Introduction to Signals and Systems

Lathi Chapt. 1

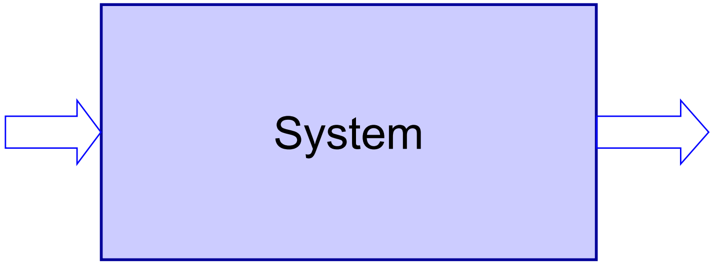
Didactic material

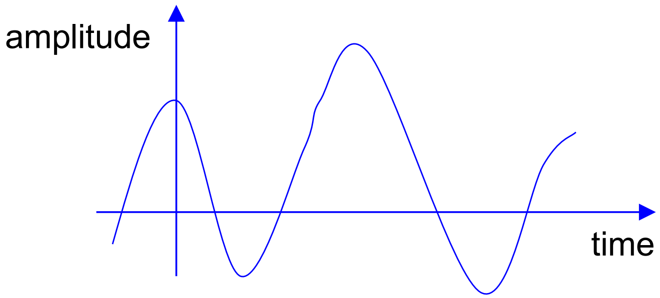
- Textbook
  - Signal Processing and Linear Systems, B.P. Lathi, CRC Press
- Other books
  - Signals and Systems, Richard Baraniuk's lecture notes, available on line
  - Digital Signal Processing (4th Edition) (Hardcover), John G. Proakis, Dimitris K Manolakis
  - Teoria dei segnali analogici, M. Luise, G.M. Vitetta, A.A. D'Amico, McGraw-Hill
  - Signal processing and linear systems, Schaun's outline of digital signal processing
- All textbooks are available at the library
- Handwritten notes will be available on demand

Signals&Systems

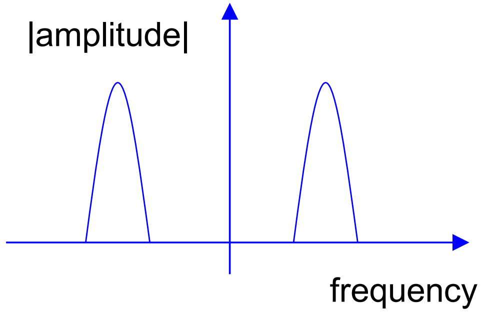
Input signal

Output signal

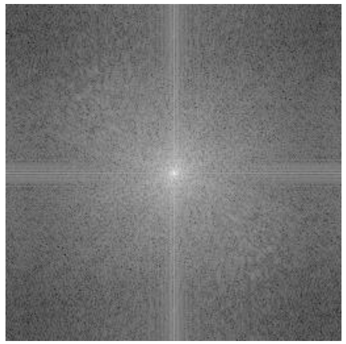




Linear time invariant systems (LTIS)



LTIS perform any kind of processing on the input data to generate output data



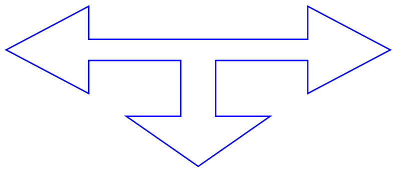
Contents

Signals

- Signal classification and representation
  - Types of signals
  - Sampling theory
  - Quantization
- Signal analysis
  - Fourier Transform
    - Continuous time, Fourier series, Discrete Time Fourier Transforms, Windowed FT
  - Spectral Analysis

Systems

- Linear Time-Invariant Systems
  - Time and frequency domain analysis
  - Impulse response
  - Stability criteria
- Digital filters
  - Finite Impulse Response (FIR)
- Mathematical tools
  - Laplace Transform
    - Basics
  - Z-Transform
    - Basics



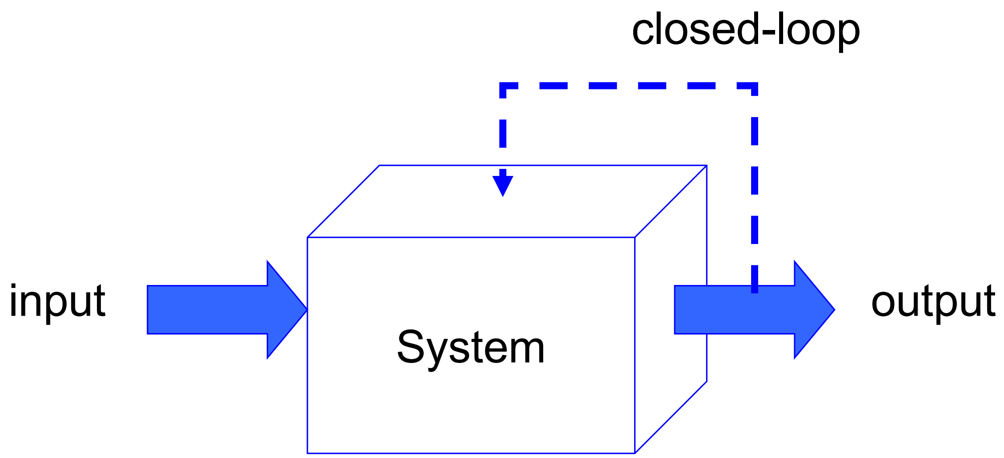
Applications in the domain of Bioinformatics

What is a signal?

- A signal is a set of information of data
  - Any kind of physical variable subject to variations represents a signal
  - Both the independent variable and the physical variable can be either scalars or vectors
    - Independent variable: time (t), space (x, $\mathbf{x}=[x_1,x_2]$, $\mathbf{x}=[x_1,x_2,x_3]$)
    - Signal:
    - Electrochardiography signal (EEG) 1D, voice 1D, music 1D
    - Images (2D), video sequences (2D+time), volumetric data (3D)

What is a system?

- Systems process signals to
  - Extract information
  - Enable transmission over channels with limited capacity (JPEG, JPEG2000, MPEG coding)
  - Improve security over networks (encryption, watermarking)
  - Support the formulation of diagnosis and treatment planning (medical imaging)
  - …….



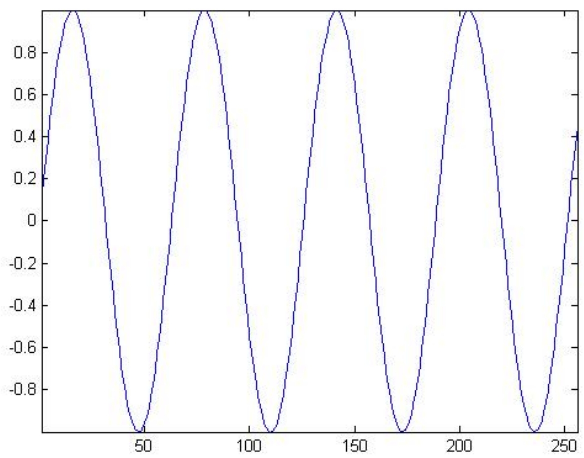
The function linking the output of the system with the input signal is called transfer function and it is typically indicated with the symbol $h(\bullet)$

Classification of signals

- Continuous time – Discrete time
- Analog – Digital (numerical)
- Periodic – Aperiodic
- Energy – Power
- Deterministic – Random (probabilistic)
- Note
  - Such classes are not disjoint, so there are digital signals that are periodic of power type and others that are aperiodic of power type etc.
  - Any combination of single features from the different classes is possible

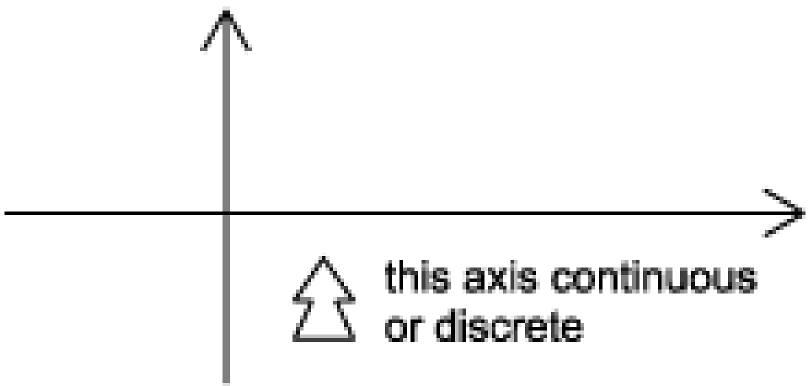
Continuous time – discrete time

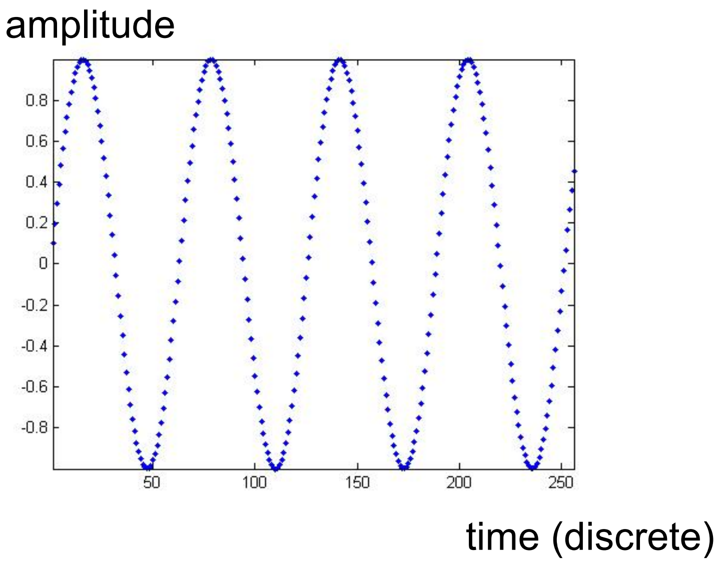
- Continuous time signal: a signal that is specified for every real value of the independent variable
  - The independent variable is continuous, that is it takes any value on the real axis
  - The domain of the function representing the signal has the cardinality of real numbers
    - Signal ↔ f=f(t)
    - Independent variable ↔ time (t), position (x)
    - For continuous-time signals: $t \in \mathbb{R}$



Continuous time – discrete time

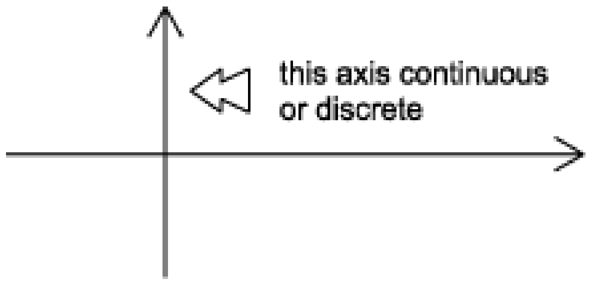
- Discrete time signal: a signal that is specified only for *discrete values* of the independent variable
  - It is usually generated by *sampling* so it will only have values at *equally spaced* intervals along the time axis
  - The domain of the function representing the signal has the cardinality of integer numbers
    - Signal ↔ f=f[n], also called "sequence"
    - Independent variable ↔ n
    - For discrete-time functions: $t \in \mathbf{Z}$





Analog - Digital

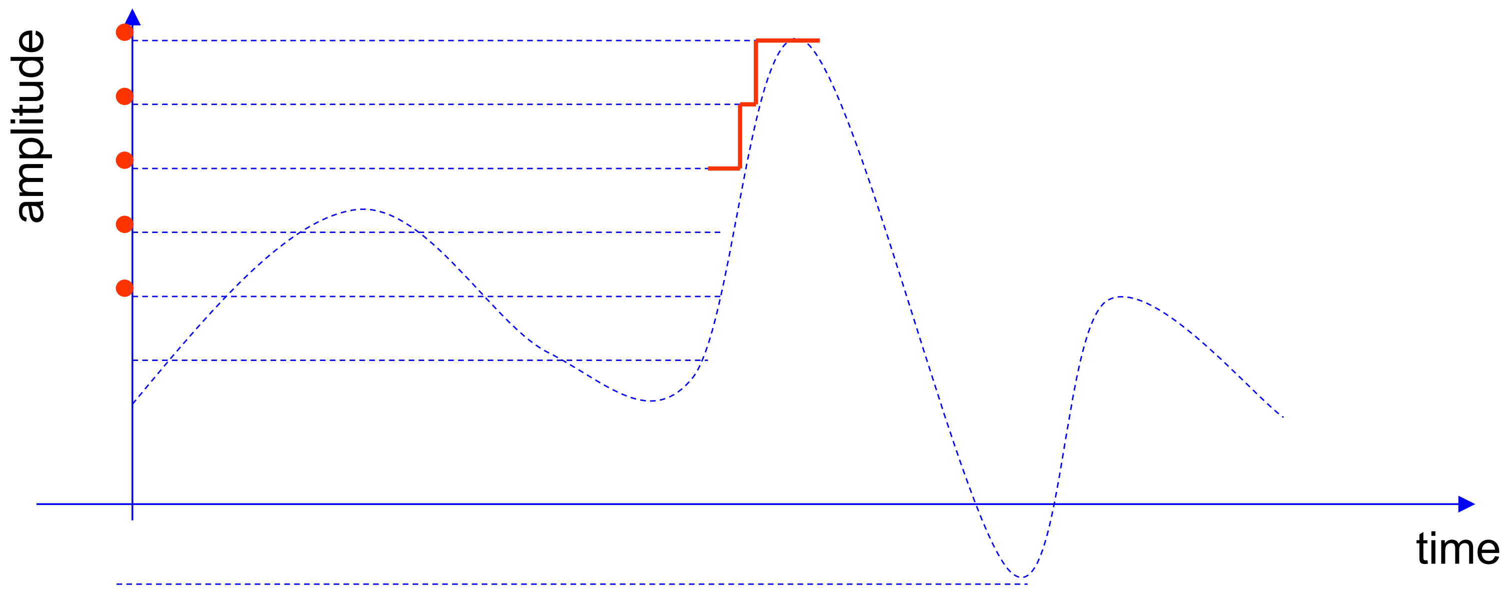
- **Analog signal**: signal whose amplitude can take on any value in a continuous range
  - The amplitude of the function f(t) (or f(x)) has the cardinality of real numbers
    - The difference between analog and digital is similar to the difference between continuous-time and discrete-time. In this case, however, the difference is with respect to the value of the function (y-axis)
  - Analog corresponds to a continuous y-axis, while digital corresponds to a discrete y-axis



- *Here we call digital what we have called quantized in the EI class*
- An analog signal can be both continuous time and discrete time

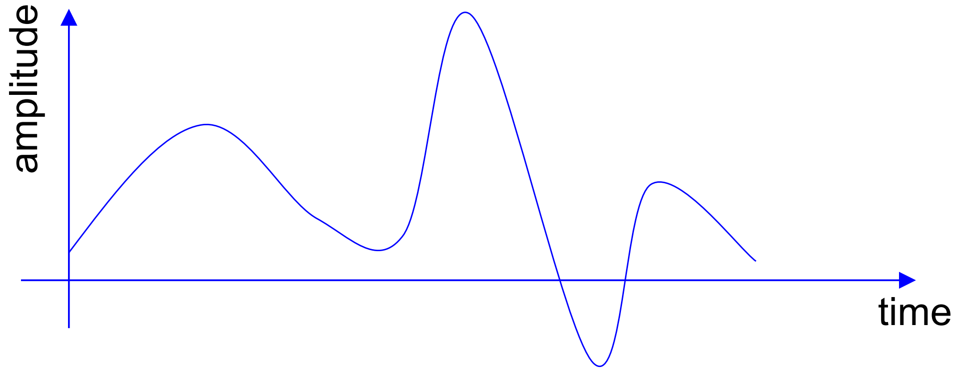
Analog - Digital

- **Digital signal**: a signal is one whose amplitude can take on only a finite number of values (thus it is quantized)
  - The amplitude of the function f() can take only a finite number of values
  - A digital signal whose amplitude can take only M different values is said to be M-ary
    - Binary signals are a special case for M=2

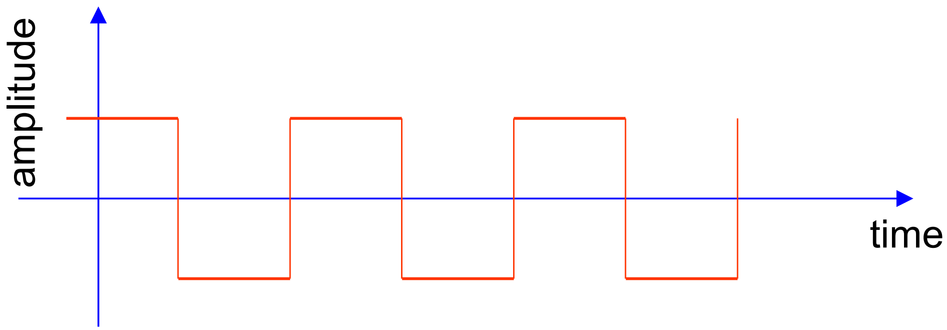


Example

– Continuous time analog

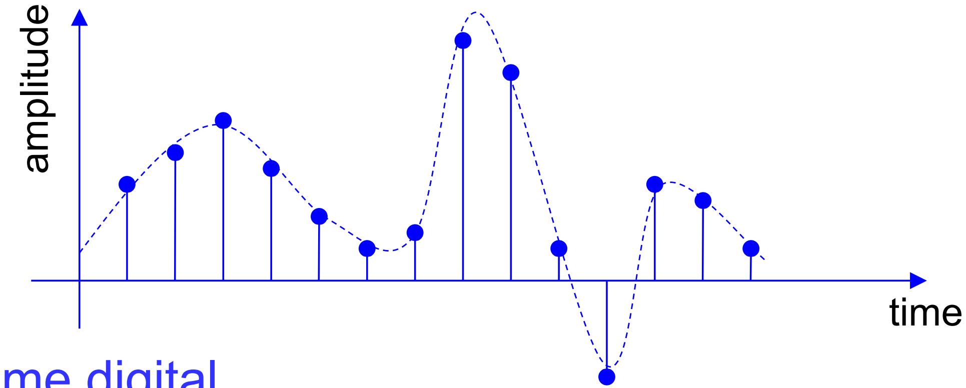


– Continuous time digital (or quantized)
  - binary sequence, where the values of the function can only be one or zero.

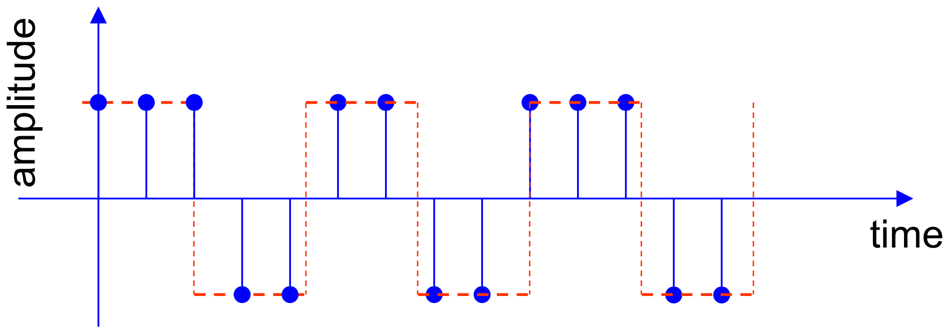


Example

- Discrete time analog



- Discrete time digital
  - binary sequence, where the values of the function can only be one or zero.



Summary

Signal amplitude/ Time or space	Real	Integer
Real	Analog Continuous-time	Digital Continuous-time
Integer	Analog Discrete-time	Digital Discrete time

Note

- In the image processing class we have defined as digital those signals that are both quantized and discrete time. It is a more restricted definition.
- The definition used here is as in the Lathi book.

Periodic - Aperiodic

- A signal f(t) is *periodic* if there exists a positive constant T_0 such that

$$f(t+T_0) = f(t) \qquad \forall t$$

  - The *smallest* value of T_0 which satisfies such relation is said the *period* of the function f(t)
  - A periodic signal remains unchanged when *time-shifted* of integer multiples of the period
  - Therefore, by definition, it starts at minus infinity and lasts forever

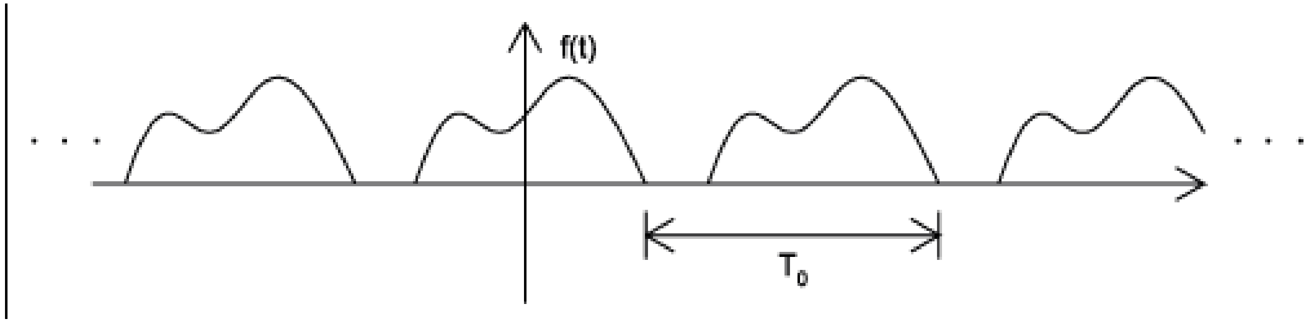
$$-\infty \le t \le +\infty \qquad t \in {}^{\circ}$$

$$-\infty \le n \le +\infty \qquad n \in \mathbf{Z}$$

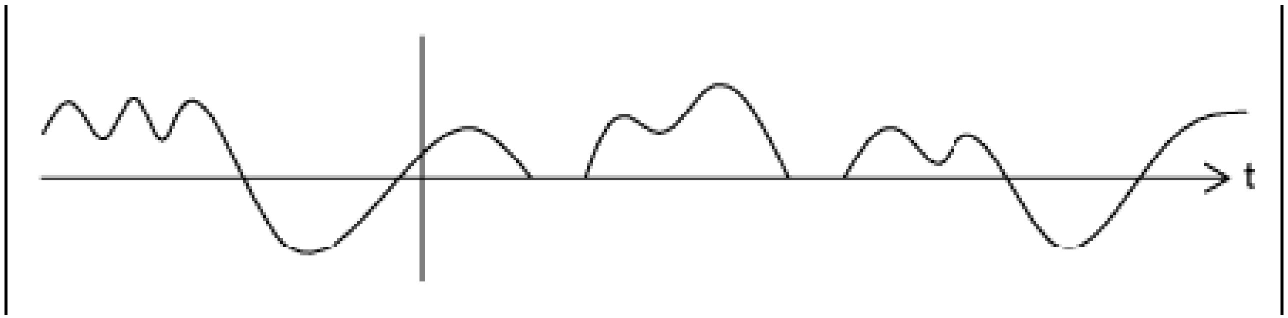
  - Periodic signals can be generated by *periodical extension*

Examples

- Periodic signal with period T_0

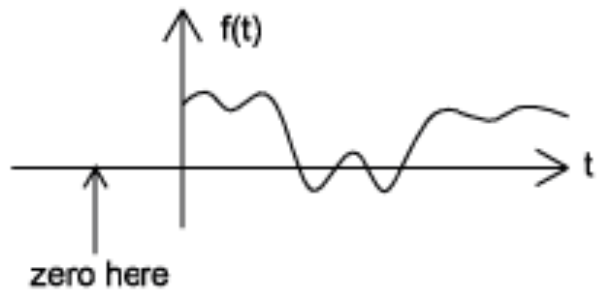


- Aperiodic signal

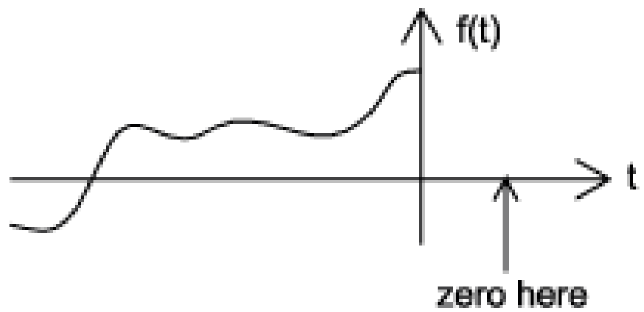


Causal and non-Causal signals

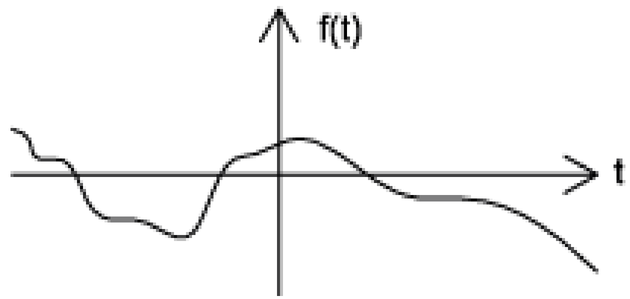
- *Causal* signals are signals that are *zero for all negative time (or spatial positions)*, while



- *Anticausal* are signals that are zero for all *positive time (or spatial positions)*.



- *Noncausal* signals are signals that have nonzero values in both positive and negative time



Causal and non-causal signals

- Causal signals

$$f(t) = 0 \qquad t < 0$$

- Anticausals signals

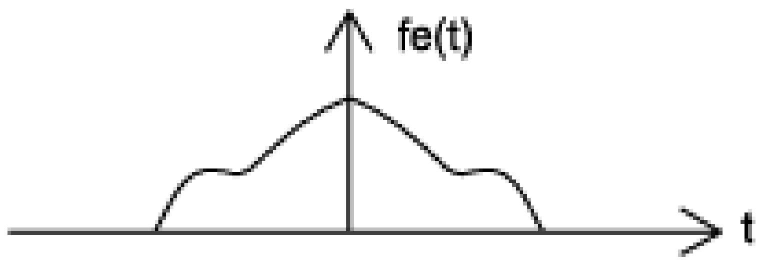
$$f(t) = 0 \qquad t \geq 0$$

- Non-causal signals

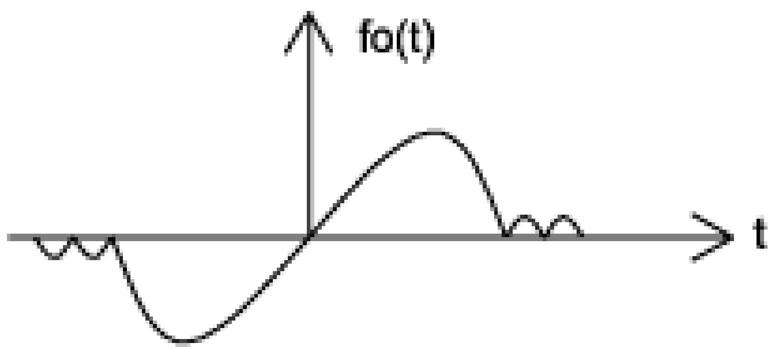
$$\exists t_1 < 0: \qquad f(t_1) = 0$$

Even and Odd signals

- An even signal is any signal f such that f (t) = f (-t). Even signals can be easily spotted as they are symmetric around the vertical axis.



- An odd signal, on the other hand, is a signal f such that f (t)= - (f (-t))



Decomposition in even and odd components

- Any signal can be written as a combination of an even and an odd signals
  - Even and odd components

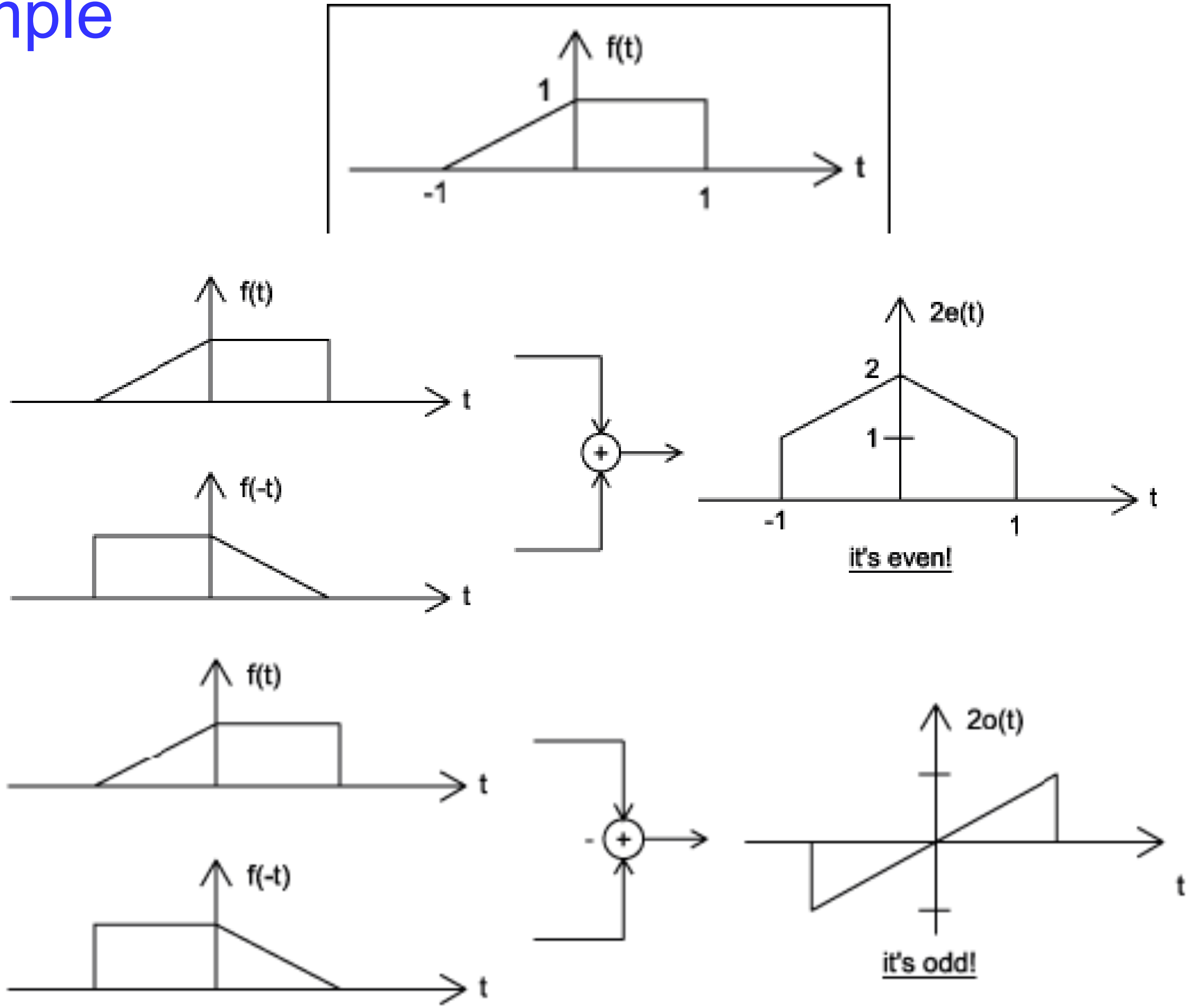
$$f(t) = \frac{1}{2}\left(f(t) + f(-t)\right) + \frac{1}{2}\left(f(t) - f(-t)\right)$$

$$f_e(t) = \frac{1}{2}\left(f(t) + f(-t)\right) \quad \text{even component}$$

$$f_o(t) = \frac{1}{2}\left(f(t) - f(-t)\right) \quad \text{odd component}$$

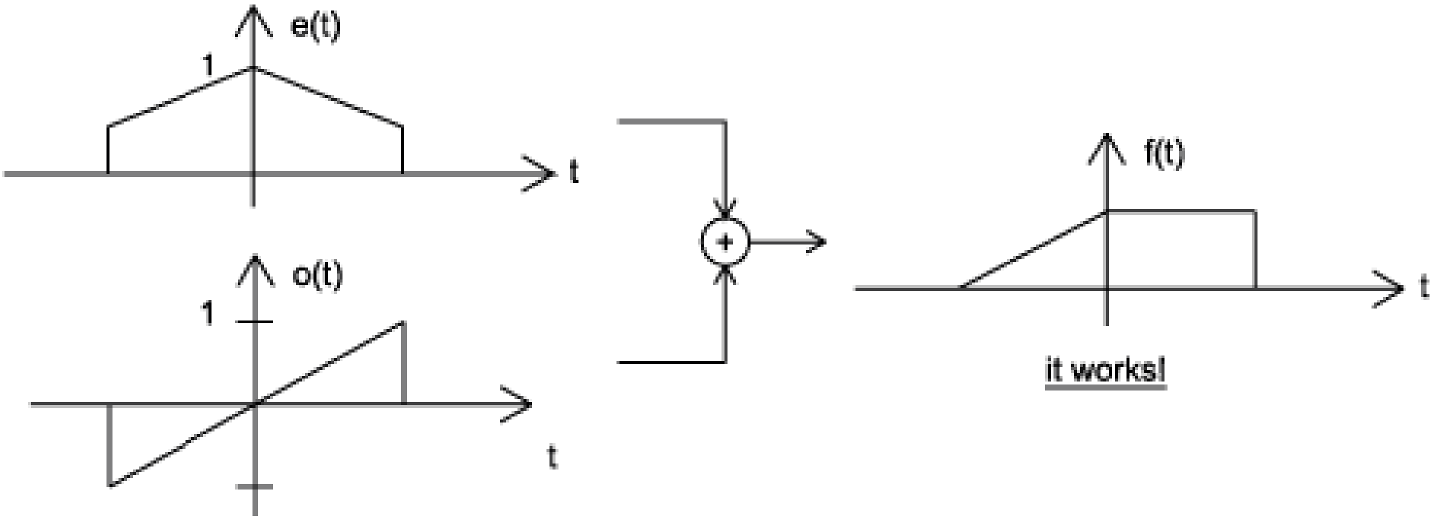
$$f(t) = f_e(t) + f_o(t)$$

Example



Example

- Proof



Some properties of even and odd functions

- even function x odd function = odd function
- odd function x odd function = even function
- even function x even function = even function
- Area

$$\int_{-a}^{a} f_e(t)\,dt = 2\int_{0}^{a} f_e(t)\,dt$$

$$\int_{-a}^{a} f_o(t)\,dt = 0$$

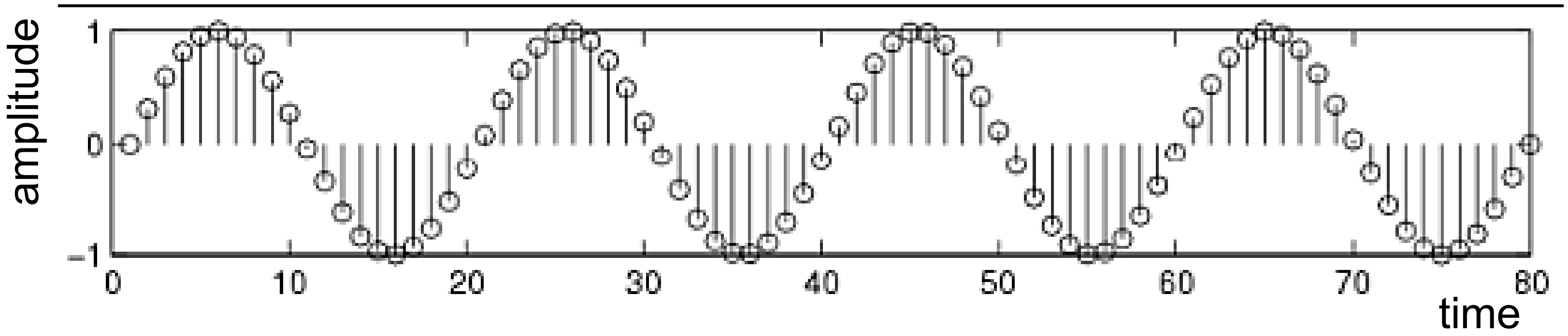
Deterministic - Probabilistic

- Deterministic signal: a signal whose *physical description* in known completely
- A deterministic signal is a signal in which each value of the signal is fixed and can be determined by a mathematical expression, rule, or table.
- Because of this the future values of the signal can be calculated from past values with complete confidence.
  - There is *no uncertainty* about its amplitude values
  - Examples: signals defined through a mathematical function or graph

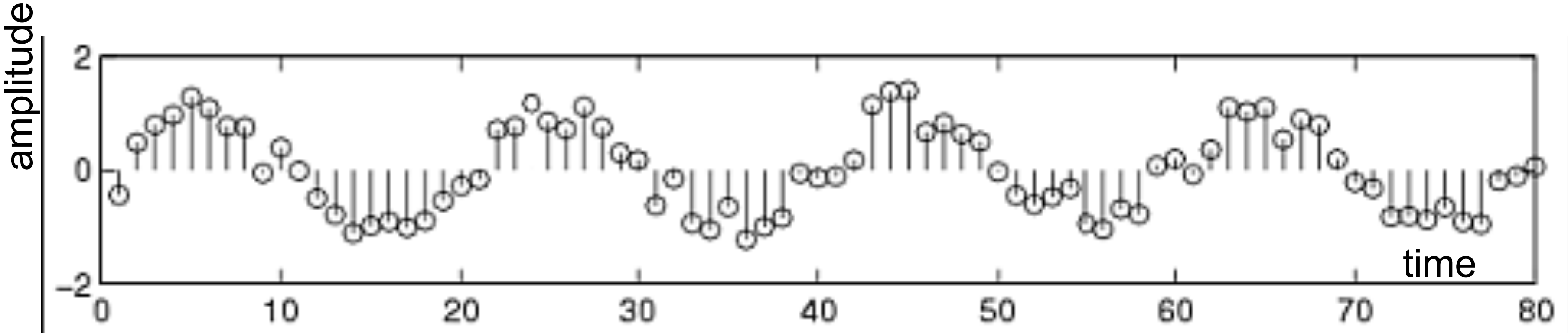
- Probabilistic (or random) signals: the amplitude values *cannot be predicted precisely* but are known only in terms of probabilistic descriptors
- The future values of a random signal cannot be accurately predicted and can usually only be guessed based on the averages of sets of signals
  - They are realization of a stochastic process for which a model could be available
  - Examples: EEG, evocated potentials, noise in CCD capture devices for digital cameras

Example

- Deterministic signal



- Random signal



Finite and Infinite length signals

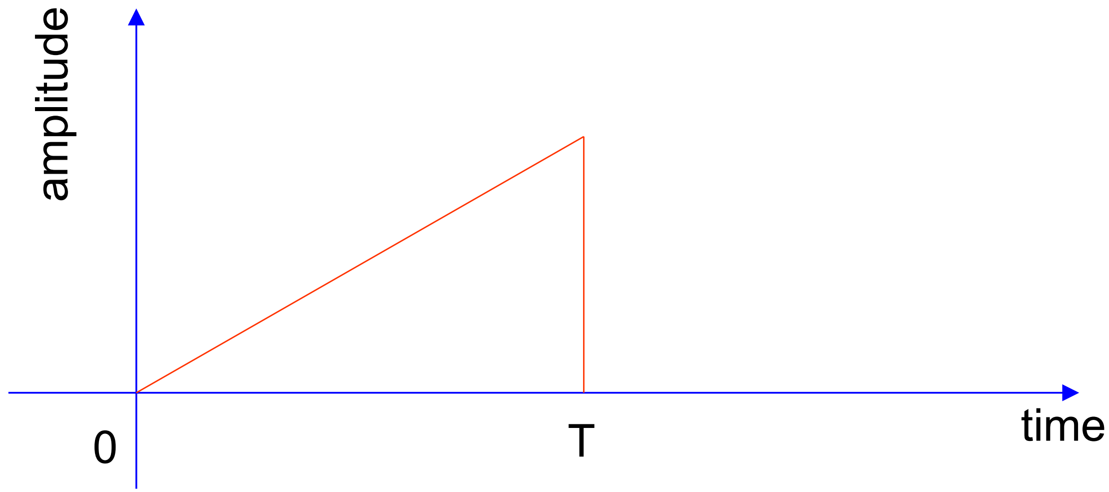
- A finite length signal is non-zero over a finite set of values of the independent variable

$$f = f(t), \forall t : t_1 \le t \le t_2$$
$$t_1 > -\infty, t_2 < +\infty$$

- An infinite length signal is non zero over an infinite set of values of the independent variable
  - For instance, a sinusoid $f(t)=\sin(\omega t)$ is an infinite length signal

Size of a signal: Norms

- "Size" indicates largeness or strength.
- We will use the mathematical concept of the norm to quantify this notion for both continuous-time and discrete-time signals.
- The energy is represented by the area under the curve (of the squared signal)



Energy

- Signal energy

$$E_f = \int_{-\infty}^{+\infty} f^2(t)dt$$

$$E_f = \int_{-\infty}^{+\infty} \left|f(t)\right|^2 dt$$

- Generalized energy : L_p norm
  - For p=2 we get the energy (L_2 norm)

$$\left\|f(t)\right\| = \left(\int \left(\left|f(t)\right|\right)^p dt\right)^{1/p}$$

$$1 \leq p < +\infty$$

Power

- Power
  - The power is the time average (mean) of the squared signal amplitude, that is the *mean-squared* value of f(t)

$$P_f = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} f^2(t) dt$$

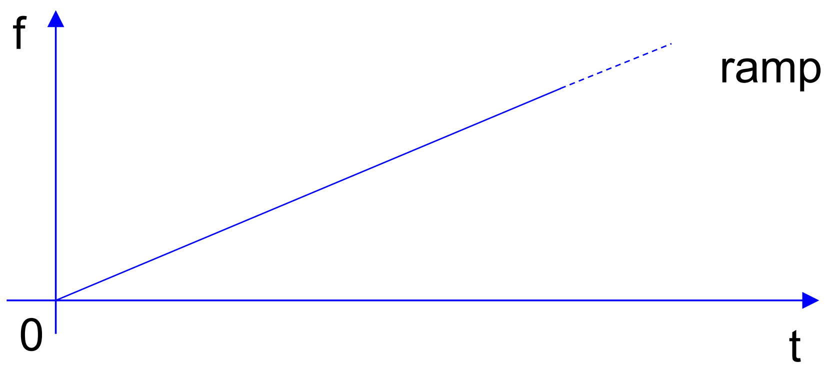
$$P_f = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} \left| f(t) \right|^2 dt$$

Power - Energy

- The square root of the power is the root mean square (*rms*) value
  - This is a very important quantity as it is the most widespread measure of similarity/dissimilarity among signals
  - It is the basis for the definition of the Signal to Noise Ratio (SNR)

$$SNR = 20\log_{10}\left(\sqrt{\frac{P_{signal}}{P_{noise}}}\right)$$

  - It is such that a constant signal whose amplitude is =rms holds the same power content of the signal itself
- There exists signals for which neither the energy nor the power are finite

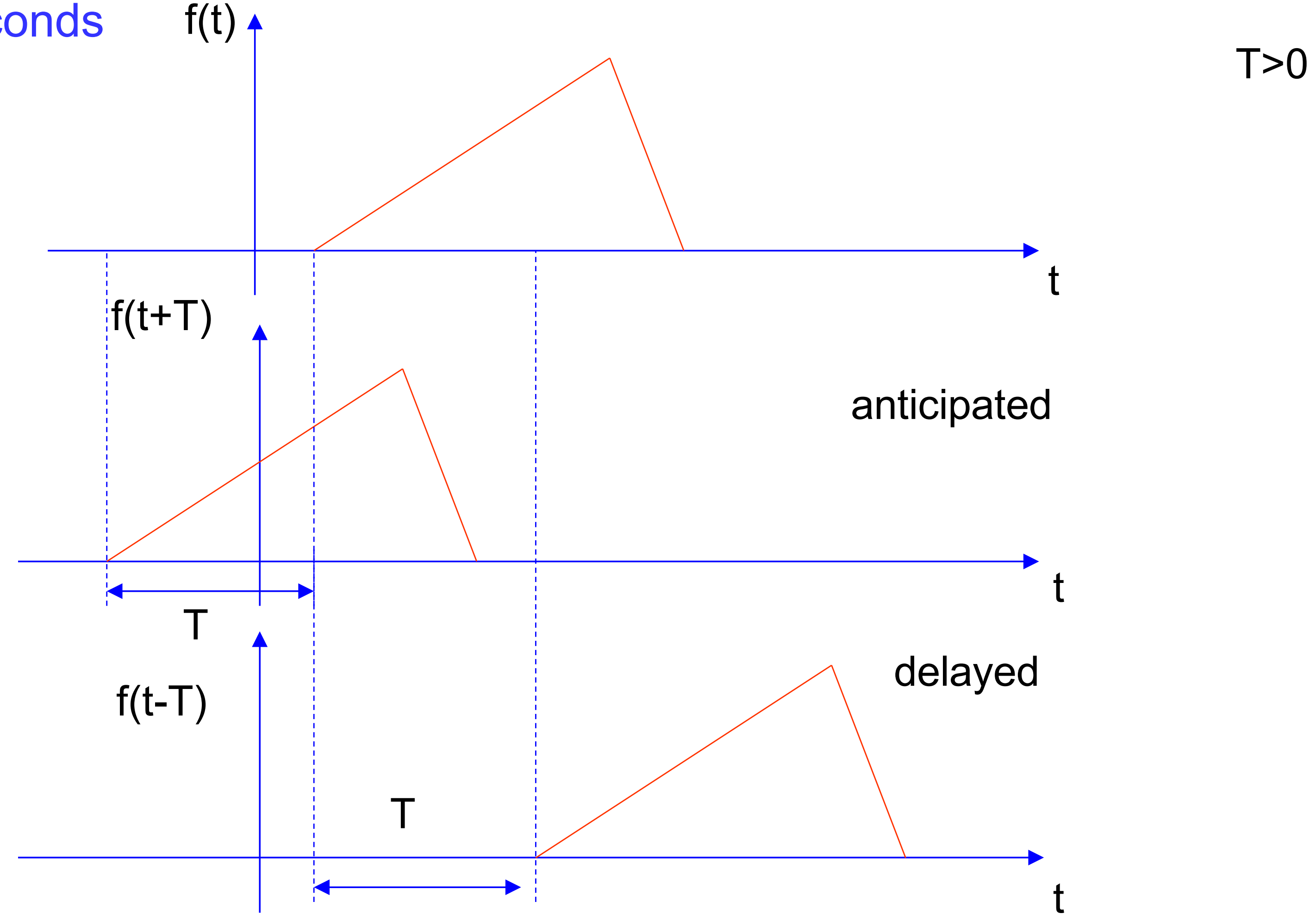


Energy and Power signals

- A signal with finite energy is an energy signal
  - Necessary condition for a signal to be of energy type is that the amplitude goes to zero as the independent variable tends to infinity
- A signal with finite and different from zero power is a power signal
  - The mean of an entity averaged over an infinite interval exists if either the entity is periodic or it has some statistical regularity
  - A power signal has infinite energy and an energy signal has zero power
  - There exist signals that are neither power nor energy, such as the ramp
- All practical signals have finite energy and thus are energy signals
  - It is impossible to generate a real power signal because this would have infinite duration and infinite energy, which is not doable.

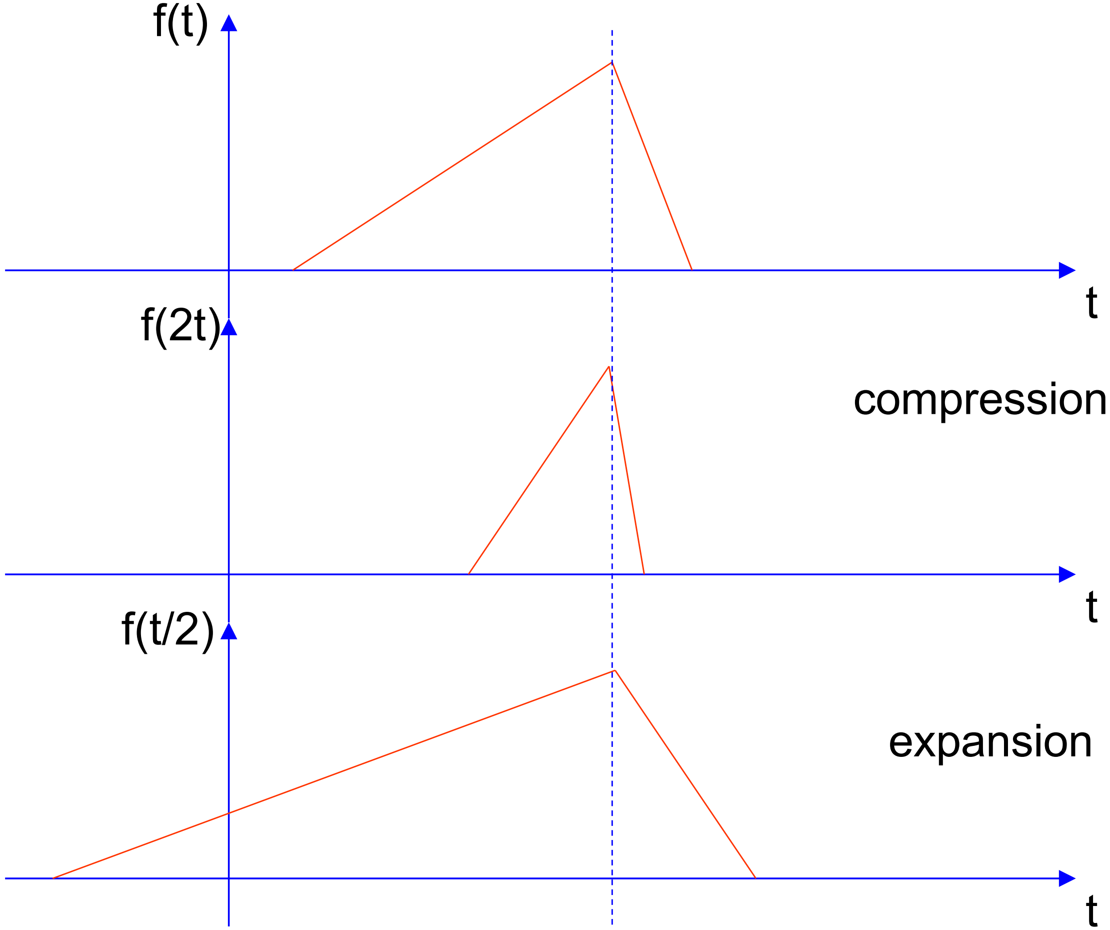
Useful signal operations: shifting, scaling, inversion

- **Shifting**: consider a signal f(t) and the same signal delayed/anticipated by T seconds



Useful signal operations: shifting, scaling, inversion

- (Time) Scaling: compression or expansion of a signal in time



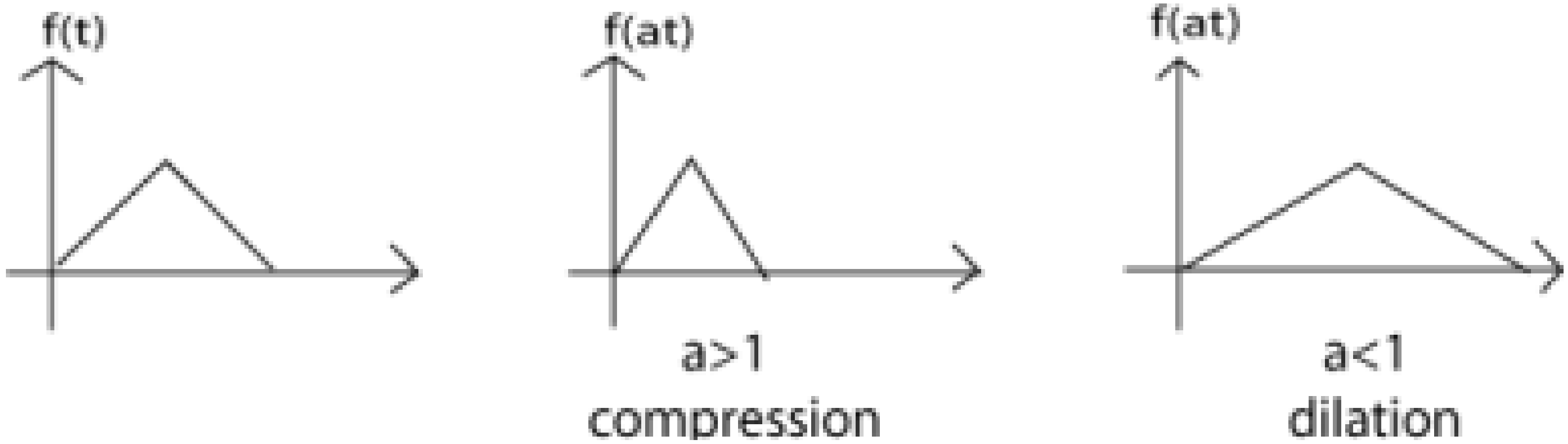
compression $\varphi(t) = f(2t)$

expansion $\varphi(t) = f(t/2)$

Useful signal operations: shifting, scaling, inversion

- Scaling: generalization

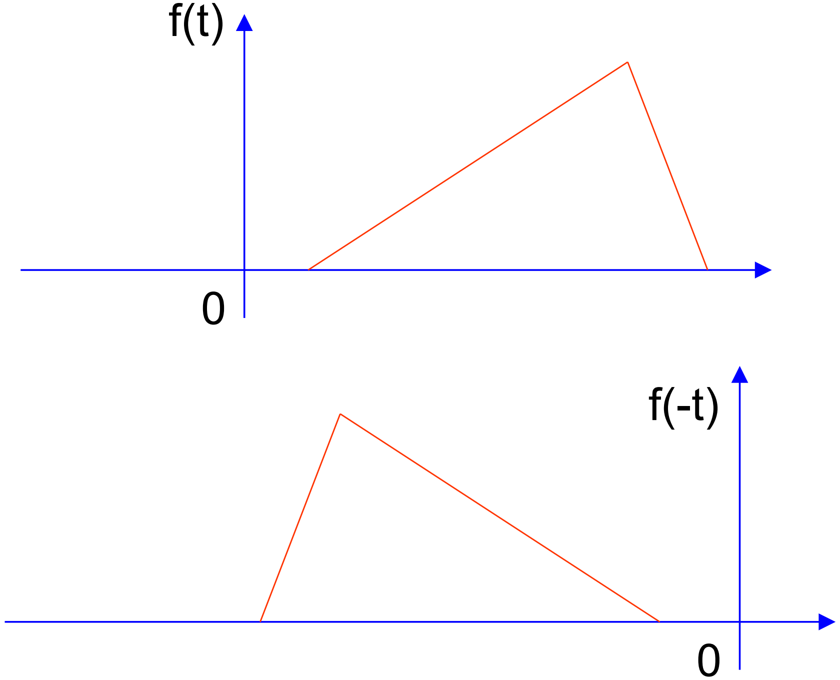
$$
\begin{aligned}
&a > 1 \\
&\varphi(t) = f(at) \;\rightarrow\; \text{compressed version} \\
&\varphi(t) = f\left(\frac{t}{a}\right) \rightarrow \text{dilated (or expanded) version} \\
&\text{Viceversa for } a < 1
\end{aligned}
$$

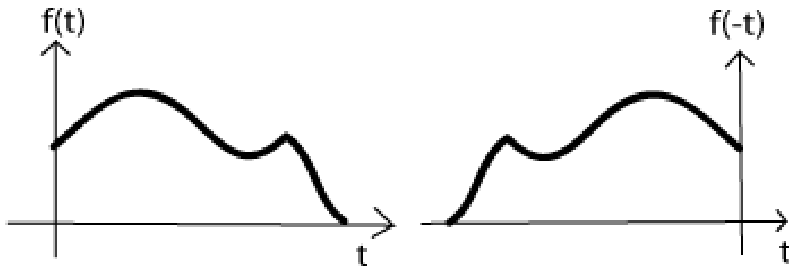


Useful signal operations: shifting, scaling, inversion

- (Time) inversion: mirror image of f(t) about the vertical axis

$$\varphi(t) = f(-t)$$





Useful signal operations: shifting, scaling, inversion qui

- Combined operations: $f(t) \rightarrow f(at\text{-}b)$
- Two possible sequences of operations

1. *Time shift f(t) by to obtain f(t-b). Now time scale the shifted signal f(t-b) by a to obtain f(at-b).*
2. *Time scale f(t) by a to obtain f(at). Now time shift f(at) by b/a to obtain f(at-b).*
   - *Note that you have to replace t by (t-b/a) to obtain f(at-b) from f(at) when replacing t by the translated argument (namely t-b/a))*

Useful functions

- The unit step function
  - Useful for representing causal signals

$$u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

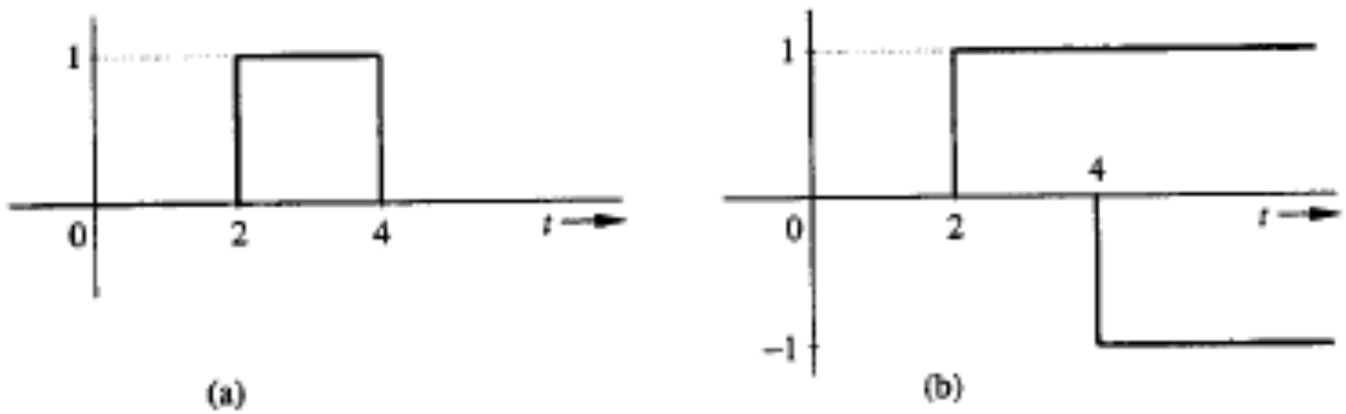


Fig. 1.15 Representation of a rectangular pulse by step functions.

$$f(t) = u(t-2) - u(t-4)$$

Unit step function

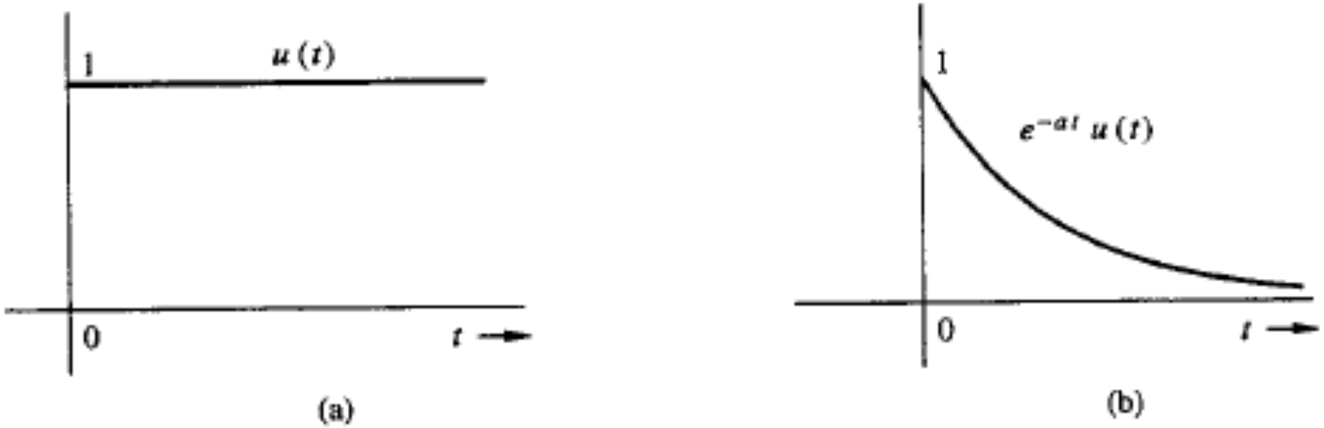
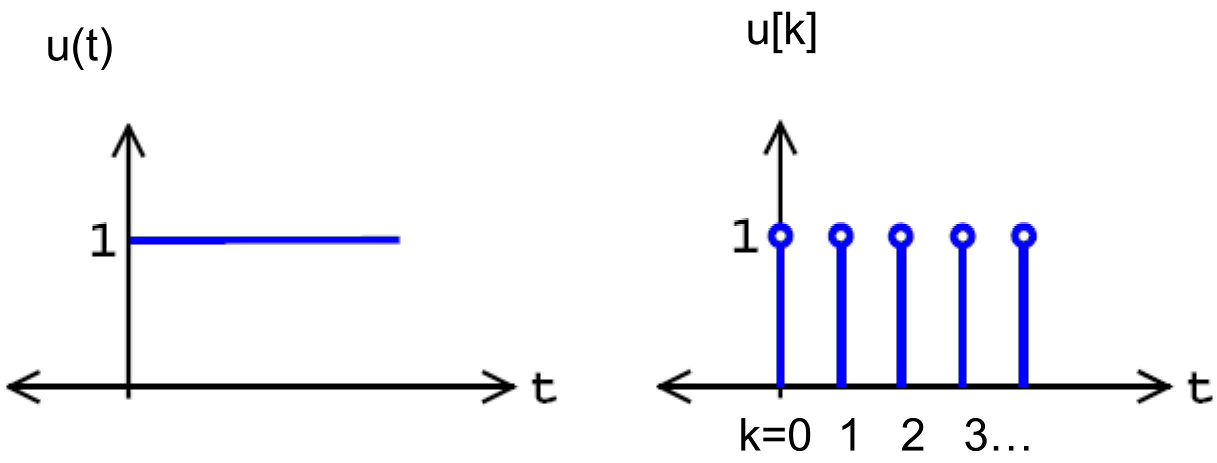


Fig. 1.14 (a) Unit step function $u(t)$ (b) exponential $e^{-at}u(t)$.

Useful functions

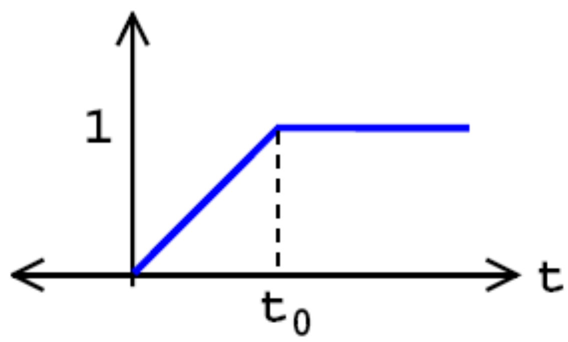
- Continuous and discrete time unit step functions



Useful functions

- The ramp function (continuous time)

$$r(t) = \begin{cases} 0 \text{ if } t < 0 \\ \frac{t}{t_0} \text{ if } 0 \leq t \leq t_0 \\ 1 \text{ if } t > t_0 \end{cases}$$

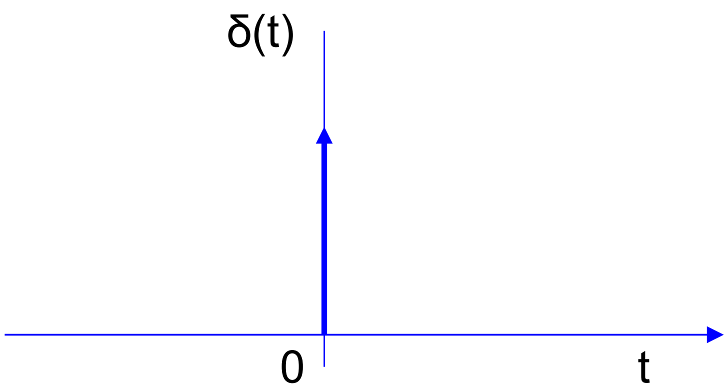


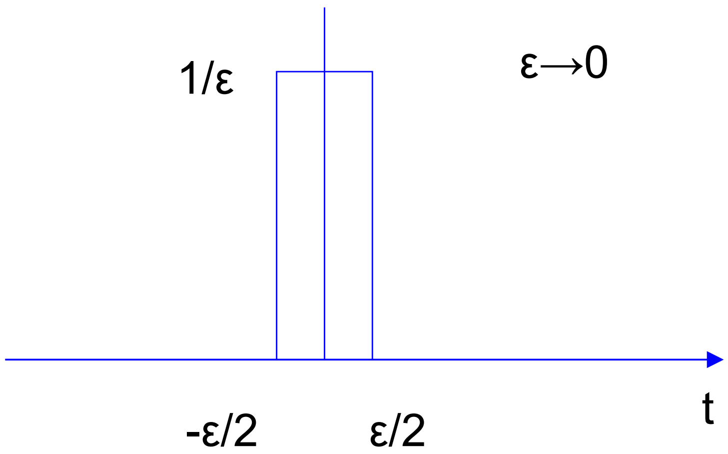
Useful functions

- The unit impulse function

$$\delta(t) = 0, t \neq 0$$

$$\int_{-\infty}^{+\infty} \delta(t)dt = 1$$





The function k δ(t) is zero for all t different from zero and has area equal to k

Properties of the unit impulse function

- Multiplication of a function by impulse

$$\phi(t)\delta(t) = \phi(0)\delta(t)$$
$$\phi(t)\delta(t-T) = \phi(T)\delta(t-T)$$

- Sampling property of the unit function

$$\int_{-\infty}^{+\infty} \phi(t)\delta(t)\,dt = \int_{-\infty}^{+\infty} \phi(0)\delta(t)\,dt = \phi(0)\int_{-\infty}^{+\infty} \delta(t)\,dt = \phi(0)$$

$$\int_{-\infty}^{+\infty} \phi(t)\delta(t-T)\,dt = \phi(T)$$

  – The area under the curve obtained by the product of the unit impulse function shifted by T and $\varphi(t)$ is the value of the function $\varphi(t)$ for $t=T$

Properties of the unit impulse function

- The unit step function is the integral of the unit impulse function

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

$$\int_{-\infty}^{t} \delta(t)\,dt = u(t)$$

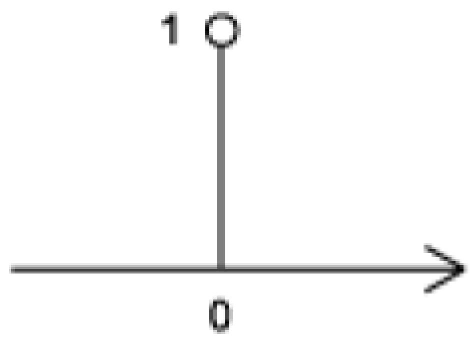
  - Thus

$$\int_{-\infty}^{t} \delta(t)\,dt = u(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

Properties of the unit impulse function

- Discrete time impulse function

$$\delta[n] = \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ otherwise} \end{cases}$$



Useful functions

- The continuous time complex exponential

$$f(t) = Ae^{j\omega t}$$

- Euler's relations

$$Ae^{j\omega t} = Acos(\omega t) + j(Asin(\omega t))$$

$$cos(\omega t) = \frac{e^{j\omega t} + e^{-(j\omega t)}}{2}$$

$$sin(\omega t) = \frac{e^{j\omega t} - e^{-(j\omega t)}}{2j}$$

$$e^{j\omega t} = cos(\omega t) + jsin(\omega t)$$

- Discrete time complex exponential
  - k=nT

$$\begin{aligned} f[n] &= Be^{snT} \\ &= Be^{j\omega nT} \end{aligned}$$

Useful functions

- Exponential function e^{st}
  - Generalization of the function $e^{j\omega t}$

$$s = \sigma + j\omega$$

Therefore

$$e^{st} = e^{(\sigma+j\omega)t} = e^{\sigma t}e^{j\omega t} = e^{\sigma t}(\cos \omega t + j \sin \omega t) \tag{1.30a}$$

If $s^* = \sigma - j\omega$ (the conjugate of s), then

$$e^{s^* t} = e^{\sigma - j\omega} = e^{\sigma t}e^{-j\omega t} = e^{\sigma t}(\cos \omega t - j \sin \omega t) \tag{1.30b}$$

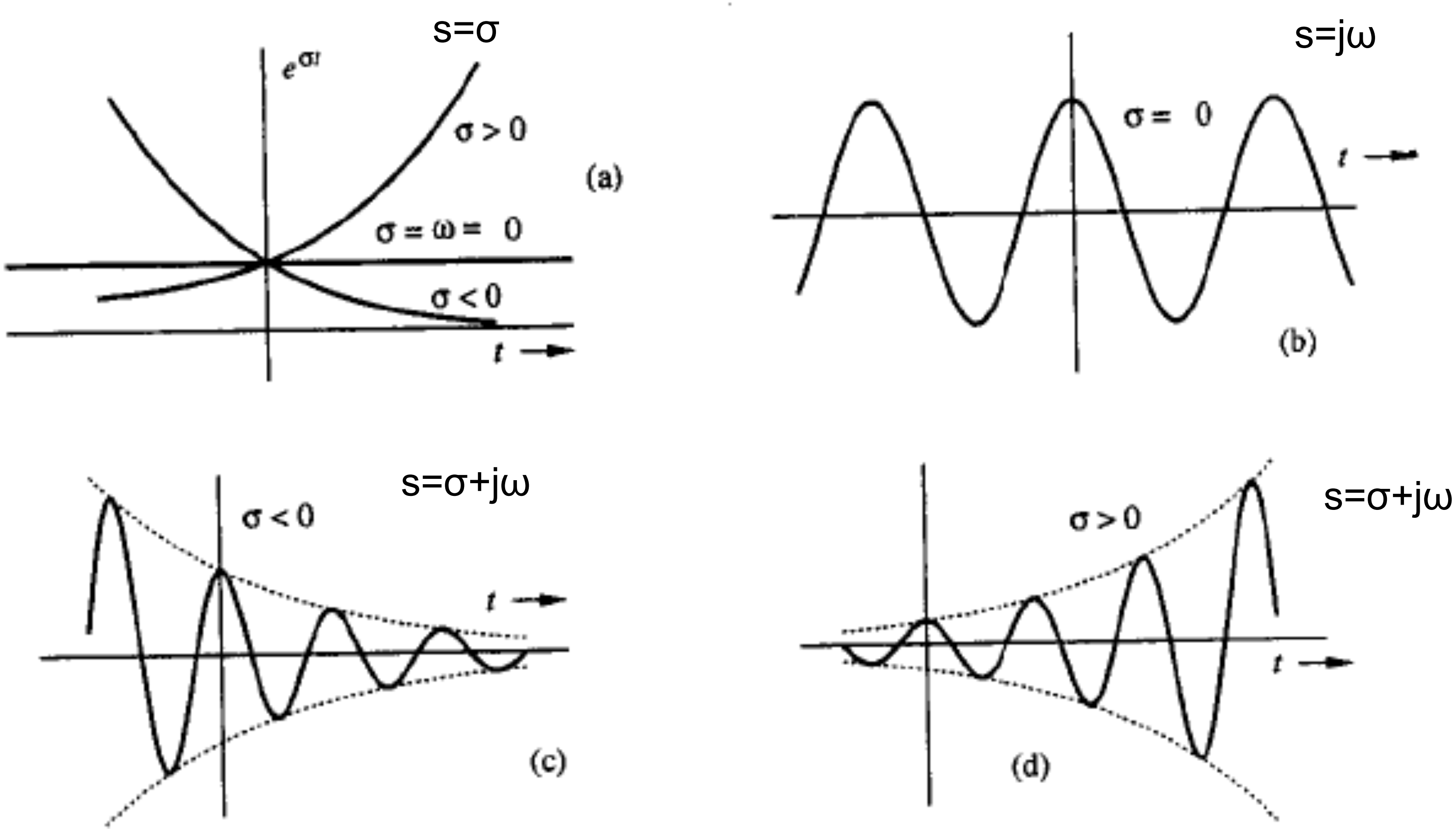
and

$$e^{\sigma t} \cos \omega t = \frac{1}{2}(e^{st} + e^{s^* t}) \tag{1.30c}$$

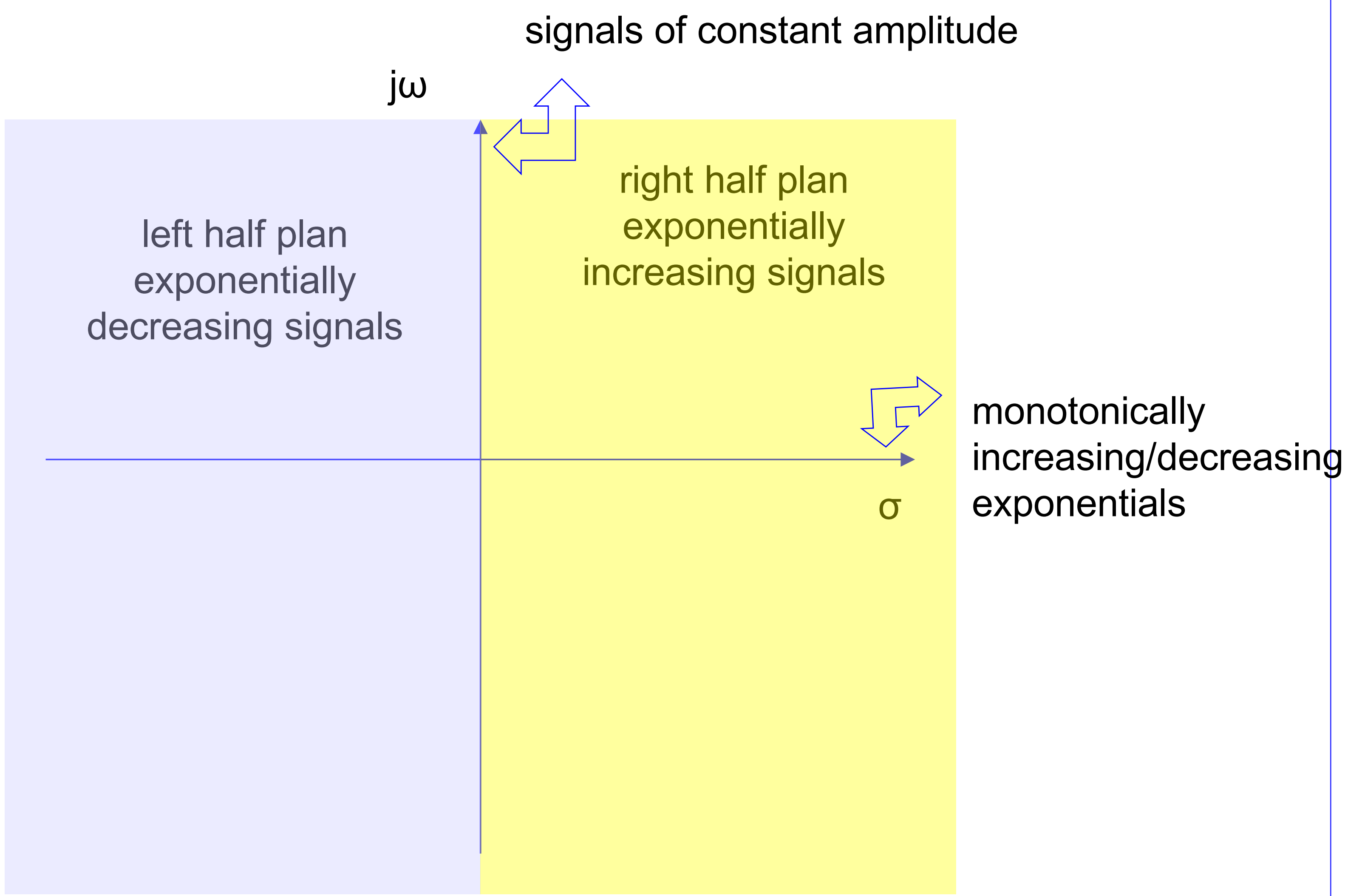
The exponential function

- Special cases
- A constant $K = K \times \exp 0$
- A monotonic exponential $\exp^{\sigma t}(\omega = 0, s = \sigma)$
- A trigonometric function $\cos(\omega t).(s = 0, \sigma = +/- j\omega)$
- An exponentially varying trigonometric function

The exponential function



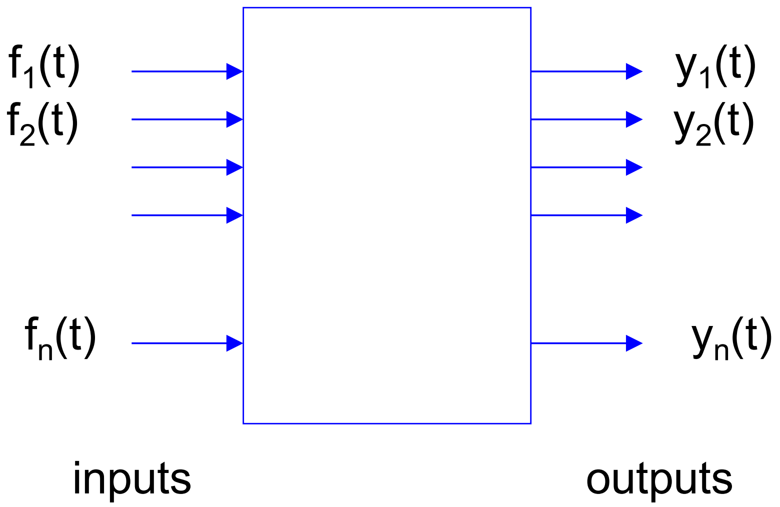
Complex frequency plan



Linear Systems

Systems

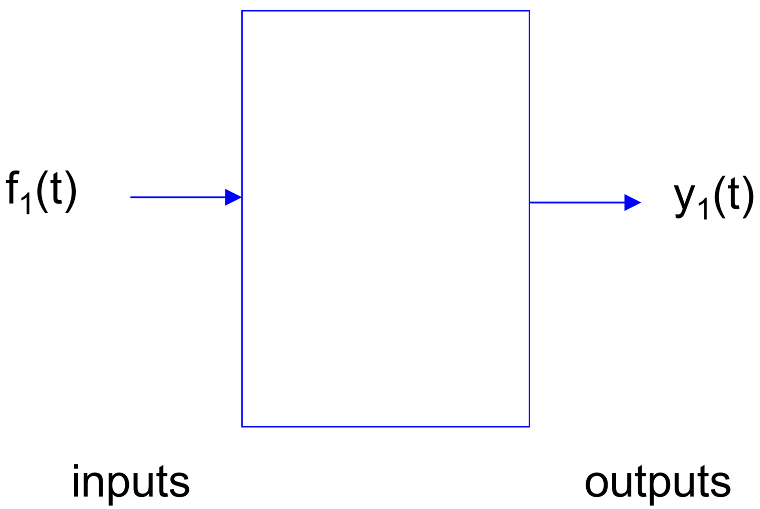
- A system is characterized by
  - inputs
  - outputs
  - rules of operation (mathematical model of the system)

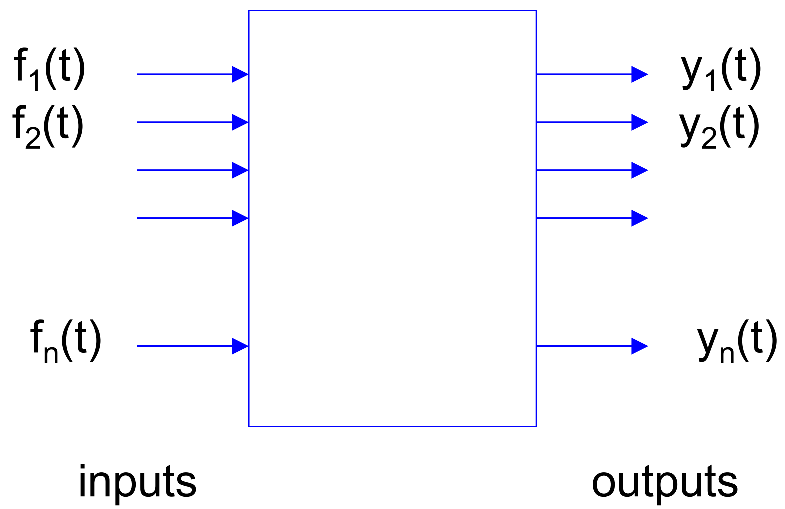


Systems

- Study of systems: mathematical modeling, analysis, design
  - Analysis: how to determine the system output given the input and the system mathematical model
  - Design or synthesis: how to design a system that will produce the desired set of outputs for given inputs

- SISO: single input single output - MIMO: multiple input multiple output





Classification of systems

1. Linear and non linear
2. Constant parameters and time-varying parameters
3. Instantaneous (memoryless) and dynamic (with memory)
4. Causal and non-causal
5. Lumped-parameters and distributed-parameters
6. Continuous-time and discrete-time
7. Analog and digital

Linear and non-linear systems

- Additivity

$$f_1 \rightarrow y_1 \quad \text{and} \quad f_2 \rightarrow y_2 \qquad \text{then} \qquad f_1 + f_2 \rightarrow y_1 + y_2$$

- Homogeneity (scaling)

$$f_1 \rightarrow y_1 \qquad \text{then} \qquad a_1 \times f_1 \rightarrow a_1 \times y_1$$

- Superposition

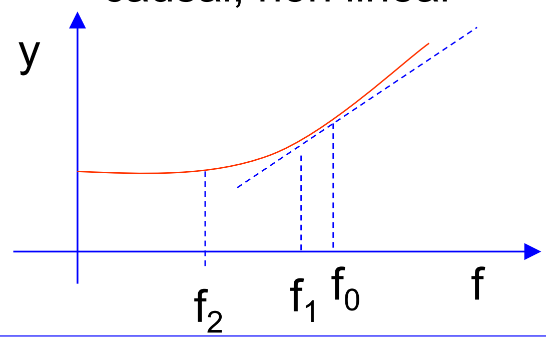
$$a_1 \times f_1 + a_2 \times f_2 \rightarrow a_1 \times y_1 + a_2 \times y_2$$

Response of a linear system

- Total response = Zero-input response + Zero-state response
  - The output of a system for t≥0 is the result of two independent causes: the initial conditions of the system (or system state) at t=0 and the input f(t) for t≥0.
  - **Because of linearity**, the total response is the **sum** of the responses due to those two causes
  - The **zero-input** response is only due to the initial conditions and the **zero-state** response is only due to the input signal
  - This is called decomposition property
- Real systems are *locally* linear
  - Respond linearly to small signals and non-linearly to large signals

causal, linear

causal, non linear



locally linear around f_0

Response of a linear system

- Because of linearity we can evaluate separately the two components of the output and calculate the total response as their sum.
  - The zero input component can be computed assuming the input to be zero and
  - the zero-state component can be computed assuming the initial conditions to be zero
- If the input can be expressed as the sum of simple components, the output can be obtained as the sum of the responses to each component

$$f(t) = a_1 \times f_1(t) + a_2 \times f_2(t) + \cdots + a_n \times f_n(t)$$

$$y(t) = a_1 \times y_1(t) + a_2 \times y_2(t) + \cdots + a_n \times y_n(t)$$

Where $y_k(t)$ is the response to $f_k(t)$.

Time-invariant and time-varying parameters

Systems whose parameters do not change with time are **time-invariant** (also **constant-parameter**) systems. For such a system, if the input is delayed by T seconds, the output is the same as before but delayed by T (assuming identical initial conditions). This property is expressed graphically in Fig. 1.28.

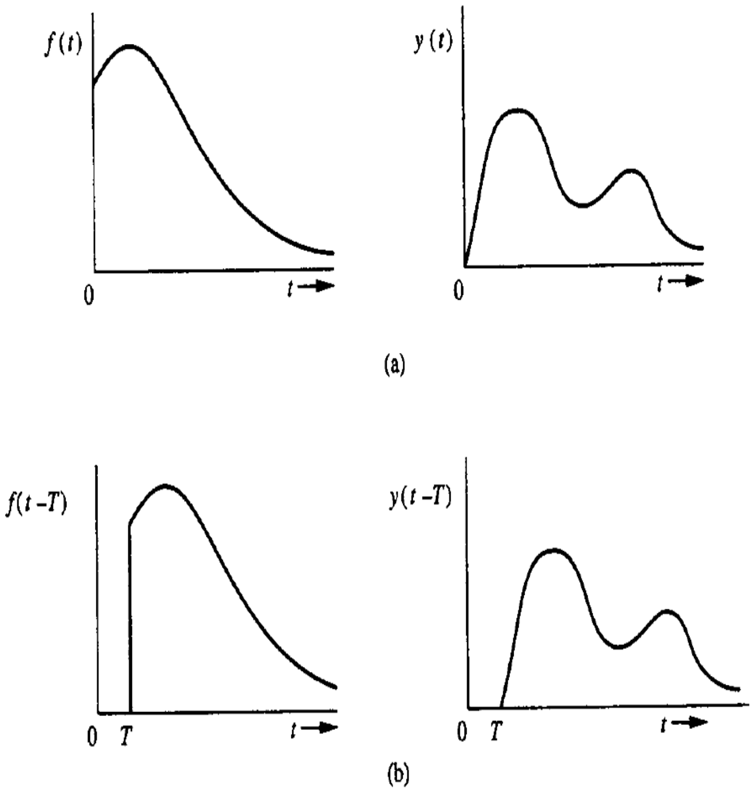


Fig. 1.28 Time-invariance property.

Instantaneous and dynamic systems

- Instantaneous (memoryless): the output at time t only depends on the input at time t
  - Special case of dynamic systems
- Dynamic: the output does not only depend on the input at the same instant
  - Finite memory: the output is completely determined by the input over the last T seconds

Istantaneous and dynamic systems

As observed earlier, a system's output at any instant t generally depends upon the entire past input. However, in a special class of systems, the output at any instant t depends only on its input at that instant. In resistive networks, for example, any output of the network at some instant t depends only on the input at the instant t. In these systems, past history is irrelevant in determining the response. Such systems are said to be **instantaneous** or **memoryless** systems. More precisely, a system is said to be instantaneous (or memoryless) if its output at any instant t depends, at most, on the strength of its input(s) at the same instant but not on any past or future values of the input(s). Otherwise, the system is said to be **dynamic** (or a system with memory). A system whose response at t is completely determined by the input signals over the past T seconds [interval from $(t - T)$ to t] is a **finite-memory system** with a memory of T seconds. Networks containing inductive and capacitive elements generally have infinite memory because the response of such networks at any instant t is determined by their inputs over the entire past $(-\infty, t)$. This is true for the RC circuit of Fig. 1.26.

Causal and non-causal systems

A **causal** (also known as a **physical** or **non-anticipative**) system is one for which the output at any instant t_0 depends only on the value of the input $f(t)$ for $t \leq t_0$. In other words, the value of the output at the present instant depends only on the past and present values of the input $f(t)$, not on its future values. To put it simply, in a causal system the output cannot start before the input is applied. If the response starts before the input, it means that the system knows the input in the future and acts on this knowledge before the input is applied. A system that violates the condition of causality is called a **noncausal** (or **anticipative**) system.

Any practical system that operates in real time† must necessarily be causal.

- Why to study non-causal systems?
  - Pre-recorded data
  - Other variable than time can be considered (es. Images)

Lumped and distributed parameters

- Lumped parameters: The system parameters are assumed to be functions of time alone. In consequence, the system equations require only one independent variable (time) and therefore are ordinary differential equations
  - This applies when the system's size is small with respect to the wavelength of the signal that propagates through the system
    - Electric current in a circuit
- Distributed parameters: The system parameters depends on both time and space leading to mathematical models that are partial differential equations
  - Examples: microwave and antennas where the system's spatial dimension is not negligible with respect to the signal wavelength
- Here we focus on lumped-parameters systems

Continuous-time and discrete-time

Distinction between discrete-time and continuous-time signals is discussed in Sec. 1.2-1. Systems whose inputs and outputs are continuous-time signals are **continuous-time systems**. On the other hand, systems whose inputs and outputs are discrete-time signals are **discrete-time systems**. If a continuous-time signal is sampled, the resulting signal is a discrete-time signal. We can process a continuous-time signal by processing its samples with a discrete-time system.

1.7-7 Analog and Digital Systems

Analog and digital signals are discussed in Sec. 1.2-2. A system whose input and output signals are analog is an **analog system**; a system whose input and output signals are digital is a **digital system**. A digital computer is an example of a digital (binary) system. Observe that a digital computer is an example of a system that is digital as well as discrete-time.

Invertible and non-invertible systems

- A signal S performs some operations on the input signal to generate the output signal. If we can obtain the input back from the output the system is said to be invertible
  - For this to be true, a given output must correspond to a given input (one-to-one relation)
  - Example of non-invertible system: the rectifier, the differentiator. For this last invertibility can be recovered by knowing one boundary condition

System model: input-output description

- The system model is a mathematical expression of a rule that satisfactorily approximates the dynamical behavior of a system
- Internal description
  - The internal structure of the system is known so that we can write equations that describe such a structure
- External description
  - The system is viewed from the input and output terminals and is considered as a black-box. In this case the system model only describes the relation between the input and the output signals irrespectively of the internal structure
  - Implicitly assumes controllability and observability

Cannot be controlled by the input

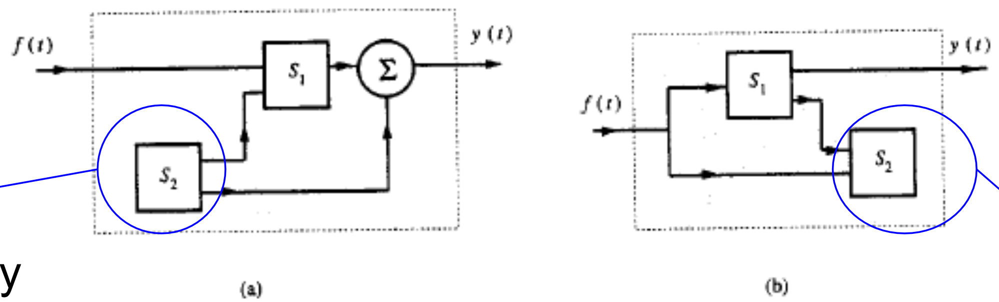


Fig. 1.35 Structures of uncontrollable and unobservable systems.

Cannot be observed from the output

Summary

- A signal is a set of information of data.
- A system processes input signals to produce output signals (response)
  - Can be a physical system or an algorothm
- Signal can be classified in different ways that are not mutually exclusive
- The linearity property implies superposition such that the response to different causes can be obtained as the sum of the responses to each cause that is calculated *as if* the others were not present