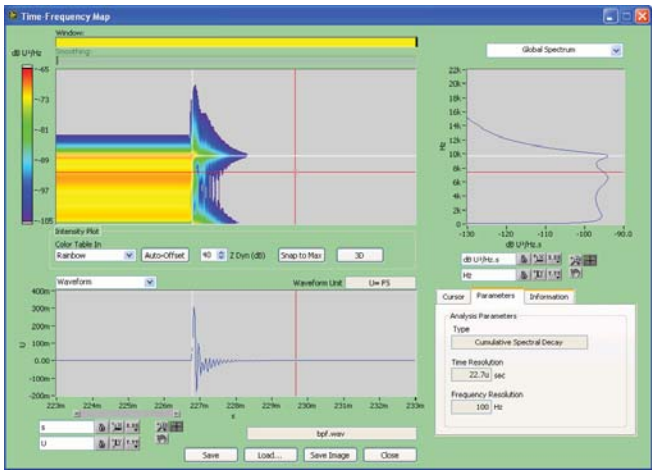

Figure 3 Impulse Response of a Band-Pass Filter (CSD)

2.5. Wigner-Ville Distribution

The WVD is a very sharp distribution however it is well-known that it generates strong interferences and negative values [2, 5-9]. Figure 4 shows an example of these interferences.

In this paper, we used a smoothed estimator of the WVD called Smoothed Pseudo Wigner Distribution (SPWVD) [12, 14] and defined by:

$$\mathrm{SW_x}(t,\omega) = \int\int \mathrm{h}(\tau)\, \mathrm{g}(u)\, \mathrm{x}\left(t+u+\frac{\tau}{2}\right) \cdot \mathrm{x}^*\left(t+u-\frac{\tau}{2}\right) \mathrm{e}^{-i\tau\omega}\, du\, d\tau \qquad (23)$$

with h, g being two window functions. These functions smooth the WVD along the frequency axis with **H**, the Fourier transform of **h**, and along the time axis with **g**:

$$\mathrm{SW_x}(t,\omega) = \int\int g(u-t)\, H(\zeta-\omega)\, W_{\mathrm{x}}(u,\zeta)\, du\, d\zeta \qquad (24)$$

The two chosen window functions are Gaussian as was the case for STFT. The estimator $\mathrm{SW_x}$ allows, by an appropriate choice of the two window sizes, attenuation of the interferences that may otherwise obscure the distribution of a complex signal (compare Figure 4 and Figure 5). The counterpart is that the marginal densities (instantaneous power, energy spectrum) are also smoothed (see Table 1 on page 11).

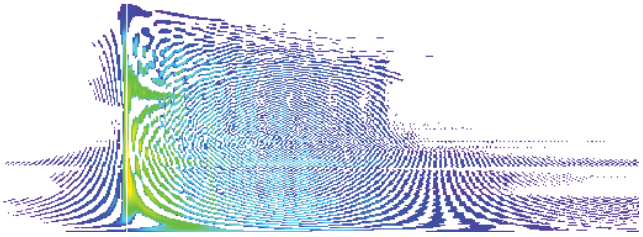

Figure 4 WVD of a Loudspeaker Impulse Response.

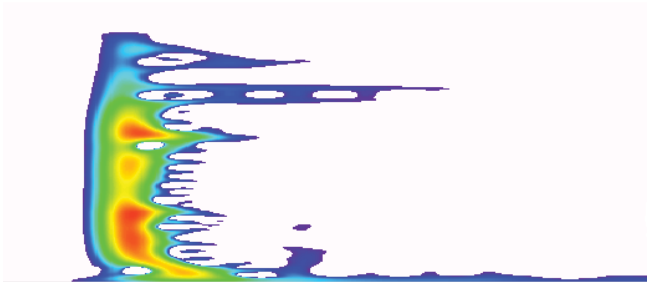

Figure 5 SPWVD of the same signal.

In addition, to eliminate the cross-terms that appear between positive and negative frequencies, the signal is stripped of its negative frequencies by analytical filtering before analysis. To remove the negative frequencies from the signal we add an imaginary part equal to the Hilbert Transform of the signal:

$$\tilde{\mathrm{x}}(t) = \mathrm{x}(t) + j\int \frac{\mathrm{x}(\tau)}{\pi(t-\tau)} d\tau \qquad (25)$$

For implementation we used the following definition in the digital domain [3]:

$$\mathrm{SW}[n,m] = 2\sum_{p=-P}^{N-P-1} h[p]\; e^{-j4\pi mp/N} \cdot \left(\sum_{q=-Q}^{Q} g[q]\, \tilde{x}[n+q+p]\, \tilde{x}^{*}[n+q-p] \right) \qquad (26)$$

With:

- g[q]: Gaussian of 2Q+1 points,
- h[p]: Gaussian of N points.

Because of the factor 4π in the exponential, we have frequency aliasing occurring at the Nyquist frequency instead of the Sampling frequency:

$$\mathrm{SW}[n, m+\frac{N}{2}] = \mathrm{SW}[n,m] \qquad (27)$$

It is possible for the negative frequencies of the signal to overlap the positive frequencies. This situation is avoided by the analytical filtering mentioned earlier.

The smoothed WVD (SPWVD) is calculated with a time increment that is usually bigger than the sampling interval. As with STFT, the threshold of completeness of the distribution is the effective duration of the time smoothing window.

The Gaussian window was chosen because of its optimal concentration in the time-frequency plane. When **g** and **h** have the same duration the SPWVD becomes identical to a STFT [8]. The practical consequence of this is that it is possible to obtain a tunable intermediate between a WVD and a STFT. Moreover, the amount of smoothing can be tuned independently along the time and frequency axis.

It is worth noting that contrary to the STFT, the measurement accuracy of the smoothed WVD is not boundary by the Heisenberg Limit: the time resolution is decoupled from the frequency resolution.

In figure 6, the SPWVD is applied on the same filter impulse response as in Figure 2 & 3. Compared to STFT and CSD, the main energy is more concentrated in a narrower time region and the resonances are sharper both in time and frequency. The main drawback is the resulting interference pattern or side-lobes from the sharper filters which can be a little distracting. However, their strength can be reduced by increasing smoothing and decreasing frequency resolution.

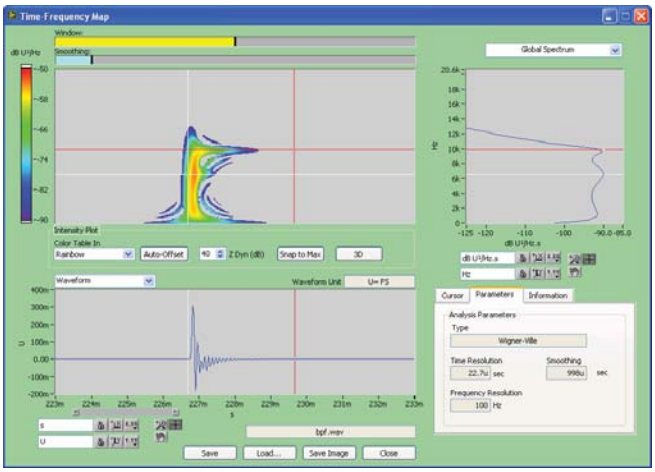

Figure 6 Impulse Response of a Band-Pass Filter (SPWVD)

2.6. Wavelet Transform

The Continuous Wavelet Transform (CWT) is defined by [10]:

$$CWT_x(t,a) = \frac{1}{\sqrt{|a|}} \int x(\tau)\, h\left(\frac{\tau - t}{a}\right) d\tau = \frac{\sqrt{|a|}}{2\pi} \int X(\omega)\, H(a\omega)\, e^{i\omega t} d\omega \qquad (28)$$

The spectral expression shows that the CWT is equivalent to a filter bank with constant percentage bandwidth.

For time-frequency analysis purposes we use the square magnitude of the spectral expression in (28) and replace the scale by the inverse of the frequency:

$$\Lambda_x(t,\omega) = \left| \frac{1}{\sqrt{\omega}} \int X(\zeta) H\left(\frac{\zeta}{\omega}\right) e^{j\zeta t} d\zeta \right|^2 \qquad (29)$$

With:

- $X(\zeta)$: spectrum of the signal to analyze
- $H(\zeta)$: Wavelet spectrum

For the wavelet **h**, we use a Gaussian analytical wavelet for its optimal time-frequency behavior. The wavelet spectrum is energy normalized to get values calibrated as spectral power density.

The main advantage of using the frequency expression to calculate the wavelet transform is that we have complete control over the central frequencies and bandwidth of each wavelet filter. To facilitate the comparison of the result with traditional measurement, we choose to follow standardized 1/n octave (ISO) frequencies.

For implementation we used the following definition in the digital domain:

$$T[n,f] = \left| \sum_{k=0}^{N-1} H_f[k] \cdot X[k] \cdot e^{j2\pi nk/N} \right|^2 \qquad (30)$$

where f is the center frequency of each band.

To reduce the size of the result, the CWT is usually calculated with a time increment bigger than the sampling interval. To achieve this, the distribution is downsampled along the time axis. In a similar manner to STFT, the threshold of completeness is the rms duration of the shortest wavelet (maximum center frequency).

It is worth noticing that the wavelet transform is subject to Heisenberg limit as is the case with STFT. However unlike STFT, for each center frequency the duration of the wavelet is inversely proportional to the bandwidth. The bandwidth is proportional to the analysis frequency, which is specific to the wavelet transform.

In figure 7, the CWT is applied to the impulse response of the same band-pass filter as in Figure 2, 3 & 6. Unlike the other transforms, CWT has more resolution at low frequencies and it is easier to see the full bandwidth of the DUT's impulse response on one time-frequency map. The broadening caused by the analysis filters, however, makes it a little more difficult to distinguish the DUT's ringing from the analysis filter's bandwidth. The ideal shape of the CWT to a perfect, Dirac impulse response is a symmetric pear shape with constant sloping sides.

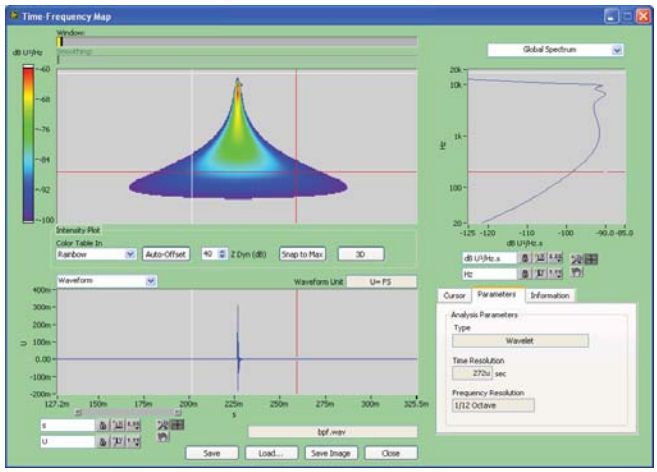

Figure 7 Impulse Response of a Band-Pass Filter (CWT)

2.7. Relationships and Properties Summary

As mentioned before, all time-frequency distributions can be expressed as the result of a time-frequency convolution between the WVD of the signal and a kernel proper to the transform (eq. 3 & 10). The kernel entirely defines the time-frequency transform and its properties. The kernel also smoothes the marginal distributions: instantaneous power and energy spectrum. The kernel sets the resolution of the distribution in both time and frequency. Table 1 gives the kernel expression for each transform and the resulting smoothing on the marginal distributions. The kernels for STFT, SPWVD, CWT are defined using Gaussian functions as described in the previous chapters.

The following figures show the different kernels in the time-frequency plane to get a visual understanding of them. There are graphical representations of the two-dimensional functions defined in the second column of Table 1. The kernel for the WVD is not shown because it is a single point. It is important to note that bandwidth and duration are decoupled for the SPWVD. The kernel of the CSD stands out with its strong side-lobes pattern. Also it is important to see how duration and bandwidth evolve with frequency for the CWT.

Note: In this four figures, k is an arbitrary constant.

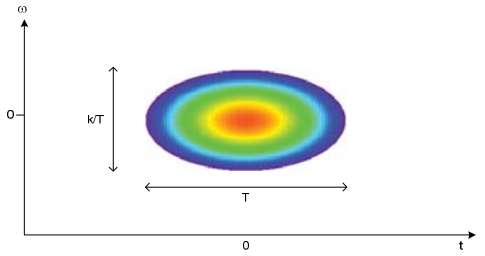

Figure 8 Kernel of the STFT.

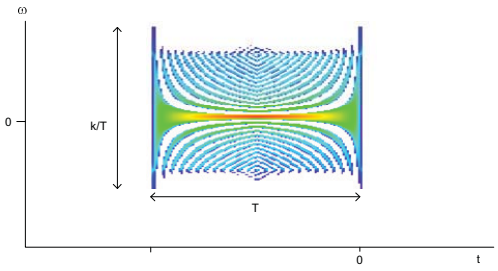

Figure 9 Kernel of the CSD

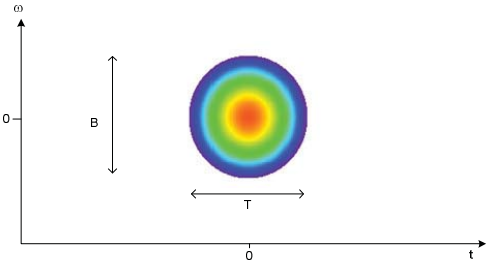

Figure 10 Kernel of the SPWVD

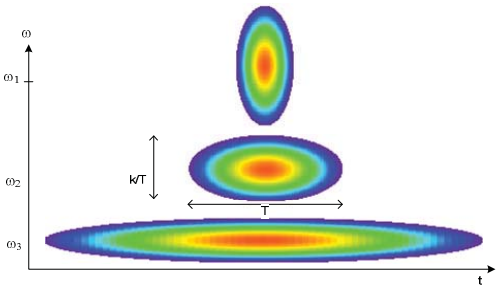

Figure 11 Kernels of the CWT for three wavelet frequencies

Transform	Time-Frequency Kernel	Instantaneous Power	Energy Spectrum	Heisenberg Limit?
WVD (Wigner)	$\frac{1}{2\pi}\,\delta(t)\,\delta(\omega)$	$\lvert x(t)\rvert^2$	$\lvert X(\omega)\rvert^2$	No
STFT (Fourier)	$W_h(t,\omega)=\frac{1}{\pi}e^{-\frac{t^2}{\sigma}-\sigma\omega^2}$ with $\sigma=2\alpha^2T^2$	$\lvert x(t)\rvert^2 * e^{-\frac{t^2}{\sigma}}$	$\lvert X(\omega)\rvert^2 * e^{-\sigma\omega^2}$	Yes. Time and frequency resolutions are coupled.
CSD (Cumulative Spectral Decay)	$\frac{1}{\pi\omega}\sin\left[\omega\left(1-\left\lvert\frac{2t+T}{T}\right\rvert\right)\right]\cdot 1_{[-T,0]}(t)$	$\lvert x(t)\rvert^2 * 1_{[-T,0]}(t)$	$\lvert X(\omega)\rvert^2 * \left\lvert\frac{2}{\omega T}\sin\left(\frac{\omega T}{2}\right)\right\rvert^2$	Yes. Time and frequency resolutions are coupled.
SPWVD (Smoothed Wigner)	$g(t)\cdot H(\omega)=\frac{1}{\pi\sqrt{\alpha\beta}}e^{-\frac{t^2}{\alpha}-\frac{\omega^2}{\beta}}$	$\lvert x(t)\rvert^2 * e^{-\frac{t^2}{\alpha}}$	$\lvert X(\omega)\rvert^2 * e^{-\frac{\omega^2}{\beta}}$	No. Time and frequency resolutions are arbitrary.
CWT (Wavelet)	$W_h\left(\eta t,\frac{\omega}{\eta}\right)=\frac{1}{\pi}e^{-\frac{\eta^2t^2}{\sigma}-\sigma\frac{\omega^2}{\eta^2}}$ with η center frequency of the wavelet.	$\int\left\lvert\int X(\zeta)e^{-\frac{\sigma}{2}\left(\frac{\zeta}{\omega}\right)^2}e^{j\zeta t}d\zeta\right\rvert^2\frac{d\omega}{\omega}$	$\int\lvert X(\eta)\rvert^2 e^{-\sigma\left(\frac{\eta}{\omega}\right)^2}\frac{d\eta}{\omega}$	Yes. Time and frequency resolutions are coupled for each wavelet.

Table 1 Summary of Time-Frequency Distributions Properties.
Note: * denotes convolution.

3. APPLICATIONS OF TIME-FREQUENCY ANALYSIS TO AUDIO SYSTEMS

All the following measurement data and analysis were performed using Listen, Inc's SoundCheck® audio measurement system and SoundMap™ time-frequency analysis software.

3.1. Loudspeaker Impulse responses

One useful application of time-frequency analysis is for studying the impulse response of transducers. The selected distributions are applied to some loudspeaker impulse responses to demonstrate how they contribute to the measurement and study of:

- Ringing and resonance frequencies
- Reflections
- Frequency-dependent acoustic center and group-delay
- Time-alignment of transducers
- Influence of crossover filters
- Non-linear distortion: time-frequency analysis of harmonic impulses responses

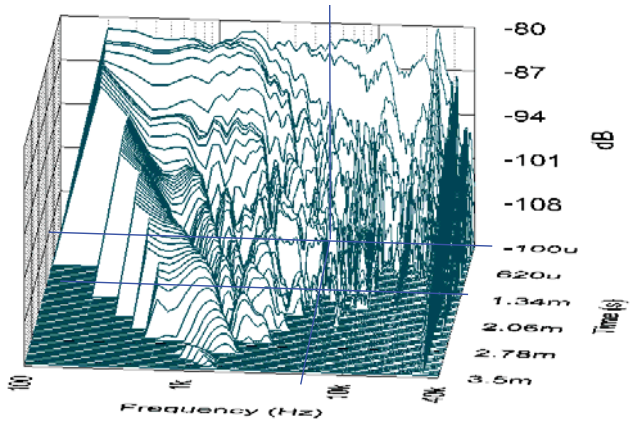

Figure 12 Traditional "Waterfall Plot" of a 2-Way Studio Monitor Loudspeaker (CSD)

All loudspeakers have resonances and a typical 2-way loudspeaker with two different drivers will have at least two natural resonance frequencies, one at a low frequency from the woofer and one at a high frequency from the tweeter. It can be interesting to see how these resonances behave in both the time and frequency domains. Typically, a low Q at resonance is desirable to minimize amplitude peaks and ringing. Figure 12 shows a traditional cumulative spectral decay "Waterfall plot" for a 2-way studio monitor loudspeaker. This particular loudspeaker has a titanium dome tweeter mounted in a wave guide to better focus high frequencies.

The waterfall plot in figure 12, shows that the tweeter has a very high Q resonance at high frequencies above 30 kHz and rings for almost 3.5ms. However, if we look at the same measurement data using a smoothed Wigner-Ville transform and plot as a 3D surface map (figure 13), we get quite a different picture.

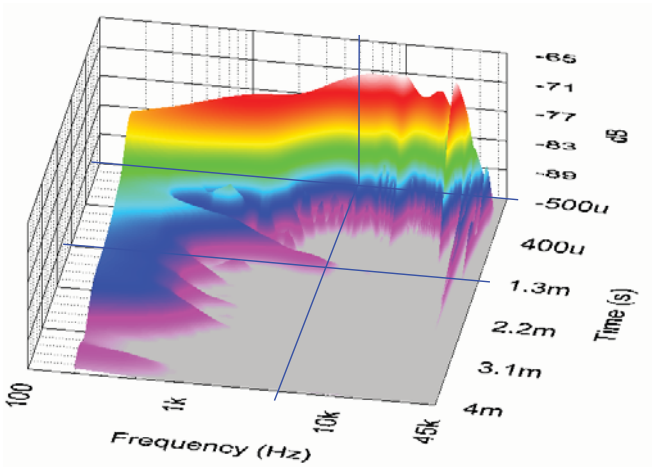

Figure 13 Surface Plot of the same 2-Way Studio Monitor Loudspeaker (SPWVD)

With the SPWVD surface map it is easier to see details of the loudspeaker's resonances and it more clearly shows reflections that might otherwise be misinterpreted as a resonance (e.g. approximately @ 1.3ms below 5 kHz). It can also be seen that the high frequency ringing of the tweeter is sharper and only lasts for about 2.2ms as opposed to 3.5ms that appears in the CSD waterfall plot. This is most likely an artifact of CSD's quasi-rectangular window which is not as smooth as the SPWVD's Gaussian window and therefore can contribute some ringing artifacts.

Figure 14 shows the time-frequency intensity map with corresponding group delay and time envelope graphs from which the 3D surface plot was derived. This map allows many more interesting details about the loudspeaker and sound measurement to be explored. For

example, the main energy band from the direct sound arrival at the far left of the intensity map is the fundamental impulse response of the loudspeaker. The more narrow and vertical the energy band, the better is the impulse response of the loudspeaker. This loudspeaker has many high frequency resonances, probably due to the stiffness of the titanium dome tweeter and the hard wave guide (i.e. horn). It is now easier to identify the two big resonances at 31 & 35 kHz. There is also a small lower frequency reflection at approximately 1.3ms after the main arrival of the direct sound from the loudspeaker. This is difficult to see in the time response. It is also challenging to time window out the reflections without windowing out the high frequency resonances. The 3D time-frequency map makes it easier to discriminate reflections from resonances.

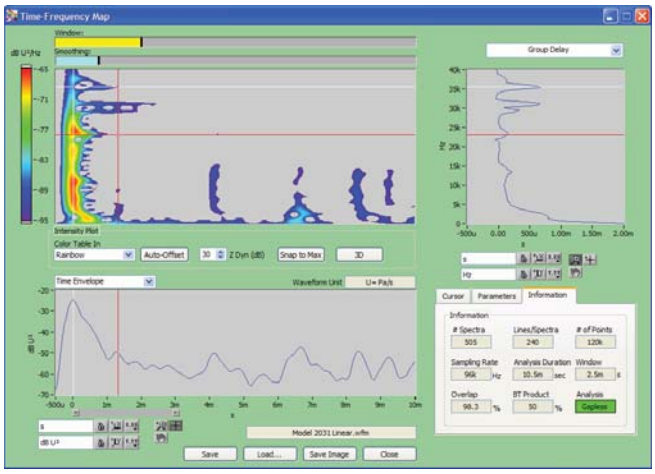

Figure 14 Intensity Map, Time Envelope, and Group Delay of the same Loudspeaker (SPWVD)

By looking more closely at the lower, left hand portion of the time-frequency map, it is obvious that the low frequency woofer and high frequency tweeter are not time aligned. Notice how the main energy area bends or is delayed at low frequencies compared to the frequencies above 3 or 4 kHz. This can also be seen from the group delay curve in the upper right graph. Ideally, the group delay should be a straight vertical line indicating constant group delay.

Some speakers use steep crossovers between the woofer and tweeter that may have a flat amplitude response but poor time response. Look at the ringing around 2 kHz in the surface plot of figure 15 due to a poorly designed crossover.

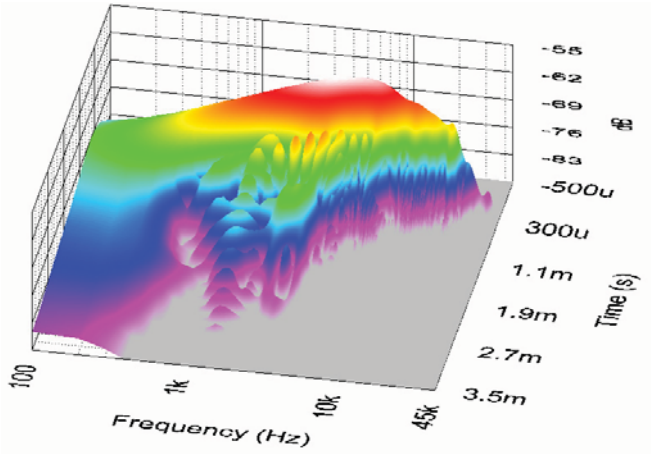

Figure 15 Different 2-Way Loudspeaker with Poor Transient Response at 2 kHz Crossover Frequency (SPWVD)

Thanks to Farina's method for measuring impulse response with a continuous log sweep, it is possible to analyze harmonic impulse responses [15]. By performing time-frequency analysis on a loudspeaker's harmonic impulse response, new characteristics never revealed before can be seen. Figure 16 shows the 3rd harmonic impulse response of the same 2-way studio monitor loudspeaker. Even the 3rd harmonic rings at the tweeter's resonant frequencies.

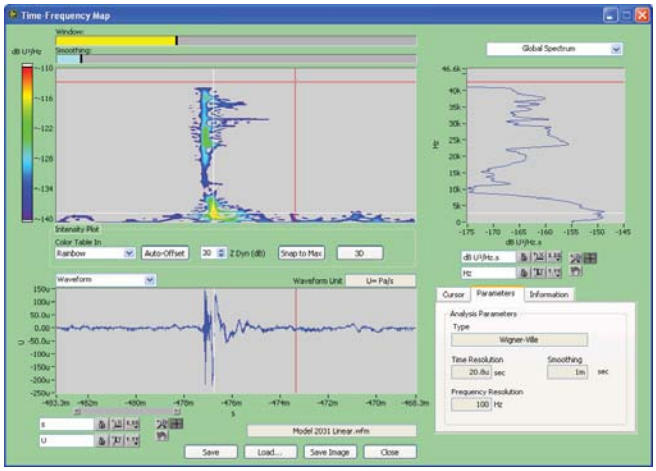

Figure 16 Studio Monitor 3rd Harmonic Impulse Response (SPWVD)

Due to the constant bandwidth analysis of both the Wigner-Ville and Cumulative Spectral Decay transforms, it is difficult to see measurement data details at both low and high frequencies on the same time-frequency map even when plotted on a log frequency scale. This tends to emphasize or magnify the high

frequencies at the expense of the low frequency information.

The Wavelet transform, however, uses constant percentage frequency resolution rather than constant resolution. Compared to STFT, this offers better time resolution at high frequencies and better frequency resolution at low frequencies. Figure 17 shows a 12th Octave Wavelet analysis of a small, full-range, single-driver loudspeaker's impulse response. The impulse response was "spliced" [16] from a low frequency, near-field measurement and a high frequency, time-windowed, far-field measurement to create a full-bandwidth, impulse response from 20 to 40 kHz (50ms long). Note that due to a limited test room size, many time-windowed impulse response measurements of loudspeakers are quite short e.g. under 10ms, in order to remove the first room reflection from the measurement. Consequently they will have poor frequency resolution e.g. 100 Hz and lose some of the advantages of the Wavelet transform. Splicing overcomes these issues and makes the Wavelet transform useful for this application. The ideal shape of a loudspeaker's impulse response using a Wavelet transform would be a symmetric, pear shape similar to figure 7.

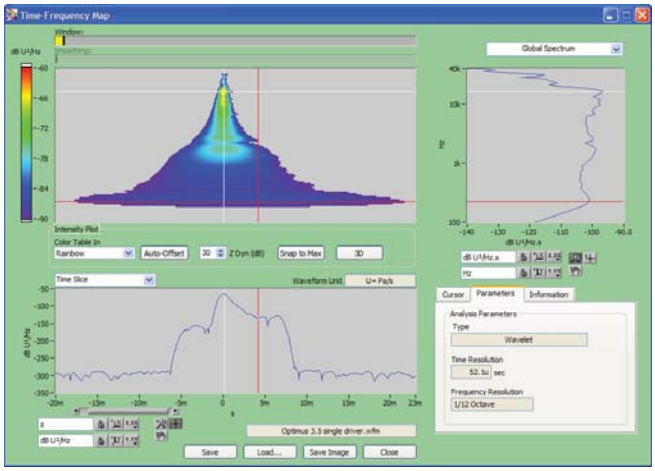

Figure 17 Analysis of a Small, Single 3-inch Driver Loudspeaker (CWT)

Analysis shows that this speaker has a high Q resonance at low frequencies, a resonance at about 2 kHz probably the first breakup mode of the cone, and a high frequency resonance around 18 kHz where the cone breakup becomes chaotic.

3.2. Loudspeaker Loose Particles and Rub & Buzz Defects

In loudspeaker production, particles of dirt, solder beads, glue, magnet or metal chips can become lodged in the speaker coil and cause a distinctive rattling sound.

These impulses occur at random positions in time, uncorrelated to the stimulus's harmonic content. As such, typical distortion tests (such as Rub & Buzz) based on harmonic analysis are not effective for loose particle detection.

Time-frequency analysis, however, is very effective at isolating loose particle defects from harmonically related Rub & Buzz defects. Some examples are shown here.

Several identical woofers with varying faults were tested with a swept-sine stimulus over the frequency range of 20 - 1000 Hz. This frequency range was chosen because loose particles tend to occur near the loudspeaker's first resonant frequency and maximum cone displacement, which in this case was around 57 Hz. Data was collected and time-frequency analysis performed using a STFT to produce the figures below.

In a speaker with no loose particle fault (Figure 18) the time-frequency map shows only the stimulus signal and a few additional harmonics.

In a speaker with a Rub & Buzz fault but no loose particles, the Rub & Buzz fault can be seen as clusters of harmonics, horizontal bands "smeared" along the time axis (Figure 19). The Rub & Buzz fault is recognized as a periodic disturbance.

In a speaker with an obvious loose particle fault, the loose particles hit randomly in time, and appear as vertical bands "smeared" along the frequency axis (Figure 20).

The map for the speaker with both faults clearly shows the difference between the Rub & Buzz and loose particle faults. There are loose particle "streaks" over a wide range of frequencies but a short span of time and Rub & Buzz "streaks" running in narrow frequency bands over longer periods of time (Figure 21).

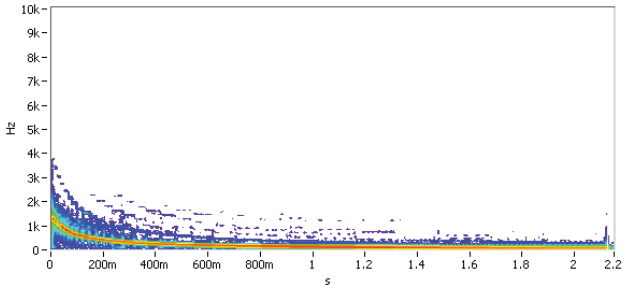

Figure 18 Good Speaker (STFT)

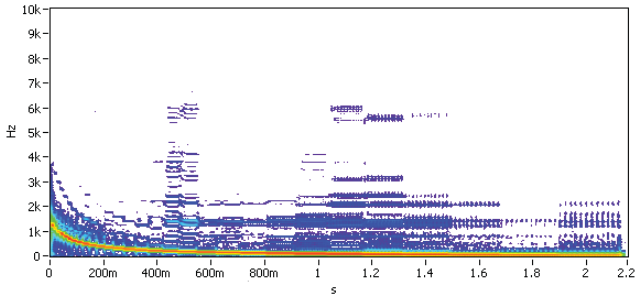

Figure 19 Rub & Buzz defect (STFT)

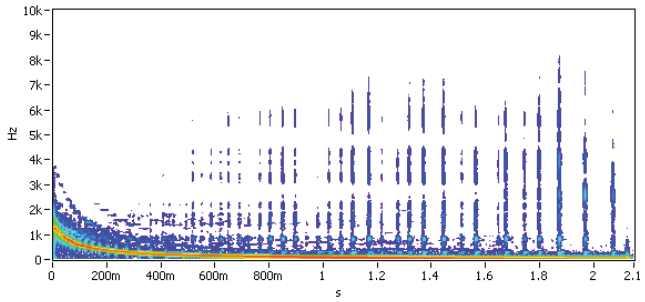

Figure 20 Loose Particle Defect (STFT)

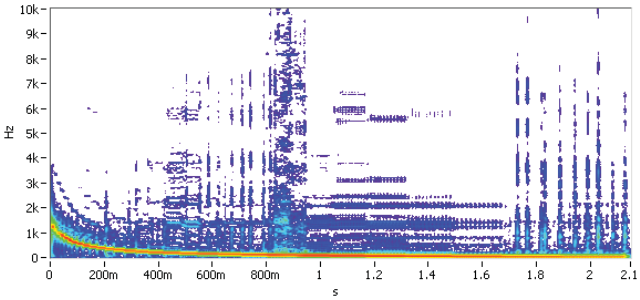

Figure 21 Rub & Buzz and Loose Particles (STFT)

Note: All Figures above use STFT with the same analysis settings: 4 ms time increment and 25 Hz frequency resolution. The dynamic range is 60 dB.

3.3. Transient and Instability Problems in Digital Devices

Modern day digital/analog devices such as Bluetooth headsets, MP3 players and VoIP telephones exhibit problematic behavior due to packet loss, jitter, sampling rate inaccuracies and perceptual codecs. These problems manifest themselves as transients or instabilities in the output of the devices and are difficult to detect in either the time or the frequency domain. The simultaneous view offered by Time-Frequency analysis provides a valuable tool for detection and analysis of such problems. All figures in this section used the Wigner-Ville distribution because of its flexibility and precision.

3.3.1. Bluetooth headset jitter

Digital audio devices with sampling rate fluctuations can introduce jitter. This is generally due to the inexpensive crystal clock used in low cost consumer goods such as Bluetooth devices. The subtlety of jitter can make it difficult to detect by visual inspection in the time or frequency domain alone. Time-Frequency Analysis allows one to detect jitter by examining a single tone as it progresses over time.

To obtain a Time-Frequency map of a jitter curve, a 10.5 second 16 kHz tone was played through a Bluetooth headset. Analysis revealed a peak to peak fluctuation of 50 Hz (Figure 22). Relative to the 44.1 kHz sampling rate, the sampling rate fluctuation is approximately 140 Hz. Time Frequency Analysis produces a clear view of Bluetooth headset jitter that is not available using traditional stationary time or frequency domain analysis techniques.

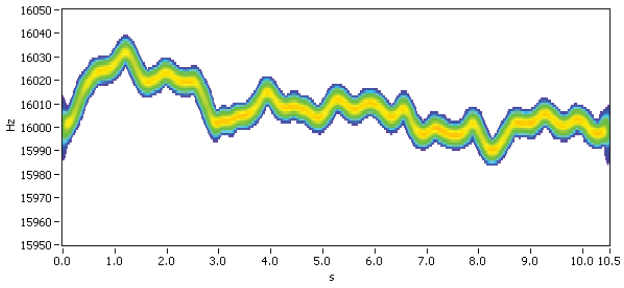

Figure 22 Bluetooth Headset Jitter versus Time using Wigner-Ville Transform (SPWVD)

3.3.2. Bluetooth headset dropouts

In addition to jitter, Bluetooth headsets suffer from signal dropouts. Dropouts are the result of packet loss

due to weak signal or RF interference. Time-Frequency Analysis allows for the detection of dropouts that might go unnoticed if examined only in the time or frequency domain.

For example, a segment of human speech (Figure 23) was recorded using a Bluetooth headset. When looking at the time waveform, it can be difficult to distinguish between dropouts and silent sections of speech. However, when viewing a Time-Frequency map of the speech in Figure 24 the dropouts appear as vertical lines where the level is significantly lower than the surrounding areas of the map. The dropouts occur at approximately 9.8 seconds and 10.1 seconds. Time-Frequency Analysis clarifies the identification of subtle Bluetooth dropouts that may be missed by visual inspection of the waveform.

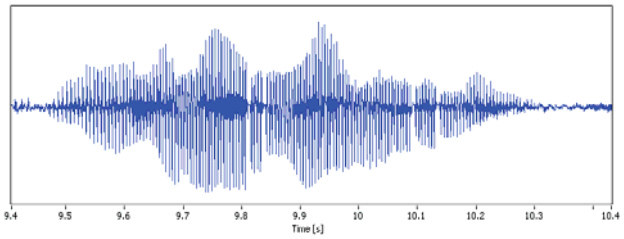

Figure 23 Speech Waveform Recorded using Bluetooth Headset

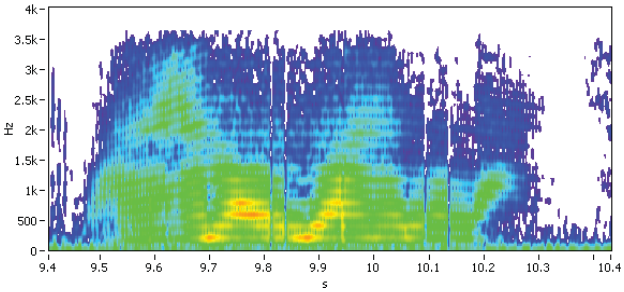

Figure 24 Time-Frequency Map of Recorded Speech Waveform Exhibiting Dropouts (SPWVD)

3.3.3.MP3 Player Impulse Response

MP3 encoders utilize perceptual coding techniques to lower the bit rate of audio waveforms while minimally degrading the sound quality. The effects of MP3 encoders on a waveform are illuminated by Time-Frequency Analysis.

For example, a simple Dirac was MP3 encoded at 96 kilobits per second (kbps). The Time-Frequency map of the Dirac (Figure 25) reveals interesting information about the encoder.

First, the encoder restricts the bandwidth of the Dirac to approximately 17 kHz. The bandwidth restriction is possible because of the ear's insensitivity at such high frequencies. The encoder assumes that the listener is not capable of hearing such high frequencies and they are therefore disposable. Second, the pre-ringing from 70ms to 75ms at 17 kHz indicates that the encoder is using FIR filters to limit the bandwidth. Finally, the MP3 encoder utilizes temporal masking to shape the noise of the waveform. Temporal masking occurs when two distinct sounds occur very close together in time. If the first sound is of a significantly higher level than the second, the second will not be heard. During the 20 ms following the Dirac, the encoder decreases the number of bits used to quantize the waveform. Temporal masking indicates that the added quantization noise will not be heard because it occurs so close in time to the Dirac. Time-Frequency analysis clearly shows that the noise was present (Figure 25).

The energy spectrum of the noise was then calculated (Figure 26). The spectrum was calculated by summing all of the spectra between the vertical cursors (77 ms to 95 ms). This shows that the noise is shaped to be significantly lower in the 3 kHz to 8 kHz range.

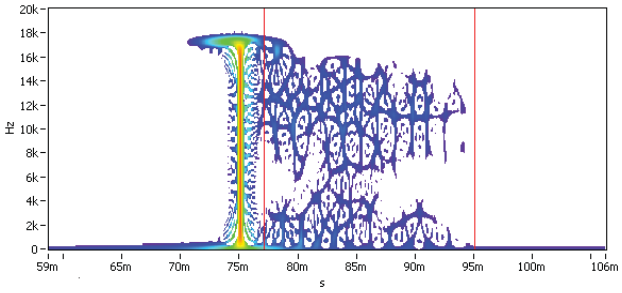

Figure 25 MP3 Encoded Dirac Showing Band Limiting, Pre-Ringing and Quantization Noise (SPWVD)

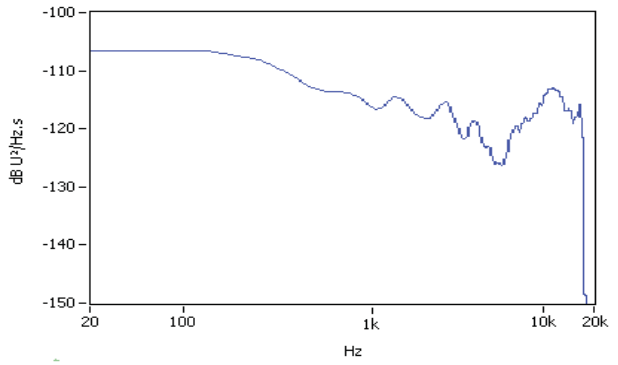

Figure 26 Energy Spectrum of the Quantization Noise

3.3.4. VoIP

Voice over IP (VoIP) is another form of digital communication that is interesting to explore with Time-Frequency Analysis. VoIP systems encode, band limit and noise gate input signals. Time-Frequency Analysis yields a complete picture of the effects of the VoIP system on input signals.

For example, a stepped sine with a frequency range of 20 Hz to 20 kHz was played and recorded through a VoIP system. The Time-Frequency map (Figure 27) shows that the signal is band limited from 75 Hz to 8 kHz. The map also shows that the input signal is aliased to 12 kHz and 16 kHz with a level 75 dB lower than that of the original signal. This indicates that that the anti-aliasing filters are of poor quality.

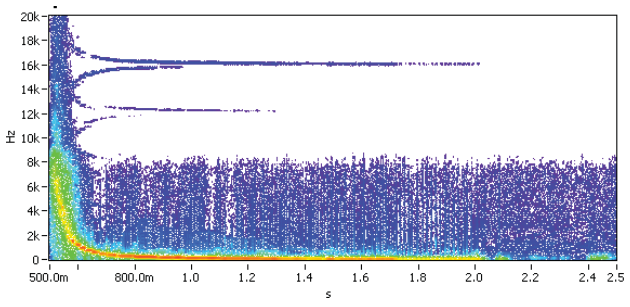

Figure 27 Stepped Sine Recorded using VoIP (SPWVD)

Another interesting example signal is the Composite Source Signal (CSS). Many VoIP protocols use noise gating to avoid the transmission of noise. CSS signals provide a way to circumvent the noise gate. The signal consists of a preconditioning signal that imitates voice, followed by pseudo random white noise. The voice segment of the CSS signal is used to open the noise gate of the VoIP system. The noise segment of the CSS signal is then transmitted before the gate is able to close. This allows the testing of VoIP systems using noise as a stimulus.

The Time-Frequency map of a VoIP recorded CSS signal (Figure 28) shows that the noise section of the stimulus successfully passed through the VoIP codec.

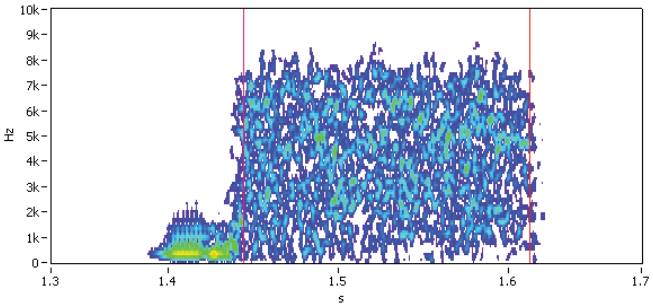

Figure 28 CSS Signal Recorded using VoIP (SPWVD)

The frequency response of the VoIP system is then determined by isolating the stationary section of the noise signal and calculating its energy spectrum (Figure 29).

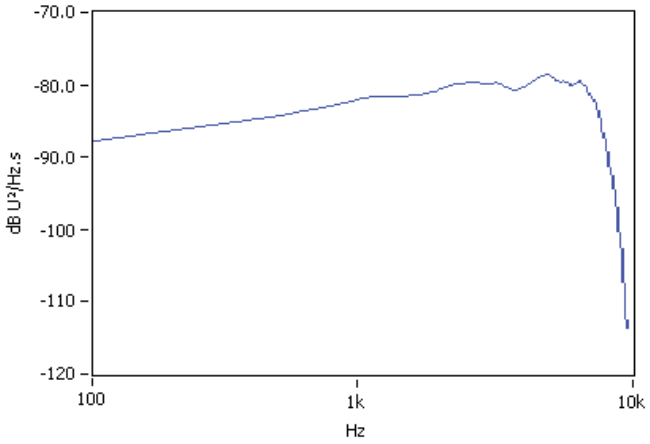

Figure 29 VoIP Frequency Response Calculated from a CSS Signal

The final test signal sent through the VoIP system was a Dirac. The Time-Frequency map (Figure 30) shows that the Dirac was converted to a pulse of noise by the VoIP system. There is noticeable ringing at 8 kHz and at approximately 100 Hz.

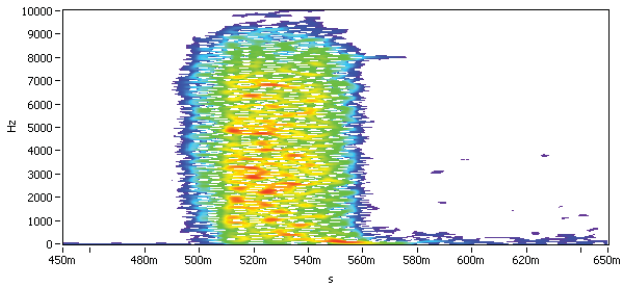

Figure 30 Dirac Recorded using VoIP (SPWVD)

3.4. Comparison of the Four Distributions

The STFT is the simplest of the distributions both in its use and its interpretation, but it suffers from the Heisenberg limit and therefore cannot show fine details of a signal. Because of its relative computational simplicity, it is a useful tool to obtain a quick overview of long signals.

The CSD is a highly specialized transform that is only useful to study resonances in impulse responses. It offers a simple analysis but can exaggerate ringing. Moreover it does not reveal group delay and instantaneous frequency. Its easy physical interpretation makes it a useful "sanity check" when using other methods.

The SPWVD is more complex to set up as smoothing is an additional parameter. It is more difficult to interpret because interferences can be confusing, even if the marginal distributions (time envelope, energy spectrum) help to ignore the interferences. However the independent choice of time and frequency smoothing results in greater flexibility and increased resolutions beyond the Heisenberg limit. This makes it the most precise of the transforms presented. It is a great complement to STFT allowing focus on short events and it is especially useful for impulse response studies.

The CWT differs from other methods in that it uses constant percentage frequency resolution rather constant resolution. Compared to STFT, this offers better time resolution at high frequencies and better frequency resolution at low frequencies. It is more psychoacoustically significant since human hearing works more on a logarithmic than linear scale and it is easy to see the entire 20-20kHz range in one picture. For Impulse Response analysis, its low frequency resolution is a plus but the spreading of low frequency energy along the time axis can make it more difficult to interpret. When used for impulse response analysis, it is advisable to have a long impulse signal in order to accurately analyze low frequencies (e.g. 1 s for 20 Hz resolution 1/3 octave).

3.5. Future Developments

Future development could focus on: Zoom Analysis, Affine WVD Smoothing, and Reassignment.

Zoom analysis would allow the user to focus on a restricted bandwidth of the signal. This would be useful for jitter analysis or to examine the low frequency content of a signal.

Affine WVD Smoothing [17] combines the flexibility and resolution of the SPWVD with a constant percentage bandwidth analysis. This method allows a gradual transition between WVD and Wavelet, the same way that SPWVD can morph into STFT. That method permits going beyond the Heisenberg Limit while keeping log frequency resolution.

Reassignment [18] improves the readability of time-frequency distributions in general by concentrating energy towards the center.

3.6. Conclusion

As usual, no one analysis tool can do it all; each time-frequency transform has its trade-offs. Some are better for certain applications (e.g. CSD is good and SPWVD is even better for analyzing transducer impulse responses). Some are easier to use, such as STFT which is very similar to traditional FFT analysis. Some transforms are more precise and more flexible but typically require more careful use. For example, the Wavelet transform is very useful for analyzing wide bandwidth signals over a long time period but looks quite different visually, and SPWVD may be difficult to interpret because of interferences. In general, it is best to try several different transforms on the same signal to see which one works best or gives different insight to what is happening in the signal.

Time-frequency analysis encompasses practically all aspects of signal analysis and is very similar to how the human brain interprets sound. Time-frequency 3D plots and intensity maps offer a quick overview of a signal both in the time and frequency domains, simultaneously, and make it easier to spot problems. The use of more traditional 2D graphing tools to slice or power sum in either the time or frequency domain,

enables the user to closely examine features and focus on possible cause and effect of these problems.

4. REFERENCES

[1] J. Ville, "Théorie et Applications de la Notion de Signal Analytique", Câbles et Transmission, 2e Année, No 1, pp.61-74, 1948.

[2] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, "The Wigner Distribution – A Tool For Time-Frequency Signal Analysis, Part I: Continuous-Time Signals", Philips J. Res., Vol.35, pp 217-250, 1980

[3] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, "The Wigner Distribution – A Tool For Time-Frequency Signal Analysis, Part II: Discrete-Time Signals", Philips J. Res., Vol.35, pp 276-299, 1980

[4] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, "The Wigner Distribution – A Tool For Time-Frequency Signal Analysis, Part III: Relation With Other Time-Frequency Signal Transformations", Philips J. Res., Vol.35, pp 372-389, 1980

[5] C.P. Janse and A.J.M. Kaiser, "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution", J. Audio. Eng. Soc., Vol.31, No 4, April 1983

[6] L. Cohen, "Time-Frequency Distributions- A Review", Proc. IEEE, Vol.77, No. 7, July 1989

[7] F. Hlawatsch and G.F. Boudreaux-Bartels, "Linear And Quadratic Time-Frequency Signal Representations", IEEE Signal Proc. Magazine, April 1992

[8] P. Flandrin, "Temps-Fréquence", Série Traitement du Signal, Hermès, Paris, 1993

[9] L. Cohen, "Time-Frequency Analysis", Signal Processing Series, Prentice Hall, 1995

[10] G. Kaiser, "A Friendly Guide to Wavelets", Birkhaüser, 1999

[11] L. Cohen, "The Scale Representation", IEEE Trans. on Signal Proc., Vol.41, No. 12, December 1993

[12] P. Flandrin and B. Escudié, "Principle And Application of Time-Frequency Analysis by Means of the Wigner-Ville Distribution", Revue Traitement du Signal, Vol.2, No2, pp 143-151, GRETSI 1985

[13] J.M. Berman and L.R. Fincham, "The Application of Digital Techniques to the Measurement of Loudspeakers", JAES, vol. 25, pp.370-384, June 1977, Appendix II

[14] W. Martin and P. Flandrin, "Wigner-Ville Spectral Analysis of Nonstationary Processes", IEEE Trans. On ASSP, Vol.33, No6, December 1985

[15] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique", AES 108th Convention, 2000, Preprint 5093

[16] C. Struck and S. Temme, "Simulated Free Field Measurements", JAES, Vol. 42, No 6, 1994 June

[17] P. Flandrin And O. Rioul, "Affine Smoothing of the Wigner-Ville Distribution", IEEE ICASSP, pp. 2455-2458, Albuquerque, April 1990

[18] F. Auger and P. Flandrin, "Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method", IEEE Trans. on Signal Proc., Vol.43, No5, May 1995