ECE 8700: Communication Systems Engineering

Module 1
Signal Space Analysis

Mojtaba Vaezi
http://www.ece.villanova.edu/ mvaezi

Villanova University

March 23, 2025

Contents

1 Baseband vs. Bandpass Signals **5**

2 Lowpass Equivalent of Bandpass Signals **12**

3 Linear Algebra: A Primer **23**

3.1 Vector Space Concepts 23

4 Orthogonal Sets **32**

5 Gram–Schmidt Orthogonalization **37**

6 Linear Vector Space and Subspace **44**

6.1 Euclidean Space 45

6.2 Euclidean Space of Complex-Valued Vectors 45

6.3 Subspaces of a Linear Space 47

6.4 Span of a Set of Vectors 48

7 Inner Product for Signals **54**

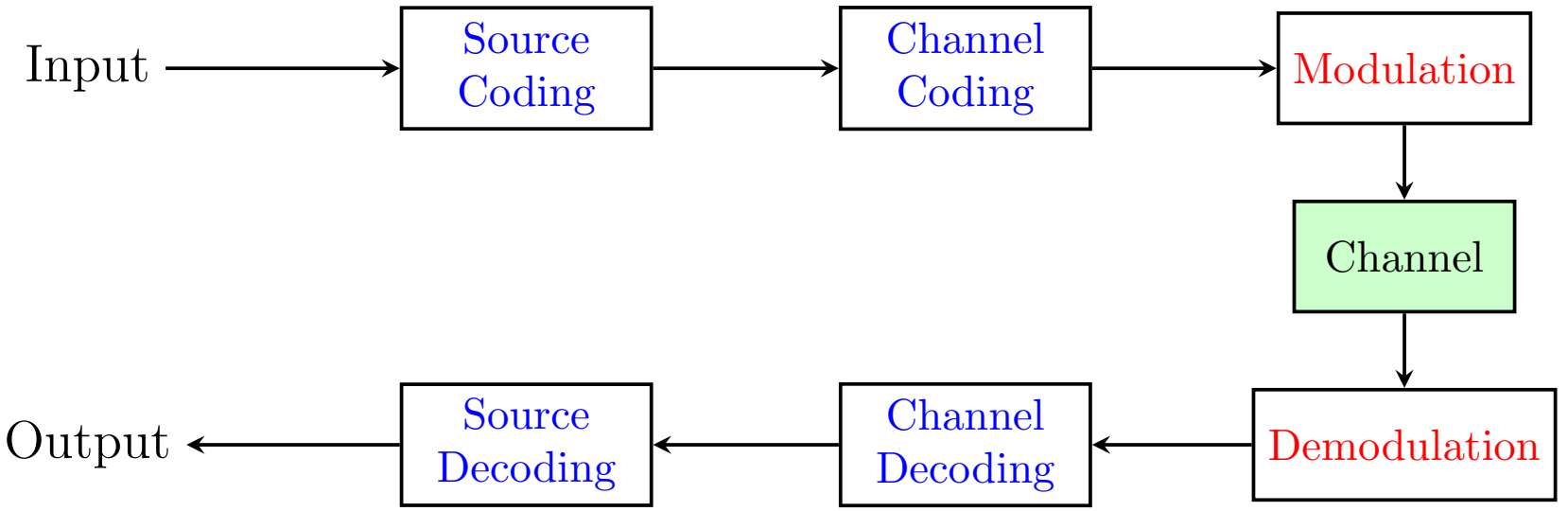
7.1 Geometric (Vectorial) Representation of Signals 57

7.2 Gram-Schmidt Orthogonalization for Signals 68

8 Useful Random Variables **78**

9 Appendix: Vector Space **83**

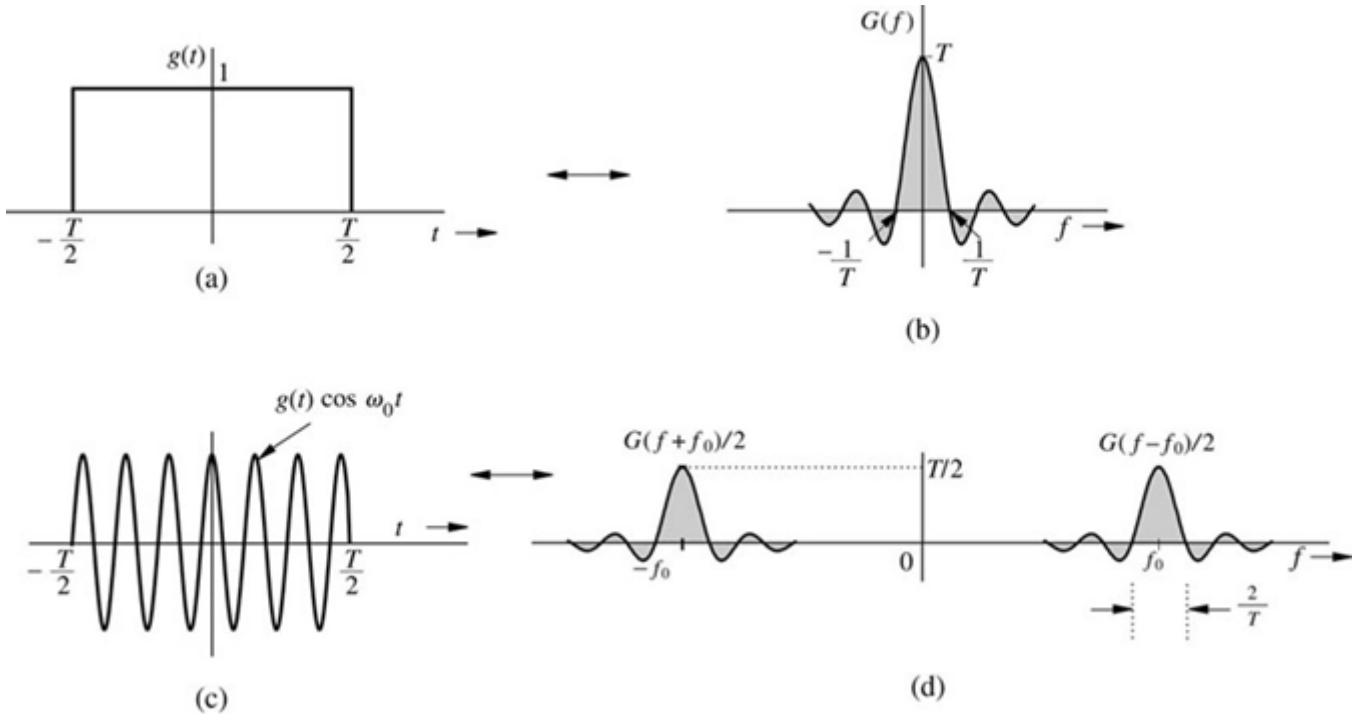
- A digital transmission scheme generally involves



- Input/output are considered digital (if analogue sampled/quantized)
- **CODEC**, **MODEM**, **channel** (transmission medium)
- Your cell-phone, for example, contains a pair of transmitter and receiver, consisting of a **CODEC** and **MODEM**

1 Baseband vs. Bandpass Signals

- If the spectrum of a signal is centered at $f = 0$, the signal is called a **baseband, or lowpass,** signal.



- If it is centered at a higher frequency, it is called a **bandpass** signal.

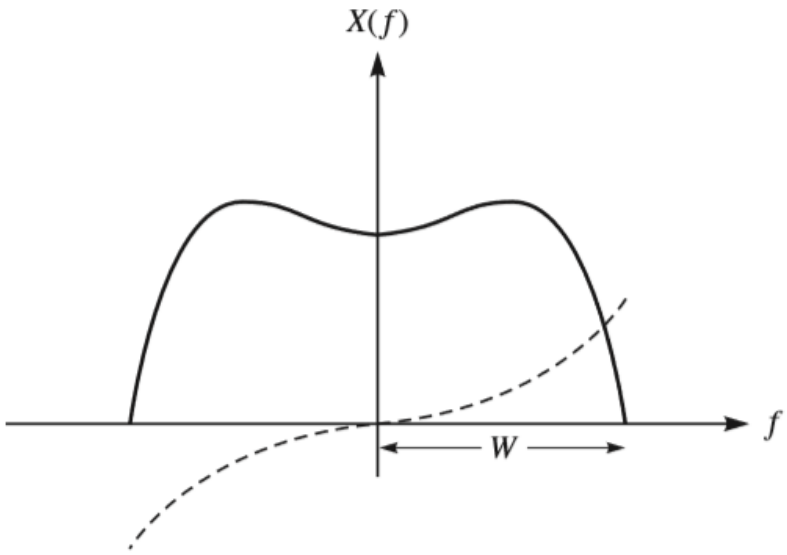


Figure 1: The spectrum of a real-valued lowpass (baseband) signal.

- A baseband signal is a signal with the spectrum centered at DC. For instance, speech, music, and video signals are all lowpass signals.
- The bandwidth of a real lowpass signal is the smallest positive W such that $X(f) = 0$ outside $[-W, W]$.

- The spectrum can be separated into two parts: positive and negative
- We define the **positive spectrum** and the **negative spectrum** of a signal $x(t)$ as

$$X_+(f) = \begin{cases} X(f) & f > 0, \\ \frac{1}{2}X(0) & f = 0, \\ 0 & f < 0, \end{cases} \qquad X_-(f) = \begin{cases} X(f) & f < 0, \\ \frac{1}{2}X(0) & f = 0, \\ 0 & f > 0, \end{cases}$$

$$X(f) = X_+(f) + X_-(f)$$

- It is also clear that

$$X_+(f) = X(f)u(f),$$
$$X_-(f) = X(f)u(-f),$$

where $u(\cdot)$ is the unit step signal.

If $x(t)$ is real-valued signal, then we get $X(f) = X^*(-f)$.
That is, $X(f)$ obeys conjugate symmetry, and we have

- $|X(f)| = |X^*(-f)| = |X(-f)|$ i.e., magnitude is even
- $\angle X(f) = \angle X^*(-f) = -\angle X(-f)$ i.e., phase is odd

A signal $x(t)$ is said to be **Hermitian** if its Fourier transform $X(f)$ satisfies the following property:

$$X(-f) = X^*(f),$$

where $X^*(f)$ denotes the complex conjugate of $X(f)$.

– For a real-valued time-domain signal $x(t)$, its Fourier transform $X(f)$ is Hermitian.

– This implies that the magnitude of $X(f)$ is even:

$$|X(-f)| = |X(f)|.$$

– The phase of $X(f)$ is odd:

$$\angle X(-f) = -\angle X(f)$$

Example 1. *Show that* $x(t) = e^{-at}u(t)$, $a > 0$ *is Hermitian, i.e.,* $X(f)$ *obeys conjugate symmetry.*

Solution. From the table, we know that $X(f) = \frac{1}{a+j2\pi f}$. Thaus,

$$X(f) = \frac{1}{a + j2\pi f} \Rightarrow X(-f) = \frac{1}{a - j2\pi f} \Rightarrow X^*(-f) = \frac{1}{a + j2\pi f}$$

Then it is easy to see that

$|X(f)| = |X^*(-f)| = |X(-f)|$ i.e., magnitude is even

$\angle X(f) = \angle X^*(-f) = -\angle X(-f)$ i.e., phase is odd

Fourier Transform Symmetries	
CT Signal	**Fourier Transform**
$x(t)$	$X(f)$
$x(-t)$	$X(-f)$
$x^*(t)$	$X^*(-f)$
$x^*(-t)$	$X^*(f)$

Table 1: Fourier Transform Symmetries for Continuous-Time Signals

Useful FT Pairs

Signal	Time Function	$\rightleftharpoons$	Fourier Transform
Delta	$\delta(t)$	$\rightleftharpoons$	1
DC	1	$\rightleftharpoons$	$\delta(f)$
Shifted delta	$\delta(t-t_0)$	$\rightleftharpoons$	$e^{-j2\pi f t_0}$
Complex exponential	$e^{j2\pi f_c t}$	$\rightleftharpoons$	$\delta(f-f_c)$
Cosine	$\cos 2\pi f_c t$	$\rightleftharpoons$	$\frac{1}{2}[\delta(f-f_c)+\delta(f+f_c)]$
Sine	$\sin 2\pi f_c t$	$\rightleftharpoons$	$\frac{1}{2j}[\delta(f-f_c)-\delta(f+f_c)]$
Rectangle	$\mathrm{rect}\left(\frac{t}{T}\right)$	$\rightleftharpoons$	$T\,\mathrm{sinc}(fT)$
Sinc	$\mathrm{sinc}(2Wt)$	$\rightleftharpoons$	$\frac{1}{2W}\mathrm{rect}\left(\frac{f}{2W}\right)$
Triangle	$\Lambda\left(\frac{t}{T}\right)$	$\rightleftharpoons$	$T\,\mathrm{sinc}^2(fT)$
Signum	$\mathrm{sign}(t)$	$\rightleftharpoons$	$\frac{1}{j\pi f}$
	$\frac{1}{\pi t}$	$\rightleftharpoons$	$-j\mathrm{sign}(f)$
Unite step	$u(t)$	$\rightleftharpoons$	$\frac{1}{j2\pi f}+\frac{1}{2}\delta(f)$
Exponential, $a>0$	$e^{-at}u(t)$	$\rightleftharpoons$	$\frac{1}{a+j2\pi f}$
Gaussian	$e^{-\pi t^2}$	$\rightleftharpoons$	$e^{-\pi f^2}$
Delta train	$\sum_{m=-\infty}^{\infty}\delta(t-mT_0)$	$\rightleftharpoons$	$\frac{1}{T_0}\sum_{n=-\infty}^{\infty}\delta(f-\frac{n}{T_0})$

- A **bandpass signal** is a real signal whose frequency content, or spectrum, is located around some frequency $\pm f_0$, which is far from zero.

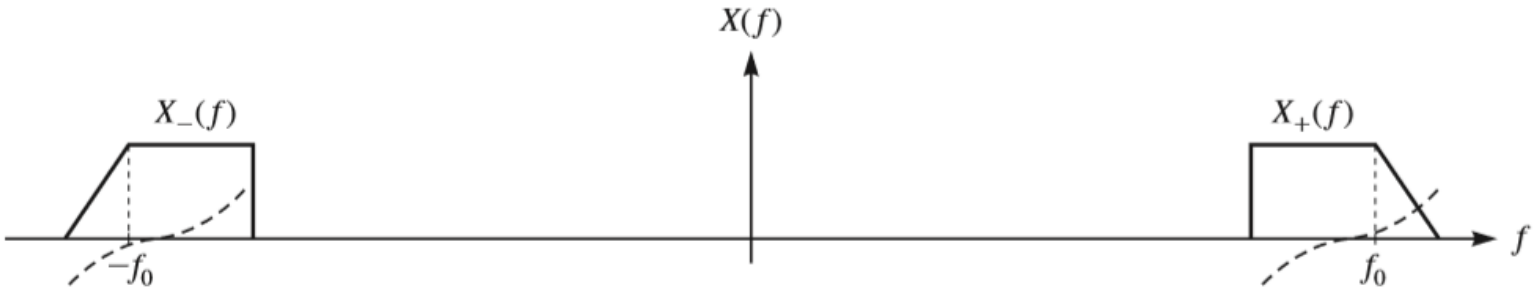


Figure 2: The spectrum of a real-valued bandpass signal.

- Due to the symmetry of the spectrum, $X^+(f)$ has all the information that is necessary to reconstruct $X(f)$. In fact, we can write

$$\begin{aligned} X(f) &= X^+(f) + X^-(f) \\ &= X^+(f) + X_+^*(-f), \end{aligned}$$

- Since any real signal is Hermitian, $X^-(f) = X_+^*(-f)$

Thus, the knowledge of $X^+(f)$ is sufficient to reconstruct $X(f)$, for real signals. This is not true for complex signals, as their spectrum is not symmetric.

2 Lowpass Equivalent of Bandpass Signals

- A bandpass signal can be represented in one of four ways,
  - Amplitude/phase
  - In-phase/quadrature
  - Complex envelope
  - Analytic signal
- A bandpass signal is "real", but may contain "unnecessary" content such as the carrier frequency f_c that has nothing to do with the digital information transmitted.
- So, it is more convenient to remove this carrier frequency and transform $x(t)$ into its lowpass equivalent signal $x_l(t)$ before "analyzing" the digital content.
- The goal is to find a time signal whose Fourier transform is $X_+(f)$, i.e.,

$$x_+(t) \iff X_+(f)$$

- Because $X_+(f)$ is non-symmetric in general, the corresponding time function is not real (is complex).

$$\begin{aligned} x_+(t) &= \mathcal{F}^{-1}[X_+(f)] \\ &= \mathcal{F}^{-1}[X(f)u(f)] \\ &= x(t) * \left(\frac{1}{2}\delta(t) + j\frac{1}{2\pi t}\right) \\ &= \frac{1}{2}x(t) + j\frac{1}{2}\hat{x}(t) \end{aligned}$$

- $\hat{x}(t)$ is called the *Hilbert transform* of $x(t)$.

$$\hat{x}(t) \triangleq \frac{1}{\pi t} * x(t)$$

$$\mathcal{F}[\hat{x}(t)] = -j\,\mathrm{sign}(f)X(f) = \begin{cases} -jX(f), & f > 0 \\ jX(f), & f < 0 \end{cases}$$

- Multiplying by j is equivalent to multiplication by $e^{j\pi/2}$, so the Hilbert transform is just a phase shifter, shifting positive frequencies by $-90°$ and negative frequencies by $90°$.

Definition 1. *We define $x_\ell(t)$, lowpass equivalent, or the complex envelope, of $x(t)$ as the signal whose spectrum is given by $2X_+(f+f_0)$, i.e.,*

$$X_l(f) = 2X_+(f+f_0) = 2X(f+f_0)u(f+f_0).$$

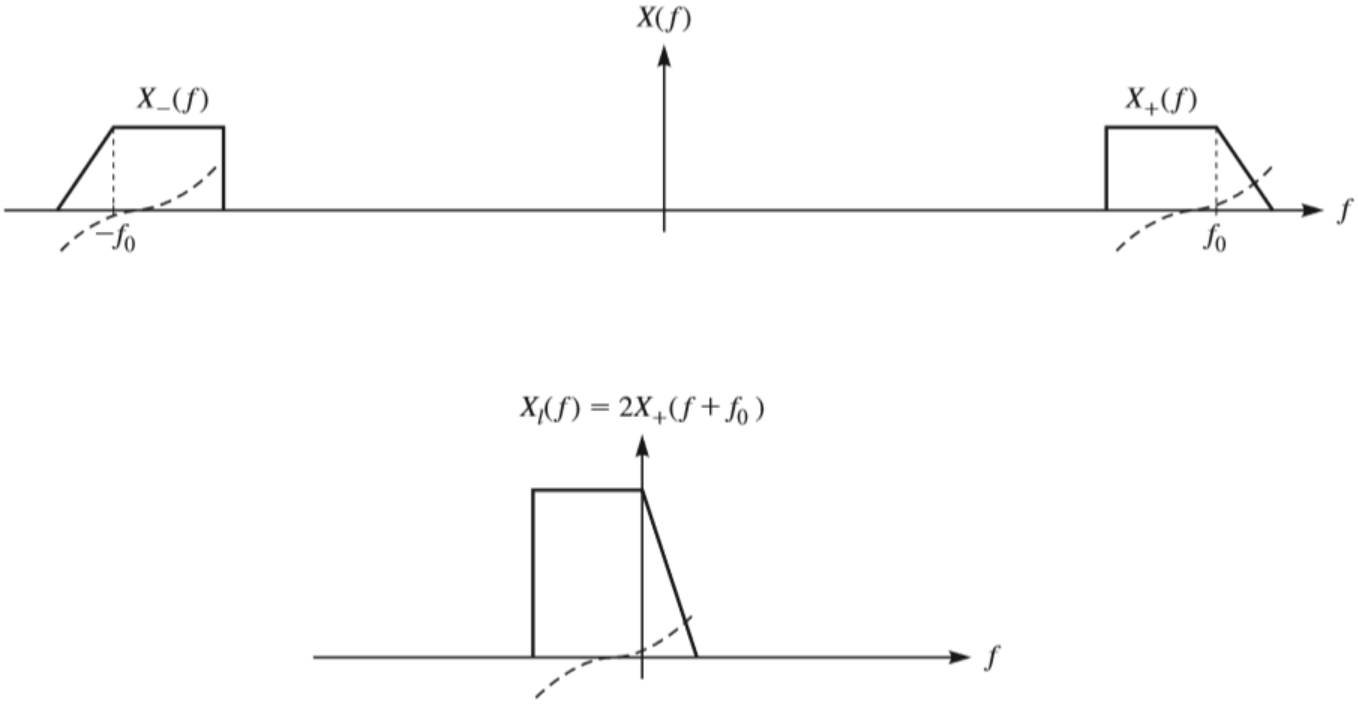
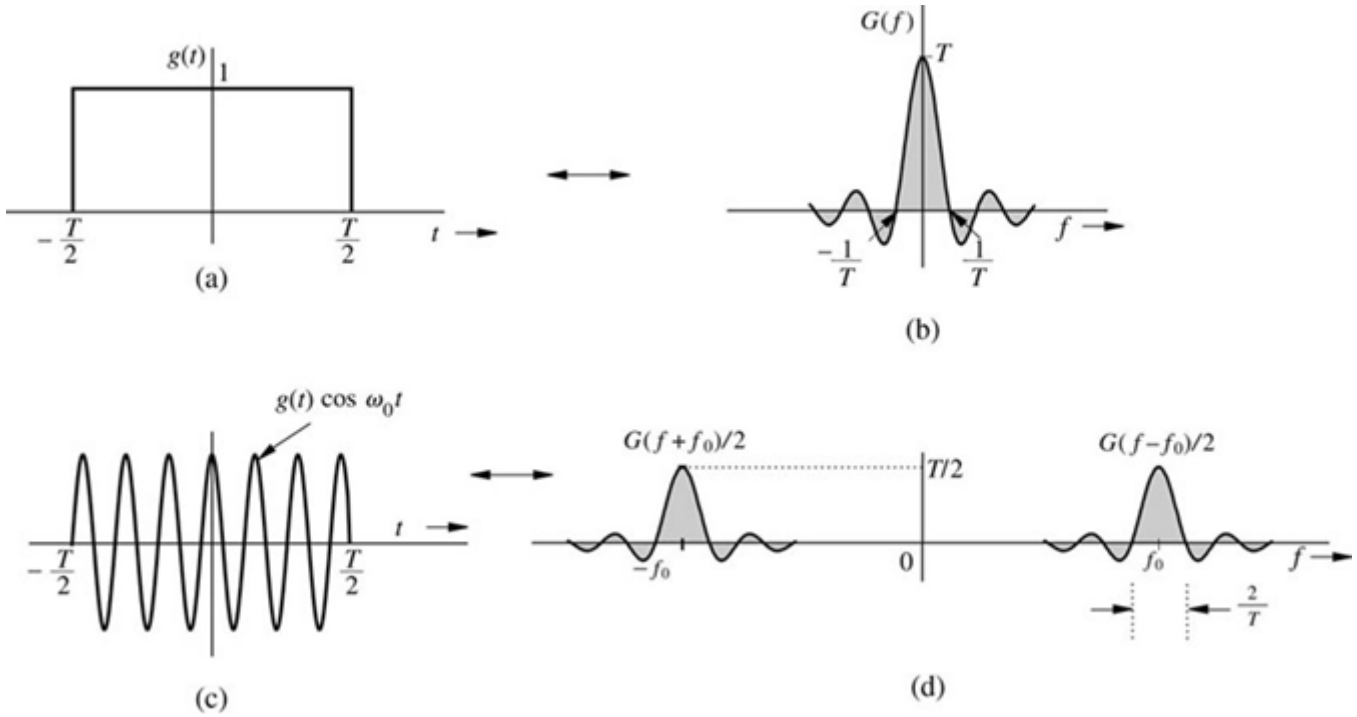


Figure 3: **(top)** The spectrum of a real-valued bandpass signal.
(bottom) The spectrum of the lowpass equivalent of the signal shown on top.

Example 2. *Show that the lowpass equivalent of an RF pulse is its envelope which is a rectangular pulse.*



- Taking the inverse Fourier transform of $X_l(f)$, we get

$$\begin{aligned} x_l(t) &= \mathcal{F}^{-1}[X_l(f)] \\ &= \mathcal{F}^{-1}[2X_+(f+f_0)] \\ &= 2x_+(t)e^{-j2\pi f_0 t} \\ &= \underbrace{[x(t)+j\hat{x}(t)]}_{\text{analytic}} e^{-j2\pi f_0 t} \end{aligned}$$

- From the above, we can write

$$\underbrace{x(t)}_{\text{bandpass}} = \text{Re}\Big[\underbrace{x_l(t)}_{\text{lowpass}} e^{j2\pi f_0 t}\Big]$$

  This relation expresses any bandpass signal in terms of its lowpass equivalent.

- Further, we can write

$$X(f) = \frac{1}{2}\left[X_l(f-f_0) + X_l^*(-f-f_0)\right]$$

Lowpass Equivalent of Bandpass Signals

Summary

- Bandpass signal

$$x(t) = \mathrm{Re}\{x_\ell(t)e^{j2\pi f_0 t}\}$$
$$X(f) = \frac{1}{2}\left[X_\ell(f - f_0) + X_\ell^*(-f - f_0)\right]$$

- Lowpass equivalent signal or complex envelope

$$x_\ell(t) = \left[x(t) + j\hat{x}(t)\right] e^{-j2\pi f_0 t}$$
$$X_\ell(f) = 2X(f + f_0)u(f + f_0)$$

I-Q Representation of Signals

- The real and imaginary parts of the complex envelope $x_l(t)$ are called the *in-phase* (I) and *quadrature* (Q) components of $x(t)$.

$$x_l(t) = x_i(t) + jx_q(t) \tag{1}$$

- Both $x_i(t)$ and $x_q(t)$ are real-valued lowpass signals.
- The lowpass equivalent was the downshifted analytic signal

$$\begin{aligned} x_l(t) &= \big(x(t) + j\hat{x}(t)\big)e^{-j2\pi f_0 t} \\ &= \big[x(t)\cos(2\pi f_0 t) + \hat{x}(t)\sin(2\pi f_0 t)\big] \\ &\quad + j\big[\hat{x}(t)\cos(2\pi f_0 t) - x(t)\sin(2\pi f_0 t)\big] \end{aligned} \tag{2}$$

- Thus, equating (1) and (2) we get:

> Any bandpass signal $x(t)$ can be expressed in terms of two lowpass signals, namely, its I and Q components:
>
> $$x(t) = x_i(t)\cos(2\pi f_0 t) - x_q(t)\sin(2\pi f_0 t)$$

I-Q Representation of Signals

- Modulator diagram (the right one is pretty common)

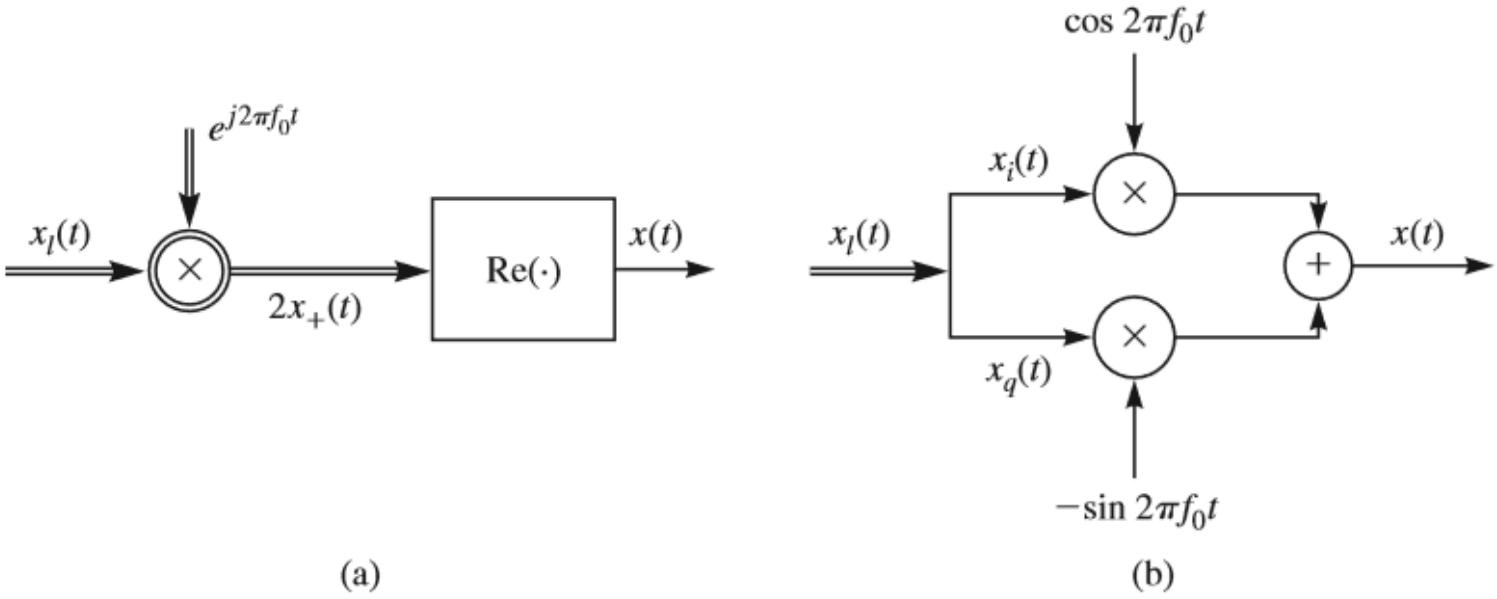
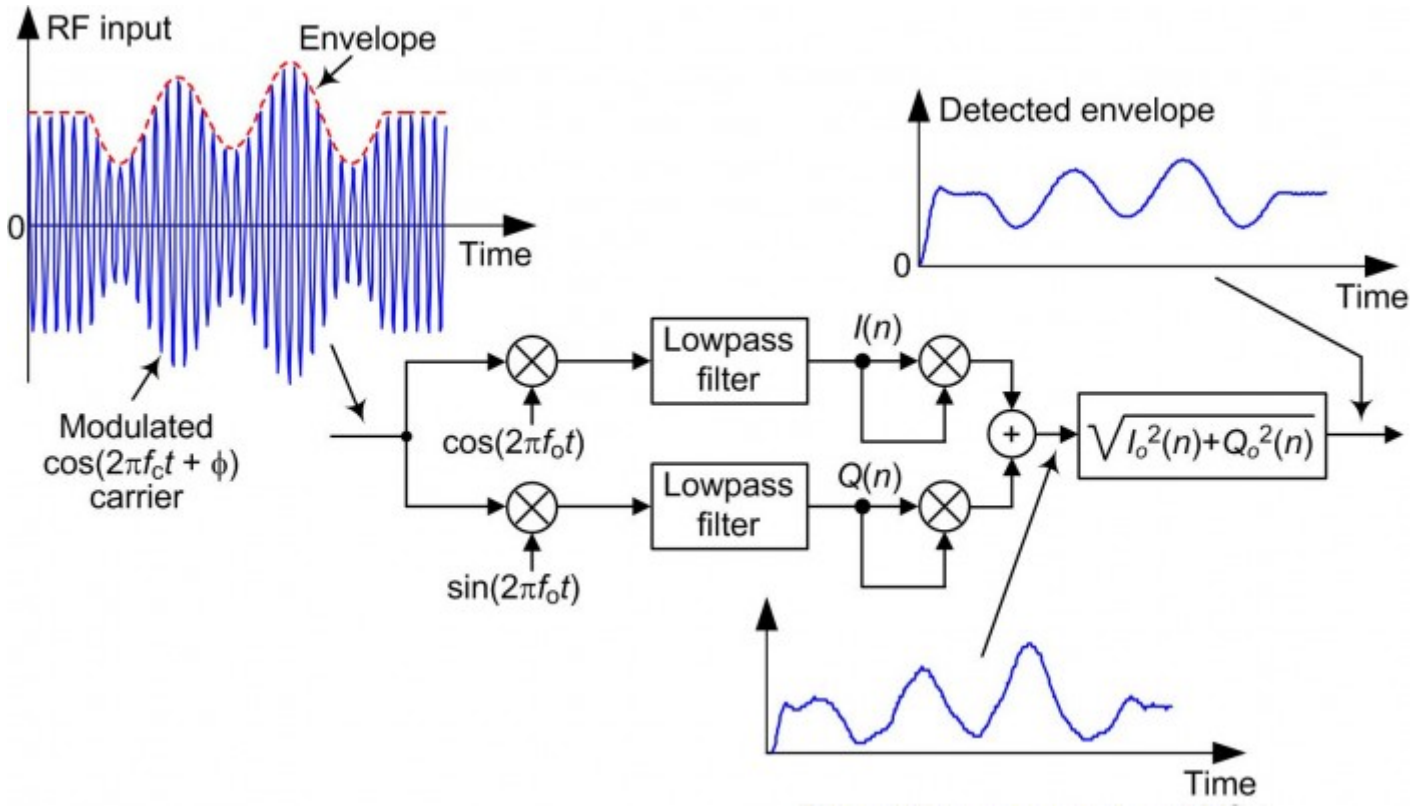


Figure 4: A complex (a) and real (b) modulator.

I-Q Representation of Signals

- Asynchronous ($f_o = f_c \pm 25\% f_c$) complex envelope detector



[Figure source]

- We define the *envelope* and *phase* of $x(t)$, denoted by $r_x(t)$ and $\theta_x(t)$, by

$$r_x(t) = \sqrt{x_i^2(t) + x_q^2(t)}$$

$$\theta_x(t) = \arctan\left(\frac{x_q(t)}{x_i(t)}\right)$$

we have

$$x_l(t) = r_x(t)e^{j\theta_x(t)}$$

Substituting this result into $x(t) = \mathrm{Re}\{x_\ell(t)e^{j2\pi f_0 t}\}$

$$x(t) = \mathrm{Re}\left[r_x(t)e^{j(2\pi f_0 t + \theta_x(t))}\right] = r_x(t)\cos\left(2\pi f_0 t + \theta_x(t)\right)$$

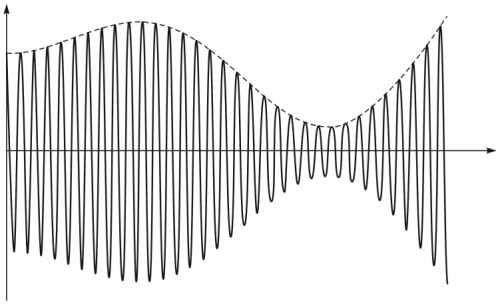
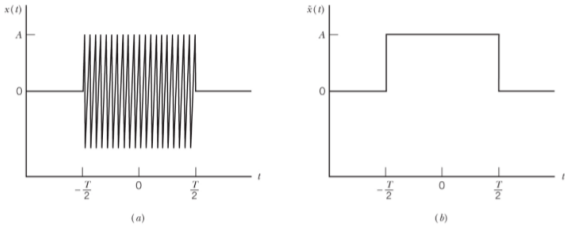


Figure 5: A bandpass signal. The dashed curve denotes the envelope.

Example 3. *Represent the following RF pulse in terms of 1) amplitude-phase, 2) complex envelope, 3) analytic signal and 4) in-phase and quadrature components.*

$$x(t) = A\,rect(t/T)cos(2\pi f_o t)$$



3 Linear Algebra: A Primer

3.1 Vector Space Concepts

The **inner product** on $\mathbb{R}^n$ generalizes the notion of the **dot product** of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ that you may already be familiar with.

Definition 2. *Let* $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ *and let* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *be vectors in* $\mathbb{R}^n$. *The* ***inner product*** *of* $\mathbf{u}$ *and* $\mathbf{v}$ *is:*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

Notice that the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ or $\mathbf{u} \cdot \mathbf{v}$ can be computed as a matrix multiplication as follows:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

The following summarizes the basic algebraic properties of the inner product.

Properties of Inner Product:
Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors in $\mathbb{R}^n$ and let α be a scalar. Then:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $(\alpha \mathbf{u}) \cdot \mathbf{v} = \alpha(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\alpha \mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = 0$

Example 4. *Let* $\mathbf{u} = (2, -5, -1)$ *and let* $\mathbf{v} = (3, 2, -3)$. *Compute* $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{v} \cdot \mathbf{u}$, $\mathbf{u} \cdot \mathbf{u}$, *and* $\mathbf{v} \cdot \mathbf{v}$.

Solution. By definition:

$$\mathbf{u} \cdot \mathbf{v} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v} \cdot \mathbf{u} = (3)(2) + (2)(-5) + (-3)(-1) = -1$$

$$\mathbf{u} \cdot \mathbf{u} = (2)(2) + (-5)(-5) + (-1)(-1) = 30$$
$$\mathbf{v} \cdot \mathbf{v} = (3)(3) + (2)(2) + (-3)(-3) = 22$$

Definition 3. *The* ***length*** *or* ***norm*** *of a vector* $\mathbf{u} \in \mathbb{R}^n$ *is defined as:*

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

A vector $\mathbf{u} \in \mathbb{R}^n$ with norm 1 will be called a *unit vector*:

$$\|\mathbf{u}\| = 1.$$

Below is an important property of the inner product.

Theorem 1. *Let* $\mathbf{u} \in \mathbb{R}^n$ *and let* α *be a scalar. Then:*

$$\|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\|.$$

By the above theorem, any non-zero vector $\mathbf{u} \in \mathbb{R}^n$ can be scaled to obtain a new unit vector in the same direction as $\mathbf{u}$. Indeed, suppose that $\mathbf{u}$ is non-zero so that

$\|\mathbf{u}\| \neq 0$. Define the new vector:

$$\mathbf{v} = \frac{1}{\|\mathbf{u}\|}\mathbf{u}.$$

Notice that $\alpha = \frac{1}{\|\mathbf{u}\|}$ is just a scalar, and thus $\mathbf{v}$ is a scalar multiple of $\mathbf{u}$. Therefore $\mathbf{v}$ is a unit vector (see Figure 6).

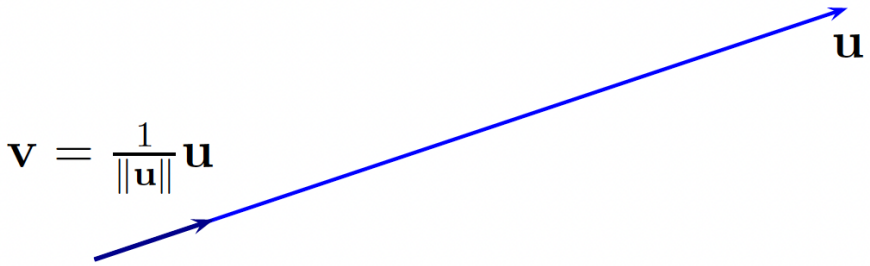


Figure 6: Normalizing a non-zero vector.

The process of taking a non-zero vector $\mathbf{u}$ and creating the new vector $\frac{1}{\|\mathbf{u}\|}\mathbf{u}$ is sometimes called **normalization** of $\mathbf{u}$.

Example 5. *Let* $\mathbf{u} = (2, 3, 6)$. *Compute* $\|\mathbf{u}\|$ *and find the unit vector* $\mathbf{v}$ *in the same direction as* $\mathbf{u}$.

Solution. By definition:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{2^2 + 3^2 + 6^2} = \sqrt{49} = 7.$$

Then the unit vector that is in the same direction as $\mathbf{u}$ is:

$$\mathbf{v} = \frac{1}{\|\mathbf{u}\|}\mathbf{u} = \frac{1}{7}\begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{3}{7} \\ \frac{6}{7} \end{bmatrix}.$$

You can easily verify that $\|\mathbf{v}\| = 1$. Thus, $\mathbf{v}$ is indeed a unit vector.

Definition 4. *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be vectors in* $\mathbb{R}^n$. *The distance between* $\mathbf{u}$ *and* $\mathbf{v}$ *is the length of the vector* $\mathbf{u} - \mathbf{v}$. *In other words:*

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

where $d(\mathbf{u}, \mathbf{v})$ *denotes the distance between* $\mathbf{u}$ *and* $\mathbf{v}$.

Example 6. *Find the distance between* $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} 7 \\ -9 \end{bmatrix}$.

Solution. We compute:

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \left\| \begin{bmatrix} 3 \\ -2 \end{bmatrix} - \begin{bmatrix} 7 \\ -9 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 3 - 7 \\ -2 - (-9) \end{bmatrix} \right\| = \left\| \begin{bmatrix} -4 \\ 7 \end{bmatrix} \right\|.$$

Now calculate the norm:

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(-4)^2 + 7^2} = \sqrt{16 + 49} = \sqrt{65}.$$

Thus, the distance between $\mathbf{u}$ and $\mathbf{v}$ is $\sqrt{65}$.

- In the context of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$, orthogonality is synonymous with perpendicularity. Below is the general definition.

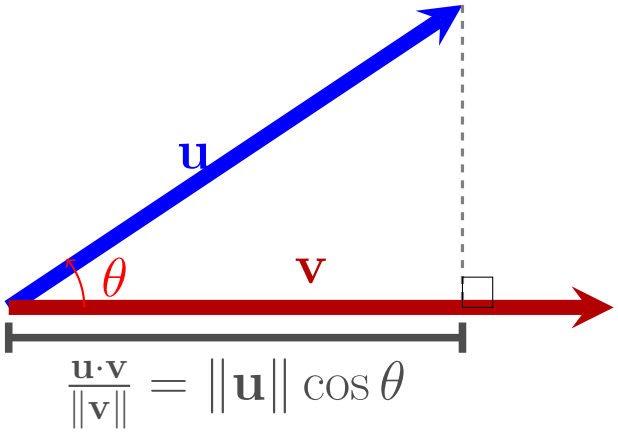
Definition 5. *Two vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^n$ *are said to be* ***orthogonal*** *if* $\mathbf{u} \cdot \mathbf{v} = 0$.

- In $\mathbb{R}^2$ and $\mathbb{R}^3$, the notion of orthogonality should be familiar to you. In fact, using

the Law of Cosines in $\mathbb{R}^2$ or $\mathbb{R}^3$, one can prove that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos(\theta) \tag{3}$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. If $\theta = \frac{\pi}{2}$, then clearly $\mathbf{u} \cdot \mathbf{v} = 0$.



- In higher dimensions, i.e., $n \geq 4$, we can use equation (3) to *define* the angle between vectors $\mathbf{u}$ and $\mathbf{v}$. In other words, the **angle** between any two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is defined to be:

$$\theta = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}\right).$$

The general notion of orthogonality in $\mathbb{R}^n$ leads to the following theorem.

Theorem 2. ***(Pythagorean Theorem):*** *Two vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are orthogonal if and only if* $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Proof. Since $\|\mathbf{u}+\mathbf{v}\| = \sqrt{(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})}$, we get

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v} \\ &= \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2.\end{aligned}$$

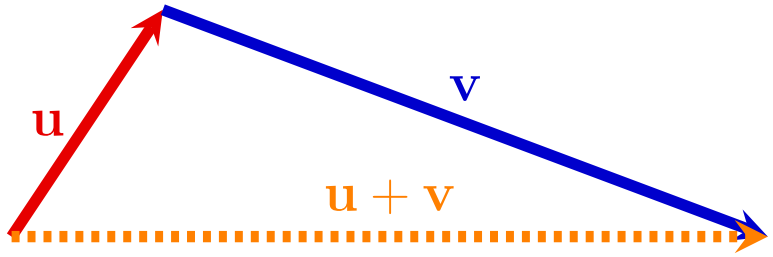
Therefore, $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if and only if $\mathbf{u}\cdot\mathbf{v} = 0$.

Useful inequalities:

Any two n-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$ satisfy the

- **Triangular inequality**: $\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$
- **Cauchy–Schwarz inequality**: $|\mathbf{u}\cdot\mathbf{v}| \le \|\mathbf{u}\|\cdot\|\mathbf{v}\|$

Euclidean Distance



$$\|\mathbf{v}+\mathbf{u}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

4 Orthogonal Sets

Definition 6. *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ *is said to be an* ***orthogonal set*** *if any pair of distinct vectors* $\mathbf{u}_i, \mathbf{u}_j$ *are orthogonal, that is,* $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ *whenever* $i \neq j$.

Example 7. *Is the set* $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ *an orthogonal set?*

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}.$$

Solution. Compute:

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{u}_2 &= (1)(0) + (-2)(1) + (1)(2) = 0, \\ \mathbf{u}_1 \cdot \mathbf{u}_3 &= (1)(-5) + (-2)(-2) + (1)(1) = 0, \\ \mathbf{u}_2 \cdot \mathbf{u}_3 &= (0)(-5) + (1)(-2) + (2)(1) = 0. \end{aligned}$$

Therefore, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set.

Definition 7. *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\} \in \mathbb{R}^n$ *is said to be* ***linearly independent*** *if no one vector can be represented as linear combination of the rmaining vectors.*

- This immediately implies that $p \leq n$, as otherwise the vectors cannot be linearly independent.

Theorem 3. *Let* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ *be an orthogonal set of non-zero vectors in* $\mathbb{R}^n$*. Then the set* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ *is linearly independent. In particular, if* $p = n$*, then the set* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ *is a* ***basis*** *for* $\mathbb{R}^n$*.*

Proof. Suppose that there are scalars $c_1, c_2, \ldots, c_p$ such that:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p = \mathbf{0}.$$

Take the inner product of $\mathbf{u}_1$ with both sides of the above equation:

$$c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) + c_2(\mathbf{u}_2 \cdot \mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p \cdot \mathbf{u}_1) = \mathbf{0} \cdot \mathbf{u}_1.$$

Since the set is orthogonal, the left-hand side of the last equation simplifies to $c_1(\mathbf{u}_1 \cdot \mathbf{u}_1)$. The right-hand side simplifies to 0. Hence,

$$c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) = 0.$$

But $\mathbf{u}_1 \cdot \mathbf{u}_1 = \|\mathbf{u}_1\|^2 > 0$ is not zero, and therefore the only way that $c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) = 0$ is if $c_1 = 0$.

Repeat the above steps using $\mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_p$ and conclude that $c_2 = \ldots = c_p = 0$. Therefore, $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is linearly independent.

- If $p = n$, then the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is automatically a basis for $\mathbb{R}^n$.

Example 8. *Is the set* $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ *a basis for* $\mathbb{R}^3$*?*

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}.$$

Solution. In Example 7 we showed $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. By Theorem 3, the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is linearly independent. To verify linear independence, we computed that:

$$\det([\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3]) = 30,$$

which is non-zero.

Now consider $\mathbb{R}^3$ (i.e., $n = 3$) and a set of 2 vectors:

Example 9 $(p \leq n)$**.**

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}$$

Here, $p = 2$ *and* $n = 3$*. This set is linearly independent because no vector in the set can be written as a linear combination of the other. There is no non-trivial solution to the equation*

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 = 0$$

other than $c_1 = c_2 = 0$*.*

Definition 8. *A set of vectors* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ *is said to be an* ***orthonormal set*** *if it is an orthogonal set and if each vector* $\mathbf{u}_i$ *in the set is a unit vector.*

Example 10. *Consider the previous orthogonal set in* $\mathbb{R}^3$*:*

$$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

It is not an orthonormal? If not, make it an orthonormal set.

Solution:

It is not an orthonormal set because none of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are unit vectors. Explicitly,

$$\|\mathbf{u}_1\| = \sqrt{6}, \quad \|\mathbf{u}_2\| = \sqrt{5}, \quad \|\mathbf{u}_3\| = \sqrt{30}.$$

However, from an orthogonal set, we can create an orthonormal set by normalizing each vector to get the orthonormal set

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{\frac{\mathbf{u}_1}{\sqrt{6}}, \frac{\mathbf{u}_2}{\sqrt{5}}, \frac{\mathbf{u}_3}{\sqrt{30}}\}$$

5 Gram–Schmidt Orthogonalization

Theorem 4 (Gram–Schmidt Orthogonalization Algorithm). *Let $\mathcal{U}$ be an m-dimensional subspace of $\mathbb{R}^n$. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, construct $\mathbf{u}_1, \ldots, \mathbf{u}_m$ successively as below:*

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{u}_1, \mathbf{v}_2 \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1, \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{u}_1, \mathbf{v}_3 \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\langle \mathbf{u}_2, \mathbf{v}_3 \rangle}{\|\mathbf{u}_2\|^2}\mathbf{u}_2, \\
&\vdots \\
\mathbf{u}_m &= \mathbf{v}_m - \frac{\langle \mathbf{u}_1, \mathbf{v}_m \rangle}{\|\mathbf{u}_1\|^2}\mathbf{u}_1 - \frac{\langle \mathbf{u}_2, \mathbf{v}_m \rangle}{\|\mathbf{u}_2\|^2}\mathbf{u}_2 - \cdots - \frac{\langle \mathbf{u}_{m-1}, \mathbf{v}_m \rangle}{\|\mathbf{u}_{m-1}\|^2}\mathbf{u}_{m-1}.
\end{aligned}$$

Then:

1. *$\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is an orthogonal basis of $\mathcal{U}$, and*
2. *$\text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for each $k = 1, 2, \ldots, m$.*

• Geometric Interpretation of Gram–Schmidt Orthogonalization

Definition 9. *The* ***vector projection*** *of a vector* $\mathbf{v}$ *on a nonzero vector* $\mathbf{u}$ *is defined as*

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ *denotes the inner product of the vectors* $\mathbf{u}$ *and* $\mathbf{v}$.

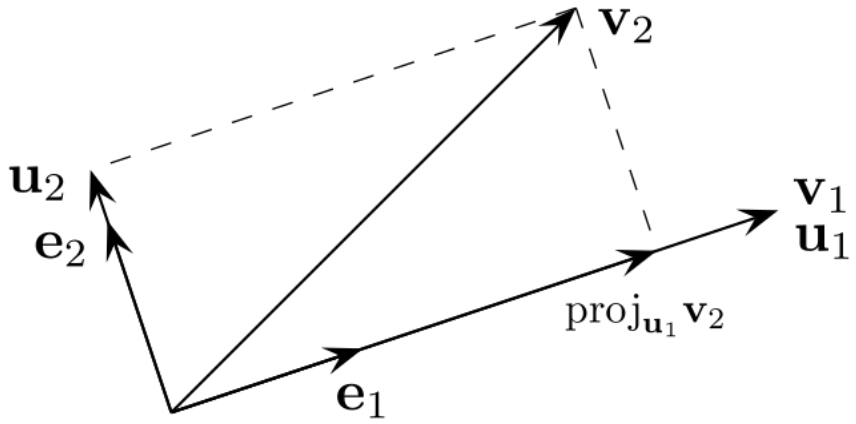


Figure 7: The first two steps of the Gram–Schmidt process. [source: Wikipedia]

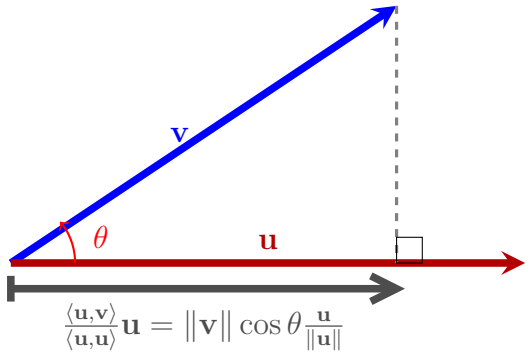


Figure 8: Vector projection of $\mathbf{v}$ over $\mathbf{u}$ in terms of inner products. We usually define $\mathbf{e} \triangleq \frac{\mathbf{u}}{\|\mathbf{u}\|}$.

In light of the above definition, it can bee sen that

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, \\
\mathbf{u}_2 &= \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2), \\
\mathbf{u}_3 &= \mathbf{v}_3 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3), \\
&\vdots \\
\mathbf{u}_m &= \mathbf{v}_m - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_m) - \text{proj}_{\mathbf{u}_2}(\mathbf{v}_m) - \cdots - \text{proj}_{\mathbf{u}_{m-1}}(\mathbf{v}_m).
\end{aligned}$$

At the end of the Gram-Schmidt process, we obtain an **orthogonal** set of vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$.

We can normalize them to get an **orthonormal** set.

$$\mathbf{e}_1 \triangleq \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$
$$\mathbf{e}_2 \triangleq \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$$
$$\vdots$$
$$\mathbf{e}_m \triangleq \frac{\mathbf{u}_m}{\|\mathbf{u}_m\|}$$

Example 11. ***Orthogonalizing Two Vectors in*** $\mathbb{R}^2$

We are given two linearly independent vectors in $\mathbb{R}^2$*:*

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Apply the Gram-Schmidt process to orthogonalize them.

1.

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

2. Find the project of $\mathbf{v}_2$ with respect to $\mathbf{u}_1$:

$$\text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) = \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 = \frac{\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$$

Therefore, $\mathbf{v}_2$ is already orthogonal to $\mathbf{v}_1$. Thus,

$$\mathbf{u}_2 = \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Example 12. ***Orthogonalizing Three Vectors in*** $\mathbb{R}^3$

We are given three linearly independent vectors in $\mathbb{R}^3$*:*

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Apply the Gram-Schmidt process to generate an orthonormal basis.

1.

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

2. Find the project of $\mathbf{v}_2$ with respect to $\mathbf{u}_1$, and then evaluate $\mathbf{u}_2$:

$$\text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) = \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 = \frac{\begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) = \begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix}$$

3. Find the project of $\mathbf{v}_3$ with respect to $\mathbf{u}_1$ and $\mathbf{u}_2$, and then evaluate $\mathbf{u}_3$:

$$\text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) = \frac{\begin{bmatrix}0 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix}}{2} \cdot \begin{bmatrix}1\\1\\0\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\1\\0\end{bmatrix}$$

$$\text{proj}_{\mathbf{u}_2}(\mathbf{v}_3) = \frac{\begin{bmatrix}0 & 1 & 1\end{bmatrix}\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix}}{\frac{3}{2}} \cdot \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix}$$

Subtracting the projections gives:

$$\mathbf{u}_3 = \mathbf{v}_3 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3) = \begin{bmatrix}0\\1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} - \frac{1}{3}\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{bmatrix} = \frac{2}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix}$$

6 Linear Vector Space and Subspace

Definition 10. *A linear space or vector space is a set of vectors together with two operators, addition of vectors and multiplication by a scalar.*

Example 13. ***Ordinary Euclidean Space.*** *The most familiar example of a linear space is the ordinary (real) Euclidean space.*

- In n-dimensional Euclidean space, a vector is specified by its n coordinates:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

- To save space, we usually represent $\mathbf{v}$ as: $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T$ where the superscript $(\cdot)^T$ denotes a **transpose**.

6.1 Euclidean Space

- There are two main operations defined in Euclidean space:
  1. **Vector addition:** Add corresponding components of two vectors.

$$\mathbf{u}+\mathbf{v}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}+\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=\begin{bmatrix}u_1+v_1\\u_2+v_2\\\vdots\\u_n+v_n\end{bmatrix}$$

  2. **Scalar multiplication:** Multiply each component of the vector by a scalar.

$$c\cdot\mathbf{v}=c\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=\begin{bmatrix}cv_1\\cv_2\\\vdots\\cv_n\end{bmatrix}$$

6.2 Euclidean Space of Complex-Valued Vectors

A space of particular importance in this course is the **Euclidean space of complex-valued vectors**.

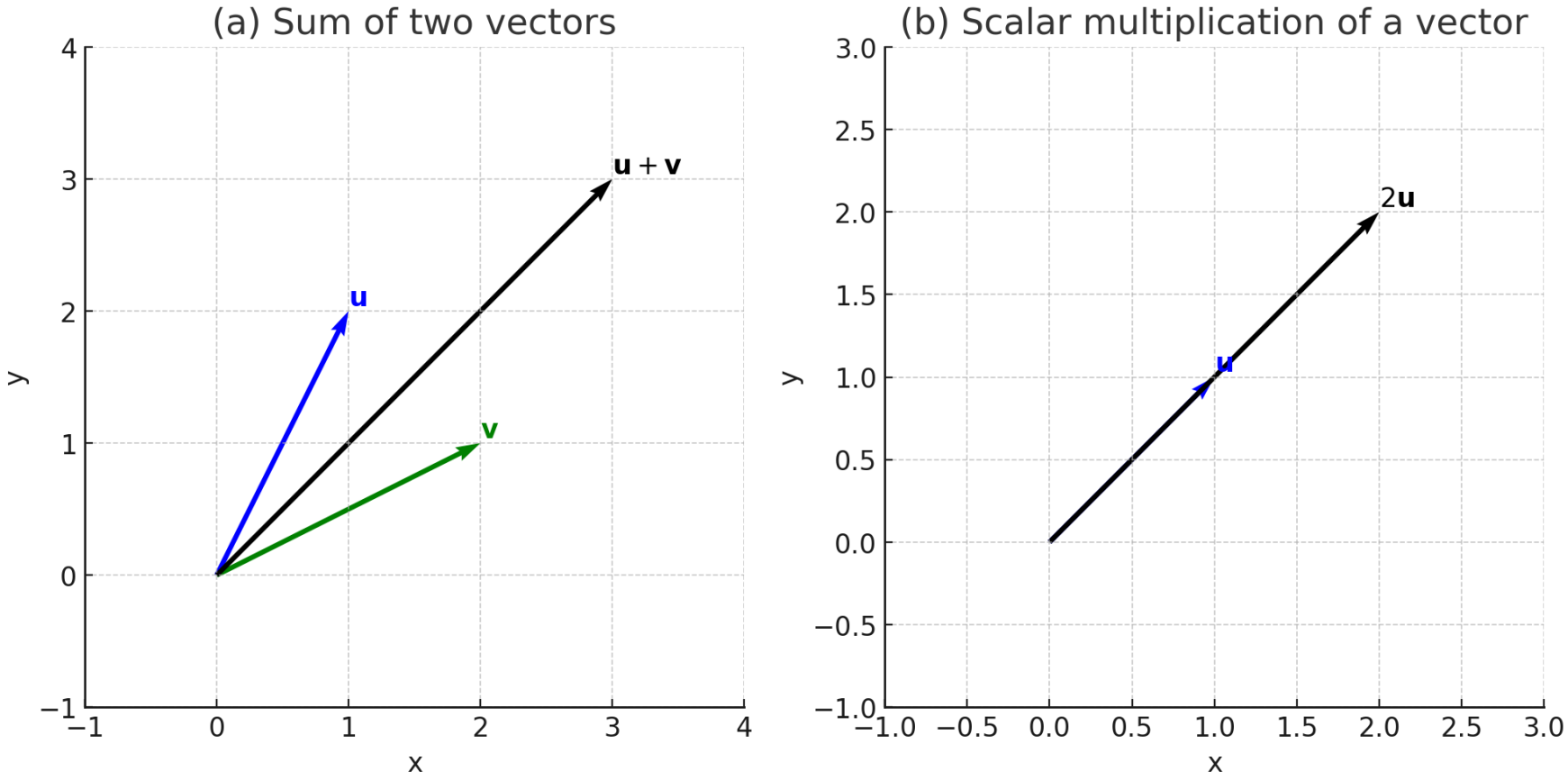


Figure 9: Elementary operations in a two-dimensional linear space. (a) Sum of two vectors. (b) Multiplication of a vector by a scalar.

- Vectors in this space are represented similarly to

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T,$$

  except that the components v_k are **complex valued**.

- Ordinary (real) Euclidean space is a special case of this, where the **imaginary parts** of the vector components are **zero**.

6.3 Subspaces of a Linear Space

Definition 11. *A **subspace** of a linear space is a subset that is itself a linear space. This means:*

- *The sum of any two vectors in the subspace is also in the subspace.*
- *The product of any vector in the subspace by any scalar is also in the subspace.*

Examples:

1. In three-dimensional Euclidean space:

- A **line** or a **plane** through the origin is a subspace.
- The vector **0** must be in the subspace.

2. A more general subspace is the set of vectors obtained by forming all possible weighted linear combinations of n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. The subspace so formed is said to be **spanned** by the set of n vectors.

6.4 Span of a Set of Vectors

Definition 12. *The* ***span*** *of a set of vectors* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *in a vector space is the set of all possible* ***linear combinations*** *of these vectors:*

$$span(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}) = \{\mathbf{w} \mid \mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n, \quad c_1, \ldots, c_n \in \mathbb{F}\}$$

Key Points:

1. The span is a **subspace** of the vector space.
2. It is the smallest subspace containing all the vectors in the set.

Gram-Schmidt Orthogonalization

- **Examples:**
  - **Single Vector:** The span of a non-zero vector $\mathbf{v}$ is a **line** through the origin:
$$\text{span}(\{\mathbf{v}\}) = \{c\mathbf{v} \mid c \in \mathbb{F}\}.$$
  - **Two Vectors:** If $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, their span is a **plane** through the origin.
  - **Three Vectors:** If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent in $\mathbb{R}^3$, their span is the entire space $\mathbb{R}^3$.
- **Geometric Interpretation:**
  - The span represents all points that can be reached by scaling and adding the vectors.
  - Linearly dependent vectors form a span confined to a lower-dimensional space (e.g., a line or plane).

Example 14. ***Span of Two 2D Vectors.*** *What is the span of* $\{\mathbf{v}_1, \mathbf{v}_2\}$ *when* $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *and* $\mathbf{v}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

Solution. The span of $\{\mathbf{v}_1, \mathbf{v}_2\}$ is the set of all possible linear combinations:

$$\text{span}(\{\mathbf{v}_1, \mathbf{v}_2\}) = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R}\}.$$

Substitute $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$\mathbf{w} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} c_1 + 3c_2 \\ 2c_1 + 4c_2 \end{bmatrix}.$$

Observations:

- If $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent (which they are in this case), their span forms a **plane** in 2D space, $\mathbb{R}^2$.
- Any vector in the plane can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Example 15. *Suppose* $W = span\{\mathbf{x}_1, \mathbf{x}_2\}$ *where*

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Find an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ *for* W.

Solution.

- $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$
- $\mathbf{v}_2 = \mathbf{x}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{x}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} - \frac{4}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$

Example 16. *Suppose* $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \in \mathbb{R}^4$ *where*

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

is a basis for a subspace W *of* $\mathbb{R}^4$*. Find an orthogonal basis for* W*.*

Solution.

- $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$
- $\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} - \frac{5}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 9 \\ 18 \\ -15 \\ 0 \end{bmatrix}$

  to make it easier to work, we can replace $\mathbf{v}_2$ with $14\mathbf{v}_2$

Gram-Schmidt Orthogonalization

- $\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{14}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} - \frac{9}{630}\begin{bmatrix} 9 \\ 18 \\ -15 \\ 0 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4 \\ -2 \\ 0 \\ 5 \end{bmatrix}$

7 Inner Product for Signals

For other linear spaces, similar definitions of the inner product and norm in Euclidean space can be extended.

- Two cases of interest:
  - Discrete-time signals
  - Continuous-time signals
- **Inner Product and Norm for Discrete-Time Signals:**
  For discrete-time signals x_k and y_k, the inner product and norm are defined as:

$$\langle x_k, y_k \rangle = \sum_{k=-\infty}^{\infty} x_k y_k^*, \qquad \|x_k\|^2 = \sum_{k=-\infty}^{\infty} |x_k|^2$$

  which is a natural extension of the finite-dimensional case
- **Inner Product and Norm for Continuous-Time Signals:**
  For continuous-time signals $x(t)$ and $y(t)$, the summation generalizes to an inte-

gral:

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) y^*(t)\, dt, \qquad \|x(t)\|^2 = \int_{-\infty}^{\infty} |x(t)|^2\, dt$$

- **Correlation and Orthogonality in Continuous-Time Signals:**
  - In the context of continuous-time signals, the inner product:

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t) y^*(t)\, dt$$

    is also referred to as the **correlation** of $x(t)$ with $y(t)$.
  - It measures how closely the two signals resemble one another.
    - When $x(t) = y(t)$, the correlation reduces to the signal's **energy**:

$$\langle x(t), x(t) \rangle = \|x(t)\|^2 = \int_{-\infty}^{\infty} |x(t)|^2\, dt \triangleq \mathcal{E}_x \quad \Rightarrow \quad \|x(t)\| = \sqrt{\mathcal{E}_x}$$

    - When $\langle x(t), y(t) \rangle = 0$, the signals are said to be **orthogonal**.

- **Observations:**
  - The inner product allows projection and comparison of signals.
  - The norm provides a measure of signal energy for both discrete-time and continuous-time signals.
- Triangular inequality:

$$\|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\|$$

- Cauchy–Schwarz inequality

$$|\langle x_1(t), x_2(t)\rangle| \leq \|x_1(t)\| \cdot \|x_2(t)\| = \sqrt{\mathcal{E}_{x_1}\mathcal{E}_{x_2}}$$

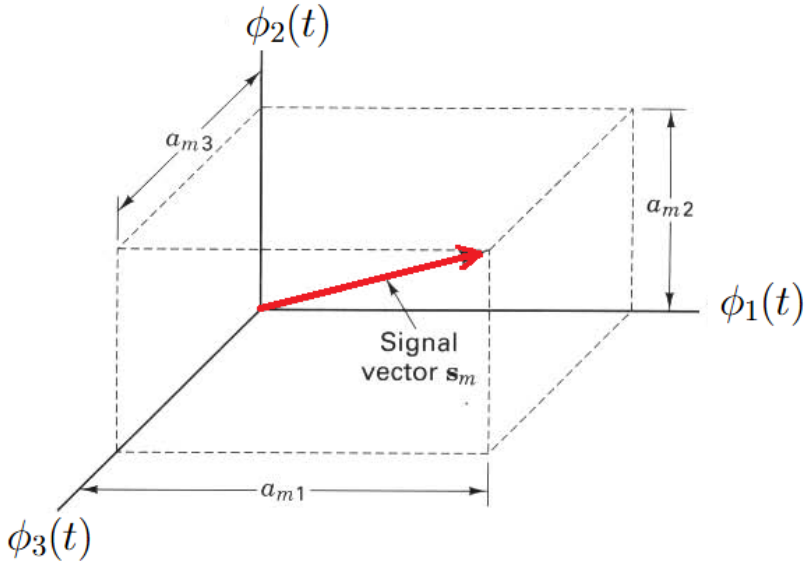
or equivalently,

$$\left| \int_{-\infty}^{\infty} x_1(t)x_2^*(t)\,dt \right| \leq \left| \int_{-\infty}^{\infty} |x_1(t)|^2\,dt \right|^{\frac{1}{2}} \cdot \left| \int_{-\infty}^{\infty} |x_2(t)|^2\,dt \right|^{\frac{1}{2}}$$

with the equality achieving when $x_1(t) = ax_2^*(t)$, and a is a complex number, in general.

7.1 Geometric (Vectorial) Representation of Signals

- In digital communications, signals can be represented geometrically as vectors in a multidimensional space, where the dimensions of the space correspond to the basis functions of the signal.
- We define an N-dimensional orthogonal space as a space characterized by a set of N linearly independent functions called basis functions.
- Then, the process of converting a time signal to a vector representation is called basis function expansion.

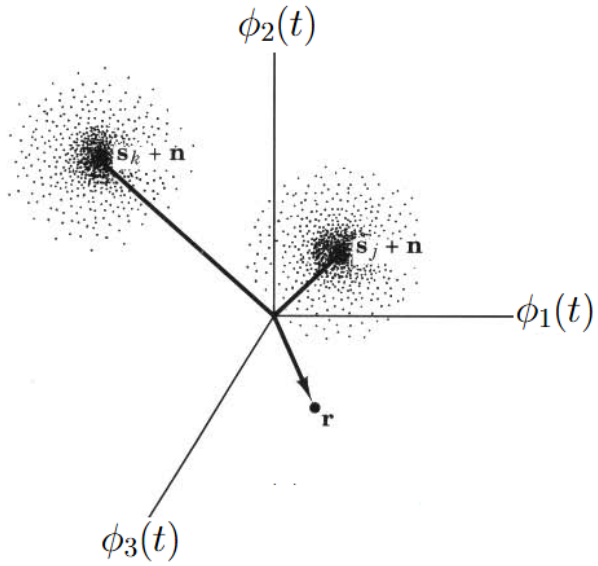


$$s_m(t) = a_{m1}\phi_1(t) + a_{m2}\phi_2(t) + a_{m3}\phi_3(t)$$

$$\mathbf{s}_m = \begin{bmatrix} a_{m1} \\ a_{m2} \\ a_{m3} \end{bmatrix}$$

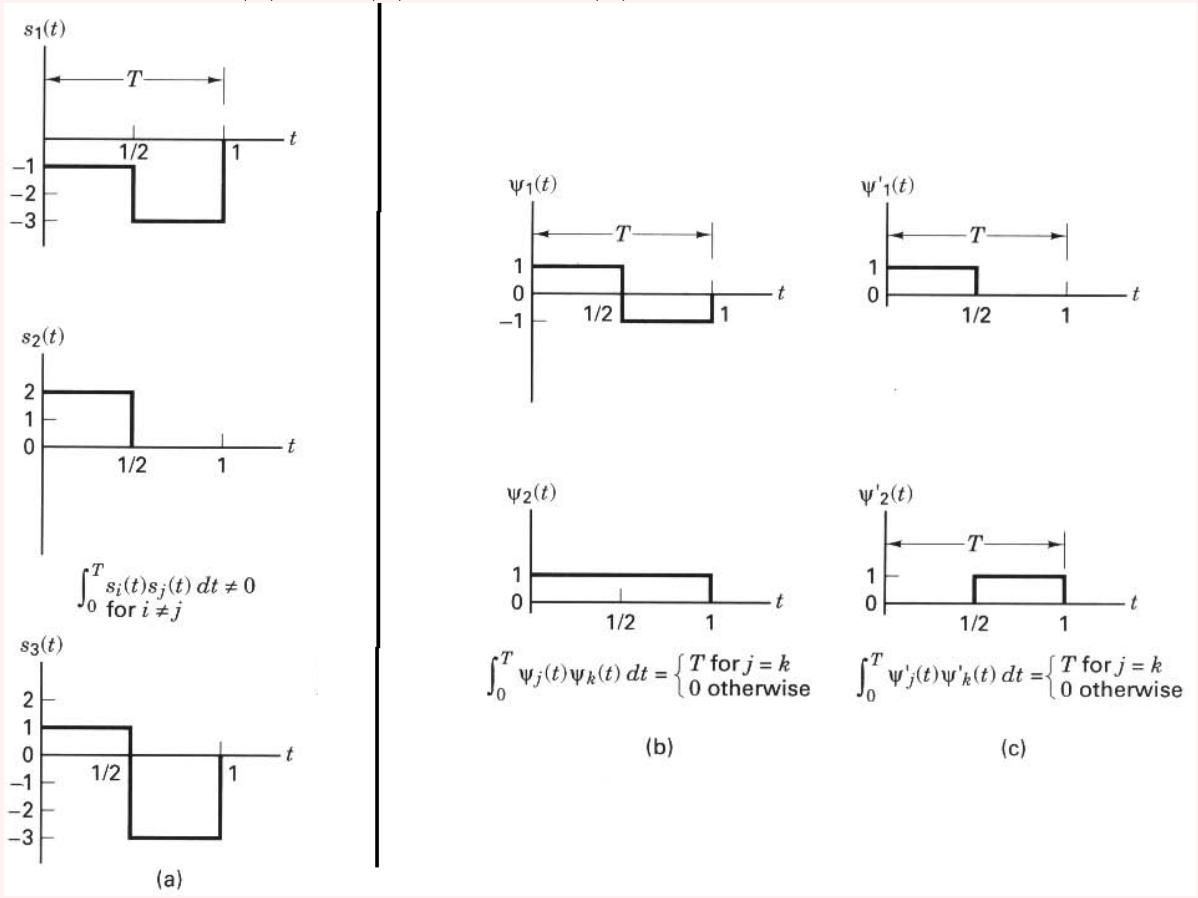
Orthogonal Expansion of Signals

- This representation allows us to analyze the relationship between different signals in terms of their distance and orientation, which are critical for understanding how signals are transmitted and detected in the presence of noise.



- Vectors $\mathbf{s}_j$ and $\mathbf{s}_k$, called signal vectors, represent two possible *reference signals* belonging to the set of M signal waveforms $\{s_m(t)\}$, and $\mathbf{n}$ is the noise vector.
- Vectors $\mathbf{r}$ represent a possible signal vector that may arrive at the receiver during some symbol interval. The task of the receiver is to decide if $\mathbf{r}$ has close resemblance to $\mathbf{s}_j$, $\mathbf{s}_k$, or another signal vector.

Example 17. *Three waveforms $s_1(t), s_2(t)$ and $s_3(t)$ are shown in figure below. 1. Show that they are not orthogonal. 2. Show that $\{\phi_1(t), \phi_2(t)\}$ and $\{\phi_1'(t), \phi_2'(t)\}$ each form an orthogonal set (basis functions). 3. Represents $s_1(t), s_2(t)$ and $s_3(t)$ based on the each of the above basis function.*



- **Solution:** Parts 1 and 2 are trivial.
  Part 3: Given the non-orthogonal set $\{s_i(t)\}$ and the orthogonal basis set $\{\psi_j(t)\}$:

$$s_1(t) = \psi_1(t) - 2\psi_2(t)$$
$$s_2(t) = \psi_1(t) + \psi_2(t)$$
$$s_3(t) = 2\psi_1(t) - \psi_2(t)$$

  For the orthogonal basis set $\{\psi'_j(t)\}$:

$$s_1(t) = -\psi'_1(t) - 3\psi'_2(t)$$
$$s_2(t) = 2\psi'_1(t)$$
$$s_3(t) = \psi'_1(t) - 3\psi'_2(t)$$

- How do we get the above expansion?

- What are the practical applications of being able to describe $s_1(t)$, $s_2(t)$, and $s_3(t)$ in terms of $\psi_1(t)$, $\psi_2(t)$, and the appropriate coefficients?
  - **Efficient Representation:** Instead of transmitting three distinct waveforms, the system can transmit the coefficients of the two basis functions.
  - **Simplified Implementation:** The transmitter and receiver only need to handle the two basis functions $\psi_1(t)$ and $\psi_2(t)$, reducing the overall complexity of the system.

Orthogonal Expansion of Signals

- Suppose there exists a set of **orthonormal** functions $\{\phi_n(t), n = 1, 2, \ldots, N\}$:

$$\langle \phi_n(t), \phi_m(t) \rangle = \int_{-\infty}^{\infty} \phi_n(t)\phi_m^*(t)\, dt = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

- Signal $s(t)$ can be approximated by a weighted linear combination of basis functions.

$$\hat{s}(t) = \sum_{k=1}^{N} s_k \phi_k(t)$$

  where

  - $\phi_k(t)$ are orthonormal basis functions,
  - s_k are the corresponding signal components.
  - $\hat{s}(t)$ is the approximation of $s(t)$

- The approximation error is given by:

$$e(t) = s(t) - \hat{s}(t)$$

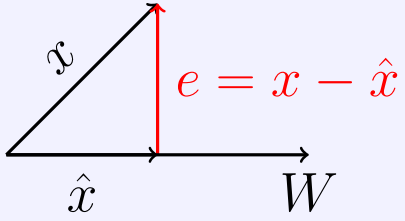
- Our goal is to select the coefficients of the basis functions such that the approximation error

$$\mathcal{E}_e = \int_{-\infty}^{\infty} |s(t) - \hat{s}(t)|^2 \, dt = \int_{-\infty}^{\infty} \left| s(t) - \sum_{k=1}^{N} s_k \phi_k(t) \right|^2 dt$$

  is minimized.

- To find s_k, differentiate $\mathcal{E}_e$ with respect to each coefficient, set it zero, and solve.

The orthogonality principle says that the error vector of the optimal estimator (in a mean square error sense) is orthogonal to any possible estimator.



Geometric representation of orthogonality where W is a one-dimensional subspace.

- Alternatively, we may use *orthogonality principle from estimation theory* based on the mean squared error which states that for optimum estimation, the estima-

tion error should be orthogonal to every basis function:

$$\int_{-\infty}^{\infty} \left[s(t) - \sum_{k=1}^{N} s_k \phi_k(t) \right] \phi_n^*(t)\, dt = 0, \quad n = 1, 2, \ldots, N$$

- Since the functions $\{\phi_n(t)\}$ are orthonormal the above reduces to

$$s_n = \langle s(t), \phi_n(t) \rangle = \int_{-\infty}^{\infty} s(t) \phi_n^*(t)\, dt \qquad n = 1, 2, \ldots, N$$

Thus, the coefficients are obtained by projecting the signal $s(t)$ onto each of the functions $\{\phi_n(t)\}$.

- Consequently, $\hat{s}(t)$ is the projection of $s(t)$ onto the N-dimensional signal space spanned by the function $\{\phi_n(t)\}$. Therefore, $\hat{s}(t)$ is orthogonal to $e(t)$ i.e.,

$$\langle e(t), \hat{s}(t) \rangle = 0.$$

- The minimum mean-squared error (MMSE) is simplified as (why?)

$$\mathcal{E}_{\min} = \int_{-\infty}^{\infty} e(t) s^*(t)\, dt = \int_{-\infty}^{\infty} |s(t)|^2\, dt - \sum_{k=1}^{N} |s_k|^2$$

- Thus, when $\mathcal{E}_{\min} = 0$, we have

$$\mathcal{E}_s = \int_{-\infty}^{\infty} |s(t)|^2 = \sum_{k=1}^{N} |s_k|^2$$

- Under the above condition, we may express $s(t)$ as

$$s(t) = \sum_{k=1}^{N} s_k \phi_k(t)$$

When every finite energy signal can be represented by a series expansion of the form

$$s(t) = \sum_{n=1}^{N} s_n \phi_k(t)$$

with $\mathcal{E}_{\min} = 0$, the set of orthonormal functions $\{\phi_n(t)\}$ is said to be complete.

Example 18. ***Trigonometric Fourier Series:*** *Consider a finite energy real signal $s(t)$ that is zero everywhere except in the range $0 \leq t \leq T$ and has a finite number of discontinuities in this interval. Its periodic extension can be represented in a Fourier series as*

$$s(t) = \sum_{k=0}^{\infty} \left(a_k \cos\left(\frac{2\pi kt}{T}\right) + b_k \sin\left(\frac{2\pi kt}{T}\right) \right)$$

where the coefficients $\{a_k, b_k\}$ that minimize the mean square error are given by:

$$a_0 = \frac{1}{T}\int_0^T s(t)\,dt$$

$$a_k = \frac{2}{T}\int_0^T s(t)\cos\left(\frac{2\pi kt}{T}\right)\,dt, \quad k = 1, 2, 3, \ldots$$

$$b_k = \frac{2}{T}\int_0^T s(t)\sin\left(\frac{2\pi kt}{T}\right)\,dt, \quad k = 1, 2, 3, \ldots$$

The set of functions $\left\{ \frac{1}{\sqrt{T}}, \sqrt{\frac{2}{T}}\cos\left(\frac{2\pi kt}{T}\right), \sqrt{\frac{2}{T}}\sin\left(\frac{2\pi kt}{T}\right) \right\}$ is a complete set for the expansion of periodic signals on the interval $[0, T]$ and hence the series expansion results in zero mean square error.

Example 19. ***Exponential Fourier Series:*** *Consider a general finite energy signal $s(t)$ (real or complex) that is zero everywhere except in the range $0 \leq t \leq T$ and has a finite number of discontinuities in this interval. Its periodic extension can be represented in an exponential Fourier series as*

$$s(t) = \sum_{n=-\infty}^{\infty} x_n e^{j2\pi \frac{n}{T} t} \tag{4}$$

where the coefficients $\{x_n\}$ that minimize the mean square error are given by

$$x_n = \frac{1}{T} \int_{-\infty}^{\infty} x(t) e^{-j2\pi \frac{n}{T} t} dt \tag{5}$$

The set of functions $\left\{ \sqrt{\frac{1}{T}} e^{j2\pi \frac{n}{T} t} \right\}$ is a complete set for expansion of periodic signals on the interval $[0, T]$, and, hence, the series expansion results in zero mean square error.

7.2 Gram-Schmidt Orthogonalization for Signals

- Similar to vectors, the Gram-Schmidt process can be used to convert a set of signals into an orthonormal set of basis functions.
- Given a set of finite-energy signal waveforms $s_1(t), s_2(t), \ldots, s_m(t)$, the Gram-Schmidt process generates an orthonormal set of functions $\phi_1(t), \ldots, \phi_n(t)$, and $n \leq m$ through the following steps:

Gram-Schmidt Orthogonalization for Signals

1. $\gamma_1(t) = s_1(t) \implies \qquad \phi_1(t) = \frac{\gamma_1(t)}{\|\gamma_1(t)\|}$
2. For $i = 2, 3, \ldots, m$

$$\gamma_i(t) = s_i(t) - \sum_{j=1}^{i-1} \langle s_i(t), \phi_j(t) \rangle \phi_j(t), \implies \qquad \phi_i(t) = \frac{\gamma_i(t)}{\|\gamma_i(t)\|}$$

- This process continues until m signal sets are exhausted and $n \leq m$ orthonormal basis are constructed.

Gram-Schmidt Orthogonalization for Signals

- After constructing the set of orthonormal waveforms $\{\phi_n(t)\}$, we can express the M signals $\{s_m(t)\}$ as linear combinations of the $\{\phi_n(t)\}$. Thus, we may write:

$$s_m(t) = \sum_{n=1}^{N} s_{mn}\phi_n(t), \quad m = 1, 2, \ldots, M \tag{6}$$

- Based on the above expression, each signal may be represented by the vector:

$$\mathbf{s}_m = \begin{bmatrix} s_{m1} & s_{m2} & \cdots & s_{mN} \end{bmatrix}^T$$

- Equivalently, each signal can be represented as a point in the N-dimensional (in general, complex) signal space with coordinates $\{s_{mn},\ n = 1, 2, \ldots, N\}$.

- Therefore, a set of M signals $\{s_m(t)\}_{m=1}^{M}$ can be represented by a set of M vectors $\{\mathbf{s}_m\}_{m=1}^{M}$ in the N-dimensional space, where $N \leq M$.

  - ◇ The corresponding set of vectors is called the **signal space representation**, or **constellation**, of $\{s_m(t)\}_{m=1}^{M}$.
  - ◇ If the original signals are real, then the corresponding vector representations are in $\mathbb{R}^N$; and if the signals are complex, then the vector representations are in $\mathbb{C}^N$.

- From the orthonormality of the basis $\{\phi_n(t)\}$, it follows that

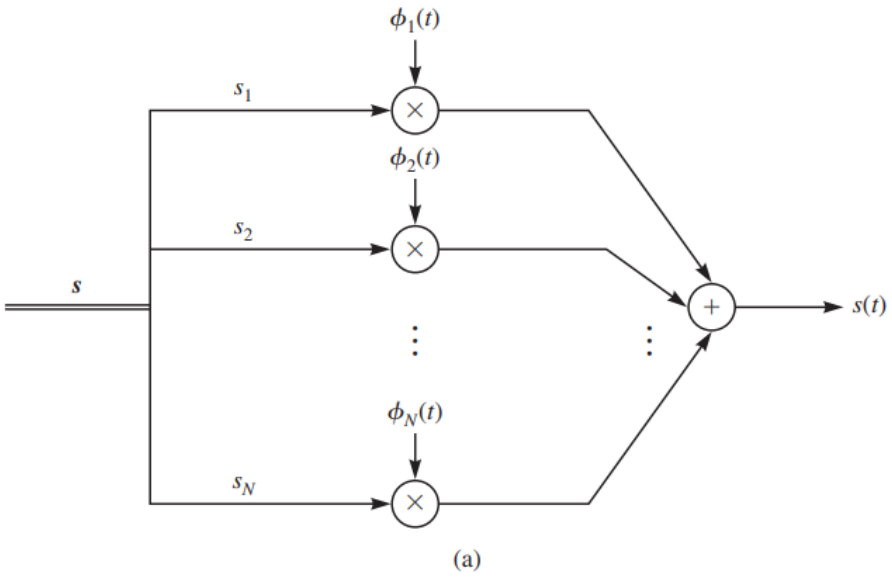
$$\mathcal{E}_m = \int_{-\infty}^{\infty} |s_m(t)|^2 \, dt = \sum_{n=1}^{N} |s_{mn}|^2 = \|\mathbf{s}_m\|^2 \tag{7}$$

- Also,

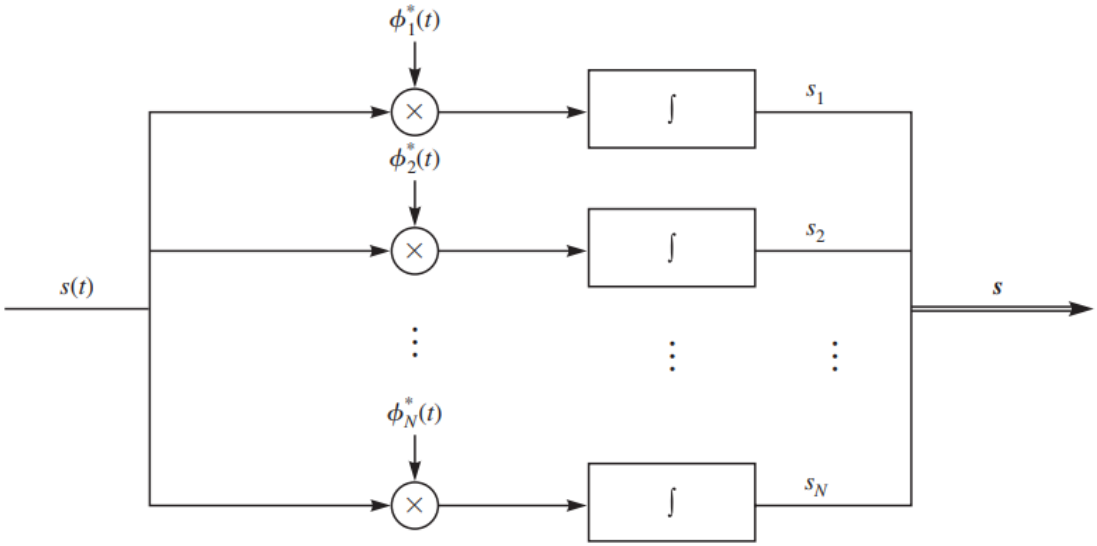
$$\begin{aligned} d_{12} &= \text{Euclidean distance between } s_1(t) \text{ and } s_2(t) \\ &= \|\mathbf{s}_1 - \mathbf{s}_2\| \\ &= \sqrt{\sum_{i=1}^{N} (s_{1i} - s_{2i})^2} \\ &= \sqrt{\int_0^T |s_1(t) - s_2(t)|^2 \, dt} \end{aligned}$$

- The process of mapping a vector to a signal and a signal to a vector is illustrated in Figures (a) and (b) on the next page, respectively.

Gram-Schmidt Orthogonalization for Signals



(a)

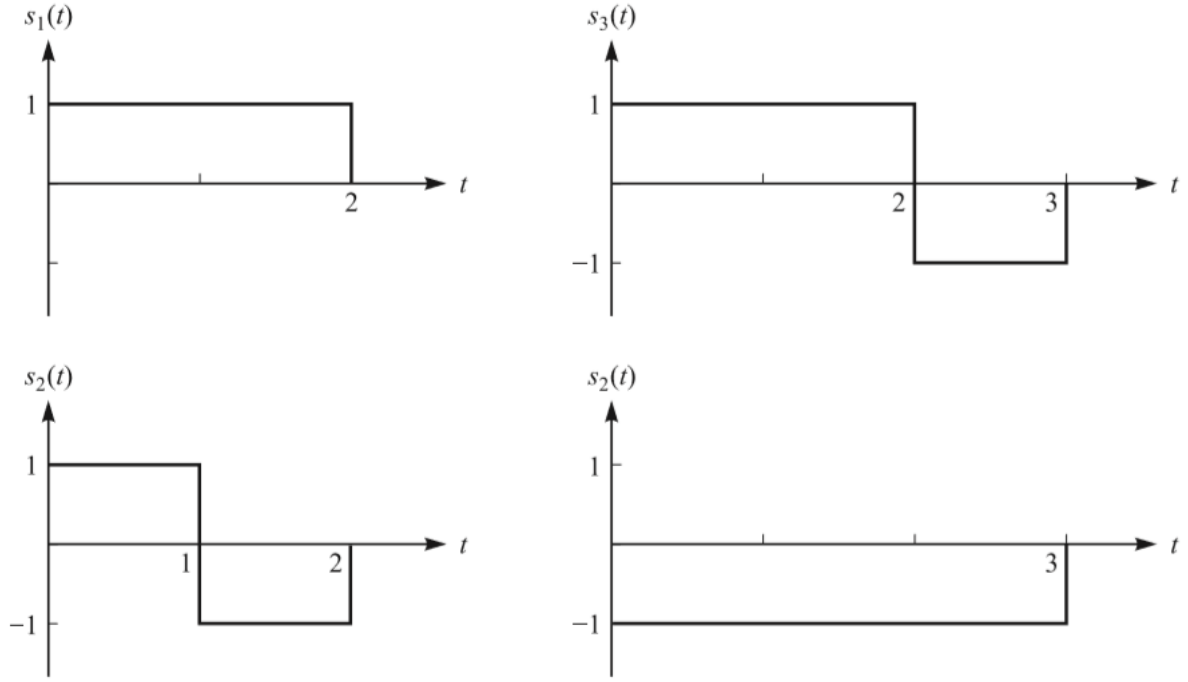


(b)

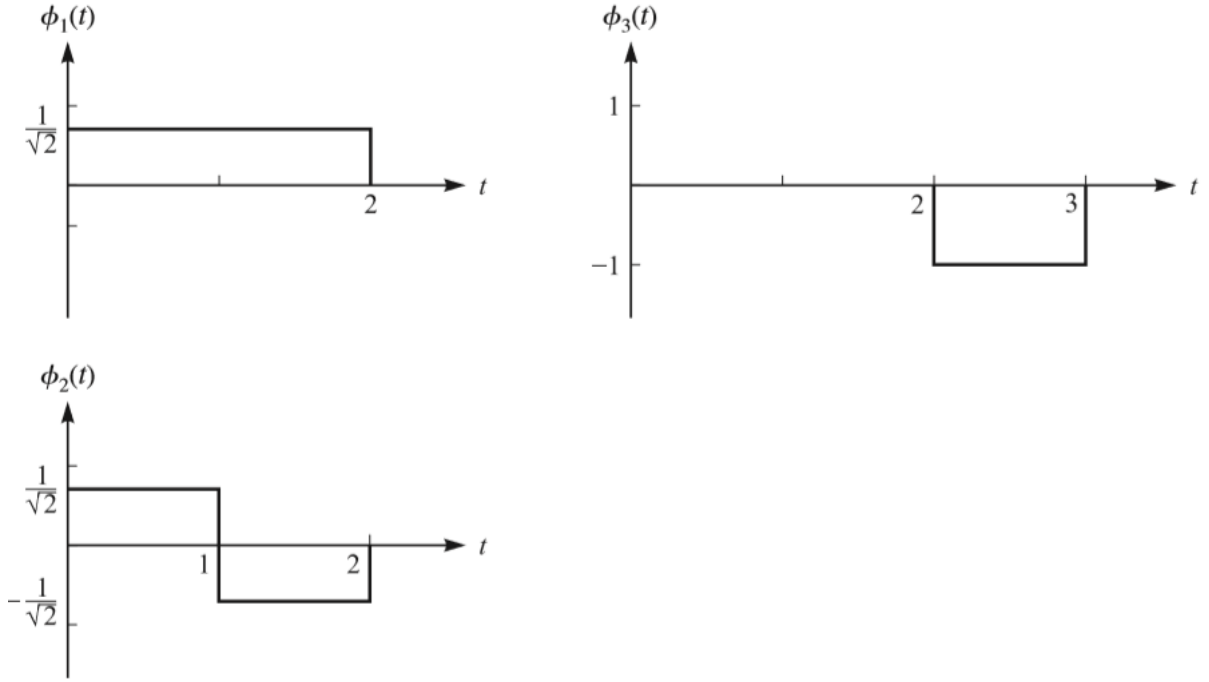
Figure (a): mapping a vector to a signal

Figure (b): mapping a signal to a vector

Example 20. *Apply the Gram-Schmidt procedure to the set of four waveforms illustrated in the figure below.*



Solutions:



- Note that the number of signals is $M = 4$ whereas the number of basis is $N = 3$

Gram-Schmidt Orthogonalization for Signals

Example 21. *For the signal waveforms in the previous example*

- *Express each signal as a linear combination of basis functions.*
- *Explain what happens to the basis if the order of projection changes?*
- *Introduce a new basis and find the signal constellation.*

Answers:

$$
\begin{aligned}
s_1(t) &= \sqrt{2}\phi_1(t) + 0\cdot\phi_2(t) + 0\cdot\phi_3(t) &&\Longrightarrow \mathbf{s}_1 = \begin{bmatrix}\sqrt{2} & 0 & 0\end{bmatrix}^T \\
s_2(t) &= 0\cdot\phi_1(t) + \sqrt{2}\phi_2(t) + 0\cdot\phi_3(t) &&\Longrightarrow \mathbf{s}_2 = \begin{bmatrix}0 & \sqrt{2} & 0\end{bmatrix}^T \\
s_3(t) &= \sqrt{2}\phi_1(t) + 0\cdot\phi_2(t) + \phi_3(t) &&\Longrightarrow \mathbf{s}_3 = \begin{bmatrix}\sqrt{2} & 0 & 1\end{bmatrix}^T \\
s_4(t) &= -\sqrt{2}\phi_1(t) + 0\cdot\phi_2(t) + 1\cdot\phi_3(t) &&\Longrightarrow \mathbf{s}_4 = \begin{bmatrix}-\sqrt{2} & 0 & 1\end{bmatrix}^T
\end{aligned}
$$

These vectors are called `signal space representations` or `constellations` of the above set of signals.

Solution:

- $\phi_1(t) = \frac{s_1(t)}{\|s_1(t)\|} = \frac{s_1(t)}{\sqrt{2}}$
- $$\gamma_2(t) = s_2(t) - \langle s_2(t), \phi_1(t)\rangle \phi_1(t)$$
$$= s_2(t) - \left(\int_0^3 s_2(t)\phi_1^*(t)\, dt \right) \phi_1(t) = s_2(t) \implies \phi_2(t) = \frac{\gamma_2(t)}{\|\gamma_2(t)\|} = \frac{s_2(t)}{\sqrt{2}}.$$
- $$\gamma_3(t) = s_3(t) - \langle s_3(t), \phi_1(t)\rangle \phi_1(t) - \langle s_3(t), \phi_2(t)\rangle \phi_2(t)$$
$$= s_3(t) - \sqrt{2}\phi_1(t) - 0 \cdot \phi_2(t) = \begin{cases} -1, & 2 \le t < 3 \\ 0, & \text{otherwise.} \end{cases} \implies \phi_3(t) = \frac{\gamma_3(t)}{\|\gamma_3(t)\|}$$
- $$\gamma_4(t) = s_4(t) - \langle s_4(t), \phi_1(t)\rangle \phi_1(t) - \langle s_4(t), \phi_2(t)\rangle \phi_2(t) - \langle s_4(t), \phi_3(t)\rangle \phi_3(t)$$
$$= s_4(t) - (-\sqrt{2})\phi_1(t) - (0)\phi_2(t) - \phi_3(t) = 0$$
- Orthonormal basis $= \{\phi_1(t), \phi_2(t), \phi_3(t)\}$, where $N = 3 \le M = 4$.

- **Remark:** The orthonormal basis is **not unique**.
  For example, for $k = 1, 2, 3$, we can define

$$\phi_k(t) = \begin{cases} 1, & k-1 \leq t < k \\ 0, & \text{otherwise.} \end{cases}$$

  Then

$$\begin{aligned} s_1(t) &\xrightarrow{\mathbf{\Phi}} \mathbf{s}_1 = (+1, +1, 0) \\ s_2(t) &\xrightarrow{\mathbf{\Phi}} \mathbf{s}_2 = (+1, -1, 0) \\ s_3(t) &\xrightarrow{\mathbf{\Phi}} \mathbf{s}_3 = (+1, +1, -1) \\ s_4(t) &\xrightarrow{\mathbf{\Phi}} \mathbf{s}_4 = (-1, -1, -1) \end{aligned}$$

Vectorial Representation of Signals: A Summary

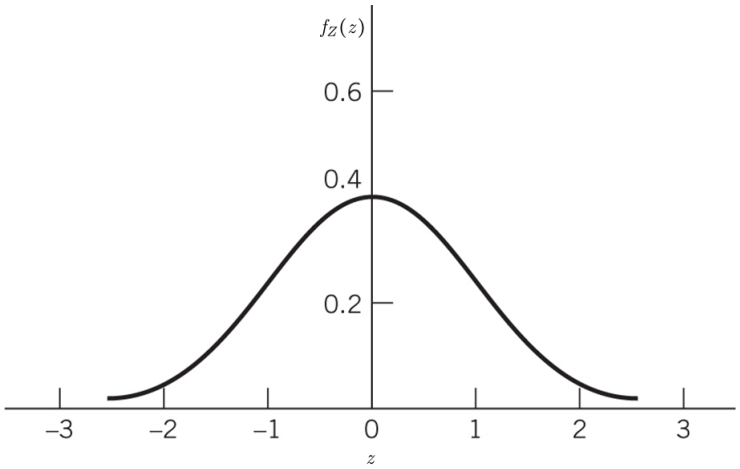
- Through the Gram-Schmidt orthogonalization, a set of M signals $\{s_m(t)\}_{m=1}^{M}$ can be represented by a set of M vectors $\{\mathbf{s}_m\}_{m=1}^{M}$ in the N-dimensional space, where $N \leq M$.
- The functions $\{\phi_n(t)\}$ obtained from the Gram-Schmidt procedure are not unique.
  - Changing the order of orthogonalization for $\{s_m(t)\}$ results in different orthonormal waveforms.
  - The vector representation of $\{s_m(t)\}$ depends on the chosen orthonormal functions $\{\phi_n(t)\}$.
- The dimensionality of the signal space, N, remains unchanged.
- The geometric configuration of the vectors $\{s_m\}$ (lengths and inner products) is invariant to the choice of $\{\phi_n(t)\}$.

8 Useful Random Variables

TABLE 2.3–3
Properties of Important Random Variables

RV (Parameters)	PDF or PMF	E[X]	VAR[X]
Bernoulli (p)	$P(X=1)=1-P(X=0)=p$ $0 \le p \le 1$	p	$p(1-p)$
Binomial (n, p)	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ $0 \le k \le n,\ 0 \le p \le 1$	np	$np(1-p)$
Uniform (a, b)	$\frac{1}{b-a},\ a \le x \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential (λ)	$\lambda e^{-\lambda x},\ \lambda > 0,\ x \ge 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gaussian (m, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}$ $\sigma > 0$	m	σ^2
Rayleigh (σ^2)	$\frac{x}{\sigma^2}e^{-x^2/2\sigma^2}$ $x, \sigma > 0$	$\sigma\sqrt{\frac{\pi}{2}}$	$\left(2-\frac{\pi}{2}\right)\sigma^2$
Jointly Gaussian ($\boldsymbol{m}$, $\boldsymbol{C}$)	$\frac{1}{(2\pi)^{n/2}\det(\boldsymbol{C})}e^{-\frac{1}{2}[(\boldsymbol{x}-\boldsymbol{m})'\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{m})]}$ $\boldsymbol{C}$ symmetric and positive definite	$\boldsymbol{m}$	$\boldsymbol{C}$ (cov. matrix)

Some Useful Random variables



- The *probability density function* (pdf) of a Gaussian random variable Z with a mean μ and a variance σ^2 is defined as

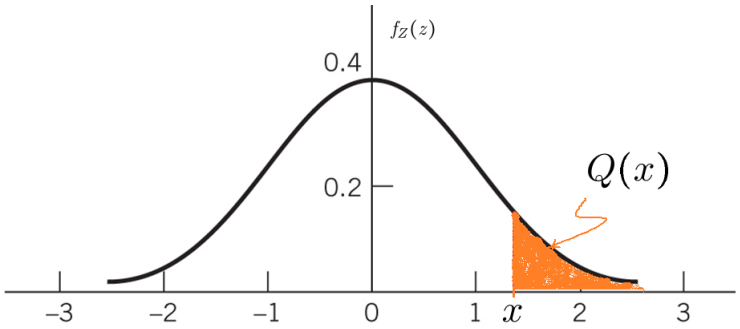
$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-m)^2}{2\sigma^2}}, \qquad -\infty < z < \infty$$

- for $m = 0$ and $\sigma^2 = 1$ we get $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ which is called *standard* Gaussian distribution, $\mathcal{N}(0, 1)$.

- For a given pdf $f_Z(z)$, by definition,

$$\boxed{P(a < Z \leq b) = \int_a^b f_Z(z) dz}$$

Some Useful Random variables



$P(a < Z \leq b)$ is the probability that Z is less than b and greater than a

- For a Gaussian pdf, such an integral cannot be evaluated in a closed form and must be evaluated numerically
- Numerical tables usually are available for *Q function* which is defined as $P[\mathcal{N}(0,1) > x)]$ which results in

$$\boxed{Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}$$

- In general, if $X \sim \mathcal{N}(m, \sigma^2)$, then

$$P[X > \alpha] = Q\left(\frac{\alpha - m}{\sigma}\right), \qquad P[X < \alpha] = Q\left(\frac{m - \alpha}{\sigma}\right)$$

- Following are some of the important properties of the Q function:

$$Q(-\infty) = 1, \qquad Q(0) = \frac{1}{2}, \qquad Q(\infty) = 0,$$

$$Q(-x) = 1 - Q(x).$$

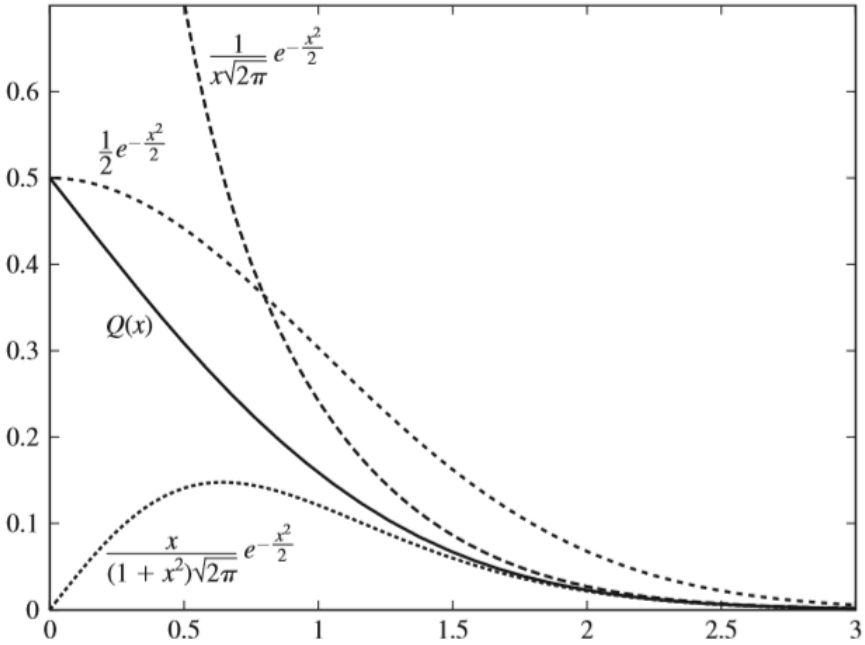
- Clearly, $Q(x)$ is a decreasing function.
- You can find Q-function using tables or MATLAB. In MATLAB, you can use
  - `qfunc(x)` to find the Q function of x, and
  - `qfuncinv(x)` to find its inverse
- Some useful bounds for the Q function for $x > 0$ are:

$$Q(x) \leq \frac{1}{2} e^{-x^2/2}$$

$$\frac{x}{(1+x^2)\sqrt{2\pi}} e^{-x^2/2} < Q(x) < \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}$$

- From the last two bounds we conclude that for large x we have

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}$$



9 Appendix: Vector Space

Definition 13 (Vector Space)**.** *A* ***vector space*** *is a collection of objects called* ***vectors****, which can be added together and multiplied by scalars to produce other vectors in the same set. It is defined over a* ***field*** *(e.g.,* $\mathbb{R}$ *or* $\mathbb{C}$*).*

- **Properties of a Vector Space:** Let V be a vector space over a field F, with $u, v, w \in V$ and $a, b \in F$. The following properties must hold:
  1. **Closure under addition:** $u + v \in V$
  2. **Commutativity of addition:** $u + v = v + u$
  3. **Associativity of addition:** $u + (v + w) = (u + v) + w$
  4. **Zero vector:** There exists $0 \in V$ such that $u + 0 = u$
  5. **Additive inverses:** For every $u \in V$, there exists $-u \in V$ such that $u + (-u) = 0$
  6. **Closure under scalar multiplication:** $a \cdot u \in V$
  7. **Distributive properties:**
     - $a \cdot (u + v) = a \cdot u + a \cdot v$
     - $(a + b) \cdot u = a \cdot u + b \cdot u$

8. **Associativity of scalar multiplication:** $a \cdot (b \cdot u) = (a \cdot b) \cdot u$
9. **Identity element:** $1 \cdot u = u$

Example 22. ***Examples of Vector Spaces:***

- $\mathbb{R}^n$*: The space of n-dimensional real vectors.*
- *The set of all polynomials of degree n or less.*
- *The set of all continuous functions from* $\mathbb{R}$ *to* $\mathbb{R}$*.*

Definition 14 (Subspace)**.** *A* `subspace` *is a subset of a vector space that is itself a vector space under the same addition and scalar multiplication.*

Conditions for a Subspace: A subset $W \subseteq V$ is a subspace of V if:

1. **Zero vector:** $0 \in W$
2. **Closed under addition:** If $u, v \in W$, then $u + v \in W$
3. **Closed under scalar multiplication:** If $u \in W$ and $a \in F$, then $a \cdot u \in W$

Example 23. ***Examples of Subspaces:***

- *The set* $\{(x, 0) \mid x \in \mathbb{R}\}$ *is a subspace of* $\mathbb{R}^2$.
- *The set of all polynomials with degree* ≤ 2 *is a subspace of the set of all polynomials.*
- *The solution space of a homogeneous system of linear equations forms a subspace.*

Key Difference Between Vector Space and Subspace:

- A **vector space** is the entire set of vectors that satisfies the properties of a vector space.
- A **subspace** is a smaller set within a vector space that also satisfies the properties of a vector space.