Gaussian Process

Definition: A real-valued continuous time process $X$ is called a Gaussian process if each finite-dimensional vector $\mathbf{X}_{t_1, t_2, \ldots, t_n}$ has the multivariate distribution $N(\mu_t, \Sigma_t)$, where $\mu_t$ and $\Sigma_t$ are the mean and covariance of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$.

For example, the Wiener process introduced earlier is an example of a Gaussian process. Gaussian processes can also be Markovian. They need not be stationary (a Wiener process is not).
Continuous-Time Linear System

Definition: Let \( x_1(t) \) and \( x_2(t) \) be deterministic linear systems. If \( x_1, x_2 \) be two scalar constants, consider a system \( T \) as shown below.

\[
x(t) \quad \boxed{T} \quad y(t) = T(x(t))
\]

The system \( T(\cdot) \) is linear if

\[
T(a x_1(t) + a_2 x_2(t)) = a_1 T(x_1(t)) + a_2 T(x_2(t)).
\]

We are interested in how a random process behaves when passed through a linear system.

Recall that a random process \( X \) is defined by

\[
X_t(\omega) \quad \text{(for fixed } \omega, X_t(\omega) \text{ is called the sample path for)}
\]

A natural definition would be to study the effect of \( T(\cdot) \) on a sample path by sample path basis.

\[
X_t(\omega) \quad \boxed{T} \quad Y_t(\omega)
\]
Theorem:
Let $X(t)$ be input to a system $L$.

\[ X(t) \xrightarrow{L} Y(t). \]

Then,

\[ E(Y(t)) = E(L(X(t))) = L E(X(t)) = L(\mu_X(t)). \]

Proof:

\[ Y(t) = \int_{-\infty}^{\infty} h(t, \tau) X(\tau) d\tau \quad \text{(this is for a)} \]

\[ \text{fixed } \omega. \]

Note that we are not assuming time-invariance; $h$ is a function of both $t$ & $\tau$.

\[ E(Y(t)) = E \left( \int_{-\infty}^{\infty} h(t, \tau) X(\tau) d\tau \right) \]

\[ = \int_{-\infty}^{\infty} h(t, \tau) E(X(\tau)) d\tau \]

\[ = \int_{-\infty}^{\infty} h(t, \tau) \mu_X(\tau) d\tau \]

\[ = L(\mu_X(t)). \]

Note: The switch of the expectation and the integral has not been justified very rigorously.
Similar results hold for the cross-correlation terms \( R_{XY} \) and \( R_{YY} \). We shall mostly be concerned with WSS processes and LTI systems so we develop results for these cases below.

(i) \( y(t) = \int_{-\infty}^{\infty} h(\tau) x(t-\tau) \, d\tau \) \( (x(t) \text{ is WSS} \), \( h(t) \text{ is LTI} \)).

\[
\Rightarrow \quad \mu_y(t) = \int_{-\infty}^{\infty} h(\tau) \mu_x \, d\tau
\]

\[
= \mu_x \left[ \int_{-\infty}^{\infty} h(\tau) \, d\tau \right] \underset{\text{conservative}}{=} \mu_x \cdot H(\omega) \bigg|_{\omega = 0}
\]

C.E. the mean of \( y(t) \) is the mean of \( x(t) \) scaled by the d.c. gain.

(ii) \( R_{YX}(\tau) = E(\{y(\tau+t) \cdot x(t)\}) \)

\[
= E\left( x(t) \int_{-\infty}^{\infty} h(\tau) x((t+\tau)-\tau) \, d\tau \right)
\]

\[
= E\left( \int_{-\infty}^{\infty} h(\tau) x(t+\tau-x) x(t) \, d\tau \right)
\]

\[
= \int_{-\infty}^{\infty} h(\tau) R_{XX}(\tau-t) \, d\tau
\]

\[
= h(\tau) \ast R_{XX}(\tau)
\]
(iii) $R_{yy}(t) = E(y(t-t_1) y(t))$

\[ = E\left[ \int_{-\infty}^{\infty} h(t) x(t-2-t_1) dt_1 \cdot \int_{-\infty}^{\infty} h(t) x(t-t_2) dt_2\right] \]

\[ = E\left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t_1) h(t_2) x(t-2-t_1) x(t-t_2) dt_1 dt_2\right] \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t_1) h(t_2) R_{xx}(t-2-t_1, t-t_2) dt_1 dt_2 \]

\[ = \int_{-\infty}^{\infty} h(t_1) \int_{-\infty}^{\infty} R_{xx}(t-t_1, t-t_2) dt_2 dt_1 \cdot \left[ h(t) \ast R_{xx}(t) \right]_{t=t-t_1+t-t_2}. \]

Note that \( \int_{-\infty}^{\infty} h(t_1) R_{xx}(t-t_1, t-t_2) dt_1 \) is the convolution of \( h(t) \ast R_{xx}(t) \) evaluated at time \( \tau = t-t_1+t-t_2 \).

Let us denote this by a function \( g(t) \).

The previous equation becomes,

\[ R_{yy}(t) = \int_{-\infty}^{\infty} h(t_1) g(t-t_1) dt_1 \]

Substitute \( x = -t_2 \) to get \( dx = -dt_2 \)

\[ \Rightarrow \int_{-\infty}^{\infty} h(-x) g(t-x) dx = \int_{-\infty}^{\infty} h(-x) g(t-x) dx \]

\[ = g(t) \ast h(-t). \]
\[ R_y(t) = h(t) \ast h(-t) \ast R_{xx}(t) \]

Note: This also shows that when a WSS process is passed through a LTI system, the output process is also WSS.

Example:

Let \( X(t) \) be a WSS process & \( h(t) \) be the differentiation operator

\[
\begin{align*}
X(t) & \quad \underbrace{h(t)} \quad X(t) = \frac{dX(t)}{dt}.
\end{align*}
\]

Then,

\[
M_H(t) = \frac{d}{dt}M_X(t) = 0, \quad \text{(since } M_X(t) \text{ is constant).}
\]

\[
R_{yy}(t) = h(t) \ast h(-t) \ast R_{xx}(t)
\]

\[
= \frac{d}{dt} \left( -\frac{d}{dt} R_{xx}(t) \right)
\]

\[
= -\frac{d^2}{dt^2} R_{xx}(t).
\]
**Power Spectral Density**

Definition: Let \( R_{xx}(t) \) be an autocorrelation function of \( X \). Then, the power spectral density (psd) of \( X \), denoted \( S_{xx}(\omega) \), is given by

\[
S_{xx}(\omega) \triangleq \int_{-\infty}^{\infty} R_{xx}(t) e^{j\omega t} dt.
\]

Note: Sometimes \( S_{xx}(\omega) \) may not exist. In this course we should not worry too much about this.

Thus, \( S_{xx}(\omega) \) is the Fourier transform of \( R_{xx} \).

\[
R_{xx}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) e^{j\omega t} d\omega.
\]

We can also define the cross-psd,

\[
S_{xy}(\omega) \triangleq \int_{-\infty}^{\infty} R_{xy}(t) e^{j\omega t} dt.
\]

**Properties of \( S_{xx}(\omega) \)**

1. \( S_{xx}(\omega) \) is real valued.

Proof:

\[
S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(t) e^{j\omega t} dt = \int_{-\infty}^{\infty} R_{xx}(t) e^{j\omega t} dt = S_{xx}(\omega) \Rightarrow \text{it is real.}
\]
(iii) If $X(t)$ is real-valued (as we have assumed so far), then $S_{xx}(\omega)$ is an even function.

\[ S_{xx}(-\omega) = \int_{-\infty}^{\infty} x(t) e^{i\omega t} dt = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt = S_{xx}(\omega) \quad \text{as it is even.} \]

(iii) $S_{xx}(\omega) \geq 0$. (shall show this later).

$S_{xx}(\omega)$ generally has an interpretation as the density function of average power vs. frequency, as we shall see later. However, $S_{xx}(\omega)$ has no such interpretation in general.

Since $R_{xx}(\omega)$ and $S_{xx}(\omega)$ are Fourier transform pairs, therefore all properties of Fourier transforms hold and can be used to great effect.

**3.5.2. Interpretation**

Let $X(t)$ be real and $x_T(t) \Delta X(t) \cdot I_{[-T,T]}(t)$

where $I_{[-T,T]}(t) = \begin{cases} 1 & -T \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$
Now, \[ F.T.\left(X_T(t)\right) = \int_{-T}^{T} X(t)e^{-j\omega t} \, dt \]

\[ \Rightarrow \left| F.T.\left(X_T(t)\right)\right|^2 = \int_{-T}^{T} \int_{-T}^{T} X(t_1)X(t_2) e^{-j\omega (t_1-t_2)} \, dt_1 \, dt_2 \]

\[ = \int_{-T}^{T} \int_{-T}^{T} X(t_1)X(t_2) e^{-j\omega (t_1-t_2)} \, dt_1 \, dt_2 \]  (since signals are real)

\[ \Rightarrow \frac{1}{2T} E\left(\left| F.T.\left(X_T(t)\right)\right|^2\right) = \frac{1}{2T} \int_{-T}^{T} \int_{-T}^{T} R_{xx}(t_1-t_2) e^{-j\omega (t_1-t_2)} \, dt_1 \, dt_2 \]

Near we introduce a change of variables

\[
\begin{bmatrix}
  s \\
  \tau
\end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} t_1 \\
  t_2
\end{bmatrix}
\]

\[ \Rightarrow J(s, \tau) = \text{needs to be found} \]

\[ \Rightarrow b_1 = \frac{s+\tau}{2}, \quad b_2 = \frac{s-\tau}{2} \]

\[ \Rightarrow J(s, \tau) = \left| \det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \right| = \frac{1}{2} \]

\[ \Rightarrow \quad 0 \quad \text{becomes} \]

\[ \frac{1}{4T} \iint_{R^*} R_{xx}(t) e^{-j\omega t} \, ds \, dt \]

where \( R^* \) is the region
\[
\begin{bmatrix}
S \\
T
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
1 & -1
\end{bmatrix}
\begin{bmatrix}
T_1 \\
T_2
\end{bmatrix}
\]

\[
\frac{1}{4T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_n(t) e^{j\omega t} (4T - 2T) \, dt = \frac{L}{4} \left( e^{j\omega T} \right)
\]
\[
+ \frac{1}{4T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_n(k) e^{j\omega t} (4T - 2T) \, dt = \int_{-\frac{T}{4}}^{\frac{T}{4}} R_n(t) e^{j\omega t} (1 - 2kt) \, dt
\]
\[ \int_{-\infty}^{0} \left\{ \int_{-\infty}^{0} R_{xx}(t) e^{j\omega t} \left[ \int_{0}^{2\pi} ds \right] dT \right\} \]

\[ + \frac{1}{4\pi} \left\{ \int_{0}^{2\pi} R_{xx}(t) e^{j\omega t} \left[ \int_{0}^{2\pi} \frac{2\pi}{\sin(\pi t)} dt \right] \right\} \]

\[ = \int_{-\infty}^{0} \left[ 1 - \frac{1-T}{2\pi} \right] R_{xx}(t) e^{j\omega t} \frac{2\pi}{\sin(\pi t)} dt \]

In the limit \( T \to \infty \), this integral tends to \( \int_{-\infty}^{\infty} R_{xx}(t) e^{j\omega t} dt \).

6. \( S_{xx}(\omega) = \lim_{T \to \infty} \frac{1}{2T} \mathbb{E} \left[ \text{I.F.T.} \left( X_T(t) \right) \right] \)

\( S_{xx}(\omega) \geq 0 \) & real and is related to the power of \( X(t) \) at \( \omega \).

Example

Let \( X(t) \) be white.

\[ R_{xx}(t) = e^{-|t|/\alpha}, \quad -\infty < t < \infty, \quad \alpha > 0. \]

\[ \mathcal{F}_{xx}(\omega) = \int_{-\infty}^{\infty} \frac{e^{-|t|/\alpha}}{\sqrt{2\pi}} e^{-j\omega t} dt \]

\[ = \int_{-\infty}^{0} e^{-(\alpha - j\omega)t} dt + \int_{0}^{\infty} e^{-(\alpha + j\omega)t} dt \]

\[ = \left[ \frac{e^{-(\alpha - j\omega)t}}{\alpha - j\omega} \right]_{-\infty}^{0} + \left[ \frac{e^{-(\alpha + j\omega)t}}{\alpha + j\omega} \right]_{0}^{\infty} \]

\[ = \frac{1}{\alpha - j\omega} + \frac{1}{\alpha + j\omega}. \]
\[
\begin{align*}
E &= \frac{1}{\omega - j\omega} + \left( 0 + \frac{1}{\omega + j\omega} \right) \\
&= \frac{2\omega}{\omega^2 + \omega^2}
\end{align*}
\]

Example
Let \( \nu(t) \) be the white noise process. \( R_{\nu\nu}(\tau) = \delta^2 \delta(\tau) \). The p.s.d. \( S_{\nu\nu}(\omega) = \delta^2 \), \(-\infty < \omega < \infty\). White noise is thus an idealization of a noise process that has equal power at all frequencies. In fact no noise process can actually be white since \( R_{\nu\nu}(0) = \infty \) \( \Rightarrow \) it is a process with infinite variance & hence power. However in communications band-limited white noise is often used as a noise model.
For a deterministic signal $x(t)$.

By Parseval's relation we have that if

$$x(t) \rightarrow X(f)$$

Then,

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(f)|^2 df$$

Now $|x(t)|^2$ has the interpretation of instantaneous power across a 1 Hz bandwidth.

$$\int_{-\infty}^{\infty} |x(t)|^2 dt$$

is the energy of the signal.

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt$$

is the average power.

Now,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |X(f)|^2 df$$

for each $f$.

Let us define a signal

$$x_T(t) = \begin{cases} x(t) & \text{for } -T < t < T \\ 0 & \text{otherwise} \end{cases}$$

and let $X_T(f)$ be the F.T. of $x_T(t)$. 
\[ \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \to \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega \quad \text{(by Parseval's relation)} \]

Accordingly, the energy of the signal in a small band of frequencies \((\omega, \omega + \Delta \omega)\) is given by

\[ \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \Delta \omega \]

For a stochastic signal, the corresponding figure would be

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbb{E}(|X(\omega)|^2) \Delta \omega \]

The limit

\[ \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \]

has the interpretation of the density of power of the signal in the range \((\omega, \omega + \Delta \omega)\).
Recall that,
\[ R_{yx}(t) = h(t) \ast R_{xx}(t) \] when \( h(t) \) is LTI and \( x(t) \) is white noise.

Taking Fourier transforms on both sides, we have

\[ S_{yx}(\omega) = H(\omega) S_{xx}(\omega) \]

Similarly,
\[ R_{yy}(t) = h(t) \ast h(-t) \ast R_{xx}(t) \]

\[ S_{yy}(\omega) = H(\omega) H^*(\omega) S_{xx}(\omega) \]

\[ = S_{xx}(\omega) |H(\omega)|^2. \]

This is because if
\[ h(t) \leftrightarrow H(\omega) \]
\[ h(-t) \leftrightarrow H^*(\omega) \]

This is a fundamental result that characterizes the effect of a linear filter on p.s.d.

\[ x(t) \]
\[ S_{xx}(\omega) \]
\[ h(t) \]
\[ y(t) \]
\[ S_{yy}(\omega) = S_{xx}(\omega) |H(\omega)|^2. \]

Example:

Let \( \mathcal{N}(\omega) \) be a AWGN process such that
\[ S_{\mathcal{N}}(\omega) = \frac{N_0}{2} \quad -\infty < \omega < \infty. \]

Let \( h(t) \) be a brick wall filter band limited to \( -\frac{\omega_0}{2} < \omega < \frac{\omega_0}{2} \).

If \( \mathcal{W}(\omega) \) is the output after filtering \( \mathcal{N}(t) \) through \( h(t) \), we have

\[ S_{\mathcal{W}}(\omega) = \begin{cases} \frac{N_0}{2} & -\frac{\omega_0}{2} < \omega < \frac{\omega_0}{2} \\ 0 & \text{otherwise} \end{cases} \]
Now, the variance of a noise sample from $W(t)$, say $W(t)$,

\( \text{given by} \quad R_W(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{S}_{Wu}(\omega) \, d\omega \)

\[ = \frac{1}{2\pi} \times \frac{N_0}{2} \times 4\pi W \]

\[ = N_0 W. \]
Theorem: Let \( X(t) \) be a WSS process. Then \( S_{XX}(\omega) \geq 0 \) and
\[
\frac{1}{2\pi} \int_{-\omega_2}^{\omega_2} S_{XX}(\omega) \, d\omega
\]
for all \( \omega_2 \geq \omega_1 \).

Proof: Define a filter with transfer function \( H(\omega) \) below & pass \( X(t) \) through it.
\[
H(\omega) = \begin{cases} 
1 & \omega \in (\omega_1, \omega_2) \\
0 & \text{otherwise} 
\end{cases}
\]

\[
S_{YY}(\omega) = \begin{cases} 
S_{XX}(\omega) & \omega \in (\omega_1, \omega_2) \\
0 & \text{otherwise} 
\end{cases}
\]

Now the output power is \( Y(t) = E[Y(t)^2] = R_Y(0) \).

\[
R_{YY}(0) = \frac{1}{2\pi} \int_{-\omega_2}^{\omega_2} S_{YY}(\omega) \, d\omega = \frac{1}{2\pi} \int_{-\omega_2}^{\omega_2} S_{XX}(\omega) \, d\omega
\]

Now choosing \( \omega_2 = \omega_1 + \Delta \omega \) we have

\[
R_{YY}(0) \geq \frac{1}{2\pi} S_{XX}(\omega_1) \Delta \omega
\]

\( \because \) the function \( S_{XX}(\omega) \) has the interpretation of a power density
The positivity was shown earlier.

This automatically leads us to a simple criterion for deciding when a function can be a valid p.s.d.
We just need a function to be:

\[ S_{xx}(\omega) \text{ real and } \geq 0. \]

since

\[
\frac{X(e)}{S_x(x)} \rightarrow S_{xx}(\omega)
\]

we can always find a filter with response \( \sqrt{F(\omega)} \), so that \( S_{xx}(\omega) = F(\omega) \).

In general this may correspond to a complex process \( X(t) \).

If \( F(\omega) \) is even then it corresponds to a real \( X(t) \).

**Cyclostationarity**

Definition. A process \( X(t) \) is said to be cyclostationary if there exists a \( T > 0 \) such that

\[ M_X(t) = M_X(t+T) \quad \text{for all } t, \]

and

\[ C_{xx}(t_1, t_2) = \text{covariance between } X_{t_1} \text{ and } X_{t_2} \]

\[ = C_{xx}(t_1+T, t_2+T) \quad \text{for all } t_1, t_2. \]

Thus the condition of stationarity behaviour is satisfied only periodically, not always.
Example

Let \( x(t) = \cos \left( 2\pi f_c t + \Theta(t) \right) \).

and
\[
\Theta[n] \overset{\Delta}{=} \begin{cases} 
\frac{\pi}{2} & B[n] = 1 \\
-\frac{\pi}{2} & B[n] = 0
\end{cases}
\]

\( B[n] \) is a Bernoulli random sequence with parameter \( \frac{1}{2} \), and
\[
\Theta[n] \overset{\Delta}{=} \Theta[k] \quad \text{for} \quad kT \leq t < (k+1)T.
\]

This corresponds to binary phase-shift keying.

\( T \) is a multiple of \( \frac{1}{f_c} \), so that there are integral number of

\[ E(x(t)) = \frac{1}{2} \cos \left( \omega_c (\sin(\pi f_c t) + \frac{1}{2} \cos(\pi f_c - n) \right) \]

carrier cycles in one bit time.

\[
E_{xx}(t_1, t_2) = E \left( \cos(2\pi f_c t_1 + \Theta(k)) \cos(2\pi f_c t_2 + \Theta(k)) \right)
\]

\[
= \text{Now} \quad \cos(2\pi f_c t_1 + \Theta(k)) = \cos(\Theta(k)) \cos(2\pi f_c t) - \sin(\Theta(k)) \sin(2\pi f_c t).
\]

Define

\[
\text{in-phase) } s_{\text{I}}(t) \overset{\Delta}{=} \begin{cases} 
\cos(\omega_c f_c t) & 0 \leq t \leq T \\
0 & \text{otherwise}
\end{cases}
\]

\[
\text{(quadrature) } s_{\text{Q}}(t) \overset{\Delta}{=} \begin{cases} 
\sin(\omega_c f_c t) & 0 \leq t \leq T \\
0 & \text{otherwise}
\end{cases}
\]

\[
\theta \left( 0, 2\pi f_c t + \Theta(k) \right) = \begin{cases} 
\cos(\Theta(k)) s_{\text{I}}(kT) & k \geq 0 \\
\sin(\Theta(k)) s_{\text{Q}}(kT) & k \leq -2
\end{cases}
\]
\[ E(\sin^2(\theta[k])) = -\frac{1}{2} E(\sin(\theta[k]))^2 = 0. \]

\[ E C_{xx}(t_1, t_2) = E \left[ \sum_{k=-\infty}^{\infty} \sin(\theta[k]) \sin(\theta[k]) \right] = \sum_{k=-\infty}^{\infty} E(\sin(\theta[k]) \sin(\theta[k])) = \sum_{k=-\infty}^{\infty} 0 = 0. \]

Now \[ E(\sin(\theta[k]) \sin(\theta[k])) = \begin{cases} 0 & \text{if } k \neq k' \\ 1 & \text{if } k = k' \end{cases} \]

\[ C_{xx}(t_1, t_2) = \sum_{k=-\infty}^{\infty} \sin(\theta[k]) \sin(\theta[k]) \sin^2(t_1-kT) \sin^2(t_2-kT) \]

Now \( \sin^2 \) has support only if \( k \in [-T, T] \), so only one of the above terms can be non-zero. Thus if \( t_1 \) and \( t_2 \) are not in the same period then the product is zero.

For \( 0 \leq t_1 \leq T \) and \( 0 \leq t_2 \leq T \)

\[ C_{xx}(t_1, t_2) = \delta(t_1, t_2) \sin^2(t_1) \sin^2(t_2) \]

and consequently \( C_{xx}(t_1+T, t_2+T) = \delta(t_1, t_2) \sin^2(t_1) \sin^2(t_2) \).

\( \Rightarrow X(t) \) is cyclostationary.
Mean square calculus for random processes

We consider random processes that can be real or complex-valued.
Most of the development here is for real-valued processes.

Stochastic Continuity

(i) a.s. continuity at $t$

\[ P \left\{ \omega \in \Omega : \lim_{s \to t} X(s, \omega) = X(t, \omega) \right\} = 1 \]

(ii) p. continuity at $t$

\[ \lim_{s \to t} P \left\{ |X(s) - X(t)| > \epsilon \right\} = 0 \]

(iii) m.s. continuity at $t$

\[ E \left[ |X(t+\epsilon) - X(t)|^2 \right] \rightarrow 0 \quad \text{as} \quad \epsilon \to 0. \]

Recall that for a deterministic function $f(x)$, we say that $f$ is continuous at $x_0$ if

\[ \lim_{x \to x_0} f(x) = f(x_0) \]

Furthermore, we say that $\lim_{x \to x_0} f(x) = f(x_0)$ if, given $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x$ with $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \epsilon$. 
We shall primarily work with definition (iii) (m.s. = continuity), since it relates closely to auto-correlation.

**Theorem**: A random process \( X(t) \) is m.s. continuous if \( \rho_X(t_1, t_2, t_3) \) is continuous at \( t_1 = t_2 = t \).

**Proof**: \( E(1X(t_1 + \varepsilon) - X(t_1)^2) = \rho_X(t_1 + \varepsilon, t_1 + \varepsilon) - \rho_X(t_1, t_1 + \varepsilon) \)

\(- \rho_X(t_1, t + \varepsilon) + \rho_X(t, t) \).

Now if \( \rho_X \) is continuous at \( t_1 = t_2 = t \) then the RHS \( \leq |\rho_X(t_1 + \varepsilon, t_1 + \varepsilon) - \rho_X(t_1, t_1 + \varepsilon)| + |\rho_X(t, t) - \rho_X(t_1, t_1)| \)

\( \to 0 \) as \( \varepsilon \to 0 \).

**Two important lemmas**

(i) **Lemma**: Suppose \( X_n \xrightarrow{m.s.} X, Y_n \xrightarrow{m.s.} Y \) then:

\[ E(X_nY_n) \xrightarrow{m.s.} E(XY). \]

*Note*: \( X_n \xrightarrow{m.s.} X \Rightarrow E(X_n^2) < \infty, Y_n \) and \( \lim_{n \to \infty} E(X_n^2) \to 0 \) as \( n \to \infty \).

**Proof**: \( (a+b)^2 \leq 2a^2 + 2b^2 \)

\[ E(X_n + Y_n) \xrightarrow{m.s.} X + Y \]

since \( E((X_n - X + Y_n - Y)^2) \leq 2E(X_n^2) + E(Y_n - Y)^2 \)

\( \to 0 \) \( \to 0 \)

Now, \( X_nY_n = \frac{1}{2} \left[ (X_n + Y_n)^2 - X_n^2 - Y_n^2 \right] \)

\[ \Rightarrow E(X_nY_n) \xrightarrow{m.s.} \frac{1}{2} (E(X + Y)^2 - EX^2 - EY^2) = E(XY). \]
Lemma: Suppose $X_n \xrightarrow{m.s.} X$. Then $E(X_n) \rightarrow E(X)$.

Take $Y_n = 1$ in the above lemma.

Now we show that if $X(t)$ is m.s. continuous at all $t$ then $R_{XX}(t_1, t_2)$ is continuous over $R \times R$ or over the domain of definition, say $(T \times T)$.

Proof:

Let $(s, t) \in T \times T$ and suppose that $(s_n, t_n) \in T \times T$ for all $n \geq 1$, such that $\lim_{n \to \infty} (s_n, t_n) = (s, t)$, i.e. $s_n \to s$ and $t_n \to t$ as $n \to \infty$.

Since $X(t)$ is m.s. continuous, we have

$$X_{s_n} \xrightarrow{m.s.} X_s \quad \text{as} \quad s_n \to s$$

$$X_{t_n} \xrightarrow{m.s.} x_t \quad \text{as} \quad t_n \to t$$

\[ \therefore \quad E(X_{s_n} X_{t_n}) \rightarrow E(X_s x_t) \quad \text{as} \quad n \to \infty. \] (Using the Lemma shown earlier)

\[ \Rightarrow \quad R_{XX}(X_{s_n}, X_{t_n}) \rightarrow R_{XX}(X_s, x_t) \quad \text{as} \quad n \to \infty. \]

Corollary: A wide sense stationary process $X(t)$ is m.s. continuous at all $t$ if $R_{XX}(T)$ is continuous at $T=0$.

We need continuity at $t_1 = t_2$ but this is the same as $T=0$, for a WSS process.
Example: The Brownian process has $E(N_t - W_t^v) = 0^2(t-s)$.

\[ W_s \rightarrow W_t \text{ as } s \rightarrow t \] \[ \text{i.e., it is mean-square continuous.} \]

The auto-correlation function $R_{W}(t,s) = \delta_{min}(t,s)$ is also continuous at the point $(t,t)$.

Example: The Poisson process with rate $\lambda$.

$E(N_t - N_s^2) = \lambda(t-s) + \lambda(t-s)^2 \rightarrow 0 \text{ as } s \rightarrow t$.

\[ = 0 \text{ as } s \rightarrow t \]

This is somewhat contrary to our idea of continuity since the Poisson process exhibits jumps at certain time instants.

\[ \text{--} \]

\[ a_1 a_2 a_3 a_4 \]

\[ = \text{keep in mind that continuity does not imply continuity of sample paths. In fact a simple argument shows that the sample paths of a Poisson process are almost surely not continuous, since} \]

\[ P(N_t \text{ is continuous on } [0,a]) = P(\text{no arrival in } (0,a)) \]

\[ = e^{-\lambda a} \rightarrow 0 \text{ as } a \rightarrow \infty. \]
Definition: The random process $X(t)$ has a mean-square derivative at $t$ if the mean square limit of \( \frac{[X(t+\varepsilon) - X(t)]}{\varepsilon} \) exists as $\varepsilon \to 0$.

This means that the mean square (m.s.) limit exists if there exists a random variable (say $Y$), such that
\[
E \left| Y - \frac{X(t+\varepsilon) - X(t)}{\varepsilon} \right|^2 \to 0 \quad \text{as } \varepsilon \to 0.
\]

Typically we denote the m.s. derivative by $X', \dot{X}^2 \text{ or } dX(t) \text{ or } \dot{X}$.  

In general we may not know what $\dot{X}$ is when we are trying to evaluate it. So we use the Cauchy convergence criterion.

Cauchy criterion for convergence of random variables:

(a) $X_n \xrightarrow{a.s.} X$ iff
\[
P \left\{ \omega : \lim_{m,n \to \infty} \left| X_m(\omega) - X_n(\omega) \right| = 0 \right\} = 1.
\]

(b) $X_n \xrightarrow{P} X$ iff
\[
\lim_{m,n \to \infty} P(|X_m - X_n| > \varepsilon) = 0, \quad \forall \varepsilon > 0.
\]

(c) $X_n \xrightarrow{m.s.} X$ iff
\[
E_x^2 < \infty, X_n \text{ and } \\
\lim_{m,n \to \infty} E \left| X_m - X_n \right|^2 = 0.
\]
Note: We say that for a sequence of numbers $a_{m,n}$, we have
\[ \lim_{m,n \to \infty} a_{m,n} = a \text{ if } \]
for each $\varepsilon > 0$, there exists an integer $N_\varepsilon$ such that for all
$(m,n)$ such that $m > N_\varepsilon$ and $n > N_\varepsilon$ we have
\[ |a_{m,n} - a| < \varepsilon. \]

Note that the Cauchy criterion allows us to determine whether
the limit can exist without explicitly evaluating it.

We now return to the problem of testing whether the
m.s. derivative exists.
This depends upon whether
\[ \lim_{\varepsilon_1, \varepsilon_2 \to 0} \frac{1}{\varepsilon_1} \int \left| \frac{x(t+\varepsilon_1) - x(t)}{\varepsilon_1} - \frac{x(t+\varepsilon_2) - x(t)}{\varepsilon_2} \right|^2 dt \rightarrow 0. \]

**Theorem:** A random process $X(t)$ with autocorrelation $R_{XX}(t_1, t_2)$
has a m.s. derivative at time $t$ if \( \frac{\partial^2 R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2} \) exists at $t_1 = t_2 = t$.

**Proof:**
\[ E \left\{ \left| \frac{x(t+\varepsilon_1) - x(t)}{\varepsilon_1} - \frac{x(t+\varepsilon_2) - x(t)}{\varepsilon_2} \right|^2 \right\} \]
\[ = E \left| \frac{x(t+\varepsilon_1) - x(t)}{\varepsilon_1} \right|^2 + E \left| \frac{x(t+\varepsilon_2) - x(t)}{\varepsilon_2} \right|^2 + 2 E \left( \left( \frac{x(t+\varepsilon_1) - x(t)}{\varepsilon_1} \right) \left( \frac{x(t+\varepsilon_2) - x(t)}{\varepsilon_2} \right) \right), \]
\[ \text{ (1) } \]
Now \[ \mathbb{E} \left[ \left( \frac{X(t+e) - X(t)}{e} \right)^2 \right] = \frac{R_{XX}(t+e, t+e) - R_{XX}(t, t+e) - R_{XX}(t+e, t) + R_{XX}(t, t)}{e^2} \]

\[ \frac{\partial}{\partial t_2} R_{XX}(t_1, t_2) \bigg|_{t_1 = t, t_2 = t} \]

To see this note that for a function\[ f(x_1, x_2) \text{ we have} \]

\[ \frac{\partial}{\partial x_2} f(x_1, x_2) = \lim_{h_2 \to 0} \frac{f(x_1, x_2 + h_2) - f(x_1, x_2)}{h_2} \]

\[ \frac{\partial^2}{\partial x_1 \partial x_2} f(x_1, x_2) = \lim_{h_2 \to 0} \lim_{h_1 \to 0} \frac{f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) - f(x_1, x_2 + h_2) + f(x_1, x_2)}{h_1 h_2} \]

For the first two terms converge to

\[ \frac{\partial}{\partial t_2} R_{XX}(t_1, t_2) \bigg|_{t_1 = t, t_2 = t} \]

The middle term is

\[ \mathbb{E} \left[ \left( \frac{X(t+e, t+e) - X(t, t+e) - X(t+e, t) + X(t, t)}{e^2} \right)^2 \right] \]

\[ \frac{\partial}{\partial t_2} R_{XX}(t_1, t_2) \bigg|_{t_1 = t, t_2 = t} \]

\[ \frac{\partial}{\partial t_2} R_{XX}(t_1, t_2) \bigg|_{t_1 = t, t_2 = t} \text{ goes to zero if } \frac{\partial}{\partial t_2} R_{XX}(t_1, t_2) \text{ exists at } t_1 = t_2 = t. \]
Example: Derivative of Wiener process.

$W(t)$ is a Wiener process with $R_{W(t)}(t_1, t_2) = \frac{t_1^2 - t_2^2}{2} = \frac{(t_1 - t_2) + (t_1 + t_2)}{2}$. We have

$$E \left[ \frac{(W(t_1) - W(t_2))^2}{t_1 - t_2} \right] = \frac{t_1^2 - t_2^2}{2} = \frac{t_1 - t_2}{2} \to \infty \quad \text{as} \quad \text{as} \quad \epsilon \to 0.$$ 

In the m.s. sense, the derivative of the Wiener process does not exist in the ordinary sense. In engineering, we get away by using the Dirac delta function, which is a generalized function.