LECTURE 10

MODELING SIGNAL SETS

(or how to characterize sets of signals in order to subsequently design best operators, in the sense of performing well for (almost) all of them)

10.1 So far we considered the signals (or images, or data vectors) as given, i.e. \( f(t) \) or \( \mathbf{f} \) was a deterministic, specific signal to be represented and acted upon. However in real life we are rarely in the position of having to deal with specific, a-priori known signals or data. Most often we have
to consider classes or types of signals and images, sets of data that may possess some common properties, and in most cases we cannot list or see all the elements of these sets at all. In this case we have to describe or characterize the signal set/class of interest by some general properties. For example we could decide that all functions we shall deal with will be "continuous, square integrable and have derivatives not too high at most places", etc. An important and very useful way to achieve class descriptions for signals is to regard them as "realizations" of a "stochastic process", i.e., as random choices from a set of signals,
or signals whose specifics are determined by some randomly selected parameters.

A random variable is a quantity that assumes values in some set with probabilities assigned to the possible values that may arise. Denoting by $X_w$ a random variable we have, formally, a set of possible indices denoted by $\Omega$ and a realization $\omega_0 \in \Omega$ has probability $p(\omega_0)$. As we have seen in LECTURE 1, we have

$$E[X_w] = \sum_{\omega_k \in \Omega} X_{\omega_k} p(\omega_k) = \mu_w$$

the average value of $w$ and the MSE in replacing $w$ by $\mu_w$ is the variance

$$E[X_w^2] - \mu_w^2 = \sigma_{xw}^2$$
A random process \( \{ \Phi(t) \}_w \in \mathbb{R}^2 \) or \( \{ \Phi(w) \}_w \in \mathbb{R} \) associates a function \( \Phi_w(t) \) to a \( w_0 \in \mathbb{Z} \) or a vector \( \Phi_w \) to an index \( w_0 \in \mathbb{Z} \), and the index \( w_0 \) may be viewed as the random choices made in selecting the "realization" of the continuous or discrete function \( \Phi_w \).

A "genie" may be visualized as selecting the random signal from a box containing the ensemble of possible signals denoted as \( \{ \Phi_w \} \).

\[ \text{a realization of the random process } \{ \Phi_w \} \]
102 Examples of Random Processes

**EXAMPLE 1**

Given a set of vectors signals of length $M$

$$S = \{ \Phi_1, \Phi_2, \ldots, \Phi_M \}$$

as the ensemble of possible signals, select one of them at random, with uniform probability over $\Omega = \{1, 2, \ldots, M\}$, i.e.

$$\Pr \{ \omega = \omega_0 \in \Omega \} = \frac{1}{M},$$

as the realization of the process i.e. $\Phi_{\omega_0}$, $\forall \omega_0 \in \Omega$.

Of course we can also have a set of $M$ signals $\{\Phi_k(t)\}$ for $k = 1, 2, \ldots, M$ selected with probability $\frac{1}{M}$.

**EXAMPLE 2**

$$\omega \in \Omega = [-1, 1], \quad p(\omega) = \frac{1}{2}$$

$$\left( \int p(\omega) d\omega = \frac{1}{2} \int 1 d\omega = 1 \right)$$

$p(\omega) = \frac{1}{2}$

$$\Phi_\omega(t) = \omega \cdot f(t)$$. 
where \( f(t) \) can be any deterministic function over \([0, 1]\), e.g.

\[
\begin{align*}
  f_1(t) &= t \\
  f_2(t) &= t^2 \\
  f_3(t) &= e^{-mt} \\
  f_4(t) &= \sin 2\pi f t / \cos 2\pi f t \\
  \text{etc.}
\end{align*}
\]

**Example 3**

\( \{ \bar{m}_w \}_{w \in \Omega} \) where for \( w_0 \in \Omega \) we select the vector \( \bar{m}_{w_0} \) as follows

\[
\bar{m}_{w_0} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{bmatrix}
\]

where \( m_k \) are realizations of independent, identically distributed random variables with zero mean and variance \( \sigma_n^2 \).

Note that here we have no explicit index set \( \Omega \)!

Suppose all variables \( m_k \) are distributed
according to a (continuous) distribution function $p_m(x)$ then we have that a specific realization of the process, i.e.

$$
\mathbf{\bar{m}}_{\omega_0} = \begin{bmatrix}
\bar{m}_1 \\
\bar{m}_2 \\
\vdots \\
\bar{m}_N
\end{bmatrix}
$$

has a probability density

$$
p(\mathbf{\bar{m}}_{\omega_0}) = \prod_{k=1}^{N} p(\bar{m}_k)
$$

If, for example all $m_k$ are Gaussian i.i.d random variables, we obtain

$$
p(\mathbf{\bar{m}}_{\omega_0}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{\bar{m}_k^2}{2\sigma^2}} = \frac{1}{\sqrt{(2\pi \sigma^2)^N}} e^{-\frac{\mathbf{\bar{m}}_{\omega_0}^T \mathbf{I} \cdot \mathbf{\bar{m}}_{\omega_0}}{2\sigma^2}} = \frac{1}{\sqrt{(2\pi \sigma^2)^N}} e^{-\frac{1}{2} \mathbf{\bar{m}}_{\omega_0}^T \Sigma^{-1} \mathbf{\bar{m}}_{\omega_0}} = \frac{1}{\sqrt{(2\pi \sigma^2)^N}} e^{-\frac{1}{2} \mathbf{\bar{m}}_{\omega_0}^T \Sigma^{-1} \mathbf{\bar{m}}_{\omega_0}}
$$

with $\Sigma = \sigma^2 \mathbf{I}$, $|\Sigma| = (\sigma^2)^N$. 

This example is very important. The process \( \{W_n\} \) defined as above is called the (Gaussian) White Noise process. It takes values in \( \mathbb{R}^{\mathbb{N}} \) (\( \mathbb{N} \)-vectors are realizations of it) and the probability density over \( \mathbb{R}^{\mathbb{N}} \) is, as we have seen, determined by the matrix \( \Sigma = \sigma^2 I \).

**Example 4**

\( \{\phi_w(t)\}_{w \in \mathbb{Z}} \) is defined as follows:

- \( \omega \) sets three random parameters \((u, a, b)\) and
- \( \phi_w(t) = \begin{cases} a & \text{if } t \leq u \\ b & \text{if } t > u \end{cases} \)

with
- \( p(u) \) - uniform over \([0, 1]\)
- \( p(a) \) - a given distribution over \([-1, 1]\)
- \( p(b) \) - a given distribution over \([-1, 1]\)

\( u, A, B \) are independent random variables with realizations \((u, a, b)\).
Examples

\{ \psi_w(t) \}_{w \in \mathbb{Z}} \text{ defined as follows}

\psi_w(t) = a \cos(2\pi ft + w)

where a, f are fixed parameters
and w is a random "phase", uniformly
distributed over the interval [0, 2\pi).

Example 6

\{ \rho_w(t) \}_{w \in \mathbb{Z}} \text{ defined as follows}

\rho_w(t) = \sum_{k \in \text{Index set}} w_k \beta_k(t)

Here w is a vector of random variables
\overline{w} = [w_1, w_2, \ldots, w_k, \ldots]. They may be
assumed to be independent or not, and one
must have the joint pdf of the variables
w_k | k \in \text{Index set}, i.e. the multivariate pdf of \overline{w}.

\{ \beta_k(t) \}_{k \in \text{Index set}} \text{ is a family of deterministic functions}
Example 7

\[
\left\{ \{ \vec{c}_w \} \right\} = \left\{ \sum_{k=1}^{M} w_k \vec{\beta}_k \right\} \quad \forall \vec{w} = [w_1, w_2, \ldots, w_M] \in \mathbb{R}^M
\]

This example is the discrete version of Example 6, with \( \vec{w} \) a random vector and \( \{ \vec{\beta}_k \} \) for \( k = 1, 2, \ldots, M \) a set of deterministic (vector) signals.

Note here that the vectors \( \{ \vec{\beta}_k \} \) for \( k = 1, 2, 3, \ldots, M \) are not necessarily orthogonal to each other (but we can assume that they are normalized, so that \( \vec{\beta}_k^* \vec{\beta}_k = \langle \vec{\beta}_k, \vec{\beta}_k \rangle = 1 \), i.e. their length \( \sqrt{2} \) is 1.) Note that we can have \( M > N \), or \( M < N \) or \( M = N \), and in each case \( \vec{\beta}_k \in \mathbb{R}^N \).
10.3 For a random process \( \{ p_w(t) \} \) or \( \{ \bar{p}_w \} \) we seldom have the
most complete descriptions as given in the examples considered. Then, we must rely
and exploit some statistics that we may have about the ensemble of signals.

The most important statistics for a random (stochastic) process are the ensemble average
and the auto-correlation. These are defined as follows.

\[(a) \quad E_w \{ p_w(t) \} = \langle p(t) \rangle = \text{average over the ensemble (i.e. w.r.t. the choices w) with the corresponding probabilities) of the values of } p_w(t) \text{ at } t \]

\[(b) \quad E_w \{ \bar{p}_w \} = \bar{p} = \text{a vector with entries} \quad E_w(p_{w(k)}) \]
\( \mathbb{E}_w \{ p_w(t_1) p_w(t_2) \} = R_p(t_1, t_2) \)

\( \equiv \) average of the product at two points in time \((t_1, t_2)\) of the values of \(p_w(t)\) w.r.t. the choices of realizations \((w)\)

\( \mathbb{E}_w \overline{p_w} \overline{p_w}^* = R_p \) - a matrix with entries \( \mathbb{E}_w \{ p_w(t) p_w(t)^* \} \)

Note that the ensemble averages are deterministic functions or vectors and we shall always assume that for a random process these are known (given, or measured a-priori) hence we shall "center" the processes, i.e. subtract their average to consider the processes of "deviations from the average", i.e. do the following:
Consider \( \phi_w(t) - \phi(t) \triangleq \phi_{\text{centered}}(t) \)

\[ \overline{\phi_w} - \overline{\phi} \triangleq \overline{\phi_{\text{centered}}} \]

and for the centered processes we have

\[ E_w \{ \phi_{\text{centered}}(t) \} \cdot = E_w \{ (\phi_w(t) - \phi(t)) \} = E_w \{ \phi_w(t) \} - E_w \{ \phi(t) \} = 0 \]

\[ E_w \{ \phi_{\text{centered}} \} \cdot = E_w \{ \overline{\phi_w} - \overline{\phi} \} = \frac{E_w \{ \overline{\phi_w} \} - \overline{\phi}}{E_w \{ \phi_w \} - \overline{\phi} = 0} \]

and

\[ E_w \{ (\phi_w(t_1) \phi(t_2)) \} \cdot = E_w \{ (\phi_w(t_1) - \phi(t_1))(\phi_w(t_2) - \phi(t_2)) \} = E_w \{ \phi_w(t_1) \phi_w(t_2) - \phi_w(t_1) \phi(t_2) - \phi(t_1) \phi_w(t_2) + \phi(t_1) \phi(t_2) \} = R(t_1, t_2) - \phi(t_1) \phi(t_2) \triangleq R_{\phi}(t_1, t_2) \]

\[ E_{\{ \phi_{\text{centered}} \} \cdot \phi_{\text{centered}}^*} \cdot = E(\overline{\phi_w} - \overline{\phi})(\overline{\phi_w} - \overline{\phi})^* \cdot = R_{\phi} - \overline{\phi} \overline{\phi} = \overline{R_{\phi_{\text{centered}}}} \]
In the future developments we shall always assume that the processes we consider are centered, i.e. have zero mean.

For the examples considered we have the following ensemble averages and autocorrelations.

**Ex 1:**
\[ E_w \{ \tilde{\omega} \tilde{\omega}^* \} = \frac{1}{M} \sum_{k=1}^{M} \tilde{\omega}_k \tilde{\omega}_k^* = R_\omega \]

**Ex 2:**
\[ E_w \{ \tilde{\omega}_1(t) \tilde{\omega}_1(t) \} = E_w \{ \omega^2 \} = \int_{-\infty}^{\infty} \frac{1}{2} \omega^2 \, d\omega = \left( \frac{1}{2} \omega^2 \right) \bigg|_{-\infty}^{\infty} = 0 \]

\[ E_w \{ \tilde{\omega}_1(t) \tilde{\omega}_2(t) \} = \int_{-\infty}^{\infty} \omega_1(t) \omega_2(t) \, d\omega = E_w \{ \omega \tilde{\omega} \} \]
Ex 3:

\[ \mathbb{E}_{\omega} \{ \bar{m}_{\omega} \} = 0 \]

\[ \mathbb{E}_{\omega} \{ \bar{m}_{\omega} \bar{m}_{\omega}^* \} = \sigma_m^2 \mathbb{I} \]

since \( \mathbb{E}_{\omega} \{ m_{\omega} m_{\omega}^* \} = \left\{ \begin{array}{ll}
\mathbb{E}_{\omega} \{ m_{\omega}^2 \} & \text{for } l = k \\
\mathbb{E}_{\omega} \{ m_{\omega} m_{\omega}^* \} \mathbb{E}_{\omega} \{ m_{\omega} \} & \text{for } l \neq k
\end{array} \right. \)

\[ = \left\{ \begin{array}{ll}
\sigma_m^2 & \\
0 & \text{since } m_{\omega} \text{ are i.i.d. with zero mean and variance } \sigma_m^2
\end{array} \right. \]

Ex 4

\[ \mathbb{E}_{\omega} \{ \phi_\omega (t) \} = \]

\[ = \mathbb{E} \{ a_j \mathbb{P}(U > t_j) + b_j \mathbb{P}(U < t_j) \} = \]

\[ = \left( \int_{t_j}^{+\infty} a_j p_A(a) \, da \right) (1,t_j) + \left( \int_{-\infty}^{t_j} b_j p_B(b) \, db \right) (1,t_j) \]

\[ = \left( \int_{t_j}^{+\infty} a_j p_A(a) \, da \right) (1,t_j) + \left( \int_{-\infty}^{t_j} b_j p_B(b) \, db \right) (1,t_j) \]

\[ \mathbb{E}_{\omega} \{ \phi_\omega (t_1) \phi_\omega (t_2) \} = \]

\[ = \mathbb{E}_{\omega} \{ \begin{cases} a^2 & \text{if } t_1 < u < t_2 \\ ab & \text{if } t_1 < u < t_2 \\ b^2 & \text{if } t_1 > t_2 > u \end{cases} \} = \]
\[
\begin{align*}
\int_0^1 & \left[ a^2 p_A(a) da \right] \left( 1 - \max(t, t_2) \right) + \\
& + \left[ \int_{-1}^1 a b p_A(a) p_B(b) dab \right] \left( \max(t, t_2) - \min(t, t_2) \right) + \\
& + \left[ \int_{-1}^1 b^2 p_B(b) db \right] \left( \min(t, t_2) \right)
\end{align*}
\]

Assuming \( E\{a_j\} = E\{b_j\} = 0 \)
and \( E\{a_j^2\} = E\{b_j^2\} = \sigma^2 \)

We have:

\[
E \{ p_w(t) \} = 0
\]

\[
R_{p(t)}(t_1, t_2) = \sigma^2 \left( 1 + \min(t_1, t_2) - \max(t_1, t_2) \right)
\]

\[
= \sigma^2 \left( 1 - t_2 - t_1 \right)
\]

Note here the important fact that
\( R_{p(t)}(t_1, t_2) \) here depends only on \( \tau = (t_2 - t_1) \)
i.e. the autocorrelation is a function of one variable \( \tau \), measuring the gap between \( t_1 \) and \( t_2 \).
For this example, we have

\[
E \{ \phi_w(t) \} = E_0 \frac{1}{2} a \cos(2\pi ft + \omega)
\]

\[
= a \int_0^{2\pi} \cos(2\pi ft + \omega) \frac{1}{2\pi} d\omega = 0
\]

(average of cosine over a period!)

and

\[
E_0 \{ \phi_w(t_1) \phi_w(t_2) \} = R_0(t_1, t_2)
\]

\[
= E_0 \left\{ a^2 \frac{1}{2\pi} \cos(2\pi ft_1 + \omega) \cos(2\pi ft_2 + \omega) d\omega =
\right.

\[
= \frac{a^2}{2\pi} \int_0^{2\pi} \left[ \frac{1}{2} \cos(2\pi ft_1 - 2\pi ft_2) + \frac{1}{2} \cos(2\pi ft_1 + 2\pi ft_2 + 2\pi) \right] d\omega
\]

\[
= \frac{a^2}{2\pi} \int_0^{2\pi} \frac{1}{2} \cos(2\pi f(t_1 - t_2)) d\omega +
\]

\[
+ \frac{a^2}{2\pi} \int_0^{2\pi} \frac{1}{2} \cos(2\omega + 2\pi f(t_1 + t_2)) d\omega =
\]

\[
= \frac{a^2}{2\pi} \cos(2\pi ft_1 - t_2) \cdot \frac{2\pi}{2} = \frac{a^2}{2} \cos(2\pi f(t_1 - t_2))
\]

Here again we see that the autocorrelation depends on

\[
R_0(t_1, t_2) = \frac{a^2}{2} \cos(2\pi ft_1 - t_2) \quad \tau = t_2 - t_1
\]
Ex 6

Here we have, in general

\[
E_w \{ \mathbf{o}_w(t) \mathbf{y} \} \Delta \mathbf{c}_{ar}(t) = E_w (\sum_k \omega_k \beta_k(t))
\]

\[
= \sum_k E_w \{ \omega_k \} \beta_k(t)
\]

and

\[
E_w \{ \mathbf{o}_w(t_1) \mathbf{o}_w(t_2) \} \Delta \mathbf{R}_w(t_1, t_2) =
\]

\[
= E_w \left( \sum_k \omega_k \beta_k(t_1) \sum_l \omega_l \beta_l(t_2) \right) =
\]

\[
= \sum_k \sum_l E_w \{ \omega_k \omega_l \} \beta_k(t_1) \beta_l(t_2)
\]

where we have used linearity.

If the \( \omega_k \) random variables are independent and have zero mean we get:

\[
E_w \{ \mathbf{c}_w(t) \mathbf{y} \} = \mathbf{c}_{ar}(t) = 0
\]

and

\[
\mathbf{R}_w(t_1, t_2) = \sum_k E_w \{ \omega_k^2 \} \beta_k(t_1) \beta_k(t_2)
\]
The most interesting case is, however, if we have:

\[ p_w(t) = \sum_{k=-N}^{N} w_k e^{i2\pi ft} \]

with \( w_k \)'s independent, zero mean random variables with variances \( \sigma_k^2 \). Then we obtain

\[ \mathbb{E}_w \{ p_w(t) \} = 0 \]

and

\[ R(t, t_2) = \sum_{k=-N}^{N} \sigma_k^2 e^{i2\pi ft_1} e^{-i2\pi ft_2} = \sum_{k=-N}^{N} \sigma_k^2 e^{i2\pi f(t_1 - t_2)} \]

Again, a process whose autocorrelation depends only on \( \tau = t_2 - t_1 \).
Ex 7

For this case we have

\[ E \{ \mathbf{\Phi}_w y \} = E \{ \sum_{k=1}^{N} \omega_k \mathbf{\beta}_k y \} = \sum_{k=1}^{N} E \{ \omega_k y \} \mathbf{\beta}_k \]

and \[ E \{ \mathbf{\Phi}_w \mathbf{\Phi}_w^* y \} = E \{ (\sum_{k=1}^{N} \omega_k \mathbf{\beta}_k^*) (\sum_{k=1}^{N} \omega_k \mathbf{\beta}_k^*)^* \} = \sum_{k=1}^{N} \sum_{l=1}^{N} E \{ \omega_k \omega_l^* y \} \mathbf{\beta}_k \mathbf{\beta}_l^* = R_6 \]

Assuming we are zero mean and independent random variables we obtain here:

\[ E \{ \mathbf{\Phi}_w y \} = \mathbf{0} \]

\[ R_6 = E \{ \mathbf{\Phi}_w \mathbf{\Phi}_w^* y \} = \sum_{k=1}^{M} \sigma_k^2 \mathbf{\beta}_k \mathbf{\beta}_k^* \]

\[ = \begin{bmatrix} \mathbf{\Sigma}_{\sigma_1^2} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{\sigma_2^2} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Sigma}_{\sigma_{M}^2} \end{bmatrix} \begin{bmatrix} -\mathbf{\beta}_1^* \\ -\mathbf{\beta}_2^* \\ \vdots \\ -\mathbf{\beta}_M^* \end{bmatrix} \]

\[ = \begin{bmatrix} \mathbf{\Sigma}_{\sigma_1^2} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_{\sigma_2^2} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Sigma}_{\sigma_{M}^2} \end{bmatrix} \begin{bmatrix} -\mathbf{\beta}_1^* \\ -\mathbf{\beta}_2^* \\ \vdots \\ -\mathbf{\beta}_M^* \end{bmatrix} \]

\[ \text{matrix} \]
If we have a process defined as
\[
\Phi_w = U \begin{bmatrix}
\omega_1 \\
\omega_2 \\
\vdots \\
\omega_N
\end{bmatrix}
\]
where $U$ is a unitary $N \times N$ matrix (with orthonormal columns, i.e. $U = [\bar{u}_1 \bar{u}_2 \ldots \bar{u}_N]$ and $\bar{u}_k \bar{u}_k = \delta_{kk}$) we get for $\omega_k$'s that are independent, zero mean random variables with variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2$ that
\[
E\{\Phi_w\} = 0
\]
\[
E\{\Phi_w \Phi_w^*\} = U \begin{bmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \sigma_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma_N^2
\end{bmatrix} U^* = R_p
\]
Therefore $U \begin{bmatrix}
\text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)
\end{bmatrix} U^*$ is the eigenvector eigendecomposition of the matrix $R_p$. 
10.5. The statistics of a random process \( \{ \dot{w}(t) \}_{w \in \mathbb{R}} \) or \( \{ \ddot{w} \}_{w \in \mathbb{R}} \) have some interesting properties.

We shall always consider centered processes, so that the ensemble average signal will be assumed to be 0. Hence we concentrate on the autocorrelation, or the so-called second-order statistics of the process.

In the continuous case we have

\[
R(t_1, t_2) = \mathbb{E}_w \{ \dot{w}(t_1) \dot{w}(t_2) \} = \mathbb{E}_w \{ \dot{w}(t_2) \dot{w}(t_1) \}
\]

\[
= R(t_2, t_1),
\]

hence the function \( R(t_1, t_2) \) is symmetric in its arguments. We also have that, if \( \{ \dot{w}(t) \}_{w \in \mathbb{R}} \) is a process that take values in \( \mathbb{R} \), that \( R(t_1, t_2) \in \mathbb{R} \) and clearly \( R(t, t) \geq 0 \) for all \( t \). We have furthermore the following property:
Consider $N$ points over $[0,1)$, i.e.

$$0 \leq t_1, t_2, \ldots, t_N < 1$$

and $N$ real numbers $x_1, x_2, \ldots, x_N$.

Then consider the linear combination of $\varphi_w(t)$-values weighted by the $x$-values as follows:

$$\sum_{k=1}^{N} x_k \varphi_w(t_k) = m(\bar{x}, \bar{t})$$

We have clearly $E_w m = 0$ and

$$E_w (m^2) = E_w \left( \sum_{k=1}^{N} x_k \varphi_w(t_k) \sum_{l=1}^{N} x_l \varphi_w(t_l) \right) =$$

$$= \sum_{k,l} x_k x_l R(t_k, t_l) > 0$$

for all vectors $\bar{x} = [x_1, x_2, \ldots, x_N]$. This means that the function of two variables $R(t, \bar{t})$ is what is called a "positive definite kernel" on $[0,1]^N$. 

(Note: If we have that \[ \int_0^1 R(t, \tau) \, dt \, d\tau < \infty \]
then the theory of integral equations and Hilbert-Schmidt operators due to Fredholm states the following:

The operator:

\[ \mathcal{K}(f(t)) = \int_0^1 R(t, \tau) f(\tau) \, d\tau \]

that maps functions square integrable on \([0, 1]\) into functions square integrable on \([0, 1]\), is positive, self-adjoint and compact and has a countable set of positive eigenvalues \(\{\lambda_k\}_{k=1}^{\infty}\) associated to a set of orthonormal eigenfunctions \(\{\beta_k(t)\}_{k=1}^{\infty}\), i.e., we have for

\[ \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq \ldots > 0 \]

\[ \sum_{k=1}^{\infty} \lambda_k^2 < \infty \quad \text{and} \quad \lambda_k \beta_k(t) = \int_0^1 R(t, \tau) \beta_k(\tau) \, d\tau \]
and the eigenfunctions $\{\beta_k^2(t)\}$ form a basis for all square integrable functions over $[0,1]$, i.e. every $f(t)$ can be written as
$$f(t) = \sum_{k=1}^{\infty} \langle f(t), \beta_k^2(t) \rangle \beta_k^2(t)$$
the equality being in the sense that
$$\lim_{N \to \infty} \| f(t) - \sum_{k=1}^{N} \langle f(t), \beta_k^2(t) \rangle \beta_k^2(t) \| \to 0.$$ Then by Mercer's Theorem we have that over $[0,1]^2$
$$R(t,\tau) = \sum_{k=1}^{\infty} \lambda_k \beta_k^2(t) \beta_k^2(\tau)$$
(where the series on the left converges "nicely" to $R(t,\tau)$)
We shall next see these properties for the discrete case!
In the discrete case we have a random process of signal vectors $\mathbf{\bar{e}}_w$ over $\mathbb{R}$ with $E_w(\mathbf{\bar{e}}_w) = 0$

$$E_w(\mathbf{\bar{e}}_w \mathbf{\bar{e}}_w^*) = \mathbf{R}_\phi = E_w(\mathbf{\bar{e}}_w \mathbf{\bar{e}}_w)^* = \mathbf{R}_\phi^*$$

The matrix autocorrelation is symmetric and also nonnegative definite. Clearly for any vector $\mathbf{x}$ we have

$$\mathbf{x}^* \mathbf{R}_\phi \mathbf{x} = \mathbf{x}^* E_w(\mathbf{\bar{e}}_w \mathbf{\bar{e}}_w^*) \mathbf{x} =$$

$$= E_w(\mathbf{x}^* \mathbf{\bar{e}}_w \mathbf{\bar{e}}_w^* \mathbf{x}) = E_w(\mathbf{\bar{e}}_w \mathbf{x}^2) \geq 0.$$  

From simple linear algebra we know that any symmetric, nonnegative definite matrix has a full set of nonnegative eigenvalues associated to ortho-normal eigenvectors yielding the following
"Spectral" decomposition of $R_6$

$$R_6 = U \Lambda U^* =$$

$$= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \cdots \\ U_N \end{bmatrix}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$ and

$\bar{U}_k^* \bar{U}_k = \langle \bar{U}_k, \bar{U}_k \rangle = \delta_{kk}$

(We see this from the fact that for $k=1, 2, \ldots, N$

$\exists k > 0, \Rightarrow R_6 \bar{U}_k = \lambda_k \bar{U}_k$ and $\langle \bar{U}_k, \bar{U}_k \rangle = 1$

and if $\bar{U}_k$ and $\bar{U}_e$ correspond to different $\lambda$'s

then $\bar{U}_k^* R_6 \bar{U}_e = \lambda_e \langle \bar{U}_k, \bar{U}_e \rangle =$

$= \bar{U}_e^* R_6 \bar{U}_k = \lambda_k \langle \bar{U}_k, \bar{U}_e \rangle$

$$\Rightarrow (\lambda_k - \lambda_e) \langle \bar{U}_k, \bar{U}_e \rangle = 0 \Rightarrow \langle \bar{U}_k, \bar{U}_e \rangle = 0.$$

If we have eigenvalues with multiplicities then
we can find a orthogonal basis for the invariant subspaces characterized by $R_6 \bar{U} = \lambda \bar{U}$.
10.6. Looking at a random process in different bases for \( \mathbb{R}^N \). Suppose we have a set of \( N \) orthonormal vectors in \( \mathbb{R}^N / \mathbb{C}^N \):

\[
\{ \bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_N \}
\]

forming a unitary matrix \( \beta = [\bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_N] \).

Suppose we "change basis" in considering the "realizations" of \( \{\bar{\xi}_w\} \), i.e., for each \( \bar{\xi}_w \) in the ensemble of possible realizations we do

\[
\bar{\xi}_w = I \bar{\xi}_w = \beta^* \bar{\xi}_w = \beta^* \cdot \begin{bmatrix}
<\bar{\beta}_1, \bar{\xi}_w> \\
<\bar{\beta}_2, \bar{\xi}_w> \\
<\bar{\beta}_N, \bar{\xi}_w>
\end{bmatrix}
\]

The vectors \( \beta^* \bar{\xi}_w \), i.e., the representations of \( \bar{\xi}_w \) in the \( \beta \)-basis are a different random process: \( \{\bar{\xi}_w\} \) as seen in the "rotated" basis \( \beta^* \):

\[
\bar{\xi}_w \triangleq \beta^* \bar{\xi}_w
\]
The process \( \{ \tilde{E}_w \} \) was found to have (second-order) statistics given by:

\[
E_w \{ \tilde{E}_w \} = E_w \{ \beta \tilde{E}_w \} = \beta^* E \{ \tilde{E}_w \} = 0.
\]

\[
E_w \{ \beta \tilde{E}_w \} = E_w \{ \tilde{E}_w \} = \beta^* E \{ \beta \tilde{E}_w \} = \beta^* E \{ \beta \tilde{E}_w \} = \beta^* \beta = \beta^* R \beta
\]

We conclude that:

\[
R_{\beta} \beta = \beta^* R \beta = \beta^* \beta = \beta^* \mathbf{1} \beta
\]

i.e. the process in the new basis has the same eigenvalues as the original one and the eigenvectors are simply rotated by \( \beta^* \).
An important example:

Suppose we have a vector (signal) process defined as follows:

\[ \{ \mathbf{M}_n \} \text{ vectors whose entries are} \]

\[ \text{i.i.d. random variables with zero mean and variance } \sigma_n^2. \]

Then, as we saw before we have:

\[ \mathbb{E} \{ \mathbf{M}_n \} = \mathbf{0} \]

\[ \mathbb{E} \{ \mathbf{M}_n \mathbf{M}_n^T \} = \sigma_n^2 \mathbf{I} \]

In this case we see that in any basis \( \beta \), the realizations of \( \mathbf{M}_n \) have the statistics:

\[ \mathbb{E} \{ \beta^T \mathbf{M}_n \} = \mathbf{0} \]

\[ \mathbb{E} \{ \beta^T \mathbf{M}_n \mathbf{M}_n^T \beta \} = \mathbf{R} \sigma_n^2 \mathbf{I} \quad \mathbf{R} = \sigma_n^2 \mathbf{I} \]

i.e. the process "looks" the same in all the o.n. bases of \( \mathbb{R}^n \), in terms of the second order statistics.
Next let us ask how a random process is changed/distorted by a linear operator $\mathcal{K}$. Suppose we consider a random process $\{\xi(t)\}_{t \in \mathbb{R}}$ or $\{\xi_w(t)\}_{w \in \mathbb{R}}$ and we apply to its realization a linear operator $\mathcal{K}$. Then we generate a new random process $\{\phi_{\xi_w}(t)\}_{t \in \mathbb{R}}$ or $\{\phi_{w}(t)\}_{t \in \mathbb{R}}$.

We know that:

$$\phi_{\xi_w}(t) = \mathcal{K} \{\xi(t)^2\}$$

or $\phi_{w}(t) = \mathcal{K} \{\xi(t)^2\}$

$\phi_{\xi_w}(t) = \int_{-\infty}^{+\infty} \xi_w(s) h(t,s) \, ds$ (linear)

or $\phi_{w}(t) = \int_{-\infty}^{+\infty} \xi(s) h(t-s) \, ds$ (linear + shift invariant)
Therefore we have

\[ E_{w_1} \mathbb{E}_{w_2} (t) = \int_0^\infty E_{\mathbb{E}_{w_2} (t)} h_2(t, s) \, ds = 0. \]

\[ E_{w_1} \mathbb{E}_{w_2} (t) \mathbb{E}_{w_2} (t) = \int_0^\infty \int_0^\infty E_{\mathbb{E}_{w_2} (t)} E_{\mathbb{E}_{w_2} (t)} h_2(t, s) h_2(t, t) \, ds \, dy. \]

\[ \mathbb{R}_{\text{cont}} (t, t_2) = \int_0^\infty \int_0^\infty R_{\mathbb{E}_{w_2} (t)} h_1(t, s) h_2(t_2, t) \, ds \, dy. \]

Suppose \( h_2(t, s) \to h_{L_1} (t-s) \) and also \( R_{\mathbb{E}_{w_2} (t)} (t_1, t_2) = R_{\mathbb{E}_{w_2} (t)} (t_1 - t_2) \) (i.e. the process is what is commonly called "wide-sense stationary".

Then:

\[ \mathbb{R}_{\text{cont}} (t, t_2) = \int_0^\infty \int_0^\infty R_{\mathbb{E}_{w_2} (t)} h_{L_1} (t-s) h_{L_1} (t_2, t) \, ds \, dy. \]

\[ = \int_0^\infty \int_0^\infty R_{\mathbb{E}_{w_2} (t)} h_{L_1} (t_2, t) \, ds \, dy. \]
\[ R_{\text{out}}(t, t_1) = \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{+\infty} R_{\text{in}}(\xi') h_{\text{LSI}}(t_1 - \xi - \xi') d\xi' \right] h_{\text{LSI}}(t_2 - \eta) d\eta \]

\[ = \int_{-\infty}^{+\infty} \left[ R_{\text{in}}(t) \otimes h_{\text{LSI}}(t) \right]_{t_1 - \eta} \left. h_{\text{LSI}}(t_2 - \eta) \right. d\eta \]

\[ = \int_{-\infty}^{+\infty} F(t_1 - \eta) h_{\text{LSI}}(t_2 - \eta) d\eta = \int_{-\infty}^{+\infty} F(\eta') h_{\text{LSI}}(t_2 - t_1 + \eta') d\eta' \]

\[ = \int_{-\infty}^{+\infty} F(\eta') h_{\text{LSI}}(t_2 - t_1 + \eta') d\eta' \]

\[ = R_{\text{in}}(t) \otimes h_{\text{LSI}}(t) \otimes h_{\text{LSI}}(-t) \bigg|_{t = t_2 - t_1} \]

\[ = R_{\text{in}}(t) \otimes \left[ h_{\text{LSI}}(t) \otimes h_{\text{LSI}}(-t) \right] \text{ at } t = t_2 - t_1 \]

Hence

\[ R_{\text{out}}(t_1, t_1) = R_{\text{out}}(t_2, t_1) \]
and
\[ R_{\text{ent}}(\tau) = R_e(\tau) \ast [h_{LSI}(\tau) \ast h_{LSI}(\tau)] \]
or in the Fourier Transform Domain
\[
\mathcal{F}\{R_{\text{ent}}(\tau)\} = \mathcal{F}\{R_e(\tau)\} \cdot H(f) \cdot \overline{H(f)} = \\
= \mathcal{F}\{R_e(\tau)^2\} \|H(f)\|^2
\]

The \( \mathcal{F}\{R(\tau)^2\} \) is called the "power spectrum" of a stationary process and is denoted by \( S_e(f) \). So in this language we proved that
\[ S_{\text{ent}}(f) = S_e^2(f) \cdot \|H(f)\|^2 \]

As before, everything is simpler for the more important case of disunited vector signals!
If a vector process \( \{ \tilde{e}_w \} \) were is "distorted" by a linear operator \( \mathcal{A} \) we have "modified" that
\[
\tilde{e}_{\text{out}} = \mathcal{A} \tilde{e}_w
\]
where \( \mathcal{A} \) is a deterministic matrix.

Then clearly
\[
E_t(\tilde{e}_{\text{out}}) = \mathcal{A} E_t(\tilde{e}_w) = \mathcal{A} \tilde{\sigma} = \tilde{\sigma}
\]
and
\[
E_t(\tilde{e}_{\text{out}}^* \tilde{e}_{\text{out}}) = \mathcal{A} E_t(\tilde{e}_w^* \tilde{e}_w) \mathcal{A}^* =
\]
\[
= \mathcal{A} R_{\tilde{\sigma}} \mathcal{A}^* = R_{\text{out}}
\]
Clearly if \( \mathcal{A} \) is a circulant matrix, i.e \( \mathcal{A} \) is linear and shift-invariant we obtain that in case \( R_{\tilde{\sigma}} \) is circulant, that \( R_{\text{out}} \) will be circulant too. When do we have \( R_{\text{out}} \)-circulant?
If \( \{\bar{E}_n\} \) is a stationary-periodic process we have
\[
\overline{\bar{E}}_0 = \begin{bmatrix} \bar{E}_0 \\ \bar{E}_1 \\ \vdots \\ \bar{E}_{N-1} \end{bmatrix}
\]
and \( \overline{\bar{E}} E_k = 0 \)
with
\[
\overline{\bar{E}} E_k \bar{E}_k = r((l-k) \mod N)
\]
i.e. \( \overline{\bar{E}} \bar{E}_0 \bar{E}_j = r(0), \overline{\bar{E}} \bar{E}_i \bar{E}_j = r((-1) \mod N) = r(N-1) \)

But obviously we also have (since we assume \( \bar{E} \) to be real valued)
\[
\overline{\bar{E}} \bar{E}_i \bar{E}_j = E \{ \bar{E}_i \bar{E}_j \} = r(1)
\]

Hence for such a process we have
\[
r(N-1) = r(1)
\]
\[
r(N-2) = r(2)
\]
\[
r(N-k) = r(k) \quad \text{for } k = 1, 2, \ldots, (N-1)
\]

Therefore the covariance matrix of \( \{\bar{E}_n\} \) looks as follows:

Hence the sequence \( r(1), r(2), \ldots, r(N-1) \) is palindromic!
\[ R_\phi = E_w \{ E_w R_\phi^* E_w \} = \]

\[
\begin{bmatrix}
V(0) & V(1) & \cdots & V(N-1) \\
V(1) & V(0) & \cdots & V(N-2) \\
\vdots & \vdots & \ddots & \vdots \\
V(N-1) & V(N-2) & \cdots & V(0)
\end{bmatrix} = \]

\[
\begin{bmatrix}
V(0) & V(1) & V(N-1) \\
V(1) & V(0) & V(N-2) \\
V(N-2) & V(N-1) & V(0) \\
V(0) & V(2) & V(N-3) & V(0)
\end{bmatrix}
\]

\[ \text{which is symmetric.} \]

\text{In this case we have clearly that}

\[ E \{ E_w R_\phi^* E_w \} = H I_{L_1} R_\phi H I_{L_1}^* = R_\text{out} \]

\text{is a product of Toeplitz circulant matrices and therefore it is also circulant.}

\text{If } R_\phi \text{ is only symmetric and Toeplitz, } R_\text{out} \text{ will not be circulant, in general.} \]