



# Audio e digitalizzazione



# Trattamento dei suoni

- Finora abbiamo discusso della fisica del suono. Molti dei principi studiati, hanno permesso agli esseri umani di migliorare l'esperienza sonora, sfruttando ad esempio i fenomeni di diffrazione o riflessione.
- Tuttavia nasce l'esigenza di «trattare i suoni», ossia memorizzarli e riprodurli all'occorrenza, oppure trasmetterli a grandi distanze.
- Vista la natura del suono, fu evidente la necessità di trasportare l'energia associata all'onda sonora utilizzando una grandezza fisica differente.



# Audio

L' **audio** è un **segnale** elettromagnetico che trasporta informazione sonora.

- La rappresentazione di un suono tramite segnali elettromagnetici è ormai diffusa e utilizzata quotidianamente.
- Rende agevoli i processi di **registrazione, riproduzione, elaborazione, trasmissione e archiviazione**. Per esempio è semplice «memorizzare un campo elettrico», mentre sarebbe impraticabile immagazzinare direttamente l'onda di pressione che costituisce un suono.
- La televisione, il telefono, la radio sono solo alcuni esempi che dimostrano l'enorme importanza dell'**audio**.



# Audio analogico e digitale

- Come per la maggior parte dei segnali, possiamo distinguere in **audio analogico** e **digitale**.
- Ricordiamoci che un segnale si dice analogico quando è a *tempo continuo* e *valori continui*. Viceversa si dice digitale se è a *tempo discreto* e *valori discreti*.
- Normalmente, i segnali analogici hanno sempre preceduto cronologicamente la controparte digitale.



# Audio multicanale

Per migliorare l'esperienza sonora e fornire una maggiore consapevolezza dello spazio, può essere utile sfruttare più flussi informativi differenti, riprodotti e mescolati insieme.

Un segnale audio costituito da più flussi informativi differenti, prende il nome di **audio multicanale**. Ogni singolo flusso informativo può essere considerato come un segnale audio a se stante. Ovviamente questa caratteristica è indipendente dal fatto che l'audio sia analogico o digitale.

- Un audio ad un solo canale si dice **Mono**.
- Un audio a due canali si dice **Stereo**.
- ...



# Audio: Rappresentazione Analogica

- E' una rappresentazione che lavora per ***analogia***
  - la curva continua nel tempo delle variazioni di ampiezza viene rappresentata da una curva continua nel tempo delle variazioni di tensione elettrica
    - Cioè, la traccia registrata segue l'andamento della curva dell'ampiezza, dopo che questa è stata convertita in segnale elettrico



# Audio analogico - Acquisizione

Nonostante non verrà approfondito l'audio analogico, è comunque necessario dare un accenno. Non solo perché ancora oggi è diffuso, ma soprattutto per analizzare il processo di creazione di un segnale audio digitale.

Come si ottiene un audio analogico a partire da un suono?

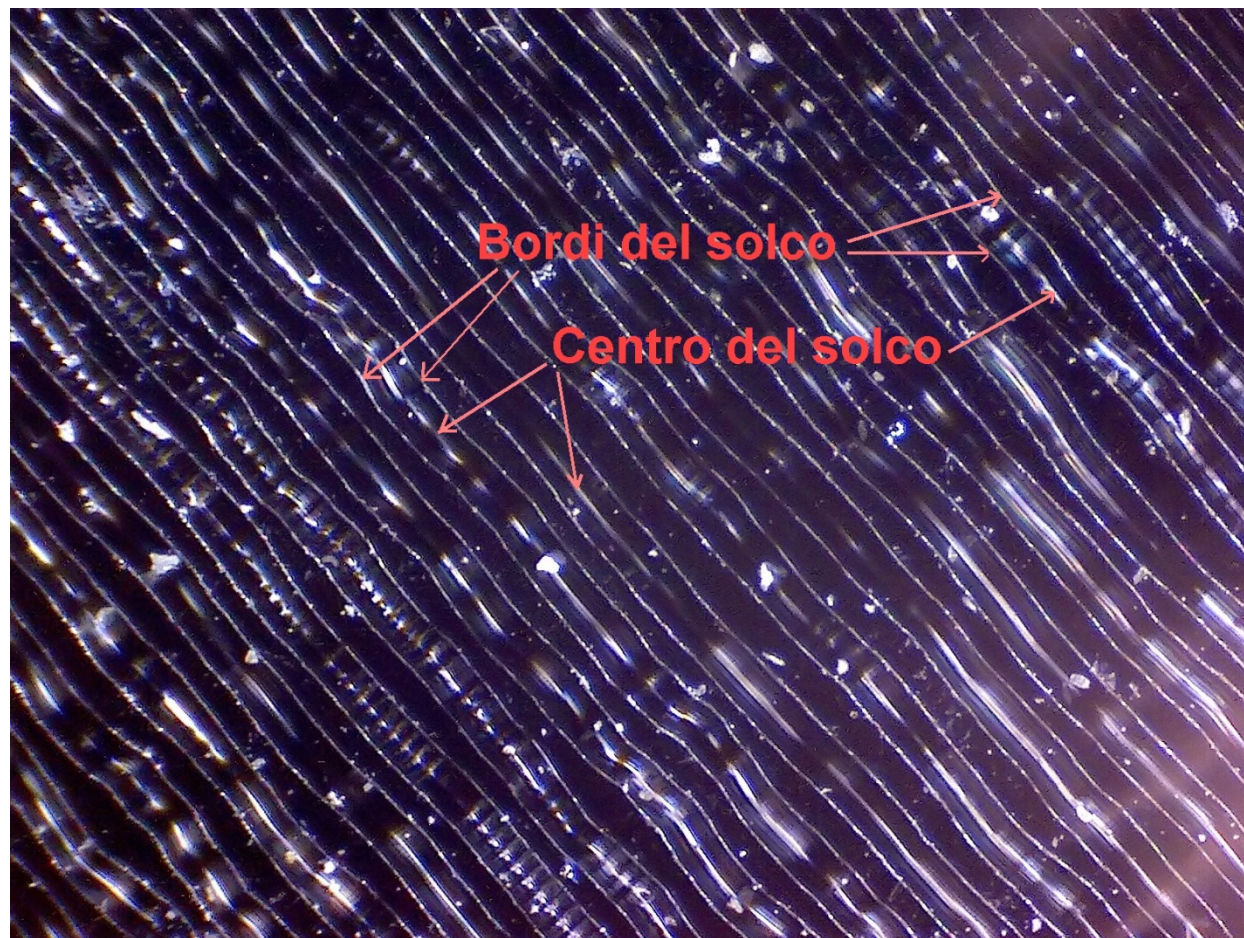
- Il suono viene rilevato da un **trasduttore** che trasforma le onde di pressione in onde elettriche (variazioni di tensione).
- Il segnale così ottenuto viene pretrattato e registrato su un supporto analogico (es: **disco in vinile, audio cassetta**).



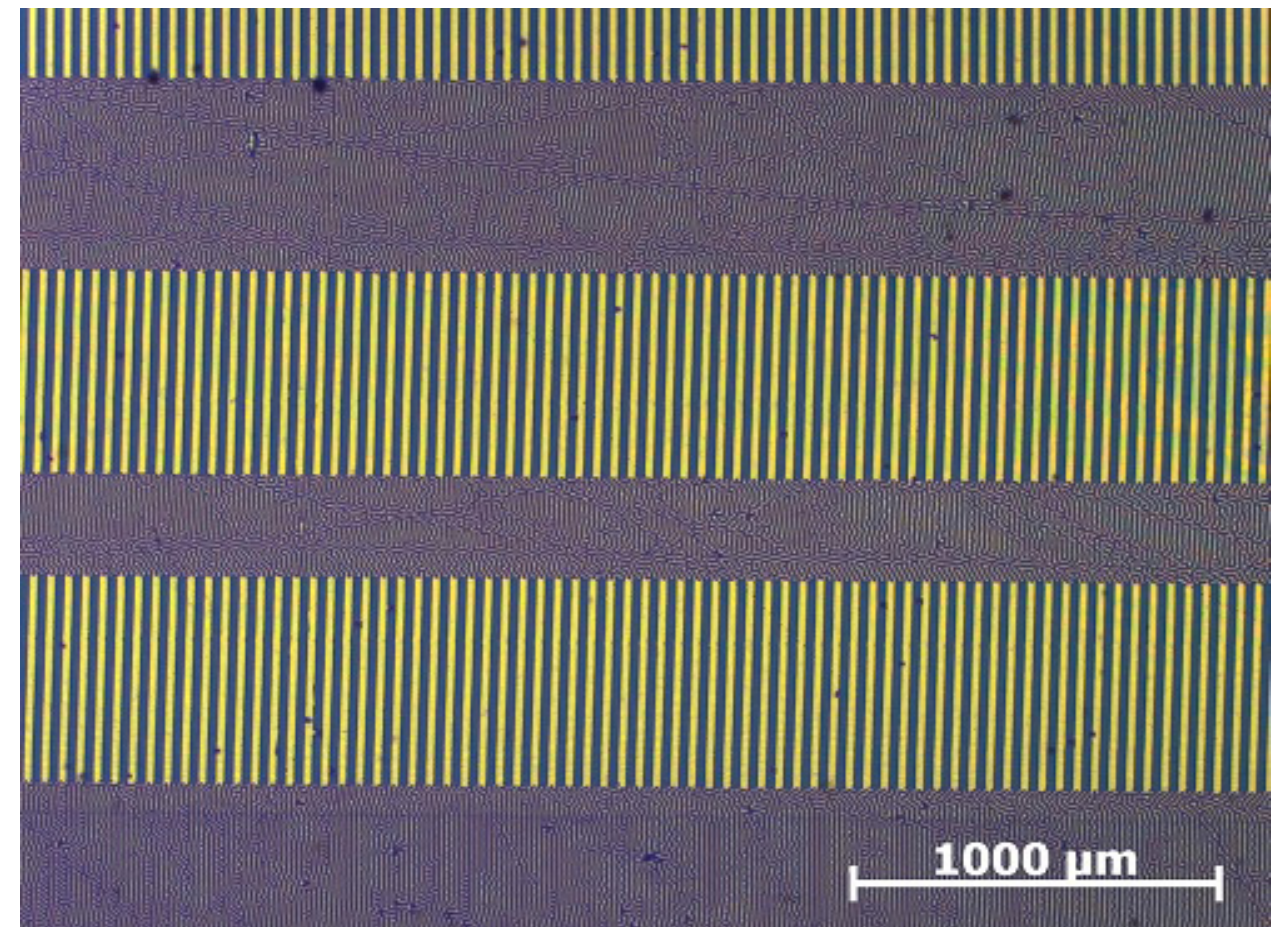


# Audio analogico - Riproduzione

Il suono originale può essere riprodotto interpretando le variazioni delle grandezze fisiche presenti sul supporto. Nel caso del disco in vinile la variazione **della profondità o irregolarità dei solchi**, mentre per le audiocassette la **variazione dell'intensità dei campi magnetici**.



Solchi di un disco in vinile

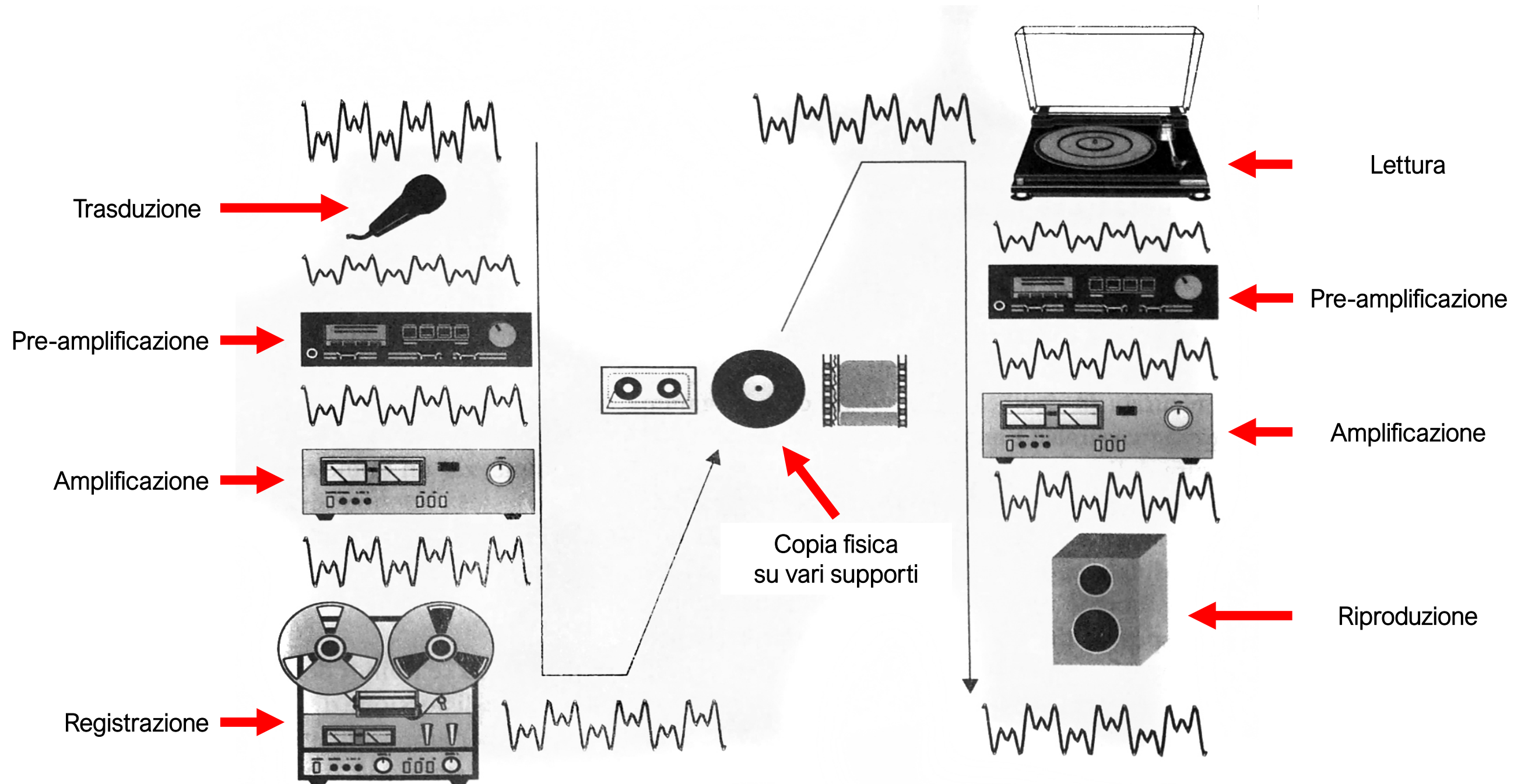


Domini magnetici allineati di un audiocassetta





# Audio analogico - Pipeline





# Audio analogico - Distorsione

Ogni trasformazione fisica a cui è sottoposto il suono originale avviene sempre con un certo errore. L'errore si propaga fino alla riproduzione dell'audio analogico, che risulterà differente rispetto al segnale di partenza. La **distorsione** introdotta viene chiamata anche **rumore**.

L'entità della distorsione può essere misurata e di solito rappresenta un indice di qualità.

Definiamo il **rapporto segnale-rumore** (Signal Noise Ratio - SNR) come:

$$SNR = \frac{S}{N}$$

Dove  $S$  è l'ampiezza massima del segnale originale e  $N$  è l'ampiezza massima della distorsione introdotta.



# Audio analogico - Distorsione

Il valore  $SNR$  può essere misurato anche in **decibel**:

$$SNR_{dB} = 10 \log_{10} \frac{S^2}{N^2} = 20 \log_{10} \frac{S}{N}$$

Ovviamente è auspicabile che il valore di  $SNR$  sia quanto più grande possibile.

I dischi in vinile usurati, introducono tic e altre imperfezioni. Alcune di queste vengono ritenute «gradevoli» dagli audiofili. Nelle audiocassette è invece tipico un fruscio di fondo.



# Audio analogico – Pro e contro

- **Pro:** permette di rappresentare facilmente gamme molto ricche di frequenze; richiede apparecchiature poco sofisticate per essere riprodotto.
- **Contro:** soggetto a distorsioni fisiche; degrado continuo dei supporti; rappresentazione dipendente dal supporto.





# Audio analogico – Pro e contro?

	Advantages	Disadvantages
Analogue	<p>Once recorded, audio is stored/archived. 2" 24 track tape is a world wide standard.</p> <p>Warm and natural sound?</p> <p>Theoretically better audio bandwidth.</p> <p>Tried and tested format.</p> <p>Editing limitations discourage constant tinkering and changing of audio.</p>	<p>Cheaper recorders suffer from distortion and tape noise/hiss.</p> <p>Tape is expensive and vulnerable to deterioration.</p> <p>Tape is becoming increasingly hard to source.</p> <p>Recorded need constant maintenance.</p> <p>Linear format - tape must be wound/rewound to the location of the recording to be heard.</p> <p>Editing of audio difficult if not impossible.</p> <p>Harder to synchronise.</p> <p>Copying deteriorates sound.</p>



# Audio digitale

- Un segnale si dice **digitale** se è *a tempo discreto* e a *valori discreti*.
- Intuitivamente ci rendiamo conto che il segnale sonoro descritto da un audio digitale, consisterà in un insieme di istanti di tempo, in cui potrà assumere uno tra un numero finito di valori di ampiezza.
- Come la maggior parte dei segnali digitali, è adatto ad essere trattato da elaboratori.



# Audio: Rappresentazione Digitale (dal testo)

- E' una rappresentazione che NON cerca di imitare la curva continua di ampiezza con una curva analoga ad essa, MA assegna dei numeri che rappresentano di volta in volta il valore dell'ampiezza in istanti successivi di tempo
  - Sarà la successione di numeri a rappresentare l'andamento della curva di ampiezza



# Audio digitale - Acquisizione

Come si ottiene un audio digitale a partire da un suono?

- Il suono viene rilevato da un **trasduttore** che trasforma le onde di pressione in onde elettriche (variazioni di tensione) → Es.: materiali *piezoelettrici* (come il quarzo)
- Il segnale così ottenuto viene pretrattato e mandato ad un **convertitore** Analogico-Digitale (**ADC**).
- In uscita dall' ADC si ottiene un segnale a tempo discreto e a valori discreti, ossia digitale.
- Il segnale digitale viene infine rappresentato in un **formato** specifico e archiviato in una memoria di massa.



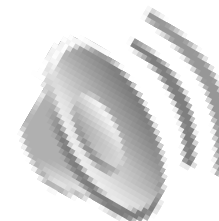




# Audio digitale - Riproduzione

## Come si ottiene il suono a partire dall'audio digitale?

- Il formato in cui è memorizzato l'audio viene interpretato e l'output inviato ad un convertitore Digitale analogico (**DAC**).
- Il **DAC** (nel nostro caso la **scheda audio**), produce un segnale elettrico che descrive il suono, ossia un audio analogico.
- Le variazioni della grandezza elettrica vengono interpretate da un dispositivo (ad esempio **casse audio**), per produrre una vibrazione, ossia la sorgente dell'onda sonora.





# Audio digitale

- Dispositivi più complessi che si occupano sia di acquisire che di riprodurre prendono il nome di dispositivi di **Data AcQuisition (DAQ)**



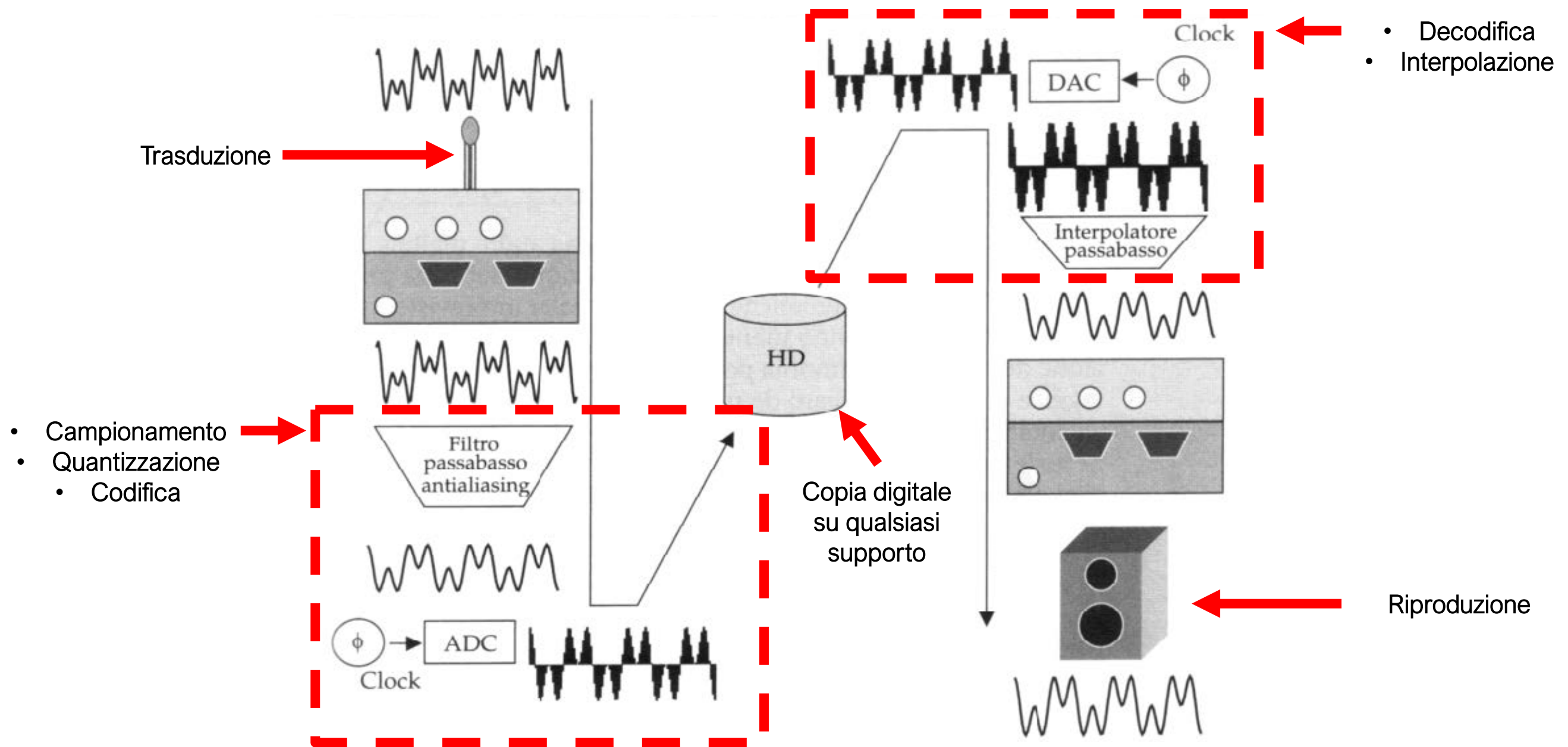


# Ma cosa c'è dentro un ADC?





# Audio digitale - Esempio Pipeline







# Audio digitale

- Prima di vedere come trattare l'audio digitale, è necessario comprendere bene come va rappresentato. Approfondiamo:
  - 1) La conversione da analogico a digitale e quindi il campionamento e la quantizzazione;
  - 2) Alcuni formati molto utilizzati;



# Conversione Analogico-Digitale

La maggior parte dei dispositivi oggi in un uso, trattano segnali digitali. Per questo motivo è necessario convertire i segnali che per loro natura sono analogici, in segnali digitali.

Questo richiede il passaggio da domini continui a domini discreti. In particolare si deve:

- Trasformare un segnale a tempo continuo in uno a tempo discreto;
- Trasformare un segnale a valori continui in uno a valori discreti.

Il primo processo prende il nome di **campionamento**, mentre il secondo **quantizzazione**.



# Campionamento

La trasformazione di un segnale a tempo continuo in un segnale a tempo discreto prende il nome di **campionamento**.

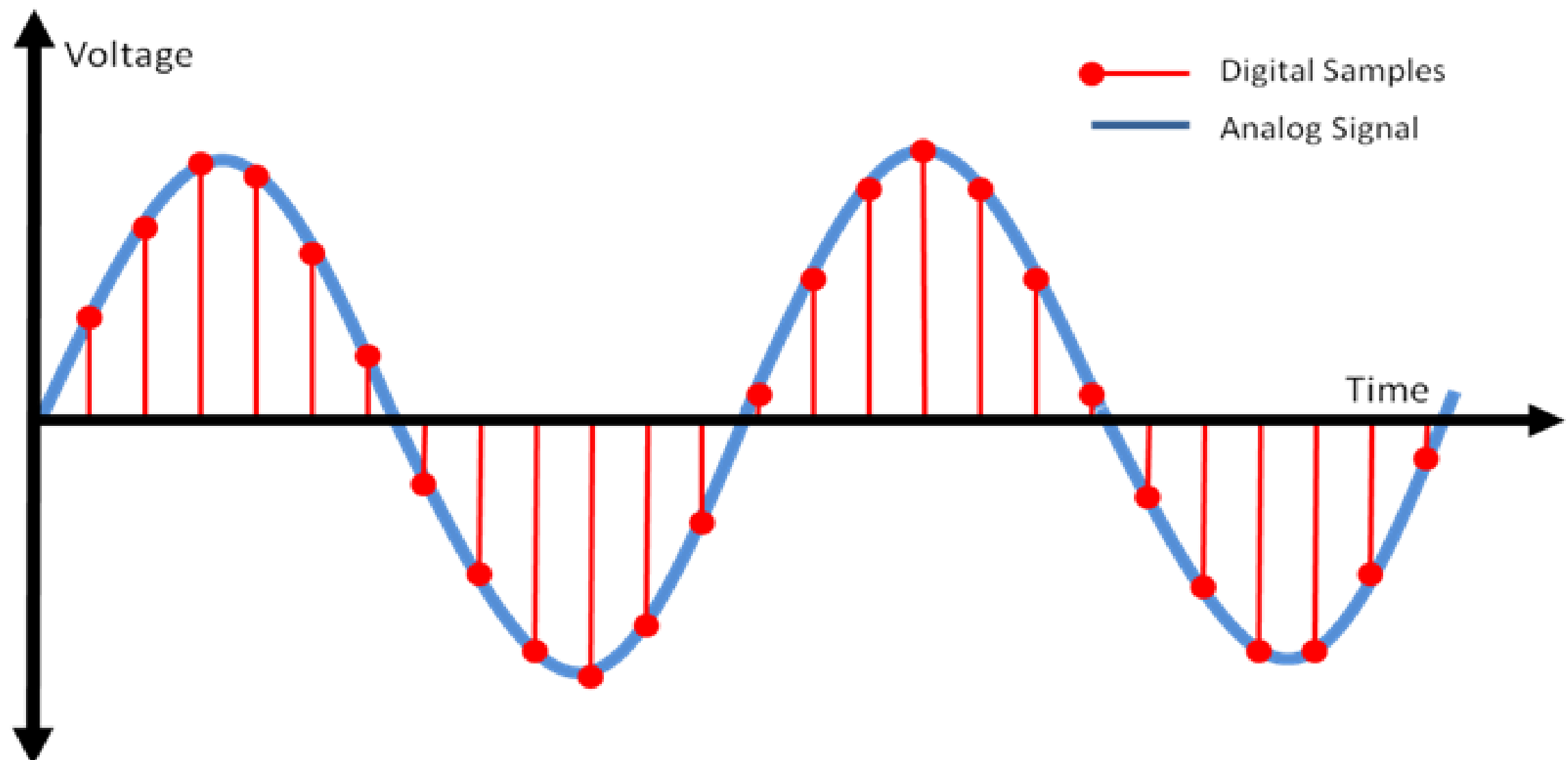
In pratica si considerano solo alcuni valori del segnale tra loro equidistanti. Si descrive cioè il segnale, utilizzando solo un numero limitato di **campioni**.

La frequenza con cui si prendono i campioni prende il nome di **sampling rate** (detta anche frequenza di campionamento).

Il problema è: quanti campioni servono per rappresentare fedelmente il segnale originale?



# Campionamento - Esempio







# Campionamento – Nyquist rate

In un segnale **periodico** e **a banda limitata**, si definisce **Nyquist rate**, il doppio della più alta frequenza presente nello **spettro** segnale.

- Per conoscere il Nyquist rate basta utilizzare la Serie di Fourier per ottenere lo spettro.
- Per segnali non periodici si può agire utilizzando la Trasformata di Fourier. Mentre, per segnali a banda non limitata (spettro non limitato), si è costretti ad usare un filtro **passa basso** per riportarsi ad un segnale a banda limitata.



# Campionamento – Teorema del campionamento di Nyquist-Shannon

Il **teorema** afferma che per poter ricostruire fedelmente un segnale campionato, è necessario che il **sampling rate**  $f_c$  sia almeno superiore al Nyquist rate  $f_N$

$$f_c > f_N$$

Attenzione!  
Con l'uguaglianza stretta si  
potrebbe ottenere il  
*Campionamento Critico.*  
*Che significa?*

Con **sampling rate** (frequenza di campionamento) ci riferiamo al numero di campioni presi per ogni secondo.



# Campionamento Critico

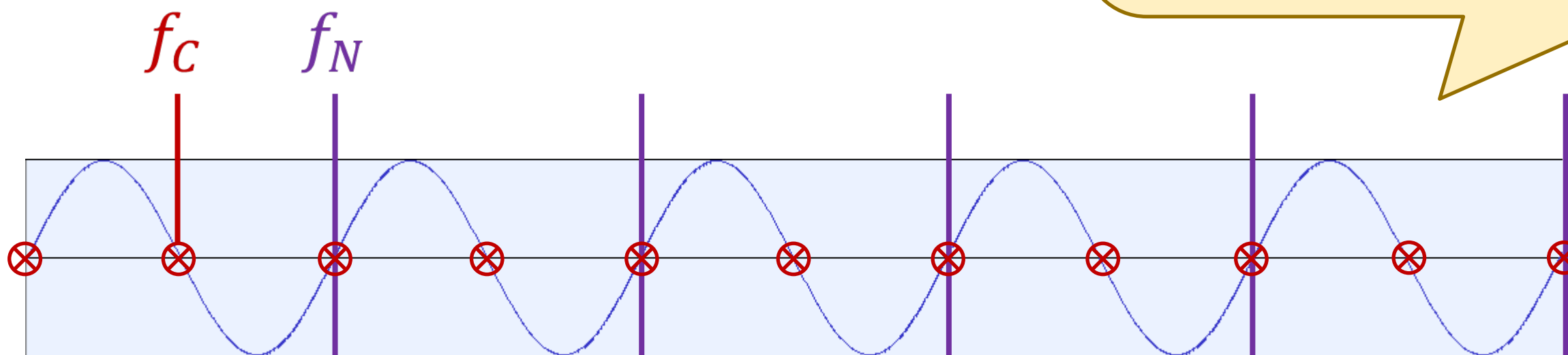
## Esempio:

### ■ Tono puro da 100Hz

□  $f_N = 100\text{Hz} \rightarrow f_C = 200\text{Hz}$

Lo stato di Campionamento Critico non porta sempre ad una ricostruzione errata del segnale. Campionare esattamente al Nyquist rate può causare questo problema, ma non è detto che accada per forza.

Ad esempio: se avessimo uno sfasamento di  $+90^\circ$  sarebbe possibile ricostruire il segnale senza nessun problema.



Come sarà il segnale ricostruito con questo campionamento?

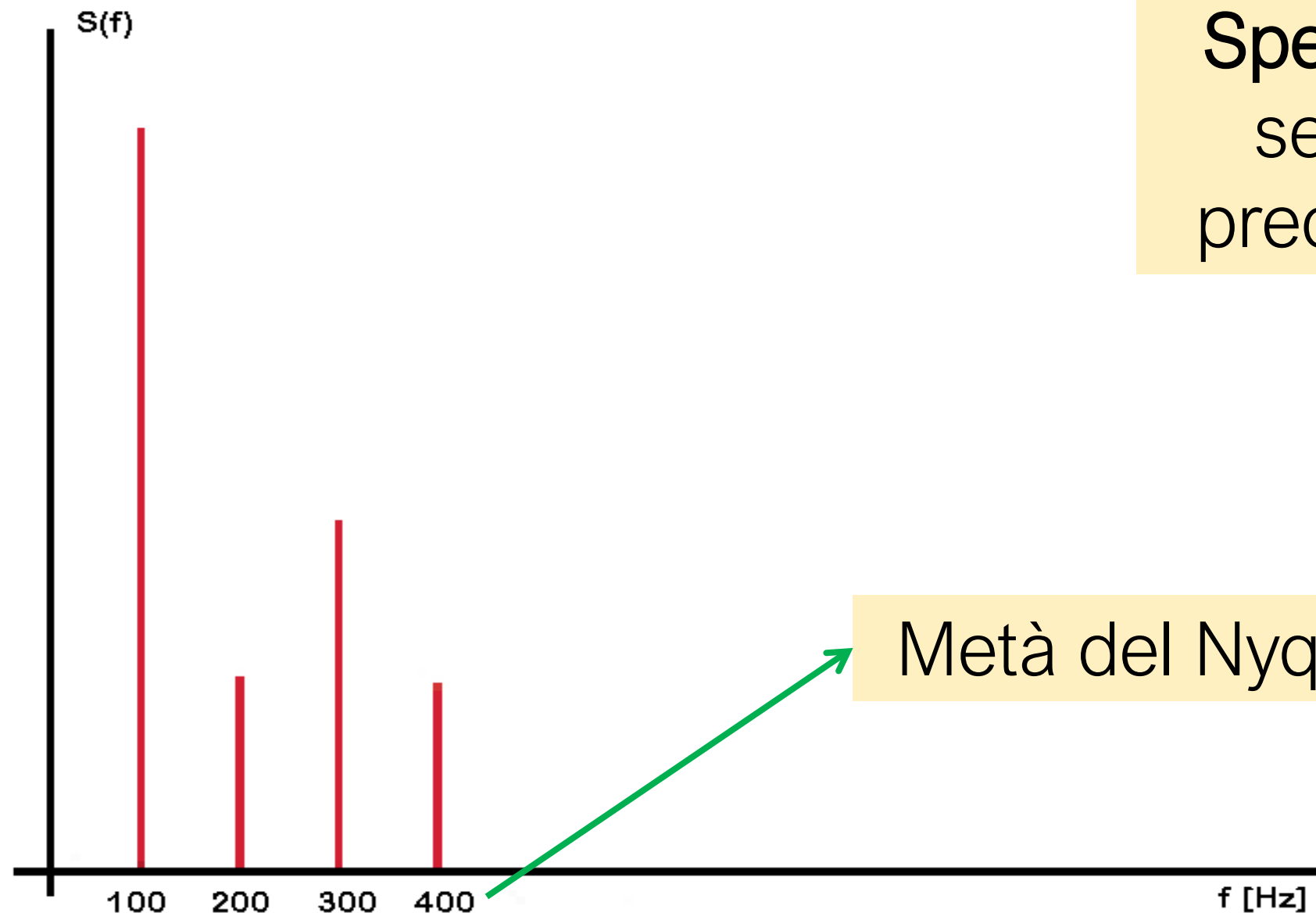


...nullo!





## ... e della frequenza



Spettro del  
segnale  
precedente

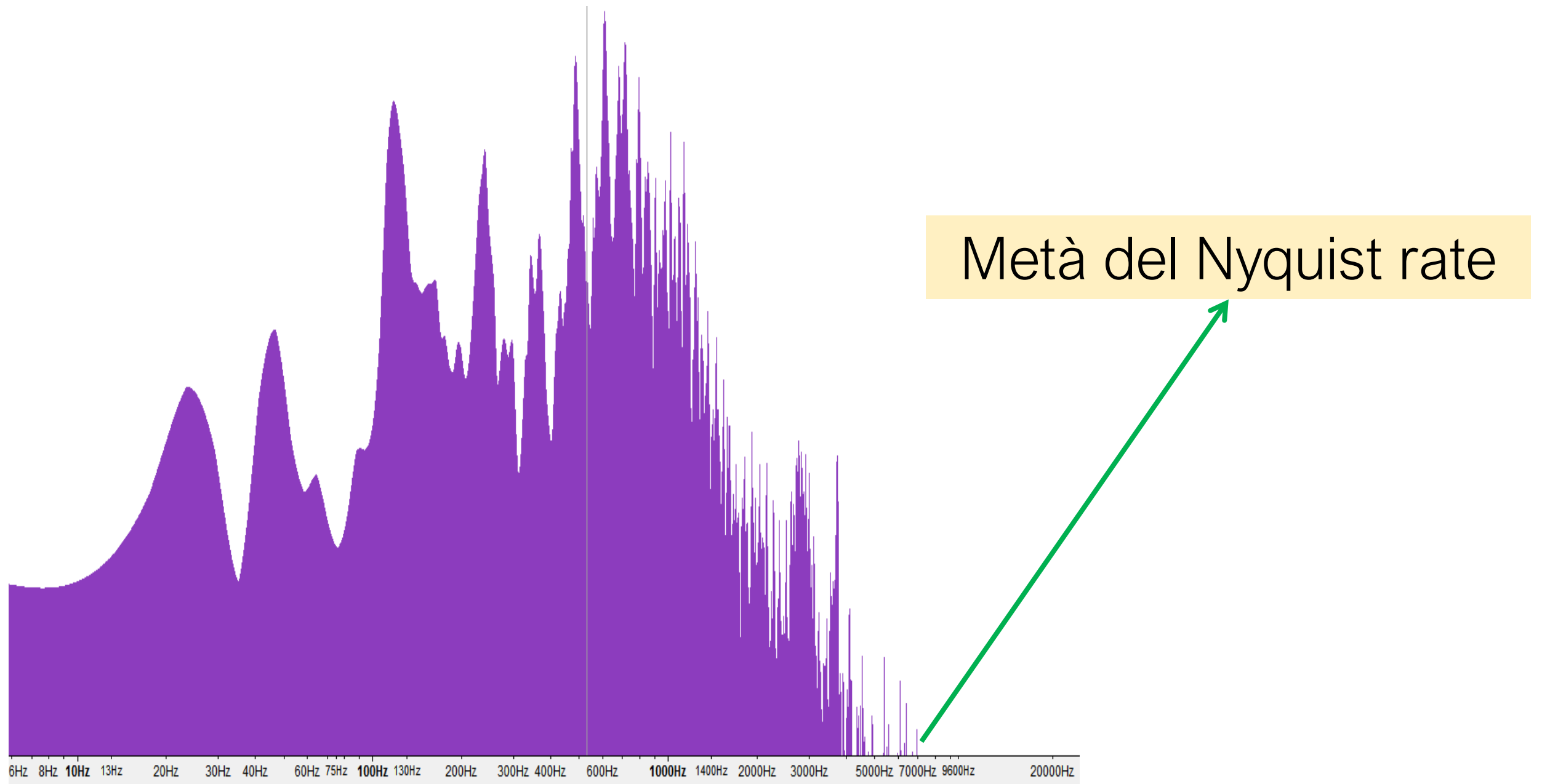
Metà del Nyquist rate

... si passa al dominio delle frequenze.





# Esempio non periodico





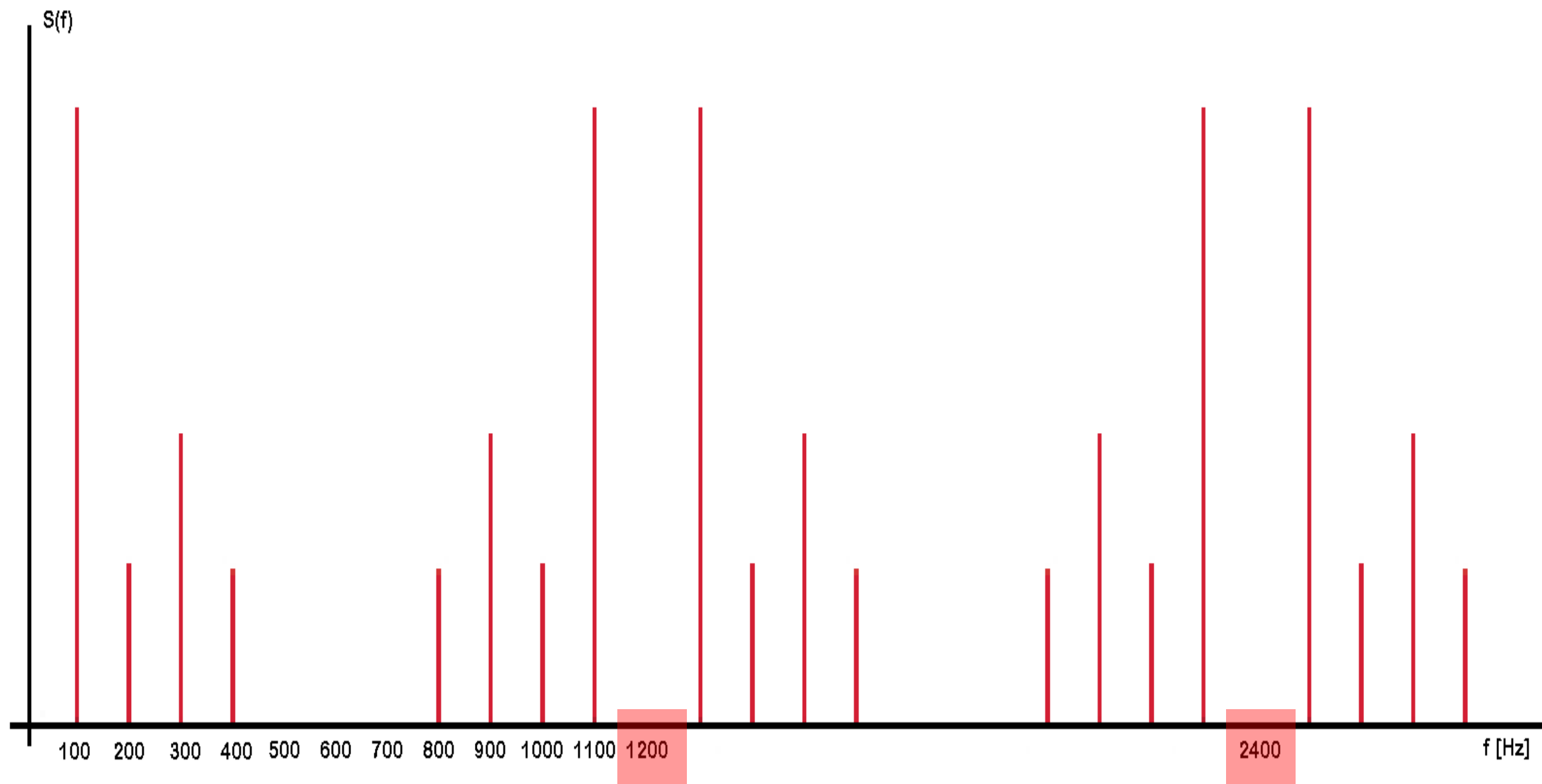
# Sampling rate

Dunque un segnale come quello precedente (periodico e a banda limitata), che ha come massima frequenza dello spettro 400 Hz, può essere ricostruito fedelmente prendendo i campioni con una frequenza di almeno 800 Hz.

In altre parole servono almeno 800 campioni al secondo.



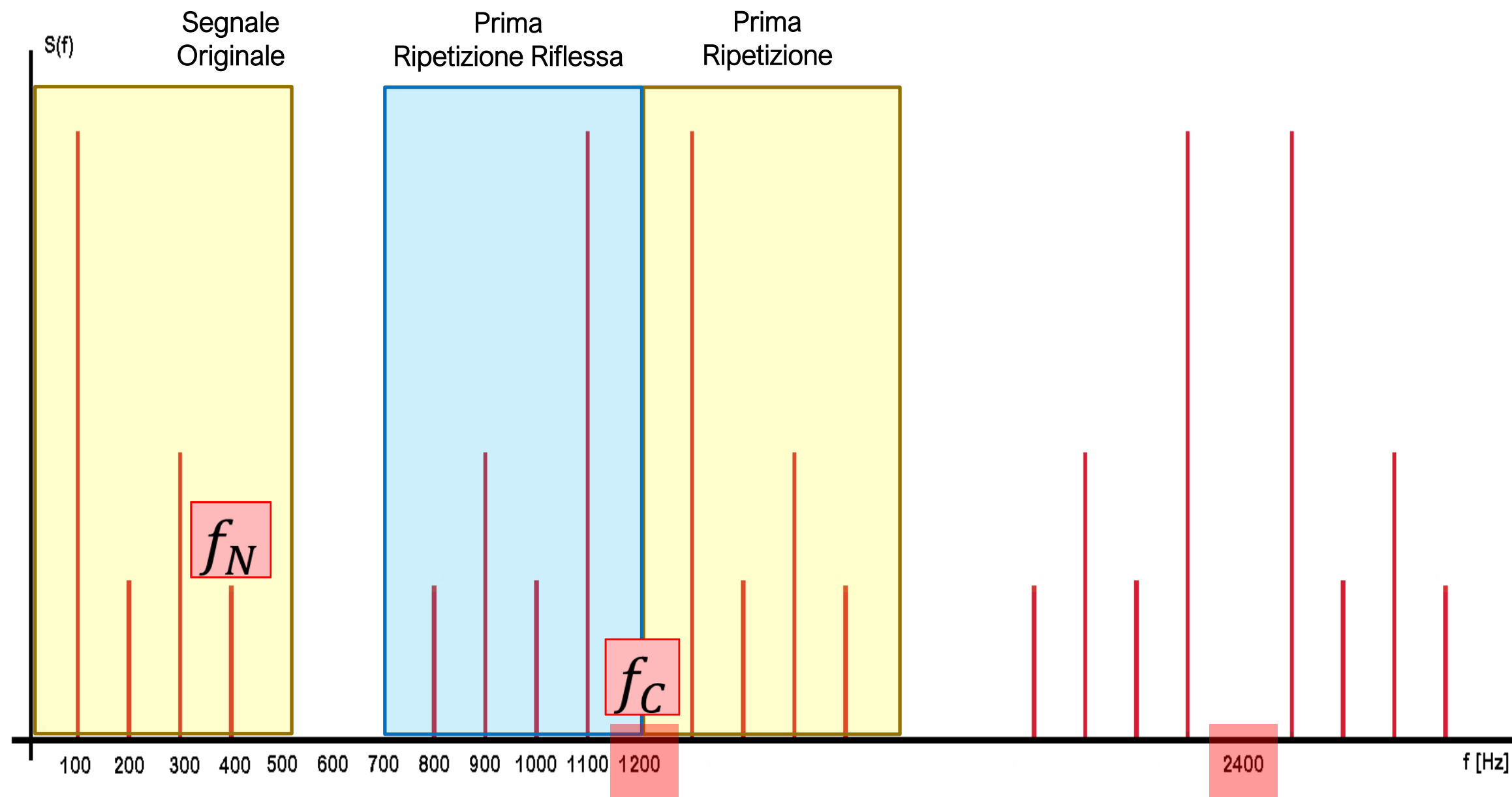
# Spettro del segnale campionato



Nello spettro del segnale campionato, lo spettro originale si ripete per multipli della frequenza di campionamento e viene anche riflesso. Nell'esempio si vede lo spettro del segnale campionato con frequenza di campionamento pari a 1200 Hz.



# Ricostruzione segnale originale



Per ricostruire il segnale originale bisogna eliminare tutte le frequenze indesiderate (non presenti in origine). Per questo si utilizza un **filtro passa basso** che elimini tutte le frequenze superiori a quella più alta del segnale originale.



# Sottocampionamento

Cosa accade se la frequenza di campionamento è troppo bassa?

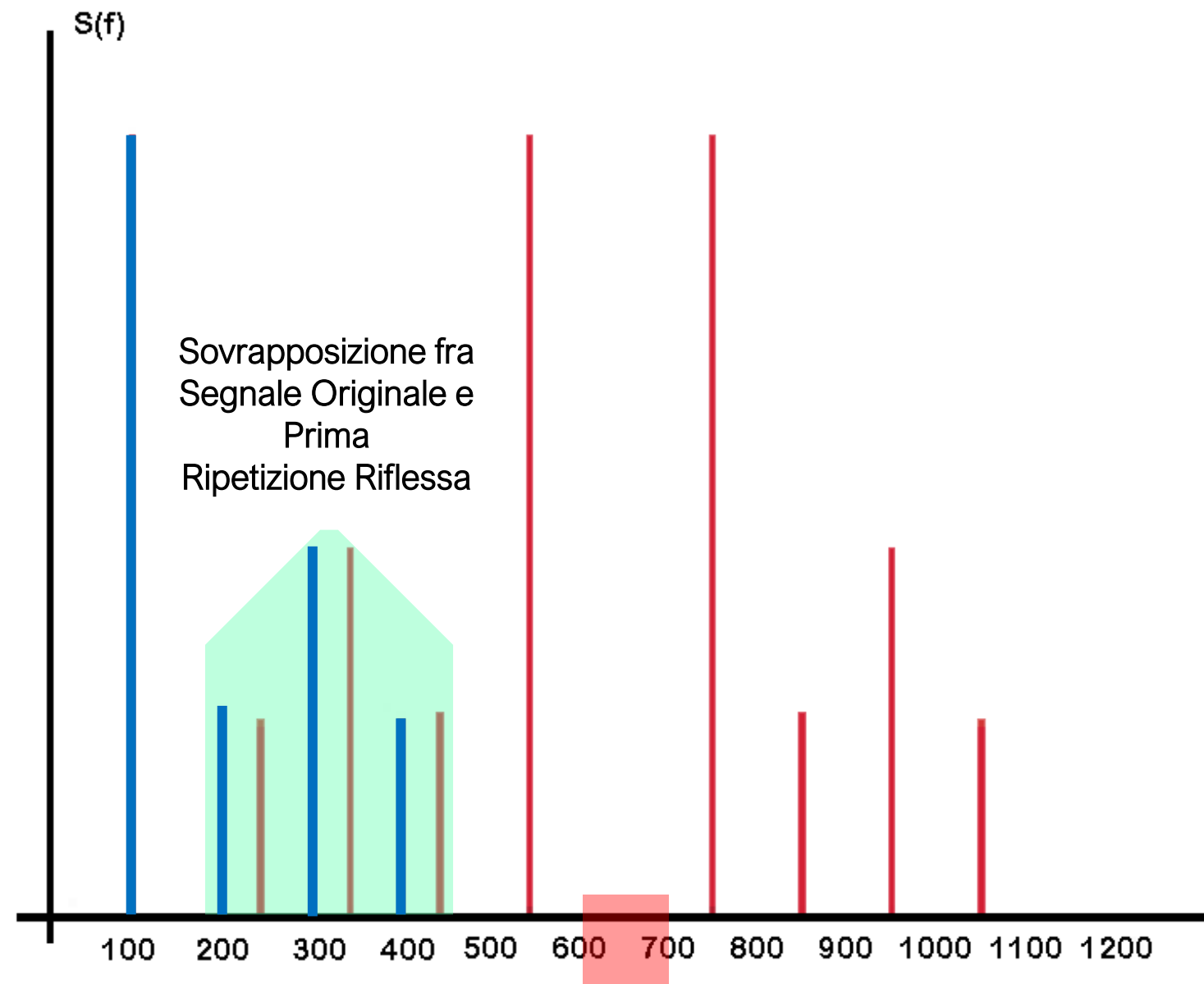
- Si perdono dettagli del segnale originale.
- Si introducono informazioni in origine non presenti, cioè una distorsione: **Aliasing**.

L'**aliasing** deriva da una sovrapposizione delle più alte frequenze dello spettro del segnale originale, con le frequenze introdotte nel segnale campionato





# Aliasing



Nell'esempio il segnale è stato campionato ad una frequenza di **650 Hz**, più piccola di quella suggerita dal teorema del campionamento. Alcune righe spettrali attorno alla frequenza di campionamento, sono andate a sovrapporsi alle frequenze originali.



# Audio digitale - Campionamento

- E' chiaro che una frequenza di campionamento alta è sempre migliore, in termini di qualità del segnale finale, rispetto ad una più bassa.
- Nonostante ciò, oltre una certa soglia, prendere troppi campioni porterebbe ad uno spreco di memoria senza effettivi miglioramenti.

Quale potrebbe essere una buona frequenza di campionamento per un suono?



# Audio digitale - Campionamento

- IDEA: gli esseri umani percepiscono frequenze fino a **20000 Hz**. Elimino tutte le frequenze superiori, così da rendere il suono a **banda limitata**.
- Una possibilità è campionare ad una frequenza di campionamento  $f_c$  pari a **44100 Hz**. Un *secondo* di audio digitale sarà rappresentato quindi da 44100 campioni.
- Questa frequenza è tipica degli audio digitali. E' un po' più del doppio di 20000, per assicurarsi un banda di sicurezza. Ricordiamo infatti che i **filtri passabasso ideali non esistono!** E' bene quindi non campionare esattamente al doppio della frequenza di Nyquist.



# Quantizzazione - Definizione

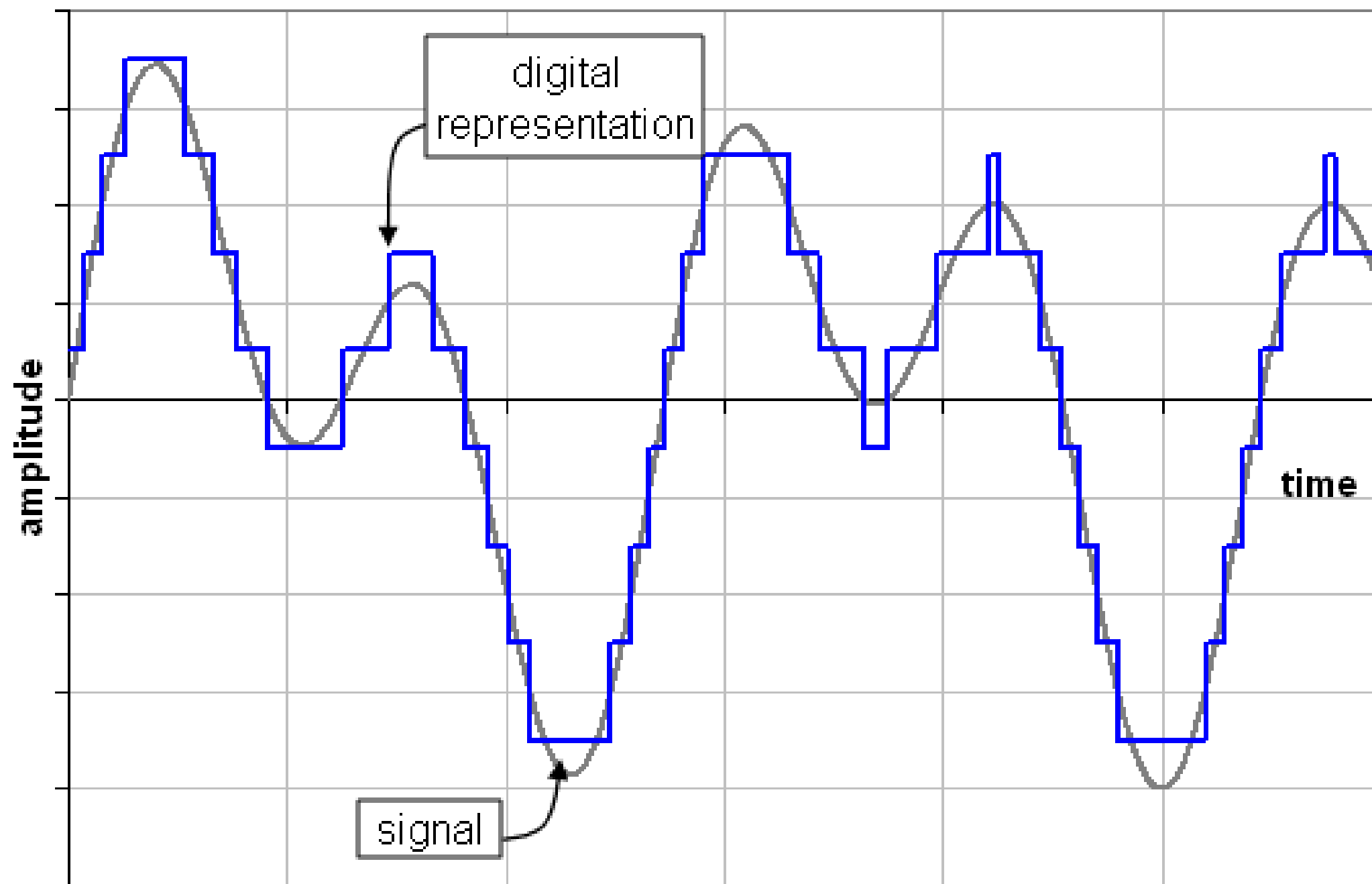
La trasformazione di un segnale a valori continui in un segnale a valori discreti, prende il nome di **quantizzazione**.

Ad ogni valore originale sarà associato un nuovo valore in un insieme discreto di livelli. Questo introdurrà un certo errore, poiché valori originariamente differenti possono collassare nello stesso livello, divenendo indistinguibili.

La **precisione** di quantizzazione è legata alla minima variazione nella grandezza originale che induce un passaggio da un livello ad un altro nel dominio quantizzato. Più è piccola la minima variazione necessaria, più sarà precisa la quantizzazione.



# Quantizzazione - Esempio





# Quantizzazione – Definizioni

**Uniforme o lineare:** la quantizzazione è uniforme quando ad intervalli di ampiezza uguale nel dominio originale, corrispondono un numero uguale di livelli di quantizzazione.

**Non Uniforme o non lineare:** la quantizzazione è non uniforme quando ad intervalli di ampiezza uguale nel dominio originale, corrispondono un numero diverso di livelli di quantizzazione. In altre parole si è più precisi nel quantizzare certi intervalli e meno precisi per altri.

**Valore di fondo scala:** è la dimensione del range che si vuole rappresentare. Ad esempio, se si vuole rappresentare una tensione tra 5 e 20 volt:

$$VFS = 20 - 5 = 15 \quad \text{ossia} \quad VFS = Max - Min$$





# Quantizzazione uniforme – Errore

La quantizzazione è un processo di approssimazione che come tale introduce una perdita e quindi una distorsione. Di norma, si associa ai valori che cadono nello stesso intervallo il valore di mezzo di tale intervallo. Per cui, nel caso di quantizzazione uniforme, l'errore massimo di quantizzazione  $E_{max}$  vale:

$$E_{max} = \frac{VFS}{2Q} \text{ dove } Q \text{ è uguale al numero di livelli di quantizzazione.}$$

Nel caso dell' audio digitale, il numero massimo di livelli (uniformi o meno) dipenderà dal **numero di bit** che decideremo di utilizzare per rappresentare i livelli di ampiezza (**profondità in bit**). Dunque se  $N$  è il numero di bit utilizzati:

$$Q = 2^N$$



# Quantizzazione uniforme – Distorsione SQNR

Come nel caso dell'audio analogico è possibile stimare la distorsione introdotta dal trattamento del segnale originale. Un indice che tiene conto della distorsione introdotta dalla **quantizzazione** è il **Signal to Quantization Noise Ratio**. L'**SQNR medio** nel caso di **quantizzazione uniforme** può essere calcolato come segue:

$$SQNR = 2^N \times \sqrt{\frac{3}{2}}$$

Dove  $N$  è la **profondità in bit**.

In decibel:

$$SQNR_{dB} = 10 \log_{10} SQNR^2 = 6,02N + 1,7609$$

E' evidente che utilizzando un numero superiore di bit si quantizzi in maniera più precisa e si introduca una più bassa distorsione. Infatti un SQNR alto è indice di maggiore qualità.



# Quantizzazione uniforme

## Distorsione SQNR (dal testo)

In decibel:

$$SQNR_{dB} = 10 \log_{10} SQNR^2 = 6,02N + 1,7609$$

Questa relazione stabilisce che ogni bit contribuisce per circa 6dB al SQNR.

Es.:

Dato  $N=8$ , l'SQNR sarà circa  $6 \cdot 8 = 48\text{dB}$

Dato un SQNR pari a circa 120dB, allora  $N = 120/6 = 20$



# Quantizzazione uniforme – Distorsione SQNR

Il valore SQNR così calcolato è inteso per l'errore di quantizzazione medio!  
Cosa vuol dire?

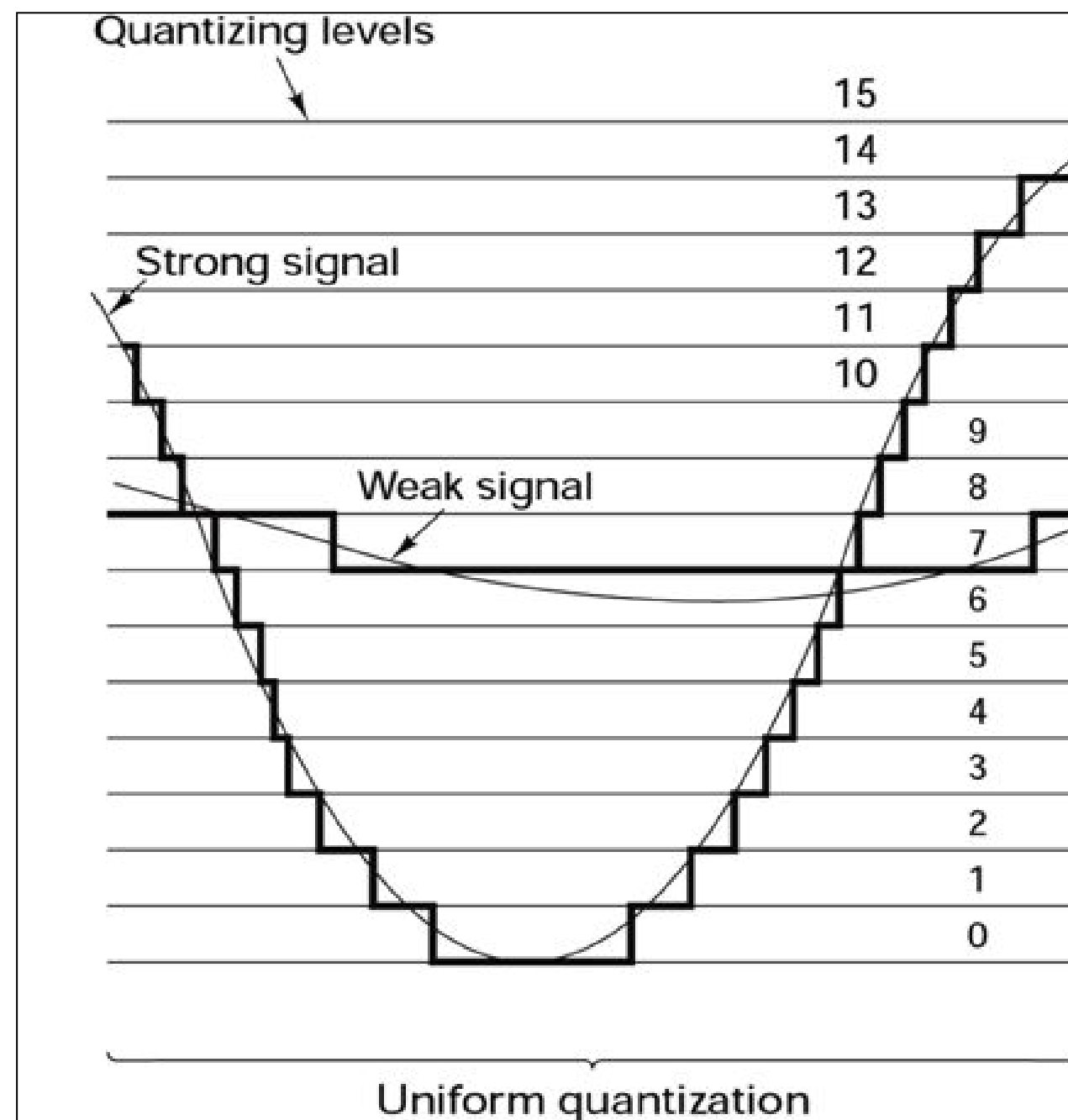
- A parità di errore di quantizzazione, le ampiezze più piccole in valore assoluto sono distorte maggiormente (in proporzione), rispetto alle ampiezze più grandi.
- Per essere precisi bisognerebbe calcolare la distorsione per ogni possibile valore di ampiezza.
- L'indice SQNR è calcolato considerando l'ampiezza media quadratica di un segnale sinusoidale e l'errore medio quadratico di un segnale a dente si sega (che spazia linearmente nel range delle proprie ampiezze).
- $SQNR > 60\text{dB}$  sono valori accettabili
- Gli standard CD a 16 bit hanno  $SQNR = 96\text{dB}$



# Quantizzazione uniforme – Distorsione SQNR

*Si può notare che l'andamento dell'audio indicato nel grafico come «Strong signal» risulta ancora riconoscibile dopo la quantizzazione.*

*Per l'audio «Weak signal», la lieve variazione precedente alla quantizzazione viene quasi totalmente persa.*





# Quantizzazione non uniforme

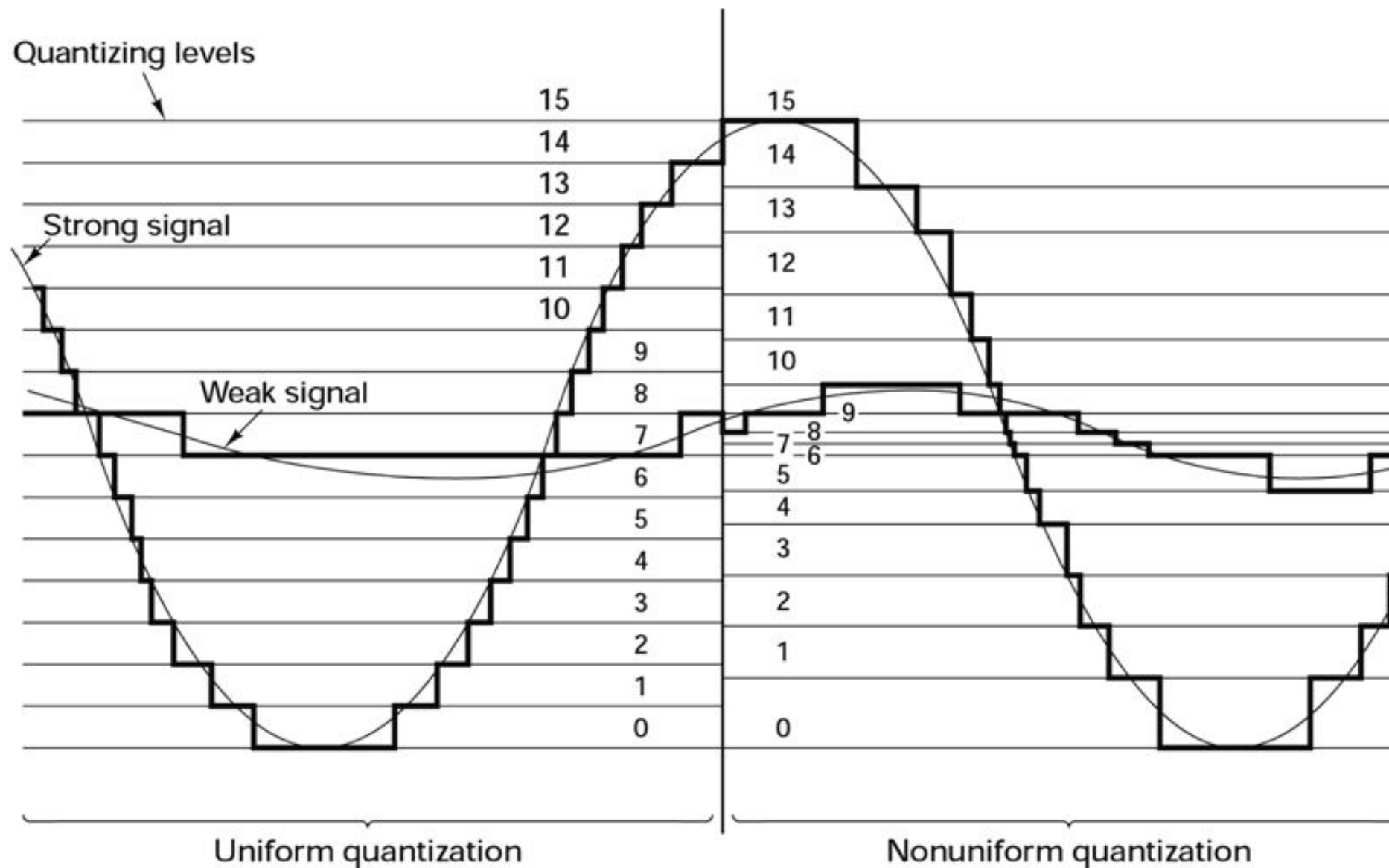
- Le considerazioni fatte per l'errore di quantizzazione, che pesa più sulle ampiezze piccole ci suggeriscono un'idea.
- Se a queste uniamo il fatto che gli esseri umani percepiscono meglio variazioni di volume che interessano le ampiezze basse...
- ... capiamo che la quantizzazione non uniforme potrebbe aiutarci a migliorare la qualità complessiva a parità di bit utilizzati.





# Quantizzazione non uniforme

## *Uniform and Non-uniform Quantization*





# Audio digitale

A questo punto abbiamo tutte le grandezze necessarie a caratterizzare un segnale audio digitale:

- ❑ **Numero di canali;**
- ❑ **Frequenza di campionamento** (numero di campioni al secondo);
- ❑ **Profondità in bit per campione** ( bit di quantizzazione per valore ).

Quanti bit sono necessari in memoria per rappresentare un segnale audio digitale?



# Audio digitale – Spazio in memoria

Sia  $f_c$  il sampling rate,  $N$  la profondità in bit,  $D$  la durata del flusso audio e  $C$  il numero di canali, allora il numero di bit necessari a rappresentare il segnale ( senza compressione ) si calcola:

$$Size = f_c \times N \times D \times C$$

Il numero di bit che fluisce nell'unità di tempo ( un secondo ) prende il nome di **bit rate**. Si misura in **bps** ( bit per secondo).

$$bitrate = f_c \times N \times C$$

E' chiaro che l'obiettivo è garantire una buona qualità utilizzando la minima quantità di memoria. I metodi di compressione rappresentano un passo successivo che permette di abbassare il **bit rate** preservando la qualità.