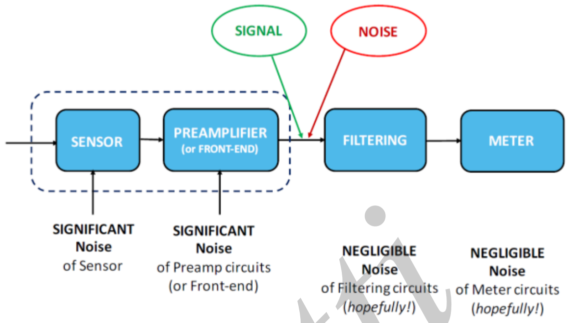
2. NOISE DESCRIPTION

To properly carry out a measurement, we must consider not only signal, but also noise. The presence of noise limits the system resolution. Therefore, it is necessary to deeply understand what noise is and how it is generated.

2.1 INTRODUCTION ON NOISE

The measurement system that we'll use to study the noise is exactly equal to that one used so far to study signals, in which the following four main blocks are typically sketched:



- Sensor:
  Transduces the physical variable of interest into an electrical quantity.

- Preamplifier:
  Is needed to make the signal level sufficiently high to be read by the following stages.
  At the preamplifier output we typically have a significant noise level superimposed to the signal, which is due to sensor intrinsic noise and to noise limitations of the gain stage.

- Filter:
  Has the task of shaping the noise and signal to improve the Signal-to-Noise Ratio (SNR).

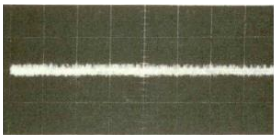
- Meter circuits

The noise of the last two blocks can be typically neglected in a well-designed system, while we have to focus on the noise coming out from the first two blocks.

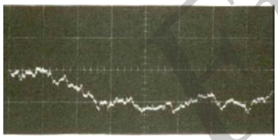
Once again, noise can be described like random disturbances which are superimposed to the useful signal. These unwanted fluctuations are generated randomly during the detection system.

2.2 STATISTICS OF NOISE SAMPLES AND PROBABILITY DISTRIBUTION

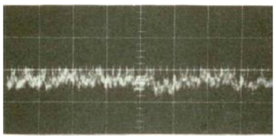
It's possible to look at noise waveforms in time using an oscilloscope. Depending on what kind of noise we are dealing with, we can visualize different noise shapes on the oscilloscope screen:



White Noise → spectrum $S = constant$



Random-Walk noise → spectrum $S = \frac{1}{f^2}$

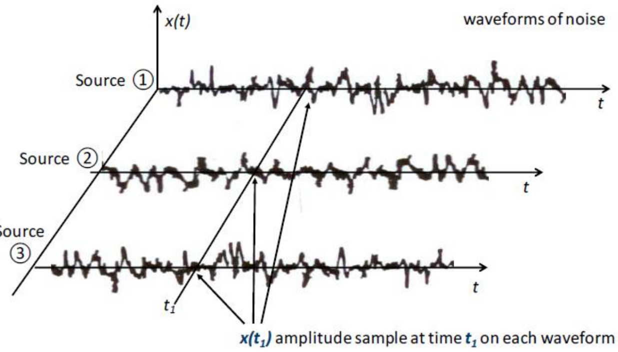


Flicker Noise → spectrum $S = \frac{1}{f}$

Speaking about noise, we have to introduce also the concept of sample.

Let's consider for instance three identical resistors. Plotting the noise waveforms of each source we immediately notice that the three waveforms are not equal. Even if the noise sources are identical, we have three different replicas of noise waveforms.

We can start evaluating the noise amplitude at a particular time instant t_1, and we end up with a set of different values.



Now, let's take one replica and let's use a scale.

Once we have sampled the noise amplitude $x(t_1)$ at the particular time instant t_1, we can then compare to a scale of discrete values x_k spaced by constant interval Δx and classify this sampled value at the nearest discrete value x_k of the scale.

Observing a high number N of noise waveforms and calling ΔN_k the number of sampled waveforms classified at x_k, we can define:

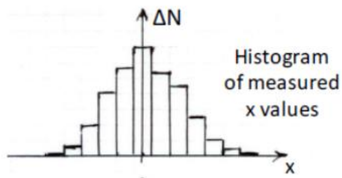
Statistical frequency of the amplitude x_k $\quad \Delta f_k = \frac{\Delta N_k}{N}$

Using this procedure, we collect N values of $x(t_1)$ in N different waveforms.

We can repeat this procedure for every time interval, obtaining:

$$\Delta N_0 \ in \ the \ central \ \Delta x \ (around \ x = 0)$$
$$\Delta N_1 \ in \ the \ first \ \Delta x \ (centerd \ in \ x_1 = \Delta x)$$
$$\dots$$
$$\Delta N_k \ in \ the \ k-th \ \Delta x \ (centered \ in \ x_k = k\Delta x)$$

We can then build a histogram of the measured x values, which give us a distribution of how many times we have a certain value in the different replicas.



Moving to differentials, so $\Delta x \to dx$, we get for an infinitesimal variation dx:

$$\Delta N_k \to dN_k = n(x_k)dx$$

Hence, for $\Delta f_k \to df_k$ we obtain:

$$df_k = \frac{dN_k}{N} = \frac{n(x_k)}{N}dx$$

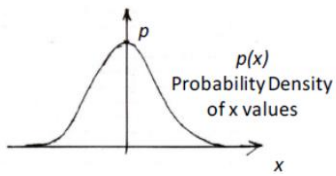
For $N \to +\infty$ we can explicit the last equation in this way:

$$df_k|_{N\to+\infty} = \frac{n(x_k)}{N}dx|_{N\to+\infty} = p(x)dx$$

in which we have defined:

Probability Density $\quad p(x) = \frac{n(x_k)}{N} \quad for \ N \to +\infty$

Instead of having a histogram, we'll have a probability density, which is a function that gives us the probability of having a certain value at the specific time t_1, over the full set of noise replicas.



2.2.1 STATIONARY AND NON-STATIONARY NOISE

More in general, we may have two kind of probability noise densities and respectively two types of noise:

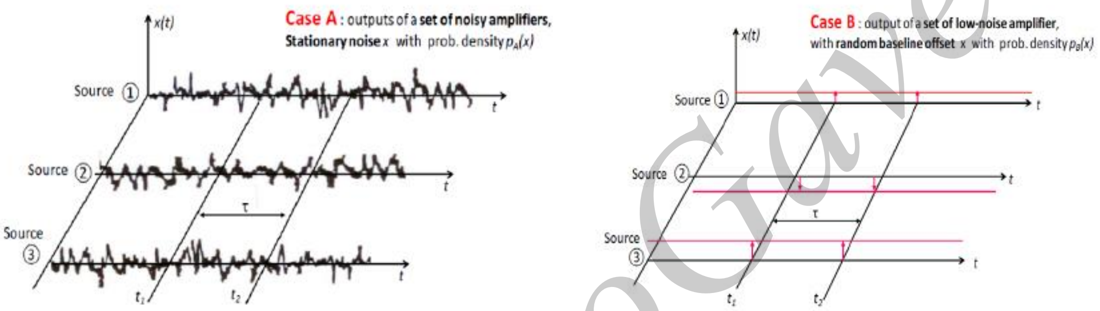
- **Stationary Noise**: the probability density is constant in time → $p = p(x)$

- **Non-Stationary Noise**: the probability density varies over time → $p = p(x, t)$

Beware:
at this point we may be induced to think that knowing the probability density corresponds to having a complete mathematical description of noise behavior, but it is absolutely not true, because the probability density alone does not give us a complete description of noise. In fact, different cases can have equal probability density.

Demonstration (example):

Let's observe the output waveforms of two different sets of amplifiers, one composed by noisy elements (set A) and the other composed by low-noise amplifiers (set B). consider having a stationary noise with a probability density $p_A(x)$ in set A and a random baseline offset with probability density $p_B(x)$ in set B.



In set A, if we evaluate the output noise amplitude at two different time instants t_1 and t_2, we end up with two random values, respectively $x(t_1)$ and $x(t_2)$ having the same probability density $p_A(x)$.

If we call τ the time interval between t_1 and t_2 ($\tau = t_2 - t_1$) and we study $x(t_1)$ and $x(t_2)$ as functions of τ, we end up with some properties of the sampled values, related to the time interval τ.

For each source:

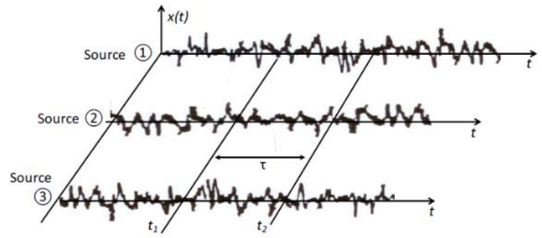
- $x(t_1)$ and $x(t_2)$ are practically identical for an ultra-short interval τ
- $x(t_1)$ and $x(t_2)$ are somewhat different for short interval τ
- $x(t_1)$ and $x(t_2)$ are different and independent for longer interval τ

In set B, instead, the situation is totally different. In this case, if we sample the output waveforms at two different time instants t_1 and t_2, we still have that $x(t_1)$ and $x(t_2)$ are random values with probability density $p_B(x)$, but now if we compare $x(t_1)$ and $x(t_2)$ sampled on the same amplifier, we see that they are identical for any time distance τ, short or long.
In other words, the value of the amplifier output in set B is different from component to component (following a statistical fluctuation), but for each amplifier the output is constant over time.
A time-constant value at the output is called offset or baseline.

Set B is different from set A, but they can have equal probability density: $p_B(x) = p_A(x)$.

■

2.3 COMPLETE NOISE DESCRIPTION WITH PROBABILITY DISTRIBUTION



The probability as we have seen in the last paragraph, which is actually called **marginal probability** $p_m(x,t)dx$, is not sufficient to fully describe noise.

To do that, we need also to know the **joint probability** $p_j(x_1,x_2,t_1,t_2)dx_1dx_2$ which gives us the probability of having a certain value x_2 at the time t_2 and a certain value x_1 at time t_1.

(A joint probability, in probability theory, refers to the probability that two events will both occur, or in other words, it represents the likelihood of two events occurring together)

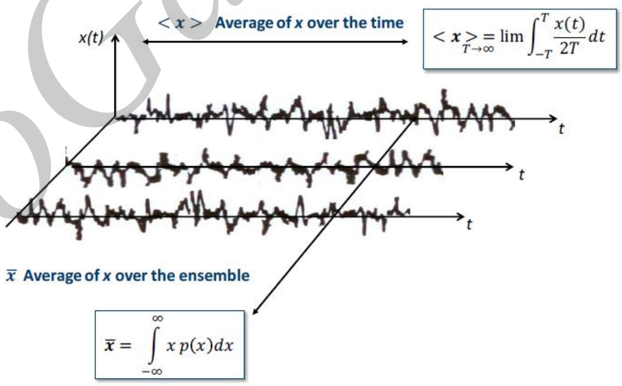
Therefore, a full description of noise is obtained by knowing both:

Marginal Probability Density $\quad p_m(x) = p_m(x,t_1)\ \forall t_1$

Notice: for stationary noise it doesn't depend on time $p_m = p_m(x)$.

Joint Probability Density $\quad p_j(x_1,x_2) = p_j(x_1,x_2,t_1,t_2) = p_j(x_1,x_2,t_1,t_1+\tau)\ \forall t_1, t_2 = t_1+\tau$

Notice: for stationary noise the joint probability p_j depends only on the time interval τ and not on the time instant t_1, $p_j = p_j(\tau)$.



Now, let's consider again the noise amplitude waveforms of many sources.

Instead of having just two axes, the time and the amplitude, here we introduce a new dimension which is the dimension of the replicas, of the ensemble.

This means that we'll have to possible ways to make the average

When we average in time, we choose a particular waveform among all the waveforms and we compute the integral:

Average of x over time $$< x > = \lim_{T\to\infty} \int_{-T}^{+T} \frac{x(t)}{2T} dt = \lim_{T\to\infty} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \frac{x(t)}{T} dt$$

On the other hand, when we average the amplitude of x over the ensemble, we integrate the x variable multiplied by its probability distribution $p(x)$:

Average of x over ensemble $\quad \overline{x} = \int_{-\infty}^{+\infty} x\,p(x)dx$

2.4 NOISE DESCRIPTION WITH 2nd ORDER MOMENTS OF POWER DISTRIBUTION

Notice: for clarity, we call here the two statistical variables x and y, instead of x_1 and x_2

A fundamental way to describe noise is using the 2nd order moment of the probability density.

Considering any two statistical variables x and y, we define:

Moments of marginal probability distribution $p(x)$ $\quad m_n = \overline{x^n} = \int_{-\infty}^{+\infty} x^n p(x)dx$

and

Moments of joint probability distribution $p(x,y)$ $\quad m_{jk} = \overline{x^j y^k} = \int_{-\infty}^{+\infty} x^j y^k p(x,y)dxdy$

- Moments m_n and m_{jk} give the information on the features of the marginal and joint probability distributions, respectively
- The higher the order n or $j+k$, the more detailed is the information

If we consider a description of noise limited to the 2nd order moment we obtain:

Mean Square Value (Variance) $\quad m_2 = \overline{x^2} = \int_{-\infty}^{+\infty} x^2 p(x)dx = \sigma_x^2$

Mean Product Value (Covariance) $\quad m_{11} = \overline{xy} = \int_{-\infty}^{+\infty} x\, y\, p(x,y)\, dxdy = \sigma_{xy}^2$

Notice:

- $m_0 = m_\infty = 1$
- $m_1 = m_{10} = \overline{x} = 0 = \overline{y} = m_{01}$

We can now evaluate the mean square value of noise amplitude waveform when sampled at a generic time instant t_1:

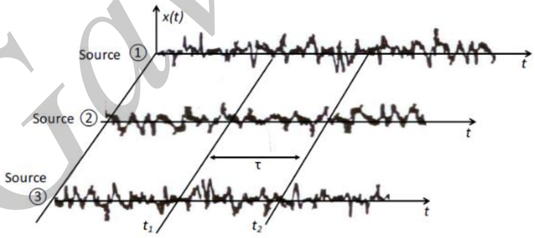
$$\overline{x^2(t_1)} = \sigma_x^2(t_1)$$

which does not depend on t_1 in case of stationary noise.

For any pair of time instants t_1 and $t_2 = t_1 + \tau$, it's possible to calculate the mean product:

$$\overline{x(t_1)x(t_2)} = \overline{x(t_1)x(t_1+\tau)}$$

which depends only on the time interval τ and not on the time position t_1, in case of stationary noise.



Looking the last equation, we can notice a likelihood with the autocorrelation, but this time we are averaging on ensemble and not on time. This suggests us that is possible to define correlation functions also for noise.

2.5 NOISE AUTOCORRELATION FUNCTION AND NOISE POWER

We can immediately define:

Noise Autocorrelation Function $\quad R_{xx}(\tau) = R_{xx}(t_1, t_1+\tau) = R_{xx}(t_1, t_2) = \overline{x(t_1)x(t_2)} = \overline{x(t_1)x(t_1+\tau)}$

We can notice that the noise autocorrelation function is always a function of the interval τ between t_1 and t_2. Furthermore, it is also a function of t_1 in case of non-stationary noise.

The autocorrelation function of a noise x, called $R_{xx}(\tau)$, is an average on ensemble, while for a signal s the autocorrelation function $k_{ss}(\tau)$ is a time average.

The noise mean square value, also called noise power, similarly to what happens with signals, corresponds to the autocorrelation function of noise at $\tau = 0$:

Noise Power $\quad \sigma_x^2 = \overline{x^2(t)} = R_{xx}(t, 0)$

Noise Power for stationary noise $\quad \overline{x^2} = R_{xx}(0) \qquad$ which is constant for any t.

2.6 NOISE POWER SPECTRUM

Noise has power-type waveforms, having divergent energy.

Considering the whole ensemble, we have statistical variations from waveform to waveform.

By averaging over the ensemble of the autocorrelations of noise waveforms, the concepts of power and power spectrum, that we introduced for a generic signal, can be extended to noise.

In other words, we take exactly the same definitions of power and power spectrum previously defined for signals and we just make an average over ensemble. as we have two dimensions: the time and the ensemble, by averaging over the ensemble we are collapsing the ensemble dimension in order to go back to our "classical" situation.

Noise Power' $$P[x(t)] = \overline{\lim_{T\to\infty} \int_{-T}^{+T} \frac{x^2(\alpha)}{2T} d\alpha} = \overline{\lim_{T\to\infty} \int_{-\infty}^{+\infty} \frac{x_T^2(\alpha)}{2T} d\alpha} = \overline{\lim_{T\to\infty} \int_{-\infty}^{+\infty} \frac{|X_T(f)|^2}{2T} df} =$$
$$= \int_{-\infty}^{+\infty} \overline{\lim_{T\to\infty} \frac{|X_T(f)|^2}{2T}} df = \int_{-\infty}^{+\infty} \lim_{T\to\infty} \frac{\overline{|X_T(f)|^2}}{2T} df$$

Noise Power Spectral Density (Noise Power Spectrum) $S_{x,P}(f) = \lim_{T\to\infty} \frac{\overline{|X_T(f)|^2}}{2T}$

Once the power spectral density of noise is defined, we can rewrite its power as:

$$P[x(t)] = \int_{-\infty}^{+\infty} S_{x,P}(f) df$$

Furthermore, by averaging over the ensemble, we can extend to noise also the second definition of power spectrum introduced for signals:

$$S_{x,P}(f) = \overline{F[K_{xx}(\tau)]} = F\left[\overline{K_{xx}(\tau)}\right] = F\left[\overline{\lim_{T\to\infty} \frac{\int_{-\infty}^{+\infty} x_T(\alpha)x_T(\alpha+\tau)d\alpha}{2T}}\right] = F\left[\lim_{T\to\infty} \frac{\overline{K_{xx,T}(\tau)}}{2T}\right] = \lim_{T\to\infty} \frac{F\left[\overline{K_{xx,P,T}(\tau)}\right]}{2T}$$

And so, we can writhe the noise power as:

Noise Power'' $P[x(t)] = \int_{-\infty}^{+\infty} S_{x,P}(f) df = \overline{K_{xx}(0)}$

For signals we introduced the concepts of power, autocorrelation and power spectral density and we linked these three elements: the autocorrelation in zero is the power of the signal, the Fourier transform of the autocorrelation of the signal is the power spectrum and finally the integral of the power spectrum is the power.

In the case of noise, we have a connection between the variance and the autocorrelation of noise in zero, then we have a connection between power and power spectrum and finally we have a connection between the power spectrum and the autocorrelation $K_{xx}(\tau)$ averaged over the ensemble.

It's worth doing some observations.

We can start observing that the mathematical spectral density $S_x(f)$ is defined over a frequency range that goes from $-\infty$ to $+\infty$, therefore we call it **bilateral noise power spectral density $S_{x,B}(f)$**. Then, the noise power is calculated as:

$$P[x(t)] = \int_{-\infty}^{+\infty} S_{x,B}(f) df$$

Since $S_{x,B}(f)$ is symmetrical $[S_{x,B}(-f) = S_{x,B}(+f)]$, we can write the noise power in this way:

$$P[x(t)] = 2\int_0^{+\infty} S_{x,B}(f) df = \int_0^{+\infty} 2S_{x,B}(f) df = \int_0^{+\infty} S_{x,U}(f) df$$

in which:

Unilateral Noise Power Spectral Density $S_{x,U}(f) = 2S_{x,B}(f)$

Is there a way to connect the $K_{xx}(\tau)$ and the $R_{xx}(\tau)$ autocorrelation functions?

Recalling the formula of $K_{xx}(\tau)$, we notice that we were making an averaging on time:

$$K_{xx}(\tau) = \int_{-\infty}^{+\infty} x(\alpha)x(\alpha+\tau)d\alpha = < x(t)x(t+\tau) >$$

So, to obtain $\overline{K_{xx}(\tau)}$ we have to do a double average: an average on time first and an average over ensemble then.

It can be demonstrated that it's possible to exchange the order of the two averages:

$$\overline{K_{xx}(\tau)} = <\overline{x(t)x(t+\tau)}> = <\overline{x(t)x(t+\tau)}> = <R_{xx}(t,t+\tau)>$$

The power spectrum is thus related to the ensemble autocorrelation function in this way:

Power Spectrum" $S_x(f) = F[<R_{xx}(t,t+\tau)>]$

where K_{xx} has been replaced with R_{xx}.

- For non-stationary noise, $S_x(f)$ can be defined with reference to the time-average of the ensemble autocorrelation function of noise.

- For stationary noise instead, there is no need of time averaging, because it's the same for each time (it depends only on the time difference τ), in fact it is simply:

$$<R_{xx}(t,t+\tau)> = R_{xx}(\tau)$$

Therefore, we can conclude that for stationary noise the power spectral density can be defined as:

$$S_x(f) = F[R_{xx}(\tau)]$$

As the integral from $-\infty$ to $+\infty$ of the power spectrum gives us the noise power and the power spectrum, in case of stationary noise, is the Fourier transform of the ensemble autocorrelation function $R_{xx}(\tau)$, so the integral from $-\infty$ to $+\infty$ of the power spectrum equals the value in zero of the anti-Fourier transform, but since the anti-Fourier transform is the ensemble autocorrelation $R_{xx}(\tau)$, we can conclude that:

$$P[x(t)] = R_{xx}(0)$$

which corresponds to the variance, as we have seen before.

<u>Notice</u>: all the theory is based on bilateral spectral density, but in real applications the unilateral one is most used.