

Random Signals, Correlation Functions, and Power Spectra

Textbook pages 841-844; 846-850

Dr D. Laurenson

18th November 2021

1 Introduction

Practical use of the techniques outlined in this course will require the ability to analyse signals with random elements. The majority of practical world signals fall into this category, to one degree or another. For instance, a digital signal may experience jitter, where the timing of transitions between logic levels is not known precisely, a signal may have an amplitude that fluctuates, or the signal may be quantized with a quantization error. One typical example of randomness is the addition of noise, from impulsive sources, or from resistances.

We will begin by defining probabilities, and a number of important properties. Following this, the theory developed for continuous random variables can be applied to the discrete-time case. We will proceed to examine the power density spectrum, before recasting the theory in a discrete-time framework, and then demonstrate how practical signals can be analysed using these techniques.

2 Random Processes

Definitions

A *random process* is the set of all possible waveforms arising from a generator of the waveforms. This is also known as an *ensemble* of time functions. Any one waveform is a *realisation* or a *sample function* of the random process.

- The ensemble is denoted by $X(t)$
- A realisation is denoted by $x(t)$

Probability functions

A set of samples from an ensemble, $X(t)$, can be obtained for a given point in time, t_0 . Those samples can be described, statistically, by a probability density function, $p(X_{t_0})$, or a probability distribution function (also known as a cumulative distribution function), $F(X_{t_0})$.

$$\begin{aligned} F(X_{t_0}) &= \Pr(X(t_0) < X_{t_0}) \\ p(X_{t_0}) &= \frac{d}{dX_{t_0}} F(X_{t_0}) \\ \Rightarrow F(X_{t_0}) &= \int_{-\infty}^{X_{t_0}} p(x) dx \end{aligned}$$

The probability, P , that $X(t_0)$ lies in the range $[A, B]$ is given by:

$$\begin{aligned} P &= F(B) - F(A) \\ &= \int_A^B p(X_{t_0}) dX_{t_0} \end{aligned}$$

Properties

- $F(-\infty) = 0$, and $F(\infty) = 1$
- $F(B) \geq F(A) \forall B > A$
- $p(x) \geq 0$
- $\int_{-\infty}^{\infty} p(x) dx = 1$

2.1 Distributions

There is a very wide number of probability distributions in use, and it is not possible in this course to address any more than a selected few. For interested readers, refer to: <http://dx.doi.org/10.1198/000313008X270448> which has a very impressive map of distribution relationships. Three of the most common, and general, distributions are the uniform, the exponential and the Gaussian (or Normal). We will look at each of these in turn.

Uniform distribution

A uniformly distributed random variable, lying between x_1 and x_2 , is defined as

$$p(x) = \begin{cases} \frac{1}{x_2-x_1} & ; x_1 \leq x \leq x_2 \\ 0 & ; \text{otherwise} \end{cases}$$

All possible values within the range occur with an equal probability. This type of distribution is very useful for characterising errors in measurements where it is only known that the true measurement lies within a specific range. For example, this is the typical form of quantization error for an analogue to digital converter.

Exponential distribution

A normalised exponentially distributed random variable, taking on any positive value, is defined as

$$p(x) = \begin{cases} \exp(-x) & ; x \geq 0 \\ 0 & ; x < 0 \end{cases}$$

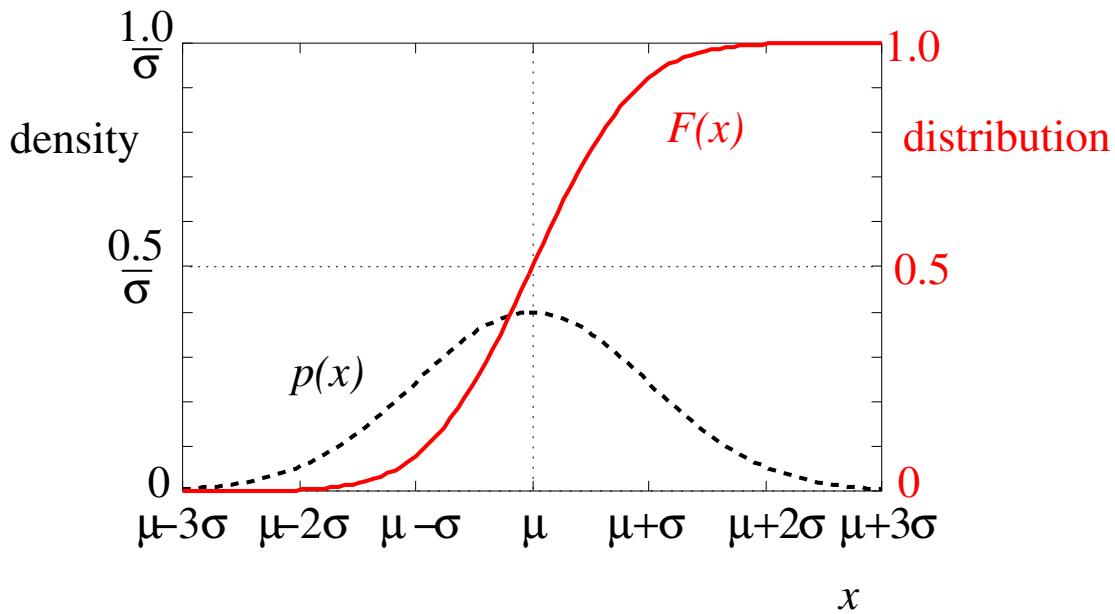
The mean value of this distribution is 1. One common application is describing the timing between two events occurring.

Gaussian distribution

The distribution is defined by

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

where μ is the mean value and σ^2 the variance. There is no closed form for the distribution function, however tables exist for the standard form with $\mu = 0$ and $\sigma^2 = 1$.



The error function, and the complementary error function, can be used when the probability distribution function for a Gaussian random process is required. They are defined as:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$$

and

$$\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$$

The error function is provided as a table in the formula sheet. The table provides values of the error function with an argument to two decimal places. The argument for a particular entry is the sum of the column heading and the row label.

To relate this to a Gaussian distribution, with non-zero mean, and non-unity variance, some manipulation is required. If the probability density function of X is given as $p(X)$, it can be deduced that the probability density function of aX , where a is a scalar constant, is $\frac{1}{a}p\left(\frac{X}{a}\right)$.

Thus, to compute the probability of a Gaussian random variable, with a mean of μ and a variance of σ^2 , lying in the range $(\mu - x, \mu + x)$, is given by $\text{erf}\left(\frac{x}{\sqrt{2}\sigma}\right)$.

Example

Let X be a Gaussian random process, with a mean, $\mu = 4$, and variance $\sigma^2 = 2$. Determine the probability of the random variable being larger than 1.

$$\Pr(X > 1) = \frac{1}{2\sqrt{\pi}} \int_1^\infty \exp\left(-\frac{(x-4)^2}{4}\right) dx$$

sub $m = x - 4$

$$= \frac{1}{2\sqrt{\pi}} \int_{-3}^\infty \exp\left(-\frac{m^2}{4}\right) dm$$

sub $t = m/2$

$$= \frac{1}{\sqrt{\pi}} \int_{-1.5}^\infty \exp(-t^2) dt$$

$$\begin{aligned} \Rightarrow \Pr(X > 1) &= \frac{1}{\sqrt{\pi}} \int_{-1.5}^{1.5} \exp(-t^2) dt + \\ &\quad \frac{1}{\sqrt{\pi}} \int_{1.5}^\infty \exp(-t^2) dt \\ &= \text{erf}(1.5) + \frac{1}{2} \text{erfc}(1.5) \\ &= \text{erf}(1.5) + \frac{1}{2}(1 - \text{erf}(1.5)) \\ &= \frac{1 + \text{erf}(1.5)}{2} = 0.9831 \end{aligned}$$

Equally, by noting that the probability of the value being greater than the mean value is 0.5, the last three lines could be written:

$$\begin{aligned} \Pr(X > 1) &= \frac{1}{\sqrt{\pi}} \int_{-1.5}^0 \exp(-x^2) dx + \frac{1}{\sqrt{\pi}} \int_0^\infty \exp(-x^2) dx \\ &= \frac{1}{2} \text{erf}(1.5) + \frac{1}{2} = 0.9831 \end{aligned}$$

Joint Probability density function

The probability density function can be extended to represent more than one variable, for example a sample function at different times. This is a joint probability distribution function, $p(x_1, x_2, \dots, x_n)$.

From a joint distribution, the *marginal distributions* can be found using integration. The marginal distribution describes the probability distribution when one of the variables is ignored.

It can be determined by integrating over the variable to be ignored:

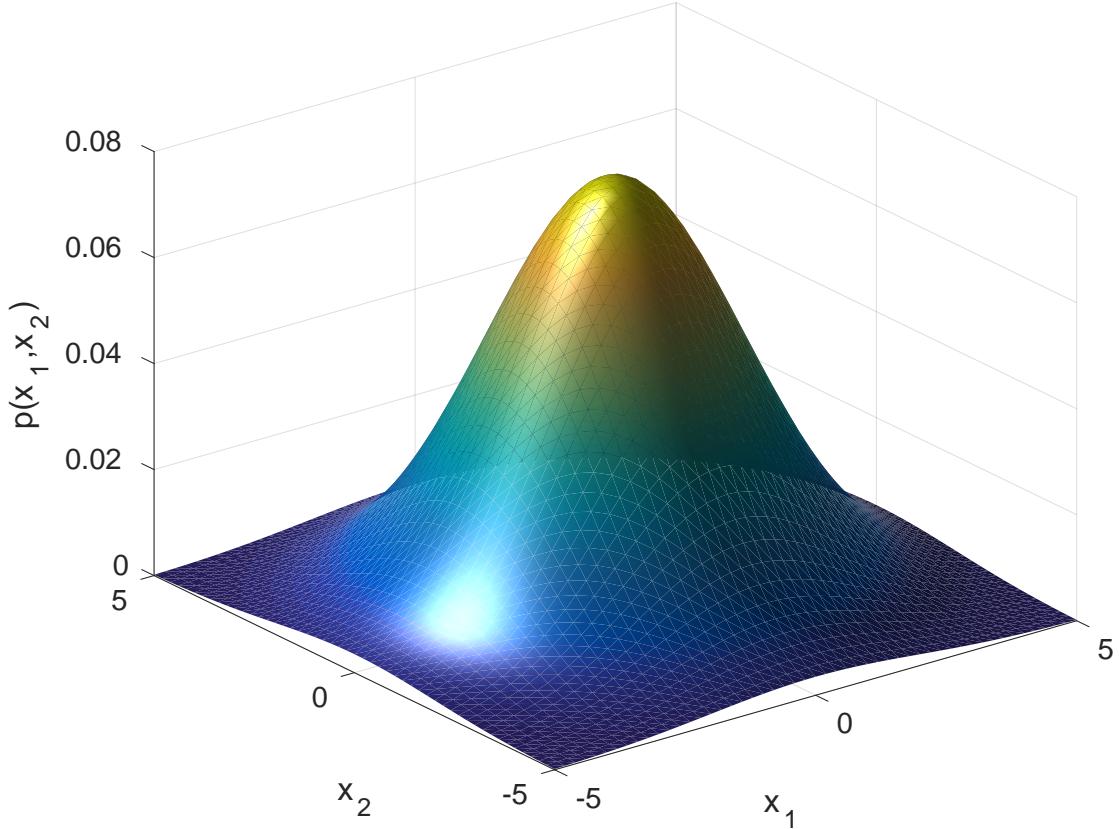
$$p(x_1) = \int_{-\infty}^{\infty} p(x_1, x_2) dx_2$$

Note that this does not include any dependence between the variables, but can be used to describe the statistics of each of the variables in turn.

If the random variables x_1 and x_2 are independent, then their joint distribution $p(x_1, x_2) = p(x_1)p(x_2)$. For example, if x_1 and x_2 are independent and drawn from the same Gaussian distribution, with mean value $\mu = 0$, and variance σ^2 ,

then their joint density function is given by:

$$p(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right)$$



Variables drawn from the same distribution that are independent, are termed independent identically distributed (iid) random variables. This is a very useful property, if it holds, as it allows a number of theories to be developed.

Central Limit Theorem

Let x_i be iid variables drawn from a distribution $p(x)$. Then, with a slight abuse of notation:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} x_i \rightarrow \mathcal{N}(\mu, \sigma^2)$$

where $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian (Normal) distribution with mean $\mu = \bar{X}$, and variance $\sigma^2 = \frac{\text{Var}(X)}{N}$. Thus, for many naturally occurring processes, if many iid variables are being summed, their summation will tend to a Gaussian distribution. This is the fundamental reason that white noise is often modelled as a Gaussian random variable.

1 Stationarity

Define a joint probability density function for a random process, $X(t)$, sampled at time instants t_i , where $i = 1, 2, \dots, n$, as

$$p(x_{t_1}, x_{t_2}, \dots, x_{t_n})$$

A process is strict-sense stationary if

$$p(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = p(x_{t_1+\tau}, x_{t_2+\tau}, \dots, x_{t_n+\tau}) \quad (1.1)$$

for all τ and n . Thus, the joint density function is invariant to any shift in time. The practical implication is that the statistics of the random variable do not depend upon the time at which they are measured. For example, if temperature is kept constant for all time, then thermal noise could be considered to be a stationary process as its statistics will not change over time. The advantage of considering a random process to be stationary is that any statistical measures, or analysis that is being performed on the signal, is independent of the absolute value of time. Thus a delay in processing has no bearing on the statistical analysis.

If a random process has statistics that change over time, then it is classed as being non-stationary. In such cases, care must be taken when finding statistics, as these will vary over time. For example, averaging a process output over too long a period will result in misleading results as the average may change during the analysis. Examples of non-stationary random processes include systems whose performance degrades or improves over time, such as analogue electronic systems, or mechanical systems. Mobile communication systems are also non-stationary due to the movement of nodes between different propagation environments. In practice, no real world systems are strict-sense stationary. However, over a specific time-frame of interest, if the joint density function is invariant, then the process can be considered to be stationary.

2 Ensemble Averages

It is helpful to be able to represent random processes by a set of metrics that can be used in analysis. One of the most straightforward metrics is the average of a process, and a second, its power. These are two examples of *moments* of a random variable, being the first and second moment respectively. Moments are evaluated using the expectation operator, denoted as $E(\cdot)$.

Moments

The l^{th} moment of a random variable is defined as the expected value of $X^l(t_i)$:

$$E(X_{t_i}^l) = \int_{-\infty}^{\infty} x_{t_i}^l p(x_{t_i}) dx_{t_i} \quad (1.2)$$

The first moment, or mean, is then:

$$E(X_{t_i}) = \int_{-\infty}^{\infty} x_{t_i} p(x_{t_i}) dx_{t_i}$$

The second moment, or power is:

$$E(X_{t_i}^2) = \int_{-\infty}^{\infty} x_{t_i}^2 p(x_{t_i}) dx_{t_i}$$

In other words, a moment is computed over all possible values of the random variable, calculating the l^{th} power of each value, and multiplying this by the probability of that value occurring. We can also define the expectation of a function of a random variable:

$$E(f(X_{t_i})) = \int_{-\infty}^{\infty} f(x_{t_i}) p(x_{t_i}) dx_{t_i} \quad (1)$$

where $f(\cdot)$ is any function of the random variable X_{t_i} at time instant t_i .

As an example, consider a uniformly distributed random variable. Let

$$p(x) = \begin{cases} 1 & ; 0 \leq x < 1 \\ 0 & ; \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} E(X^l) &= \int_{-\infty}^{\infty} x^l p(x) dx = \int_0^1 x^l dx \\ &= \left[\frac{1}{l+1} x^{l+1} \right]_0^1 = \frac{1}{l+1} \end{aligned}$$

The mean value is $\frac{1}{2}$, the power is $\frac{1}{3}$.

Now consider a function of a random variable. Let $x(n, \Theta) = \cos(\alpha n + \Theta)$, where Θ is a uniformly distributed random variable in the range $\Theta \in [0, \frac{\pi}{2}]$. Then

$$\begin{aligned} E(x(n, \Theta)) &= \int_{-\infty}^{\infty} x(n; \theta) p(\theta) d\theta \\ &= \int_0^{\pi/2} \cos(\alpha n + \theta) \frac{2}{\pi} d\theta \\ &= \frac{2}{\pi} [\sin(\alpha n + \theta)]_0^{\pi/2} \\ &= \frac{2}{\pi} \{\cos(\alpha n) - \sin(\alpha n)\} \end{aligned}$$

The resulting average is one that is time dependent as it is a function of n . Thus, this random variable is a non-stationary one as its statistics depend upon time.

3 Correlation

Correlation function

Correlation functions for sequences have already been defined ($r_{xy}(l)$ and $r_{xx}(l)$). It is also possible to define the correlation of random variables through use of the joint probability density.

Consider two random variables, X and Y at two points in time, t_1 and t_2 , then their crosscorrelation can be defined as:

$$\begin{aligned} \gamma_{xy}(t_1, t_2) &= E(X_{t_1} Y_{t_2}) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} y_{t_2} p(x_{t_1}, y_{t_2}) dx_{t_1} dy_{t_2} \end{aligned}$$

Unlike the correlation of sequences, this definition of correlation specifies fixed points in time, and finds the relationship between the two random variables. It can be considered to be the result of averaging an ensemble of experiments. In each experiment, a value for X_{t_1} and a value for Y_{t_2} are obtained. These are then compared by taking the product, and the correlation is the average of the resulting comparisons over the ensemble of experiments.

Autocorrelation function

In the same way autocorrelation can be defined:

$$\begin{aligned} \gamma_{xx}(t_1, t_2) &= E(X_{t_1} X_{t_2}) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} x_{t_2} p(x_{t_1}, x_{t_2}) dx_{t_1} dx_{t_2} \end{aligned}$$

This is the autocorrelation function of the random process. Note that in general, x_{t_1} and x_{t_2} are not necessarily independent, so the joint probability density function cannot be represented as the product of two separate density functions for each of the sample points in time. Thus, in general,

$$p(x_{t_1}, x_{t_2}) \neq p(x_{t_1})p(x_{t_2})$$

Relationship to stationarity

If we define $t_2 = t_1 - \tau$, then we can write

$$\gamma_{xy}(t_1, t_1 - \tau) = E(X_{t_1} Y_{t_1 - \tau})$$

If X and Y are jointly stationary, then $\gamma_{xy}(t_1, t_1 - \tau)$ is invariant to changes in t_1 , and instead is only dependent upon the time delay variable τ .

It is then possible to define

$$\gamma_{xy}(\tau) = \gamma_{xy}(t_1, t_1 - \tau)$$

Likewise, if the process is stationary, then

$$\begin{aligned} \gamma_{xx}(\tau) &= E(X_t, X_{t-\tau}) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_t x_{t-\tau} p(x_t, x_{t-\tau}) dx_t dx_{t-\tau} \end{aligned}$$

Autocorrelation properties

For real valued random processes:

- $\gamma_{xx}(-\tau) = \gamma_{xx}(\tau)$
- $\gamma_{xx}(0) = E(X_t^2)$, which is the average power of the random process
- For naturally occurring, non-periodic stationary signals,

$$\lim_{\tau \rightarrow \infty} \gamma_{xx}(\tau) = [E(X)]^2$$

- For periodic signals, the average of $\gamma_{xx}(\tau)$:

$$\overline{\gamma_{xx}(\tau)} = [E(X)]^2$$

Wide-sense stationarity

The definition of strict-sense stationarity is a very tight definition. This can be relaxed through examining the autocorrelation function. If $\gamma_{xx}(t_1, t_1 - \tau)$ is time invariant, i.e. it does not depend upon t_1 , then the process is said to be wide-sense stationary.

It results in the first and second moment being time invariant, and is a definition often used to model practical signals that are stationary over a specified period of time.

Autocovariance function

The autocovariance function is defined as the autocorrelation with the mean value removed:

$$\begin{aligned} c_{xx}(t_1, t_2) &= E\{[X_{t_1} - E(X_{t_1})][X_{t_2} - E(X_{t_2})]\} \\ c_{xx}(t_1, t_2) &= \gamma_{xx}(t_1, t_2) - E(X_{t_1})E(X_{t_2}) \end{aligned}$$

For a stationary random process, $c_{xx}(0)$ is the variance.

In the practical case, we do not know the joint probability function for all pairs of time values, t_1 and t_2 , in advance. Instead of deriving the autocorrelation function from the equations above, the relationship can be defined analytically if an expression for the time series is available, or through an averaging process over a number of realisations of the random process.

Example

Consider the signal $x(t) = \cos(2\pi f_0 t + \theta)$, where θ is an unknown random variable uniformly distributed between $[-\pi, \pi]$. The autocorrelation can be re-expressed in terms of θ as:

$$\begin{aligned} \gamma_{xx}(t_1, t_2) &= \int_{-\pi}^{\pi} \cos(2\pi f_0 t_1 + \theta) \cos(2\pi f_0 t_2 + \theta) p(\theta) d\theta \\ &= \frac{1}{2} \int_{-\pi}^{\pi} (\cos(2\pi f_0(t_1 + t_2) + 2\theta) + \\ &\quad \cos(2\pi f_0(t_1 - t_2))) \cdot \frac{1}{2\pi} d\theta \\ &= \frac{1}{4\pi} \left[\frac{1}{2} \sin(2\pi f_0(t_1 + t_2) + 2\theta) \right]_{\theta=-\pi}^{\pi} + \\ &\quad \frac{\cos(2\pi f_0(t_1 - t_2))}{2} \\ &= \frac{\cos(2\pi f_0(t_1 - t_2))}{2} \end{aligned}$$

This depends only on the time difference between t_1 and t_2 , thus is (at least) a wide-sense stationary random variable.

The Fourier transform of a random process does not exist, as a random process is an infinite-energy signal, and the Fourier Series cannot be used as a random signal is not necessarily periodic. Using the Wiener-Khintchine theorem, which related correlations to density spectra, a frequency domain representation of a random process can be obtained from its autocorrelation function.

Wiener-Khintchine theorem

Recall that $S_{XX}(F) \xleftarrow{F} r_{xx}(m)$. We can derive the power density spectrum of a random process based on the autocorrelation function as:

$$\Gamma_{xx}(F) = \int_{-\infty}^{\infty} \gamma_{xx}(\tau) e^{-j2\pi F\tau} d\tau$$

Similarly, a cross-power density spectrum can be defined relating two random variables:

$$\Gamma_{yx}(F) = \int_{-\infty}^{\infty} \gamma_{yx}(\tau) e^{-j2\pi F\tau} d\tau$$

These are analogous to their non-random counterparts, $S_{xx}(F)$ and $S_{xy}(F)$ respectively.

$\Gamma_{xx}(F)$ describes the power of a random variable in terms of its frequency content. The power in a given frequency band, F_1 to F_2 can be determined using integration:

$$\text{Power} = \int_{-F_2}^{-F_1} \Gamma_{xx}(f) df + \int_{F_1}^{F_2} \Gamma_{xx}(f) df$$

For real valued random variables $\Gamma_{xx}(-F) = \Gamma_{xx}(F)$, thus

$$\text{Power} = 2 \int_{F_1}^{F_2} \Gamma_{xx}(f) df$$

The inverse Fourier relationships automatically hold, so

$$\gamma_{xx}(\tau) = \int_{-\infty}^{\infty} \Gamma_{xx}(F) e^{j2\pi F\tau} dF$$

and

$$\gamma_{xy}(\tau) = \int_{-\infty}^{\infty} \Gamma_{xy}(F) e^{j2\pi F\tau} dF$$

If the random process is a discrete-time one, then the Discrete-time Fourier transform, as opposed to the Fourier transform, provides the necessary relationships.

Example

A random process, X , has an autocorrelation defined by

$$\gamma_{xx}(\tau) = \beta e^{-\alpha|\tau|}$$

Find its density spectrum.

$$\begin{aligned} \Gamma_{xx}(F) &= \int_{-\infty}^{\infty} \gamma_{xx}(\tau) e^{-j2\pi F\tau} d\tau \\ &= \int_{-\infty}^0 \beta e^{\alpha\tau} e^{-j2\pi F\tau} d\tau + \int_0^{\infty} \beta e^{-\alpha\tau} e^{-j2\pi F\tau} d\tau \\ &= \left[\frac{\beta}{\alpha - j2\pi F} e^{(\alpha - j2\pi F)\tau} \right]_{-\infty}^0 \\ &\quad - \left[\frac{\beta}{\alpha + j2\pi F} e^{-(\alpha + j2\pi F)\tau} \right]_0^{\infty} \\ &= \frac{\beta}{\alpha - j2\pi F} + \frac{\beta}{\alpha + j2\pi F} \\ &= \frac{2\alpha\beta}{\alpha^2 + (2\pi F)^2} \end{aligned}$$

Power density spectra properties

A few of the key properties are listed below:

- $\int_{-\infty}^{\infty} \Gamma_{xx}(F) dF = \gamma_{xx}(0) = E(X^2)$
- For real valued random processes, $\Gamma_{xx}(-F) = \Gamma_{xx}(F)$.
- For all random processes, even if they are complex, $\text{Im}(\Gamma_{xx}(F)) = 0$
- $\Gamma_{xy}^*(F) = \Gamma_{yx}(-F)$

Definitions

A discrete-time random signal, $X(n)$, has an ensemble of realisations, $x(n)$. It can be considered to be a sampled version of $X(t)$, where $X(n) = X(n\Delta t)$.

$$\begin{aligned} E((X(n))^l) &= \int_{-\infty}^{\infty} (x(n))^l p(x(n)) dx(n) \\ \gamma_{xx}(n, k) &= E(X(n), X(k)) \\ c_{xx}(n, k) &= \gamma_{xx}(n, k) - E(X(n))E(X(k)) \end{aligned}$$

Thus, the statistical parameters of the discrete-time random signals are analogous to their continuous time counterparts. For wide-sense stationary processes, the autocorrelation and autocovariance can be expressed as $\gamma_{xx}(m)$ and $c_{xx}(m)$, where $m = n - k$.

Summation of random variables

Very often analysis of a mixture of random variables is required, for example a signal plus noise, where $Y(n) = X(n) + W(n)$. Then, because of the linearity of integration:

$$E(Y(n)) = E(X(n)) + E(W(n))$$

Thus, the average of the sum is the sum of the individual averages.

Assuming that $X(n)$ and $W(n)$ are jointly wide-sense stationary:

$$\begin{aligned} \gamma_{yy}(m) &= E((x(n) + w(n))(x(n-m) + w(n-m))) \\ &= E(x(n)x(n-m) + x(n)w(n-m) \\ &\quad + w(n)x(n-m) + w(n)w(n-m)) \end{aligned}$$

As the expectation of a sum is the sum of expectations:

$$\begin{aligned} &= E(x(n)x(n-m)) + E(x(n)w(n-m)) \\ &\quad + E(w(n)x(n-m)) + E(w(n)w(n-m)) \\ &= \gamma_{xx}(m) + \gamma_{ww}(m) + \gamma_{wx}(m) + \gamma_{xw}(m) \end{aligned}$$

This implies that the average power, $E((Y(n))^2) = \gamma_{yy}(0)$ is given by:

$$\begin{aligned} E((Y(n))^2) &= E((X(n))^2) + E(W(n)^2) \\ &\quad + \gamma_{wx}(0) + \gamma_{xw}(0) \end{aligned}$$

A special case arises when $X(n)$ and $W(n)$ are independent:

$$\gamma_{xw}(m) = E(x(n))E(W(n))$$

Thus,

$$\begin{aligned} E((Y(n))^2) &= E((X(n))^2) + E((W(n))^2) \\ &\quad + 2E(X(n))E(W(n)) \end{aligned}$$

In practice we can often evaluate problems involving random variables without having to explicitly deal with the joint probability function. By using the expectation operation, and using the principles of superposition, and independence where it is appropriate, integration may be avoided altogether.

Example

Let $x(n) = \cos\left(\frac{2\pi n}{N} + \theta\right) + w(n)$, where $w(n)$ is additive white Gaussian noise, which may be assumed to be independent of the cosine, and θ is drawn from a uniform distribution over $[0, 2\pi]$.

$$\gamma_{xx}(m) = E(x(n)x(n-m))$$

To begin with, the expression for $x(n)$ is substituted into the expression, and multiplied out.

$$\begin{aligned}\gamma_{xx}(m) &= E\left(\cos\left(\frac{2\pi n}{N} + \theta\right)\cos\left(\frac{2\pi(n-m)}{N} + \theta\right) + \right. \\ &\quad \cos\left(\frac{2\pi n}{N} + \theta\right)w(n-m) + \\ &\quad w(n)\cos\left(\frac{2\pi(n-m)}{N} + \theta\right) + \\ &\quad \left.w(n)w(n-m)\right)\end{aligned}$$

Using the property that $E(A+B) = E(A)+E(B)$, this can be split into a sum of expectations. Also noting that $\cos\left(\frac{2\pi n}{N} + \theta\right)$ and $W(n)$ are independent, then the expectation of their product can be expressed as the product of their expectations.

$$\begin{aligned}\gamma_{xx}(m) &= \frac{1}{2}E\left(\cos\left(\frac{2\pi(2n-m)}{N} + 2\theta\right) + \right. \\ &\quad \left.\cos\left(\frac{-2\pi m}{N}\right)\right) + \\ &\quad E\left(\cos\left(\frac{2\pi n}{N} + \theta\right)\right)E(W(n-m)) + \\ &\quad E(W(n))E\left(\cos\left(\frac{2\pi(n-m)}{N} + \theta\right)\right) + \\ &\quad E(W(n))E(W(n-m))\end{aligned}$$

To simplify further, the expectation of the cosine functions need to be calculated. This involves a simple integration. It should also be noted that $E\left(\cos\left(\frac{-2\pi m}{N}\right)\right) = \cos\left(\frac{-2\pi m}{N}\right)$ as the expression is independent of the random variable.

$$\begin{aligned}E\left(\cos\left(\frac{-2\pi m}{N} + \theta\right)\right) &= \int_0^{2\pi} \cos\left(\frac{-2\pi m}{N} + \theta\right) \frac{1}{2\pi} d\theta \\ &= \frac{1}{2\pi} \left[\sin\left(\frac{-2\pi m}{N} + \theta\right) \right]_0^{2\pi} = 0\end{aligned}$$

Similarly for the other cos terms

$$\Rightarrow \gamma_{xx}(m) = \frac{1}{2} \cos\left(-\frac{2\pi m}{N}\right) + \gamma_{ww}(m)$$

Power Density Spectrum

The power density spectrum is defined as the Fourier transform of the autocorrelation function, thus:

$$\Gamma_{xx}(f) = \sum_{m=-\infty}^{\infty} \gamma_{xx}(m) e^{-j2\pi f m} \quad (1.28)$$

and

$$\gamma_{xx}(m) = \int_{-1/2}^{1/2} \Gamma_{xx}(f) e^{j2\pi f m} df \quad (1.29)$$

For the sum of random variables, $Y(n) = X(n) + W(n)$, assuming they are jointly wide-sense stationary,

$$\Gamma_{yy}(f) = \Gamma_{xx}(f) + \Gamma_{ww}(f) + \Gamma_{xw}(f) + \Gamma_{wx}(f)$$

If $X(n)$ and $W(n)$ are independent, then

$$\Gamma_{yy}(f) = \Gamma_{xx}(f) + \Gamma_{ww}(f) + 2E(X(n))E(W(n))\delta(f)$$

where $\delta(f)$ is a pulse train:

$$\delta(f) = \begin{cases} 1 & ; f \in \mathbb{Z} \\ 0 & ; \text{otherwise} \end{cases}$$

(Reminder, the Fourier transform of a constant is an impulse. Thus, the DTFT, which is periodic in the frequency domain, of a constant is a periodic pulse train).