Lecture 4: Dealing with the measurement noise of a sensor

S. Dilhaire[1], D. Maillet[2],

[1]LOMA, Université Bordeaux and CNRS, Bordeaux, France
E-mail: stefan.dilhaire@u-bordeaux.fr
[2] LEMTA, Université de Lorraine and CNRS, Vandoeuvre-lès-Nancy, France
E-mail: denis.maillet@univ-lorraine.fr

Abstract. In this lecture, the electronic noise of a measurement system and its physical origin is presented, as well as its main characteristics (see also lectures L2 and L5 for temperature measurements). The main properties of random signals (joint probability density, expectation, variance-covariance matrix) are also briefly recalled. These notions are essential for the assessment of residual error in linear (see lecture L3) and non-linear parameter estimation problems (see lecture L7) and for function estimation problems (see lecture L8),

Introduction

Measurement noise constitutes a critical point that is crucial to understand in parameter estimation, that is when comparing measurements and a model: the knowledge of noise properties allows an assessment of the level of the estimation errors (bias and dispersion) in this type of experimental inverse problems. Section 1 is devoted to the characterization of the measurement noise and a to brief description of the different origins of measurement noise an “inverter” (that is an experimentalist combined with a model-builder) can meet. Section 2 presents the vector-matrix representation of these characteristics, for further use in parameter estimation problems.

1. Physics of measurement noise

1.1. Definition

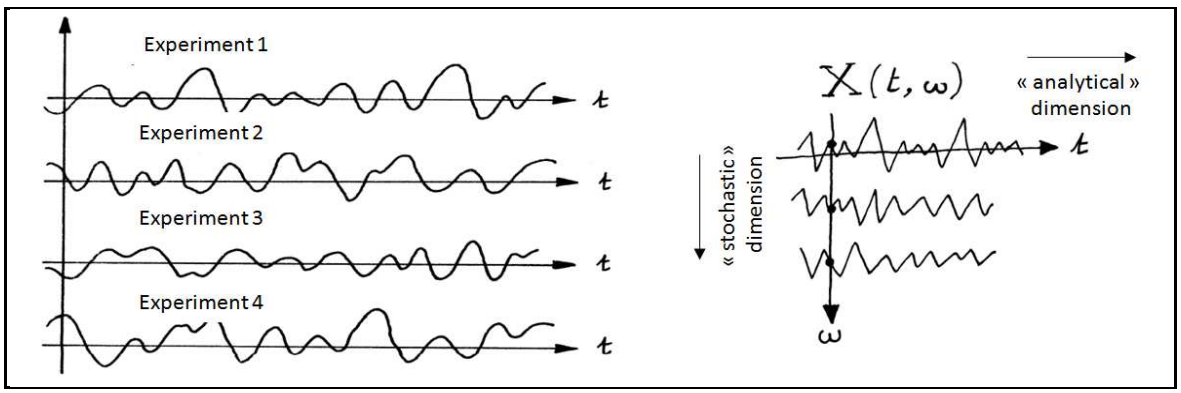

Figure 1 – Random signal

A random signal (or stochastic process) is a signal that does not remain unchanged when the experiment it stems from is repeated (see Fig. 1).

It is noted $X(t,\omega)$where ω is a test (that is a random variable that contains the outcome of a random draw).

.

$x(t, \omega_i)$is a realization of $X(t,\omega)$for a specific draw ω_i.

1.2. Statistical descriptors

- **1rst order descriptor**

  Expectation = Ensemble mean :

$$m_X(t) = E\{ X(t, \omega) \} \tag{4.1}$$

  It is an indicator of the average position of the signal at time t (see Fig. 2)

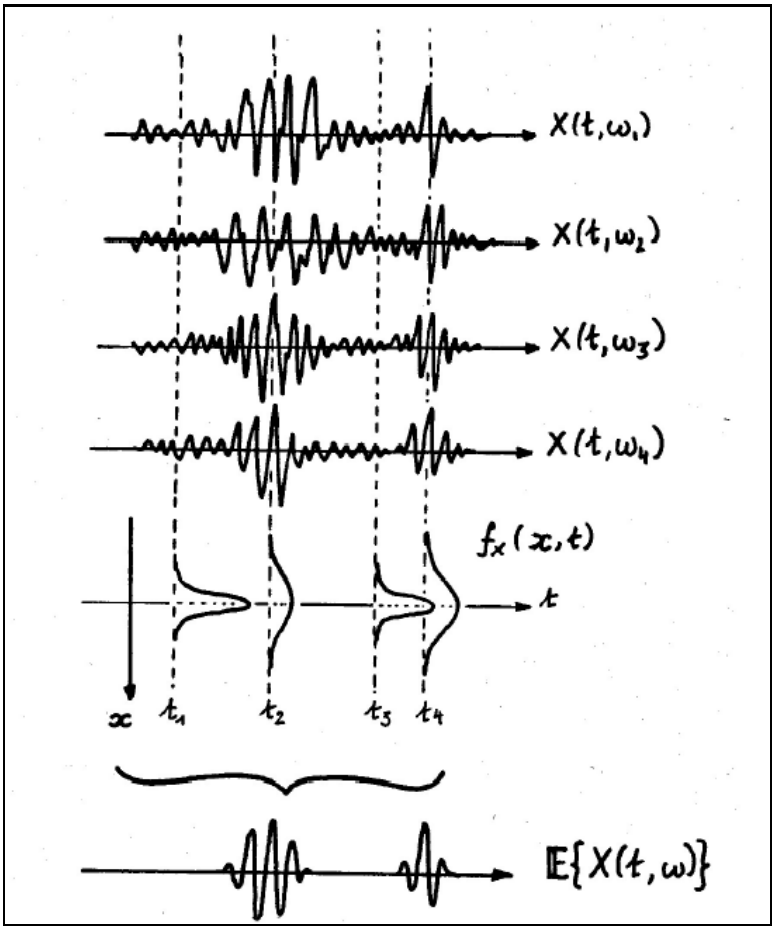

Figure 2 – Realizations of a random, local probability density function and expectancy variation

- **2nd order descriptors**

Instant power = Quadratic mean :

$$P_X(t)=\mathrm{E}\{X(t,\omega)^2\} \tag{4.2}$$

It is a quantification of the average power of the signal at a given time *t*

Calculation of the instant power of signal at time *t* :

$$P_X(t) \underset{\Delta t\to 0}{=} \frac{1}{\Delta t}\int_t^{t+\Delta t} \mathrm{E}\left\{\left|X(t,\omega)\right|^2\right\}\mathrm{d}t \tag{4.3}$$

Variance

$$V_X(t)=\mathrm{E}\{\,|X(t,\omega)-m_X(t)|^2\} \tag{4.4}$$

It quantifies the instant power of random fluctuations around the mean (dispersion indicator).

Standard deviation

$$\sigma_X(t)=\sqrt{V_X(t)} \tag{4.5}$$

It has the same physical meaning as the variance, but its physical units are the same as the signal.

The previous descriptors characterize the signal behavior (average position, dispersion) at a time *t*.

They do not allow for any analysis of the relationships (dependence) that exist between samples.

An indicator is needed for quantifying the level of the « force » that makes a sample at time *t* + *T* depend on its value at time t.

- **The autocorrelation function**

$$R_X(t,\tau)=\mathrm{E}\left\{X(t)\;X(t+\tau)\right\} \tag{4.6}$$

It quantifies the correlation (or the scalar product or the projection, in the stochastic meaning) between $X(t+\tau)$ and $X(t)$ (see Fig. 3).

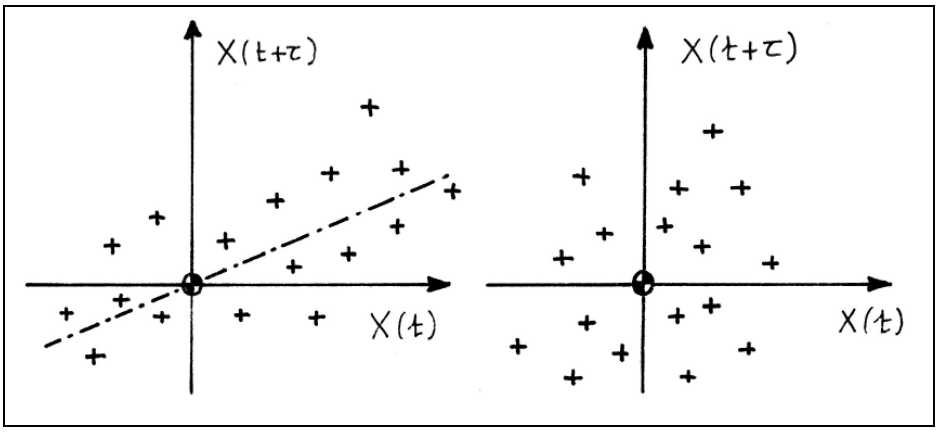

Figure 3 – Correlation between signal for a given time lag ⏐ · high level (left) and low level (right)

Interpretation of the autocorrelation function:

A signal very highly correlated with itself exhibits very slow fluctuations (with a “smooth” appearance), see Fig. 4, top.
A signal lowly correlated with itself exhibits very rapid fluctuations (with a “chaotic” appearance), see Fig. 4, bottom.

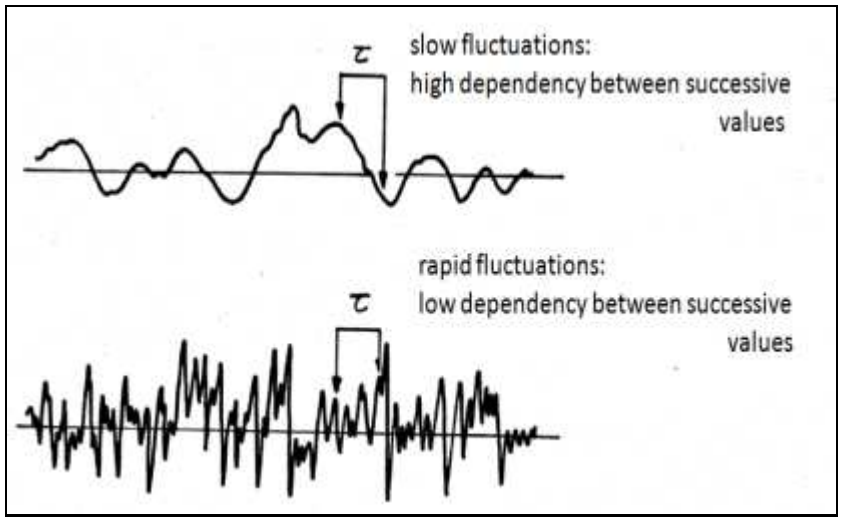

Figure 4 – Typical slow (top) and fast (bottom) fluctuations of signal

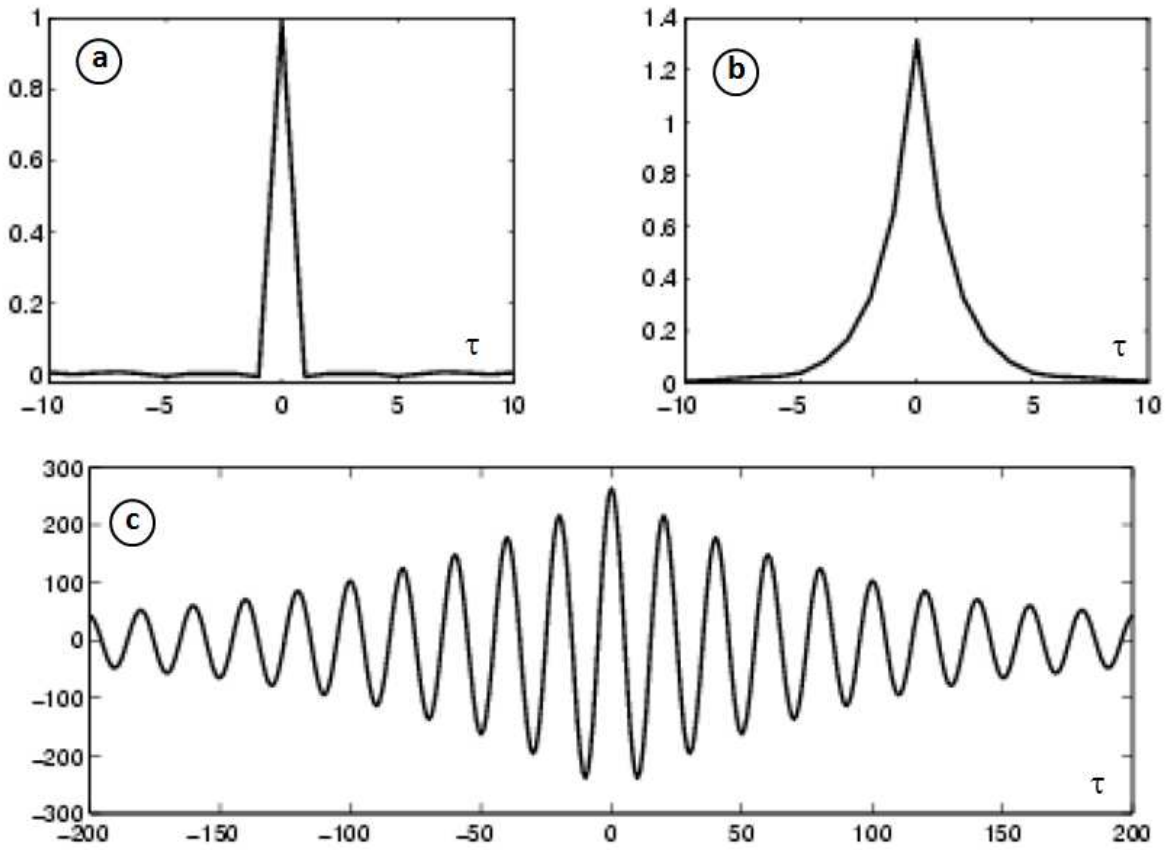

Figure 5 – $R_X(\tau)/R_X(\tau=0)$ for a) a discrete white noise
b) a thermal noise and c) a noise of narrow frequency interval

1.3 Properties: stationarity, ergodicity

- **The stationarity property of a signal**

Definition: A stationarity random signal is a signal whose statistics do not depend on time:

$$m_X(t) = m_X \quad ; \quad V_X(t) = V_X \quad ; \quad P_X(t) = P_X \qquad (4.7a, b, c)$$

Remark 1: In fact, this assumption is nearly compulsory, in order to be able to estimate the characteristics of the signal;

Remark 2: The stationarity property of a random signal is analogous to the periodicity condition of a deterministic signal. In any case, it just an idealization.

- **The ergodicity property of a signal**

Definition: A signal is ergodic (in the strong sense), see Figure 6, if:

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} X(t,\omega)\, \mathrm{d}t = \mathrm{E}\left\{ X(t,\omega) \right\} \tag{4.8}$$

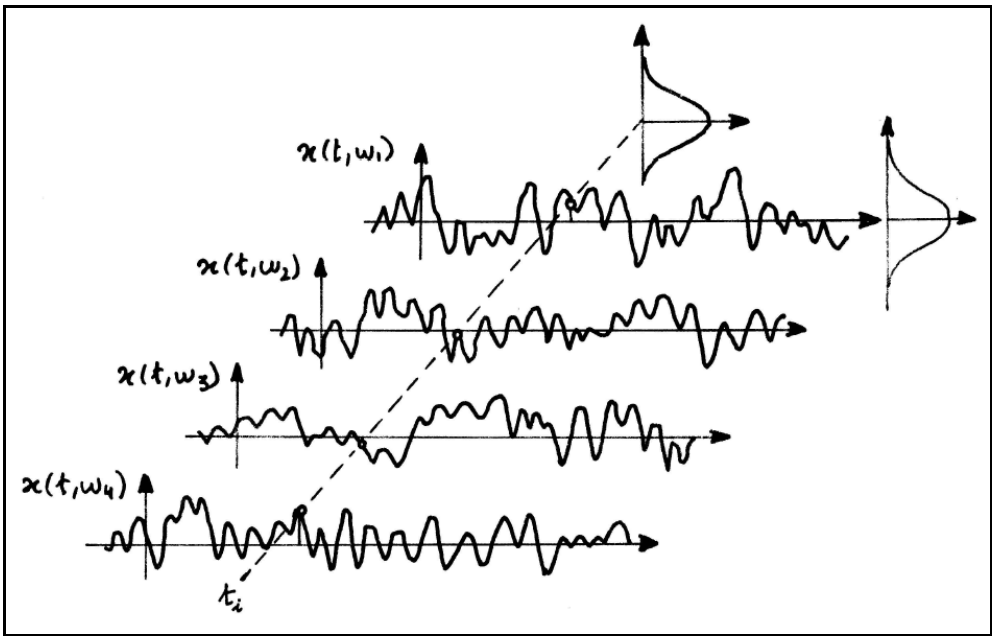

Figure 6 –Ergodic signal

As a particular consequence, it yields:

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} X(t,\omega)\, \mathrm{d}t = m_X \tag{4.9}$$

which means that its time average is equal to the ensemble average.

1.4. Examples

Different types of noise exist:

- White noise: constant spectrum (Figures 7, 9 and 10))
- Pink noise: spectrum in $1/f$
- Brown noise (brownian), sometimes called « red »: spectrum in $1/f^2$ Figure 11)
- Blue noise: spectrum in f
- Violet noise: spectrum in f^2
- Grey noise: blue + pink

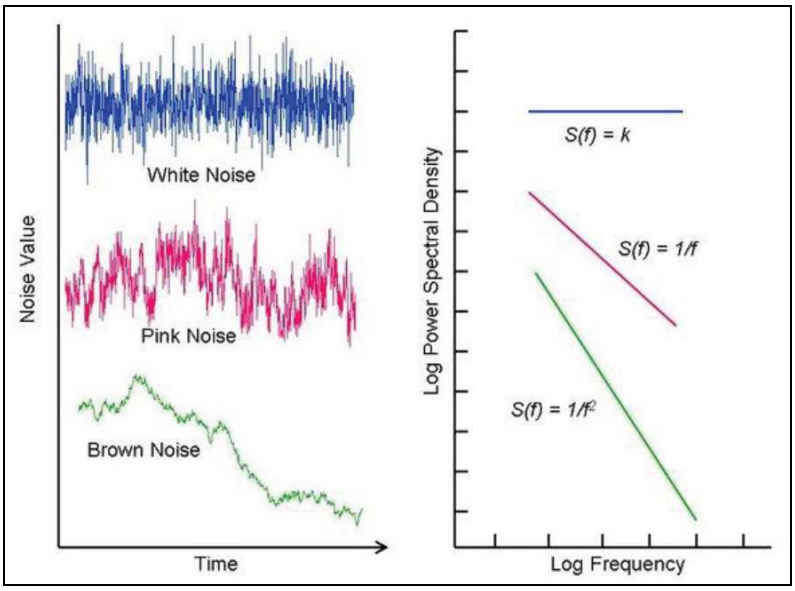

Figure 7 –Different types of noise

Other types of noise:

- Thermal noise (Johnson)
- Flicker noise
- Shot noise (Figure 12)
- Burst noise
- Avalanche noise
- Popcorn noise (see Figure 8)...

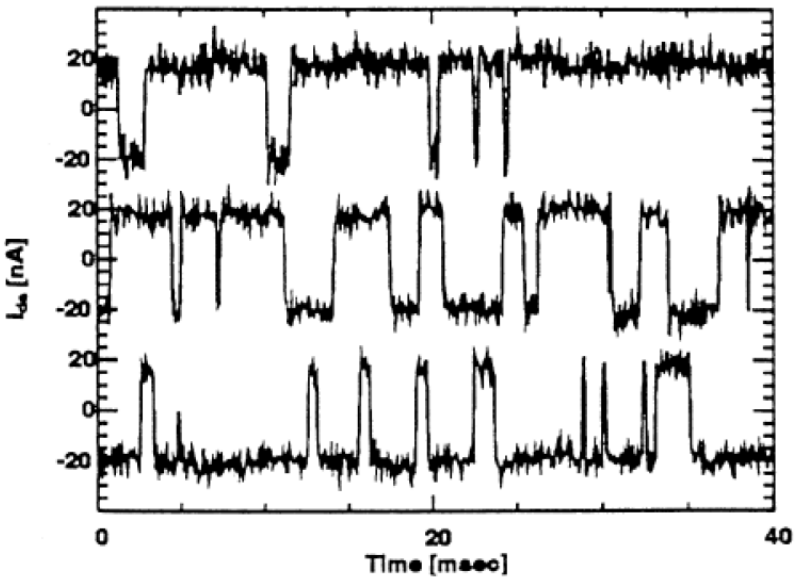

Figure 8 –Popcorn noise

The Popcorn noise (random telegraph signal) corresponds to a discrete modulation of the channel current caused by the capture and emission of a channel carrier

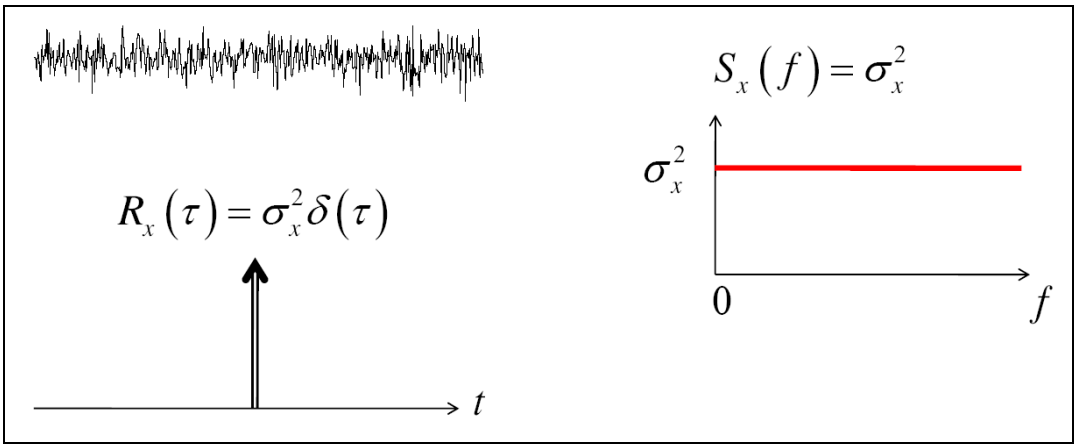

Figure 9 –White noise: autocorrelation (left) and frequency spectrum (right)

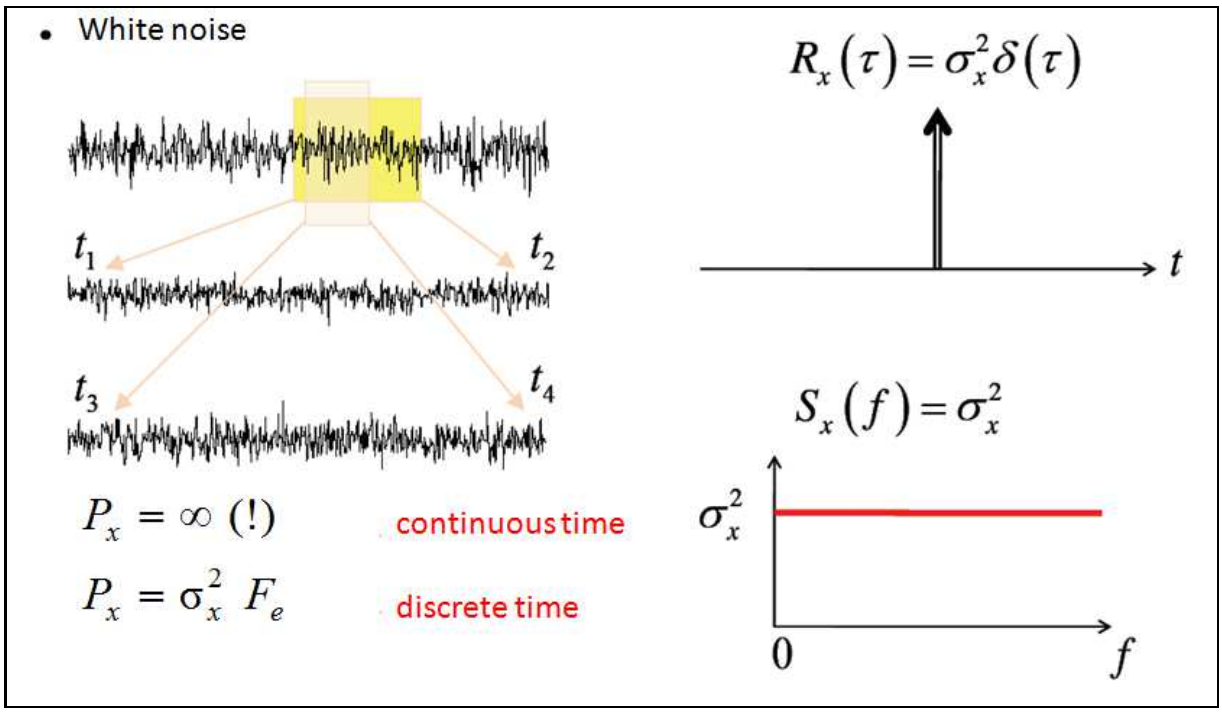

Figure 10 –White noise

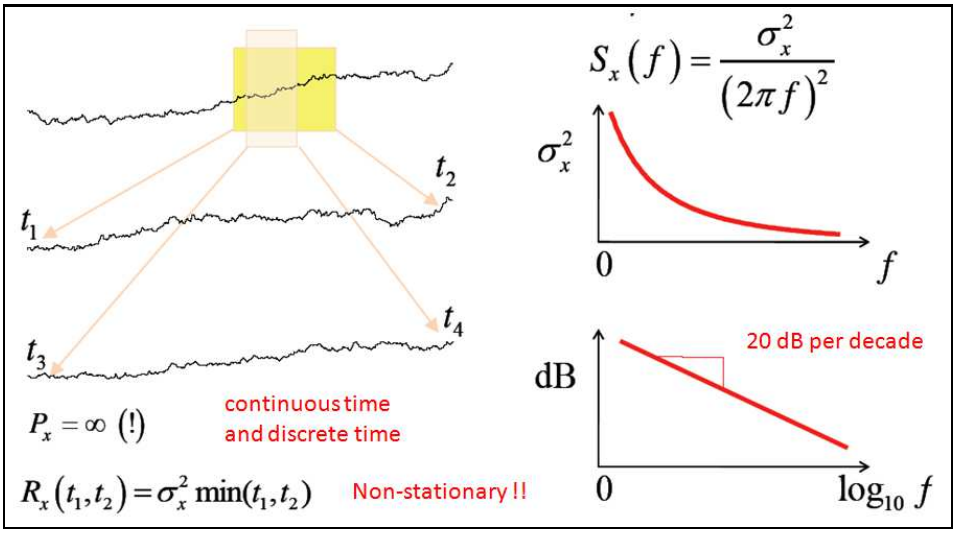

Figure 11 –Brownian noise (non-stationary)

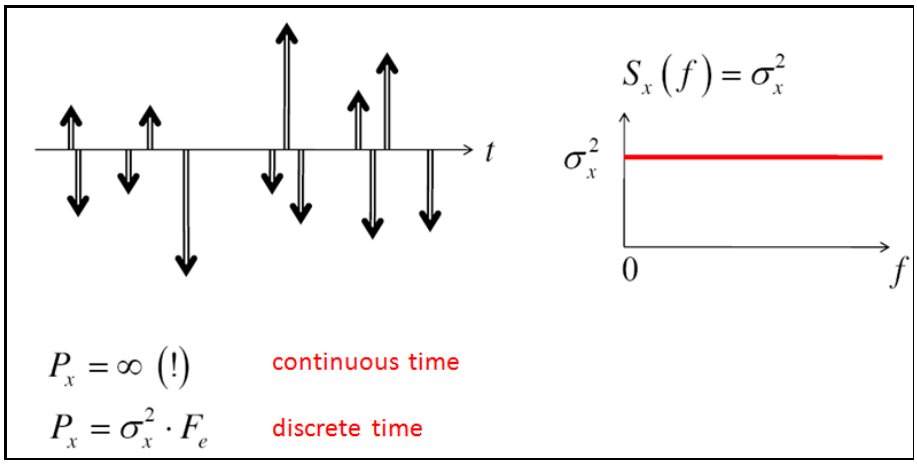

Figure 12 –Shot noise

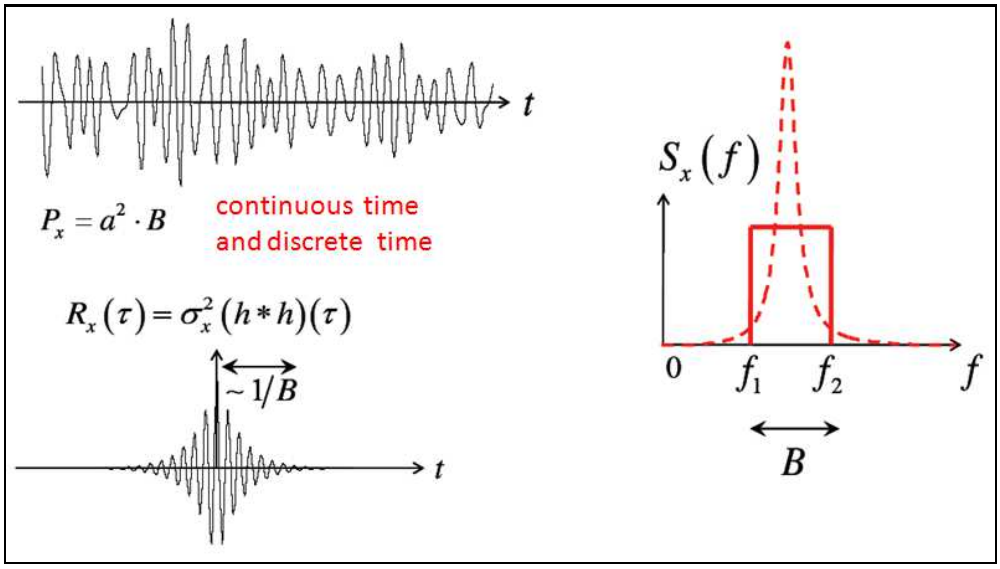

Figure 13 –Noise of limited spectral interval

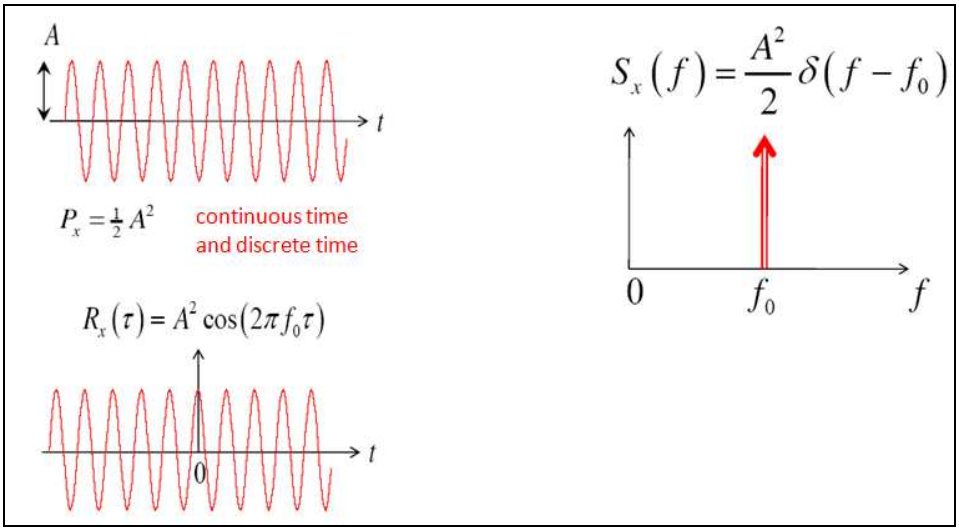

Figure 14 –Sine signal with random phase

1.5 Noise in 1/*f*^{a}

This type of noise lies in between a white noise and a brownian noise :

- It is a process with a long memory.
- Many natural signals exhibit a spectrum in 1/*f*. This is based on repeated observations. However no clear theoretical explanation is yet available. It does not stem from any differential equation.
- The longest series of information of data is the level of the Nile river measured by Egyptian over a 3500 years period: it follows a law in 1/*f*.

• This type of noise is called by different names :

- pink noise (with a law) in 1/*f*,
- Flicker noise in $1/f^a$ (in the special case where $a = 1$)

It is met in electronics, see Figures 15 and 16.

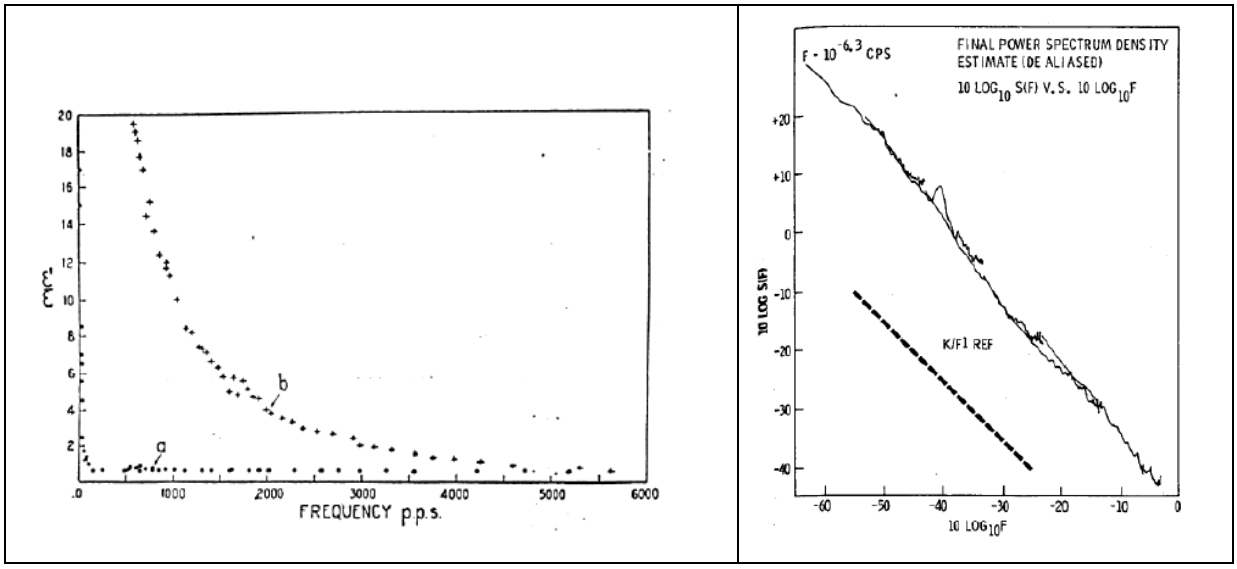

Figure 15 – Electronic noise
- left: spectrum (linear scale), B. Johnson, Phys. Rev. 26 (1925) 71
- right: noise spectrum of an amplifier over 3 months (log-log scale), M.A. Caloyannides – J. Appl. Phys. 45 (1974) 307.the heat flux density q that flows through the wall.

Other examples of pink noise are shown in Figure 16.

Flicker noise:

- is a generalization of the pink noise
- its spectrum is in $1/f^a$, with $1/2 < a < 3/2$

Thermal noise of Johnson-Nyquist

– the electronic noise has been observed first by Johnson in 1926 and it has been explained by his colleague, Nyquist.

- it results from the Brownian motion of electrons at constant temperature (thermalization).

- in its most general form, it is an elementary stochastic model of noise:

• white noise
• pink noise
• Brownian noise

Noise in a measurement chain

The noise in measurement chain based on a multi-stage measurement chain results from noises at different level of this system, see Figure 16.

Noise for a multi-stage circuit

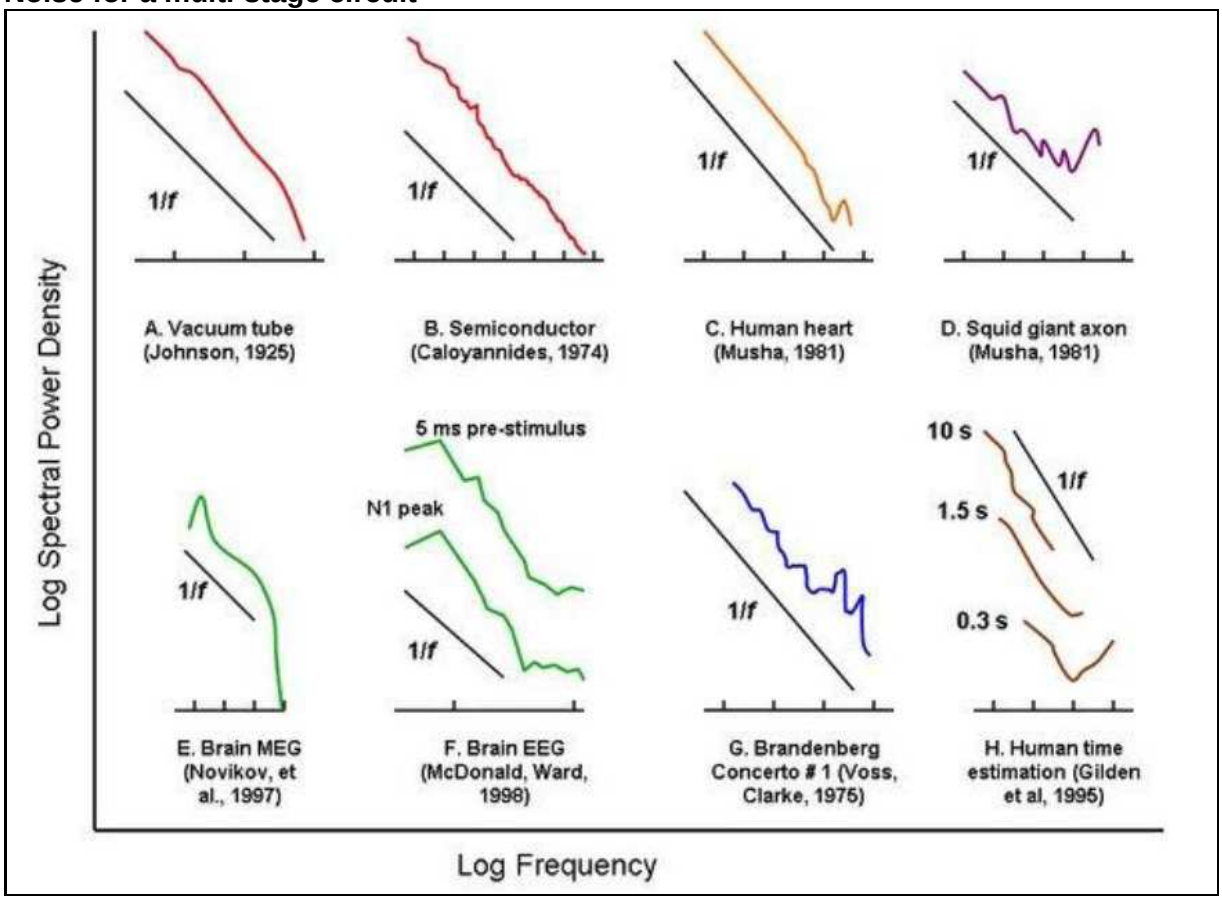

Figure 16 – Pink noise: other examples

Noise in a measurement chain

The noise in measurement chain based on a multi-stage measurement chain results from noises at different level of this system, see Figure 17.

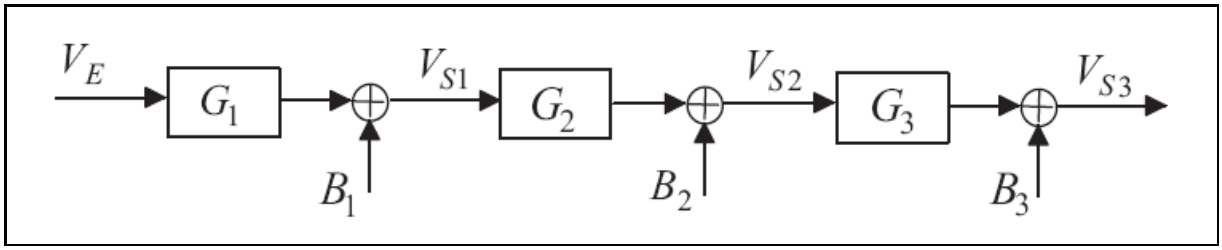

Figure 17 - Noise for a multi-stage circuit

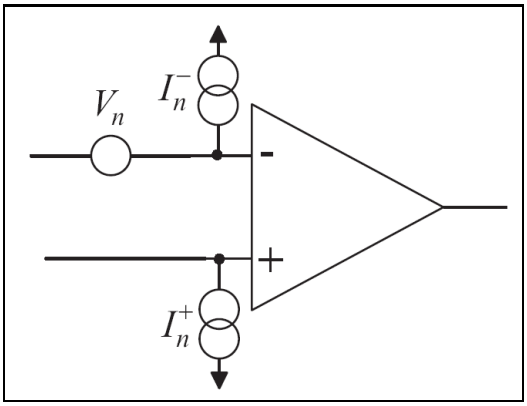

Figure 18 – Current and voltage modeling of the noise of an operational amplifier

At the local level, one has to model the behavior of each stage, see Figure 18.

The correlation function and the spectral density can be used as tools for the characterization of the system.

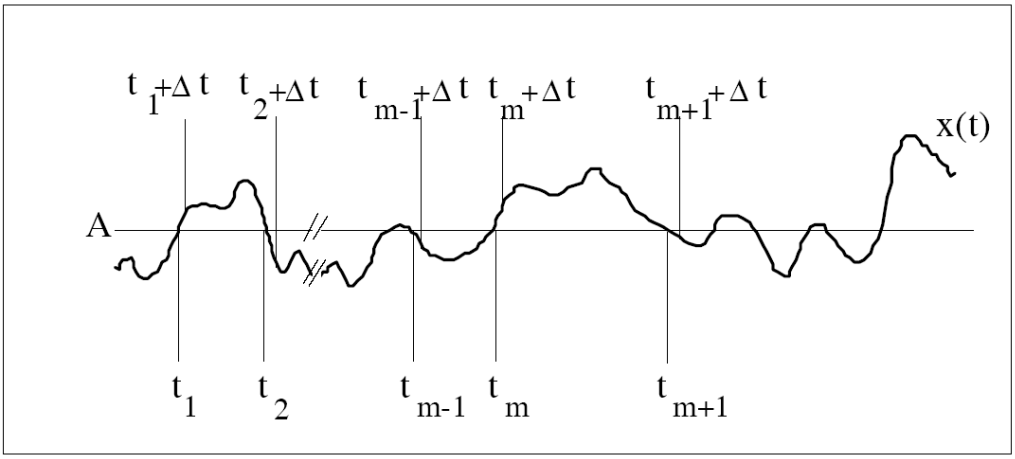

Figure 18 – Relationships between signal values at different measurement times

One has to answer the following question:

- are any values of $f(t)$ at times $t_1 + \Delta t$, $t_2 + \Delta t$, ..., $t_{m-1} + \Delta t$, $t_m + \Delta t$, $t_{m+1} + \Delta t$ possible ?

It is possible to plot the autocorrelation function (see Figure 19):

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t)\, x(t-\tau)\, \mathrm{d}t \qquad (4.10)$$

The value of this function for a zero time lag | is:

$$C_{xx}(0) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\, x(t)\, \mathrm{d}t = \lim_{T\to\infty} \frac{1}{T} \int_0^T x^2(t)\, \mathrm{d}t \tag{4.11}$$

And the sign of the time lag does not matter in its definition since this function is symmetric:

$$C_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\, x(t-\tau)\, \mathrm{d}t = \lim_{T\to\infty} \frac{1}{T} \int_{-\tau}^{T-\tau} x(t'+\tau)\, x(t')\, \mathrm{d}t' = C_{xx}(-\tau) \tag{4.12}$$

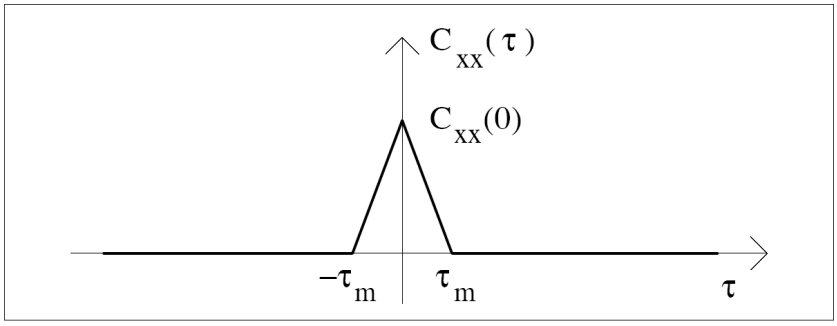

Figure 19 – Example of the shape of an autocorrelation function

The energy of the signal is defined by:

$$e_m = \lim_{T\to\infty} \frac{1}{T} \int_0^T s^2(t)\, \mathrm{d}t \tag{4.13}$$

It can be decomposed on a frequency basis, see Figure 20:

$$e_m \text{ (in Fourier domain)} = X(f)\, \mathrm{d}f \tag{4.14}$$

The quadratic mean in the time domain is equal to the quadratic value in the frequency domain.

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T x^2(t)\, \mathrm{d}t = \int_0^T X^2(f)\, \mathrm{d}f \tag{4.14}$$

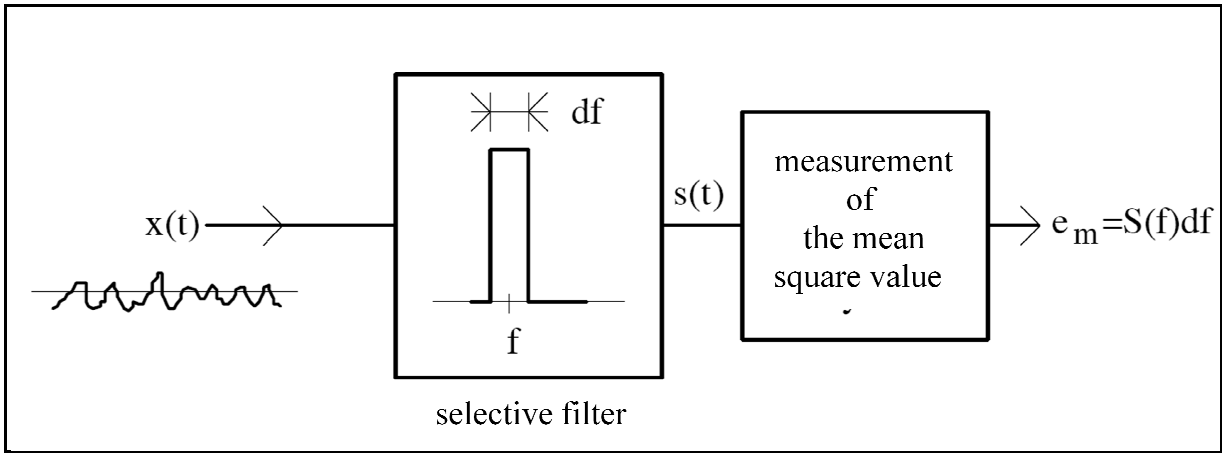

Figure 20 – Measurement of the spectral density

2. Mathematical characterization of noise for parameter estimation problems

2.1 A reminder of vector random variables

2.1.1 Static model and estimation of the parameters of a probability law

We assume here that the model output is a constant μ (it does not depend on the independent variable, called 'times' her), which means that $y_{mo} = \mu$. This means that its output is 'static' and repetition of measurements y of y_{mo} corresponds to a simple random sampling process, without replacement here, in order to construct a sample (y_1, y_2, ... , y_m).

This very classical procedure is used to get a statistical estimation of the stochastic mean μ and/or variance σ^2 of the distribution of y (the infinite number of 'potential' measurements) starting from the real discrete measurements of y in the sample.

With these assumptions, it can be easily shown that, as soon as the number of measurements m is high enough, higher than 30 in practice, the sample mean $\overline{Y}$ tends to follow a normal law, noted $\mathcal{N}$ here (Central limit theorem), and the sample variance, noted S_s^2, after a scaling by σ^2 / m, follows a Chi-2 law, noted χ^2 here, with (m – 1) degrees of freedom :

$$\begin{aligned} &\overline{Y} : \mathcal{N}\,(y_{mo},\, \sigma^2 / m) \qquad \text{and} \qquad m\, S_s^2 / \sigma^2 : \chi^2\,(m-1) \\ &\text{with :} \qquad \overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \quad \text{and} \quad S_s^2 = \frac{1}{m}\sum_{i=1}^{m} (Y_i - \overline{Y})^2 \end{aligned} \tag{4.15}$$

This allows to find non-biased estimations of μ and σ^2, using the 'hat' notation (^) to designate an estimated value of the corresponding parameter :

$$\hat{y}_{mo} = \bar{y} \qquad \text{et} \qquad \hat{\sigma}^2 = \frac{m}{m-1} s^2 \tag{4.16}$$

with $\bar{y}$ and s^2 the observed values of $\bar{Y}$ and of S_s^2, replacing random variables Y_i by the corresponding measured values y_i.

2.1.2 Measurements for a dynamical model and random vector

We consider now a dynamical model, where measurements y_i of $y_{mo} = \eta\,(t_i, \boldsymbol{x})$ at m times t_i (for $i = 1$ to m) constitute a sampling operation but now the expectation of the y_i measurements varies with times. This means that this sampling is made for a *dynamical* population.

- The first natural idea is to consider each measurement as the realization of a scalar random variable Y_i, of expectation $\mathrm{E}(Y_i) = y_i^{perfect} = \eta_{exact}\,(t_i, \boldsymbol{x}^{exact})$ that is the 'perfect' (noiseless) output of the model whose structure $\eta_{exact}\,(t_i, .)$ is exact as well as the set of parameters gathered in a column vector $\boldsymbol{x}^{exact}$ (see equation 3.1 in Lecture 3 of this series):

$$y_i = y_i^{perfect} + \varepsilon_i \tag{4.17}$$

where ε_i is the measurement noise, that is the difference between the measured signal and the perfect output of the model that corresponds exactly to this measurement.

Let us note that it is impossible de to discriminate in a measurement y_i the contribution of the exact model output $y_i^{perfect}$ from the noise ε_i.

That is why the measurement noise is considered as a random variable. The expectation of the noise is equal to zero for a good measurement chain, and one says then that the (direct) measurement y_i of $y_i^{perfect}$ is unbiased. So, because of (4.17), the experimental signal y_i (of expectation equal to $y_i^{perfect}$) is also a random variable.

One calls $\mathcal{L}$ the probability law $\mathcal{L}$ followed by noise ε_i. One good measuring instrument provides a noise whose standard deviation σ is constant. This probability law is characterized by several parameters: its expectation (that is its stochastic mean, equal to zero here), its variance σ^2 as well as possibly other parameters required to define the probability density function (pdf) $f_{\varepsilon i}$ of this random variable. If one limits oneself to the mean and variance (case of a normal law by example), one notes:

$$\varepsilon_i : \mathcal{L}\,(0, \sigma^2) \;\Rightarrow\; Y_i : \mathcal{L}\,(y_i^{perfect}, \sigma^2) \tag{4.18}$$

A capital character has been used here to designate the random variable Y_i whose

realization at time t_i is the meaured signal y_i : so this measurement is clearly a random variable too.

If the expectation of y_i is different from $y_i^{perfect}$, either the model is wrong (case of a model bias, but we have excluded this assumption above), or the measurement is biased, which means that in average (that is with repetitions) the sensor (with its acquisition chain and its calibration law) does not yield the exact value it is supposed to measure.

- The second idea is to consider the whole set of measurements (a multi-dimensional sample) as a column vector $\boldsymbol{y} = [y_1 \;\; y_2 \;\; \cdots \;\; y_m]^t$, that is the realization of a vector random variable $\boldsymbol{Y} = [Y_1 \;\; Y_2 \;\; \cdots \;\; Y_m]^t$. So, equation (4.17) can be given a vector form:

$$\boldsymbol{y} = \boldsymbol{y}^{perfect} + \boldsymbol{\varepsilon} \tag{4.19}$$

where the noise vector $\boldsymbol{\varepsilon}$ is also a random vector.

In the general case, vector noise $\boldsymbol{\varepsilon}$ defined in (4.19) has a probability law noted $\mathcal{L}$:

$$\boldsymbol{\varepsilon} : \mathcal{L}\ (\mathrm{E}(\boldsymbol{\varepsilon})=0,\ \mathrm{cov}\ (\boldsymbol{\varepsilon}),...) \tag{4.20}$$

where the (symmetrical) variance-covariance matrix $\mathrm{cov}\ (\boldsymbol{\varepsilon})$ is defined by :

$$\mathrm{cov}\ (\boldsymbol{\varepsilon}) = \mathrm{cov}\ (\boldsymbol{Y}) = \begin{bmatrix} \mathrm{var}\ (\varepsilon_1) & \mathrm{cov}\ (\varepsilon_1, \varepsilon_2) & \cdots & \mathrm{cov}\ (\varepsilon_1, \varepsilon_m) \\ & \mathrm{var}\ (\varepsilon_2) & \ddots & \mathrm{cov}\ (\varepsilon_2, \varepsilon_m) \\ & & \ddots & \vdots \\ \text{symétrique} & & & \mathrm{var}\ (\varepsilon_m) \end{bmatrix} \tag{4.21}$$

With this second point of view, the joint probability density function $f_{\boldsymbol{Y}}$ of $\boldsymbol{Y}$ can be defined as followed:

$$\begin{aligned} \mathrm{Prob}\ (y_1 \le Y_1 < y_1 + \mathrm{d}y_1,\ y_2 \le Y_2 < y_2 + \mathrm{d}y_2, \cdots,\ y_n \le Y_n < y_n + \mathrm{d}y_n) = \\ f_{\boldsymbol{Y}}\ (y_1, y_2, \cdots, y_n)\ \mathrm{d}y_1\ \mathrm{d}y_2 \cdots \mathrm{d}y_n = f_{\boldsymbol{Y}}\ (\boldsymbol{y})\ \mathrm{d}y_1\ \mathrm{d}y_2 \cdots \mathrm{d}y_n \end{aligned} \tag{4.22}$$

2.1.3 Example of a binormal joint distribution

We consider here the particular case of a Gaussian distribution. This allows a graphical plot of the joint pdf f_Y, see Figure 21, whose expression is given below:

$$f_Y\ (y_1, y_2) = \frac{1}{2\ \pi\ \sigma_1\ \sigma_2\ \sqrt{1-\rho_{12}^2}}\ exp\left[-\frac{1}{2}\ Q\ (y_1, y_2)\right]$$

with :

$$Q(y_1,y_2)=\frac{1}{(1-\rho_{12}^2)}\left(\frac{y_1-y_{mo}\,(t_1,\boldsymbol{x})}{\sigma_1}\right)^2+\left(\frac{y_2-y_{mo}\,(t_2,\boldsymbol{x})}{\sigma_2}\right)^2-2\,\rho_{12}\left(\frac{y_1-y_{mo}\,(t_1,\boldsymbol{x})}{\sigma_1}\right)\left(\frac{y_2-y_{mo}\,(t_2,\boldsymbol{x})}{\sigma_2}\right) \tag{4.23}$$

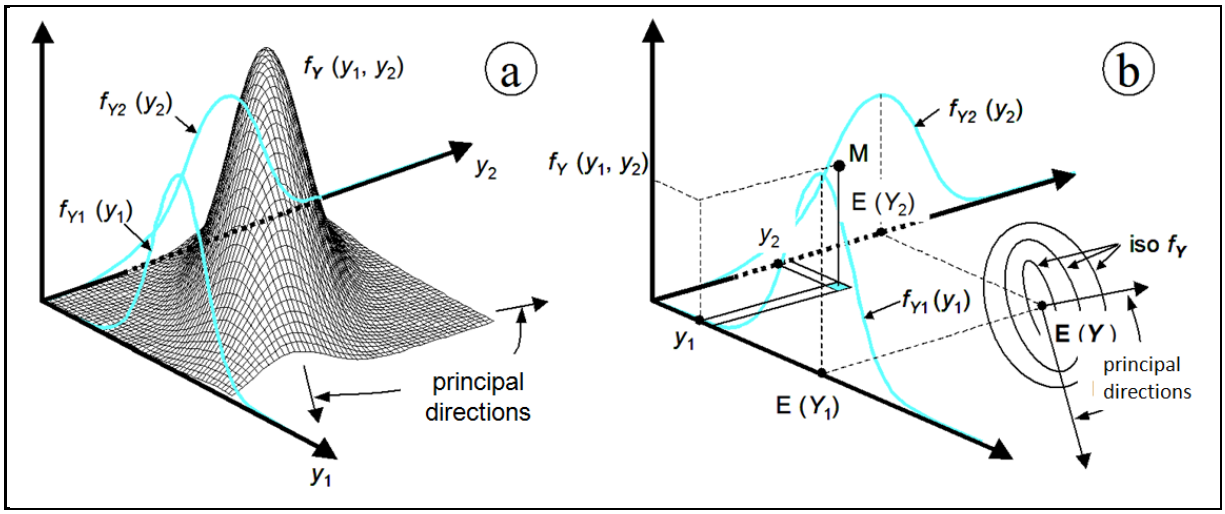

Figure 21 – Binormal law, case of 2 correlated measurements

Three parameters are present in this law : the variances σ_1^2 et σ_2^2 of the two individual random variables Y_1 et Y_2 and their correlation coefficient ρ_{12} ($-1 \leq \rho_{12} \leq 1$). The variance-covariance matrix of $\boldsymbol{Y}$ is defined in Table 1 in the very general case of a random vector $\boldsymbol{Y}$ of size 2 and of joint probability density function $f_Y\ (y_1, y_2)$. This can be easily generalized in the case $m > 2$.

Table 1– Definitions and quantities associated with two random variables

- **Expectation**

$$\mathrm{E}(Y_k) = \int_{-\infty}^{\infty} y_k \, f_{Yk}(y_k) \, \mathrm{d}Y_k = y_{mo}(t_k, \boldsymbol{x}) \quad \text{pour} \quad k = 1 \text{ ou } 2$$

- **Expectation of a random vector**

$$\mathrm{E}(\boldsymbol{Y}) = \begin{bmatrix} \mathrm{E}(Y_1) \\ \mathrm{E}(Y_2) \end{bmatrix} = \begin{bmatrix} y_{mo}(t_1, \boldsymbol{x}) \\ y_{mo}(t_2, \boldsymbol{x}) \end{bmatrix}$$

- **Marginal distributions (probability density functions)**

$$f_{Y1}(y_1) = \mathrm{Prob}\left(y_1 \le Y_1 < y_1 + \mathrm{d}y_1\right) = \int_{-\infty}^{\infty} f_{\boldsymbol{Y}}(y_1, y_2) \, \mathrm{d}y_2$$

$$f_{Y2}(y_2) = \mathrm{Prob}\left(y_2 \le Y_2 < y_2 + \mathrm{d}y_2\right) = \int_{-\infty}^{\infty} f_{\boldsymbol{Y}}(y_1, y_2) \, \mathrm{d}y_1$$

- **Specific case of the binormal law :**

$$f_{Yk}(y_k) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{y_k - \mathrm{E}(Y_k)}{\sigma_k} \right)^2 \right) \qquad \text{for} \quad k = 1 \text{ or } 2$$

- **Variances**

$$\mathrm{var}(Y_k) = \sigma_k^2 = \int_{-\infty}^{\infty} \left(y_k - \mathrm{E}(Y_k)\right)^2 f_{Yk}(y_1, y_2) \, \mathrm{d}y_k \quad \text{for} \quad k = 1 \text{ or } 2$$

- **Covariance**

$$\mathrm{cov}(Y_1, Y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(y_1 - \mathrm{E}(Y_1)\right)\left(y_2 - \mathrm{E}(Y_2)\right) f_{\boldsymbol{Y}}(y_1, y_2) \, \mathrm{d}Y_1 \, \mathrm{d}Y_2$$
$$= \mathrm{E}\left(\left[Y_1 - \mathrm{E}(Y_1)\right]\left[Y_2 - \mathrm{E}(Y_2)\right]\right) = \mathrm{cov}(\varepsilon_1, \varepsilon_2)$$

- **Correlation coefficient**

$$\rho_{12} = \frac{\mathrm{cov}(Y_1, Y_2)}{\sigma_1 \, \sigma_2}$$

- **Variance – covariance matrix**

$$\mathbf{cov}(\boldsymbol{Y}) = \begin{bmatrix} \mathrm{var}(Y_1) & \mathrm{cov}(Y_1, Y_2) \\ \mathrm{cov}(Y_1, Y_2) & \mathrm{var}(Y_2) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho_{12} \, \sigma_1 \, \sigma_2 \\ \rho_{12} \, \sigma_1 \, \sigma_2 & \sigma_2^2 \end{bmatrix}$$

Let us note that, by definition, the expectation of a real noise should be equal to zero ($\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$). This vector noise will be :

- non correlated (or independent), if its variance-covariance matrix is diagonal : $\mathrm{cov}(\varepsilon_i, \varepsilon_k) = \sigma_i^2 \, \delta_{ik}$, where δ_{ik} is Kronecker symbol (null if $i \neq k$, and equal to 1 if $i = k$), where σ_i^2 is the variance of ε_i .

- independent and identically distributed (i.i.d.) if this noise is uncorrelated with a constant variance : $var(\varepsilon_i) = \sigma^2$. In this case its variance-covariance matrix is spherical : $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \, \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix of dimensions (m x m). In this case each component of $\boldsymbol{\varepsilon}$ is independent and follows the same probability law.

2.2 Properties of random vectors

The notion of random vector has been introduced above. In the type of applications concerned by this school, it is the column vector of measurements $\boldsymbol{Y}$, of dimensions (m x 1).

This vector has an expectation, noted E ($\boldsymbol{Y}$), that is a vector of same size, whose coefficients are the expectations of the corresponding coefficients of $\boldsymbol{Y}$. It has also a variance-covariance matrix, noted cov ($\boldsymbol{Y}$), of dimensions (m x m) whose coefficients are the covariances of the coefficients of $\boldsymbol{Y}$:

$$\left[\mathrm{E}(\boldsymbol{Y})\right]_i = \mathrm{E}(Y_i) \qquad \text{and} \qquad \left[\mathrm{cov}(\boldsymbol{Y})\right]_{ik} = \mathrm{cov}(Y_i, Y_k) \tag{4.23}$$

It is always possible to linearly transform signal $\boldsymbol{Y}$ using a linear transformation whose deterministic coefficients are set in a matrix sont rangés $\boldsymbol{G}$ of size (p x m), in order to get a transformed signal $\boldsymbol{Z} = \boldsymbol{G}\,\boldsymbol{Y}$ of size (p x 1). Since $\boldsymbol{Y}$ is a random vector, such is also the case for vectoriel $\boldsymbol{Z}$. The expectation and the variance-covariance matrix of $\boldsymbol{Z}$ depend on the same properies of $\boldsymbol{Y}$:

$$\boldsymbol{Z} = \boldsymbol{G}\,\boldsymbol{Y} \qquad \Rightarrow \qquad \mathrm{E}(\boldsymbol{Z}) = \boldsymbol{G}\ \mathrm{E}(\boldsymbol{Y}) \qquad \text{et} \qquad \mathrm{cov}(\boldsymbol{Z}) = \boldsymbol{G}\ \mathrm{cov}(\boldsymbol{Y})\ \boldsymbol{G}^T \tag{4.24}$$

3. Conclusions

In this short course, the different characteristics of an electronic noise have been detailed and the resulting effect on the signal, that can be considered as the realization of a random vector has been given. The stochastic properties of this random signal vector completely depend on the same properties of the noise vector, if the model used is unbiased. So their knowledge is important if one wants to characterize the estimation error in any experimental inverse problem (see lectures 3, 7 and 8 of this series, for example).