Fundamentals of Signals and Systems

Chapter 1 - Introduction

Outline of Topics

What is a signal ?

A *signal* is actually any quantity that varies with one or more independent variables such as time and space.

For example, the electricity that lights our nights, and

- Audio signals - speech and music signals - representing *air pressure* (varying with time).

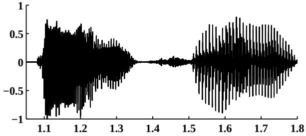


Fig. 1.1: A recording of the speech for “di qiu” in Chinese (meaning “earth” in English) spoken by a Chinese male.

- Video signals - picture and image signals - representing *brightness intensity/gray scale level* (varying with plane and time)

Fig. 1.2: The scene of *female teacher* in San Qing Mountain.

A picture is worth a thousand words. This English aphorism reminds us of the importance of images.

What is the use of signals?

Signals are used to carry/transmit *information* (also referred to as *messages*).

Examples of communications:

- Smoking used by our Chinese long long time ago;
- The gesture of *V* and Traffic lights;
- Trees of message;
- Hand-phones, $\cdots$.

What is a system?

A *system* is an *entity* that is used to achieve a specified function.

Consider a rectifier used to convert AC to DC:

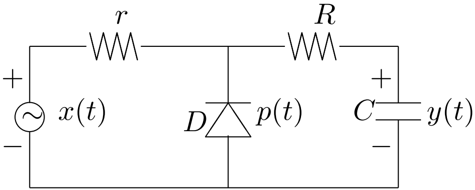


Fig. 1-3: Block diagram of a rectifier circuit.

A system is an interconnection of components or parts with terminals or access ports through which signals can be applied and extracted.

System representation

- Input/Output

$$x(t) \rightarrow y(t)$$

- Block-diagram/black box

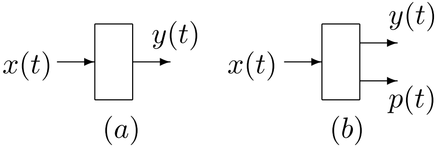


Fig. 1.4: Black box representations of the circuit by Fig. 1.3.

More sophisticated systems include

Fig. 1.5: A block diagram of communication systems.

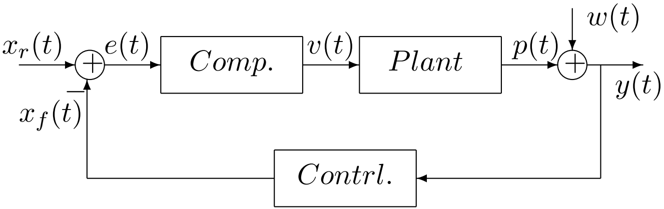


Fig. 1.6: A block diagram for a class of feedback control systems.

Description of signals

A signal is a function and usually described using

- figures/plots (a set of data)

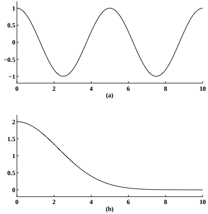


Fig. 1.8: Two signals presented by their plot.

- and very often by closed-form expressions (math. formulae). In Fig. 1.8, $s(t) = cos(0.4\pi t)$ and $y(\tau) = 2e^{-0.1\tau^2}$.

Classes of signals include:

A. Continuous-time (CT) v.s Discrete-time (DT)

Depending on the set $\mathcal{R}_\xi$ of values taken by the independent variable ξ, there are

Continuous-time (CT) signals or *analog signals*: $\mathcal{R}_\xi$ is continuous.

For example,

$$\begin{cases} s(t) &= cos(0.4\pi t),\ t \in \mathcal{R} \\ y(\tau) &= 2e^{-0.1\tau^2},\ \tau \in \mathcal{R} \end{cases} \tag{1}$$

Note: $s(t)$ and $s(\tau)$ represent one and the *same* signal as long as both t, τ are *independent* though different notations are used.

Discrete-time (DT) signals: $\mathcal{R}_\xi$ is discrete: $\mathcal{R}_\xi \in \mathcal{Z} \triangleq \{\cdots, -1, 0, 1, \cdots\}$.

Two examples of discrete-time signals defined on $\mathcal{Z}$ are given as

$$\begin{cases} s[n] = 10cos(0.125\pi n - 0.5), & n \in \mathcal{Z} \\ y[k] = e^{-0.5k^2}, & k \in \mathcal{Z} \end{cases} \tag{2}$$

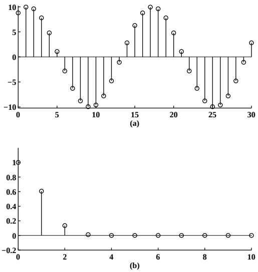


Fig. 1.9: Two discrete-time signals.

How DT signals are generated ?

- Inherently discrete (ex., no. of students in a class):

  For example, the cost of food taken *v.s* the number of students.

- Continuous in nature but sampled for digital processing.

  By sampling: $s[n] = s(t_n)$.

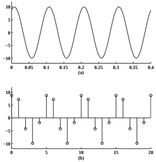


Fig. 1.11: (a) $s(t) = 10cos(20\pi t - 0.5),\ t \in [0\ \ 0.4]$. (b) $s[n] = s(t_n)$ with $t_n = n/50$.

More details on sampling will be given in a later chapter.

Notes:

To distinguish CT signals from DT signals:

- Variable notations: $t, \tau, f, \cdots$ for CT signals, $n, m, k, \cdots$ for DT signals.
- More importantly, *parentheses* $(.)$ are used for CT signals, while *brackets* $[.]$ for DT signals.

B. Energy v.s Power

How to quantify/measure signals? This is really the very essential issue for signal analysis.

Magnitude is a quantity to measure how *big* a signal is but can not give the overall view when the signal is time-varying.

Let $v(t)$ be the *voltage* across a resistor of unit resistance (i.e., one *ohm*). How to measure $v(t)$?

As well known from physics and circuit theory,

- *Instantaneous power*:

$$p(t) \triangleq i(t)v(t) = v^2(t)$$

- *Energy*: Expended/consumed over (t_1, t_2):

$$\int_{t_1}^{t_2} p(t)dt = \int_{t_1}^{t_2} v^2(t)dt$$

- *Average power* over (t_1, t_2):

$$\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} p(t)dt = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} v^2(t)dt$$

Borrowing these concepts for any signals, we have

- *Energy*:

$$E_x \triangleq \lim_{T \to +\infty} \int_{-T}^{T} |x(t)|^2 dt = \int_{-\infty}^{+\infty} |x(t)|^2 dt \qquad - \ CT\ signals$$

$$E_x \triangleq \lim_{N \to +\infty} \sum_{n=-N}^{N} |x[n]|^2 = \sum_{n=-\infty}^{+\infty} |x[n]|^2 \qquad - \ DT\ signals$$

- *Power*: CT signals DT signals

$$P_x \triangleq \lim_{T \to +\infty} \frac{1}{2T} \int_{-T}^{T} |x(t)|^2 dt \qquad P_x \triangleq \lim_{N \to +\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x[n]|^2$$

A signal x is said an *energy signal* if $0 < E_x < +\infty$, while a *power signal* if $0 < P_x < +\infty$

Key points: To calculate

$$E_x(T) \triangleq \int_{-T}^{T} |x(t)|^2 dt, \quad E_x[N] \triangleq \sum_{n=-N}^{N} |x[n]|^2$$

The definitions given above are also for complex-valued signals [1] $c(t) \in \mathcal{C}$.

[1] In the real world, all the signals are real-valued, but complex-valued signals are used in analysis. This will be seen in the sequel.

Important formulae

Cartesian form: $c = c_r + jc_i,$ where $j \triangleq \sqrt{-1}$ and

$$c_r \triangleq \mathcal{R}_e(c), \quad c_i \triangleq \mathcal{I}_m(c) \in \mathcal{R}$$

Polar form: $c = \rho e^{j\theta},$ where

$$\rho = \sqrt{c_r^2 + c_i^2}, \quad tan\theta = c_i/c_r$$

Euler's formula:

$$\boxed{e^{j\theta} = cos\theta + j sin\theta}$$

How to express $cos\theta$ and $sin\theta$ in terms of $e^{j\theta}$, $e^{-j\theta}$?

Example 1.1: Compute the energy and power of each of the following signals

- $x_1(t) = cos(0.125t), \ \forall\, t \in \mathcal{R}$.
- $x_2[n] = \begin{cases} r^n, & 0 \le n < N_0 \\ 0, & otherwise \end{cases}$, where $r \neq 0$ is constant and $N_0 > 0$ is a finite integer.
- $x_3(t) = cos(0.125t) + j2cost, \ \forall\, t \in \mathcal{R}$.

Solution:

- As $x_1(t)$ is a continuous-time signal, we have

$$\begin{aligned} E_{x_1}(T) &= \int_{-T}^{T} |x_1(t)|^2 dt = \int_{-T}^{T} cos^2(0.125t)dt \\ &= \int_{-T}^{T} \frac{1}{2}[1 + cos(2 \times 0.125t)]dt = T + 4sin(0.25T) \end{aligned}$$

Therefore,

$$E_{x_1} = \lim_{T \to +\infty} E_{x_1}(T) = +\infty, \quad P_{x_1} = \lim_{T \to +\infty} \frac{E_{x_1}(T)}{2T} = 1/2$$

So, $x_1(t)$ is a power signal.

- It is easy to see that for $N \geq N_0 - 1$,

$$E_{x_2}[N] = \sum_{n=-N}^{N} |x_2[n]|^2 = \sum_{n=0}^{N_0-1} r^{2n}$$

It follows from the famous formula for the sum, also known as geometric series, of a *geometric progression/sequence*

$$\boxed{\sum_{k=0}^{N-1} \alpha^k = \begin{cases} N, & \alpha = 1 \\ \frac{1-\alpha^N}{1-\alpha}, & \alpha \neq 1 \end{cases}} \qquad (3)$$

that $E_{x_2}[N] = \begin{cases} N_0, & r^2 = 1 \\ \frac{1-r^{2N_0}}{1-r^2}, & r^2 \neq 1 \end{cases}$. Clearly, $E_{x_2} = E_{x_2}[N_0 - 1]$ is positive and finite, and hence $P_{x_2} = 0$. So, $x_2[n]$ is an energy signal.

- Note that $|x_3(t)|^2 = cos^2(0.125t) + (2cost)^2 = cos^2(0.125t) + 4cos^2t$. Using the same procedure as that for $x_1(t)$, we can show that $x_3(t)$ is a power signal and $P_{x_3} = 1/2 + 4 \times 1/2 = \frac{5}{2}$.

End

It should be noted that if x is a power signal, then it can not be an energy, and vice versa. Is $x(t) = e^t$, $t \in \mathcal{R}$ a power or energy signal? (Answer: neither).

C. Periodic v.s Aperiodic

Observe the signal plotted in Fig. 1.12.

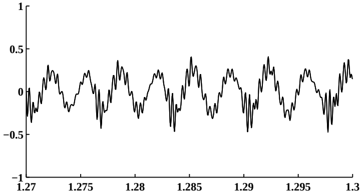


Fig. 1.12: A segment of the speech signal shown in Fig. 1.1.

What did you observe from it?

It repeats itself for every 0.0075 seconds!

Such a signal is said *periodic*.

A continuous-time signal $x(t)$ is said *periodic* is there exist a *positive* constant, say T, such that

$$x(t) = x(t+T), \quad \forall\, t \in \mathcal{R} \tag{4}$$

Clearly, $x(t) = x(t+\tilde{T}) \;\; \Rightarrow \;\; x(t) = x(t+k\tilde{T})$ for all positive integers k.

Fundamental period: $\quad T_0 \triangleq \min T \quad such\; that \;\; (4) \;\; holds.$

Fundamental frequency: $f_0 \triangleq \frac{1}{T_0}$

The frequency f_0 has a unit of hertz (Hz), i.e., cycles per second. Very often, we use *angular frequency* which is defined as

$$\omega_0 \triangleq 2\pi f_0 \qquad (in\; radians\; per\; second) \tag{5}$$

Example 1.2 : As well known, $x(t) = Acos(t)$ repeats itself every 2π, i.e., $x(t+2\pi) = x(t)$, it is periodic with $T_0 = 2\pi$. Suppose that both ω and ϕ_0 are constant. Is $s(t) = Acos(\omega t - \phi_0)$ periodic?

Solution: Note that

$$s(t+T) = Acos(\omega(t+T) - \phi_0) = Acos(\omega t - \phi_0) = s(t)$$

as long as

$$\omega T = 2\pi m \;\; \Rightarrow \;\; T = m\,\frac{2\pi}{\omega}$$

for any integer m. Therefore, $s(t)$ is periodic and

$$T_0 = \frac{2\pi}{\omega} \;\; \Rightarrow \;\; f_0 = \frac{\omega}{2\pi}$$

End

Remark: Noting that $x(t) = \kappa = \kappa cos(2\pi\, 0\, t) \;\; \Rightarrow \;\; f_0 = 0$. A constant is considered as a special periodic signal of $f_0 = 0$.

A $x[n]$ is said *periodic* if there exists a *positive integer* N such that

$$x[n] = x[n+N], \quad \forall\, n \in \mathcal{Z} \tag{6}$$

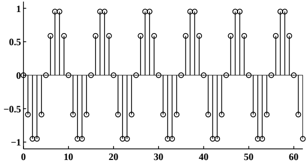


Fig. 1.14: A periodic discrete-time signal.

Similarly, the smallest value, say N_0, of such N is called the *fundamental period* of $x[n]$ and

$$F_0 \triangleq \frac{1}{N_0}, \quad \Omega_0 = 2\pi F_0 \tag{7}$$

are *fundamental freq.* and *angular freq.* for this periodic signal $x[n]$.

Example 1.3: For each of the following signals, determine whether it is periodic or not, and if it is periodic, find out the fundamental period.

- $x_1(t) = cos^2(2t)$
- $x_2(t) = e^{-\alpha^2 t} cos(2\pi t)$, where $\alpha \neq 0$ is a real number.
- $x_3[n] = (-1)^n$
- $x_4[n] = x(nT_s)$, where $x(t)$ is periodic with fundamental period $T_0 = 3T_s$.

Solution:

- Noting that $x_1(t) = \frac{1}{2}[1 + cos(4t)]$ and that $cos(4t)$ is periodic with a fundamental period $T_0 = \frac{2\pi}{4} = \frac{\pi}{2}$, we can see that

$$x_1(t + T_0) = x_1(t), \; \forall \; t$$

- If $x_2(t)$ is periodic, then there must exist a $T > 0$ such that $x_2(t) = x_2(t+T)$, that is $e^{-\alpha^2 t} cos(2\pi t) = e^{-\alpha^2 (t+T)} cos[2\pi(t+T)]$, leading to

$$cos(2\pi t) = e^{-\alpha^2 T} cos[2\pi(t+T)]$$

for all $t \in \mathcal{R}$, which is impossible as for $t = 0$ the left side of the equation is one, while the right side is smaller than one. This indicates that $x_2(t)$ is aperiodic.

- It is easy to see from that the plot of $x_3[n]$ that it is periodic with a fundamental period $N_0 = 2$. In fact, $x_3[n+N] = x_3[n] \;\Rightarrow\; 1 = (-1)^N$ and the smallest positive N is 2.
- It follows from $x(t) = x(t+T_0)$ that

$$x(nT_s) = x(nT_s+T_0) = x(nT_s+3T_s) \;\Rightarrow\; x_4[n] = x_4[n+3], \;\forall\, n \in \mathcal{Z}$$

which implies that $x_4[n]$ is periodic with a period not bigger than 3.

End

D. Deterministic v.s random

Look at $x(t) = Acos(2\pi ft + \phi)$, characterized with A, f, ϕ.

A signal is said *deterministic* if its characteristics are completely known. E.g., $x(t)$ with $A = 10, f = 50, \phi = \pi/3$.

From the view-point of communications, a deterministic signal contains no information. Why ?

A *random* signal is a signal whose characteristics obey a certain *probabilistic law*. E.g., $A = 10,\ f = 50$, but ϕ uncertain.

For example, you can prescribe with your beloved that

$$\phi = \begin{cases} \pi/4, & for \text{ ``}I\ love\ you\text{''} \\ -\pi/4, & for \text{ ``}I\ hate\ you\text{''} \end{cases}$$

$$\phi \ \rightarrow \ x(t) \rightarrow \ Channel \ \rightarrow \ \hat{x}(t) \ \rightarrow \ detection \ \rightarrow \ \hat{\phi}$$

E. Elementary signals

Unit step signals:[2]

$$u(t) \triangleq \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases}, \quad u[n] \triangleq \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \tag{8}$$

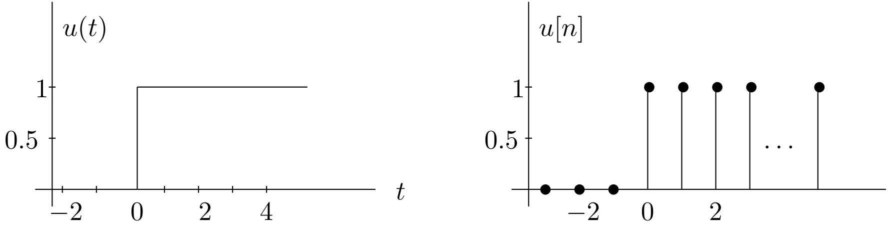


Fig. 1.15: The unit step signals $u(t)$ and $u[n]$.

[2]It is noted that unlike $u[n]$, the continuous-time unit step signal $u(t)$ is not defined at the origin $t = 0$.

Fig. 1.16 show the waveforms of the time shifted unit step signals:

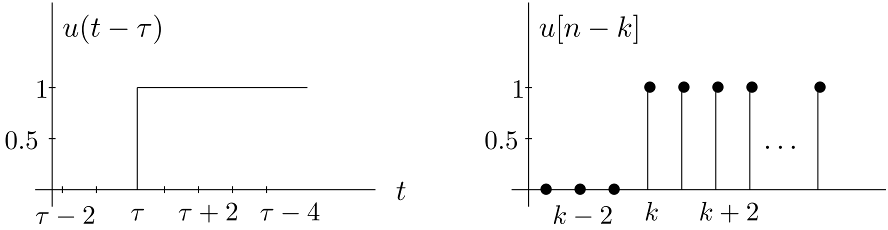


Fig. 1.16: The time shifted unit step signals $u(t-\tau)$ and $u[n-k]$.

The *sign signal* $sgn(t)$, another popularly used signal, is defined in terms of $u(t)$ as

$$sgn(t) \triangleq u(t) - u(-t) \tag{9}$$

One of the purposes of using elementary signals is to simplify expressions of signals.

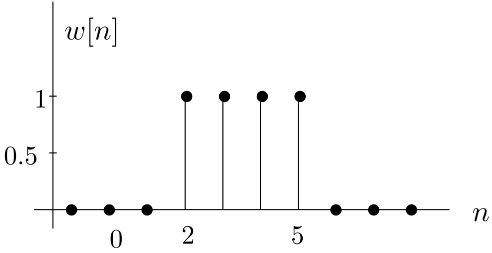
E.g., a piece-wise described signal

$$x(t) = \begin{cases} 10cos(\omega_0 t + \pi/4) & t > -1 \\ 0 & t < -1 \end{cases}$$

can be rewritten as

$$x(t) = 10cos(\omega_0 t + \pi/4)u(t+1)$$

in a closed-form.



Express $w[n]$ in terms of $u[n]$, please.

Window signals:

For CT case, the window signal is defined as

$$w_\tau(t) \triangleq \begin{cases} 1 & , \quad -\tau/2 < t < \tau/2 \\ 0 & , \quad otherwise \end{cases} \tag{10}$$

where $\tau > 0$ is a *positive* constant. See Fig. 1.18.

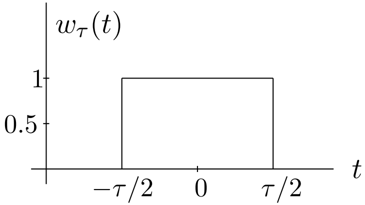


Fig. 1.18: The window signal $w_\tau(t)$.

Express $w_\tau(t)$ in terms of $u(t)$.

Unit impulse signals : The discrete-time *unit impulse signal* is defined as

$$\delta[n] \triangleq \begin{cases} 1 \quad , & n = 0 \\ 0 \quad , & otherwise \end{cases} \tag{11}$$

Such a signal, also called the *unit sample signal*, is shown in Fig. 1.19.

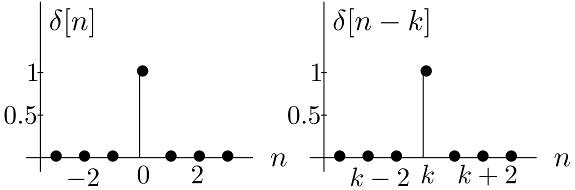


Fig. 1.19: The unit sample signal $\delta[n]$ and its shifted version $\delta[n-k]$.

A very important application:

$$\begin{aligned} x[n] &= \cdots + x[-1]\delta[n+1] + \cdots + x[m]\delta[n-m] + \cdots \\ &\triangleq \sum_{m=-\infty}^{+\infty} x[m]\delta[n-m] \end{aligned} \tag{12}$$

The CT unit impulse signal $\delta(t)$ is defined as

$$\delta(t) \triangleq \lim_{\tau \to 0} \frac{1}{\tau} w_\tau(t) \tag{13}$$

also referred to as *Dirac* function.

Fig. 1.20 depicts $\frac{1}{\tau} w_\tau(t)$ for different τ.

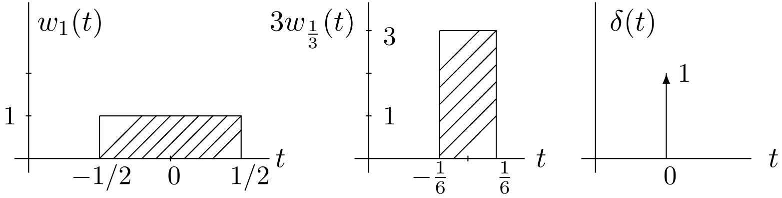


Fig. 1.20: Demonstration of $\delta(t)$ as a limit of $\frac{1}{\tau} w_\tau(t)$ for $\tau = 1,\ \frac{1}{3}$ and 0.

As observed, the surfaces of the shadowed areas are all equal to one. Therefore,

$$\int_{-\infty}^{+\infty} \delta(t)dt = \int_{-\sigma}^{\sigma} \delta(t)dt = 1, \;\; \forall \;\; \sigma > 0 \tag{14}$$

Graphically, $\kappa\delta(t-\tau)$ is represented by a vertical line starting at $t=\tau$ with the amplitude κ indicated near the arrow of the line.

Based on the definition (13), it can be shown (see **Problem 1.10**) that for any constant $\alpha \neq 0$

$$\delta(\alpha t) = \frac{1}{|\alpha|}\delta(t) \tag{15}$$

and particularly, $\delta(-t) = \delta(t)$.

Furthermore, for any given t_0

$$x(t)\delta(t-t_0) \equiv x(t_0)\delta(t-t_0) \tag{16}$$

It follows from (16) that

$$\begin{aligned}\int_{-\infty}^{+\infty} x(\xi)\delta(t-\xi)d\xi &= \int_{-\infty}^{+\infty} x(t)\delta(t-\xi)d\xi = x(t)\int_{-\infty}^{+\infty}\delta(\tau)d\tau \\ &= x(t) \end{aligned} \tag{17}$$

for any signal $x(t)$ as long as it is well defined at the instance t,[3] yielding

$$u(t) = \int_{-\infty}^{t} \delta(\tau)d\tau \tag{18}$$

Equivalently,

$$\delta(t) = \frac{du(t)}{dt} \tag{19}$$

[3] In a more rigorous manner, the *Dirac* function is defined as a function $\delta(t)$ such that (17) holds for any well-defined signal $x(t)$.

$\boxed{\textbf{Example 1.4}}$: With $x(t)$ sketched in Fig. 1.21(a), compute $\frac{dx(t)}{dt}$.

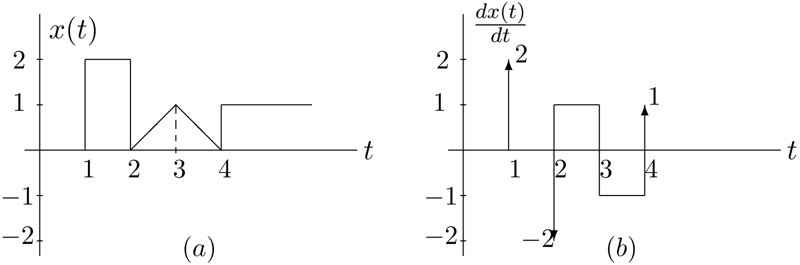


Fig. 1.21: Wave-forms of $x(t)$ and $\frac{dx(t)}{dt}$ in Example 1.4: (a) $x(t)$; (b) $\frac{dx(t)}{dt}$.

Solution: Note that

$$\begin{aligned} x(t) &= 2w_1(t-1.5)+(t-2)w_1(t-2.5) \\ &\quad +(-t+4)w_1(t-3.5)+u(t-4) \end{aligned}$$

Using the equality $f(t)\delta(t-t_0) = f(t_0)\delta(t-t_0)$, one has

$$
\begin{aligned}
\frac{dx(t)}{dt} &= 2\frac{dw_1(t-1.5)}{dt} + w_1(t-2.5) \\
&\quad +(t-2)\frac{dw_1(t-2.5)}{dt} - w_1(t-3.5) \\
&\quad +(-t+4)\frac{dw_1(t-3.5)}{dt} + \delta(t-4) \\
&= 2\delta(t-1) - 2\delta(t-2) + w_1(t-2.5) \\
&\quad -w_1(t-3.5) + \delta(t-4)
\end{aligned}
$$

End

Real exponential and sinusoidal signals:

Real-valued CT exponential signals are of form

$$x(t) = e^{\alpha t}, \quad \alpha \in \mathcal{R} \tag{20}$$

The behavior of such a signal is determined by α:

- when $\alpha = 0$, $x(t)$ is constant;
- when α is positive, $x(t) = e^{\alpha t}$ is exponentially growing as t increases (such as chain reactions in atomic explosions and complex chemical reactions).
- For $\alpha < 0$, $x(t) = e^{\alpha t}$ decays with time t (such as the voltage across the resistor in a resistor-capacitor (RC) circuit (in series) excited by a constant voltage source).

The class of exponentially amplitude-modulated sinusoidal signals is defined as

$$x(t) = e^{\alpha t} \, cos(2\pi f t + \phi) \tag{21}$$

When $\alpha < 0$, such a signal is called an *exponentially damped sinusoidal*. See Fig. 1.23.

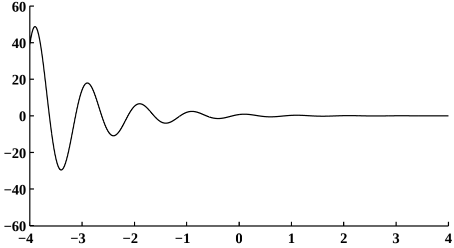


Fig. 1.23: Waveform of a damped sinusoidal signal $x(t) = e^{-t} cos(2\pi t - \pi/4)$.

$\boxed{\textit{Complex exponential signals}}$:

Let γ and A be two complex numbers. Then

$$x(t) = Ae^{\gamma t}$$

is called a continuous-time *complex exponential* signal.

Clearly, the real-valued $x(t) = A_r e^{\alpha t} cos(\beta t + \phi_0)$ can be rewritten as

$$\begin{aligned} x(t) &= \frac{A_r}{2} e^{\alpha t} [e^{j(\beta t + \phi_0)} + e^{-j(\beta t + \phi_0)}] \\ &= \frac{A_r}{2} e^{j\phi_0} e^{(\alpha + j\beta)t} + \frac{A_r}{2} e^{-j\phi_0} e^{(\alpha - j\beta)t} \triangleq A_1 e^{\gamma_1 t} + A_2 e^{\gamma_2 t} \quad (22) \end{aligned}$$

which implies that a real-valued continuous-time exponential sinusoidal signal can be represented *mathematically* using two complex exponential signals.

The discrete-time exponentially amplitude-modulated sinusoidal signals are defined as

$$x[n] \quad = \quad \rho^n cos(\Omega n + \phi) \tag{23}$$

where $\rho > 0, \Omega$ and ϕ are all real and constant.

A discrete-time *complex exponential* signal is of form

$$x[n] = A\gamma^n$$

where γ and A are two complex constants.

Show that for any (real-valued) exponential sinusoidal signal

$$x[n] = \rho^n cos(\Omega n + \phi) = A_1\gamma_1^n + A_2\gamma_2^n$$

where $\quad A_k =? \qquad \gamma_k = \ ? \qquad k = 1, 2$

The last elementary signal introduced in this section is the $sinc$ function:

$$s_\nu(t) \triangleq \frac{sin(t\nu/2)}{t\nu/2} \quad (24)$$

where ν is constant.

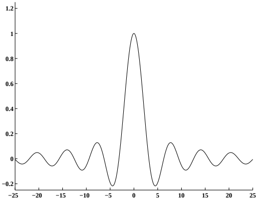


Fig. 1.24: Waveform of the sinc function $s_\nu(t)$ for $\nu = 2$.

Note: Although all the signals generated from physical phenomena can be well represented with real-valued functions, complex functions are sometimes found very useful in signal analysis.

A system is characterized by a series of operations on the input signals.

Elementary systems: $x \;\rightarrow\; y$

$\boxed{\textit{Time shifting systems}}$:

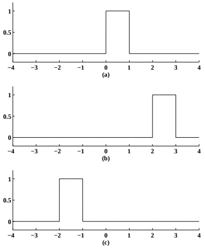


Fig. 1.25: Two time shifting systems: (a) $x(t)$. (b) $y(t) = x(t-2)$. (c) $y(t) = x(t+2)$.

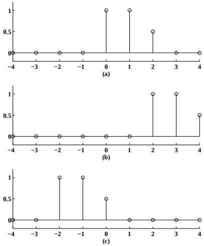


Fig. 1.26: (a) $x[n]$. (b) $y[n] = x[n-2]$. (c) $y[n] = x[n+2]$.

Mathematically described as

$$\begin{cases} x(t) & \rightarrow & y(t) = x(t-\tau) \\ x[n] & \rightarrow & y[n] = x[n-n_0] \end{cases} \qquad (25)$$

$$\tau > 0,\ n_0 > 0 \ \Rightarrow \ forward, \quad \tau < 0,\ n_0 < 0 \ \Rightarrow \ backward$$

$\boxed{\textit{Time scaling systems}}$:

$$x(t) \;\to\; y(t) = x(\alpha t),\; \alpha \neq 0 \tag{26}$$

Fig. 1.27 shows $x(\alpha t)$ with $x(t) = (1-t)w_1(t-1/2)$ for different α.

Particularly, such an operation is called *time reversal* when $\alpha = -1$.

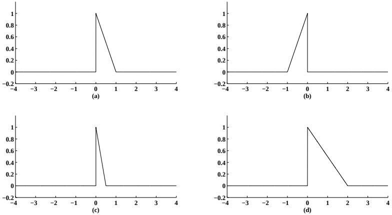


Fig. 1.27: (a) $x(t)$. (b) $x(-t)$. (c) $x(2t)$. (d) $x(t/2)$.

Combining the time shifting and the time scaling, we have a more general system:

$$x(t) \rightarrow \;\; y(t) = x(\alpha t - \beta)$$

Noting $y(t) = x(\alpha(t - \beta/\alpha))$, we can see that such an operation can be achieved by a cascade connection of a time scaler with a factor α, yielding $w(t) = x(\alpha t)$, and a time shifter with β/α, giving $y(t) = w(t - \beta/\alpha)$. See Fig. 1.28.

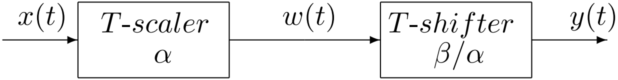


Fig. 1.28: Block-diagram of time shifting-scaling systems.

$\boxed{\textbf{Example 1.5}}$: Fig. 1.29 shows a signal $y(t)$. If $y(t) = x(-2t+1)$, sketch $x(t)$.

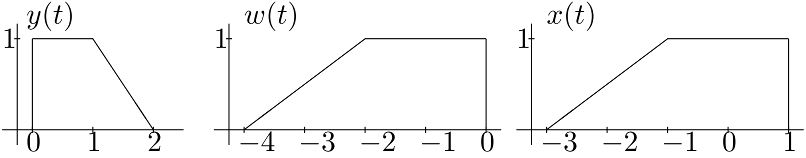


Fig. 1.29: Waveforms for Example 1.5.

Solution: Let $\tau = -2t+1$, then $t = -\frac{1}{2}(\tau - 1)$ and hence

$$y(t) = x(-2t+1) \;\Leftrightarrow\; x(\tau) = y(-\frac{1}{2}(\tau - 1))$$

Equivalently, $x(t) = y(-\frac{1}{2}(t-1))$, which leads to

$$w(t) \triangleq y(-\frac{1}{2}\,t) \;\rightarrow\; x(t) = w(t-1)$$

The discrete-time systems that realize the time scaling are defined as

$$x[n] \;\rightarrow\; y[n] = x[\alpha n] \tag{27}$$

where the constant α, unlike that in (26), takes on values which are a non-zero integer only.

Fig. 1.30 shows a signal $x[n]$ and its time scaled versions $x[2n]$ and $x[-2n]$.

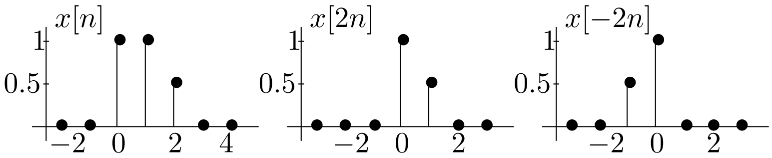


Fig. 1.30: Waveforms for $x[\alpha n]$.

$\boxed{\textit{Arithmetical operations}}$: Another three elementary operations are shown in Figure 52.

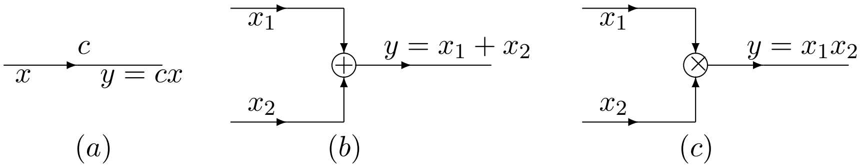


Fig. 31: Three arithmetical operations on signals. (a) amplitude scaling; (b) signal addition; (c) signal multiplication.

- An example of *amplitude scaling* is the *Ohm's law*: $v(t) = Ri(t)$
- Audio mixers are an example of *signal addition*, which add music and voice signals.
- The amplitude modulation (AM) radios are the result of *signal multiplication*

$$y(t) = s(t)cos(2\pi f_c t + \phi)$$

System modelling

- Studying a system actually means to investigate the relationship between the input and output signals: $x \ \rightarrow \ y$.
- This relationship is given by a collection of equations, called *system model*.
- The procedure of finding the equation is referred to as *system modelling*.
- The set-up of these equations is usually based on the *physical laws* such as *Newton*'s laws, *Ohm*'s law and *Kirchhoff*'s laws.

System I - RLC circuit

Look at the resistor-inductor-capacitor (RLC) (in series) network shown in Fig. 1.32.

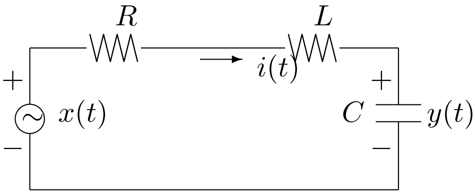


Fig. 1.32: A resistor-inductor-capacitor (RLC) (in series) circuit.

First of all, *Kirchhoff* tells $x(t) = v_R(t) + v_L(t) + y(t)$. Knowing that

$$v_R(t) = Ri(t), \;\; v_L(t) = L\frac{di(t)}{dt}, \;\; i(t) = C\frac{dy(t)}{dt}$$

we have

$$LC\frac{d^2y(t)}{dt^2} + RC\frac{dy(t)}{dt} + y(t) = x(t) \tag{28}$$

System II - Mechanical system

$x(t)$ - the applied force; $y(t)$ - the displacement of the mass M; $x(t)$ - the external force, K, D - the spring constant and the damping constant.

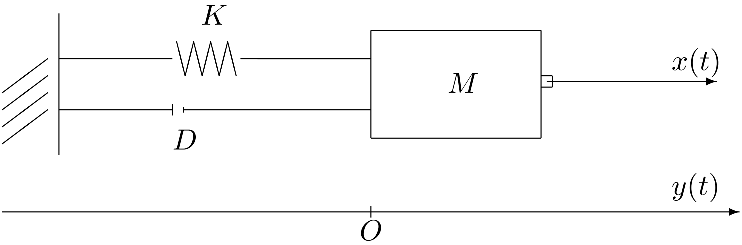


Fig. 1.33: A mechanical system.

Based on the *force balance* law, we get

$M\frac{d^2y(t)}{dt^2} = x(t) - Ky(t) - D\frac{dy(t)}{dt}$, i.e.,

$$M\frac{d^2y(t)}{dt^2} + D\frac{dy(t)}{dt} + Ky(t) = x(t) \tag{29}$$

Observation:

It is interesting to note that though *System I* and *System II* are two very different physical systems, both are described with the same system model, i.e., 2nd order linear differential equation of form

$$\alpha_2 \frac{d^2y(t)}{dt^2} + \alpha_1 \frac{dy(t)}{dt} + \alpha_0 y(t) = x(t) \tag{30}$$

Such a mathematical model can be used to describe many different physical systems.

System IV - A digital filter

A class of 1st order digital filters is described with

$$y[n] = \alpha y[n-1] + \frac{1-\alpha}{1+\alpha}\left(x[n] + \alpha x[n-1]\right) \tag{31}$$

with $y[-1] = 0$.

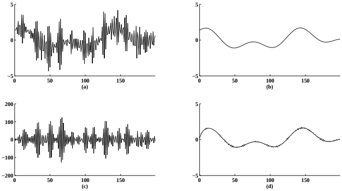


Fig. 1.34: (a) $x[n]$ - signal with noise; (b) $x_0[n]$ - the desired signal; (c) $y[n]$ - $\alpha = -0.75$; (d) $y[n]$ - with $\alpha = 0.75$.

Why $y[n] \approx x_0[n]$ with $\alpha = 0.75$, while it is a totally different story with $\alpha = -0.75$???

A. Memoryless and with memory

A system is said *memoryless* if for *any* input signal, the output at *any* given time depends *only* on the value of the input signal at that time.

- A resistor with resistance R is memoryless: $i(t) = \frac{v(t)}{R}$ as $i(t_0)$ depends only only the value of $v(t)$ at $t = t_0$.
- This is not the case for an inductor of inductance L as

$$i(t) = \frac{1}{L}\int_{-\infty}^{t} v(\tau)d\tau$$

implying for any t_0, $i(t_0)$ depends on all the values of $v(t)$ for $t < t_0$.

- What about the discrete-time system $\mathcal{S}$?

$$x[n] \;\rightarrow\; y[n] : y[n] = x[n] - x^2[n]$$

B. Causality

A system is said to be *causal* if for *any* input signal, the output at *any* given time depends *only* on values of the input *up to* that time.

- The system $y(t) = x^2(t-1)$ is causal, while $y(t) = x(t+1)$ is non-causal.
- The time reversal system $y[n] = x[-n]$ is another example of non-causal systems.

We have the following important statements:

Suppose that $x_k[n] \rightarrow\ y_k[n],\ \ k=1,2$. If the system is causal, then for any n_0 the following if-$then$ holds

$$x_1[n]=x_2[n],\ \forall\, n\leq n_0\ \ \Rightarrow\ \ y_1[n]=y_2[n],\ \forall\, n\leq n_0 \tag{32}$$

Similarly, if a continuous-time system $x(t)\ \rightarrow\ y(t)$ is causal and

$$x_k(t)\rightarrow\ y_k(t), k=1,2$$

then

$$x_1(t)=x_2(t),\ \forall\, t\leq t_0\ \ \Rightarrow\ \ y_1(t)=y_2(t),\ \forall\, t\leq t_0 \tag{33}$$

holds for any t_0 given.

C. Invertibility

A system $\mathcal{S}$ is said to be *invertible* if there exists a system $\mathcal{S}_{inv}$ such that *any* input of $\mathcal{S}$ can be recovered in the way shown in Fig. 1.36.

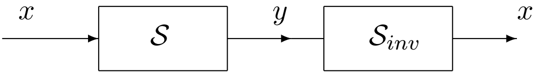


Fig. 1.36: System invertibility and inverse system

E.g., the system $\mathcal{S}$ with $x(t) \to y(t)$: $y(t) = 2x(t-3)$, is invertible as the $\mathcal{S}_{inv}$ specified with $v(t) \to w(t)$: $w(t) = \frac{1}{2}v(t+3)$, satisfies $w(t) = x(t)$ when $v(t) = y(t)$.

Another example of invertible system is the *integrator*

$$y(t) = \int_{-\infty}^{t} x(\xi)d\xi$$

since $x(t) = \frac{dy(t)}{dt}$ — the inverse system is the *differentiator*.

Question: Is the differentiator invertible?

There exist non-invertible systems. The simplest system of this kind may be the one $x(t) \rightarrow y(t)$: $y(t) = 0$, for which there is no way to determine the input which leads to $y(t) = 0$.

The concept of invertibility is of importance in many contexts. In communication systems, the received signal is generally different from the transmitted signal that propagates through the channel. If the channel is not an invertible system, there is no way to correct the distortion caused by the channel.

$\boxed{\textit{E. Stability}}$

A system is said to be *unstable* if it is out of work for an input signal of finite magnitude.

A famous example of a unstable system is the first Tacoma Narrows suspension bridge, collapsed on Nov. 7, 1940, due to wind-induced vibrations that coincided with the *inherent frequency* of the bridge.

This bridge is unstable !

How to define *stability*?

A signal x is said to be *bounded* if there exists a finite positive constant M such that $|x| < M$ holds on $\mathcal{R}$ or $\mathcal{Z}$.

An example of bounded signal is $x(t) = cos(2\pi ft)$ as

$$|x(t)| < 2, \; \forall \, t \in \mathcal{R}$$

A system $\mathcal{S}$ defined by $x \; \rightarrow \; y$ is said to be *stable* if for *any* bounded input x, the corresponding output y is bounded. Mathematically,

$$|x| \leq M_x < +\infty \;\Rightarrow\; |y| \leq M_y < +\infty \tag{34}$$

$\boxed{\textbf{Example 1.6}}$: Check the stability for each of the following systems:

- The system $x(t) \rightarrow y(t) : y(t) = 10e^{-t^2}x(t)$ is stable, because for any bounded $x(t)$ by M_x, that is $|x(t)| < M_x < +\infty$, $|y(t)| = |10e^{-t^2}x(t)| \leq 10|x(t)| < 10M_x \triangleq M_y < +\infty$ for all t, i.e., $y(t)$ is bounded.
- The system $x[n] \rightarrow y[n] : y[n] = 1.2y[n-1] + x[n], y[-1] = 0$ is unstable, because for $x[n] = \delta[n]$, a particular bounded input signal, $y[n] = 1.2^n u[n]$, which can be bigger than any given number.

Design of systems, called *controllers*, to stabilize unstable plants is one of the cores for control engineering discipline.

F. Time-invariance

A CT system is said to be *time-invariant* if $x(t) \to y(t)$, then

$$\hat{x}(t) = x(t - t_0) \;\to\; \hat{y}(t) = y(t - t_0),\ \forall\, t \in \mathcal{R}$$

holds for *any* input $x(t)$ and *arbitrary* constant $t_0 \in \mathcal{R}$.

Look at the resistor network shown in Fig. 1.38.

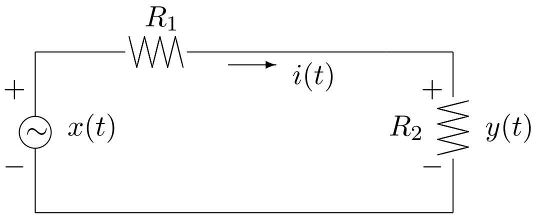


Fig. 1.38: A resistor network.

Clearly, the following expression is always true: $y(t) = \frac{R_2}{R_1+R_2}\, x(t)$.

- If $R_1, \; R_2$ all constant, the system is TI.
- If any of R_1, R_2 of this system is varying with time, it is not TI. Say

$$R_1 = 500, R_2 = 250(1 + cos(\pi t))$$

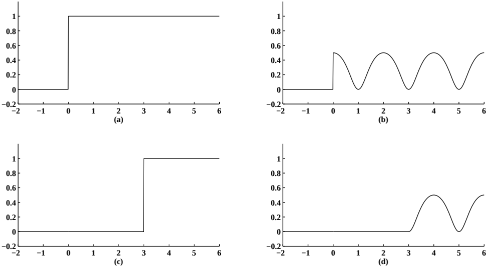


Fig. 1.39: Resistor network depicted in Fig. 1.38 with $R_1 = 500$, $R_2 = 250(1 + cos(\pi t))$: (a) $x(t) = u(t)$; (b) $y(t)$; (c) $\hat{x}(t) = u(t-3)$; (d) $\hat{y}(t)$.

Similarly, $x[n] \ \rightarrow \ y[n]$ is said to be *time-invariant* if

$$\hat{x}[n] = x[n - n_0] \ \rightarrow \ \hat{y}[n] = y[n - n_0], \ \forall \ n \in \mathcal{Z}$$

holds for for *any* input $x[n]$ and *arbitrary* constant $n_0 \in \mathcal{Z}$.

Example 1.7: Check if each of the two systems is time-invariant or not.

- The system: $y(t) = x^2(t-1)$ is TI, because for any constant t_0,

$$\begin{aligned} \hat{x}(t) \triangleq x(t - t_0) \ \rightarrow \ \hat{y}(t) &= \hat{x}^2(t-1) = x^2((t-1) - t_0) \\ &= x^2((t - t_0) - 1) = y(t - t_0) \end{aligned}$$

- How about the system by $y[n] = 0.5^n x^3[n]$?

G. *Linearity*

Let $x \rightarrow y$ be a system. Assume $x_k \rightarrow y_k, k = 1, 2$. The system is said to be *linear* if

$$\hat{x} \triangleq \alpha_1 x_1 + \alpha_2 x_2 \rightarrow \hat{y} = \alpha_1 y_1 + \alpha_2 y_2 \tag{35}$$

holds for *any* constants $\alpha_1, \alpha_2 \in \mathcal{C}$ and *arbitrary* signals x_1, x_2.

Example 1.8 : Let us examine the linearity of the following three systems.

- The system: $y[n] = x^2[n-1]$, is non-linear for the reason below. Let $x_k[n] \rightarrow y_k[n] = x_k^2[n-1], \ k = 1, 2$, and $\hat{x}[n] \triangleq \alpha_1 x_1[n] + \alpha_2 x_2[n]$. Then $\hat{x}[n] \Rightarrow \hat{y}[n] = \hat{x}^2[n-1]$, that is

$$\hat{y}[n] = \alpha_1^2 x_1^2[n-1] + 2\alpha_1\alpha_2 x_1[n-1]x_2[n-1] + \alpha_2^2 x_2^2[n-1]$$

while

$$\alpha_1 y_1[n] + \alpha_2 y_2[n] = \alpha_1 x_1^2[n-1] + \alpha_2 x_2^2[n-1]$$

Therefore, generally speaking,

$$\hat{y}[n] \neq \alpha_1 y_1[n] + \alpha_2 y_2[n] \;\; \Rightarrow \;\; non-linear$$

- Consider the system: $y(t) = x(2t)$.

  Assume $x_k(t) \;\rightarrow\; y_k(t), k = 1, 2$. Then

$$\alpha_1 y_1(t) + \alpha_2 y_2(t) = \alpha_1 x_1(2t) + \alpha_2 x_2(2t)$$

  while $\hat{x}(t) = \alpha_1 x_1(t) + \alpha_2 x_2(t) \;\rightarrow\; \hat{y}(t)$ with

$$\hat{y}(t) = \hat{x}(2t) = \alpha_1 x_1(2t) + \alpha_2 x_2(2t) = \alpha_1 y_1(t) + \alpha_2 y_2(t)$$

  which holds for *any* $x_k(t), \alpha_k \in \mathcal{C}$, we conclude that the system is linear.

- The system by

$$y(t) = x(2t) + 10$$

is non-linear because with $x_k(t) \rightarrow y_k(t) = x_k(2t) + 10,\ k = 1, 2,$ we have on the one hand,

$$\begin{aligned} \alpha_1 y_1(t) + \alpha_2 y_2(t) &= \alpha_1[x_1(2t) + 10] + \alpha_2[x_2(2t) + 10] \\ &= \alpha_1 x_1(2t) + \alpha_2 x_2(2t) + 10(\alpha_1 + \alpha_2) \end{aligned}$$

On the other hand, for the input $\hat{x}(t) = \alpha_1 x_1(t) + \alpha_2 x_2(t)$ the corresponding output is $\hat{y}(t) = \hat{x}(2t) + 10$, i.e.,

$$\hat{y}(t) = \alpha_1 x_1(2t) + \alpha_2 x_2(2t) + 10$$

Obviously, $\hat{y}(t) \neq \alpha_1 y_1(t) + \alpha_2 y_2(t)$ for those α_1, α_2 satisfying $\alpha_1 + \alpha_2 \neq 1$.

Generally, we have

$$x = \sum_k \alpha_k x_k \;\rightarrow\; y = \sum_k \alpha_k y_k \tag{36}$$

where the number of terms in the summations can be *finite* or *infinite*.

(36) is known as *superposition* property.

The interpretation is graphically given in Fig. 1.40.

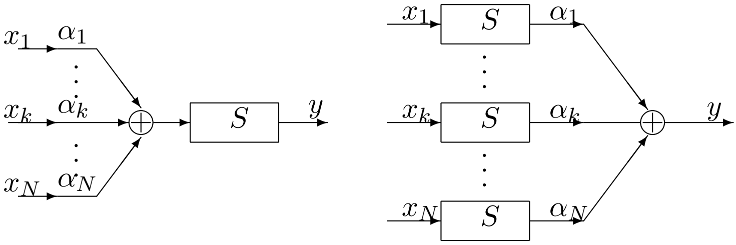


Fig. 1.40: Interpretation of superposition principle.

Below are some properties that hold for the class of linear systems:

- *A zero input leads to a zero output*, i.e., $x = 0 \;\rightarrow\; y = 0$. Because

$$x = 0 = 0\,\tilde{x} \;\rightarrow\; y = 0\,\tilde{y} = 0$$

where $\tilde{x} \;\rightarrow\; \tilde{y}$ for any $\tilde{x}$. See the 3rd system: $y(t) = x(2t) + 10$ of Example 1.8.

- A linear system $x \;\rightarrow\; y$ is causal *if and only if* the following holds

$$\begin{cases} x(t) = 0,\ \forall\ t \leq t_0 & \rightarrow \quad y(t) = 0,\ \forall\ t \leq t_0 \\ & \quad (continuous - time\ systems) \\ x[n] = 0,\ \forall\ n \leq n_0 & \rightarrow \quad y[n] = 0,\ \forall\ n \leq n_0 \\ & \quad (discrete - time\ systems) \end{cases} \tag{37}$$

for *any* $t_0 \in \mathcal{R}$, $n_0 \in \mathcal{Z}$ and all such input signals x. Referred to as the *condition of initial rest* and the proof found in **Appendix A**.

$\boxed{\textbf{Example 1.9}}$: Let $h_k[n]$ be the output of a linear system excited by $e_k[n]$, i.e., $e_k[n] \; \to \; h_k[n]$ for $k = 1, 2$. Assume that

$$\begin{aligned} e_1[n] + 2e_2[n] \quad &\to \quad f[n] \\ 2e_1[n] - e_2[n] \quad &\to \quad g[n] \end{aligned}$$

Find out $h_1[n], h_2[n]$ in terms of $f[n]$ and $g[n]$.

Proof: As the system is linear, we have

$$\begin{aligned} e_1[n] + 2e_2[n] \quad &\to \quad h_1[n] + 2h_2[n] = f[n] \\ 2e_1[n] - e_2[n] \quad &\to \quad 2h_1[n] - h_2[n] = g[n] \end{aligned}$$

Solving the equations yields $h_1[n] = \frac{f[n]+2g[n]}{5}$ and $h_2[n] = \frac{2f[n]-g[n]}{5}$.