Introduction to noise in laser systems

Xavier Audier

Friday 24^{th} January, 2020

Contents

1 Aim

The goal of this lecture is to introduce the notions of time measurement, noise, power spectral density, noise equivalent power, signal to noise ratio, shot noise, and excess laser noise.

The context is that of laser intensity measurements. In a lot of measuring devices, particularly in spectroscopy, the raw "signal" is the intensity of a laser measured with a photodiode. In order to provide the most precise measurements, one needs to understand what is "signal" and what is "noise". Lasers have specific noise properties that will be discussed, and very precise measurements can be achieved by the use of "lock-in detection". This is the final application discussed here.

Some mathematical elements have to be introduced to fully understand the equations presented here. Mainly Stochastic Processes and Fourier Transforms. They each represent a class in themselves, so the description will be extremely brief. A first reader may want to start with section 2 and read supplementary sections as needed for further understanding of the mathematical tools used.

2 Modeling laser noise

As described above, a number of measuring devices actually measure the intensity of a laser (or other source of light), by means of a photodiode (or other photodetector). A spectrometer, a camera, a position sensor, a CD-DVD reader, optical fiber communication networks, etc ...

We define the optical intensity of the laser as the energy of photons arriving per unit of time: $\mathcal{I}_{opt}(t)$ We can assume that photons arrive in "bursts". In the case of a pulsed laser this is particularly true. For continuous lasers, since the data acquisition device will take discrete time samples we can simplify by assuming that photons arrive together at regular times t_k. The following will therefore assume a pulse laser, but the results are general.

The number of electrons generated by the optical pulse $k \in \mathbb{Z}$ is a random variable X_k which follows a Poisson distribution of mean $N(t_k)$. A brief reminder on random variables and stochastic processes is given in section 5.

The average number of electron generated per pulse is linked to the optical intensity $\mathcal{I}_{opt}(t)$ via the photon-to-electron conversion equation:

$$N(t_k) = \frac{\eta}{h\nu} \int_{pulse\,k} \mathcal{I}_{opt}(t)\mathrm{d}t \tag{1}$$

Where η is the detector quantum efficiency, ν the optical frequency, and h the Plank constant. In addition to the photo-detection events, the detector current noise is modeled with a stochastic current $\epsilon(t)$. This current noise encompasses the detector dark current, the Johnson-Nyquist noise of the load resistor, and any other sources of electrical noise that is generated at the output of the photodiode independently of the optical intensity. With this model, the electrical current $I(t)$ can be expressed by:

$$I(t) = q\Big[\sum_k X_k \delta(t - t_k)\Big] + \epsilon(t) \tag{2}$$

where q is the electric charge produced by a single photo-detection event. The electrical noise $\epsilon(t)$ is considered to be independent from the $\{X_k\}$, and for simplification it will be assumed that the electrical noise is centered around zero: $\langle\epsilon(t)\rangle = 0$. Here, $\langle\cdot\rangle$ stands for the ensemble average, meaning the average over all possible realizations of the measurement given the exact same system in the exact same state. This is analogous to having a large number of identical systems all performing the same measurement (see section 5). For such ensemble of systems, each will record different values for X_k, but these values will be distributed in a Poisson law of average $N(t_k)$. The laser intensity fluctuations over time have therefore two origins:

- The first one is the fluctuation of $N(t_k)$, which arises from generation of optical pulses that are not perfectly identical, and in this lecture is referred to as classical noise (or excess noise).

- The second is the randomness on the measurement and is called shot noise. It causes a fluctuation of the value X_k even when pulses would be perfectly engineered ($N(t_k) = \text{constant}$).

- In addition, electrical noise will pollute the measurement, setting a lower limit to the laser intensity that can be measured.'

From the model described by equation 2, one can compute the power contribution of the DC component and frequency components of the electrical current $I(t)$ through a load resistor R.

3 Signal and Noise in the frequency domain

Mathematically, Fourier transforms (see section 6) can initially be applied only to functions that verify certain properties (such that the integral in Eq 28 converges for all ω). In particular, functions that are regular and non-zero only on a finite domain will work well. However, the Fourier transform can be extended from functions to *distributions* to generalize the results, even when the integral in Eq 28 does not converge. As a basic example, the constant function $h(t) = 1$ does not have a Fourier transform in the sense of functions, but has one in the sense of distributions, and it is the Dirac distribution $\delta(\omega)$ (Fig 7).

When dealing with physical systems, the quantity that is most interesting to watch is energy. Parseval's theorem states that the time signal and its frequency domain equivalent have the same total energy. However, when dealing with functions (signals) that last for an infinite time, the total energy is always going to be infinite, making it hard to manipulate. The workaround is to use *power* instead of energy, as it will be bounded regardless of how long the signal lasts. For that, the Fourier transform is "re-normalized", by dividing it by the duration of the measurement, which then tends to infinity.

3.1 Definitions

Let us consider a quantity A that is a stochastic process. In the following, definitions are made using the general quantity A, although A will be replaced by later by a current I for calculations. Similar derivations could be done using voltages but the discussion, in particular in terms of power, signal, and noise, would be identical.

The finite-time Fourier transform of a quantity $A(t)$ is defined as:

$$\mathcal{F}_T\{A(t)\}(f) \equiv \hat{A}_T(f) \equiv \int_{-T/2}^{T/2} A(t)e^{-2i\pi ft}\mathrm{d}t \tag{3}$$

The time average of $A(t)$ is defined as:

$$A_{avg} \equiv \lim_{T\to\infty} \frac{1}{T}\langle \hat{A}_T(0)\rangle \tag{4}$$

The total power of quantity A is defined as the time average of A^2:

$$P_{tot}[A] \equiv (A^2)_{avg} \tag{5}$$

With such definition, the constant function $h(t) = 2$ has an average value of 2, and a total power of 4.

Using the definition of the time variance $Var[A] \equiv (A^2)_{avg} - A_{avg}^2$, the total power can be split into two components:

$$P_{tot}[A] = A_{avg}^2 + \mathrm{Var}[A] \tag{6}$$

These two terms have two distinct origins, the first is the DC power, the second is the power carried by the fluctuations of A. Typically, when measuring the value of A, the DC power corresponds to the signal power and the power of the fluctuations are referred to as noise power P_{Noise}.

Note that "power" as defined here differs by a constant from the usual electrical power expressed in Watts. The power defined by equation 5 has the dimension of $[A]^2$, where $[A]$ is the unit of A. To obtain an electrical power in W, one needs to multiply by the appropriate factor, such as R if A is a current or $1/R$ if A is a voltage. For instance, the DC power and noise power of a voltage V generating a current I passing through a load resistor R are:

$$P_{DC} = V_{avg}^2/R = R\, I_{avg}^2 \tag{7}$$

$$P_{Noise} = \mathrm{Var}[V]/R = R\,\mathrm{Var}[I] \tag{8}$$

The noise power expressed through the variance does not provide information on the frequency at which the fluctuations of A are happening. This is the purpose of the power spectral density.

3.1.1 Power Spectral Density (PSD)

In order to study how the different frequencies contribute to $\mathrm{Var}[A]$, and therefore to the noise power, the double-sided (positive and negative frequencies) power spectral density of a quantity $A(t)$ is defined as:

$$S_A(f) \equiv \lim_{T\to\infty} \frac{1}{T} \langle |\hat{A}_T(f)|^2 \rangle \tag{9}$$

Because $A(t)$ is usually a real quantity, S_A is an even function of frequency. For this reason, only the positive frequencies are usually considered and the single-sided power spectral density of a quantity $A(t)$ is defined as:

$$S_A^+(f) \equiv 2S_A(f); f > 0 \tag{10}$$

As mentioned above, the name "power spectral density" is ambiguous, as it usually does not have the dimension of a power density, but rather of $[A]^2/Hz$, where $[A]$ is the unit of A. The PSD measures the amount of electrical power (variance) in the signal per unit of bandwidth. For instance, the amount of power coming from the frequency range with width Δf centered around f_0 is given by:

$$P_{\Delta f}(f_0) = \int_{f_0 \pm \Delta f/2} R\, S_I^+(f) \mathrm{d}f \tag{11}$$

3.1.2 Example of PSD

A typical signal and the associated PSD are presentedin Fig 1. The value of the current has a certain average I_{avg} that one is trying to measure. It also has fluctuations around the average value, in the form of a probability distribution of variance $\mathrm{Var}[I]$.

In terms frequency, the variance has components at all frequencies but with different amplitude for each. Typically a laser will have stronger fluctuations at low frequencies, meaning that its power can oscillate significantly over relatively long periods of time. This *low frequency noise* typically drops as $1/f$, meaning that at high frequencies (between consecutive photon packets) the variations will be of much lower amplitude.

At high enough frequencies (typically ≥ 1 MHz) and for stable lasers, the power spectral density can reach the *shot noise*. This is the limit where fluctuations of the laser intensity are due to the stochastic nature of photons. From one sample to the next, the intensity fluctuates not because of the average number of photons changing, but rather because for the same number of average photons the Poisson distribution will give different values.

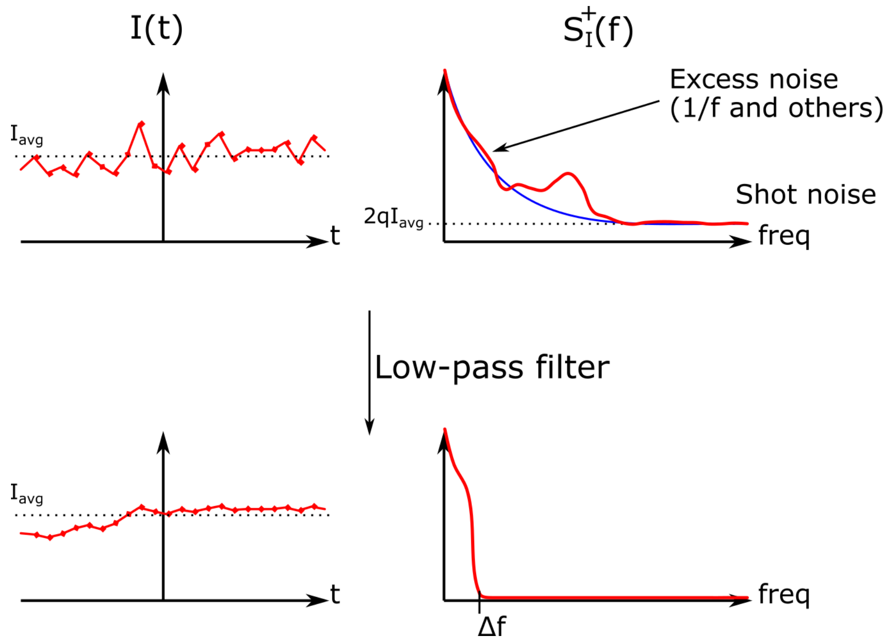

Figure 1: Top: Typical signal from a photodetector and its power spectral density. Bottom: Effect of a low-pass filter.

3.1.3 Relative Intensity Noise (RIN)

The power spectral density divided by A^2_{avg} gives the relative intensity noise (RIN) of quantity A:

$$\mathcal{RIN}_A(f, A_{avg}) \equiv \frac{S_A^+(f)}{A^2_{avg}} \tag{12}$$

The RIN is expressed in "per unit bandwidth" $1/Hz$ and quantifies the relative contribution of each spectral component to the total signal power.

3.1.4 Signal to Noise Ratio (SNR)

When measuring DC value of the quantity $A(t)$, the signal power is given by the DC electrical power A^2_{avg}. The power of the noise is given by equation 11 where the integral covers the bandwidth Δf of the measurement system (typically $f \in [0, \Delta f]$, to measure the DC component). The signal to noise ratio can therefore be expressed as:

$$\mathcal{SNR} = \frac{A^2_{avg}}{\int_0^{\Delta f} S_A^+(f)\mathrm{d}f} \tag{13}$$

3.2 Filtering

Since we are interested in the average value of I, the fluctuations are a nuisance that should be minimized. The simplest approach to optimizing the signal to noise ratio is to reduce the bandwidth

Δf of the system. This can be done by applying an electrical low-pass filter for instance. The effect is illustrated in Fig 1. By removing the noise components away from the DC value ($f = 0$), one significantly reduces the variance of the measurement.

There are two drawbacks to this approach: First, reducing the bandwidth means that the signal cannot change quickly anymore. If one needs to measure a dynamic process (with an intensity that evolves over time) it will be impossible to measure fast changes, as they will be filtered out. Second, the laser noise is highest around $f = 0$, therefore the laser excess noise will always be limiting the measurement, while the theoretical limit is much lower, as given by the shot noise.

Conventional lasers have intensity fluctuations of less than a percent at low frequencies. While this is sufficient for some applications, very precise measurements ($<10^{-4}$) require a different approach (see section 4).

3.3 Calculations

3.3.1 Average and relative intensity

Under the assumption that the electrical noise has a null time average, the time average current I_{avg} can be computed (equation 2, 3, and 4) in terms of the time average number of electrons generated N_{avg}:

$$\begin{aligned}
I_{avg} &= \lim_{T\to\infty} \frac{1}{T} \langle \int_{-T/2}^{T/2} \Big(q \Big[\sum_k X_k \delta(t - t_k) \Big] + \epsilon(t) \Big) e^{-2i\pi f t} \mathrm{d}t \rangle |_{f=0} \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \sum_k q \langle X_k \rangle \delta(t - t_k) \mathrm{d}t \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t_k \in [-T/2, T/2]} q N(t_k) \\
&= q N_{avg}
\end{aligned} \tag{14}$$

The average photocurrent I_{avg} delivered by the photodiode can either be measured directly with an oscilloscope, or derived from the laser average intensity $\mathcal{I}_{opt,avg}$ using the following relationship:

$$I_{avg} = \frac{q\eta}{h\nu} \mathcal{I}_{opt,avg} \tag{15}$$

The relative intensity $\alpha(t)$ is defined as the expected current divided by the average current:

$$\alpha(t) \equiv \frac{\langle I(t) \rangle}{I_{avg}} = \frac{N(t)}{N_{avg}} \tag{16}$$

3.3.2 Power spectral density

The photocurrent power spectral density $S_I(f)$ can be computed by considering the photodetection events statistically independent (see 5.2):

$$\langle X_k X_l \rangle = N(t_k) N(t_l) + N(t_k) \delta_{k,l} \tag{17}$$

Limitation	Electronic	Shot Noise	Excess (Classical) Noise
$\mathcal{PSD} \approx$	$S_\epsilon^+(f)$	$2qI_{avg}$	$S_\alpha^+(f)\, I_{avg}^2$

Table 1: The three measurement regimes and associated dominant power spectral densities.

The concise calculation using the notations used here gives:

$$\begin{aligned}
S_I(f) &= \lim_{T\to\infty} \frac{1}{T} \langle |\hat{I}_T(f)|^2 \rangle \\
&= \lim_{T\to\infty} \frac{1}{T} \langle | q[\sum_k X_k e^{-2i\pi t_k f}] + \hat{\epsilon}_T(f)|^2 \rangle \\
&= S_\epsilon(f) + \lim_{T\to\infty} \frac{q^2}{T} \sum_k \sum_l \langle X_k X_k \rangle e^{-2i\pi(t_k - t_l)f} \\
&= S_\epsilon(f) + \lim_{T\to\infty} \frac{q^2}{T} [\sum_k \sum_l N(t_k) N(t_l) e^{-2i\pi(t_k - t_l)f} + \sum_k N(t_k)] \\
S_I(f) &= S_\epsilon(f) + q^2 [S_N(f) + N_{avg}]
\end{aligned} \tag{18}$$

Using the relative intensity $\alpha(t)$ from equation 16, one obtains a simple expression of the power spectral density for a laser detected on a photodiode:

$$\begin{aligned}
S_I^+(f) &= 2S_\epsilon(f) + 2q^2 N_{avg} + 2q^2 N_{avg}^2 S_{N/N_{avg}}(f) \\
&= S_\epsilon^+(f) + 2qI_{avg} + I_{avg}^2 S_\alpha^+(f)
\end{aligned} \tag{19}$$

Equation 19 illustrates the dependence of the electrical PSD at the output of the detector with respect to both frequency and average current from the detector. The three terms on the right of equations 19 are linked to the electrical noise, the laser shot noise, and the laser excess (classical) noise, respectively (Table 1). The impact of these three terms will be developed further, and measured in the following sections.

At low frequencies ($f < 1\,\text{MHz}$) lasers usually features fluctuations in intensity significantly above the shot noise $S_\alpha^+ \gg 2q/I_{avg}$. Frequently, this noise renders impossible the direct measurements of small intensity fluctuations ($10^{-4} - 10^{-6}$). For this reason, lock-in amplification is typically used to increase the SNR by moving the measurement towards higher frequencies.

3.4 Noise Equivalent Power (NEP)

The electronic noise component $S_\epsilon^+(f)$ defines the limit set by the detector. It is usually relatively constant ($S_\epsilon^+(f) = S_\epsilon^+$). For very low number of photons, the signal amplitude will be on the same level as the electronic noise. Conventionally, detector manufacturers specify amount of average laser intensity $\mathcal{I}_{opt,avg}$ required to achieve a SNR of 1. As can be seen in equation 13, the SNR depends not only on the noise level, but also the bandwidth allowed. For this reason, the noise equivalent power (NEP) is given in Watt per $\sqrt{\text{Hz}}$, and corresponds to the amount of laser intensity required "per unit of bandwidth" of the detector:

$$\text{NEP} = \mathcal{I}_{opt,avg}(SNR = 1)/\sqrt{\Delta f} \tag{20}$$

When the detection is limited by the electronic noise equation 13 with an SNR of 1 gives the relationship between NEP and S_ϵ^+:

$$\text{NEP} = \frac{h\nu}{q\eta} \sqrt{S_\epsilon^+} \tag{21}$$

The power spectral density S_ϵ^+ is rarely used in practice, but the NEP is provided on all devices. The NEP is easier to interpret, as it directly gives, in Watts, the amount of optical power that can be detected by the detector (for a given bandwidth, at a certain wavelength).

4 Lock-in detection : taking advantage of high frequency

We have seen that measuring the average intensity of a laser is sub-optimal, as the DC value is surrounded (in the frequency domain) by high amplitude noise. An interesting way to obtain a more precise measurement is to capitalize on the fact that high frequencies are typically less noisy, and can reach the shot noise limit (Fig 1, top right).

We can use a band-pass filter to select only high frequency components. However, the signal will still sit at $f = 0$ (DC value). The solution is to make the signal oscillate at high frequency.

Example : Stimulated Raman Scattering (SRS) In a standard SRS system (Figure 2), a pump laser beam with optical intensity $\mathcal{I}_p$ and a Stokes beam with optical intensity $\mathcal{I}_s$ are sent through a sample. When collecting $\mathcal{I}_p$ or $\mathcal{I}_s$ after their interactions with the sample, one can measure an intensity loss $\Delta\mathcal{I}_p$ on the pump beam, and an intensity gain $\Delta\mathcal{I}_s$ on the Stokes beam. Without loss of generality, the following will assume that the SRS Stokes beam is collected by the photodiode, while the SRS Pump beam is discarded using an optical filter. After interaction with the sample, the Stokes beam has gained a relative intensity β, proportional to the number of active molecular bonds N in the probed volume, their stimulated Raman cross section σ, and the optical intensity of the SRS pump beam:

$$\beta \equiv \frac{\Delta\mathcal{I}_s}{\mathcal{I}_s} \tag{22}$$

$$\beta \propto N\sigma\mathcal{I}_p \tag{23}$$

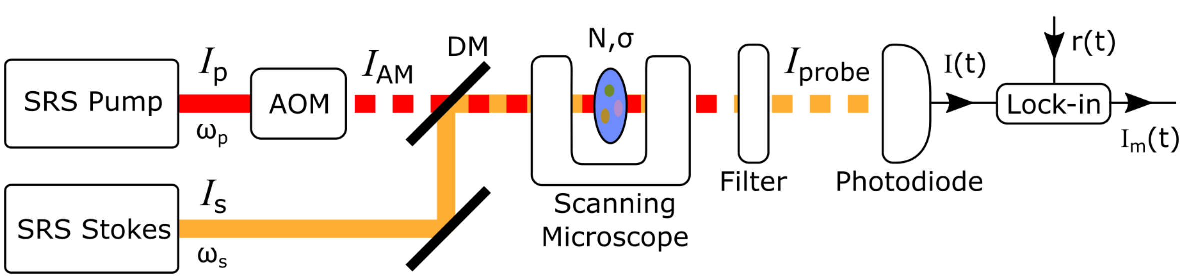

Figure 2: Typical Stimulated Raman Scattering Scanning Microscope system. The SRS pump laser is modulated using an acousto-optic modulator (AOM). It recombines with the Stokes beam using a dichroic mirror (DM) and both are sent on the sample through a scanning microscope. The modulated beam is filtered out using an optical filter and the intensity of the other beam is measured with a photodiode. The signal from the photodiode I is mixed with a reference signal r, filtered and amplified using a lock-in amplifier, which generated an output current I_m.

The amplitude of the relative SRS gain β is typically 10^{-4} to 10^{-6}. In most experimental settings, the limiting noise will be introduced by the probe laser.

As a direct result of equations 22 and 23, if an intensity modulation at frequency f_0 is applied on the SRS pump beam (by the AOM), the change in intensity $\Delta\mathcal{I}$ of the Stokes beam will also be modulated at the same frequency. By applying a band-pass filter, one can discard the noise at low frequency and obtain a much more accurate measurement. The amplitude of the signal then comes in the amplitude of a sinusoidal signal, rather than a DC value (Fig 3).

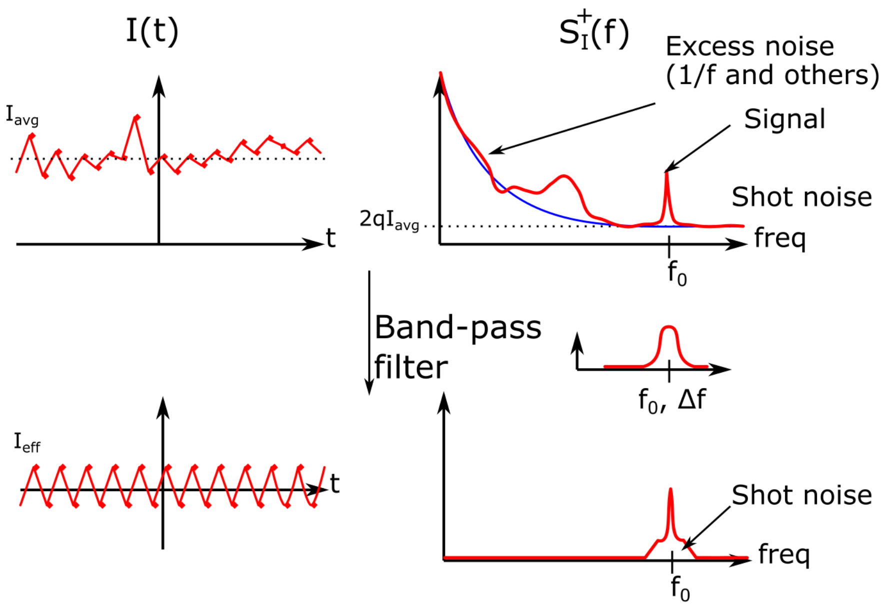

Figure 3: In a lock-in detection SRS system, the signal comes in the form of a modulation at high frequency. By applying a band-pass filter, one can discard the noise at low frequency and obtain a much more accurate measurement.

With such measurement system, the accuracy of a measurement is limited by the shot noise. The smallest variation in amplitude that can be detected is given by

$$\beta_{min} = 4\sqrt{q\Delta f / I_{avg}} \tag{24}$$

For a 10 mW of laser light at 1000 nm (about 5 mA of current), and 1 second of time averaging ($\Delta f \approx 1$Hz):

$$\beta_{min} \approx 10^{-8}$$

Such accuracy cannot be obtained with the previous method (direct DC measurement) as no laser is stable enough.

5 Supplement 1: Stochastic Processes

5.1 Theory

Definition:

> In probability theory and related fields, a stochastic or random process is a mathematical object usually defined as a family of random variables.

Notation : $\{X_t\}_{t\in T}$

$X_0 \dots X_t \dots$, are each a random variable, each following a certain distribution. A stochastic process is called *stationarity* when the random variables are identically distributed, meaning that they share the same distribution. A non stationary process is shown in Fig 4. Random variables can be correlated, for instance, in a Markov process the probability distribution of random variable X_t depends on the value of X_{t-1}.

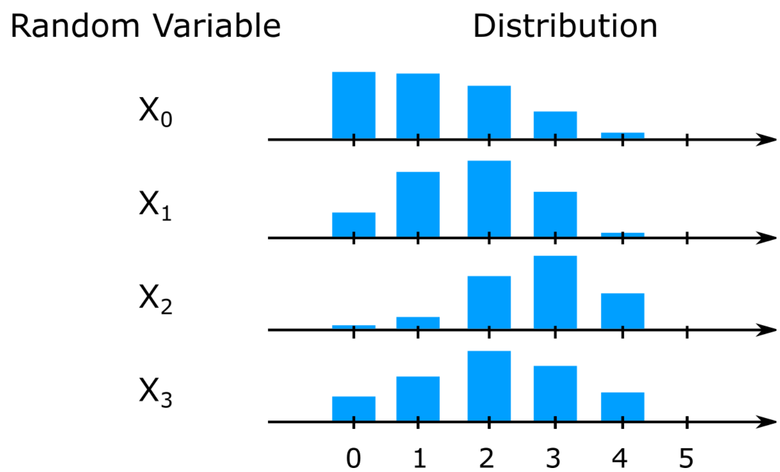

Figure 4: Example of a non-stationary stochastic process

Stochastic processes are used all the time when dealing with measurements. One measurement (series of values) corresponds to a single draw of the stochastic process. The stochastic process corresponds to all possible outcomes of the measurement and their respective probability. The concept of stochastic process can be broken down by imagining that there is an infinite number of exactly identical systems performing the exact same measurement in parallel. A measurement is just one draw in this large number of system working in parallel.

With this concept, one can take the *ensemble average* (usually noted $\langle\cdot\rangle$), meaning the average measurement over all possible outcomes. If the $\{X_t\}_{t\in T}$ are independent of each other, this is the time series made of the mean value of each distribution $\{E[X_t]\}_{t\in T}$.

Another type of averaging can be done by extending the measurement over a long time and performing a *time average*.

The time average and ensemble average may have very different properties, leading to the notion of *ergodicity*. The definition of ergodicity is complicated, and beyond the scope of this lecture, but generally speaking a process is called *ergodic* when measuring its statistical properties can be deduced from a single, sufficiently long, random sample of the process. Meaning that taking one measurement for a long time is equivalent to performing a large number of parallel measurements. For instance, a process made of identically distributed, uncorrelated random variables would be ergodic, since repeated measurements over time is equivalent to performing parallel measurements. An example of a non-ergodic process is presented a the end of this section.

5.2 Examples

Poisson Point Process The process has a given parameter λ. Each X_t follows the statistical distribution :

$$P(X_t = n) = \frac{\lambda^n}{n\,!}\mathrm{e}^{-\lambda} \tag{25}$$

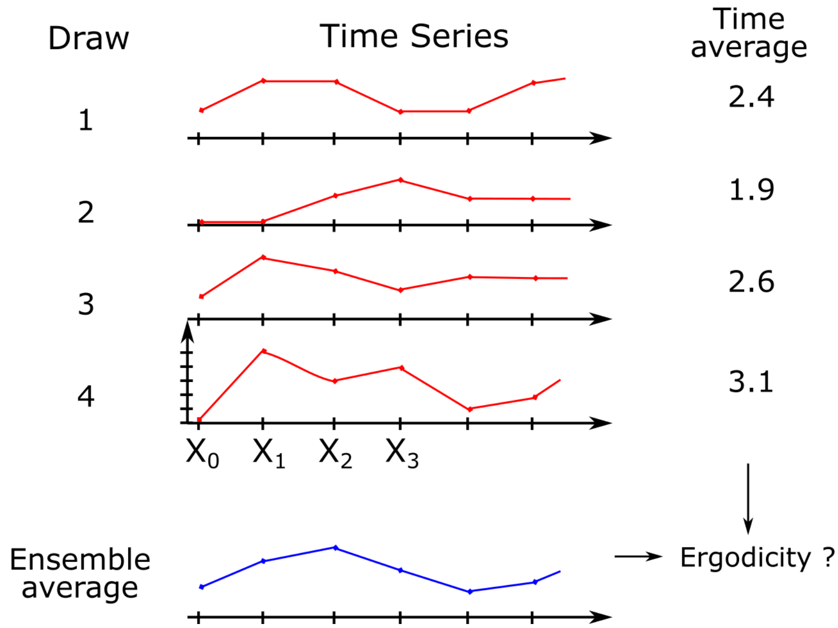

Figure 5: Illustration of different measurements.

The mean value of each random variable is:

$$\forall t \in T, \mathrm{E}[X_t] = \lambda \tag{26}$$

The variance of each random variable is:

$$\forall t \in T, \mathrm{Var}[X_t] = \lambda \tag{27}$$

We can also calculate the value of $\mathrm{E}[X_u X_v]$:

$$\begin{aligned}\forall (u,v) \in T^2, \mathrm{E}[X_u X_v] &= \mathrm{E}[X_u^2] = \lambda^2 + \lambda \qquad \text{if}\, u = v \qquad (\mathrm{E}[X^2] = \mathrm{E}[X]^2 + \mathrm{Var}[X]) \\ &= \mathrm{E}[X_u]\mathrm{E}[X_v] = \lambda^2 \qquad \text{otherwise (independant variables)} \\ &= \lambda^2 + \lambda \delta_{u,v}\end{aligned}$$

Markov Chains A Markov process is a process for which one can make predictions for its future based solely on its present state. A Markov process has a set of N possible states $\{x_i\}_{1 \le i \le N}$. The state of the Markov process at time t (X_t) can be any of those states. The probability of hoping to another state depends only on the current state as well as a transition matrix $\mathrm{A} = \{a_{i,j}\}_{1 \le i,j \le N}$

Markov chains have a lot of interesting properties as well as endless applications. They are used to model a variety of dynamic systems.

Other common processes Gaussian process, Bernoulli process, Random walks, Wiener process (Brownian Motion), etc...

Example of non ergodic process Suppose that we have two coins: one coin is fair and the other has two heads. We choose (at random) one of the coins first, and then perform a sequence of independent tosses of our selected coin. Let X[n] denote the outcome of the nth toss, with 1 for heads and 0 for tails. Then the ensemble average is 1/2(1/2+1) =3/4; yet the long-term average is 1/2 for the fair coin and 1 for the two-headed coin. So the long term time-average is either 1/2 or 1. Hence, this random process is not ergodic in mean.

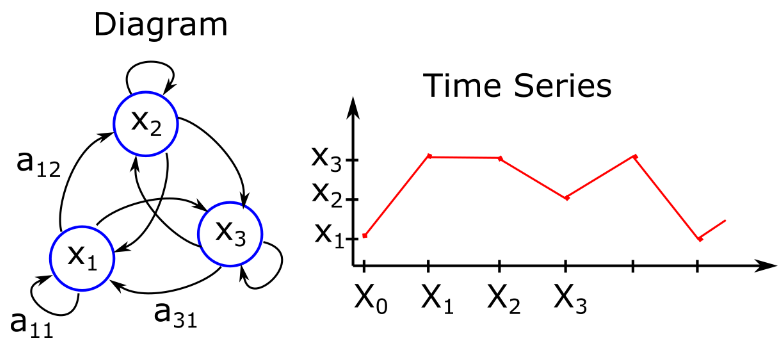

Figure 6: Illustration of a Markov process.

6 Supplement 2: Fourier Transforms

One cannot study time series without a good understanding of Fourier transforms. The Fourier transform (FT) decomposes a function of time into its constituent frequencies. Ideally, several hours should be dedicated to Fourier transform, but the following gives a short summary.

Definition For a given function h, its Fourier transform $\mathcal{F}[h]$ (or $\hat{h}$) is a function defined by:

$$\hat{h}(f) = \int_{-\infty}^{\infty} h(t)\mathrm{e}^{-i2\pi ft}\mathrm{d}t \tag{28}$$

Where $f \in \mathbb{R}$ is called the frequency. Examples of common Fourier transforms are presented in Fig 7. $\omega = 2\pi/f$, the angular frequency, is often used to simplify notations.

Properties

- The Fourier transform is linear:

$$\forall (a,b) \in \mathbb{C}^2, (h,g) \in (\mathbb{R} \to \mathbb{C})^2 \qquad \mathcal{F}[ah+bg] = a\mathcal{F}[h] + b\mathcal{F}[g]$$

- Translating in time corresponds to multiplying by a phase:

$$\mathcal{F}[h(t-t_0)](\omega) = \mathrm{e}^{-j\omega t_0}\hat{h}(\omega)$$

- The Fourier transform swaps function multiplication and convolution:

$$\mathcal{F}[hg] = \mathcal{F}[h] * \mathcal{F}[g] \qquad \mathcal{F}[h*g] = \mathcal{F}[h]\mathcal{F}[g]$$

- Parseval's theorem (the total energy is identical in time domain and frequency domain):

$$\int_{-\infty}^{\infty} |h(t)|^2\mathrm{d}t = \int_{-\infty}^{\infty} |\hat{h}(\omega)|^2\mathrm{d}\omega$$

Recommended links For visualizing complicated math, in particular Fourier Transforms: Youtube Channel "3 Blue 1 Brown"

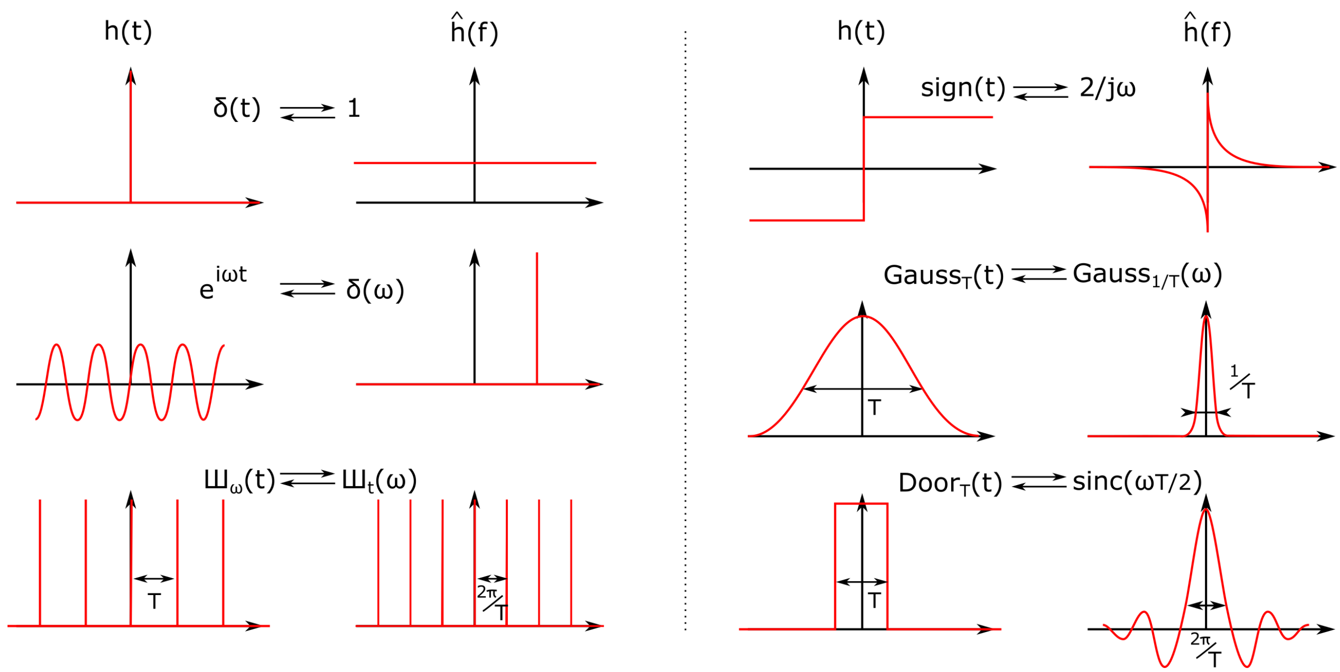

Figure 7: Fourier transforms of common functions. Some multiplicative constants have been omitted for clarity.