# Module 4

# Signal Representation and Baseband Processing

# Lesson
# 16
# Representation of Signals

## After reading this lesson, you will learn about

- ***Representation of signals following the Gram-Schmidt orthogonalization procedure;***
- ***Signal space and signal constellation;***
- ***Use of signal space for signal detection;***
- ***The fundamental detection problem in a receiver;***

As mentioned earlier in Module #1, a digital modulator is supposed to accept stream of information-bearing symbols (usually bits) and represent them appropriately with or without the help of a carrier. So, a very important issue in-between is to represent information symbols in suitable energy signals so that the signals can be modulated, amplified and transmitted. For a continuous stream of input of information sequence what kind of strategy should we take to represent them as signals? One may think of multiple alternatives including the following:

- Consider one symbol at a time and design a signal for the symbol.
- When several bits make one symbol, consider one bit at a time and design for the symbol.
- Consider a larger group of symbols in a sequence and design signals spread over long time duration [sequence based modulation - demodulation strategy].

Let us consider a systematic approach to identify M symbols from the input information sequence. If the format of input information is known, this is not a difficult task. For example, if the information sequence is binary and if we choose M = 2, we can identify '1' as one symbol and '0' as the other. Else, if we choose M = 4 for the same binary information sequence; we may consider a group of two bits at a time to define one symbol. The duration of a symbol now is twice the duration of one information bit. If the rate of incoming information is $R_b$ bits/sec, the symbol rate is $R_b/2$ symbols per second. Usually, for practical considerations, M is so chosen that $M = 2^m$, where 'm' is a positive integer.

The next issue is to design 'M' energy signals for these M symbols such that the energy of each signal is limited within the symbol duration. This problem is addressed in general by a scheme known as Gram-Schmidt Orthogonalization

## Gram-Schmidt Orthogonalization

The principle of Gram-Schmidt Orthogonalization (GSO) states that, any set of M energy signals, $\{s_i(t)\}$, $1 \le i \le M$ can be expressed as linear combinations of N orthonormal basis functions, where $N \le M$.

If $s_1(t)$, $s_2(t)$, ….., $s_M(t)$ are real valued energy signals, each of duration 'T' sec,

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\Phi_j(t)\,; \quad \begin{cases} 0 \le t \le T \\ i = 1, 2, ...., M \ge N \end{cases} \qquad 4.16.1$$

where,

$$s_{ij} = \int_0^T s_i(t)\varphi_j(t)dt \quad ; \begin{cases} i = 1,2,......,M \\ j = 1,2,......,N \end{cases} \qquad 4.16.2$$

The $\varphi_j(t)$-s are the basis functions and '$s_{ij}$'-s are scalar coefficients. We will consider real-valued basis functions $\varphi_j$ (t) - s which are orthonormal to each other, i.e.,

$$\int_0^T \varphi_i(t).\varphi_j(t)dt = \begin{cases} 1, if \ i = j \\ 0, if \ i \neq j \end{cases} \qquad 4.16.3$$

Note that each basis function has unit energy over the symbol duration 'T'. Now, if the basis functions are known and the scalars are given, we can generate the energy signals, by following **Fig. 4.16.1**. Or, alternatively, if we know the signals and the basis functions, we know the corresponding scalar coefficients (**Fig. 4.16.2**).

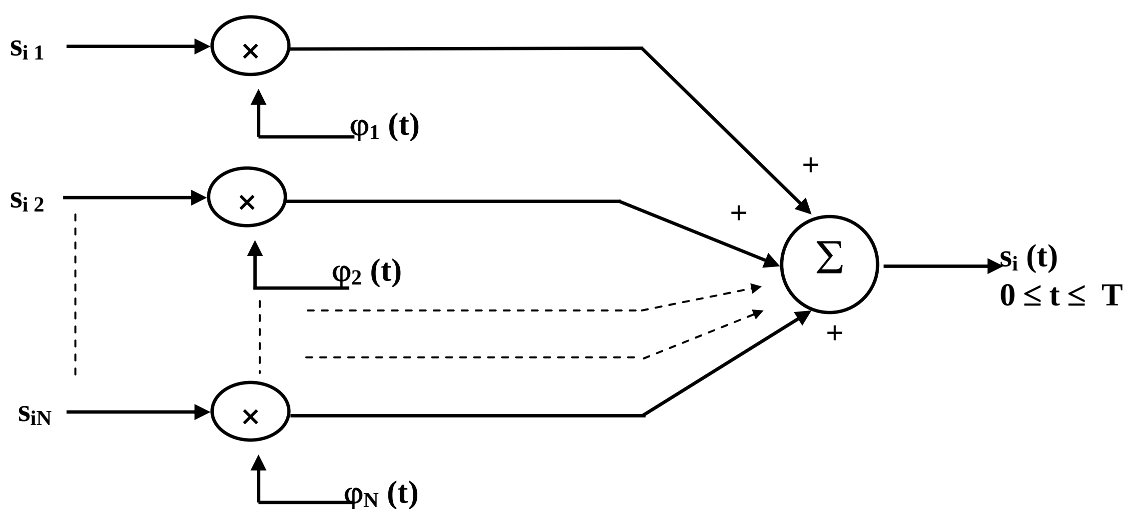


**Fig. 4.16.1** *Pictorial depiction of Equation 4.16.1*

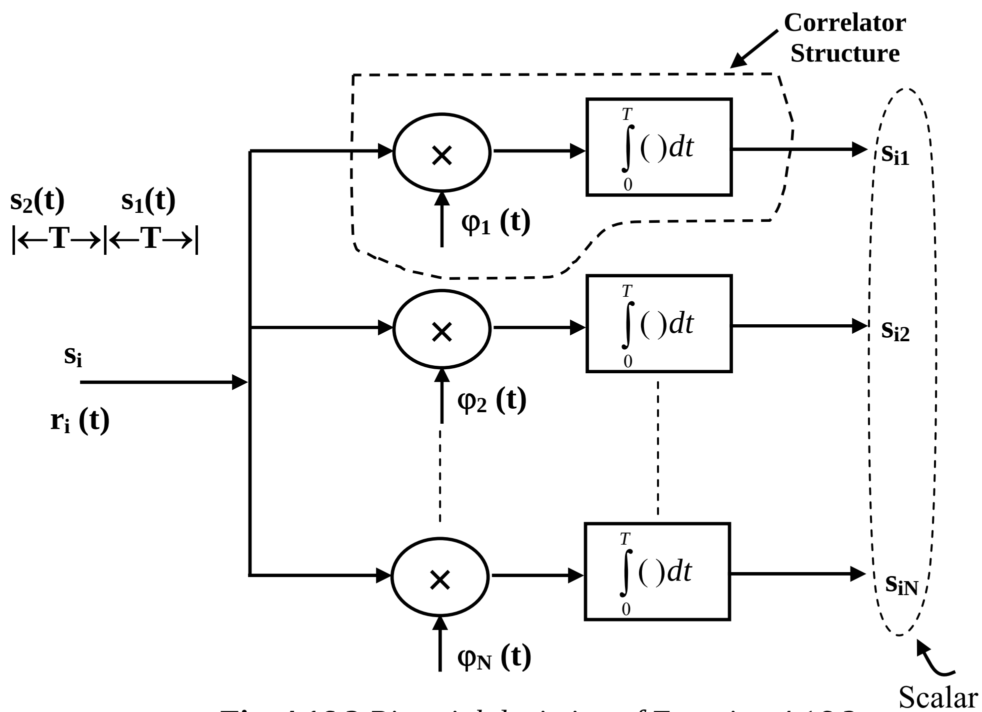


**Fig. 4.16.2** *Pictorial depiction of Equation 4.16.2*

# Justification for G-S-O procedure

**Part – I:** We show that any given set of energy signals, $\{s_i(t)\}$, $1 \le i \le M$ over $0 \le t < T$, can be completely described by a subset of energy signals whose elements are linearly independent.

To start with, let us assume that all $s_i(t)$ -s are not linearly independent. Then, there must exist a set of coefficients $\{a_i\}$, $1 < i \le M$, not all of which are zero, such that,

$$a_1 s_1(t) + a_2 s_2(t) + \ldots\ldots + a_M s_M(t) = 0, \quad 0 \le t < T \qquad 4.16.4$$

Verify that even if two coefficients are not zero, e.g. $a_1 \neq 0$ and $a_3 \neq 0$, then $s_1(t)$ and $s_3(t)$ are dependent signals.

Let us arbitrarily set, $a_M \neq 0$. Then,

$$s_M(t) = -\frac{1}{a_M}\left[a_1 s_1(t) + a_2 s_2(t) + \ldots\ldots + a_{M-1} s_{M-1}(t)\right]$$
$$= -\frac{1}{a_M}\sum_{i=1}^{M-1} a_i s_i(t) \qquad 4.16.5$$

Eq.4.16.5 shows that $s_M(t)$ could be expressed as a linear combination of other $s_i(t)$ – s, i = 1, 2, .., (M – 1).

Next, we consider a reduced set with (M-1) signals $\{s_i(t)\}$, i = 1,2,….., (M – 1). This set may be either linearly independent or not. If not, there exists a set of $\{b_i\}$, i = 1,2…, (M – 1), not all equal to zero such that,

$$\sum_{i=1}^{M-1} b_i s_i(t) = 0, \quad 0 \le t < T \qquad 4.16.6$$

Again, arbitrarily assuming that $b_{M-1} \neq 0$, we may express $s_{M-1}(t)$ as:

$$s_{M-1}(t) = -\frac{1}{b_{M-1}} \sum_{i=1}^{M-2} b_i s_i(t) \qquad 4.16.7$$

Now, following the above procedure for testing linear independence of the remaining signals, eventually we will end up with a subset of linearly independent signals. Let $\{s_i(t)\}$, i = 1, 2, …., N ≤ M denote this subset.

**Part – II :** We now show that it is possible to construct a set of 'N' orthonormal basis functions $\varphi_1(t), \varphi_2(t), \ldots, \varphi_N(t)$ from $\{s_i(t)\}$, i = 1, 2, ….., N.

Let us choose the first basis function as, $\varphi_1(t) = \frac{s_1(t)}{\sqrt{E_1}}$, where $E_1$ denotes the energy of the first signal $s_1(t)$, i.e., $E_1 = \int_0^T s_1^2(t)dt$:

$$\therefore s_1(t) = \sqrt{E_1}.\varphi_1(t) = s_{11}\varphi_1(t) \qquad 4.16.8$$

Where, $s_{11} = \sqrt{E_1}$

Now, from Eq. 4.16.2, we can write

$$s_{21} = \int_0^T s_2(t)\phi_1(t)dt \qquad 4.16.9$$

Let us now define an intermediate function:

$$g_2(t) = s_2(t) - s_{21}\varphi_1(t); \quad 0 \le t < T \qquad 4.16.10$$

Note that,

$$\int_0^T g_2(t)\varphi_1(t)dt = \int_0^T s_2(t)\varphi_1(t)dt - s_{21}\int_0^T \varphi_1(t)\varphi_1(t)dt$$

$$= s_{21} - s_{21} \quad = 0 \rightarrow g_2(t) \text{ Orthogonal to } \varphi_1(t);$$

So, we verified that the function $g_2(t)$ is orthogonal to the first basis function. This gives us a clue to determine the second basis function.

Now, energy of $g_2(t)$

$$= \int_0^T g_2^2(t)dt$$

$$= \int_0^T \left[ s_2(t) - s_{21}\varphi_1(t) \right]^2 dt$$

$$= \int_0^T s_2^{\,2}(t)dt - 2.s_{21}\int_0^T s_2(t)\varphi_1(t)dt + s_{21}^{\,2}\int_0^T \varphi_1^{\,2}(t)dt$$

$$= E_2 - 2.s_{21}.s_{21} + s_{21}^{\,2} = E_2 - s_{21}^{\,2} \quad (Say) \qquad 4.16.11$$

So, we now set,

$$\varphi_2(t) = \frac{g_2(t)}{\sqrt{\int_0^T g_2^{\,2}(t)dt}} = \frac{s_2(t) - s_{21}\varphi_1(t)}{\sqrt{E_2 - s_{21}^{\,2}}} \qquad 4.16.12$$

and $E_2 = \int_0^T s_2^{\,2}(t)dt$: Energy of $s_2(t)$

Verify that:

$\int_0^T \varphi_2^{\,2}(t)dt = 1$, i.e. $\varphi_2(t)$ is a time limited energy signal of unit energy.

and $\int_0^T \varphi_1(t).\varphi_2(t)dt = 0$, i.e. $\varphi_1(t)$ and $\varphi_2(t)$ are orthonormal to each other.

Proceeding in a similar manner, we can determine the third basis function, $\varphi_3(t)$. For i=3,

$$g_3(t) = s_3(t) - \sum_{j-1}^{2} s_{3j}\varphi_J(t); \qquad 0 \le t < T$$

$$= s_3(t) - \left[ s_{31}\varphi_1(t) + s_{32}\varphi_2(t) \right]$$

where,

$$s_{31} = \int_0^T s_3(t)\varphi_1(t)dt \quad \text{and} \quad s_{32} = \int_0^T s_3(t)\varphi_2(t)dt$$

It is now easy to identify that,

$$\varphi_3(t) = \frac{g_3(t)}{\sqrt{\int_0^T g_3^{\,2}(t)dt}} \qquad 4.16.13$$

Indeed, in general,

$$\varphi_i(t) = \frac{g_i(t)}{\sqrt{\int_0^T g_i^{\,2}(t)dt}} = \frac{g_i(t)}{\sqrt{Eg_i}} \qquad 4.16.14$$

for i = 1, 2,….., N, where

$$g_i(t) = s_i(t) - \sum_{j=1}^{i-1} s_{ij}\varphi_j(t) \qquad 4.16.15$$

and $$s_{ij} = \int_0^T s_i(t).\varphi_j(t)dt \qquad 4.16.16$$

for i = 1, 2,….., N and j = 1, 2, …, M

Let us summarize the steps to determine the orthonormal basis functions following the Gram-Schmidt Orthogonalization procedure:

If the signal set $\{s_j(t)\}$ is known for j = 1, 2,….., M, $0 \le t < T$,

- Derive a subset of linearly independent energy signals, $\{s_i(t)\}$,i = 1, 2,….., $N \le M$.
- Find the energy of $s_1(t)$ as this energy helps in determining the first basis function $\varphi_1(t)$, which is a normalized form of the first signal. Note that the choice of this 'first' signal is arbitrary.
- Find the scalar '$s_{21}$', energy of the second signal ($E_2$), a special function '$g_2(t)$' which is orthogonal to the first basis function and then finally the second orthonormal basis function $\varphi_2(t)$
- Follow the same procedure as that of finding the second basis function to obtain the other basis functions.

## Concept of signal space

Let, for a convenient set of $\{\varphi_j(t)\}$, j = 1,2,…,N and $0 \le t < T$,

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\varphi_j(t), \quad \text{i= 1, 2,……, M and } 0 \le t < T\text{, such that,}$$

$$s_{ij} = \int_0^T s_i(t)\varphi_j(t)dt$$

Now, we can represent a signal $s_i(t)$ as a column vector whose elements are the scalar coefficients $s_{ij}$, j = 1, 2, ….., N :

$$\overline{s_i} = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix}_{1\times N} \quad \text{; i= 1, 2,……, M} \qquad 4.16.17$$

These M energy signals or vectors can be viewed as a set of M points in an N – dimensional Euclidean space, known as the '*Signal Space'* (**Fig.4.16.3**). *Signal Constellation* is the collection of M signals points (or messages) on the signal space.

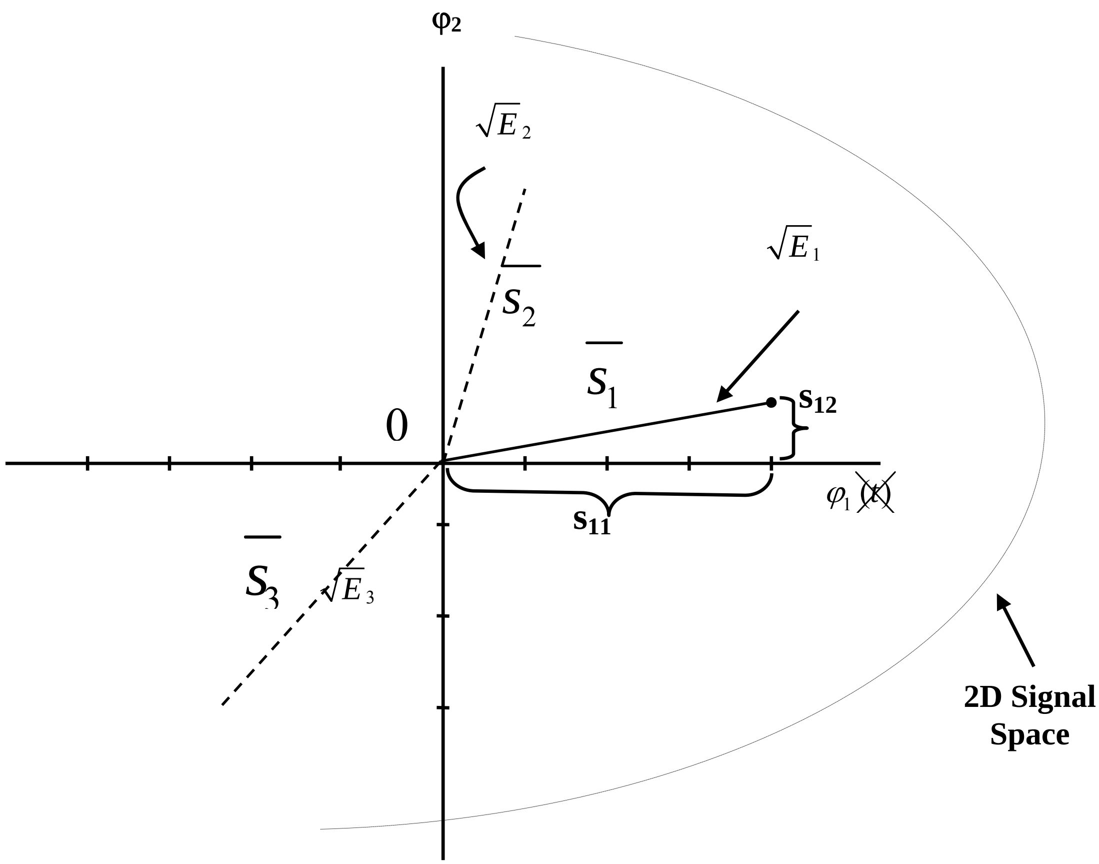


**Fig. 4.16.3** *Sketch of a 2-dimentional-signal space showing three signal vectors* $\overline{s_1}$, $\overline{s_2}$ and $\overline{s_3}$

Now, the length or *norm* of a vector is denoted as $\left\|\overline{s_i}\right\|$. The squared norm is the inner product of the vector:

$$\left\|\overline{s_i}\right\|^2 = \left(\overline{s_i}, \overline{s_i}\right) = \sum_{j=1}^{N} s^2{}_{ij} \qquad 4.16.18$$

The cosine of the angle between two vectors is defined as:

$$\cos\left(\text{angle between } \vec{s_i} \text{ \& } \vec{s_j}\right) = \frac{\left(\vec{s_i}, \vec{s_j}\right)}{\left\|\vec{s_i}\right\| \left\|\vec{s_j}\right\|} \qquad 4.16.19$$

$\therefore \vec{s_i}$ & $\vec{s_j}$ are orthogonal to each other if $\left(\vec{s_i}, \vec{s_j}\right) = 0$.

If $E_i$ is the energy of the i-th signal vector,

$$E_i = \int_0^T s_i^2(t)\,dt = \int_0^T \left[\sum_{j=1}^{N} s_{ij}\varphi_j(t)\right]\left[\sum_{k=1}^{N} s_{ik}\varphi_k(t)\right] dt$$

$$= \sum_{j=1}^{N}\sum_{k=1}^{N} s_{ij}s_{ik}\int_0^T \varphi_j(t)\varphi_k(t)\,dt \quad \text{as } \{\varphi_j(t)\} \text{ forms an ortho-normal set}$$

$$= \sum_{j=1}^{N} s_{ij}^2 = \left\|\vec{s_i}\right\|^2 \qquad 4.16.20$$

For a pair of signals $s_i$ (t) and $s_k$(t), $\left\|\vec{s_i}-\vec{s_k}\right\|^2 = \sum_{j=1}^{N}\left(s_{ij}-s_{kj}\right)^2 = \int_0^T \left[s_i(t)-s_k(t)\right]^2 dt$

It may now be guessed intuitively that we should choose $s_i$ (t) and $s_k$ (t) such that the Euclidean distance between them, i.e. $\left\|\vec{s_i}-\vec{s_k}\right\|$ is as much as possible to ensure that their detection is more robust even in presence of noise. For example, if $s_1$ (t) and $s_2$ (t) have same energy E, (i.e. they are equidistance from the origin), then an obvious choice for maximum distance of separation is, $s_1(t) = -s_2(t)$.

## Use of Signal Space for Signal Detection in a Receiver

The signal space defined above, is very useful for designing a receiver as well. In a sense, much of the features of a modulation scheme, such as the number of symbols used and the energy carried by the symbols, is embedded in the description of its signal space. So, in absence of any noise, the receiver should detect one of these valid symbols only. However, the received symbols are usually corrupted and once placed in the signal space, they may not match with the valid signal points in some respect or the other. Let us briefly consider the task of a good receiver in such a situation. Let us assume the following:

1. One of the M signals $s_i(t)$, i=1,2,….,M is transmitted in each time slot of duration 'T' sec.
2. All symbols are equally probable, i.e. the probability of occurrence of $s_i(t)$ = 1/M, for all 'i'.
3. Additive White Gaussian Noise (AWGN) processes W (t) is assumed with a noise sample function w (t) having mean = 0 and power spectral density $\frac{N_0}{2}$ [$N_0$: single sided power spectral density of additive white Gaussian noise. Noise is discussed more in next two lessons]
4. Detection is on a symbol-by-symbol basis.

Now, if R (t) denotes the received random process with a sample function r (t), we may write,

$$r(t) = s_i(t) + w(t) \quad ; \quad 0 \le t < T \text{ and } i = 1,2,\ldots,M.$$

The job of the receiver is to make "best estimate" of the transmitted signal $s_i$ (t) (or, equivalently, the corresponding message symbol $m_i$) upon receiving r(t). We map the received sample function r (t) on the signal space to include a 'received vector' or 'received signal point'. This helps us to identify a noise vector, w(t) ,also. The detection problem can now be stated as:

'*Given an observation / received signal vector ($\bar{r}$), the receiver has to perform a mapping from $\bar{r}$ to an estimate $\hat{m}$ for the transmitted symbol $m_i$ in a way that would minimize the average probability of symbol error'.*

Maximum Likelihood Detection scheme provides a general solution to this problem when the noise is additive and Gaussian. We discuss this important detection scheme in Lesson #19.

## Problems

Q4.16.1) Sketch two signals, which are orthonormal to each other over 1 sec. Verify that Eq4.16.3 is valid.

Q4.16.2) Let, $S_1(t) = \text{Cos}2\pi ft$, $S_2(t) = \text{Cos}(2\pi ft + \pi/3)$ and $S_3(t) = \text{Sin}\, 2\pi ft$. Comment whether the three signals are linearly independent?

Q4.16.3) Consider a binary random sequence of 1 and 0. Draw a signal constellation for the same.