Spectrum Analysis

Introduction

Why measure a spectra?

In electrical engineering we are most often interested how a signal develops over time. For this time-domain measurement we use the Oscilloscope. Like the rise time of pulse. However there are also many measurements where the frequency-domain measurement is more useful. For instance we can examine in which frequency bands the electrical noise lies, or we can quantify how distorted a sinusoidal signal is, as these have components at integer multiple frequencies of the signal frequency (*harmonics*). Further it can be important to know in telecommunication that a transmission of data lies only in a small frequency band designated for this purpose not disturbing other transmission channels, like different radio broadcasting channels.

In this script we will examine the basic design of spectrum analyser which should help us to make best use of this measurement tool.

Basic concepts of spectrum analysers

There are two basic concepts for designing a spectrum analyser. The first one consists of measuring the signal directly in the frequency range by the principle of superposition and with help of a band-pass filter. This we refer to a *swept-tuned* or *superheterodyne* spectrum analyser. The second type digitally samples the data in the time-domain and then mathematically transforms it into the frequency domain by the FFT (Fast Fourier Transform) Algorithm. In this concept the signal is processed in digital form from relatively early on in the measurement chain in contrast to the swept tuned analyser which is only converted to digital near the end, or not at all with old analysers. Generally the swept tune analysers are more sensitive and have a higher frequency range, but FFT- Analysers also have advantages. For example *single shot* phenomena which can not be examined with the *swept tuned* Analysers. Also with the current improvements made in the field of digital signal processing, FFT-Analysers gaining on the swept-tune Analysers. Often modern spectrum analysers are a combination of the two concepts.

Swept-Tuned Spectrum Analysis

The basic concept

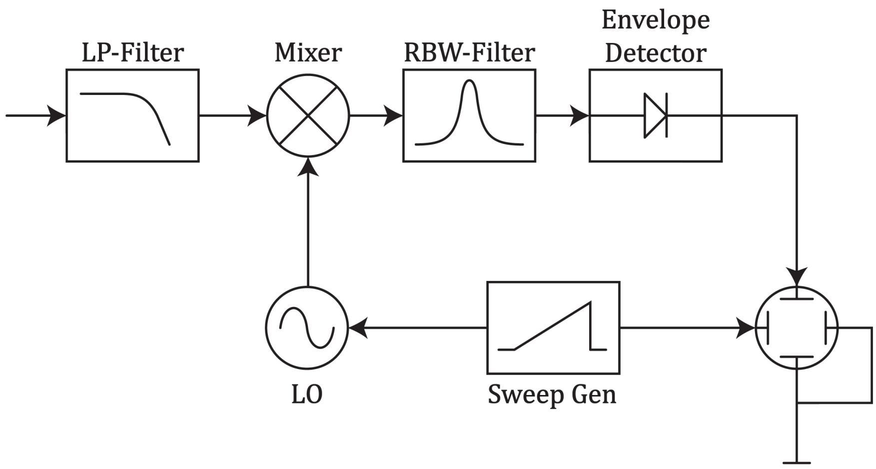

Figure 1: Block diagram of swept tuned analyser

As already mentioned the *swept-tuned* or *superheterodyne* spectrum analysers examines the Signal directly in the frequency domain. Theoretically one could build an spectrum analyser by using a bandpass-filter with tunable passband frequency. By Sweeping the frequency of the filter through the whole frequency band of interest one can associate the Amplitude at the output of the filter to a frequency value. As the demanded characteristics of such a tunable filter is technically not achievable, one goes the opposite way by "sweeping" the signal frequency trough the frequency band of interest and keeping the band pass frequency constant.

Figure 1 shows the basic block diagram of a *swept-tuned* spectrum analyser. In this design the signal frequency is mixed with a internally generated frequency in the analyser (*Local Oscillator*) creating mixing products. The two mixing products of greatest amplitude appear at $f_{LO}+f_{Sig}$ and $f_{LO}-f_{Sig}$. The lower one of these two frequencies we can make use for our analysis. By sweeping f_{LO} we can sweep the mixing product past a fixed frequency bandpass-filter. As we know the current frequency of the *Local Oscillator* and we know the the passband frequency of the filter we can calculate the frequency of the signal observed at the output of the bandpass filter $f_{IF}=f_{LO}-f_{Sig}$. By sweeping the f_{LO} over the whole span we can plot the frequencies of all the signals in the range of the analyser resulting in a spectrum.

On the display of the *spectrum analyser* we would like to see the maximum Amplitude of a sinusoidal signal at a certain frequency. The signal at the output of the *resolution bandpass filter* however is the instantaneous value of a high frequency sine wave. So the average of this is zero. In other words you first have to rectify and lowpass filter the signal to display just the envelope of the signal. This is what the *envelope detector* does.

Mixing the signals

A *frequency mixer* is a device that combines two signals. Mathematically speaking an ideal mixer multiplies the signals with each other in contrast to an additive mixer where the instantaneous values of the signals are added together. If you look at the spectrum of the additive mixer we will see that both frequencies of the original signals are still existent in the output signal. In the *frequency mixer* however you will see mixing products at $f_{out}=f_1 \pm f_2$ and weaker higher order products $f_{out}=2 \cdot f_{1,2} \pm f_{2,1}$ etc. In a real world mixer where the mixing is achieved with nonlinear components you will also have to contend with the harmonics of the original signals as well as the harmonics of the mixing products. So $2 \cdot f_1, 2 \cdot f_2, etc.$ as well as the original input signals them self at the output.

Electrically the mixing is done by sending the signals through a nonlinear device, like in a most basic design a diode. This means that low voltage signal values are amplified less (or here dampened more) than high voltage signal values. The result is a multiplication of the signals.

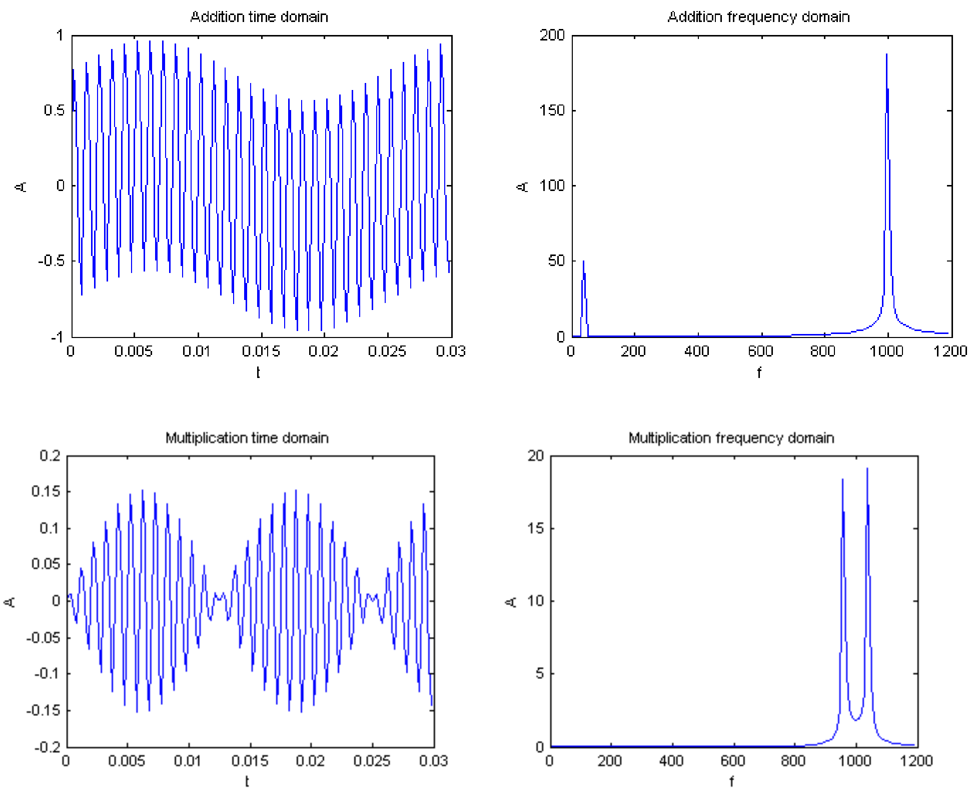

Figure 2: Addition and Multiplication of frequencies in time and frequency domain

Translating the frequency

By mixing the signal with a higher f_{LO} we can work in a higher frequency range were it is easier to build a narrow band *resolution bandwidth filter* tuned to a fixed frequency the *Intermediate frequency*. By sweeping the f_{LO} the mixing product of different input signal frequencies appear at the output of the *resolution bandwidth filter.* As for every moment we know the frequency value of the *LO as it is* set by a slowly increasing ramp signal we can use the current value of the ramp to define the x-axis of the signal point on the screen while the output of the *resolution bandwidth filter defines the y-value.*

The *intermediate oscillator* frequency is chosen higher than the topmost input frequency of the analyser.

$f_{LO} > f_{Sig,max}$

Where the *local oscillator* frequency sweeps from $f_{IF} + BW$ till

$f_{IF} + 2 \cdot BW$, BW being the bandwidth of the analyser input.

To prevent unwanted frequencies appearing at the entrance of the *resolution bandwidth filter* it is necessary to cut off any signal above the maximum input frequency with a lowpass filter before the mixer. Otherwise even with a perfect mixer, frequencies at the Input higher than f_{LO} would be mixed down to f_{IF} as mixing also works the other way round:

$f_{IF} = f_{Sig} - f_{LO}$ This is undistinguishable by the analyser from a frequency in the frequency band which it was designed for.

Resolution BW

Resolution Bandwidth Filter is set to a fixed Frequency the *Intermediate Frequency.* The Bandwidth can be adjusted, this determines the resolution of the Analyser. A small Bandwidth allows for a higher resolution, but needs more swept time, as the filter has to have enough time to tune in for a precise measurement. If the Bandwidth is set to wide, signals with similar frequencies can be overseen, especially if they have a big difference in Amplitudes. In Figure 3 and Figure 4 this effect is shown.

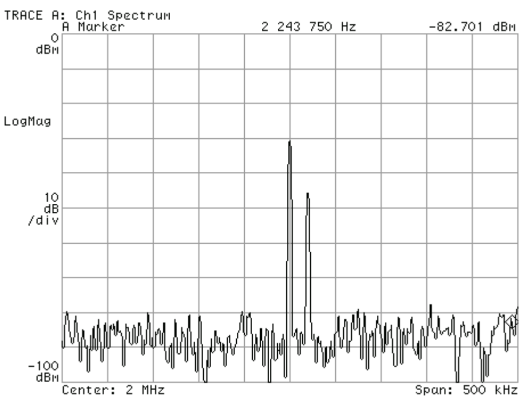

Figure 4: RBW-filter set to 30 kHz bandwidth

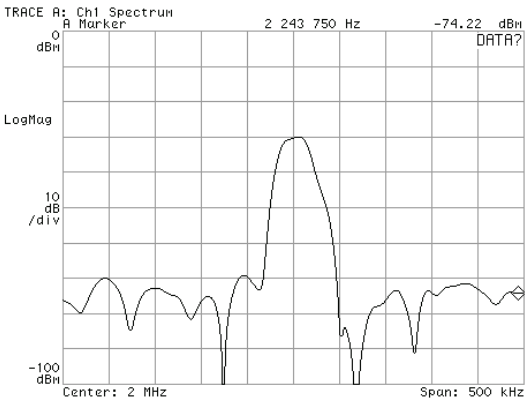

Figure 3: RBW-filter set to 30 kHz bandwidth

The
The

The envelope detector

The *IF* signal contains the amplitude information of the input signal as an envelope. To detect this amplitude we have to rectify the signal, as we only want to see the magnitude of a certain frequency component, so no time/phase domain information of the signal. Further - to detect the amplitude we have to lowpass-filter the *IF*-signal to get rid of the high frequency components. This filter is referred to as the *Video Bandwidth* filter which can be set by the user. Normally it is set to about the Bandwidth of the *RBW-Filter* however by reducing the Bandwidth one can reduce noise and so achieve a more stable display. However when the Video Filter has a smaller Bandwidth than the *RBW-Filter* it increases the Sweep time like when the *RBW-Filter* Bandwidth is decreased. For measuring the Amplitude of pulsed signals the *Video Filter* should be set to a higher Bandwidth than the *RBW-filter,* as pulsed signals have a high peaks but low averages.

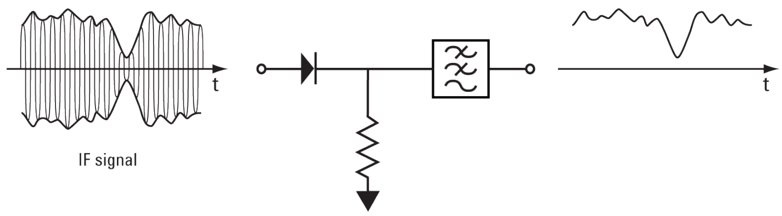

Figure 5: Envelope Detector

FFT Spectrum Analysis

The basic concept

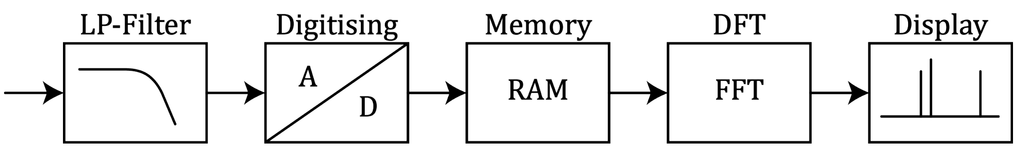

Figure 6: block shema of FFT-Analyser

FFT-Analysers digitally sample the Data in the time domain by means of Analog-to-Digital Converters (ADC) in the same way Digital Oscilloscopes sample the Data. This data is then transformed mathematically to frequency range through the Fourier-Transformation and in most cases further digitally processed.

Fourier transformation

The mathematical Fourier transformation enables the transformation of a time-domain signal into the frequency-domain by correlation the time domain signal with a sinusoidal function from negative infinity to positive infinity, this done for every single frequency point of interest. Every possible shaped signal can be decomposed into sinusoidal components a various frequencies. This is done by the Fourier transformation.

As we are dealing with digital data in this case it is a Discrete Fourier Transformation (DFT).

$$X(k)=\sum_{n=0}^{N-1} x(n)e^{j2\pi k n/N}$$

N is the total sample number, *x(n)* is the sampled version of the input signal *x(t)*, *X(k)* is the magnitude of the component at frequency Ω and *k is NΩ/2π.* Ω is the radian frequency.

To be noted is that we are doing the transformation over a finite number of samples, not from negative infinity to positive infinity. This can well be exact enough but can also cause some problems (see later)

In stead of directly calculating the values for each frequency of the DFT there are Algorithms which can do this with less calculation steps so more efficiently by using intermediary results in the calculations. Most often the Algorithm used is the Fast Fourier Transform (FFT).

Analog-to-Digital Conversion

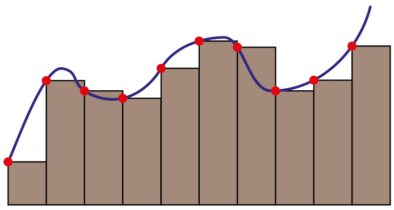

Figure 7: Digitisation

ADC's are devices that convert a analog voltage into a digital number proportional to the voltage on the input. By this process we are losing resolution by forcing the signal into discrete steps this is the quantisation error. The higher the resolution of the ADC the more Bits it possesses and the smaller the quantisation error is. The conversion however needs a certain amount of time. Oscilloscopes and FFT-Spectrometers have high speed ADC's that have a comparator for every single Bit (expensive & power consuming), so working in parallel. But even so the higher the resolution the longer one has to wait for the conversion. It is a compromise between speed and resolution. The ADC is the main limitation of Bandwidth and Resolution for the FFT-Analysers.

Anti Aliasing

When digitising analog signals in general the ADC has a smaller bandwidth than the signal at the input. In that case signals of higher amplitude are recognised but are misinterpreted as signals with a lower frequency. In actual fact following the *Nyquist–Shannon Sampling Theorem* the sampling frequency has to be at least double the frequency of the input bandwidth.

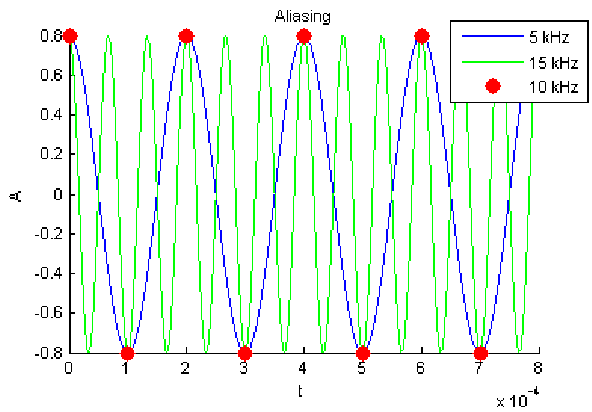

Figure 8: Aliasing

$$f_A \geq 2 \cdot B_{Sig}$$

In Figure 8 one can see what happens. The low frequency signal (blue) with half the sampling frequency can correctly be reconstructed out of the sampled data (red points) the high frequency signal (green) at 1.5 times the sampling frequency would be interpreted incorrectly as the same frequency as the low frequency signal.

Windowing

The signal frequency can be calculated exactly with FFT if the input signal we are examining is periodic and period is a multiple of the sampling period. Neither of this cases are normally going be true. This means that we have divergence from the exact result. In other words in DFT we are not doing the complete infinite Fourier Transformation but are multiplying the time-domain signal with a rectangle window function and calculating the Fourier transform just of this window, treating it as a signal repeating periodically for ever. If however the period of the signal is not a exactly a multiple of the sampling period repeating causes discontinuities which in the frequency domain transforms into additional spectral components not contained in the actual signal at he input. This spectral leakage shows itself as a line broadening and non-zero spectral components over a wide spectral range as well as diverging from the correct amplitude value.

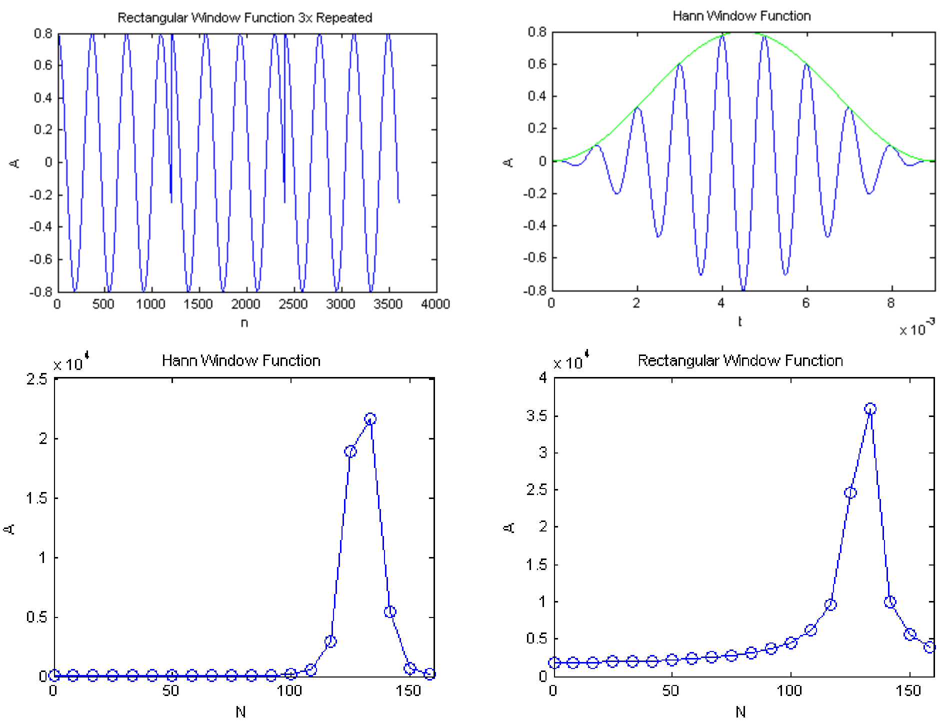

Figure 9: Hann- and Rectangular window function

A way of reducing spectral leakage is to multiply the time-domain signal with a different window function that shows a reduced of the signal at both ends of the function reducing discontinuities with a following repeated period of the function. For example the Hann- or the Hamming Window Function

Further Reading

- *Grundlagen der Spektrumanalyse* [2000]. C.Rauscher. Rohde & Schwarz GmbH & Co. KG
- *Fundamentals of Real-Time Spectrum Analysis* [2008]. Tektronix
- *Agilent Spectrum Analysis Basics* [2006]. Agilent Technologies, Inc.
- *Fundamentals of Modern Spectral Analysis* [2010]. Matthew T. Hunter et. al. University of Central Florida