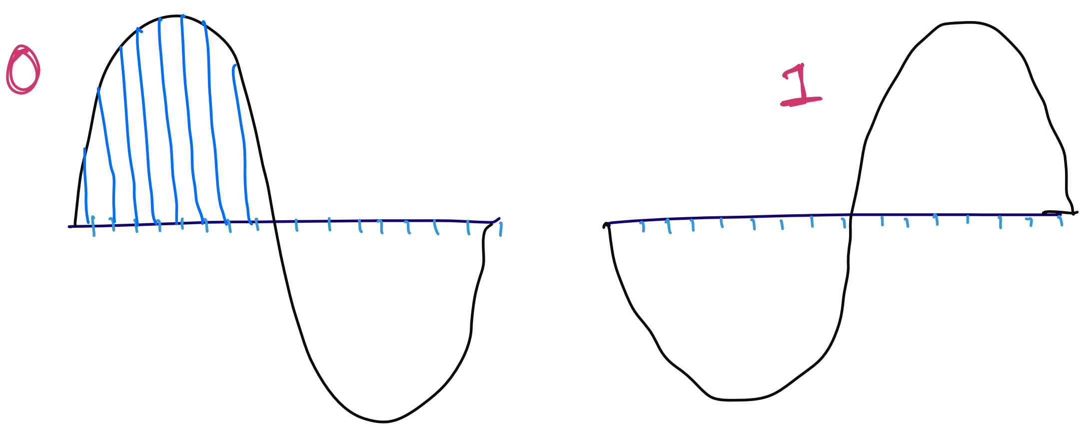
03/12/23 | **Lecture-5** | 24/03/21

→ Continuation of discussion of Stochastic processes

We saw the description of stochastic process as been characterised by the random variables at different time stamps. But it is unrealistic for $n \to \infty$

ex: BPSK signal (Binary phase shift key)



This signal consists of a sequence of waveforms and choose one waveform to transmit a One and another to transmit Zero

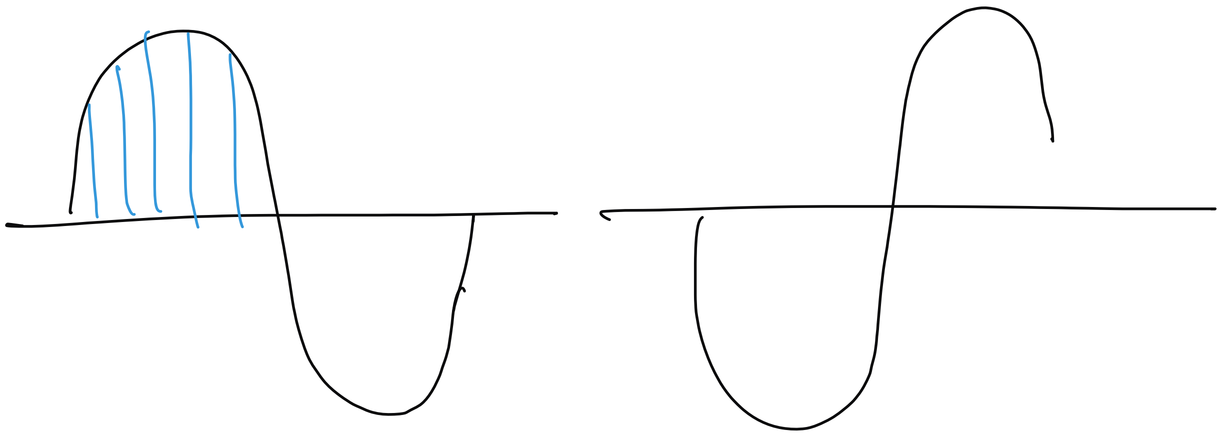
usually in BPSK there are two stretches of sinusoids that can easily be distinguished.

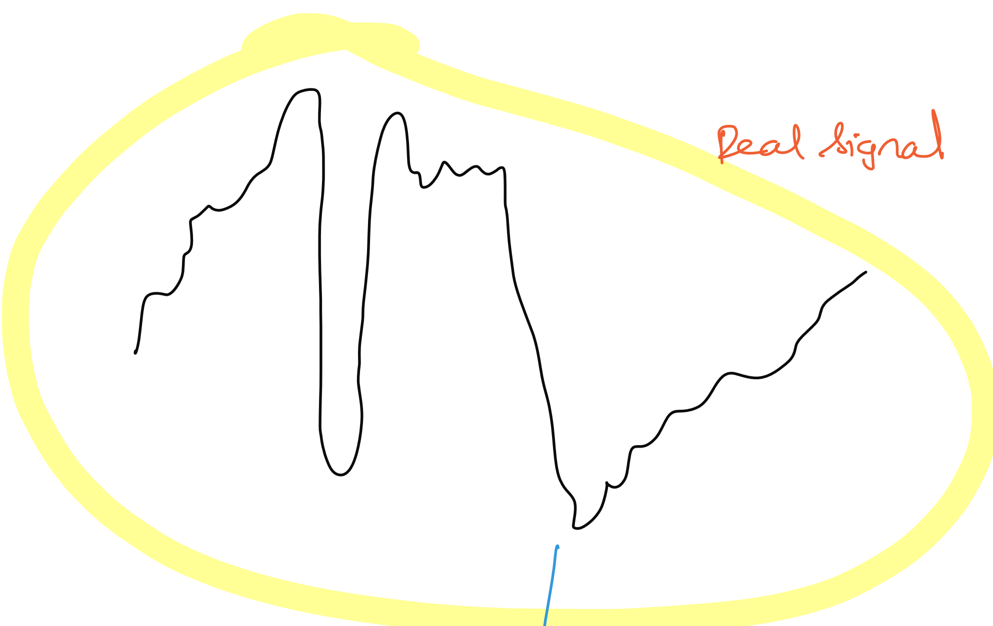
$\therefore$ Our signal is usually a sequence of Time Windows each of it is a bit time and each time window our profile cone Up-down (or) Down-Up

→ The true signal that is present on the channel.

If we are transmitting a Radio Signal using this tech-nique'

It will be very different because superImposed on this there will be Noise.





This can be a Real zero for BPSK, it seems very corrupted but more or less the UP Down trend is followed.

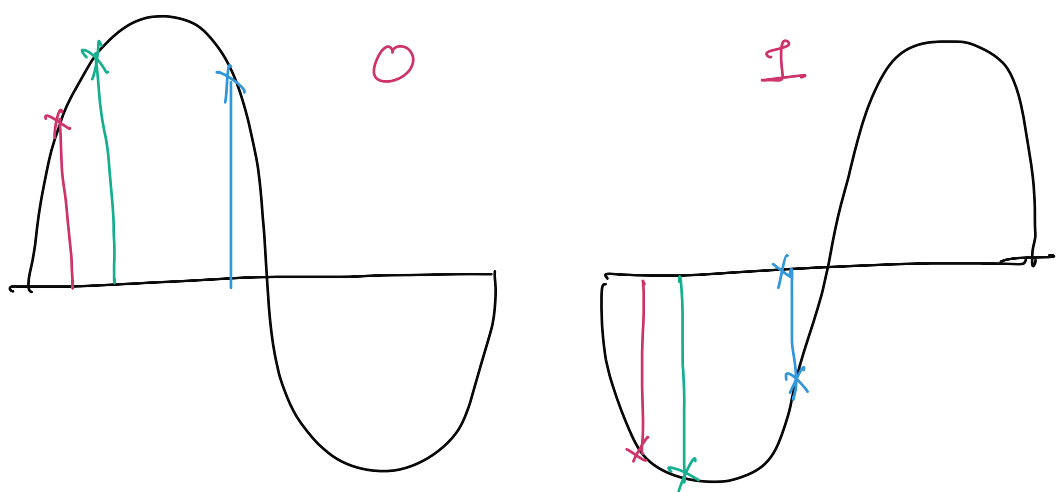
But even the UP trend has a sharp Null in time.

→ So everything we have seen above in the realtime signal is a stochastic process, we never know what kind of Noises & at what Amplitude & phase would corrupt the original signal.

→ Everytime we transmit some information that we don't know from the beginning ∴ It is a stochastic process for even for the information it carries

→ let's focus on the second kind of Randomicity each one attached to a certain time stamp

ex:



- +ve value when we are transmitting 0
- -ve value when we are Transmitting 1

∵ we don't know from the start what waveform we are transmitting so from the beginning there is a

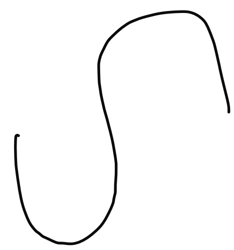
probability distribution associating to that time instant.
A RV may have as values +ve & −ve distribution

The same happens when we choose Green or Blue time instants

and RV at each time instant will be characterized by a different PDF.
This is offcourse not enough and PDFs of different RVs will have some link.

*

probability = $\frac{1}{2}$



probability = $\frac{1}{2}$

→ We have to make all the different PDFs of the RVs to agree to describe the common profile of the transmitted signal

* i.e The PDF in all colored time instances agree with each other.

• → $f_{X(t')}$

• → $f_{X(t'')}$

$\rightarrow f_{X(t''')}$

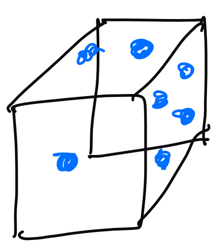
$\therefore$ once we know something about $f_{X(t')}$ then we can say something about the $f_{X(t'')}$

→ This means that we have bunch of PDF^s but I have a lot of link b/w them to describe the waveform.
$\because$ This waveform has been designed beforehand, so it is an artificial one, if we design it in a more complicated way the no. of links b/w PDF^s will grow and it will be
* even more complicated.
* This makes the characterisation with JPDF very inefficient.

Q) What is the Alternative point of view?

Ans Having completely different Approach.

→ consider the example of Rolling a Die



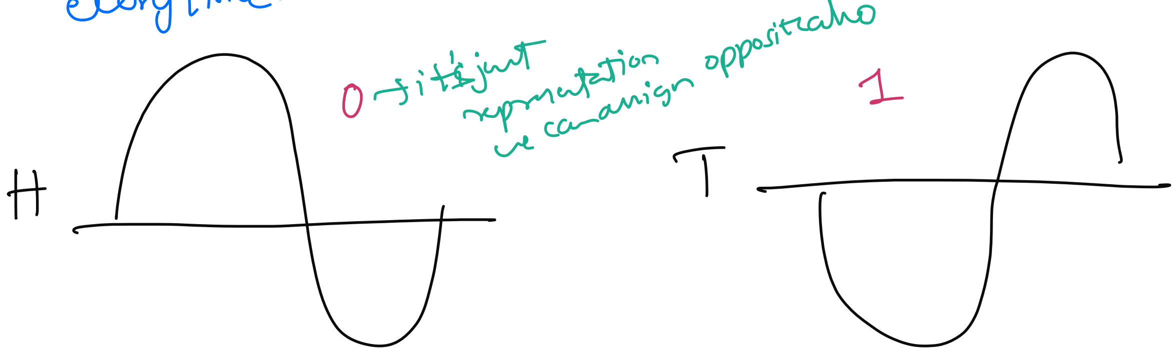
theoretical Mapping →

Here $2 \in \mathbb{R}$

The number '2' has a probability of $1/6$ if the Die is a Fair one.

* → * When you can associate probabilities to Numbers why can't we associate probabilities to all wave forms.

It would be simple in our BPSK case, because you can take a coin (fair one) and Toss it and we have H/T everytime.



→ Here we are assigning probabilities to all waveforms.

* This New point of view will allow us to interpret quite well a lot of signals expressed with all the waveforms

BPSK is equivalent to Coin Tossing

→ Therefore, our RV will not just give us Random values but It will be whole waveforms in a Time window with different probabilities.

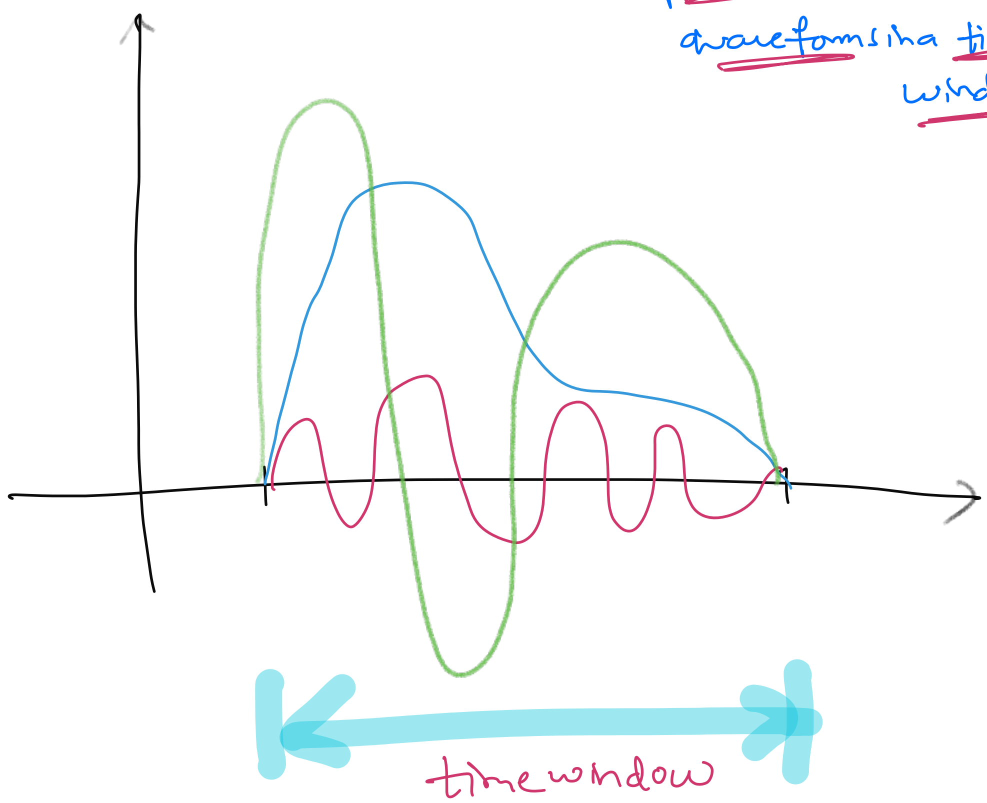
Waveforms in a time window are a Vector space!

$$\boxed{x(t) = \sum_J \alpha_J \phi_J(t)}$$

fixed-non Random Waveforms

scalars

• We have to attach probability to all the waveforms in a time window.



→ Q) Do we have a Basis for the Waveform vector space?

Ans The answer to the question is No

Because, the Dimension of this vector space is Not Finite.

But in General, we can always take waveforms that we are interested, as just a pair of sinusoids i.e as in BPSK case sinusoid Up-down & Down-up can still be produced by the Linear Combination of some fundamental waveform that I assume to be fixed.

So, Our Model would be set of waveforms that are fixed.

$$\boxed{X(t) = \sum_J \alpha_J \phi_J(t)}$$

fixed waveforms

Non-Random

Scalars

Random

Random

→ This is a good way of attaching Random phenomenon to a waveform.

→ If we change α_J we are not only changing the value of F.W. at one place but we change the value at all the are spanned by ϕ_J

ex: 256 - QAM modulation

↓

$$\alpha_0 \cos(2\pi f t) + \alpha_1 \sin(2\pi f t)$$

QAM is able to transmit in a single waveform upto 256 different symbols.

It does so carefully by choosing this 256 waveforms to which it attaches a probability of course, But these 256 waveforms one for every combination of 8 bits are defined in terms of cosine & sine on the same Bandwidth.

∴ Given a bit frequency and Bit time that freq The frequency of the sinusoid, what we are focusing is bit time 't' and choosing

Coefficients α_0, α_1. So that the Linear Combination of two waveforms gives me the final waveform that I associate to each of the possible Combination of the 8 bits.

→ So, we have to choose 256

pairs of (α_0, α_1) to encode 8 bits

each of the waveforms will have a probability attached to them by means of RV's α_0, α_1 because ϕ_0, ϕ_1 are fixed.

→ If we accept this point of view all we have to do is to characterize separately the

Random values of α and function ϕ_J

We should try and understand, How this ϕ_J should be made to be useful in the subsequent calculations.

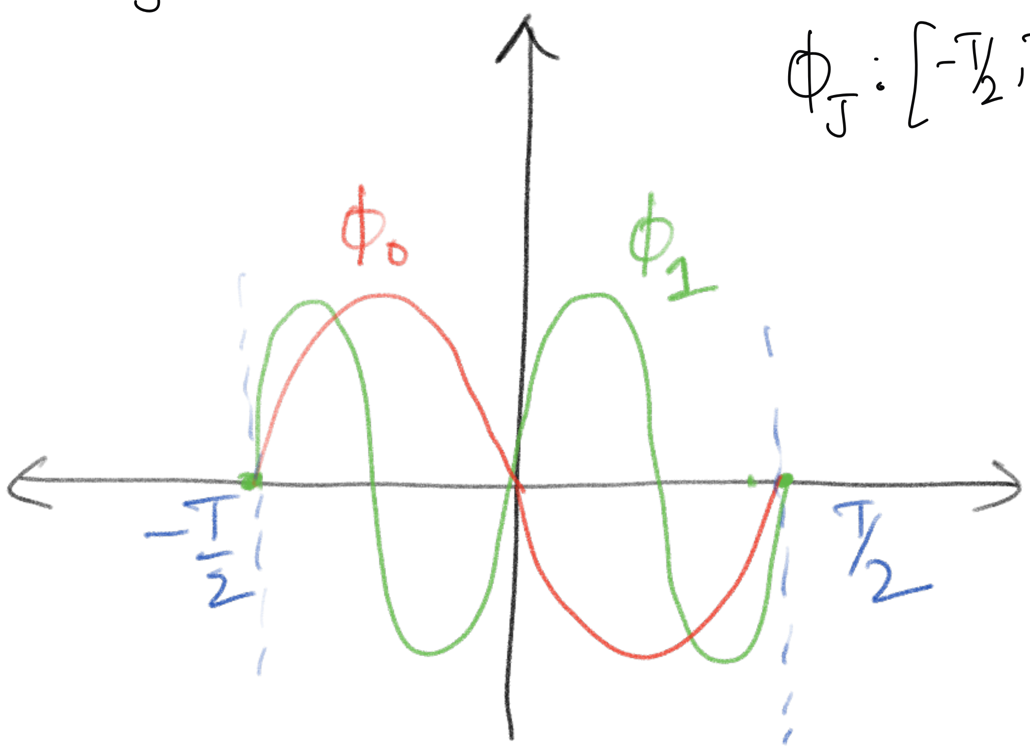
Properties of ϕ_J:

- We would like ϕ_J to be

Orthogonal & possibly Orthonormal

ϕ_J are waveforms on a certain Time interval

$\phi_J : [-T/2, T/2] \to \mathbb{R}$



$\to$ $\alpha(t), \beta(t)$

$$\boxed{\langle \alpha, \beta \rangle = \int_{-T/2}^{T/2} \alpha^*(t)\, \beta(t)\, dt}$$

Scalar product of α, β

Scalar product of Two functions defined over an Interval.

$*$ $\to$ Conjugation is need for the possibility of α, β being Complex values

→ Now that we have scalar product defined b/w the waveforms in time interval $[-T/2, T/2]$

→ We can say two waveforms are orthogonal when the scalar product defined b/w them in a time interval is zero

ex: Two sinusoids with different frequencies

→ Beyond orthogonality we would also like them to be orthoNormal. i.e their lengths are also equal to 1 apart from being orthogonal

Def:

$$\langle \phi_J, \phi_K \rangle = \begin{cases} 0 & J \neq K \\ 1 & J = K \end{cases}$$

$$\langle \phi_J, \phi_J \rangle = \|\phi_J\|^2 = 1$$

L2 Norm

→ Beautiful consequences of Ortho Normality.

let a known waveform

$$x(t) = \sum_J \alpha_J \phi_J(t)$$

- We want to know the coefficients of the waveform i.e α_k for example

$$\langle \phi_k , x \rangle = \langle \phi_k , \sum_J \alpha_J \phi_J \rangle$$

Scalar product of a single waveform ϕ_k with the Global waveform x

$$\Rightarrow \int_{-T/2}^{T/2} \phi_k^*(t) \sum_J \alpha_J \phi_J(t)\, dt$$

$$\Rightarrow \sum_J \alpha_J \underbrace{\int_{-T/2}^{T/2} \phi_k^*(t)\, \phi_J(t)\, dt}$$

This Integration has only two values

$$\langle \phi_k , \phi_J \rangle = \begin{cases} 0 & J \neq k \\ 1 & j = k \end{cases}$$

$$\boxed{\langle \phi_k , x \rangle = \alpha_k}$$

Note: If we want to know what is the coefficient of ϕ_k to give a certain waveform $x(t)$

we may simply Project our waveform $x(t)$

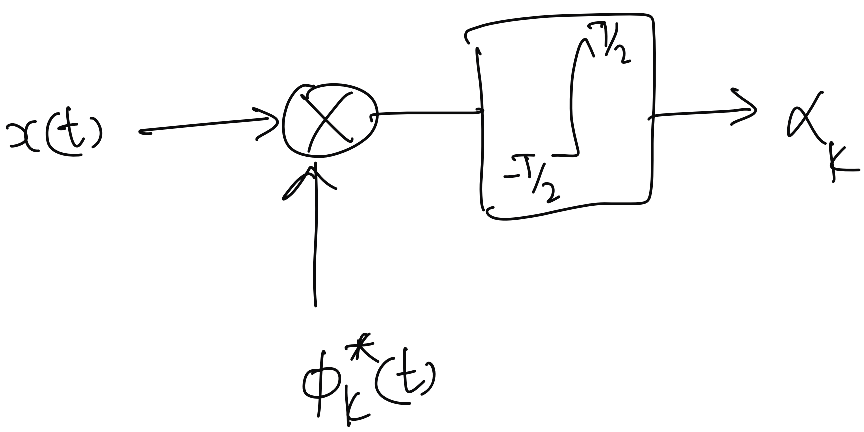
over the fixed waveform $\phi_k(t)$ that we use to express all our signals.

Fixed waveforms to have the above property should be orthonormal to each other

→ The operation we did above can be called

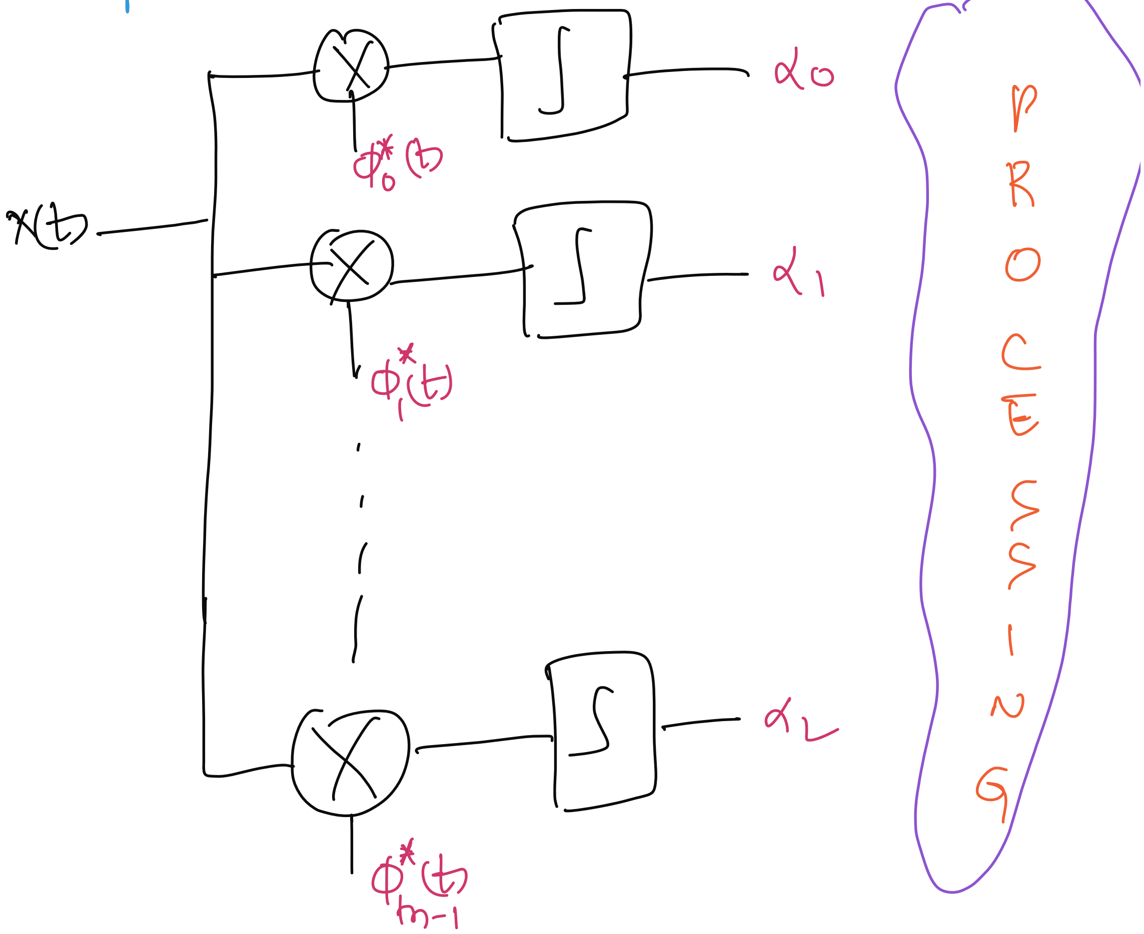
Correlator | projector

Correlator/projector is a block that performs the above integral.



But our waveform $x(t)$ is characterized not only by one standard waveform but many other standard waveform from the Basis.

∴ we will have a lot of projection in general to calculate the value of all the coefficients of the standard waveforms



Q)

→ From this situation on, we may ask ourselves beyond orthonormality is there anything we would like to have for our $\phi_j(t)$?

Which would make easier to Model our process.

The idea is, Correlator / projector have a Cost because we have to deploy specific $\phi_j(t)$ to calculate it's coefficient.

Ideally we would like to employ least amount of projectors. without loosing any useful information.

In principle our Basic Criteria would be okay assuming we have deployed a correlator with respect to $\phi_0(t)$ and you want to decide whether it is sensible to deploy another correlator w.r.to $\phi_1(t)$. one we have α_1 we must have some New piece of information if not so we would be satisfied by having α_0 w.r.to $\phi_0(t)$.

Q) How do we know a New piece of information has come in, when you know α_1 w.r.to α_0 ?

Ans we can resort to the concept of Prediction Because starting from α_0, if we know how to predict α_1 without deploying any correlator w.r.to $\phi_1(t)$. It would be Great

we would deploy a correlator for some specific $\phi_J(t)$ if for sure we know that we can't predict α_J from α_0

In principle we would like to deploy.

correlators that give us ~~Statistically Independent~~

α_J's (coefficients)

** This Regrettably is too complicated to enforce

But we can resort to a limited form of Inde-
-pendence b/w both which is the property of

zero covariance.

→ We have seen earlier that, if we have two Random Variables with Zero Covariance it is not possible to predict one value from another if we limit ourselves to linear predictors.

→ Therefore, our limited criteria would be α_J couldnot be linearly predicted from α_K

i.e $\gamma_{\alpha_J \alpha_K} = 0 \quad \forall \quad J \neq K$

$$\gamma_{\alpha_J \alpha_J} \Rightarrow E\left[(\alpha_J - m_{\alpha_J})^* (\alpha_K - m_{\alpha_K})\right]$$

assume

$\alpha_J, \alpha_K \in \mathbb{R}$

$$\alpha_J \Rightarrow \int_{-T/2}^{T/2} \phi_J(t)\, x(t)\, dt$$

$$m_{\alpha_J} = E[\alpha_J] = E\left[\int \phi_J(t)\, x(t)\, dt\right]$$

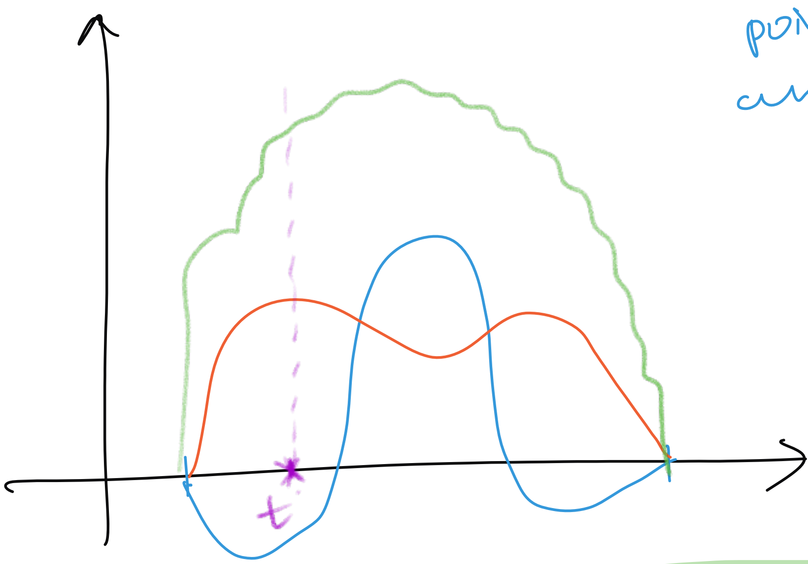
(Average)

$$= \int \underline{\phi_J(t)}\, E\underbrace{[x(t)]}_{m_{x(t)}}\, dt$$

Since they are chosen to be fixed

$$\boxed{\alpha_J = \int \phi_J(t)\, m_{x(t)}\, dt}$$

is the average of all points our x(t) might assume at time instant 't'



Similarly

$$\boxed{\alpha_K = \int \phi_K(t)\, m_{x(t)}\, dt}$$

Substituting α_J, α_K in the expression of covariance.

$$\gamma_{\alpha_J \alpha_K} = E\left[\left(\underbrace{\int \phi_J(t)\, x(t)\, dt}_{\alpha_J} - \underbrace{\int \phi_J(t)\, m_{x(t)}\, dt}_{m_{\alpha_J}}\right)\left(\underbrace{\int \phi_K(t)\, x(t)\, dt}_{\alpha_K} - \underbrace{\int \phi_K(t)\, m_{x(t)}\, dt}_{m_{\alpha_K}}\right)\right]$$

$$= E\left[\int \phi_J(t)\left(x(t) - m_{x(t)}\right) dt \int \phi_K(t)\left(x(t) - m_{x(t)}\right) dt\right]$$

$$\Rightarrow E\left[\iint \phi_J(t)\, \phi_K(s)\left(x(t) - m_{x(t)}\right)\left(x(s) - m_{x(s)}\right) dt\, ds\right]$$

Here we have just exchanged the Notation.

$$\Rightarrow \iint \phi_J(t)\, \phi_K(s)\, E\left[\left(x(t) - m_{x(t)}\right)\left(x(s) - m_{x(s)}\right)\right] dt\, ds$$

$$K_X(t,s)$$

(Second order Covariance function)

$$\gamma_{\alpha_J \alpha_K} \Rightarrow \int \phi_J(t) \underbrace{\left(\int K_X(t,s)\, \phi_K(s)\, ds \right)}_{\psi(t)} dt$$

Here there is a Statistical Characteristic of our process which is the 2nd order Covariance function of our process.

& we would like this Covariance

$$\gamma_{\alpha_J \alpha_K} \begin{cases} = 0 & \text{when } J \neq K \\ = \sigma^2_{\alpha_J} & \text{when } J = K \end{cases}$$

variance

This would make it impossible to predict α_K if we know α_J by linear prediction.

→ The constrain of Zero Covariance is we would like to have when we want to design our $\phi_J(t)$ (set of base functions) i.e we want set of $\phi_J(t)$ should be orthonormal

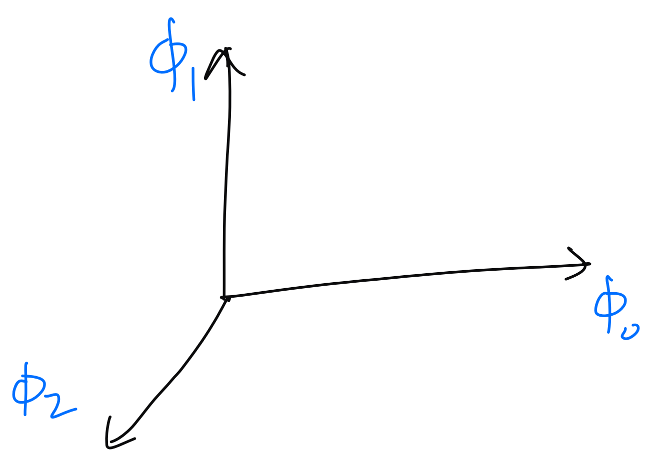
& covariance = 0

example:

$$\gamma_{\alpha_J \alpha_K} = \int \phi_J(t) \underbrace{\left[\int K_X(s,t)\, \phi_K(t)\, ds \right]}_{\psi(t)} dt$$

let's say we have 3 ϕ

$\phi_0 \quad \phi_1 \quad \phi_2 \qquad$ & $K=0$ i.e $\phi_K = \phi_0$



let's look at the expression of covariance above

$$\underbrace{\int \phi_J(t)\,\psi(t)\,dt = 0}_{} $$ This is the expression of Scalar product

$$\Downarrow$$

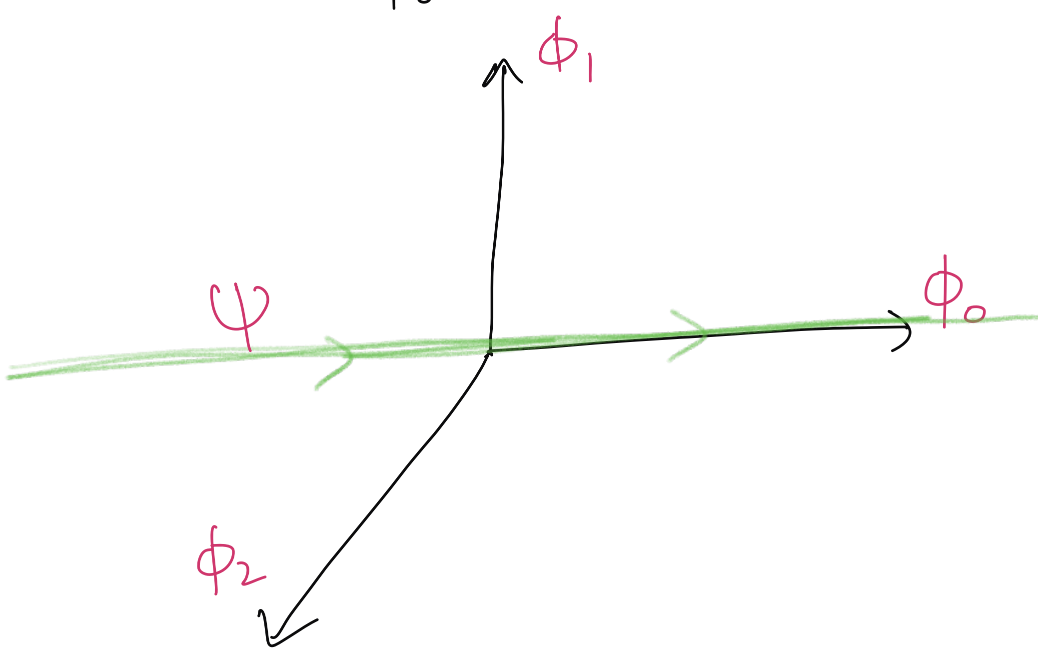
it includes ϕ_0

$$\langle \phi_J, \psi \rangle = 0$$

for $J = 1, 2$

i.e ϕ_0 should be orthogonal to ϕ_1 & ϕ_2

'ψ' can be though of a Scaled version of ϕ_0



$$\Rightarrow \boxed{\underbrace{\int k_x(t,s)\,\phi_k(s)\,ds}_{\psi} = \lambda\,\phi_k(t)}$$

Eigenvalue

Eigenfunction

ψ is a Multiple of ϕ is the language of vectors and is equal to the Multiple of

EIGEN EQUATION

Eigenvalue

$$\boxed{\underbrace{\int k_x(t,s)\,\phi_k(s)\,ds} = \lambda\,\phi_k(t)}$$

Eigenfunction

What we see here is a generalization of the application of Matrix to a vector

i.e $\boxed{AV = \lambda V}$ $\Rightarrow$ Eigen Equations

$\rightarrow$ We can say $\phi_k(t)$ is an Eigenfunction corresponding to λ is Eigen value

→ Therefore for our Design: Requirement

$$\boxed{\text{The projections } \alpha_J \text{ & } \alpha_K \text{ are not predictable by linear means by one and other}}$$

** We need to have the functions we are choosing ϕ_J & ϕ_K are also not only ORTHONORMAL but also be the EIGEN FUNCTIONS of the Covariance function $K_X(t,s)$ of the process.

→ The problem is There is no guarantee That Eigen values & Eigen vectors exist for all K_X (covariance function)

Note: The Best ϕ_J^{s} are the Eigen Functions of the Covariance function of the process. (i.e the Random process)

**

→ What we are gonna do is to look at a Fundamental theorem from Functional Analysis

But firstly let's look at the Covariance equation for $J=K$ which is is equal to its Variance.

$$\boxed{\gamma_{JJ} = \int \phi_J \underbrace{\int K_X(t,s)\, \phi_J(s)\, ds}\, dt = \sigma_J^2}$$

$$\Rightarrow \int \phi_J t \left[\int \underset{\text{scalar}}{\lambda}\, \phi_J(t)\right] dt = \sigma_J^2$$

$$\Rightarrow \lambda \underbrace{\int \phi_J(t)\, \phi_J(t)\, dt}_{\text{scalar product of } \phi_J \text{ with itself}} = \sigma_J^2$$

$$\Rightarrow \boxed{\lambda \cdot 1 = \sigma_J^2}$$

- The Eigen Function corresponding to ϕ_J is nothing but it Variance of σ_J

$$\boxed{\underline{\lambda_J = \sigma_J^2}}$$

∴ once we put set of Eigenfunction of the covariance function then we will have the (1) orthonormality
(2) Zero covariance b/w the projections
(3) Eigenvalues equal to the variance of the α_J

→ But still we have to address existence of Eigenfunctions
(Q) Do they exist? if so under what conditions?

$$\boxed{K_X(t,s) = E\left[\left(x(t) - m_{x(t)}\right)^* \left(x(s) - m_{x(s)}\right)\right]}$$

our $K_X(t,s)$ is simply Hermetian

i.e $K_X(t,s) = K_X^*(s,t)$

→ If we go back to our Linear Algebra lemmas when a Matrix is Hermetian in Real domain it meant Symmetric. The Eigen values & Eigenvectors are there

→ To Check about the existence of Eigenfunctions in functional space, we have to resort to a theorem that is analogous to Spectral Decomposition Theorem.

Earlier during Spectral Decomposition Theorem we assume finite dimension.

* In this case it's weaker due to Theorem

(i) Additional Assumption instead of K_X being Hermetian

$$\int\int |K_X(t,s)|^2 \, dt \, ds \leq +\infty$$

finite Energy Covariance

Then for

$$\exists \, \phi_J \Rightarrow \boxed{\int K_X(t,s)\, \phi_J(s)\, ds = \lambda_J \phi_J(t)}$$

(for some eigen functions)

But if

$\exists \, \psi$ such that $\forall_J$

$$\underbrace{\int \phi_J^*(t)\, \psi(t)\, dt}_{\text{Scalar product}}$$

$\psi(t)$ is orthogonal to all the ϕ_J

i.e if the scalar product is zero then

$$\Rightarrow \quad \psi \in (\ker)\, K_X$$

i.e ψ transform K_X to zero

$$\Rightarrow \int K_X(t,s)\,\varphi(s)\,ds = 0$$

∴ We can say some Eigenfunction might exist with Assumption-I

(ii) Theorem about Eigenfunction of $K_X(t,s)$ (stronger)

if $K_X(t,s) = K_X^*(s,t)$

$\Downarrow$ and

$$\iint |K_X(t,s)|^2\,ds\,dt < +\infty$$

and

$K_X(s,t)$ is Strictly Positive Definite

Then ϕ_J and allows us to express any process $X(t)$ s.t $\int |X(t)|^2 dt < +\infty$

** If $K_X(s,t)$ is Strictly Positive Definite

Then Quadratic form

$$\forall g \neq 0 \quad \boxed{\iint K_X(t,s)\, g^*(t)\, g(s)\, dt\, ds > 0}$$

In fact if $\exists \psi$ s.t $\psi \perp \phi_J$

Then by the 1st condition

$$\Downarrow$$

$$\int K_X(t,s)\,\psi(s)\, ds = 0$$

$$\Downarrow$$

$$\Rightarrow \int \psi^*(t) \int K_X(t,s)\,\psi(s)\, ds\, dt = 0$$

we built a Quadratic form above by multiplying $\int \psi^*(t)\, dt$ to

$$\int K_X(t,s)\,\psi(s)\, ds = 0$$

$$\Rightarrow \iint K_X(t,s)\,\psi^*(t)\,\psi(s)\, ds\, dt = 0$$

$\Rightarrow$ But since $K_x(t,s)$ is positive definite

$$\varphi(t) = 0$$

because null vector is $\perp$lar to all the vectors

→ we are addressing:

1) every projector we lay down must be useful

2) Minimizing the No. of projectors, include only projectors that provide something new our

→ Even if the Theorem-2 provides everything we desire about ϕ_J

Usually, the no. of ϕ_J's required to express something is Infinite. and this is a Non-sense from an Engineering point of view because we cannot make any Infinite Machine.

Therefore it would be beneficial for us to regard ϕ_J's which give us most of the information & discard others which only provide Insignificant information.

Task: is to identify the MOST IMPORTANT ϕ_J^s

In general:

$$\boxed{x(t) = \sum_{J=0}^{\infty} \alpha_J \phi_J(t)}$$

identify only useful ϕ_J^s

$$\boxed{\begin{gathered} \mathcal{J} \subset \{0, 1, \ldots\} \\ \hat{x}(t) = \sum_{J \in \mathcal{J}} \alpha_J \phi_J(t) \end{gathered}}$$

ex: $\mathcal{J} = \{0, 1, 2, 3\}$

$$\boxed{\hat{x}(t) = \sum_{J=0}^{3} \alpha_J \phi_J(t)}$$

$\hat{x}(t)$ would only be an estimate of $x(t)$

→ To know $\hat{x}(t)$ is good, The best way is do Error of Reconstruction.

$$\varepsilon = E\left[\int \left| x(t) - \hat{x}(t) \right|^2 dt\right]$$

Since, we have waveforms the Error must be computed by an Integral.

$$= E\left[\int \left| \sum_{j=0}^{\infty} \alpha_J \phi_J(t) - \sum_{j \in \mathbb{J}} \alpha_J \phi_J(t) \right|^2 dt\right]$$

$$\varepsilon \Rightarrow E\left[\int \left| \sum_{j \notin \mathbb{J}} \alpha_J \phi_J(t) \right|^2\right]$$

$$\Rightarrow E\left[\int \sum_{j \notin \mathbb{J}} \sum_{k \notin \mathbb{J}} \alpha_J^* \alpha_K \phi_J^*(t) \phi_k(t)\, dt\right]$$

$$\Rightarrow E\left[\sum_{j \notin \mathbb{J}} \sum_{k \notin \mathbb{J}} \alpha_J^* \alpha_K \underbrace{\int \phi_J^*(t) \phi_K(t)\, dt}_{}\right]$$

$J \neq K \rightarrow 0$

$J = K \rightarrow 1$

$$\Rightarrow E\left[\sum_{J \notin \mathbb{J}} \alpha_J^* \alpha_J\right]$$

$$\Rightarrow E\left[\sum_{J \notin \mathbb{J}} (\alpha_J)^2\right]$$

$$\boxed{\varepsilon = \sum_{J \notin \mathbb{J}} E\left[(\alpha_J)^2\right]}$$

exchanging Expectation with Sum

$\Rightarrow$ Usually projections are made zero-mean

i.e $m_{\alpha_J} = 0$

Variance

$$\sigma_{\alpha_J}^2 \Rightarrow E\left[|\alpha_J - m_{\alpha_J}|^2\right] = E\left[|\alpha_J|^2\right]$$

$$\therefore \quad \boxed{\mathcal{E} = \sum_{J \notin \mathbb{J}} \sigma_{\alpha_J}^2}$$

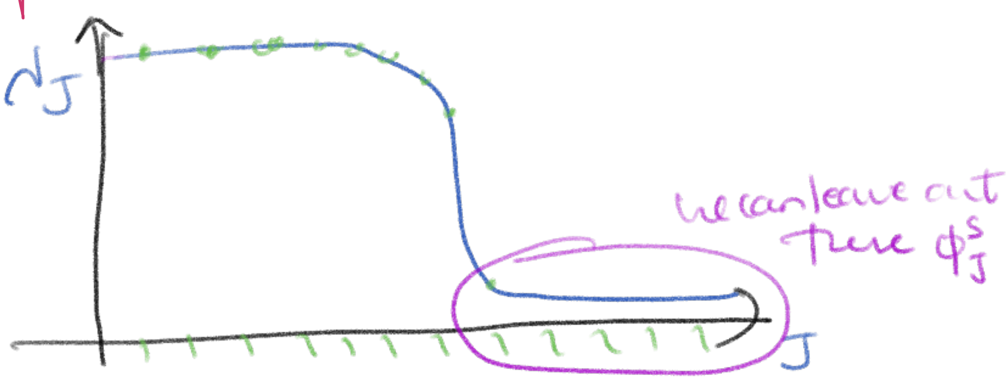
Recall, $\sigma_{\alpha_J}^2$ are nothing but λ_J

$$\Rightarrow \quad \mathcal{E} = \sum_{J \notin \mathbb{J}} \lambda_J$$

ERROR, is nothing but Sum of Eigenvalues of ϕ_J that we are not taking.

Conclusion: Avoid to take ϕ_J's corresponding to small λ_J (Eigenvalues)

Emperical considerations



Assuming we have a strictly positive definite Covariance function.

→ The above consideration leads us to Kahrunen – Löve approximation (or) Compression

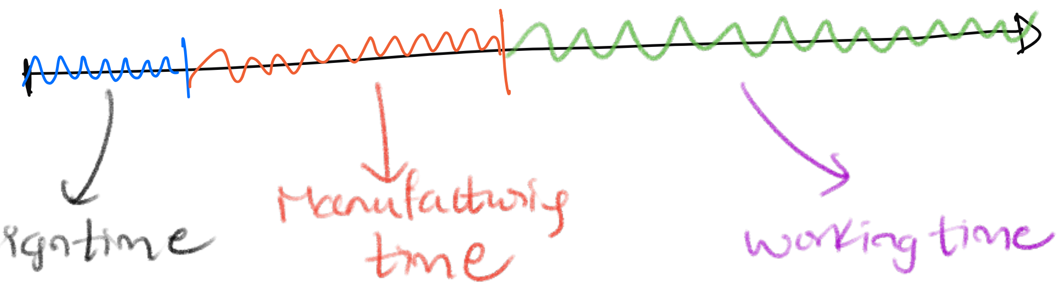
This is finishes part of the course where we are trying to characterize stochastic process

* Stationarity, Ergodicity, Mixingness

All are related concepts, They are tough concepts but the strategy would be is look it from Engineering application point of view and then develop the Math.

The Aim of all Engineering Courses is to build Systems that process information, it may be Analog s/s
Digital s/s
Mp & MC

→ In the Life time of a Device There are atleast 3 phases



→ Design time, is the time which Engineers are mostly concerned.

→ For sure design time should be as small as possible, it is disjoint from the working time.

→ It is not sure that the random singles we use during the Design time of a s/s may not be similar to what they might be subjected to the same in the Real world during working time

Regardless what Random Signals our Designed System is subjected too it should work properly.

If the statistics that our System is subjected to in the '3' different time periods. it's operation should remain good.

ex:

But if it is a Mp and it has been designed for Low power consumption.

* Then it's power consumption would be larger in one of the above '3' regions. which is undesirable.

∴ The Statistical Features (inputs) of the process which our System is designed to address are not constant Then we would be in Trouble.

* What prevents This from happening is an assumption That always we have in our processes which is STATIONARITY.

Stationarity:

It is much easier to define stationarity wrto the Joint PDF characterization of Stochastic process.