Interest & Utility of Time Frequency and Time Scale Transforms in the Partial Discharges Analysis

Cyrille Caironi(1)(2)(4) David Brie(3) Lionel Durantay(1) Abderrezak Rezzoug(2)

(1) Alstom moteurs.4, rue de la Rompure54250 Champigneulles, France, cyrille.caironi@powerconv.alstom.com
(2) GREEN, UHP Nancy 1, BP 239, 54506 Vandoeuvre les Nancy, abderrezak.rezzoug@green.uhp-nancy.fr
(3) CRAN, UHP Nancy 1, BP 239, 54500 Vandoeuvre les Nancy, david.brie@cran.uhp-nancy.fr
(4) ERIN, UHP, Esstin, 2 rue Jean Lamour 54500 Vandoeuvre

Abstract: **The development of Time-Frequency and Time-Scale Transforms has been progressed quickly for 20 years. This development and the progression of the Numerical acquisition enable the utilization of these mathematical tools to analyze Partial Discharges (PD) in the induction motors insulation. Some papers have been written about the Wavelet tools in order to analyze the partial discharges. This paper is attended as a tutorial overview of the Time Frequency and Time Scale Transforms applied to the PD analysis on the stators of the induction motors. In the first part two Time-Frequency Transforms are introduced : Short Time Fourier Transform and Wigner Distribution. The mathematical tools for the implementation of the STFT are given. An example of extraction of a PD signal from numeric noise is given and its results discussed. In the second part the Wavelets are presented on the same case study as for the Time Frequency Transforms.**

INTRODUCTION

Statistical tools and Phase Resolved Partial Discharges (PRPD) Pattern are currently used to analyze the PD [1]. The importance of frequency data for the PD analysis [2] and the problems of noise in the PD measurement are both principal reasons to work in the frequency domain. The classical transforms like the fast Fourier transform (FFT) enable us to go from time domain to frequency domain and inversely, but, don't enable to work on both domains. During the 20th century, Time-Frequency Transforms were developed. The first one is the Short Time Fourier Transforms (STFT). This transform is frequently used, especially for the voice treatment. It is easy to implement and gives good results. We present also the Wigner distribution and its advantages and disadvantages. The latest tool is Wavelet or time-scale-transform. It appeared around 20 years ago.

The first part of this paper presents the time frequency transforms and gives an example to illustrate it. The second part gives examples of single PD treatment and gives results of PD signal denoising. The last part of this paper presents the Wavelet Transforms and discuss it on the same case studies as Time Frequency Transforms.

THEORY AND IMPLEMENTATION OF TIME-FREQUENCY TRANSFORMS

The FFT enables to analyze the signal in the time domain or in the frequency domain. The transforms that we describe in this chapter enable to work in both domains. To do that we work on the complex form of the signal. This form is

$$\tilde{s}(t) = s(t) + jH[s(t)] \tag{1}$$

where H[s(t)] is the Hilbert transform of the real signal.

We present the Short Time Fourier Transform (STFT) and the Wigner Distribution (WD). They are described with test signals to evaluate them and to make a choice for the PD analysis. The signals chosen to do this test are sum of sinusoids and linear chirp :

$$x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t) \tag{2}$$

where f1=200 Hz and f2=400 Hz

$$y(t) = \sin(2\pi f t) \tag{3}$$

where f ranges from 50 Hz to 350 Hz. The last test signal is the sum of x and y :

$$z(t) = x(t) + y(t) \tag{4}$$

Short Time Fourier Transform

To study the properties of a signal at time t, one emphasizes the signal at that time and suppresses the signal at other times.

This is achieved by multiplying the signal by a window function [3], centered at t, to produce the modified signal.

$$s_t(\tau) = s(\tau).h(\tau - t) \quad (5)$$

The modified signal is a function of two times, the fixed time we are interested in, t, and the running time, τ. The window function is chosen to leave the signal more or less unaltered around time t. every point outside this window is put at zero to suppress the signal for time distant from the time of interest.
The term "window" comes from the idea that we are seeking to look at only a small piece of the signal as when we look out of a real window and see only a relatively small portion of the scenery. In this case we want to see only a small portion.
Since the modified signal emphasizes the signal around time t, the Fourier transform will reflect the distribution of the frequency around that time,

$$S_t(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau}.s_t(\tau).d\tau \quad (6)$$

where ω is the pulsation of the signal, ω=2πf (f is the frequency in Hz). Following the formula (1),

$$S_t(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau}.s(\tau).h(\tau - t).d\tau \quad (7)$$

the energy density spectrum at time t is therefore

$$P_{SP}(t,w) = |S_t(w)|^2 = \left| \frac{1}{\sqrt{2\pi}} \int e^{-jw\tau}.s(\tau).h(\tau - t).d\tau \right|^2 \quad (8)$$

for each different time we get a different spectrum and the totality of these spectra is the time-frequency distribution, P_{SP}. It goes under many names, depending on the field. In this paper, we use the most common phraseology, "spectrogram". The figures 1b, 2b and 3b give examples of test signal defined as above.

Since we are interested in analyzing the signal around the time t, we have chosen a narrow window peaked around t. Hence the modified signal is short and its Fourier transform, equation (7), is called Short-time Fourier transform. When we want to estimate time properties for a given frequency we do not take short times but long ones. In this case the windows function on the frequency is short to respect the uncertainty principle of Heisenberg. We can define a Short Frequency Fourier Transform by

$$s_\omega(t) = \frac{1}{\sqrt{2\pi}} \int e^{jwt}.S(\omega').H(\omega' - \omega).d\omega' \quad (9)$$

If we relate the window function in the time h(t) and the windows function in the frequency H(ω) by

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int h(t).e^{-j\omega t} dt \quad (10)$$

then

$$S_t(\omega) = e^{-j\omega t}.s_\omega(t) \quad (11)$$

The short time Fourier transform is the same as the Short Frequency Fourier Transform except for the phase factor $e^{-j\omega t}$. Since the distribution is absolutly square, the phase factor does not enter into and either the short-time Fourier transform or the Short Frequency Fourier Transform can be used to define the distribution,

$$P(t,w) = |S_t(\omega)|^2 = |s_\omega(\omega)|^2 \quad (12)$$

this shows that the spectrogram can be used to study the behavior of time properties at a particular frequency. This is done by choosing a narrow window in frequency or likewise by taking a broad window in time. Therefore we can make a narrowband or a wideband spectrogram by choosing size of the window in the time domain.

Wigner Distridution

The Wigner distribution (WD) is the prototype of distributions that are qualitatively different from the spectrogram. Wigner's original motivation for introducing it was to be able to calculate the quantum correction to the second virial coefficient of a gas. The Wigner distribution was introduced into signal analysis by Ville. The Wigner transform in terms of the signal is

$$W(t,\omega) = \frac{1}{2\pi} \int s^*(t - \frac{\tau}{2}).s(t - \frac{\tau}{2}).e^{-j\omega\tau}.d\tau \quad (13)$$

The Wigner distribution is said to be bilinear in the signal because the signal enters twice in its calculation. It is to be noticed that to determine the properties of the Wigner distribution at a time t we mentally fold the left part of the signal over the right one to see if there is any overlap.

As we can see on the signal y(t) defined by the equation 3 (figure 2a) the Wigner distribution (figure 2c) has a better contrast than the Short Time Fourier Transform (figure 2b). It is interesting for the strongly non stationary signals. The figure 1 presents the sum of sinusoids x(t) defined by the equation 2. The STFT gives only both frequencies but with broad band. The WD gives the both frequencies with a narrow band but we can see a lot of interference terms especially in the band of 300 Hz. These terms of interference are mathematical artifacts generated by the distribution.

The sum of both signals is defined as the equation (4) of the signal z(t). Its Wigner distribution is

$$W_z(t,\omega) = W_x(t,\omega) + W_y(t,\omega) + W_{xy}(t,\omega) + W_{yx}(t,\omega) \quad (14)$$

The cross Wigner distribution is complex. However, $W_{xy} = W_{yx}^*$ and, therefore, $W_{xy}(t,\omega) + W_{yx}(t,\omega)$ is real. Hence

$$W_z(t,\omega) = W_x(t,\omega) + W_y(t,\omega) + 2.Re\{W_{xy}(t,\omega)\} \quad (15)$$

We see that the WD of a sum of two signals is not the sum of the WD of each signal but has the additional term2.Re{Wxy(t,ω)}. It is our term of interference.

On one hand the signal z(t) (equation (4) figure 3a) is easily readable with the spectrogram (figure 3b). On the other hand the artifacts generated by the terms of interference blur the information on figure 3b. the work inside the classes of Cohen and the treatment of the ambiguity function enable to reduce this problem. But it is not the objective of this paper.

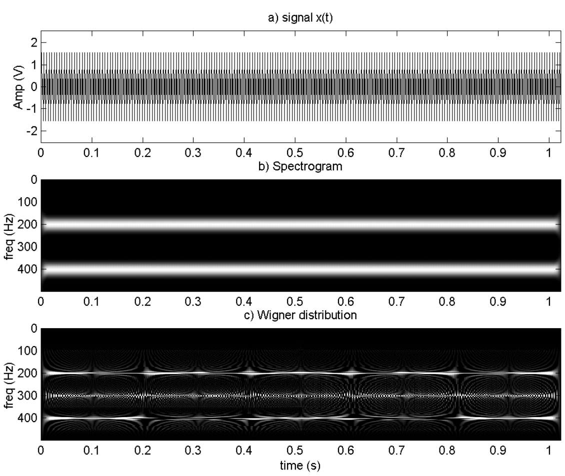

Figure 1 : Spectrogram and Wigner distribution of the signal x(t)

We can see that the easier tools in the time frequency is the Time Frequency space is the Short Time Fourier Transform. If the results of this transform are enough, it isn't necessary to use others. But, we have to remember the disadvantages of this transform. The first one is the non orthogonality of the transform and the impossibility of perfect reconstruction of the signal. The second one is the difficulty of the choice of the windows.

Nevertheless, we decided to use Short Time Fourier Transform for the Partial Discharges Analysis. Indeed, we had the better results with this tool in term of facility and readability.

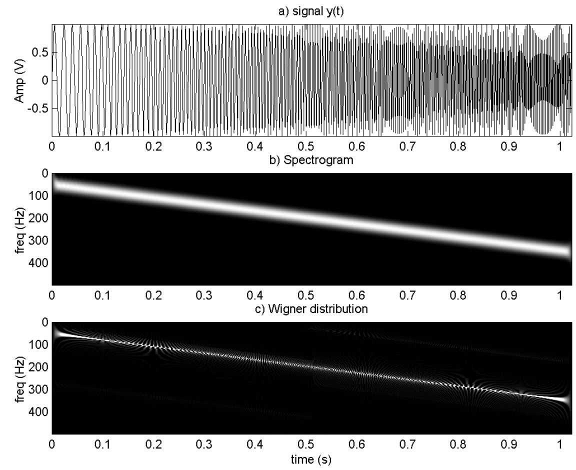

Figure 2 : Spectrogram and Wigner distribution of the signal y(t)

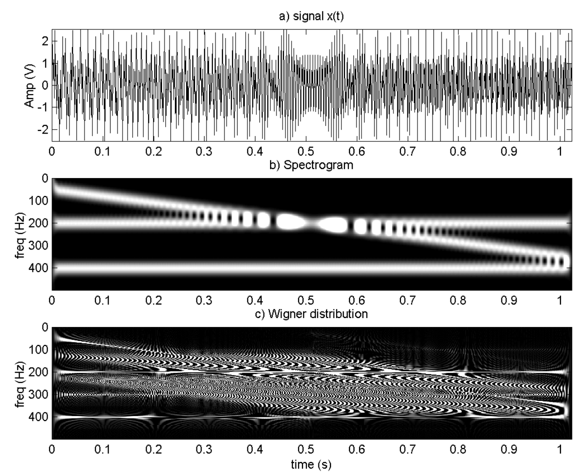

Figure 3 : Spectrogram and Wigner distribution of the signal z(t)

Single PD in the Time-frequency Space

The first application that we propose is the study of one single PD in the time-frequency space. This PD was measured by a capacitive probe on one sample composed of a polyester-mica tape. The details of the sample are not so important in this paper. The amplitude of the noise around the PD is relatively small (figure 4). The PD is measured with a frequency of 250 MS/s on 256 samples. The bandwidth of the PD is around 70 MHz. The STFT (figure 5) presents a readability until 60 MHz for two reasons, first one is the quality of the graphic, second one is the choice of the window. The WD is more efficient (figure 6). The frequency and the position on the time are more readable. On the figure 6 we can see the interference terms near the self terms. The sum of the interference terms is null. The noise is not readable on the figure 5 and 6 because its amplitude is so less important than the PD amplitude.

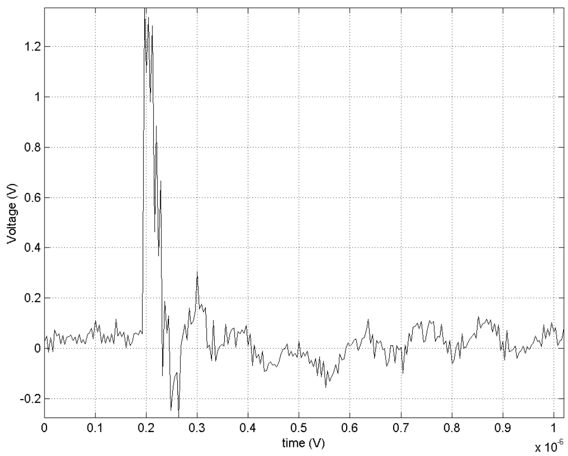

Figure 4 : Single PD Signal

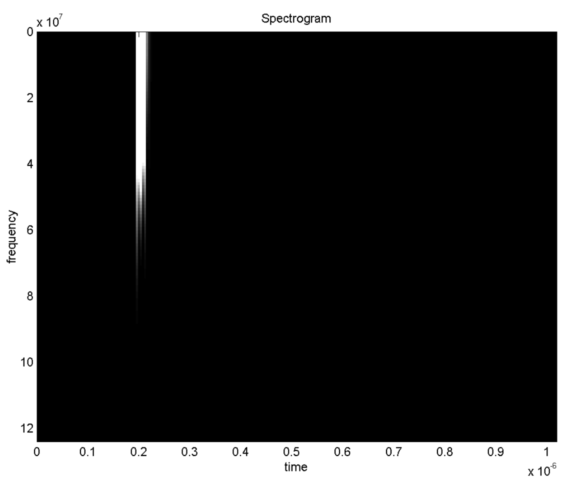

Figure5 : STFT of the single PD single

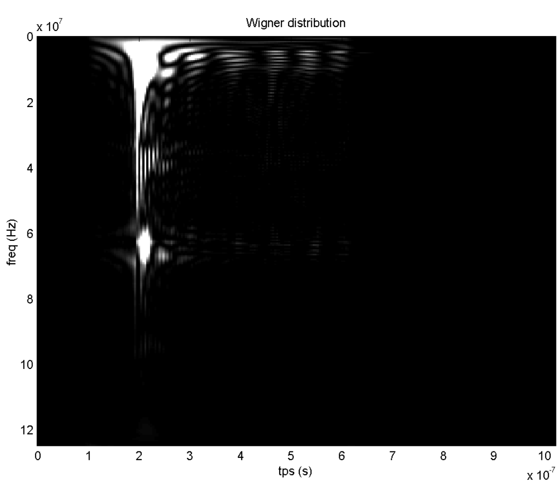

Figure 6 : WD of the single PD single

Time frequency PD analysis on one supply period

One classical process to analyze the partial discharge is the examination of the PD on one supply period. The figure 8 presents the PD measurement on the same sample as the single PD measurement. This measurement is made at 34 MS/s. Thanks to the single PD analysis, we know that this frequency band is not adequate and the PD signal has a lot of chance to look like Dirac signal.

To test the efficiency of the time-frequency transform, we add numeric noise at the measured signal (figure 9). The added signal is composed by white noise, sinusoidal noise and choke noise. From this new signal we must analyze and denoise it with mathematical tools in order to obtain as well as possible the original signal (figure 10).

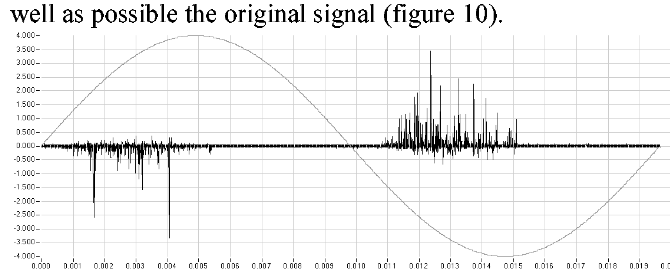

Figure 8 : PD signal measured

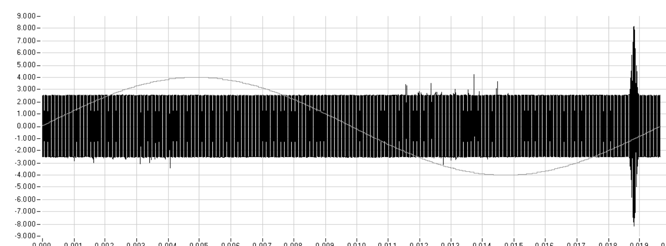

Figure 9 : Noised signal

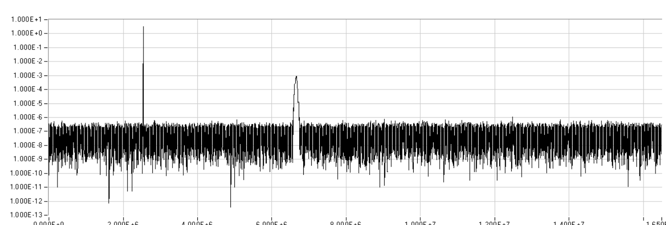

Figure 10 : FFT of the noised signal

On the spectrogram (Figure 11), we can see different zones (even if it is difficult to read a spectrogram in black and white). The bands in white on the zone N° 1 are periodic components of the signal they are present at every point of the time. We can see them in the FFT (Figure 10). We can eliminate them by filtering. The zone N° 2 is the chock. We can also see it in the spectrum but the best solution to eliminate it is in the time domain. It is the same for the white noise (zone N° 3). The zone N° 4 contains white noise and PD signal.

The figure 12 presents the result of the denoising. This result is obtained from recursive calculations in time domain and frequency domain. The denoised signal obtained can be used to do statistical treatment like PRPD pattern [1]. Of course, we loose one part of the information. But, in this case, we loose less information by numerical treatment than by the frequency band of the acquisition card.

The time-frequency distribution enables to make an analyze to determine the best tools to extract the signal from the noise. After this analysis, we can implement the tools used for the extraction on the machine to make an on line treatment. However, the time-frequency transform are difficult to implement on line because the time of treatment can be too long.

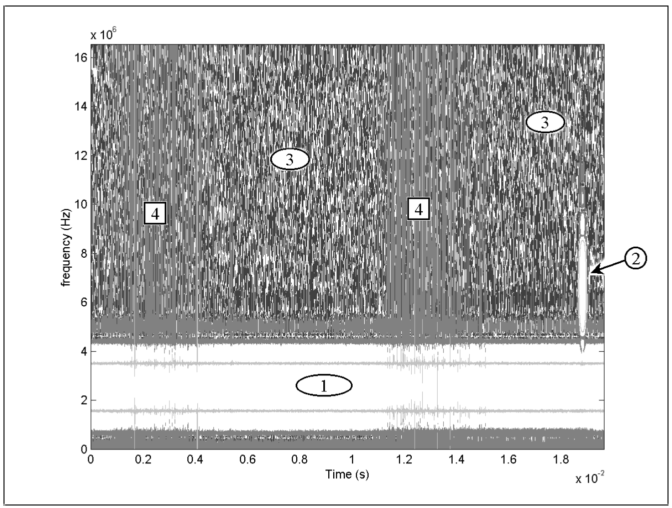

Figure 11 : Spectrogram of the Noised Signal

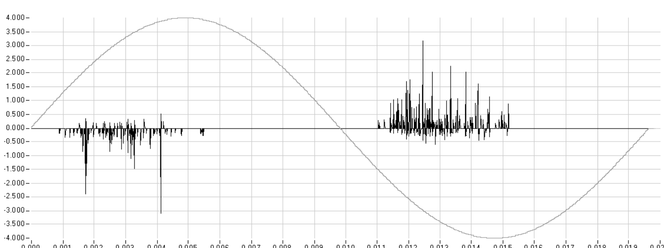

Figure 12 : Denoised signal

Automatic filtering

The work to obtain the denoised signal of the picture 12 is so heavy and it isn't a solution for the On line measurement. In order to we built an automatic filter which use the time frequency duality. This filter is adaptive filter optimized by simplex method. This type of filter already exists [10].

THEORY AND IMPLEMENTATION OF TIME-SCALE TRANSFORMS

The main difference between the Time-Frequency transforms and the time-scale transforms (or Wavelet Transforms) is treatment of the analyzing function the signal. While the Time frequency compromise between time and frequency information can be useful, the drawback is that once we choose a particular size for the time window, that window is the same for all frequencies. Many signals like PD require a more flexible approach. For example one where we can vary the window size to determine more accurately either time or frequency. It is one of the advantages of the Wavelet Analysis: a windowing technique with variable sized regions. A Wavelet is a waveform of effectively limited duration that has an average value of zero.

Continuous Wavelet Transform

The continuous Wavelet transform is the scalar product of the signal with a scaled and positioned function :

$$X(a,b) = \frac{1}{\sqrt{a}} \int x(t).\psi\left(\frac{t-b}{a}\right).dt \qquad (16)$$

Where ψ is the mother wavelet, a is the scale coefficient and b the position coefficient.
We lose the frequency reference because we don't work about the frequency of the analyzing function but on the size of the window. however we can still estimate the frequency by the value of the coefficient a. If a is low the wavelet is compressed and the frequency is high, if a is high the wavelet is stretched and the frequency is low. We can see on figure 13 that the readability of the graphic representation of a continuous wavelet is worse than the time frequency transforms

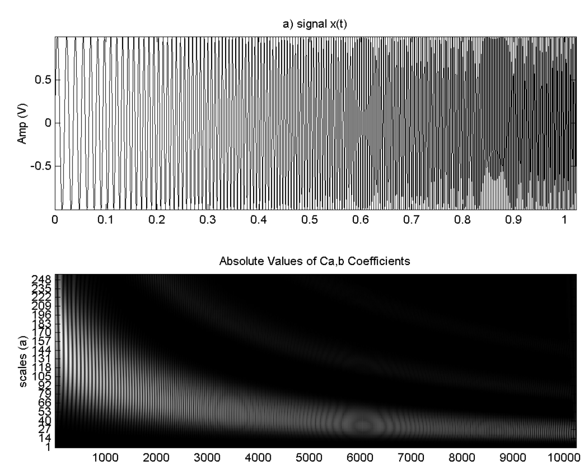

Figure 13 : continuous Wavelet transform of the signal y(t)

But the main utility of the Wavelet transform in signal processing isn't the graphic presentation. Bigger problem of the continuous Wavelet transform is the redundancy. Calculating Wavelet Coefficients at every possible scale is a fair amount of work and it generates an awful lot of data.

Discrete Wavelet Transform

In order to eliminate the redundancy we choose a exponential method in order to make the discrete transform :

$$\begin{cases} a = a_0^m \\ b = n.b_0.a_0^m \end{cases} \qquad (17)$$

Therefore, the discrete Wavelet transform is :

$$X(m.n) = \int x(t).\psi(a_0^{-m} - nb_0).dt \qquad (18)$$

In this paper we use the dyadic case where $a_0 = 2$ and $b_0 = 1$.
If the transform is discrete, the mother Wavelet ψ stay continuous. There are a lot of Wavelet: Morlet, Haar, Daubechie, Mexican hat, etc… In our study we used only Haar and Daubechie (it is interesting to note that the Haar function is just the Daubechie function of order 1).

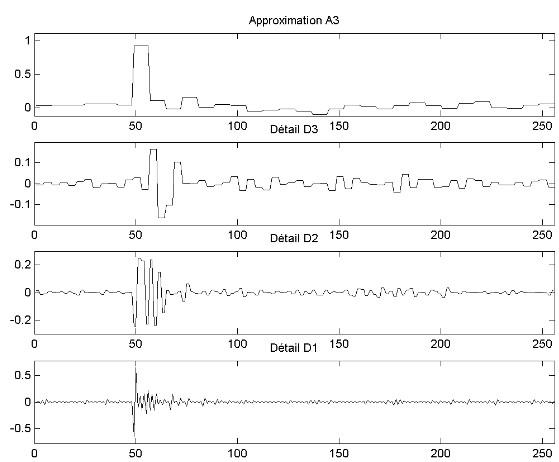

Figure 14 :Decomposition of the individual PD (figure 4) with Haar Wavelet

The interest of this type of transform, if the mother Wavelet is orthogonal, is the possibility to make the partial or total reconstruction of the signal. The figure 14 presents the decomposition of the individual PD of the figure 4. The level 3 of the approximation shows a good representation of the PD. This type of information is useful for the noise rejection and for a statistic work.

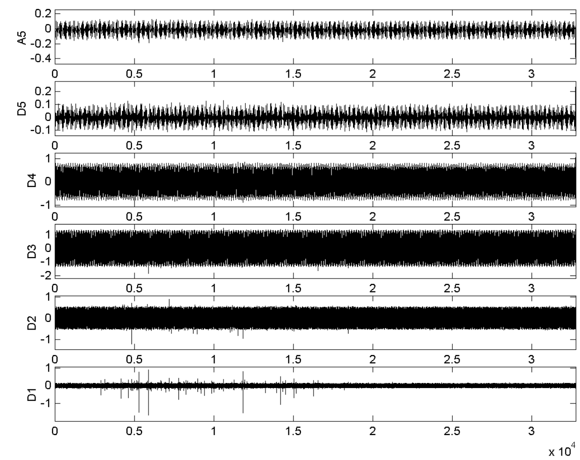

Figure 15: Decomposition of the positive phase of the noised signal (figure 9) with an order 2 Daubechie Wavelet.

In order to do the decompositions we use Mallat algorithm [xx] with different Wavelets. The figure 15 presents the decomposition on 5 level of the positive phase of the noised signal (figure 9). We have to work on every coefficient to extract the PD signal. This work is more or less similar at the time frequency extraction.

The choice of the mother wavelet is important for the analysis. The figure 16 presents the level 1 of detail reconstructed. The decomposition with an order 4 Daubechie Wavelet more readable than the decomposition with Haar Wavelet. The extraction of the signal versus noise will be easier with a Daubechie 4.

Another utility of the wavelet is the 2D analysis in order to make a statistical analysis. Presently we aren't working on this type of analysis.

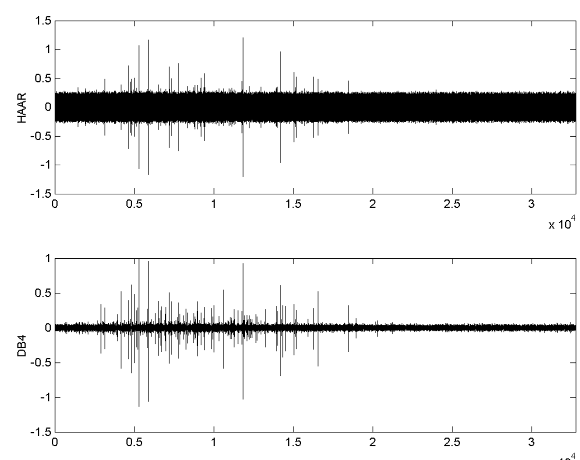

Figure 16: Comparison of decomposition between Haar and Daubechie 4 wavelets

CONCLUSION

In this paper we have presented and discussed time-frequency and time-scale transforms. We used this type of transform mainly to do signal versus noise extraction.

We don't use this type of transforms to make a default diagnostic. Even if this type of work already exists [10]. We need an industrial validation of the numerical tools to use them and we have only validated the noise rejection process. The next step may be the automatic diagnostic.

REFERENCES

[1] Fruth B., Niemeyer L., "The importance of the statistical characteristics of partial discharge data", IEEE Transaction on dielectrics and electrical insulation, Vol. 27, N° 1, Feb. 1992.

[2] Stone G, "Importance of the Bandwidth in PD measurement in operating motors and generators" IEEE Transaction on dielectrics and electrical insulation, Vol. 7, N° 1, Feb. 2000.

[3] Cohen L., "Time-frequency analysis", Prentice Hall, 1995

[4] Truchetet F., "Ondelettes pour le signal numérique", Hermes, 1998

[5] Caironi C., " Contribution a la maintenance prédictive des machines tournantes par mise en place d'outils de surveillance et d'évaluation des décharges partielles", Thesis IDPE, 25 oct. 2000.

[6] E.M. LALITHA, L. SATSH, "Wavelet Analysis for classification of multi-source PD Pattern", IEEE Transaction on dielectrics and electrical insulation, Vol. 7, N° 1, Feb. 2000.

[7] R. BARTNIKAS, "Corona measurement and interpretation" Engineering Dielectrics Vol. 1, STP 669, 1979

[8] Caironi C. and Brie D., "Contribution of Time-Frequency and Time-Scale transforms to the partial discharges analysis" Proc. IRMC'01 June 25-28, 2001, Alexandria USA.

[9] Mallat S., "A wavelet tour of signal processing" Academic press, 1999.

[10] Cavallini A., Conti M., Contin A., Montanari G.C., "Automatic Noise rejection Tools in On-Field PD Measuremant" Volta colloquium on Partial Discharges..