



# Tecnologie digitali per il suono e l'immagine 2020/21

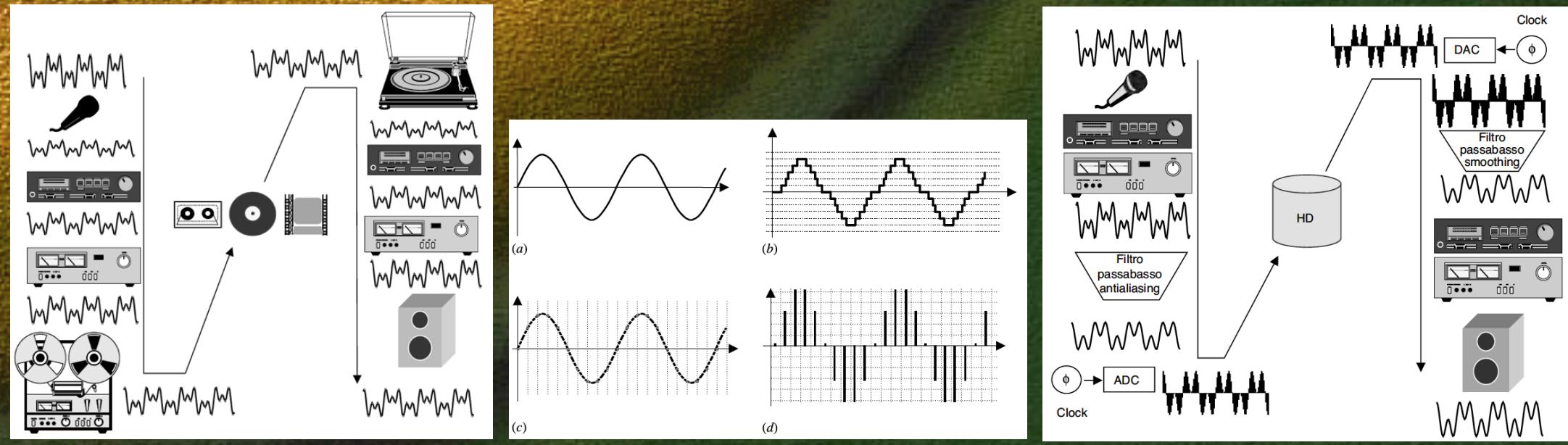
Vincenzo Lombardo

Corso di Laurea in DAMS

Università di Torino

Mutuato in parte da Elaborazione audio e musica  
(Laurea Magistrale di Informatica)

# Rappresentazione digitale del suono



# Rappresentazione del suono

- Trasmettere a distanza, tempo e spazio
- Registrazione, riproduzione, elaborazione segnale
- Consumer e professional

... flessibilità, velocità ... costi

# Analogico VS. Digitale

Indipendenza  
Segnale / Supporto

# Vantaggi del digitale

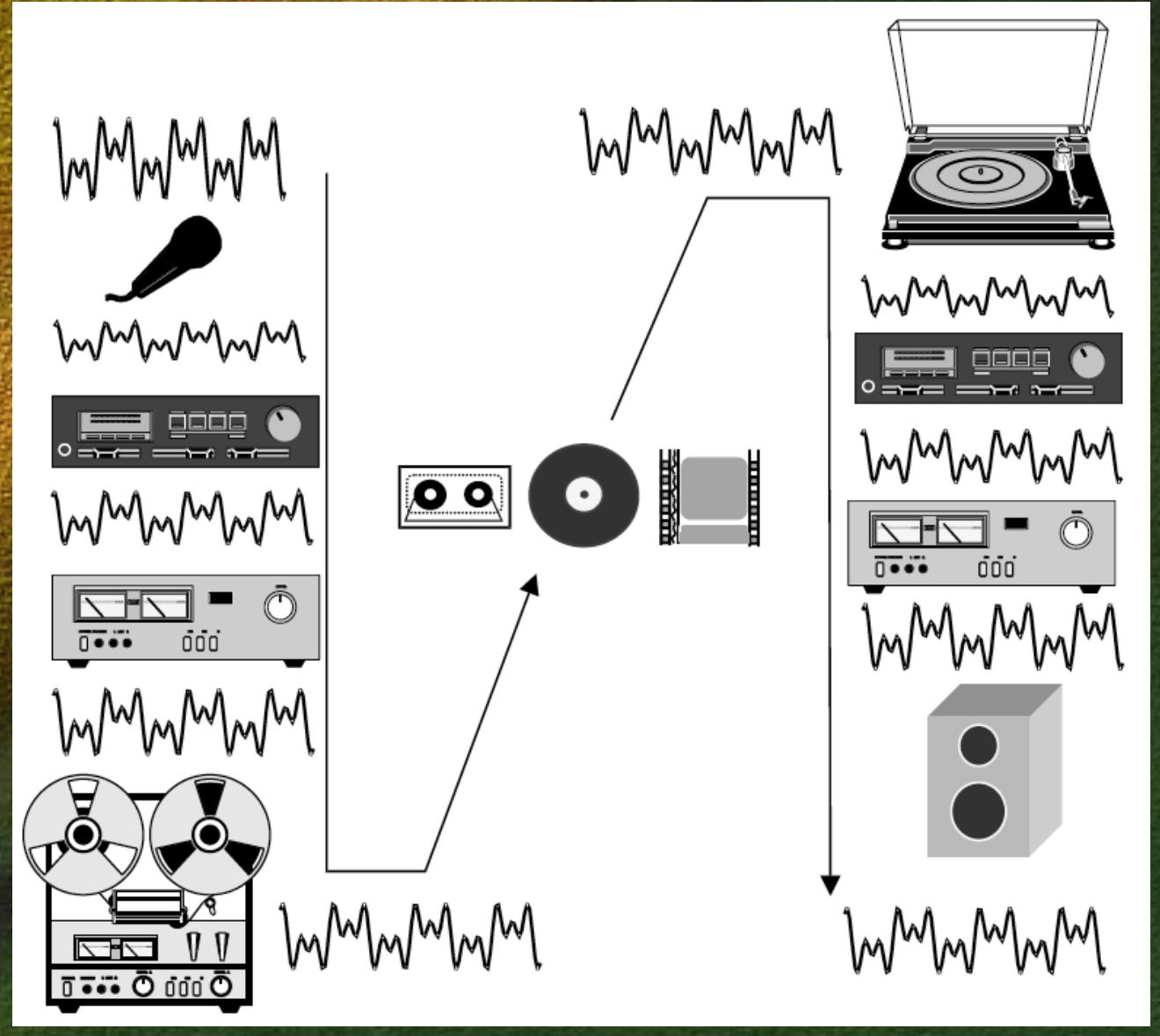
- copia del segnale identica all'originale
- non degradazione del segnale
- manipolazioni del segnale = operazioni aritmetiche “pulite”
- maggiori garanzie dai supporti standard (rilevamento/correzione errori)

# Svantaggi del digitale

- Problemi dello scambio di dati con l'esterno dell'elaboratore
- Grandi capacità e ampiezza di banda per memorizzazione e trasmissione

# Sistemi di rappresentazione

# Il suono analogico



# Rapporto segnale/rumore

## Signal-to-noise ratio (SNR)

$$SNR = \frac{\text{ampiezza\_segnale\_utile}}{\text{ampiezza\_rumore}}$$

$$SNR_{dB} = 20 \log \left( \frac{\text{ampiezza\_segnale\_utile}}{\text{ampiezza\_rumore}} \right)$$

$$SNR_{dB} = 20 [\log (\text{amp. segnale utile}) - \log (\text{amp. rumore})]$$

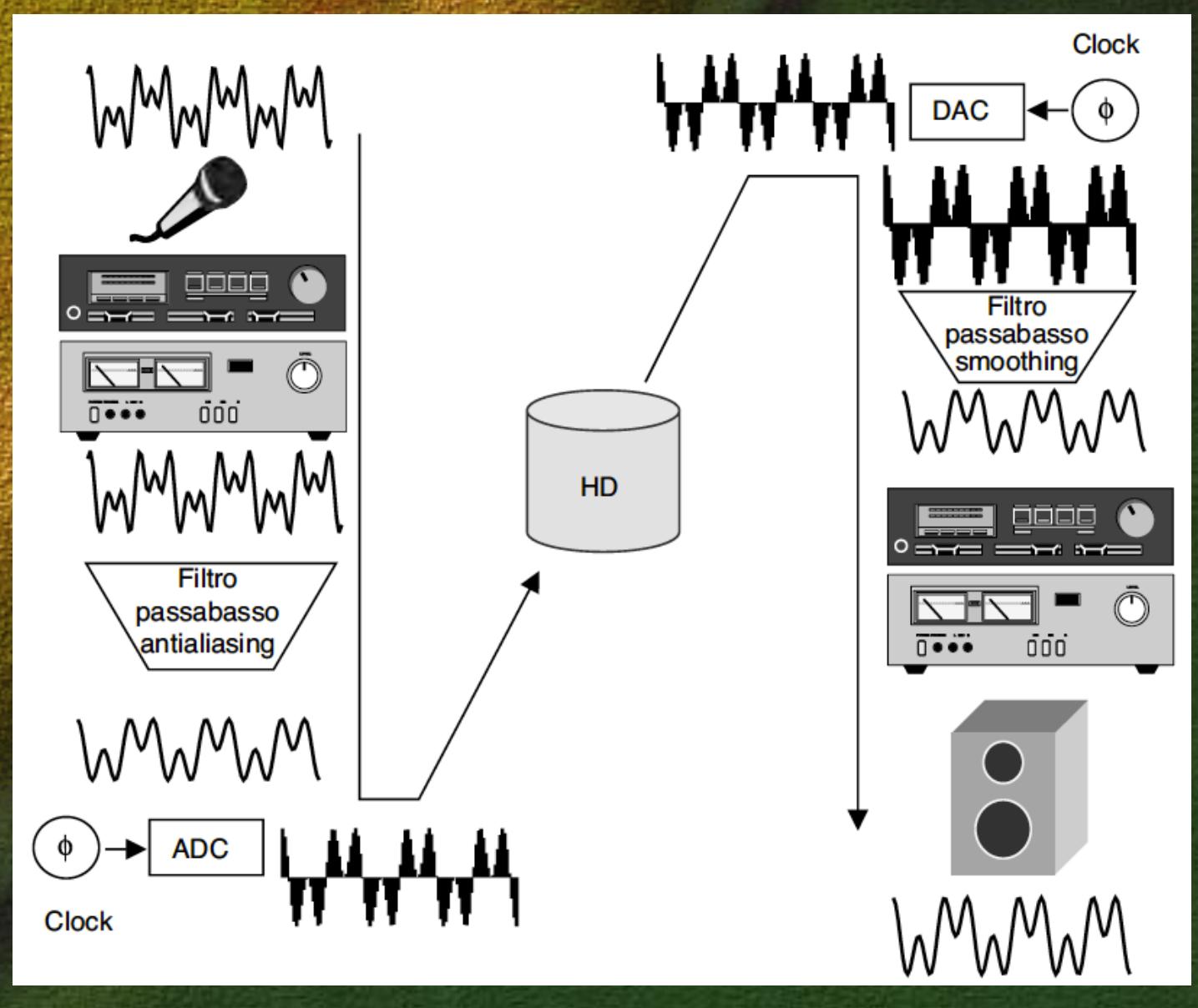
$$SNR_{dB} = \text{intensità segnale utile}_{dB} - \text{intensità rumore}_{dB}$$

# Gamma dinamica - Dynamic Range (DR)

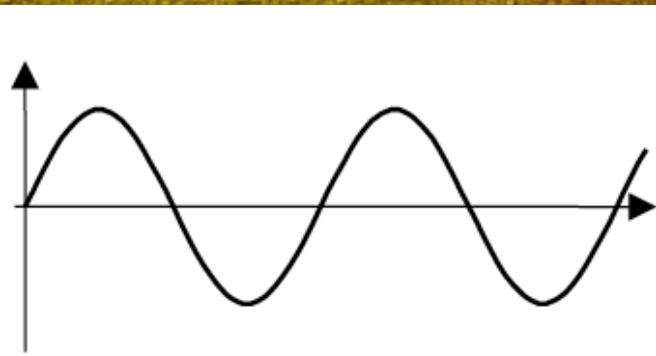
- Differenza in dB tra le ampiezze massima e minima del segnale utile
- ... dato un sistema, intervallo tra picco che causa clip e rumore di fondo

$$\text{Ampiezza}_{\text{max}}_{\text{dB}} - \text{Ampiezza}_{\text{min}}_{\text{dB}}$$

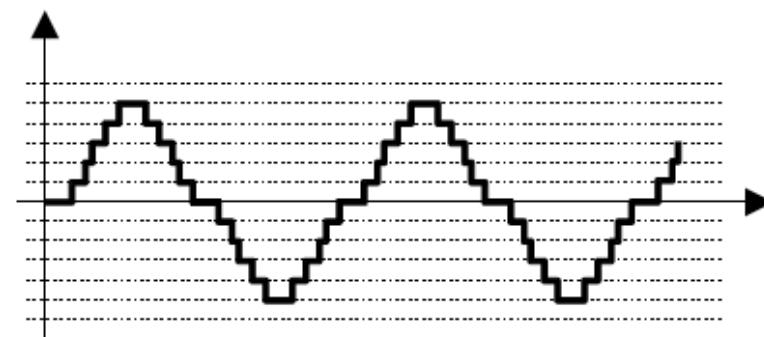
# Catena dell'audio digitale



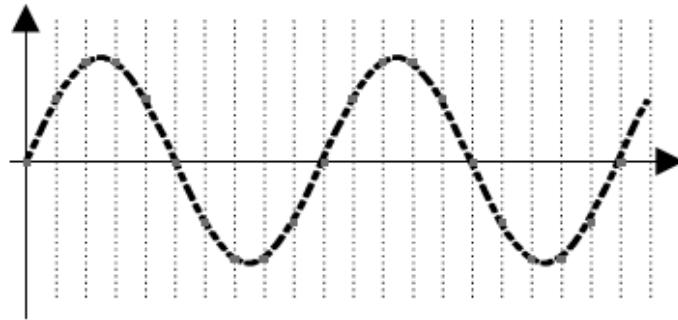
# Campionamento e quantizzazione



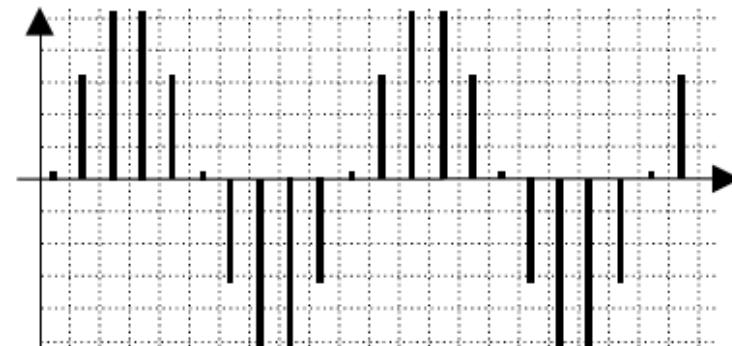
(a)



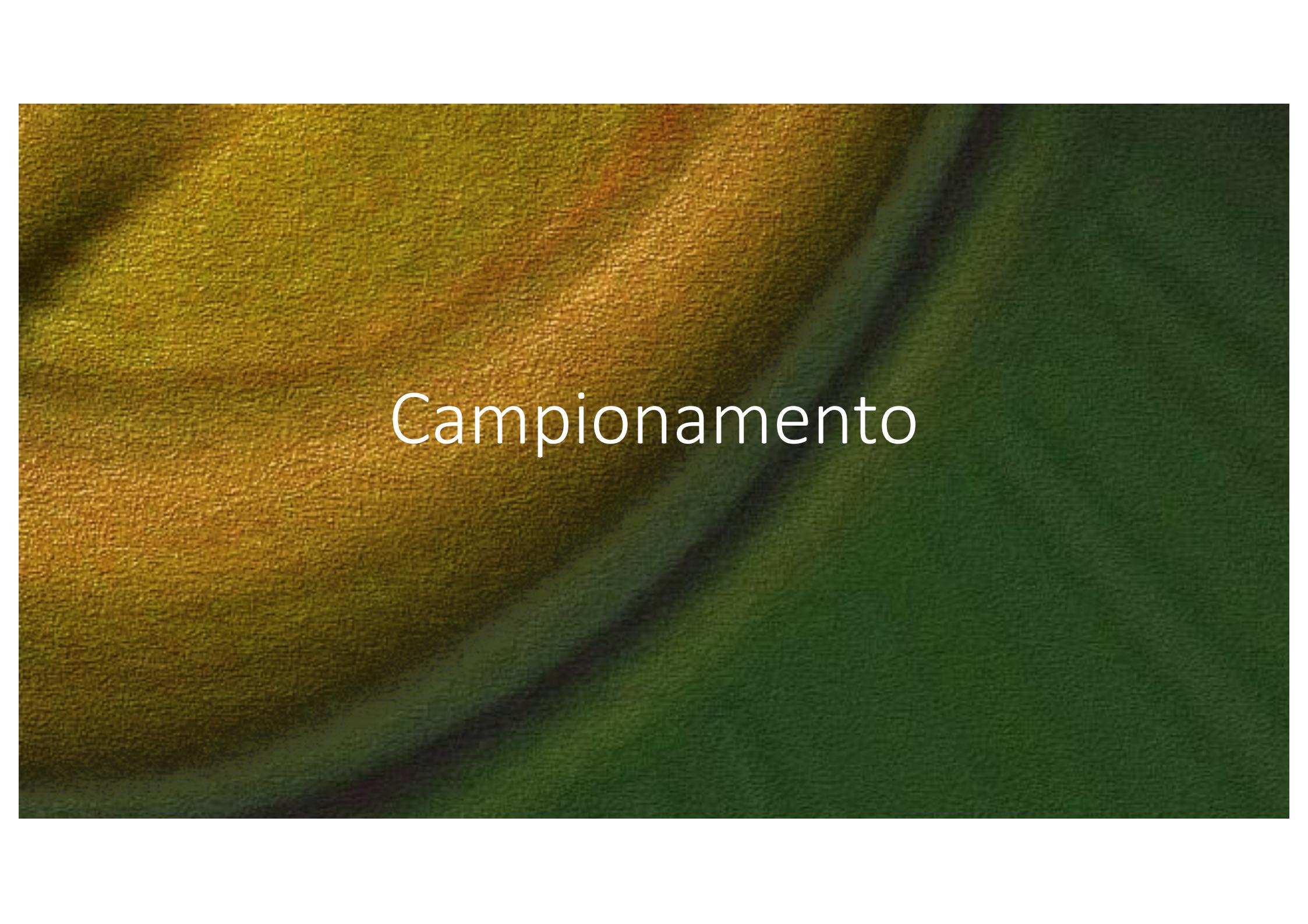
(b)



(c)



(d)

A blurred background image of a soccer field with a yellow-to-green gradient overlay.

Campionamento

# Problema del campionamento

- Discretizzare il tempo (periodo di campionamento)
- Segnali audio variano rapidamente

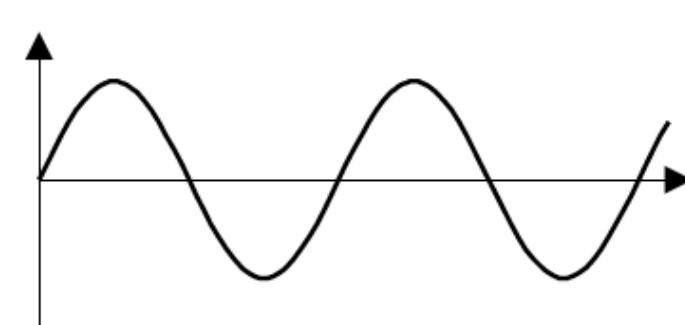
## Problema

- Con che velocità si prelevano i campioni?

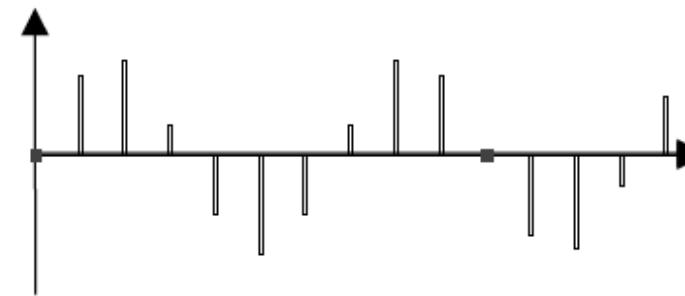
## Intuitivamente

- Prelevamento campioni deve seguire ogni variazione segnale
- Velocità di variazione dipende da componente armonica più alta
- Frequenza max determina periodo di campionamento
- Al limite segnali analogico e digitale coincidono

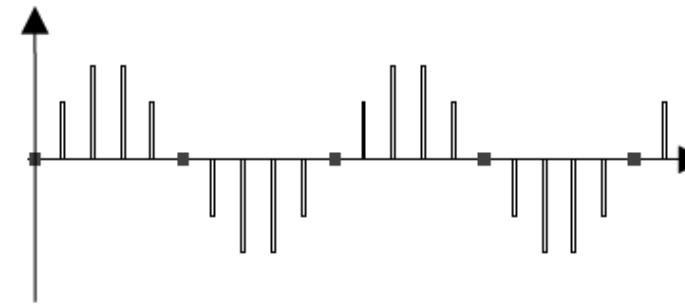
# Tasso di campionamento



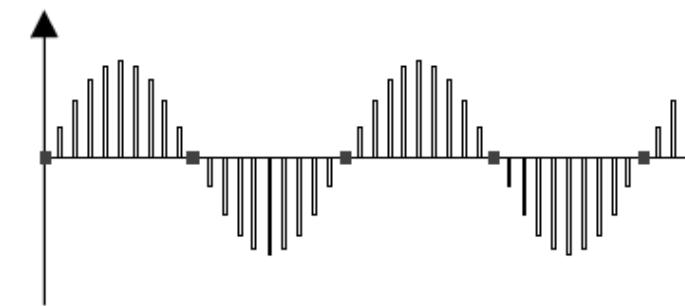
(a)



(b)



(c)



(d)

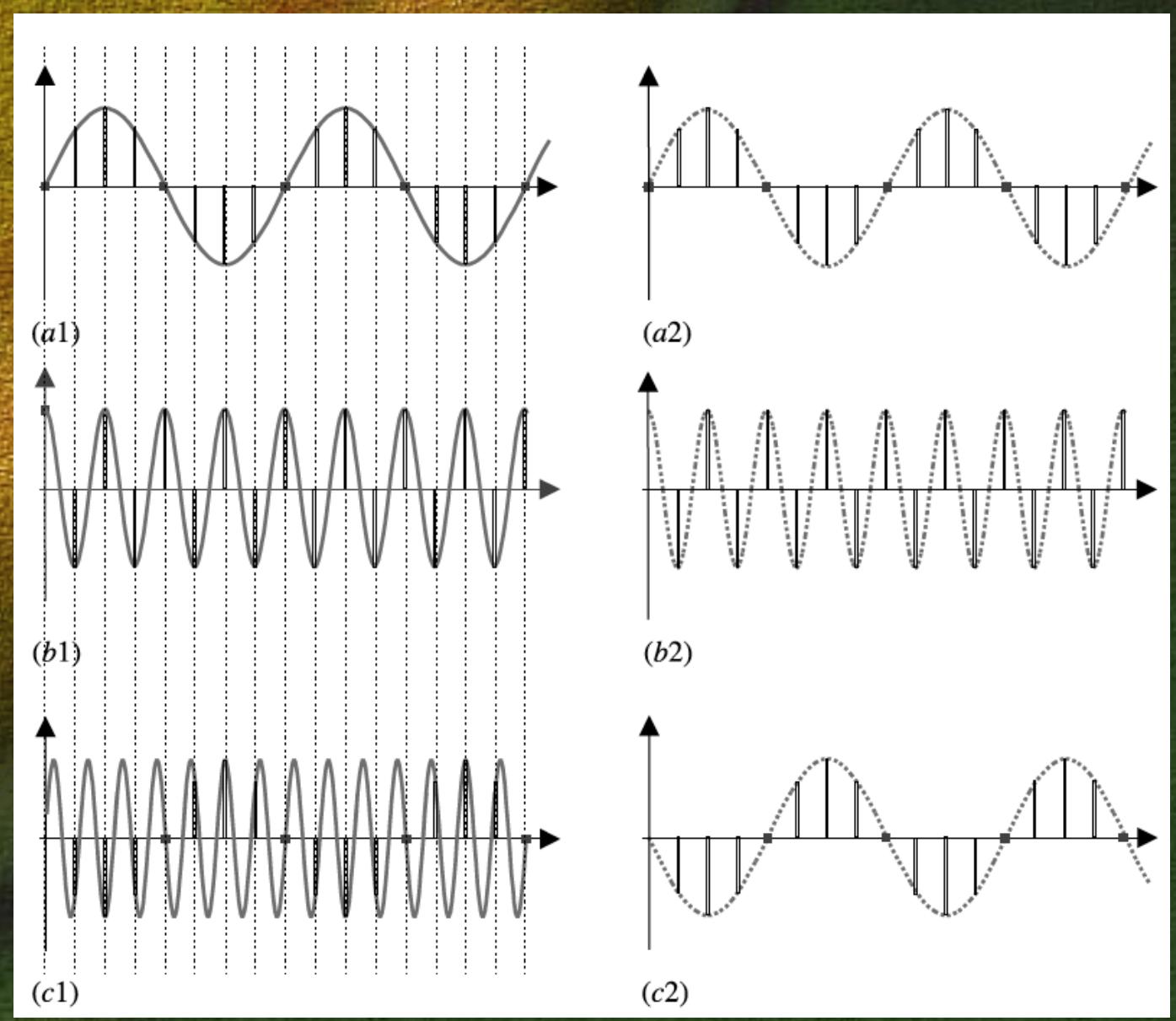
# Dimensione degli intervalli

- Maggiore frequenza di campionamento, più accurata descrizione del segnale
- Come si fa a non avere perdita di informazione ?
- Qual è il minimo valore della frequenza di campionamento ?

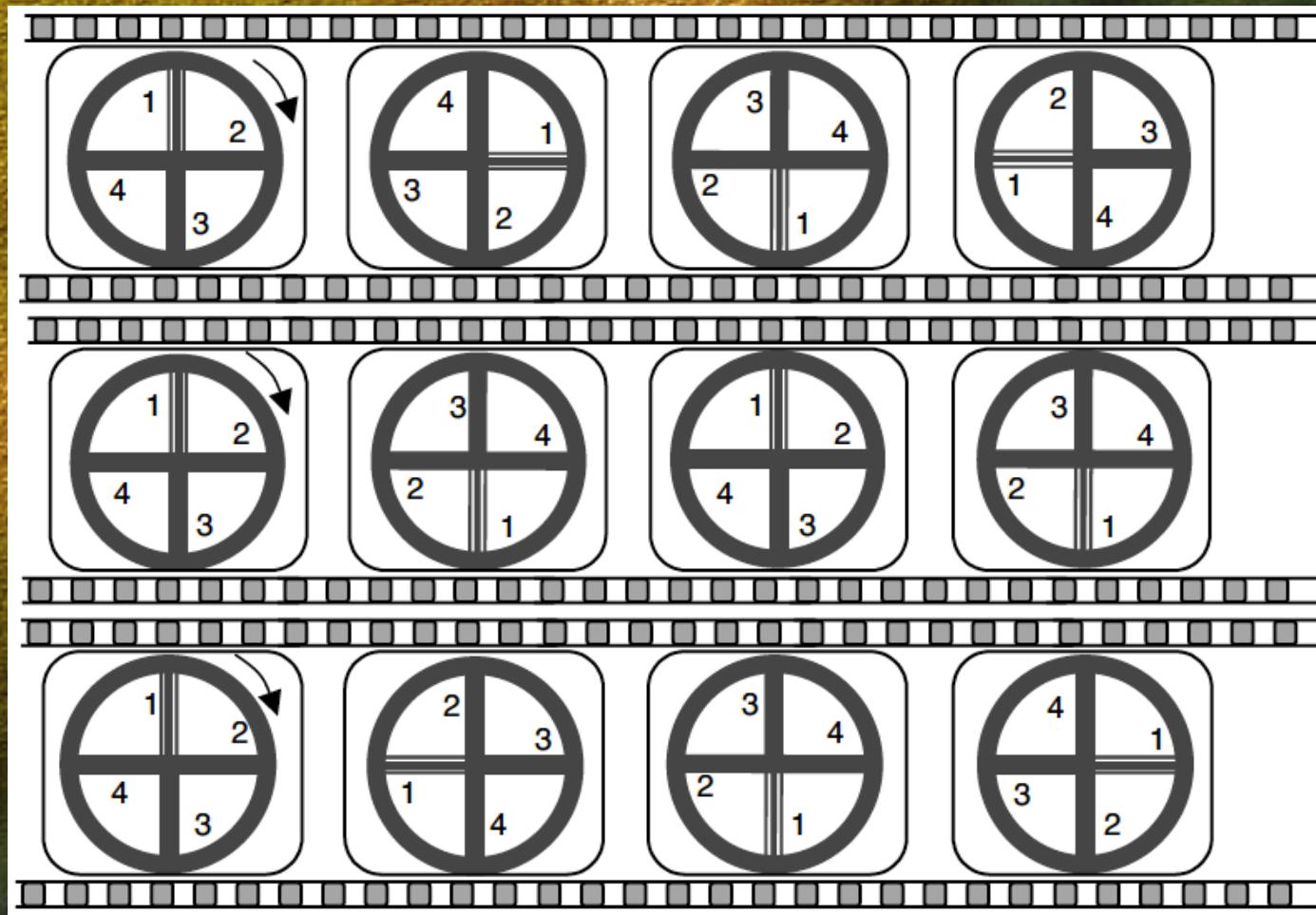
Sovracampionamento

Campionamento critico

Sottocampionamento



# Campionamento al cinema



# Il foldover o aliasing

- Sia  $f_c$  la frequenza di campionamento
- Sia  $f$  una frequenza presente nel segnale in ingresso tale che  $f > f_c / 2$
- Si “inventa” una frequenza nuova (alias)
- Come avviene ciò?

# La frequenza di alias

- frequenza ricostruita  $f_r = f + kf_c$
- $k$  intero, tale che  $-f_c/2 < f + kf_c < f_c/2$

due casi:

- $f$  nell'intervallo:  $-f_c/2 < f < f_c/2$ ,  $k=0$   $f_r=f$
- $f$  fuori intervallo:  $f < -f_c/2$ ,  $f > f_c/2$ ,  $k=?$

Esempi

$$f_r = f + kf_c, \quad -f_c/2 < f + kf_c < f_c/2$$

- $f_c=10000 \text{ Hz}, f=4000 \text{ Hz}, k=0, f_r = 4000 \text{ Hz}$   
–  $5000 < 4000 < 5000$
- $f_c=10000 \text{ Hz}, f=6000 \text{ Hz}, k=-1, f_r = -4000 \text{ Hz}$   
–  $5000 < -4000 < 5000$
- $f_c=955 \text{ Hz}, f=1000 \text{ Hz}, k=-1, f_r = 45 \text{ Hz}$   
–  $472,5 < 45 < 472,5$

## Esempio di aliasing: risultanza nei battimenti



- $k = -[2f / f_c]$ ,  $f_r = f - [2f / f_c] f_c$
- Frequenza di campionamento 8000 Hz
- Segnale sinusoidale 2000 Hz
- Segnale sinusoidale 5998 Hz ...
- Per  $k=-1$ ,  $f_r = f_c - f = 8000 - 5998 = 2002$

# Esempio di aliasing: spurio trattamento del glissando

- Da 20 Hz a 30.000 Hz
- Durata 30 secondi
- Frequenza di campionamento 44.100 Hz



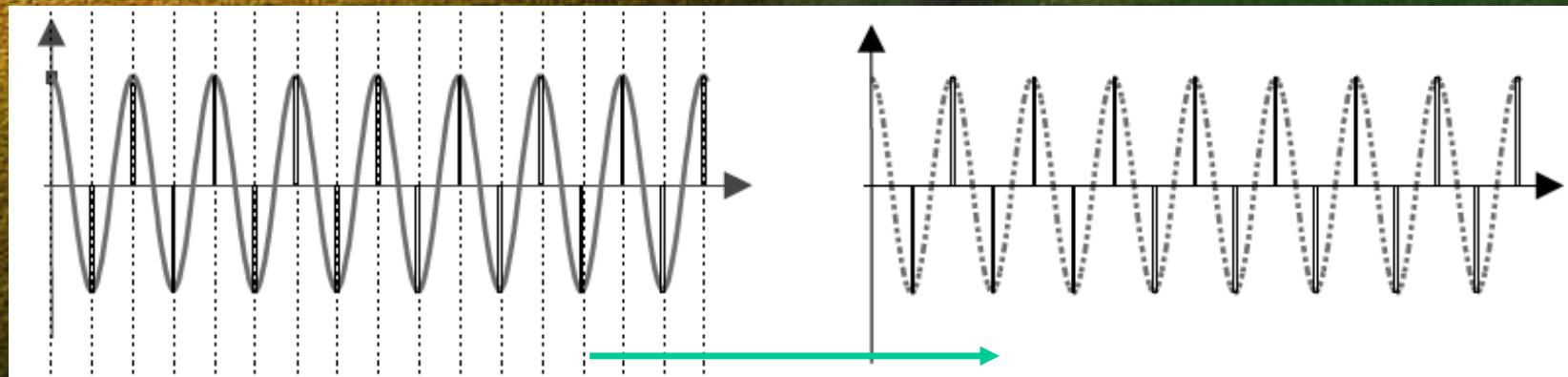
Glissando lineare



Glissando log

# La frequenza di Nyquist

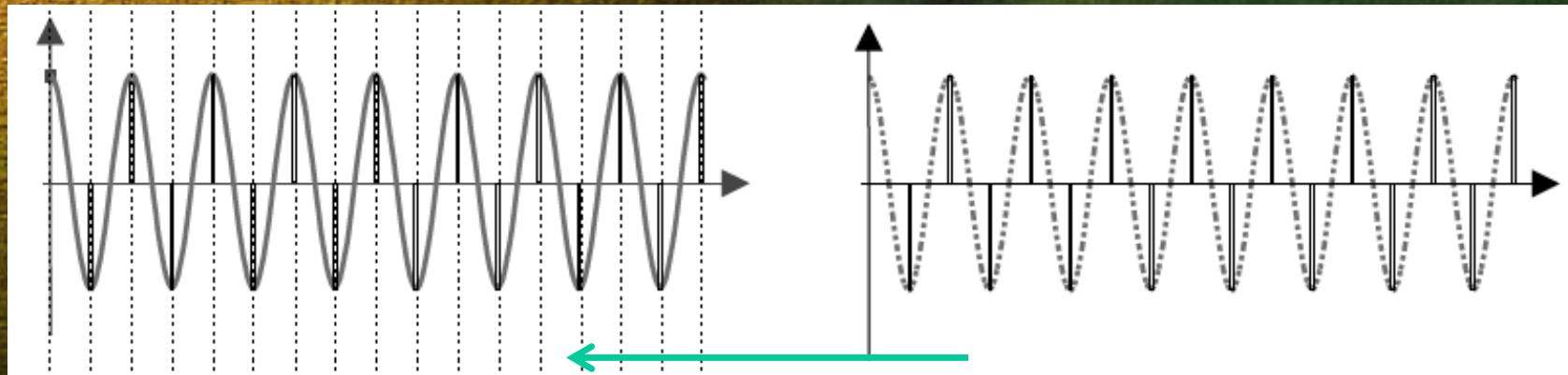
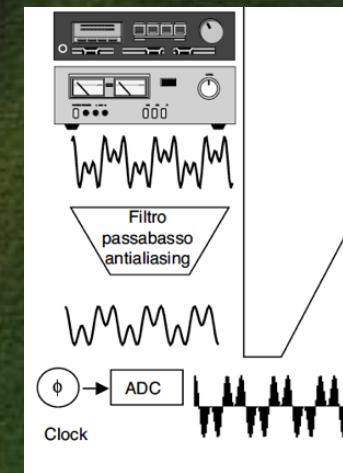
- In un periodo, il segnale cambia direzione due volte: occorrono almeno due campioni per periodo
- Frequenza di campionamento almeno il doppio della frequenza max presente



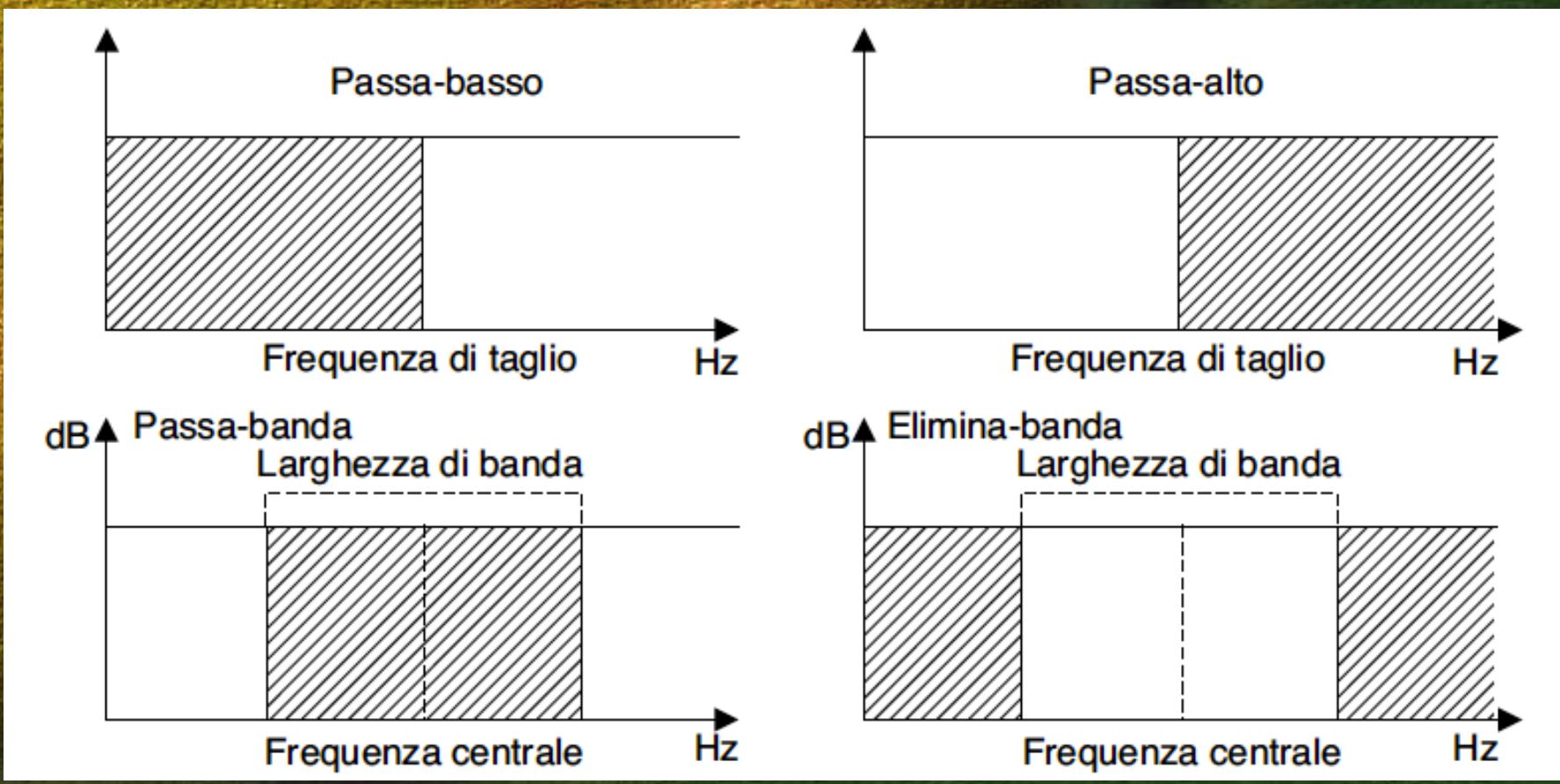
# La frequenza di Nyquist

Operativamente:

- Si fissa la frequenza di Nyquist
- Si eliminano frequenze superiori alla metà della frequenza di Nyquist

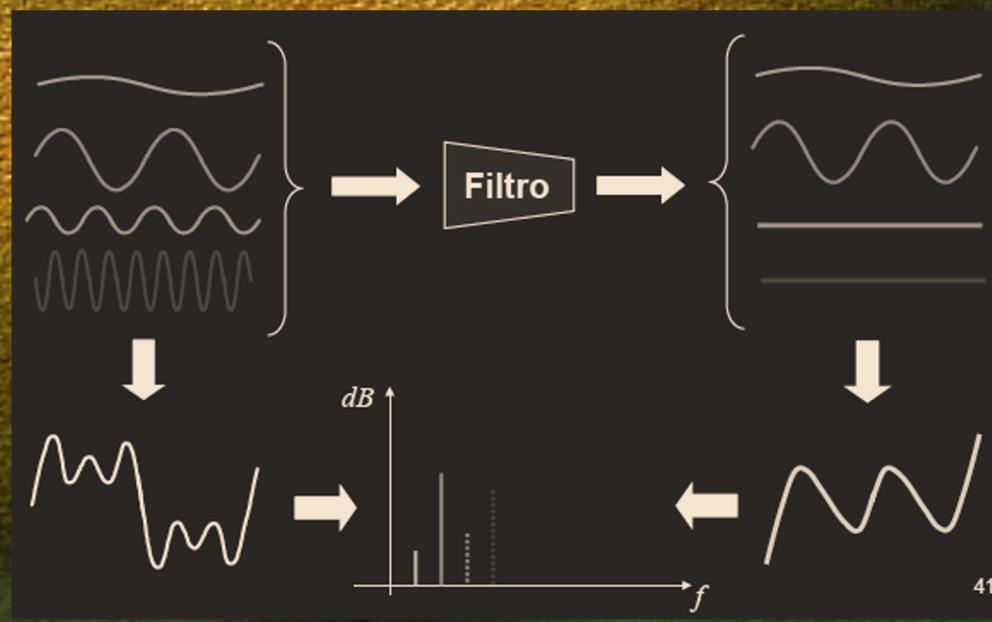


# Filtro passabasso



# Filtro passabasso

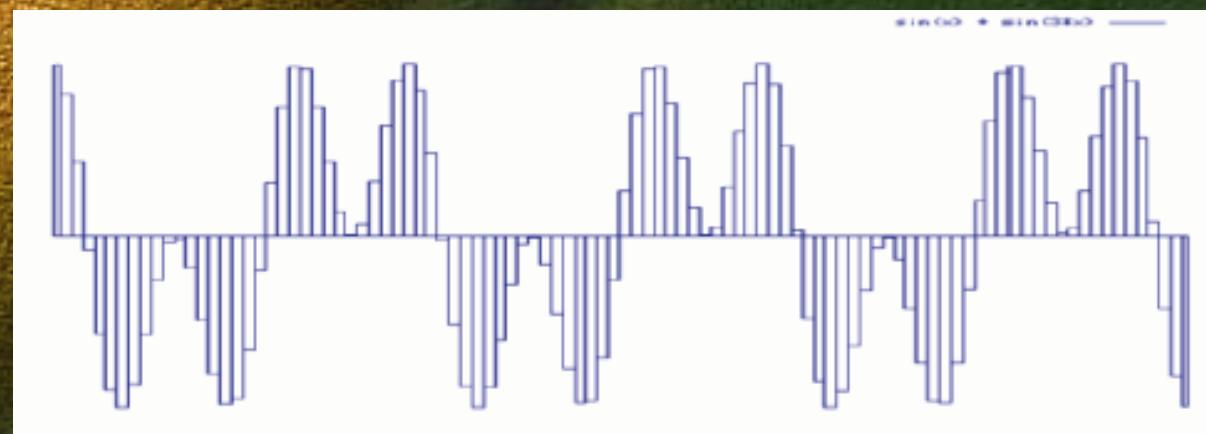
Elimina frequenze > soglia (frequenza di cut-off, metà della frequenza di campionamento)



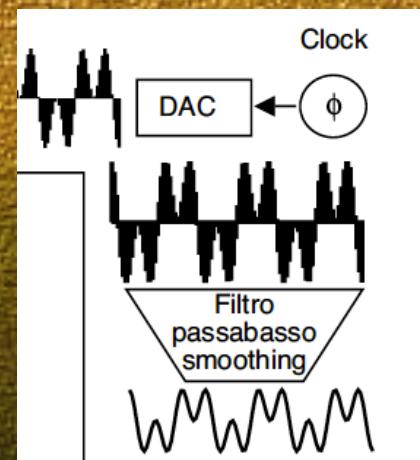
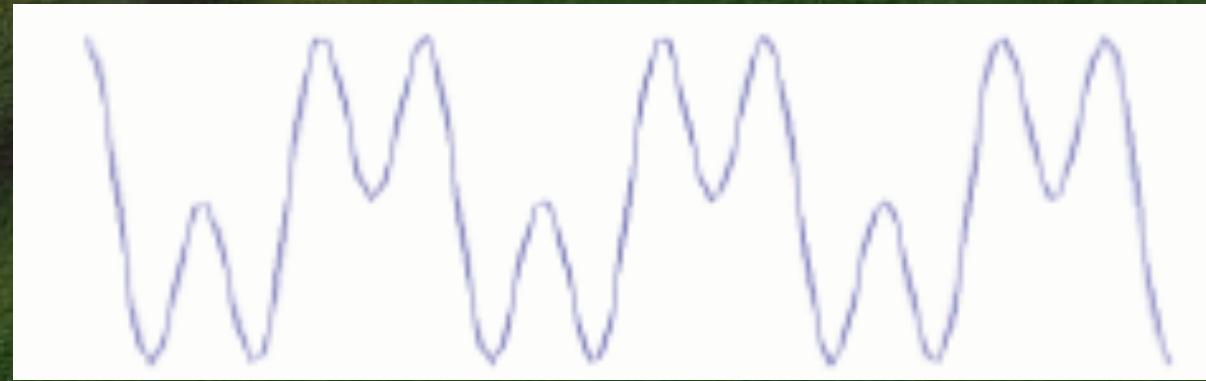
# Interpolazione dei campioni



DAC



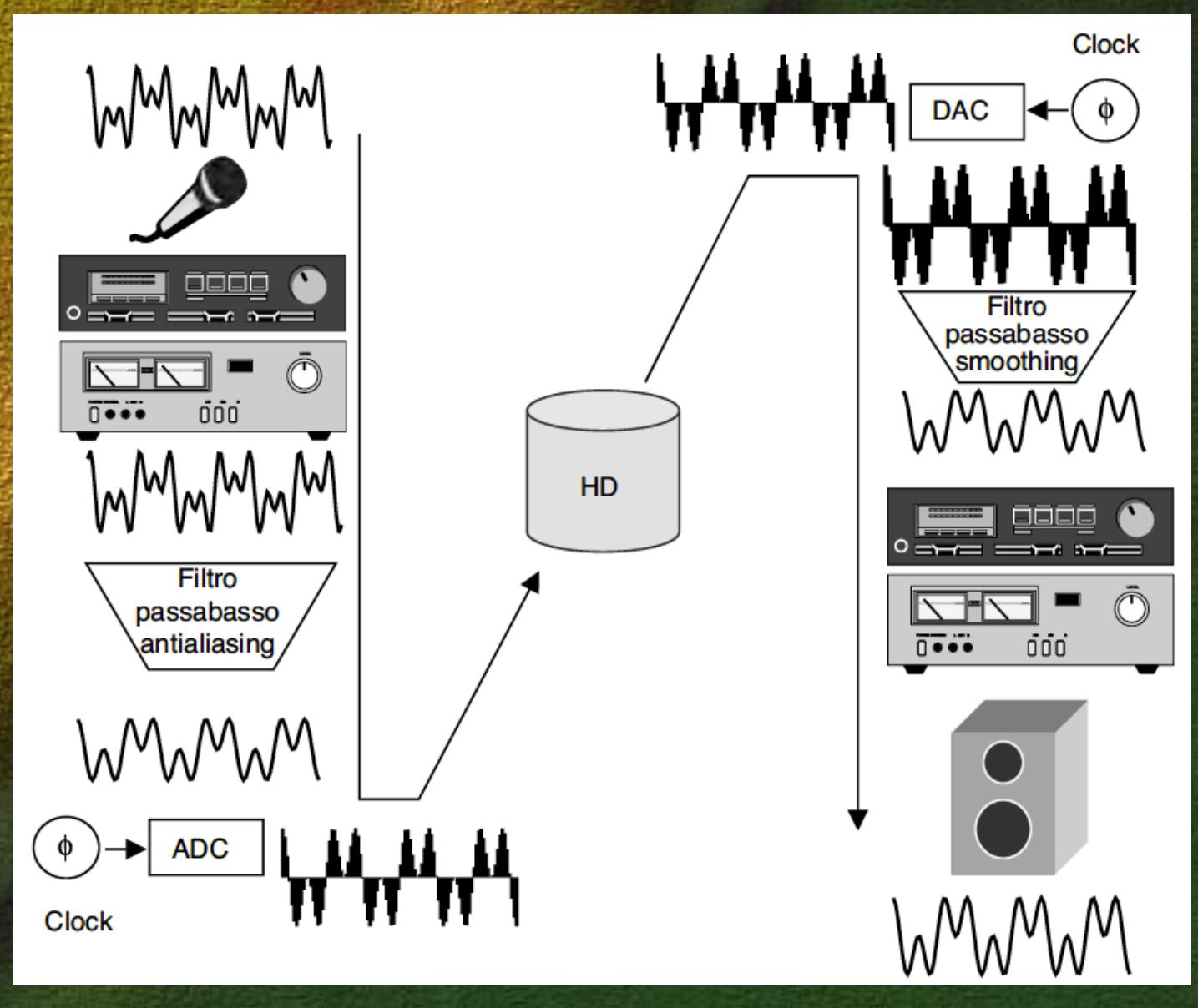
Filtro  
passabasso

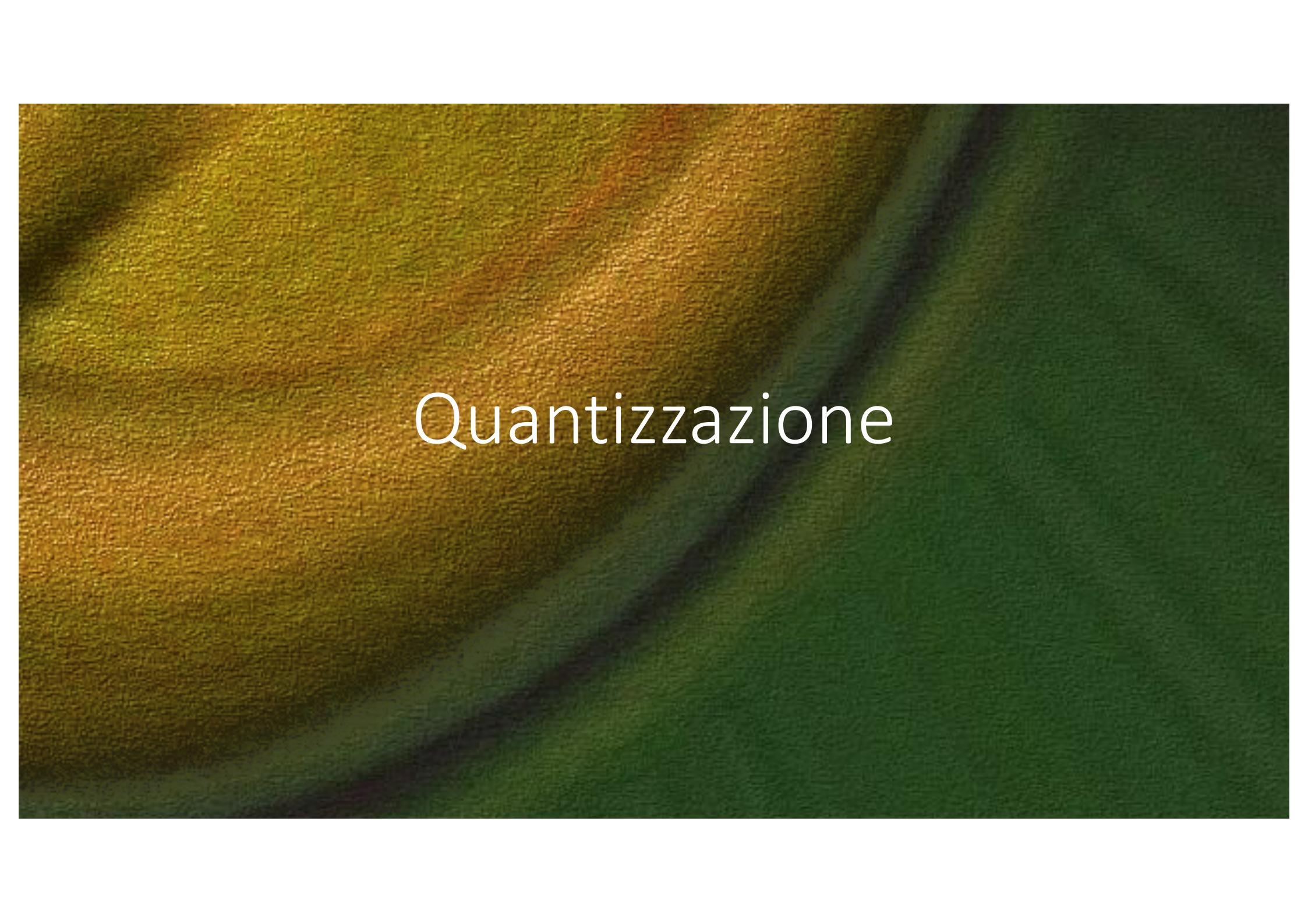


## Dati pratici per il campionamento

- Max frequenza udibile 20 KHz
- campionamento oltre i 40 KHz
- 44,1 kHz è la frequenza usata per il CD

# Catena dell'audio digitale





Quantizzazione

# Problema della quantizzazione

- da segnale a successione di numeri
- ogni numero esprime il valore di ampiezza di un campione

Di quante cifre è composto il dato numerico?

# Parole binarie

- Sequenze di bit (lunghezza  $n$ )
  - può assumere  $2^n$  configurazioni diverse
  - cioè  $2^n$  valori diversi

Esempi:

- $n=2$ ,  $2^2=4$  valori (00, 01, 10, 11)
- $n=3$ ,  $2^3=8$  (000, 001, 010, 011, 100, ...)
- ...

# Qual è il numero di cifre corrette?

numero di bit

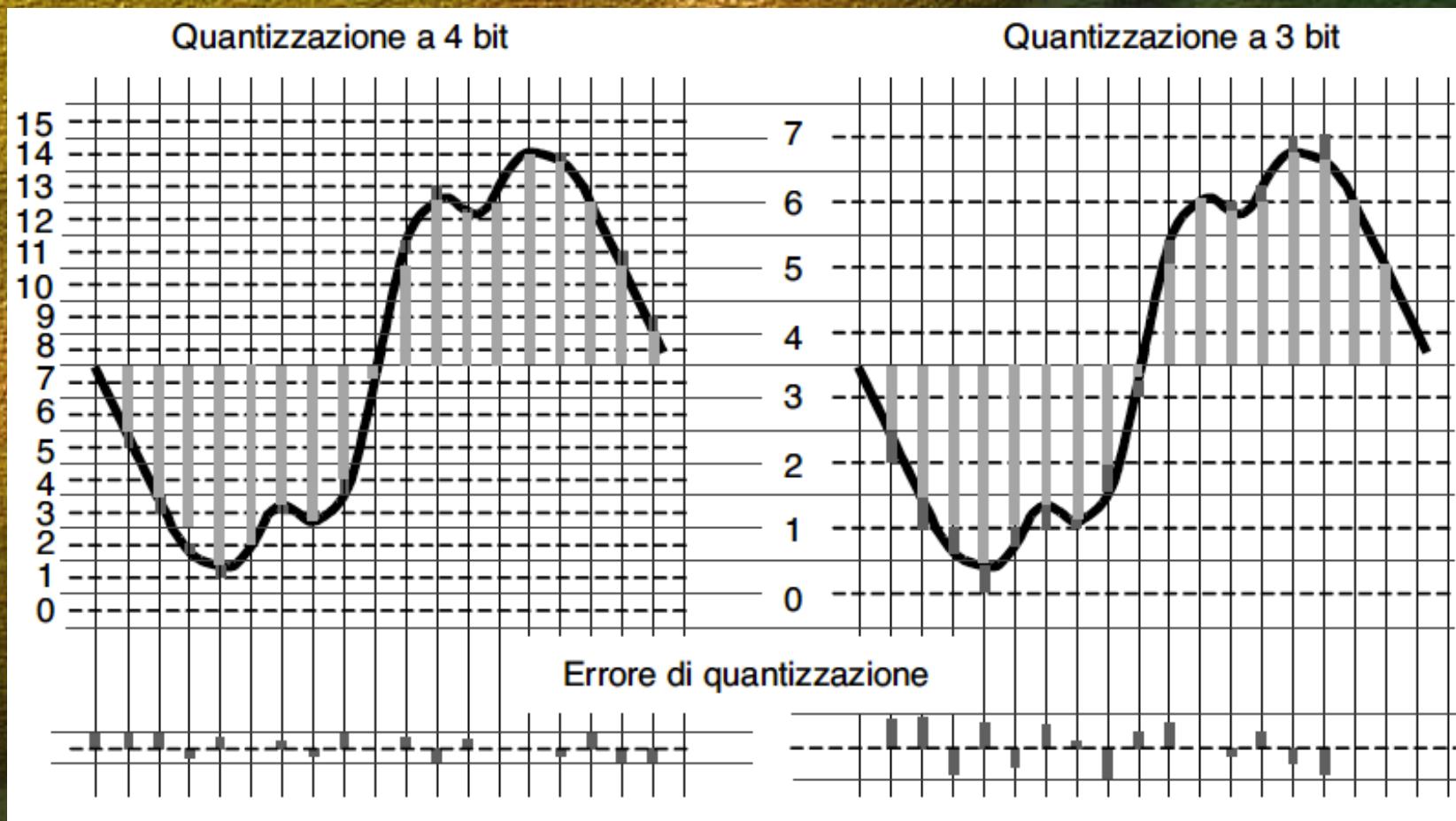
accuratezza

errore di quantizzazione

# Livelli di quantizzazione

<b>Valori di tensione</b>	<b>Parola binaria</b>
-5/-4,961	0000 0000
-4,960/-4,922	0000 0001
-4,921/-4,883	0000 0010
-4,882/-4,844	0000 0011
-4,843/-4,805	0000 0100
...	...
...	...
4,805/4,844	1111 1011
4,845/4,883	1111 1100
4,884/4,922	1111 1101
4,923/4,961	1111 1110
4,962/5	1111 1111

# Quantizzazione a 4 e 3 bit



# Limite in precisione

- Segnale rappresentato = segnale effettivo + rumore
- Analogico: differenze con il segnale effettivo = “rumore” non limitato
- Digitale: “rumore” di quantizzazione “limitato”

# Errore di quantizzazione

- Differenza tra ampiezza reale e ampiezza quantizzata
- Aampiezza quantizzata = metà della regione (si usa in riproduzione)
- L'errore di quantizzazione è al più la metà della regione di quantizzazione

# Rumore di quantizzazione

- Errore di quantizzazione distribuito casualmente
- Segnale in aggiunta al segnale utile
- Si parla di “rumore di quantizzazione”

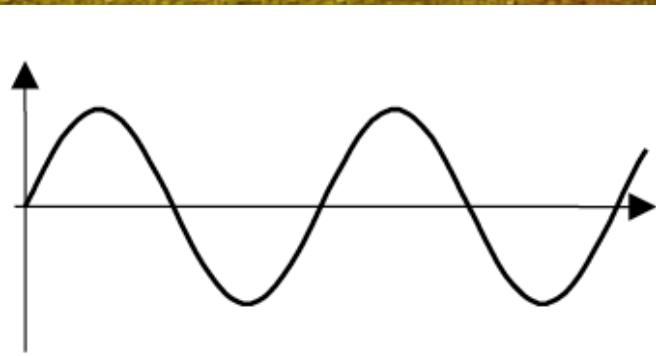


Quantizzazione con 16 bit

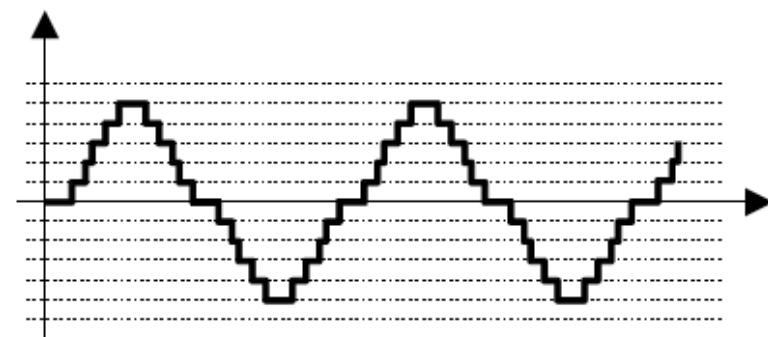


Quantizzazione con 8 bit

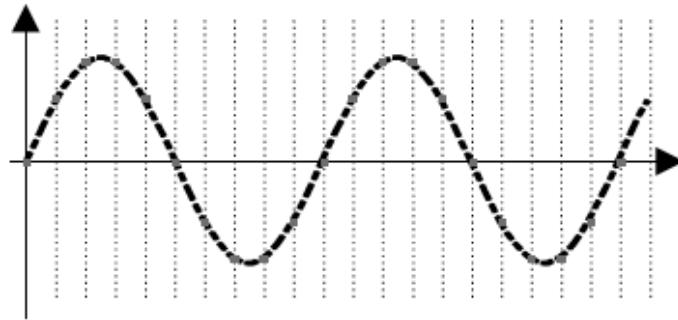
# Campionamento e quantizzazione



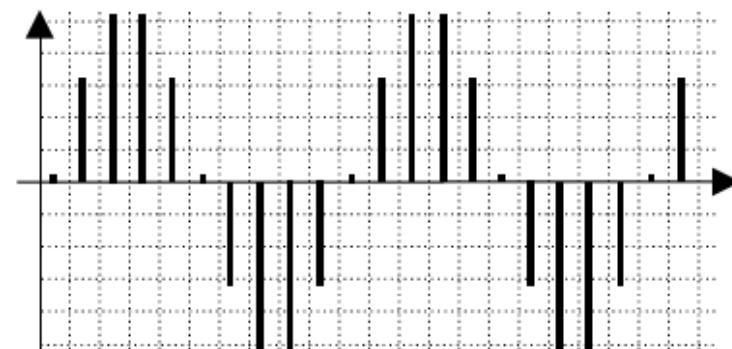
(a)



(b)



(c)



(d)

# Valutazione della rappresentazione digitale

Rapporto segnale-rumore

Gamma dinamica

# Contributo di un bit a SQNR

- Se il numero di bit è almeno 6 o 7 ...

$$SQNR = 2^N$$

In decibel, ogni bit contribuisce con circa 6 dB

$$SQNR = 20 \log 2^N \text{ dB} = N * 20 \log 2 \text{ dB} = 6.02 * N \text{ dB}$$

- Qualità CD:

- $N = 16$  bit
- $SQNR = 16 * 6 \text{ dB} = 96 \text{ dB}$

# Contributo di un bit alla gamma dinamica

L'aggiunta di un bit

- raddoppia la gamma dinamica dei valori
- dimezza l'errore di quantizzazione



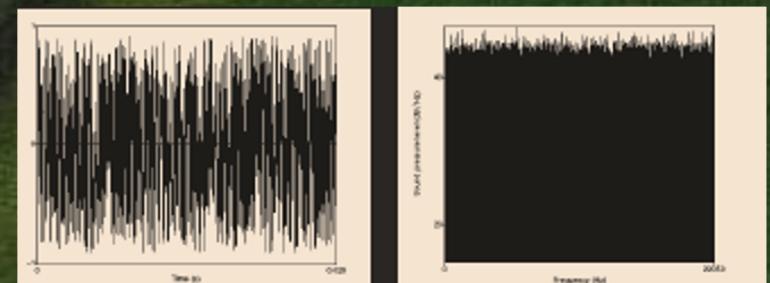
# Dithering

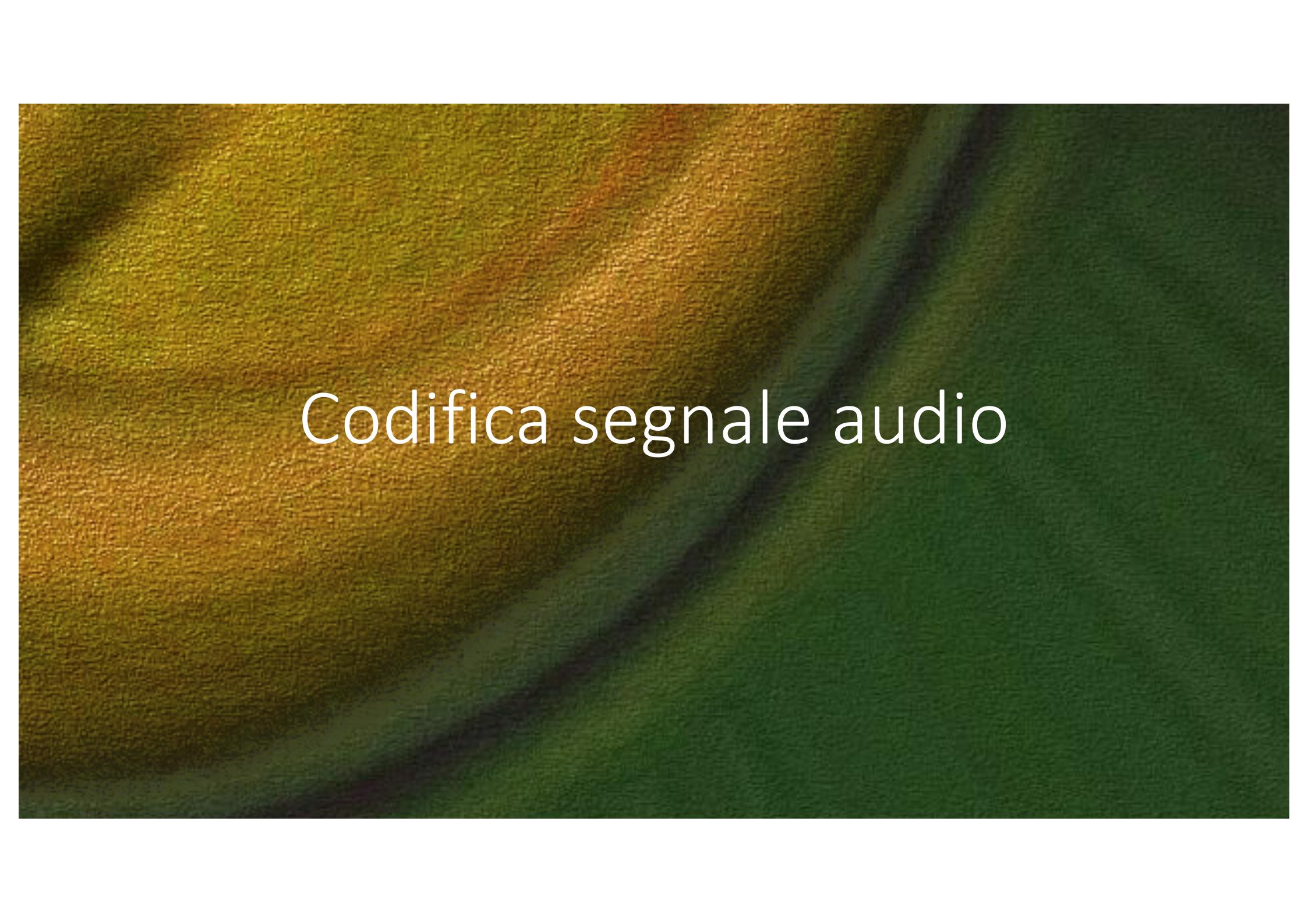
- Problemi a livelli vicini allo 0 (non nulli)
- FLIP periodico dell'ultimo bit, onda quadra
- Due soluzioni
  - Forza bruta: Aumento del numero di bit
  - Uso di tecniche a basso costo (dithering)



# Dithering

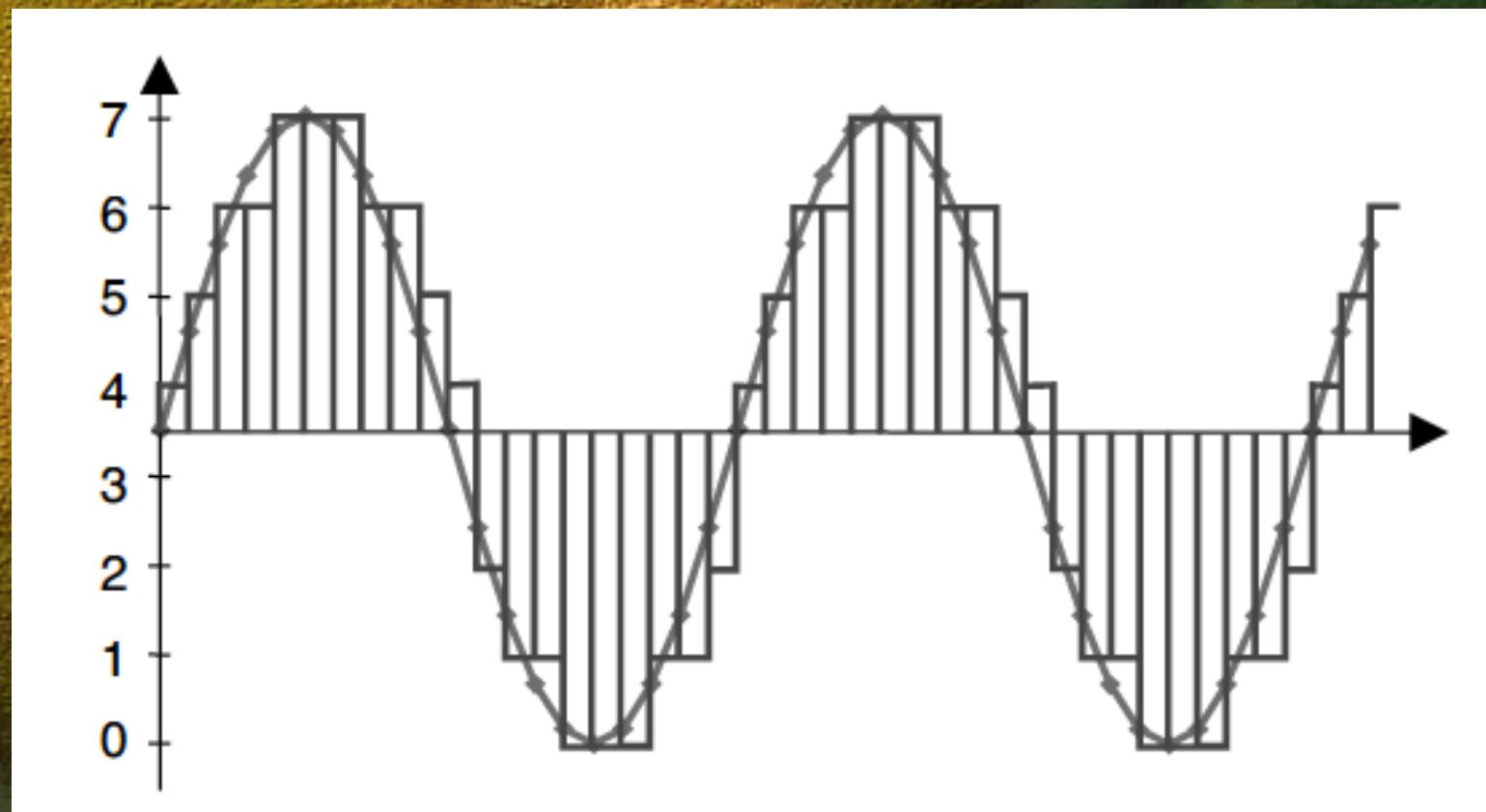
- Prima della conversione AD, *rumore di eccitazione – dither noise*
- Livello minimo (es. 3 dB, circa l'errore)
- Rumore indipendente dal segnale
- Sistemi audio di buona qualità a basso costo



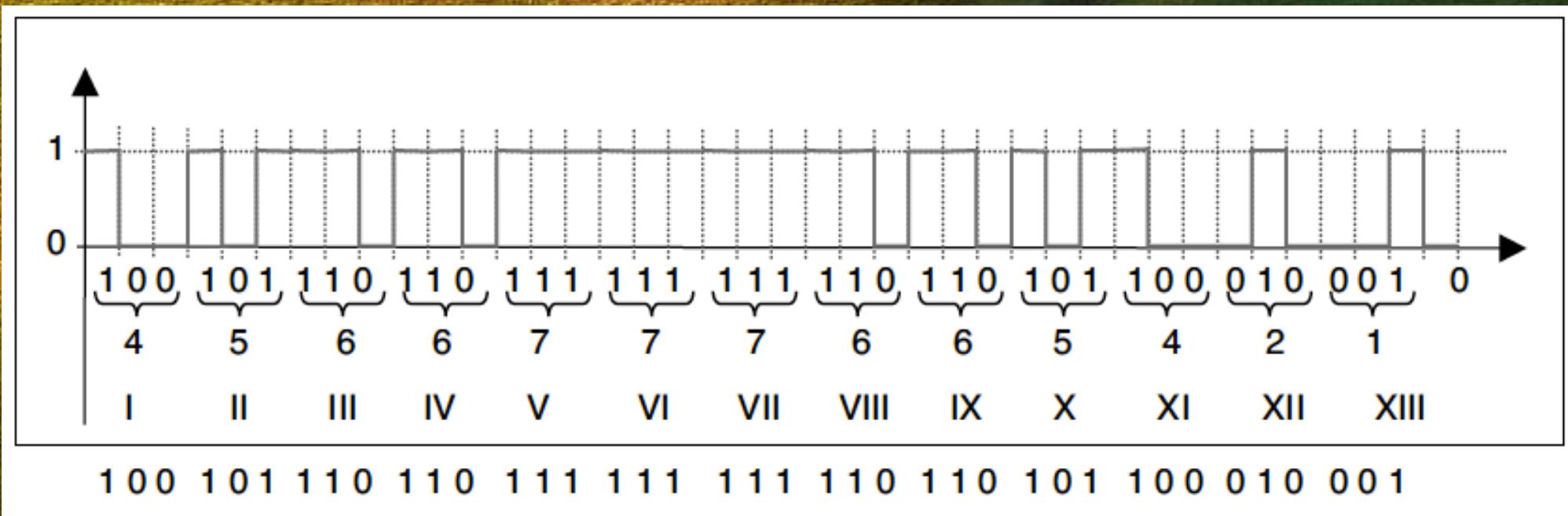


Codifica segnale audio

# Pulse amplitude modulation



# Pulse Code Modulation

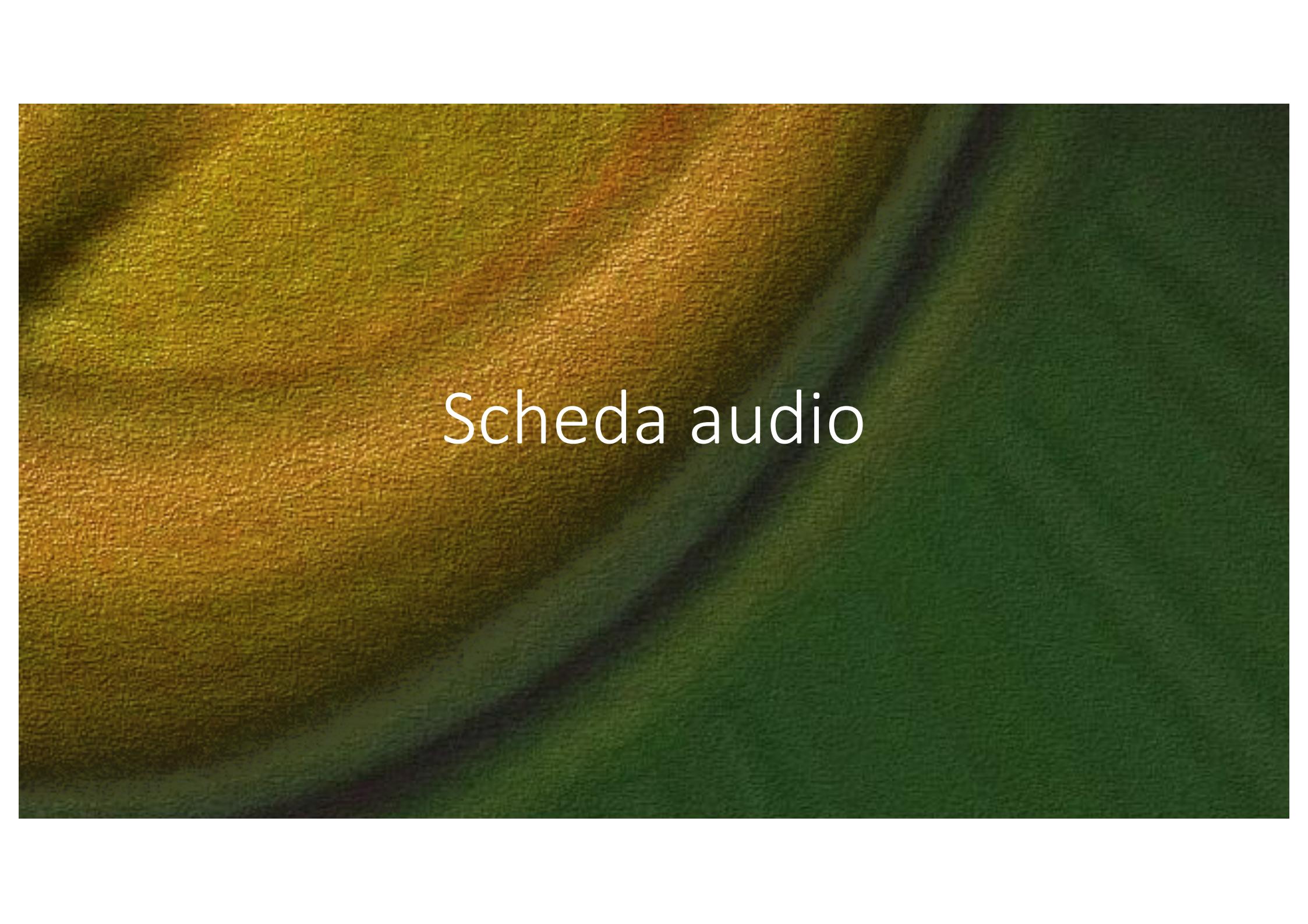


# Codifiche binarie

Unipolare

Bipolare

	Offset binary	Complemento a due	Segno e magnitudo
15	1111	0111	0111
14	1110	0110	0110
13	1101	0101	0101
12	1100	0100	0100
11	1011	0011	0011
10	1010	0010	0010
9	1001	0001	0001
8	1000	0000	0000
7	0111	1111	1000
6	0110	1110	1001
5	0101	1101	1010
4	0100	1100	1011
3	0011	1011	1100
2	0010	1010	1101
1	0001	1001	1110
0	0000	1000	1111

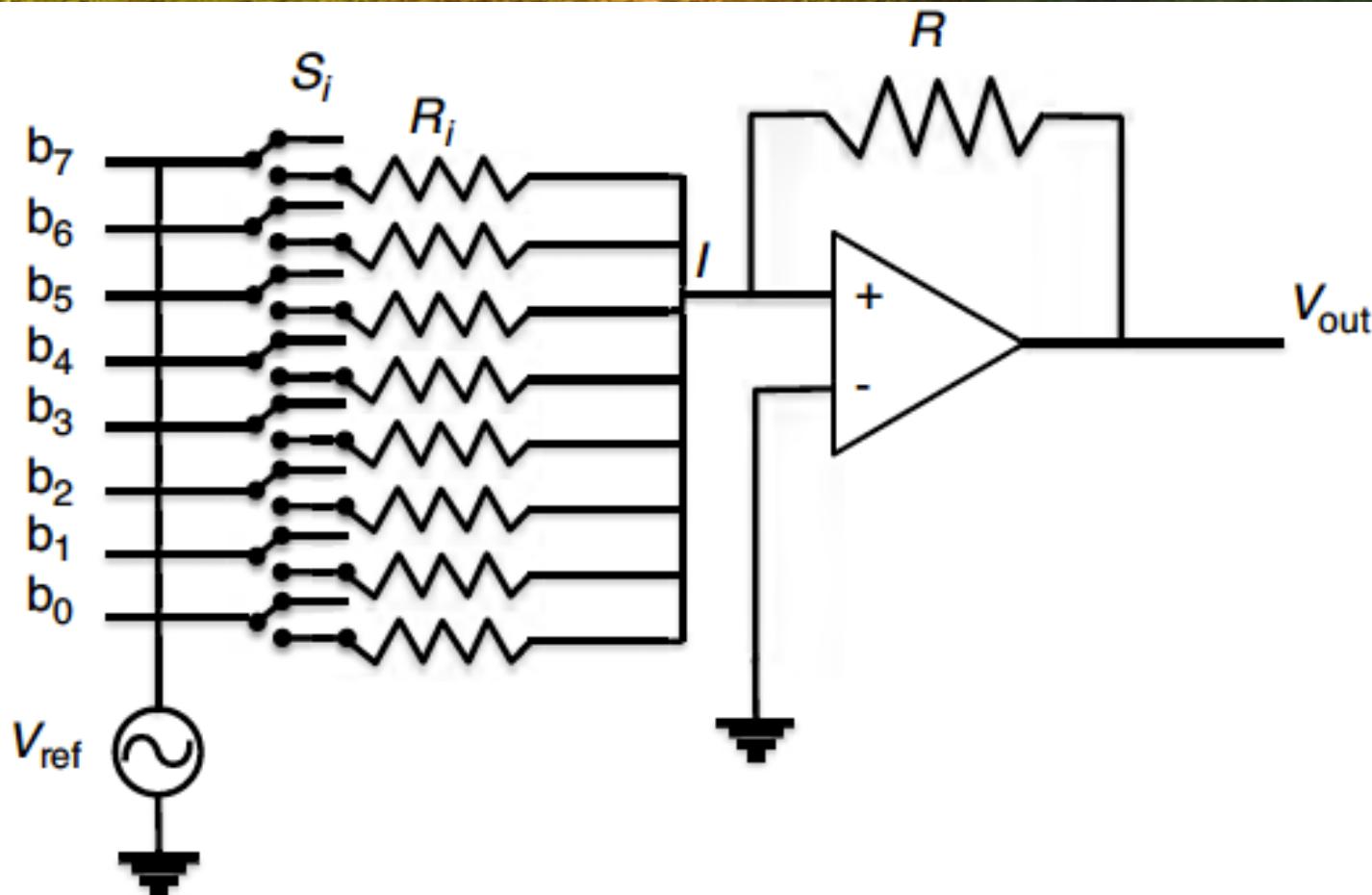


Scheda audio

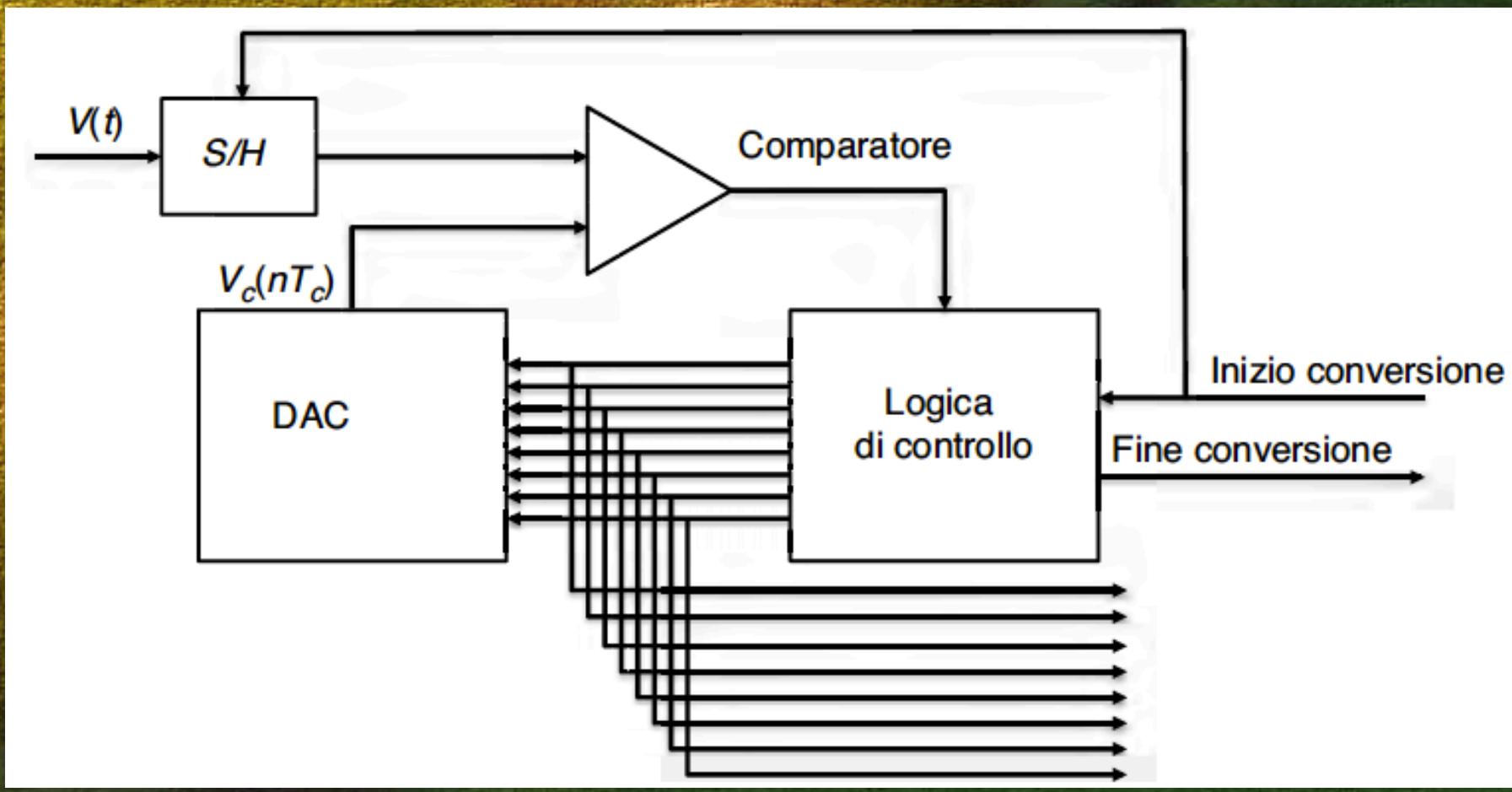
# La scheda audio

- dispositivo hardware specificamente dedicato all'audio
- Tre funzioni:
  - input/output dei segnali
  - elaborazione (DSP)
  - sintesi

# Scheda audio: DAC



# Scheda audio: ADC



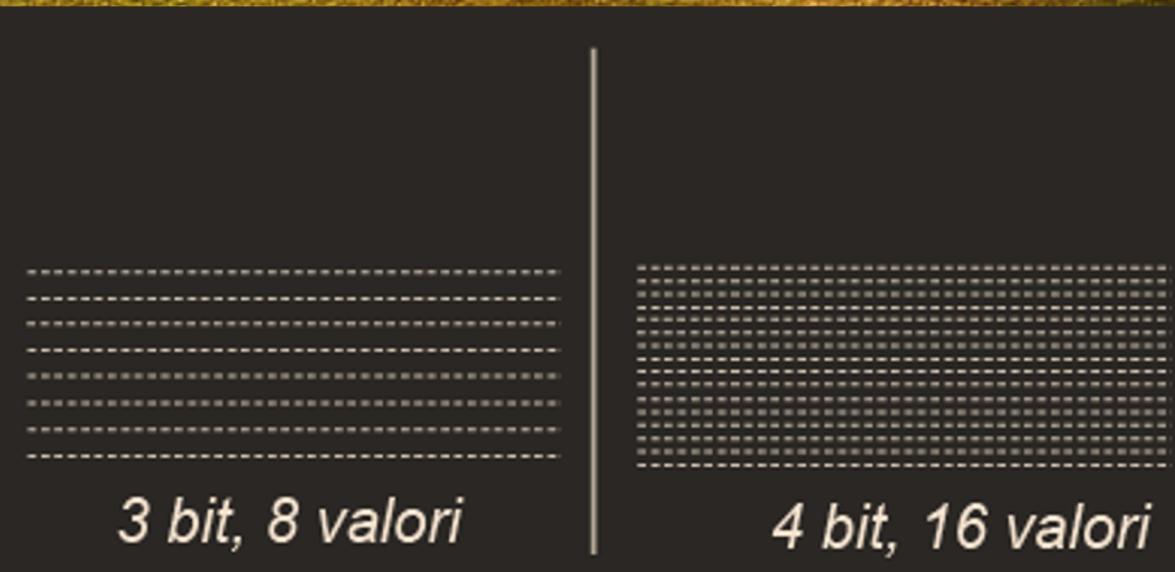
# Parametri scheda audio

- tempo di conversione (o settling time)
- rumore di fondo, misurato in dB
- Ampiezza SNR, misurato in dB (> 80-85 dB)
- Ampiezza DR (correlata a SNR)
- Uniformità risposta in frequenza (intervallo frequenze in Hz e variazione d'ampiezza in dB)
- THD (*Total Harmonic Distortion*),  $\leq 0.05 \%$
- IM (*InterModulation*) percentuale di nuove componenti spettrali non armoniche
- distorsione AM/FM, vibrato e jitter



Quantizzazione non  
uniforme

# Quantizzazione uniforme



Valori di tensione	Parola binaria
-5/-4,961	0000 0000
-4,960/-4,922	0000 0001
-4,921/-4,883	0000 0010
-4,882/-4,844	0000 0011
-4,843/-4,805	0000 0100
<b>8 bit, 256 valori</b>	
4,805/4,844	1111 1011
4,845/4,883	1111 1100
4,884/4,922	1111 1101
4,923/4,961	1111 1110
4,962/5	1111 1111

# Rapporto segnale/rumore

- SQNR = rapporto tra
  - ampiezza max segnale
  - ampiezza media errore di quantizzazione
- Ampiezza media errore di quantizzazione
  - costante e indipendente dall'ampiezza
  - no correlazione segnale/digitalizzazione

## PCM lineare

- SQNR diminuisce con l'ampiezza
- Segnali deboli degradati

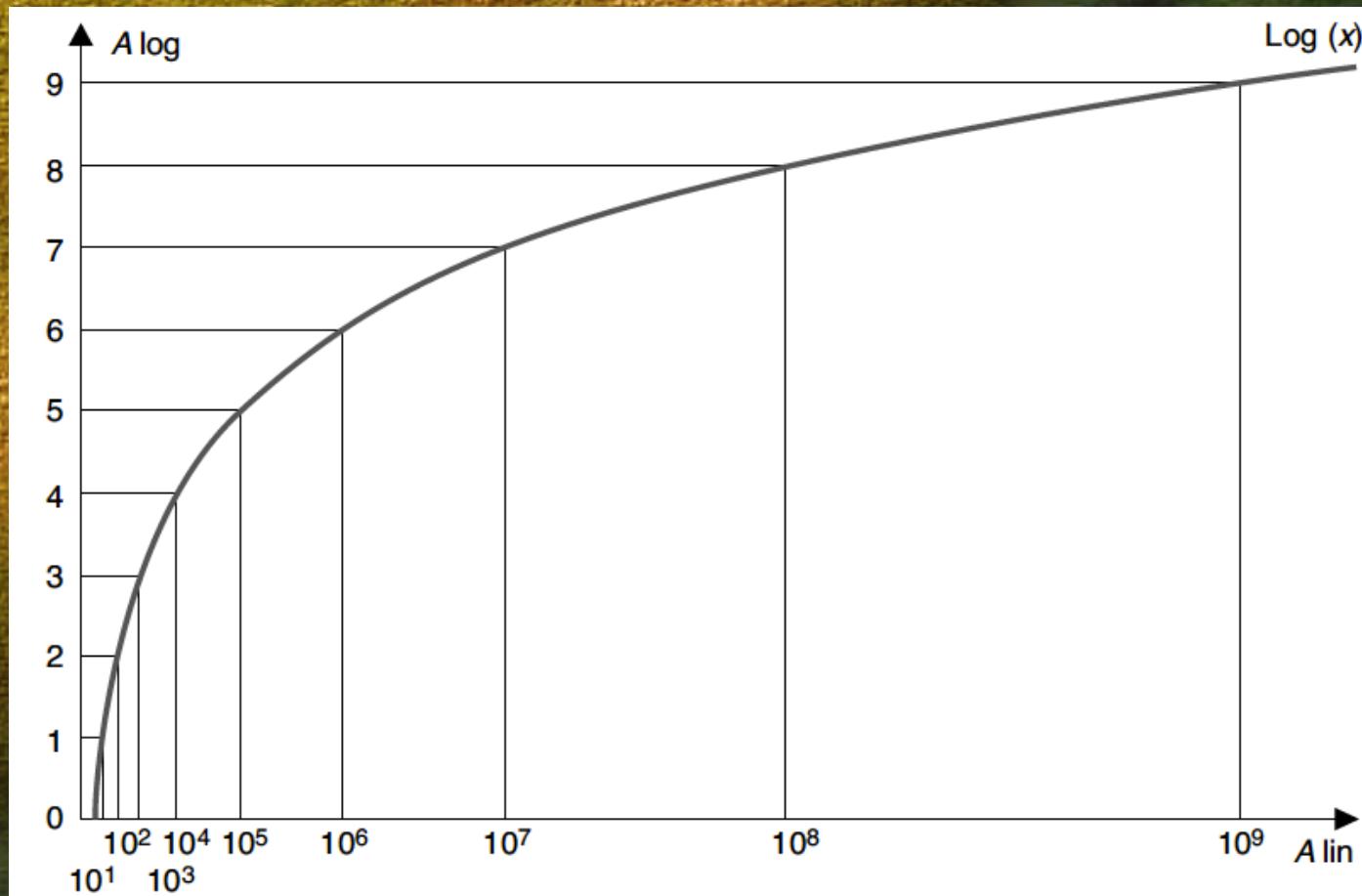
# Quantizzazione uniforme e non uniforme



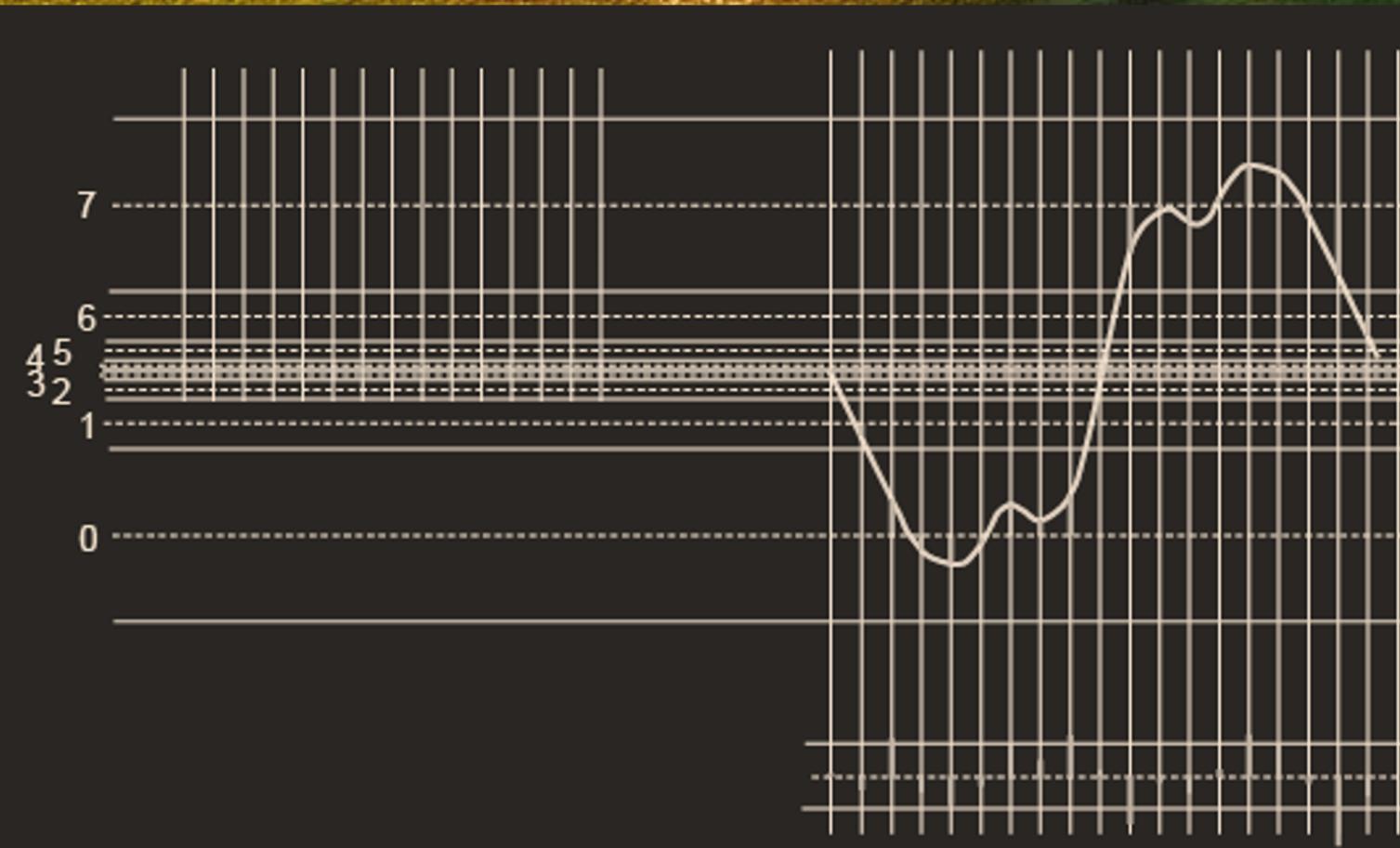
# Quantizzazione non uniforme e ampiezze ridotte

- regioni di quantizzazione ...
  - spaziate meno per ampiezze deboli
  - spaziate di più per ampiezze elevate (maggiore errore di quantizzazione)
- la quantizzazione non uniforme slega SNR e gamma dinamica

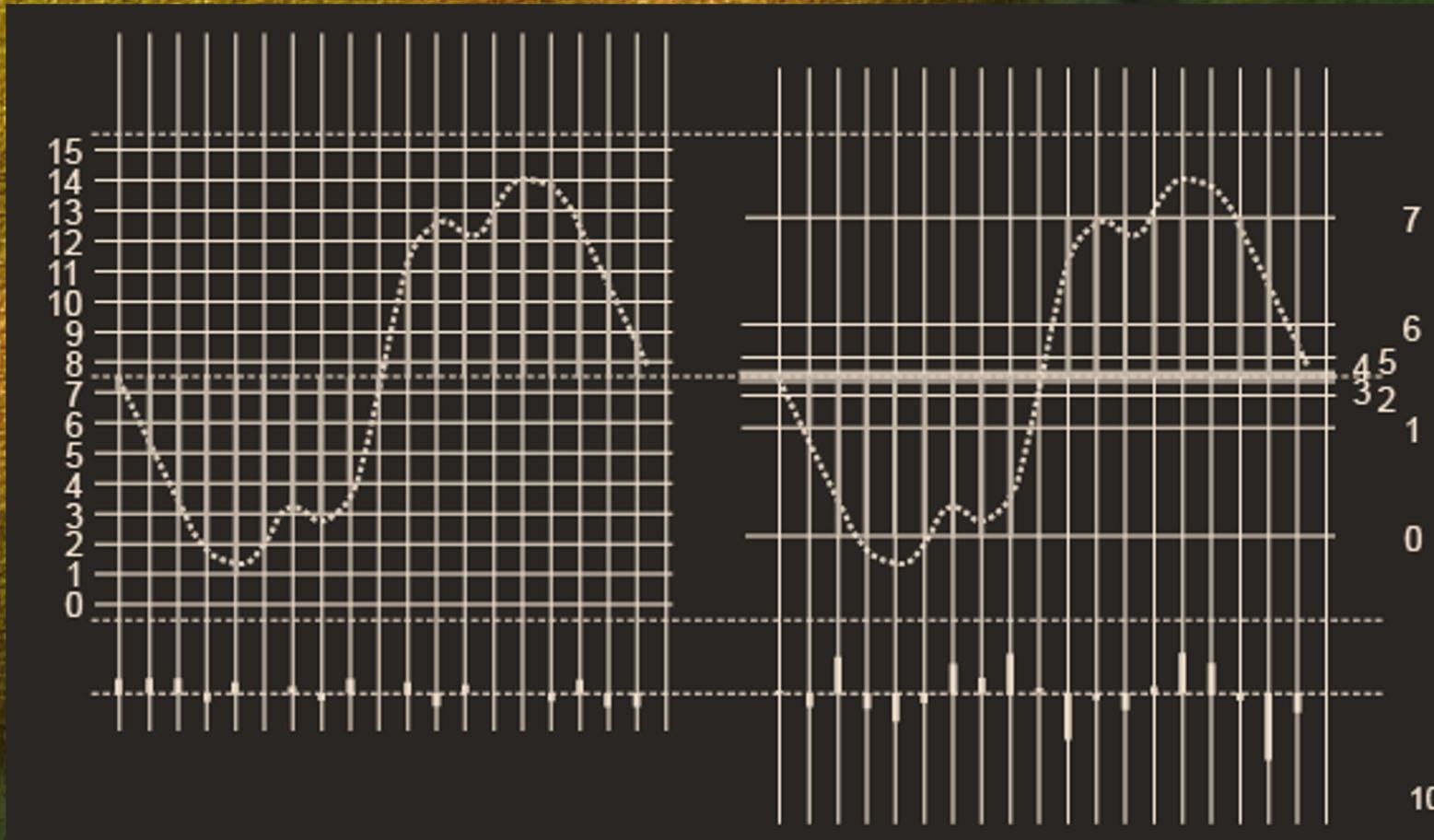
# Quantizzazione logaritmica



# Nuovo errore di quantizzazione



# 4 bit linear VS. 3 bit log



# Qualità quantizzazione logaritmica

- Gamma dinamica: 8 bit log = 13-14 bit lin
- SNR: convertitore 8-bit log ...
  - ... meglio di convertitore 8-bit lin per ampiezze deboli
  - ... peggio di convertitore 8-bit lin per ampiezze forti



[Quantizzazione con 16 bit lin](#)

[Quantizzazione con 8 bit log](#)

[Quantizzazione con 8 bit lin](#)

# Quantizzazione con virgola mobile (floating point)

- Bit extra ai bit di quantizzazione lineare
- Si aggiunge un bit extra al sistema lineare (9 bit)
  - bit extra = 0, 0 (MSb) + 8 bit (meno significativi)
  - bit extra = 1, 8 bit (più significativi) + 0 (LSb)
- bit extra è un traslatore dei bit lineari (0/1 shift)

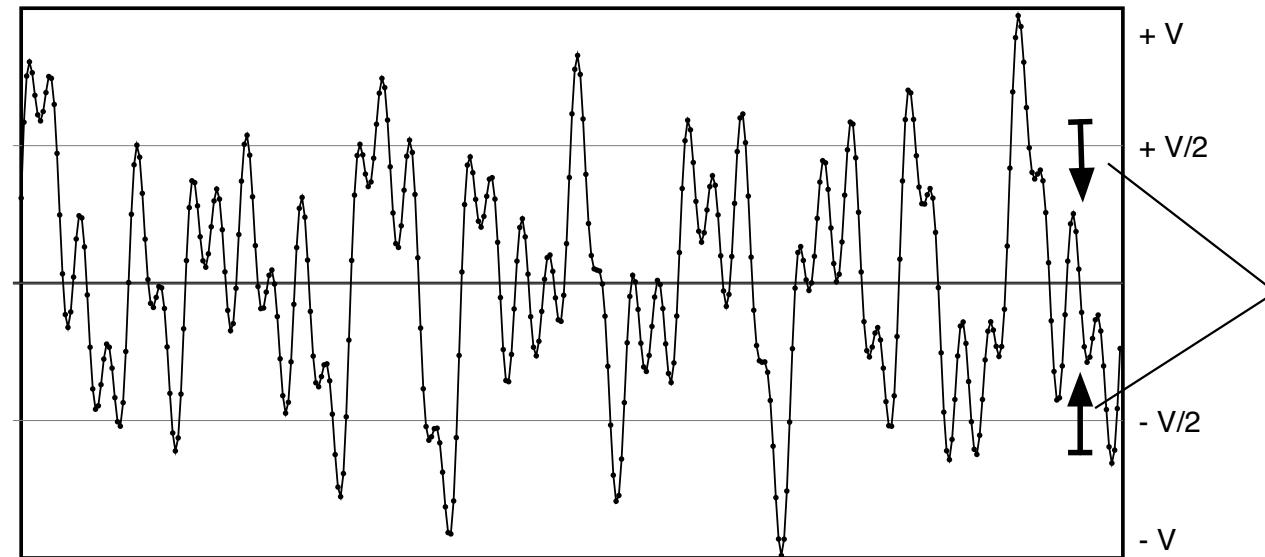
# Esempio

00001101 (bit lineari) + 0 (bit extra) – **000001101** (9 bit)

(traslazione di 0 posti)

00001101 (bit lineari) + 1 (bit extra) – **000011010** (9 bit)

(traslazione di 1 posto)

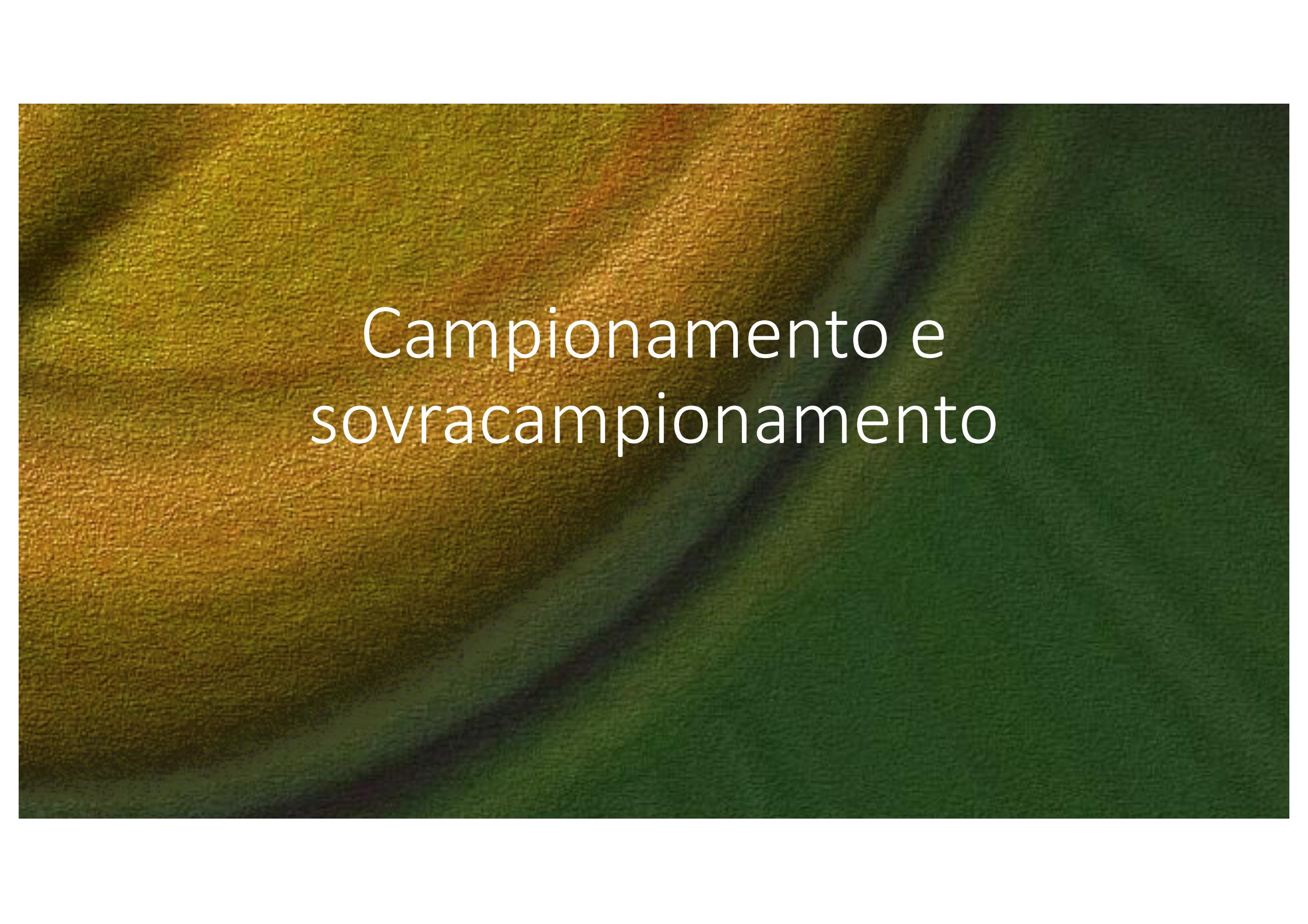


# Considerazioni

- Campioni a 8 bit sono diventati a 9 bit
  - campioni dentro intervallo (ampiezze deboli): stessa codifica (+0 iniziale)
  - campioni fuori intervallo (ampiezze forti): nuova codifica traslando i bit di un posto a sx (+ 0 finale)
- Sistema a 9 bit, mai codifiche che iniziano e finiscono con 1
  - Spaziatura regioni elevate doppia rispetto alle deboli
  - Rapporto SQNR invariato, ma DR estesa a 54 dB

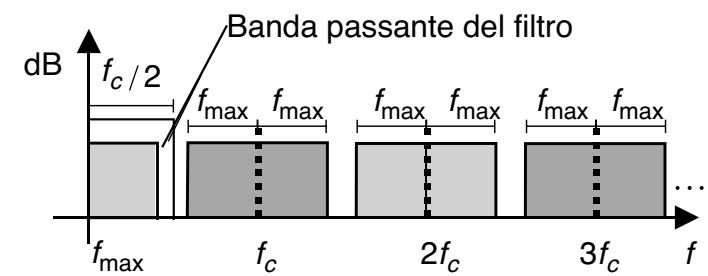
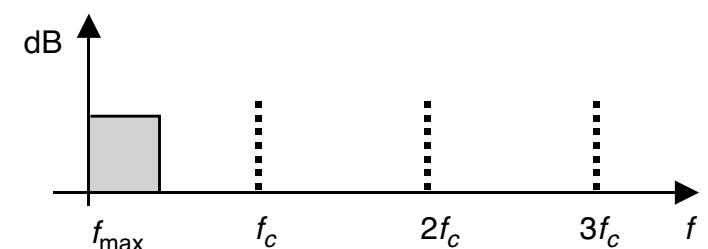
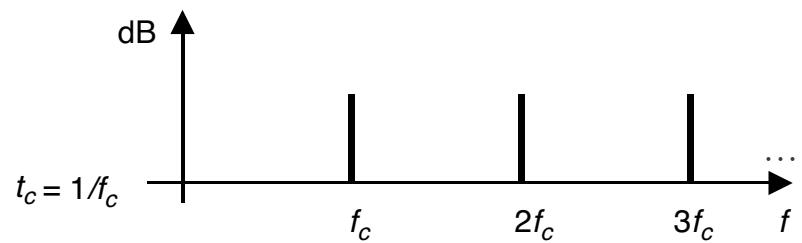
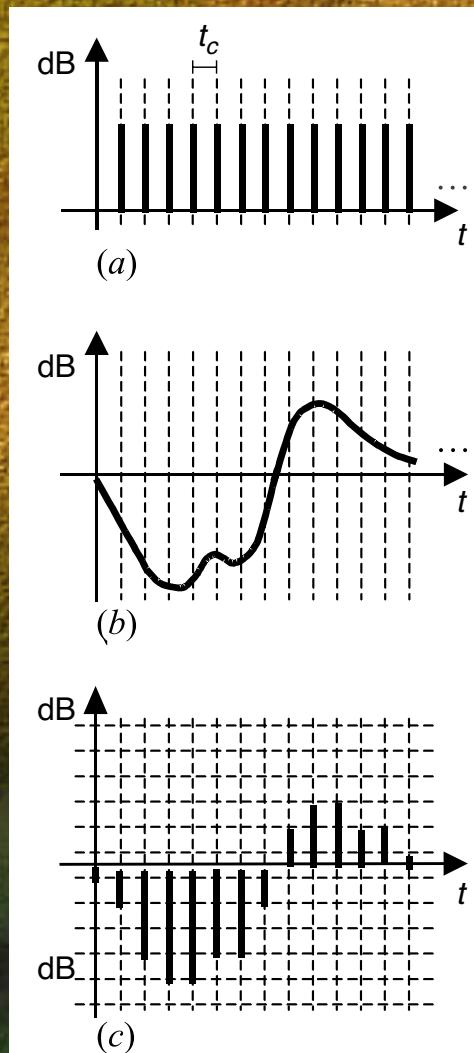
# In generale

- Si possono incrementare i bit extra (o esponente)
  - 2 bit extra 4 possibili shift, con 3 8 shift, ...
- N bit quantizzazione e P bit traslazione
  - $DR = 6N + 6P$  dB invece di  $6N$  dB del sistema lineare
- In riproduzione
  - amplificare i valori attenuati durante conversione A/D
  - amplificazione proporzionale alla attenuazione
  - *compansion* con prospettiva invertita

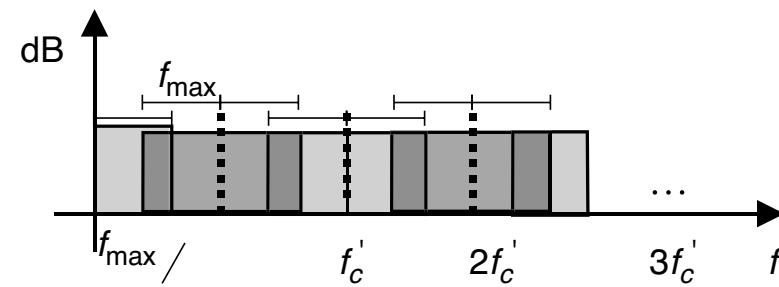
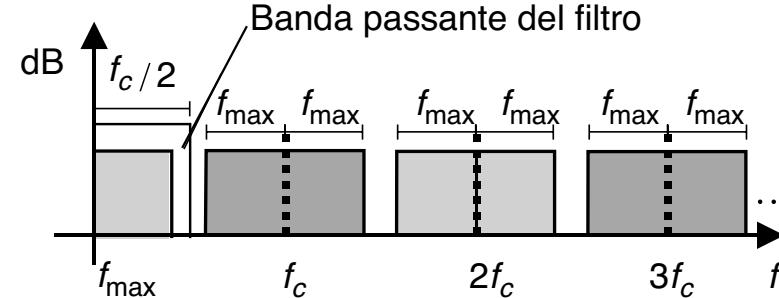
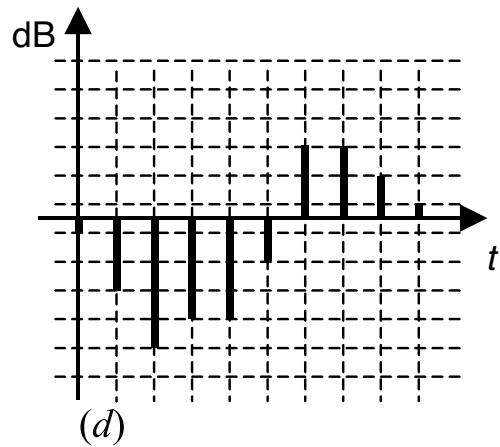
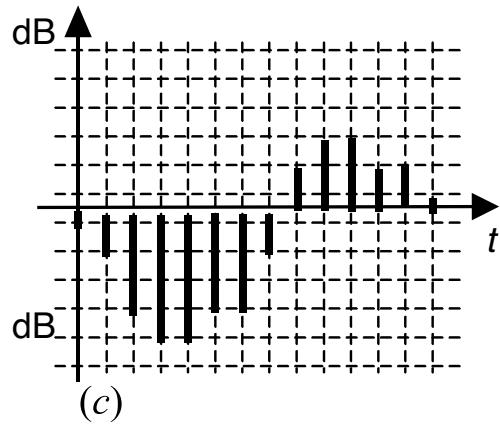


Campionamento e  
sovracampionamento

# Campionamento: modulazione di forma d'onda su treno di impulsi



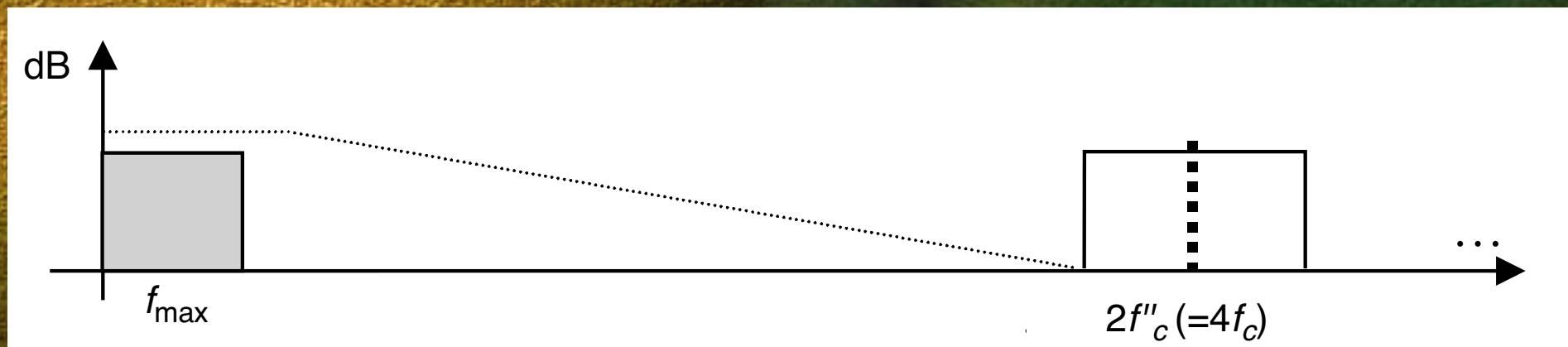
# Sottocampionamento



# Sovracampionamento



Campionamento 4×



# La teoria dell'informazione

Per comunicare una certa quantità di informazione

- larghezza di banda del segnale (aumentare campionamento)
- rapporto segnale/rumore (aumentare quantizzazione)

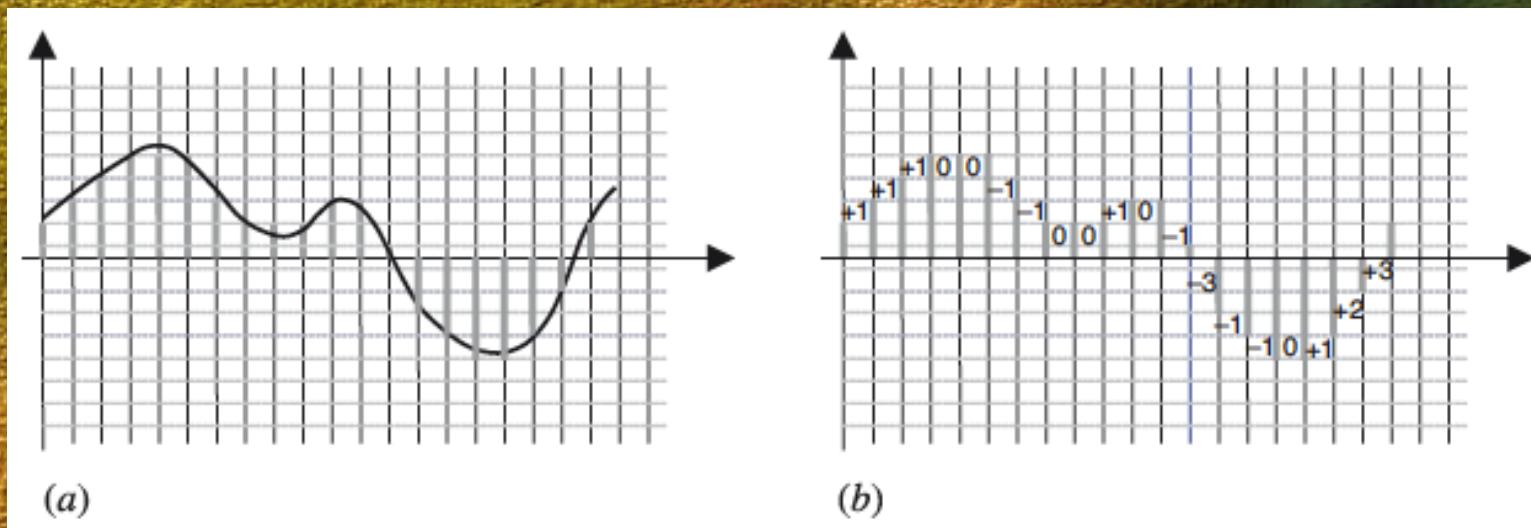
# Meglio dei due mondi

a parità di rapporto segnale/rumore, conviene

- sovraccampionare con altissimi tassi di campionamento
- codificare con parole binarie standard mediante un processo successivo di ri-quantizzazione

Modulazione  $\Sigma-\Delta$

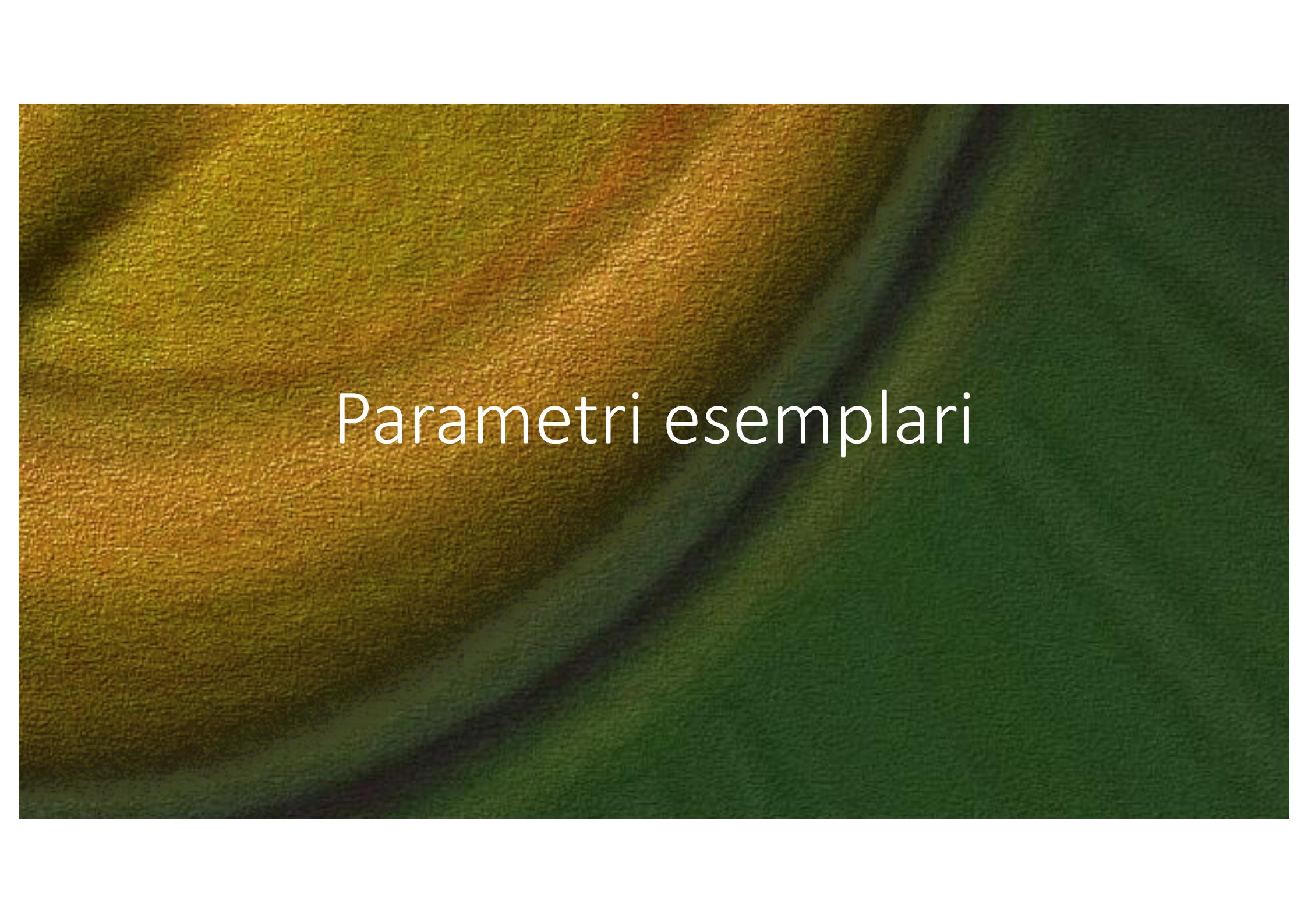
# Codifica della differenza (DPCM)



- Ridondanze temporali tra i campioni
- Differenza tra due campioni a x-bit con meno di x bit
- Si memorizza la differenza (non il campione)

# Modulazione $\Sigma$ - $\Delta$

- $\Delta$  MOD = metodo differenze con 1 bit di quantizzazione
- $\Sigma$  (somma) integrazione delle differenze per formare campioni completi
- implementa noise shaping e oversampling
- Bitstream (Philips) o Pulse Density Modulation (PDM nella comunità scientifica)



Parametri esemplari

# Tassi di campionamento

- 8 kHz (8000 campioni/sec), mu-law
- 11 kHz (in realtà 11:025 campioni/sec):  $\frac{1}{4}$  CD
- 16 kHz: G.722 (VoIP di Skype)
- 22 kHz:  $\frac{1}{2}$  CD
- 32 kHz: video, long play DAT, NICAM; pre-DAB
- 44 kHz: CD
- 48 kHz: video, DAT-audio, DAB, DV, DVD, film
- 88 kHz: 2 x CD
- 96 kHz: DVD-Audio, audio Blu-ray, HD DVD, 48 kHz
- 176 kHz: attrezzatura professionale per CD (x4)
- 192 kHz: DVD-Audio, Blu-ray, HD DVD (4x 48)
- 352 kHz: Digital eXtreme Definition (SACD), 8 x CD
- 2.8 MHz: oversampling SACD, DSD 1-bit, Modulazione  $\Delta-\Sigma$
- 5.6 MHz: oversampling SACD, 2 x DSD, registratori DSD prof

# Transfer rate

- Parametri CD
  - campionamento = 44,100 c/s
  - quantizzazione = 16 bit
  - canali = 2
- $TR = 2 \text{ can} * 44.100 \text{ c/s} * 16 \text{ bit/c} =$   
 $1.411.200 \text{ bit/s}$

Con codici correzione errore (CD): 4.321.800 bit/s

## Combinazioni standard di parametri

- 8 kHz, 8-bit,  $\mu$ -law, quant. logaritmica, mono;
- 22 kHz, 8-bit lineari senza segno, mono e stereo;
- 44,1 kHz, 16-bit lineari con segno, mono e stereo



Grazie dell'attenzione