POLITECNICO
MILANO 1863

Multimedia Signal Processing
1st Module
Fundamentals of Multimedia Signal Processing

Marco Marcon

Dipartimento di Elettronica,Informazione e Bioingegneria (DEIB)

Course overview

This module covers the fundamental tools for digital signal processing. In particular it addresses topics of signal analysis and filtering concerning audio and video data.

Some aspects of Statistical Signal Processing will also be presented during the course.

Participants will learn basic tools for digital signal filtering and analysis acquiring the knowledge to design specific filters and . In particular the program will be articulated in the following parts:

- **Review of Analog Signal Processing theory:**
  - Some fundamentals of Analog Signal Processing will be recalled (sampling, quantization, Fourier Transform and Series, noise description…).
- **Introduction to Discrete Signals Transforms:**
  - The z-transform, the Discrete Time Fourier Transform and the Discrete Fourier Transform will be described with their properties and application to filter design and analysis.
- **Introduction to Digital Filters:**
  - This course part will cover Time-domain filter representations, transfer function analysis, frequency response analysis, Finite and Infinite impulse response implementation; stability analysis.

Course overview

- **Windowing and Short Time Fourier Transform:**
  - Overview of windows and real-time processing; overlap and add, overlap and save, Short Time Fourier Transform for real time processing
- **Introduction to Multirate Processing:**
  - Downsampling, upsampling, polyphase filters, perfect reconstruction filter banks.
- **Elements of 2D Signal Processing:**
  - Image Processing, filtering, adjustment, histogram processing.

- **Laboratory activities and lectures examples**
  - The proposed examples of applications and exercises, cover audio, and image processing.
  - It will be adopted Matlab®, Simulink® and the related toolboxes (Digital Signal Processing System Toolbox, Audio System Toolbox, Image acquisition toolbox, Image Processing Toolbox).

Course overview

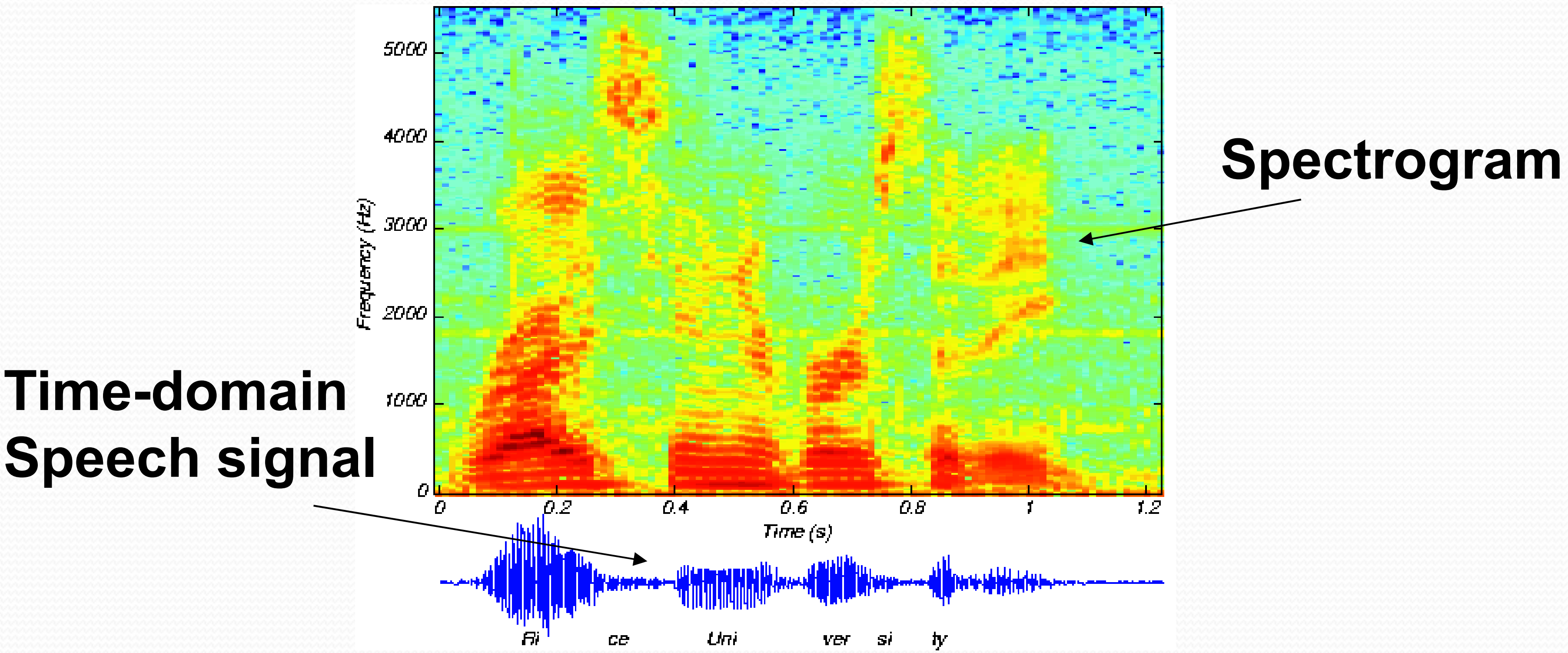
- **Website**
  - **marcon.net**
- **E-mail**
  - **marco.marcon@polimi.it**

Suggested books

- "Digital Signal Processing: System Analysis and Design", *Paulo S. R. Diniz, Eduardo A. B. da Silva, Sergio L. Netto*
- DAFX: Digital Audio Effects, 2nd Edition, Editor: Udo Zolzer
- "Lecture notes", *Marco Tagliasacchi* (Available on the website).
- Matlab, Simulink and Toolboxes online guides and example on the Mathworks website.

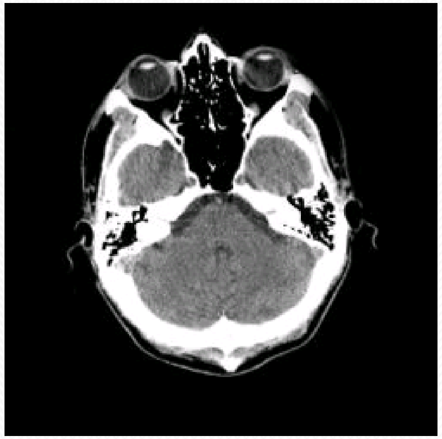
Signals

- Signals are functions of independent variables that carry information. For example:
  - Electrical signals --- voltages and currents in a circuit
  - Acoustic signals --- audio or speech signals (analog or digital)



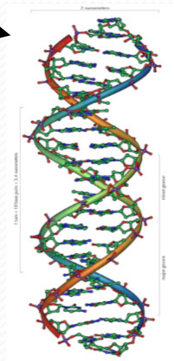
Signals

- Video signals --- intensity variations in an image (e.g. a CAT scan)



- Biological signals --- sequence of bases in a gene

DNA



Its signal representation

THE INDEPENDENT VARIABLES

What are the independent variables in these signals?

(i) Speech -> Time

(ii) CT scan image -> Space

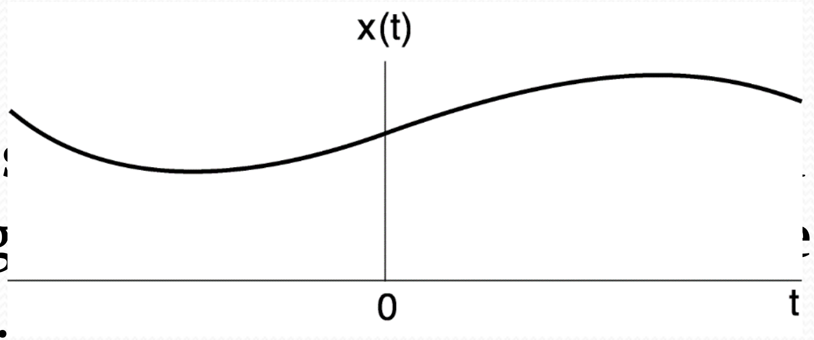
(iii) DNA sequence -> Position along the DNA helix

For this course: Focus on a single (1-D) independent variable which we call "time". Some elements will be provided for 2-D signals.

Continuous-Time (CT) signals: $x(t)$, t — continuous values

Discrete-Time (DT) signals: $x[n]$, n — integer values only

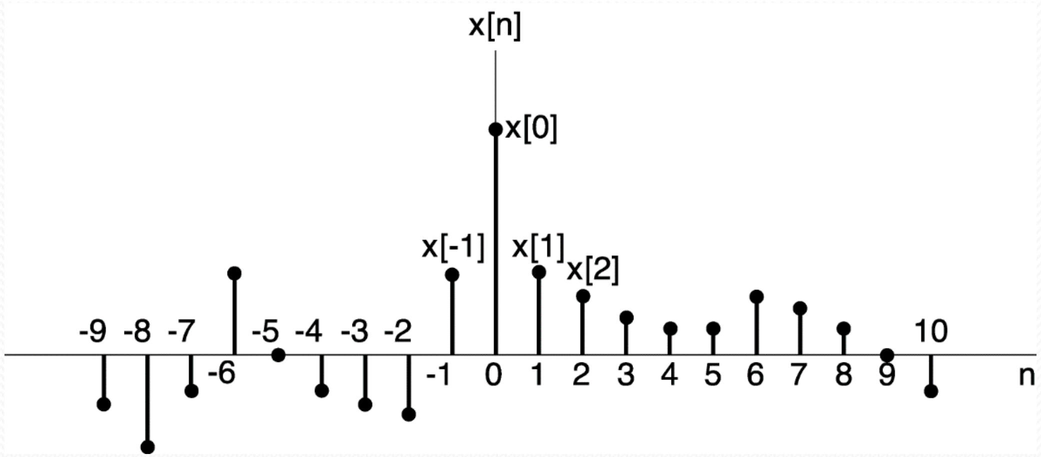
CT Signals

- Most of the s


 are CT signals—E.g ., temperature, velocity, etc.

DT Signals

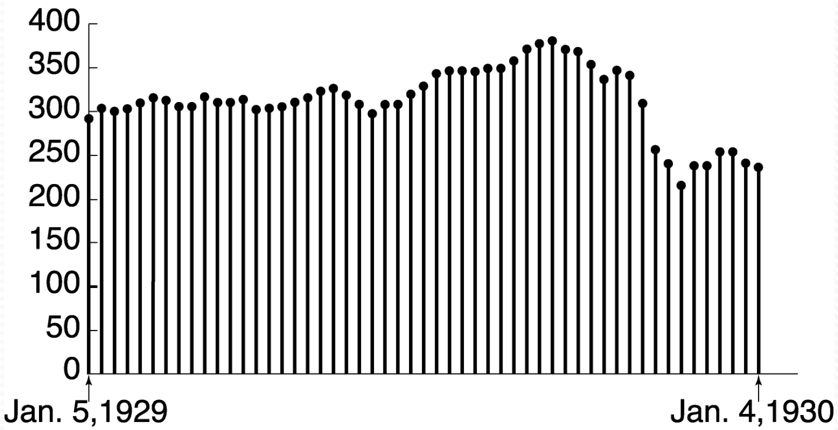
- ***x*[*n*], *n* — integer, time varies discretely**



- Examples of DT signals in nature:
  — DNA base sequence
  — Population of the *n-th* generation of certain species
  ⋮

Many human-made signals are DT Signals

Ex.#1 Weekly Dow-Jones industrial average



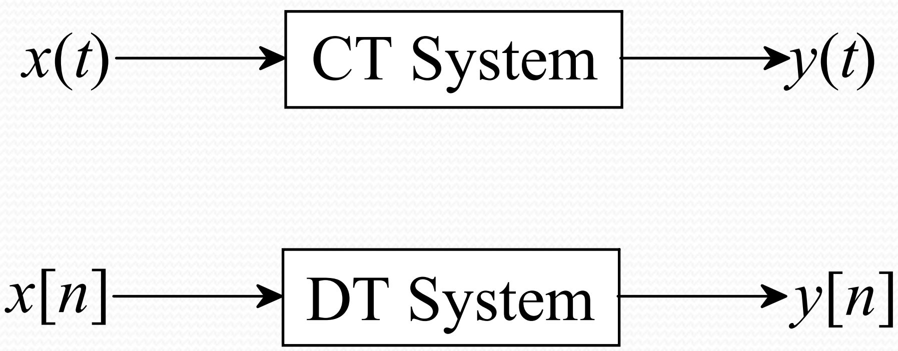
Ex.#2 digital image

Why DT? —
Can be processed by modern digital computers and digital signal processors (DSPs).

SYSTEMS

For the most part, our view of systems will be from an input-output perspective:

A system responds to applied input signals, and its response is described in terms of one or more output signals



EXAMPLES OF SYSTEMS

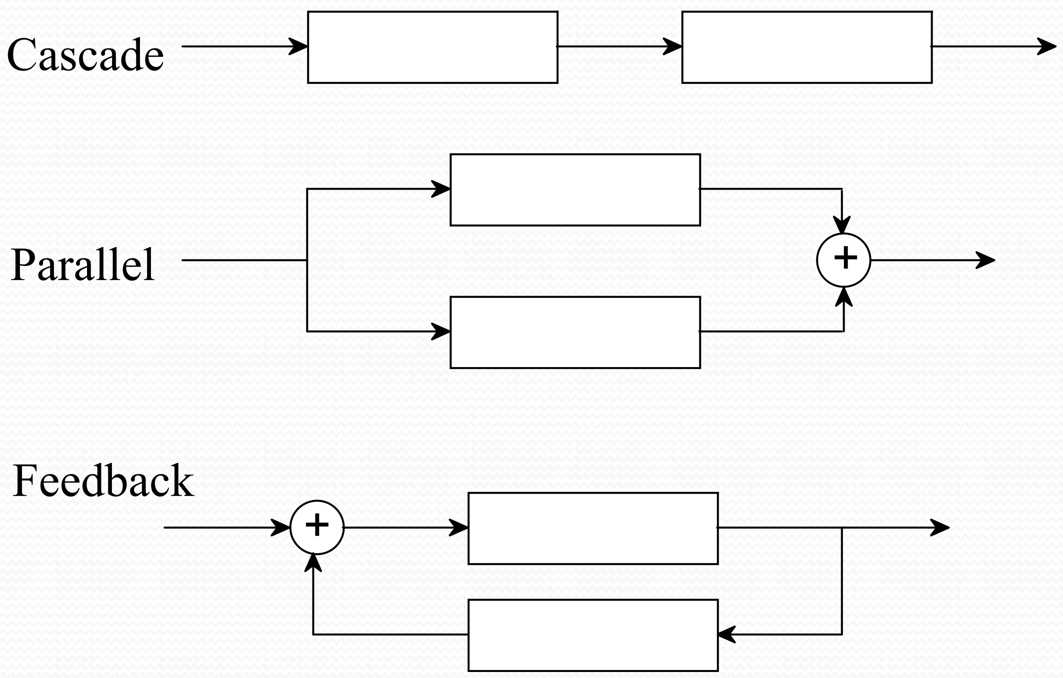
- Dynamics of an aircraft or space vehicle
- An algorithm for analyzing financial and economic factors to predict bond prices
- Audio filters
- An algorithm for post-flight analysis of a space launch
- An edge detection algorithm for medical images

What are the inputs and what are the outputs in above examples?

SYSTEM INTERCONNECTIONS

- An important concept is that of interconnecting systems
  - To build more complex systems by interconnecting simpler subsystems
  - To modify response of a system

- Signal flow (Block) diagram

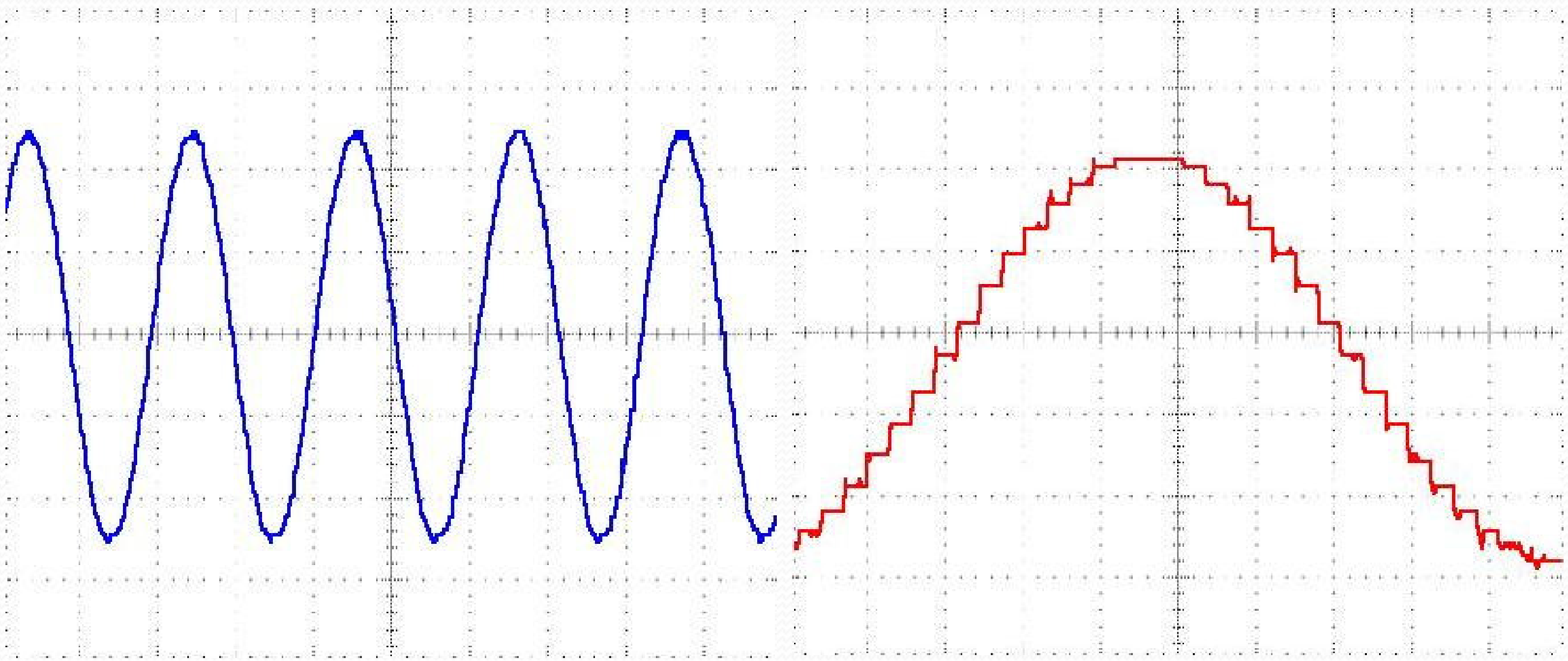


Moving from Analog to Digital World

- What is DSP?
- Converting Analog into Digital
  - Electronically
  - Computationally
- How Does It Work?
  - Faithful Duplication
  - Resolution Trade-offs

What is DSP?

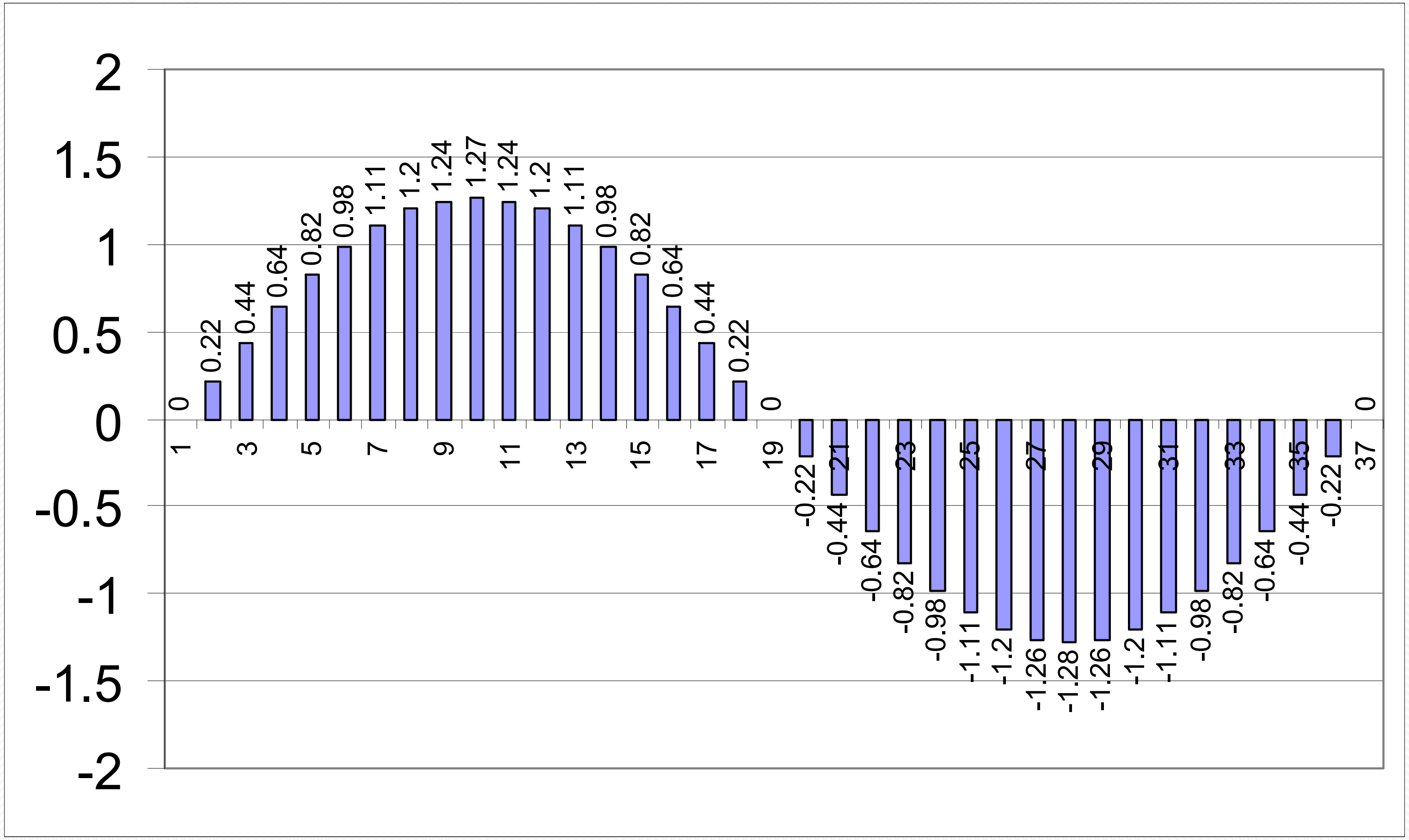
- Converting a continuously changing waveform (analog) into a series of discrete levels (digital)



What is DSP?

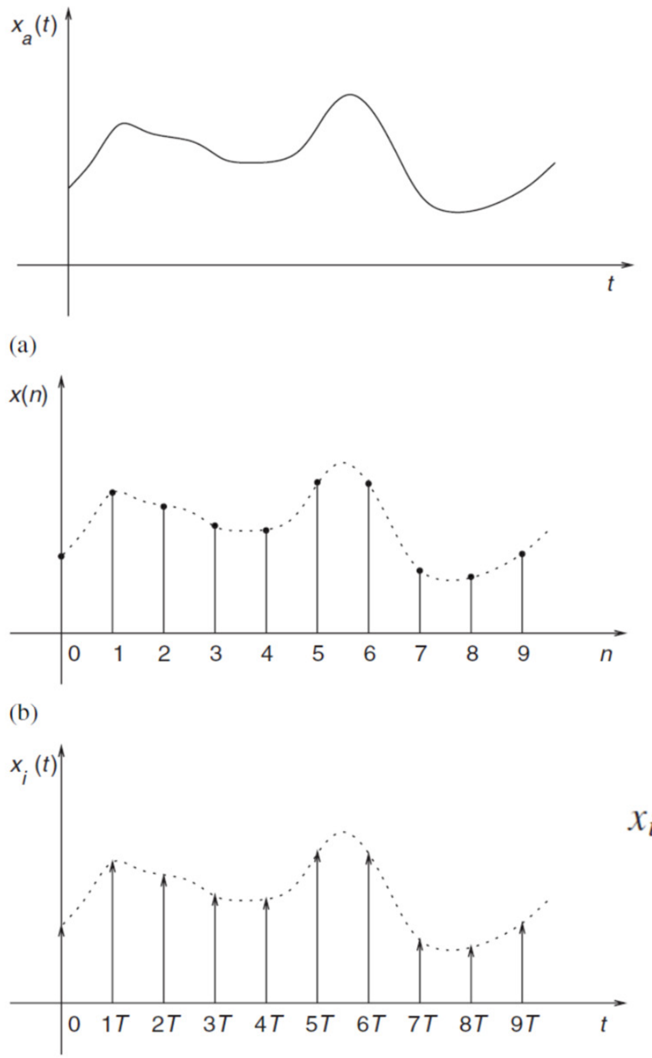
- The analog waveform is sliced into equal segments and the waveform amplitude is measured in the middle of each segment
- The collection of measurements make up the digital representation of the waveform

What is DSP?



2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
0.22
0.44
0.64
0.82
0.98
1.11
1.2
1.24
1.27
1.24
1.2
1.11
0.98
0.82
0.64
0.44
0.22
0
-0.22
-0.44
-0.64
-0.82
-0.98
-1.11
-1.2
-1.26
-1.28
-1.26
-1.2
-1.11
-0.98
-0.82
-0.64
-0.44
-0.22
0
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37

Sampling in time domain



$$x_i(t) = \sum_{n=-\infty}^{\infty} x(n)\delta(t - nT)$$

$$x_i(t) = \sum_{n=-\infty}^{\infty} x_a(nT)\delta(t - nT) = x_a(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) = x_a(t)p(t)$$

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$$

The convolution theorem

- Theorem: $x(t)*h(t) \leftrightarrow X(\omega)H(\omega)$
- Demonstration:

$$y(t)=x(t)*h(t)=\int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau$$

$$Y(\omega)=\int_{-\infty}^{\infty} y(t)e^{-j\omega t}dt=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau e^{-j\omega t}dt=$$

$$=\int_{-\infty}^{\infty} x(\tau)\int_{-\infty}^{\infty} h(t-\tau)e^{-j\omega t}dtd\tau \quad \text{let} \quad u=t-\tau$$

$$=\int_{-\infty}^{\infty} x(\tau)\int_{-\infty}^{\infty} h(u)e^{-j\omega(u+\tau)}dud\tau=\int_{-\infty}^{\infty} x(\tau)\int_{-\infty}^{\infty} h(u)e^{-j\omega u}e^{-j\omega\tau}dud\tau=$$

$$=\int_{-\infty}^{\infty} x(\tau)e^{-j\omega\tau}d\tau\int_{-\infty}^{\infty} h(u)e^{-j\omega u}du=X(\omega)H(\omega)$$

Effect of sampling in the frequency domain

- The multiplication in time domain corresponds to convolution in frequency domain and viceversa:
- The convolution in time domain corresponds to the multiplication in the frequency domain.
- Since sampling in time domain is equivalent to multiply the continuous signal with a train of impulses, then, in the frequency domain, sampling is equivalent to the convolution of the Fourier transform of the continuous signal with the Fourier transform of the train of impulses.

Effect of sampling in the frequency domain

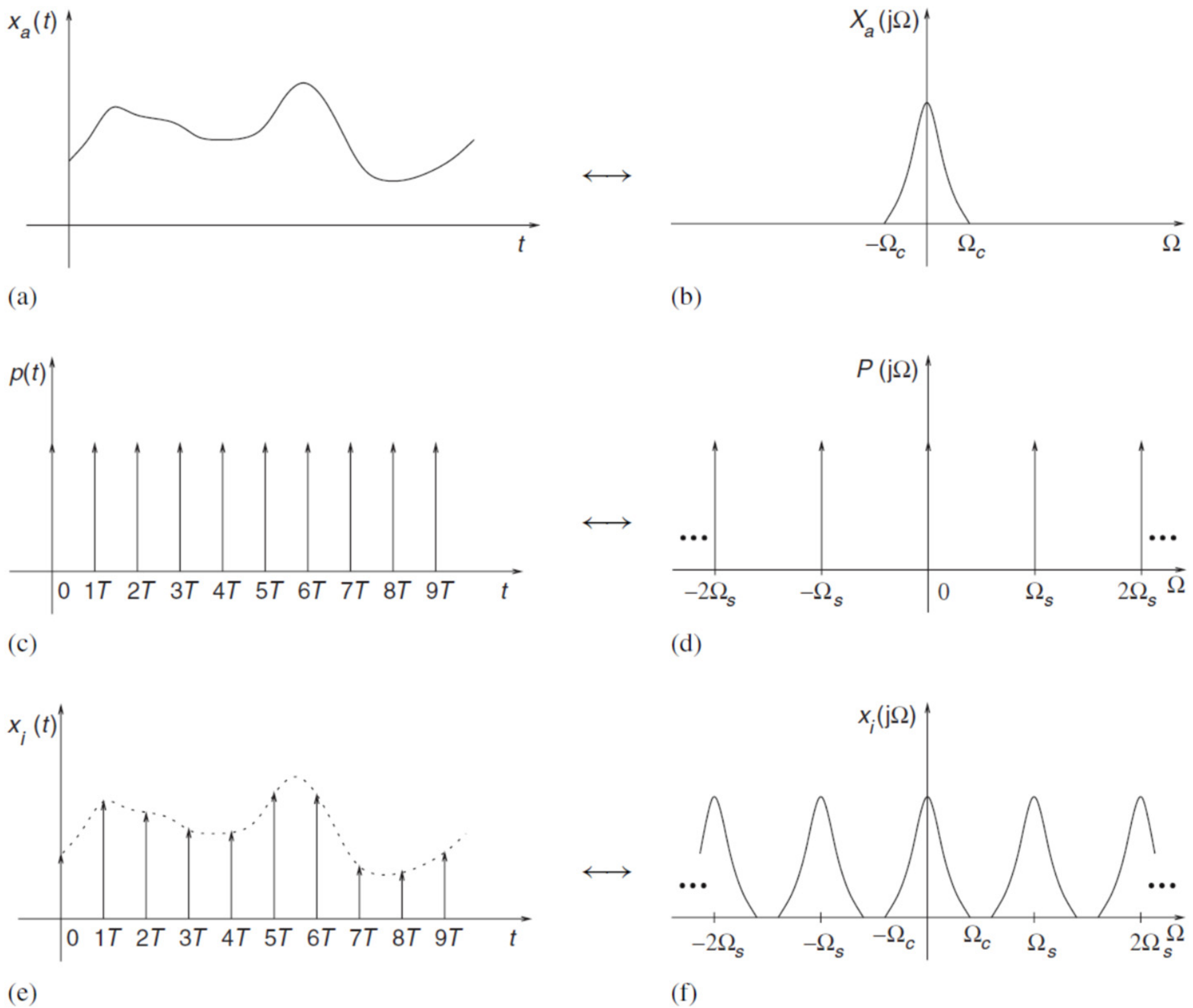
$$X_i(j\Omega) = \frac{1}{2\pi} X_a(j\Omega) * P(j\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(j\Omega - j\Omega') P(j\Omega') d\Omega'$$

- The Fourier Transform of a train of impulses in time domain is a train of impulses in the frequency domain.

$$P(j\Omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta\left(\Omega - \frac{2\pi}{T}k\right)$$

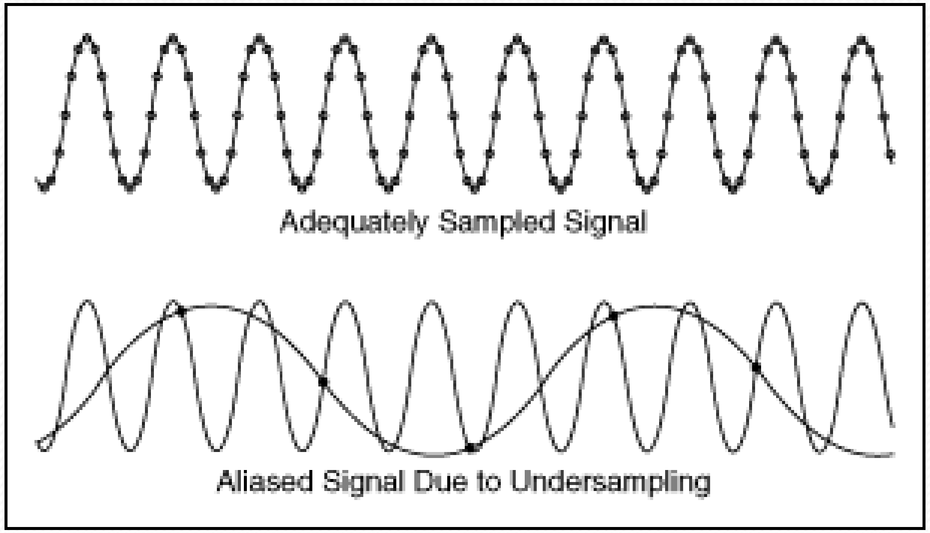
$$\begin{aligned} X_i(j\Omega) &= \frac{1}{2\pi} X_a(j\Omega) * P(j\Omega) \\ &= \frac{1}{T} X_a(j\Omega) * \sum_{k=-\infty}^{\infty} \delta\left(\Omega - \frac{2\pi}{T}k\right) \\ &= \frac{1}{T} \sum_{k=-\infty}^{\infty} X_a\left(j\Omega - j\frac{2\pi}{T}k\right) \end{aligned}$$

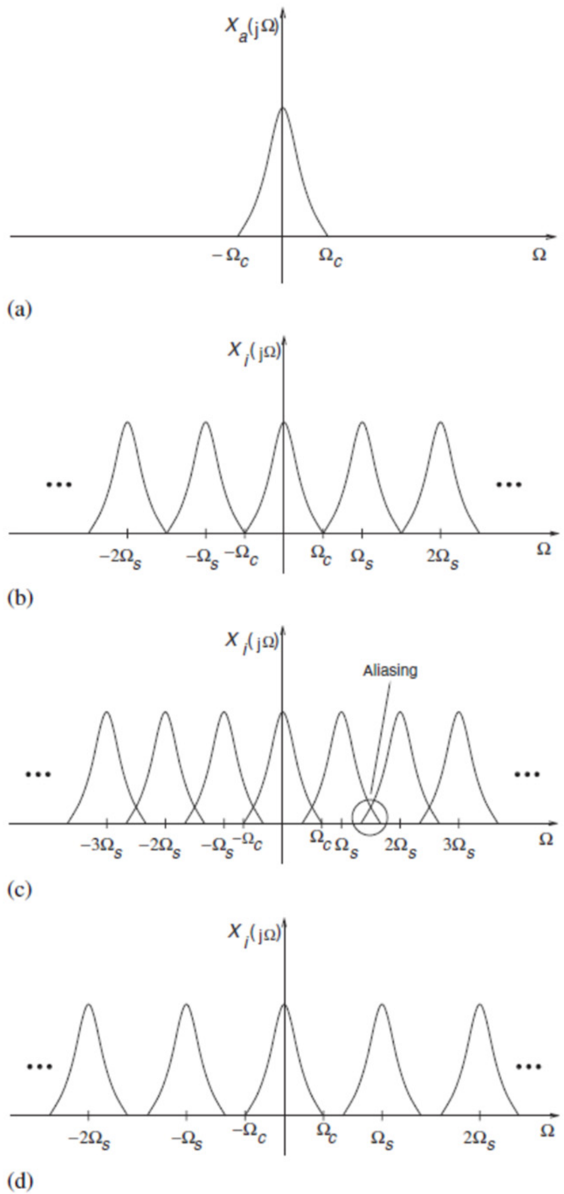
Sampling in frequency domain



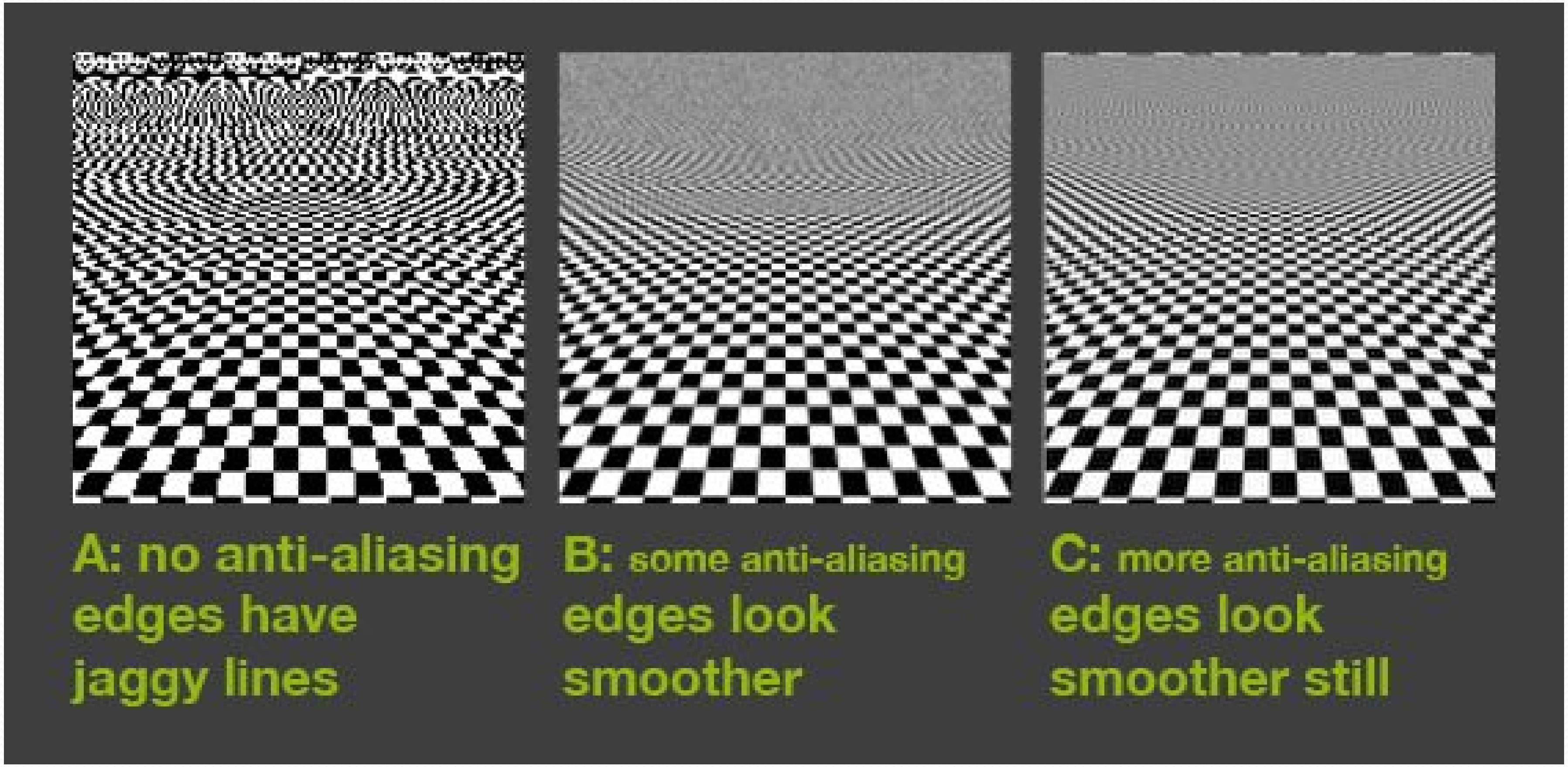
Aliasing

- Frequencies not honoring the sampling theorem creates aliasing.

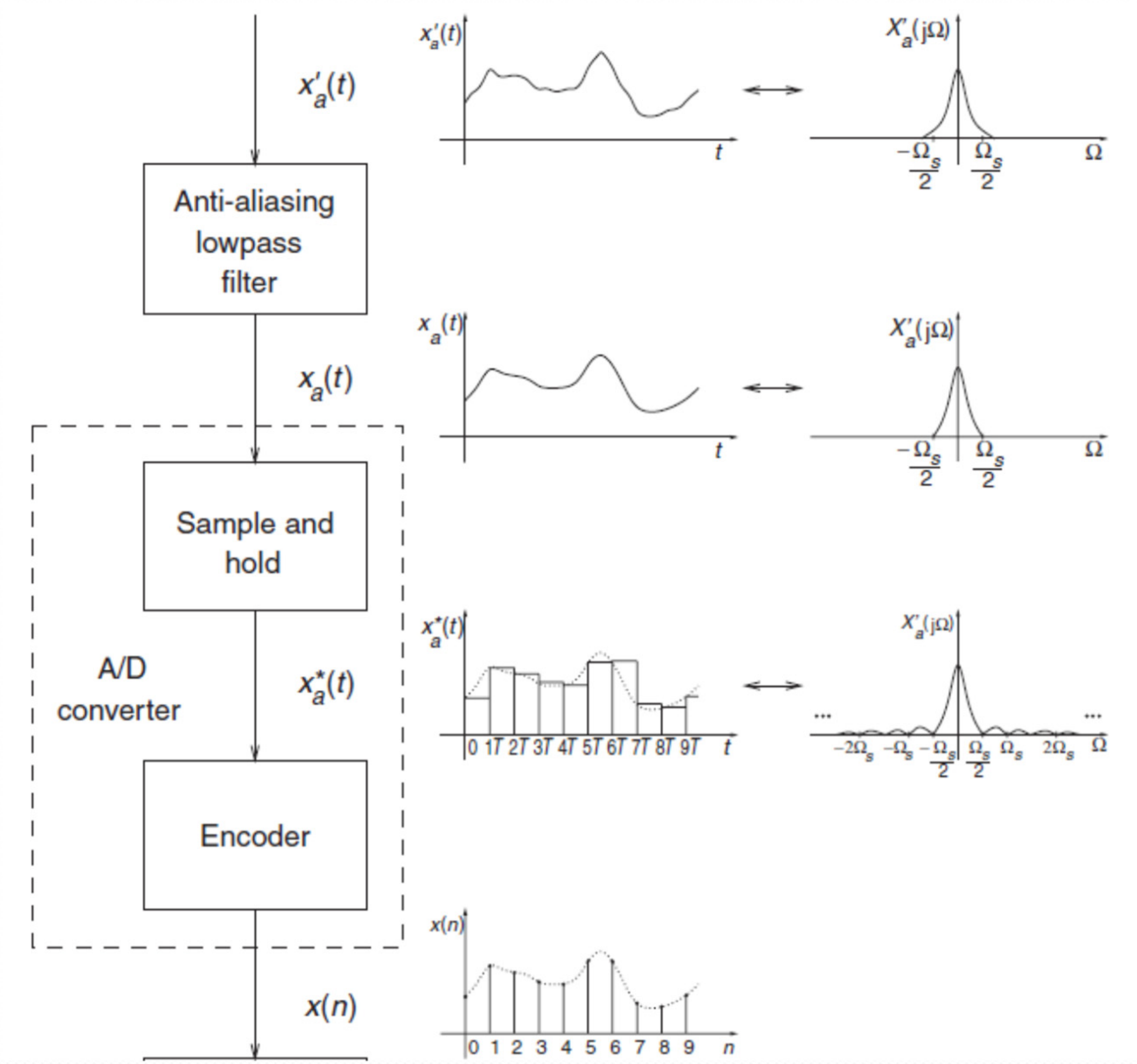




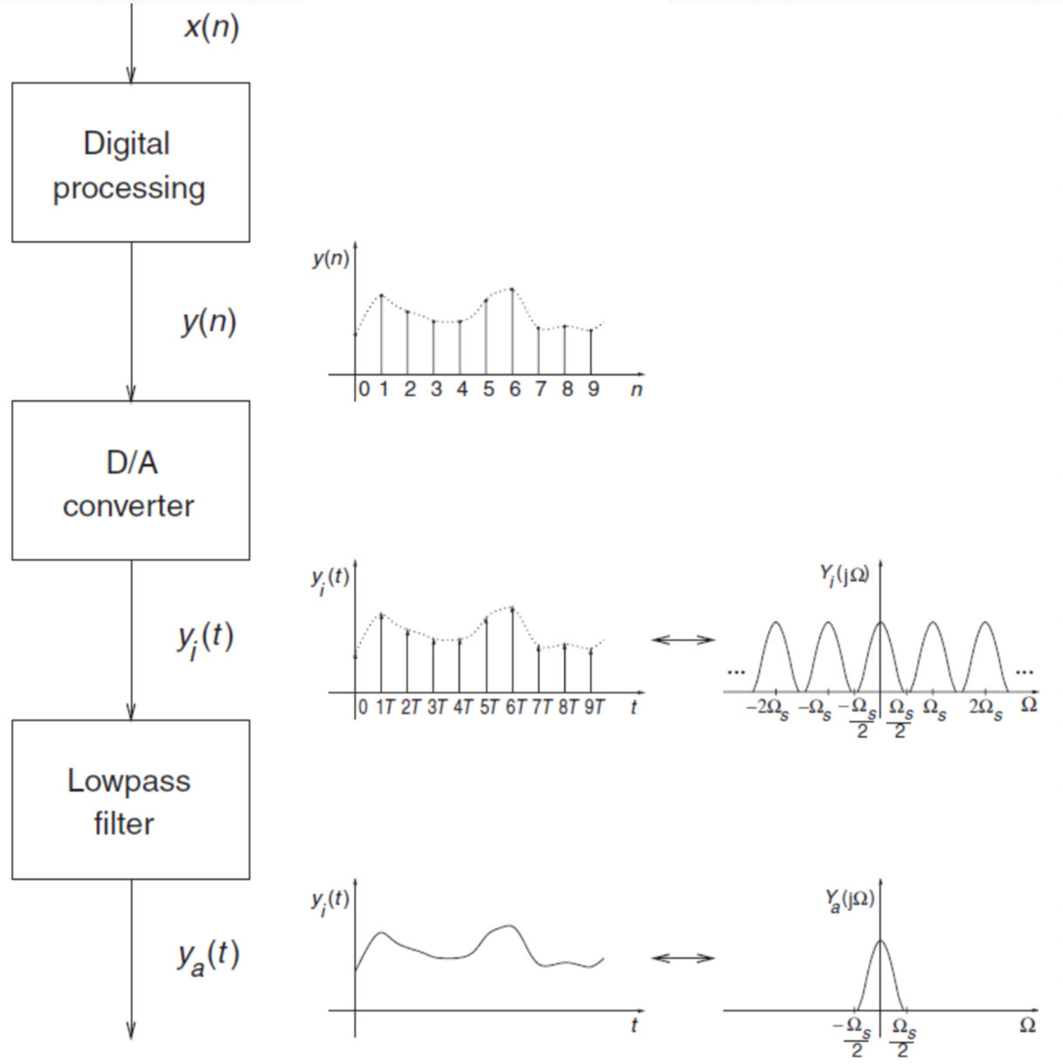
Aliasing for 2D signals



Analog -> Digital



Digital -> Analog



Converting Analog into Digital

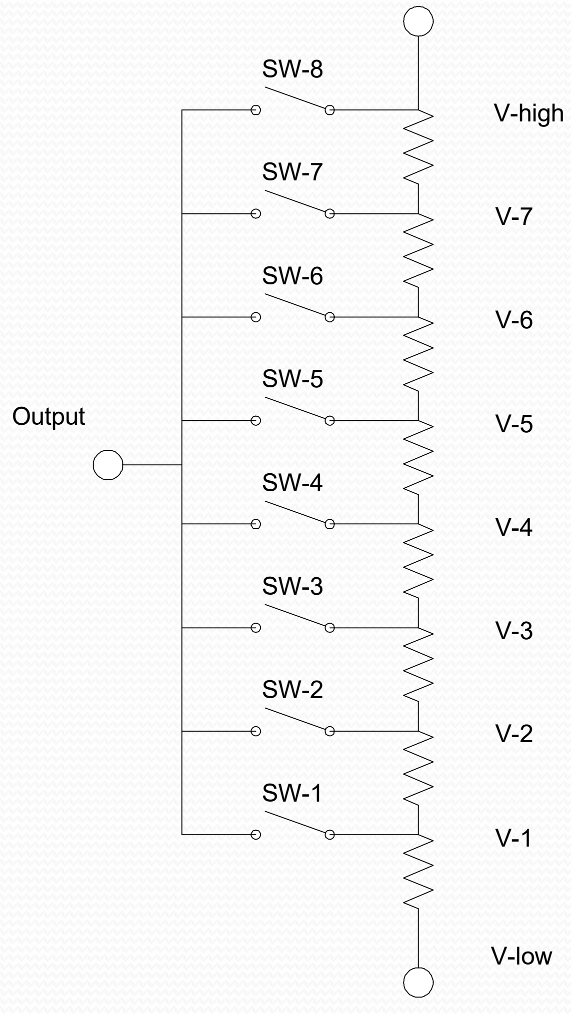
Electronically

- The device that does the conversion is called an Analog to Digital Converter (ADC)
- There is a device that converts digital to analog that is called a Digital to Analog Converter (DAC)

Converting Analog into Digital

Electronically

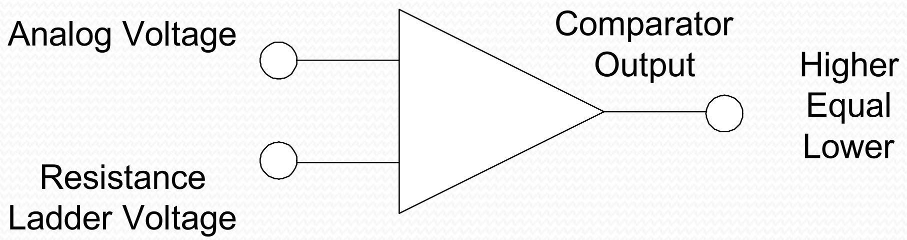
- The simplest form of ADC uses a resistance ladder to switch in the appropriate number of resistors in series to create the desired voltage that is compared to the input (unknown) voltage



Converting Analog into Digital

Electronically

- The output of the resistance ladder is compared to the analog voltage in a comparator
- When there is a match, the digital equivalent (switch configuration) is captured.



Converting Analog into Digital

Computationally

- The analog voltage can now be compared with the digitally generated voltage in the comparator
- Through a technique called binary search, the digitally generated voltage is adjusted in steps until it is equal (within tolerances) to the analog voltage
- When the two are equal, the digital value of the voltage is the outcome

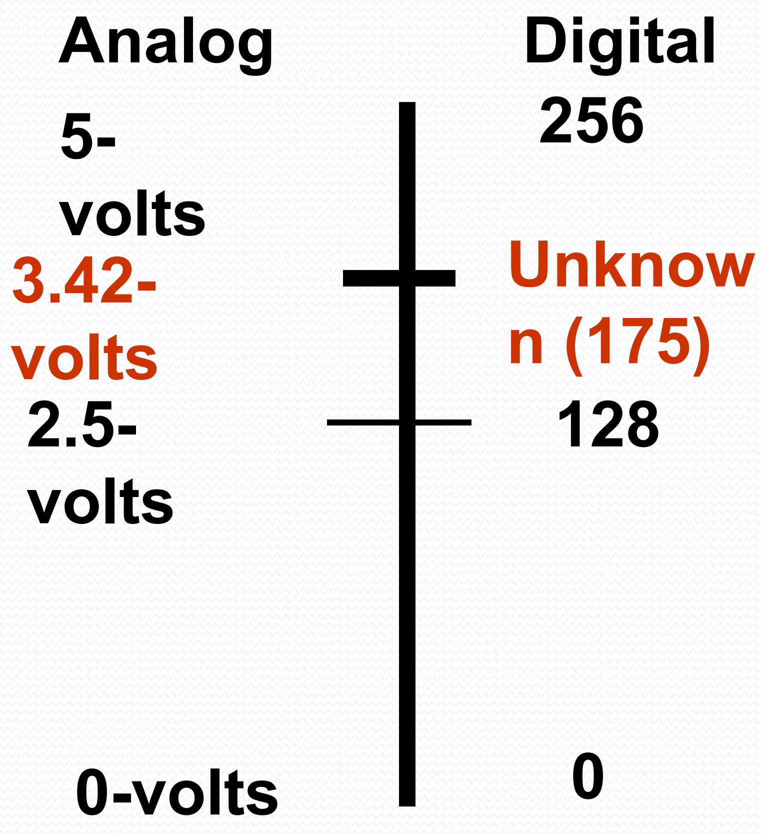
Converting Analog into Digital

Computationally

- The binary search is a mathematical technique that uses an initial guess, the expected high, and the expected low in a simple computation to refine a new guess
- The computation continues until the refined guess matches the actual value (or until the maximum number of calculations is reached)
- The following sequence takes you through a binary search computation

Binary Search

- Initial conditions
  - Expected high 5-volts
  - Expected low 0-volts
  - 5-volts 256-binary
  - 0-volts 0-binary
- Voltage to be converted
  - 3.42-volts
  - Equates to 175 binary



Binary Search

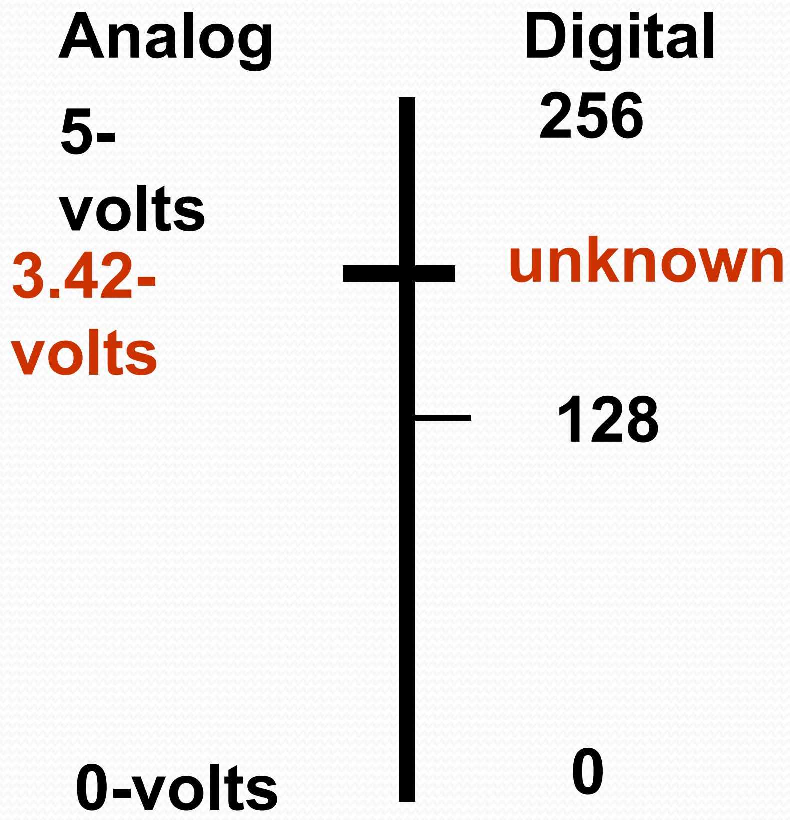
- Binary search algorithm:

$$\frac{High - Low}{2} + Low = NewGuess$$

- First Guess:

$$\frac{256 - 0}{2} + 0 = 128$$

Guess is Lower

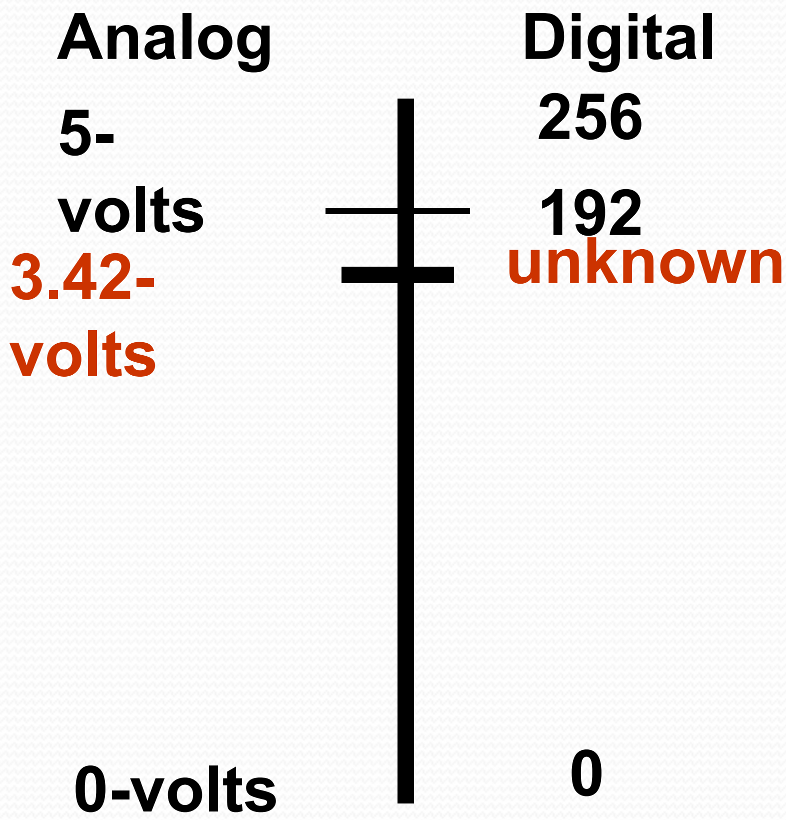


MSB

Binary Search

- New Guess (2):

$$128+\frac{256-128}{2}=192$$



Guess is Higher

1	0						

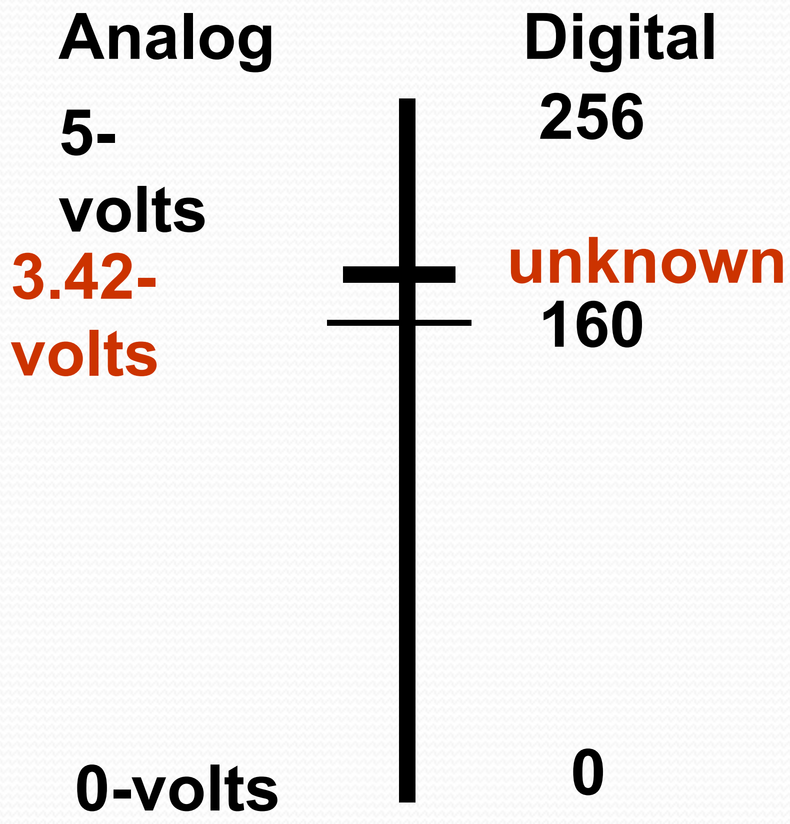
Binary Search

- New Guess (3):

$$128+\frac{192-128}{2}=160$$

Guess is Lower

1	0	1					



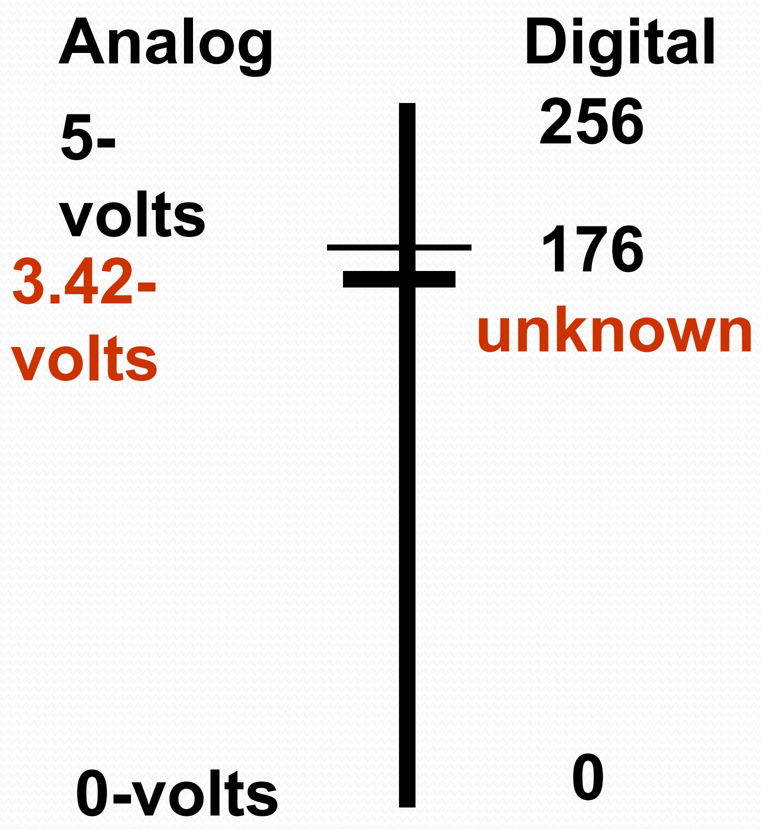
Binary Search

- New Guess (4):

$$160 + \frac{192 - 160}{2} = 176$$

Guess is Higher

1	0	1	0				



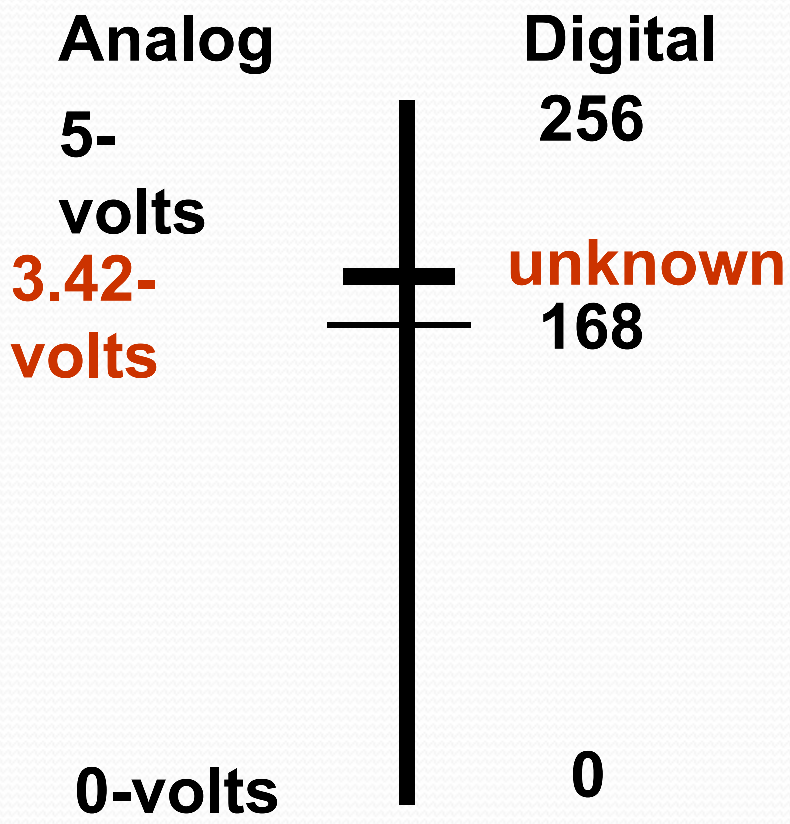
Binary Search

- New Guess (5):

$$160 + \frac{176 - 160}{2} = 168$$

Guess is Lower

1	0	1	0	1			



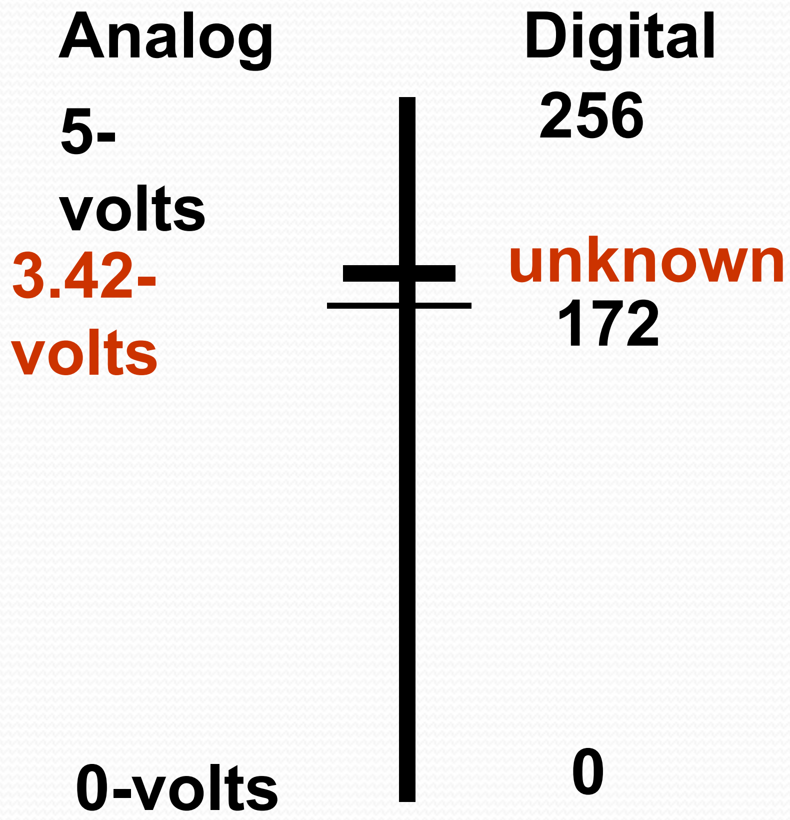
Binary Search

- New Guess (6):

$$168 + \frac{176 - 168}{2} = 172$$

Guess is Low (but getting close)

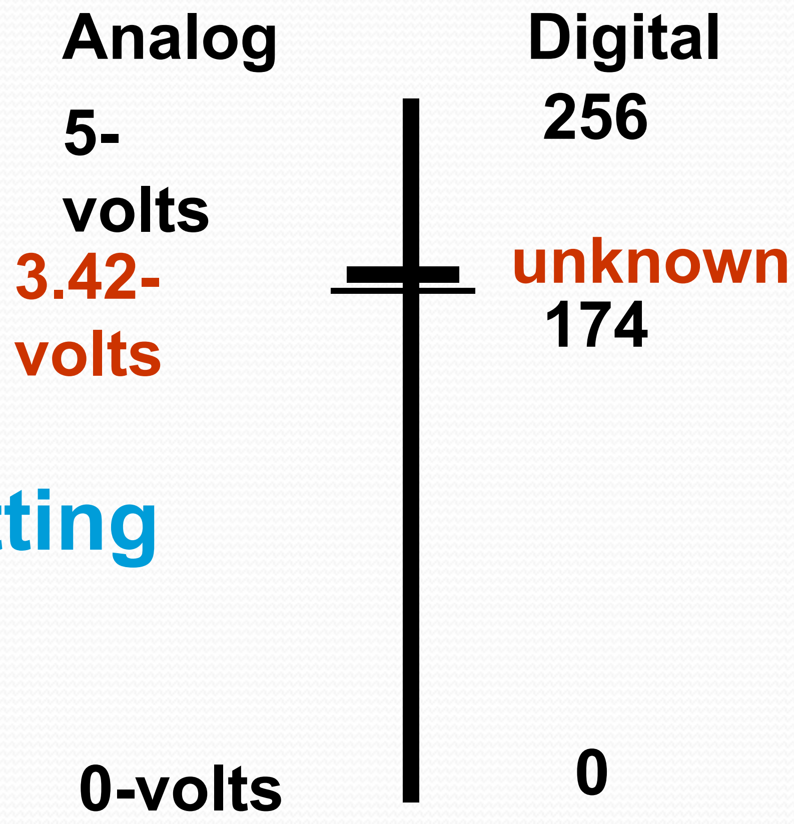
1	0	1	0	1	1		



Binary Search

- New Guess (7):

$$172 + \frac{176 - 172}{2} = 174$$



Guess is Low (but getting really, really, close)

1	0	1	0	1	1	1	

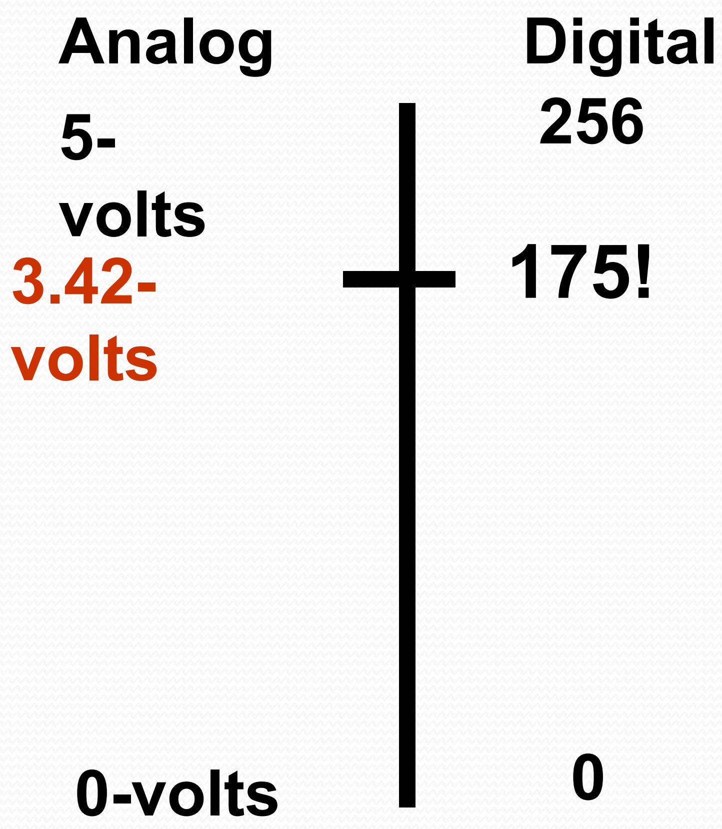
Binary Search

- New Guess (8):

$$174+\frac{176-174}{2}=175$$

Guess is Right On

1	0	1	0	1	1	1	1	← **LSB**



Binary Search

- The speed the binary search is accomplished depends on:
  - The clock speed of the ADC
  - The number of bits resolution
  - Can be shortened by a good guess.

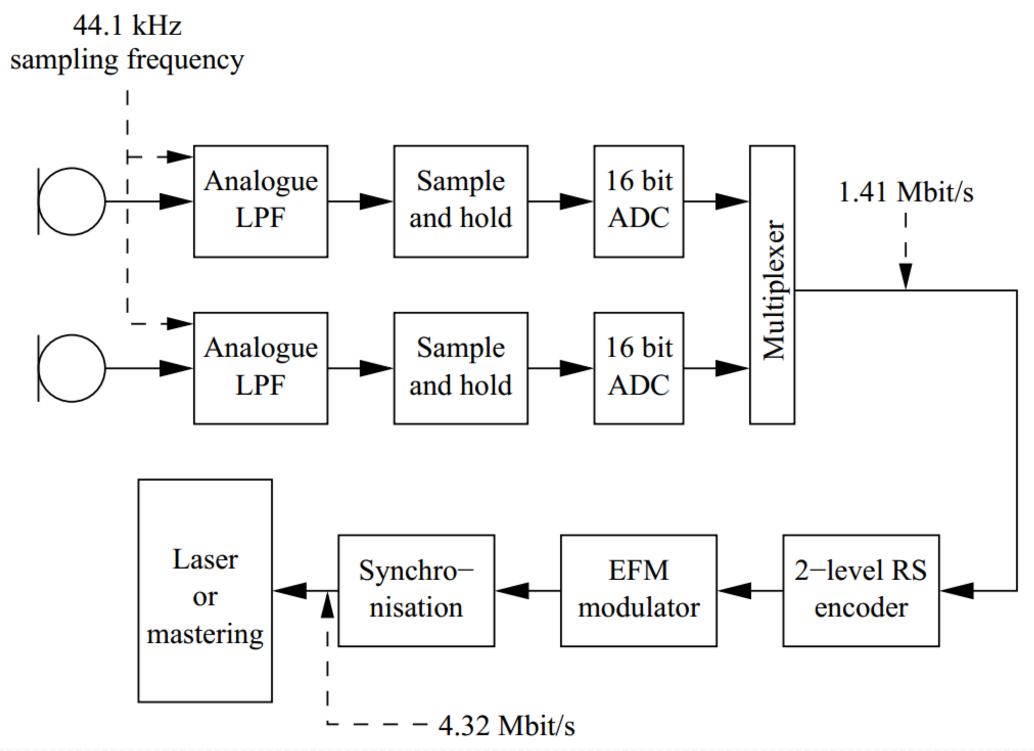
How Does It Work?

Resolution Trade-offs

Bit Resolution	High Bit Count	Good Duplication	Slow
	Low Bit Count	Poor Duplication	Fast
Sample Rate	High Sample Rate	Good Duplication	Slow
	Low Sample Rate	Poor Duplication	Fast

An example: CDs

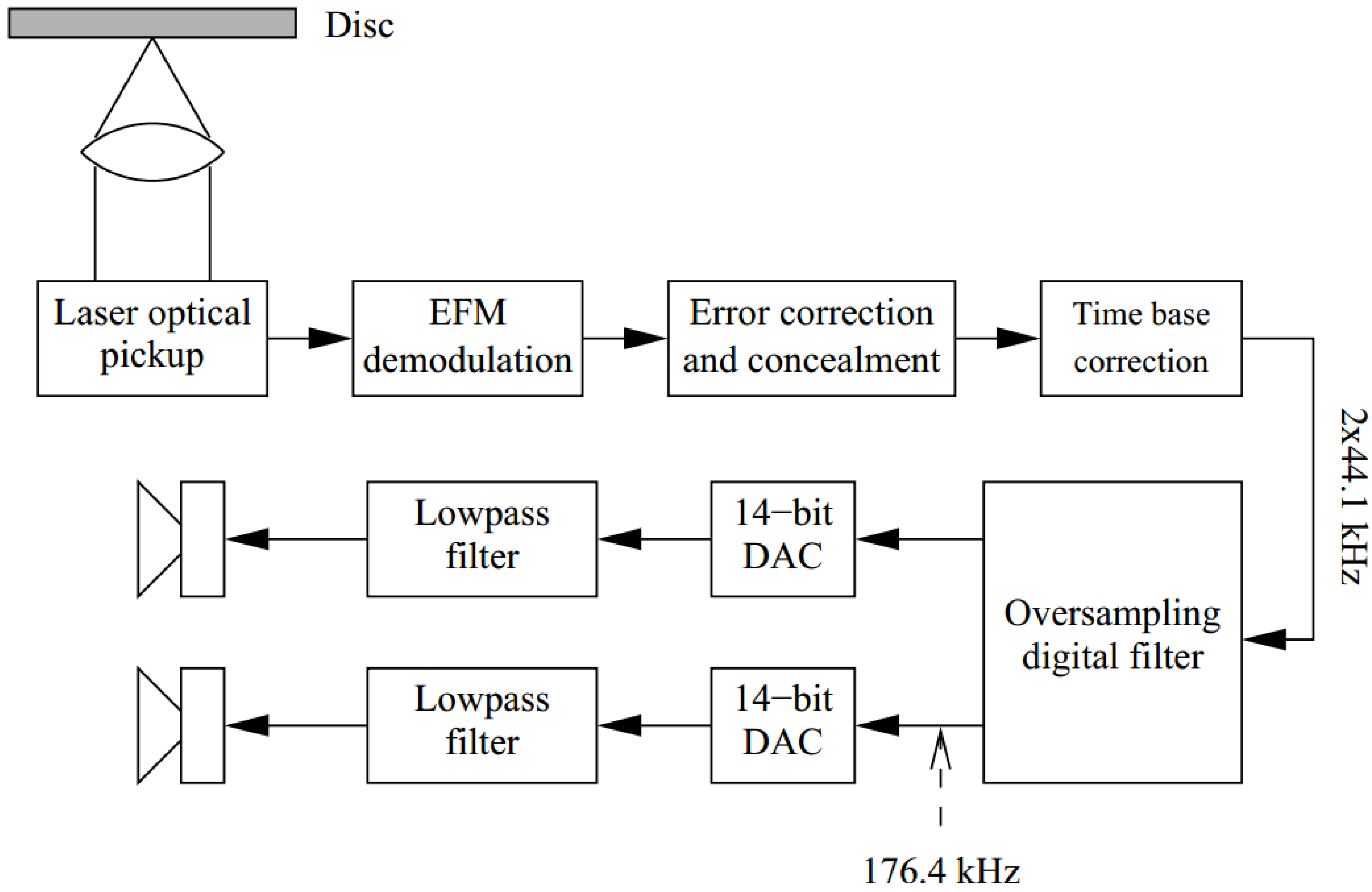
- Information on a compact disk is recorded on a spiral track as a succession of pits (10^6 bits/mm^2).
- The recording or mastering process is depicted below:



CD Recording process

- Analogue signal in each stereo channel is sampled at 44.1 kHz and digitised to 16 bits (96 dB dynamic range), resulting in 32 bits per sampling instant.
- Encoded using a two-level Reed-Solomon code to enable errors to be corrected or concealed during reproduction.
- An EFM (eight-to-fourteen) modulation scheme translates each byte in the stream to a 14 bit code, which is more suitable for disc storage (eliminates adjacent 1's, etc.)
- The resulting bit stream is used to control a laser beam, which records information on the disc.

CD audio signal reconstruction



CD Audio signal reconstruction (cont.)

- Track optically scanned at 1.2 m/s
- Signal is demodulated, errors detected and (if possible) corrected. If correction is not possible, errors are concealed by interpolation or muting.
- This results in a series of 16 bit words, each representing a single audio sample. These samples could be applied directly to a DAC and analogue lowpass filtered

CD Audio signal reconstruction (cont.)

- However, this would require high specification lowpass filters (20 kHz frequencies must be reduced by 50 dB), and the filter should have linear phase. To avoid this, signals are upsampled by a factor of 4. This makes the output of the DAC smoother, simplifying the analogue filtering requirements.
- The use of a digital filter also allows a linear phase response, reduces chances of intermodulation, and yields a filter that varies with clock rate.