

SOUND DESIGN

Alessandro Fiordelmondo
alessandro.fiordelmondo@labatrentino.it

2T

Introduzione al suono digitale

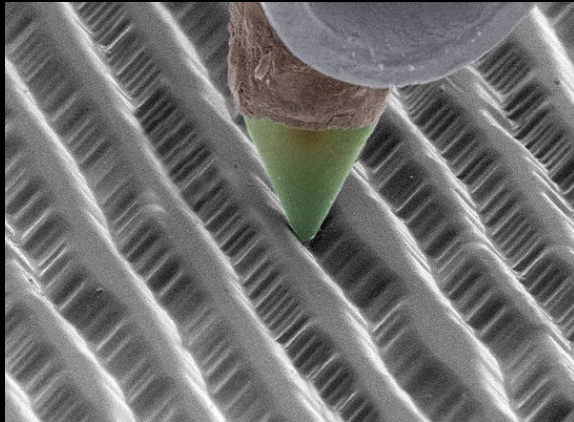
LABA Libera Accademia Belle Arti - Trentino
2021/2022

SUONO ANALOGICO

La rappresentazione analogica del suono lavora per **analogia**.

La curva continua nel tempo delle variazioni di ampiezza viene rappresentata da una curva nel dominio del tempo dalle variazioni di tensione elettrica

La curva della variazione di tensione elettrica può essere memorizzata sui solchi di un vinile o nel campo magnetico di un nastro



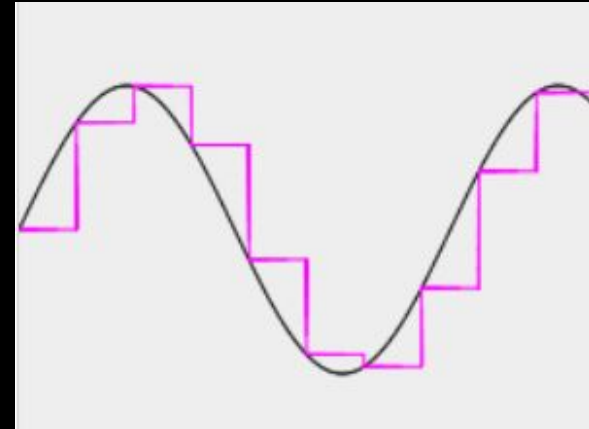
ingrandimento 1:1000
della puntina sui
solchi del vinile

SUONO DIGITALE

La rappresentazione digitale del suono è una **discretizzazione** del segnale.

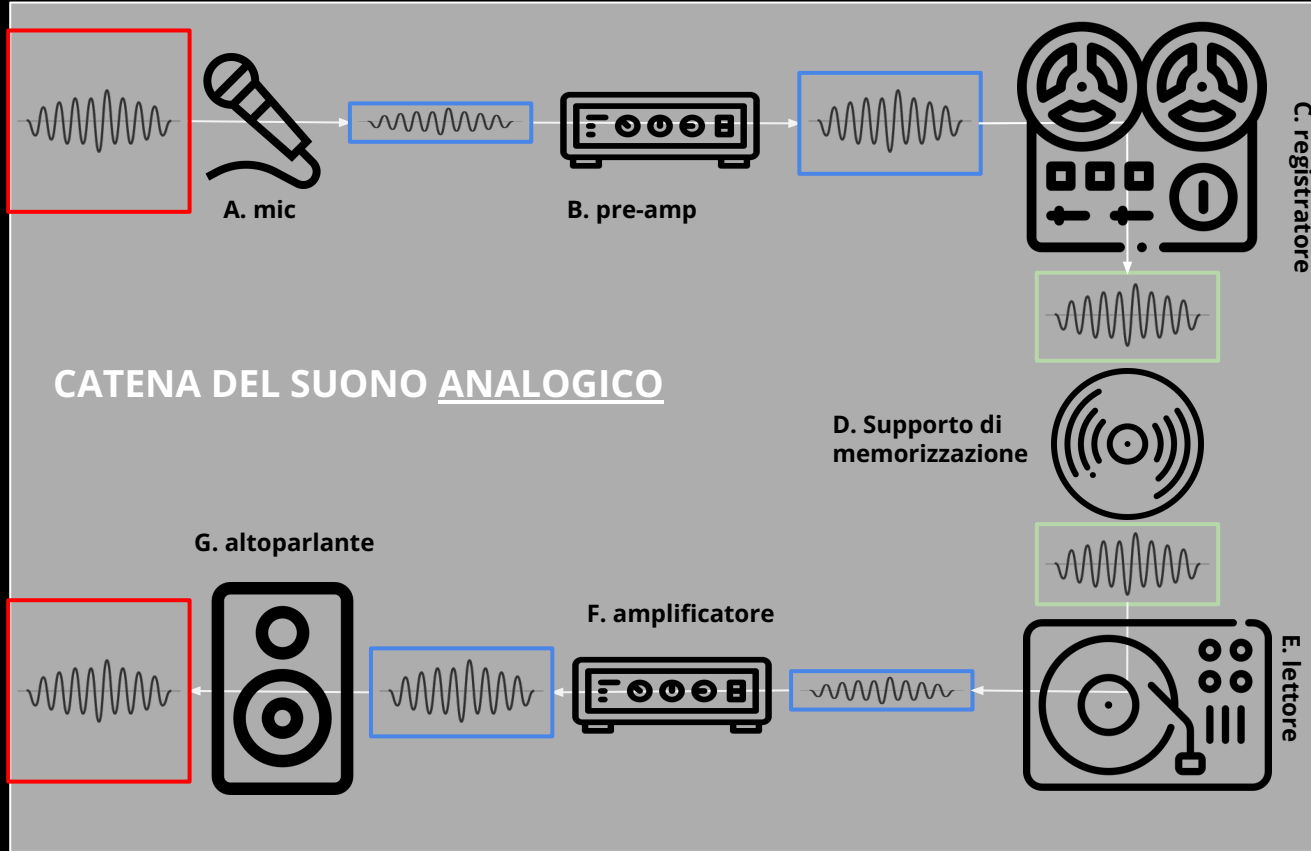
Con la digitalizzazione non abbiamo un'imitazione "continua" del segnale ma una sua rappresentazione attraverso dei numeri che approssimano l'ampiezza del suono a precisi istanti di tempo.

Suono digitale



Tensione elettrica

Variazione di
pressione
dell'aria



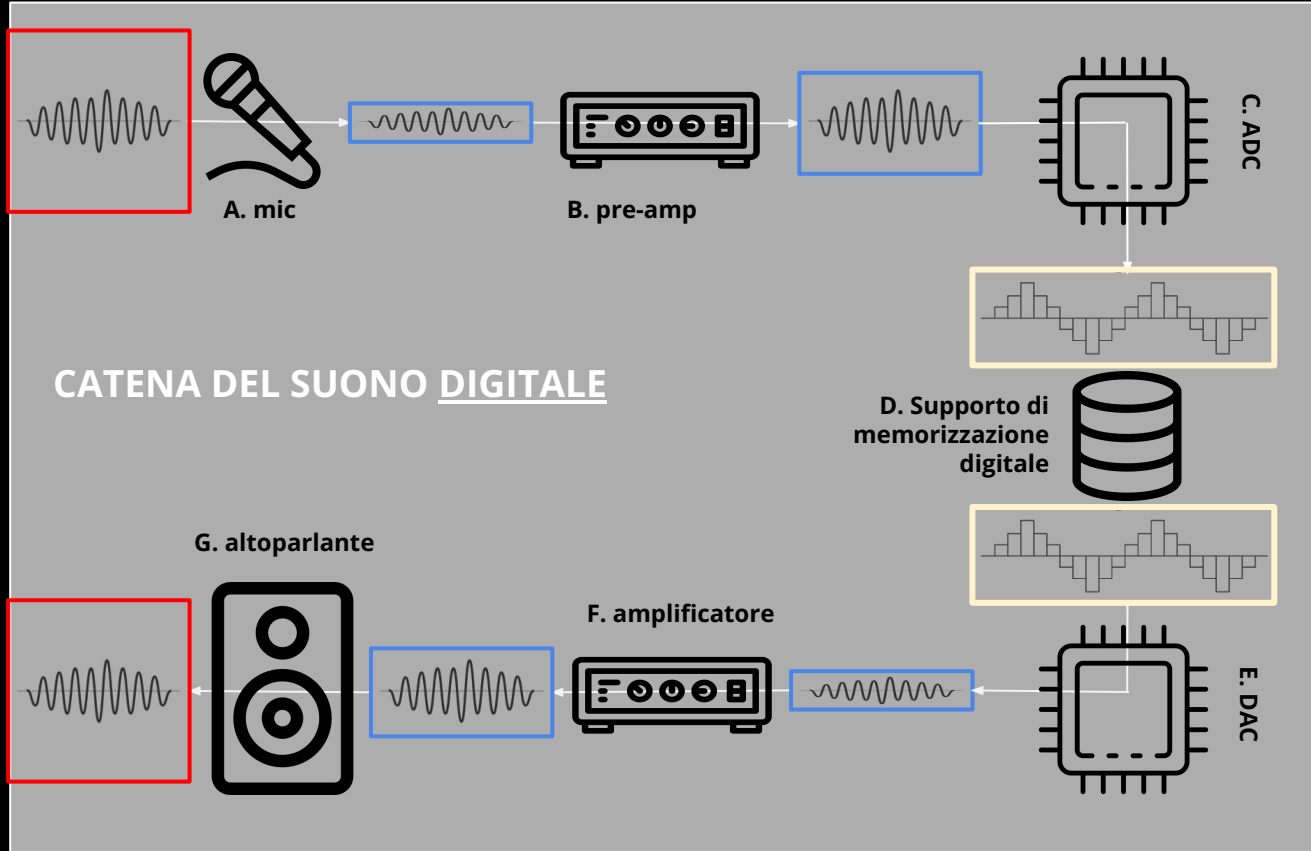
Solchi / campo magnetico
(memorizzazione tensione elettrica)

Tensione elettrica

Tensione elettrica

ADC - Analogue Digital Converter

Variazione di
pressione
dell'aria



Rappresentazione digitale del suono
Discretizzazione del segnale

Tensione elettrica

DAC - Digital Analogue Converter

Variazione di
pressione
dell'aria

QUALITÀ DEI SISTEMI

Gli elementi della catena possono introdurre *rumore* e *distorsione*.

Il *rumore* è un elemento fisso per l'intero segnale che presenta uno spettro molto ampio e che viene introdotto dai sistemi analogici.

Passaggi che introducono il rumore nella catena analogica: $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G$

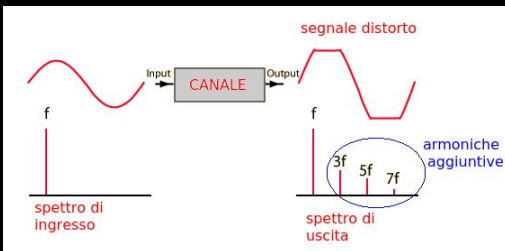
Passaggi che introducono il rumore nella catena digitale: $A \rightarrow B / F \rightarrow G$

Per stimare l'ammontare di rumore introdotto dal sistema analogico si utilizza il *rapporto segnale-rumore* (*Signal-to-Noise-Ratio* - SNR)

$$\text{SNR} = \text{max ampiezza segnale} / \text{ampiezza rumore}$$

Maggiore SNR migliore è la qualità del sistema.

Collegato alla SNR, un altro parametro che misura la qualità di un sistema è la gamma dinamica che è il rapporto tra ampiezza minima e ampiezza massima del segnale



Per calcolare sia la SNR e la gamma dinamica bisogna considerare la *massima ampiezza utile*. Per i segnali con ampiezza maggiori, i dispositivi analogici possono introdurre nel segnale la *distorsione*, che consiste nella modifica non desiderata della forma d'onda del segnale (e quindi dello spettro). Per *massima ampiezza utile* si intende la massima ampiezza per la quale non avviene l'effetto della distorsione

Analogico vs Digital

La digitalizzazione del segnale comporta una discretizzazione e cioè un'approssimazione del segnale con perdita di informazione => perdita di qualità del segnale.

I vantaggi del segnale digitale sono:

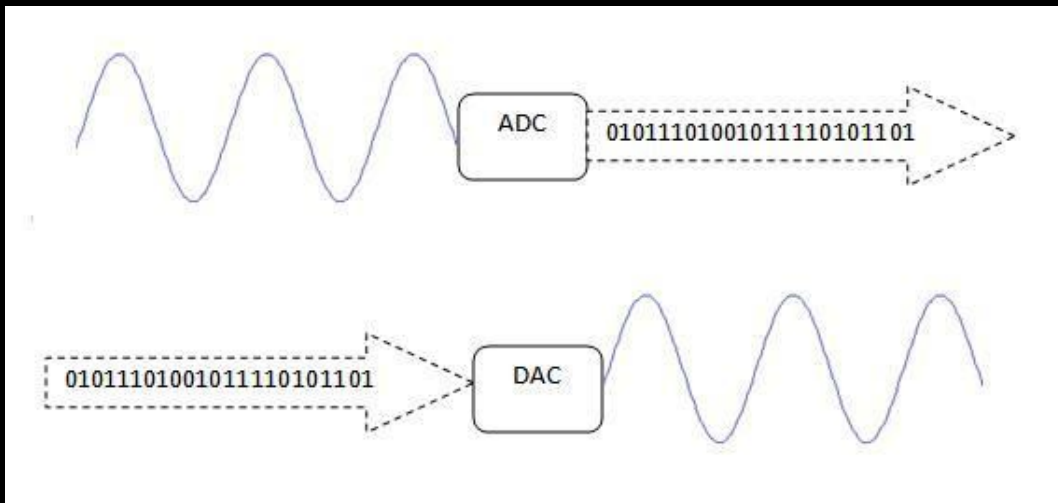
- 1) Maggiore SNR - miglior comportamento nei confronti del rumore
- 2) Facilità di elaborazione del segnale
- 3) Facilità di memorizzazione
- 4) Integrazione dei sistemi di trasmissione
- 5) Convenienza economica

II SUONO DIGITALE - ADC e DAC

I componenti fondamentali del sistema digitale sono:

- l'**ADC** (*Analogue Digital Converter*) con il quale vengono discretizzati nel tempo i valori di ampiezza del segnale (da segnale continuo a segnale discreto in ampiezza e nel tempo)
- l'**DAC** (*Analogue Digital Converter*) con il quale vengono convertite il segnale digitale (valori discreti di ampiezza in precisi istanti di tempo) in variazione di tensione elettrica (segnale continuo)

I due componenti sono integrati nella **scheda audio**.



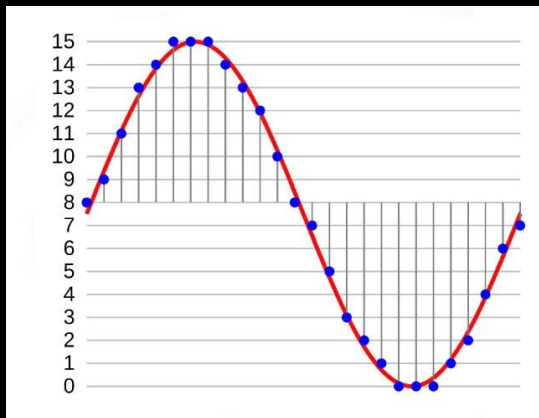
Il **campionamento** e la **quantizzazione** sono le operazioni che consentono di digitalizzare un segnale e dipendono dai componenti di conversione ADC e DAC.

II SUONO DIGITALE - CAMPIONAMENTO

Il **campionamento** è la discretizzazione del segnale analogico nel tempo.

Dato che il segnale è rappresentato dalla variazione continua delle ampiezze (variazione di pressione atmosferica) nel tempo, il campionamento individua gli istanti di tempo in cui vengono prelevate e convertite le ampiezze del segnale analogico in numeri che rappresentano le ampiezze. La conversione delle ampiezze prelevate in numeri è data dall'operazione di *quantizzazione* che si vedrà successivamente..

Il campionamento lavora sulle ascisse (l'asse delle x) della rappresentazione nel dominio del tempo.



Ogni ampiezza prelevata può essere chiamata **campione**.

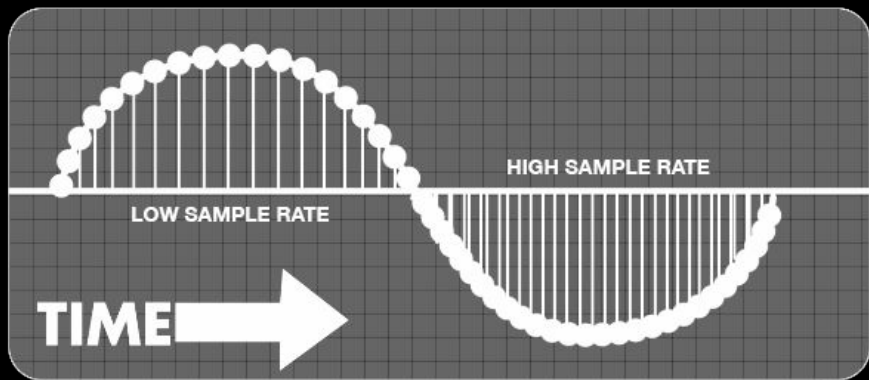
II SUONO DIGITALE - CAMPIONAMENTO - *Frequenza di campionamento (Sample rate)*

L'individuazione delle ampiezze da prelevare viene fatta in modo periodico → la distanza tra le ampiezze prelevate è sempre uguale.

Il campionamento e quindi l'estrazione dei campioni deve avvenire ad una velocità sufficientemente elevata in modo da ricostruire tutte le variazioni del segnale.

La velocità di prelevamento dei campioni è definito come la **frequenza di campionamento** (*sample rate*) e si misura in *Hz*, ossia quanti campioni vengono prelevati ad ogni secondo (campioni/secondo). Il **periodo di campionamento** è la distanza in secondi tra un campione ad un altro ($1/\text{frequenza di campionamento}$). Ad ogni periodo di campionamento si preleva un campione di ampiezza.

Più è breve il periodo di campionamento, migliore è la rappresentazione del suono campionato e maggiore è il numero di campioni e la quantità di informazione da memorizzare.

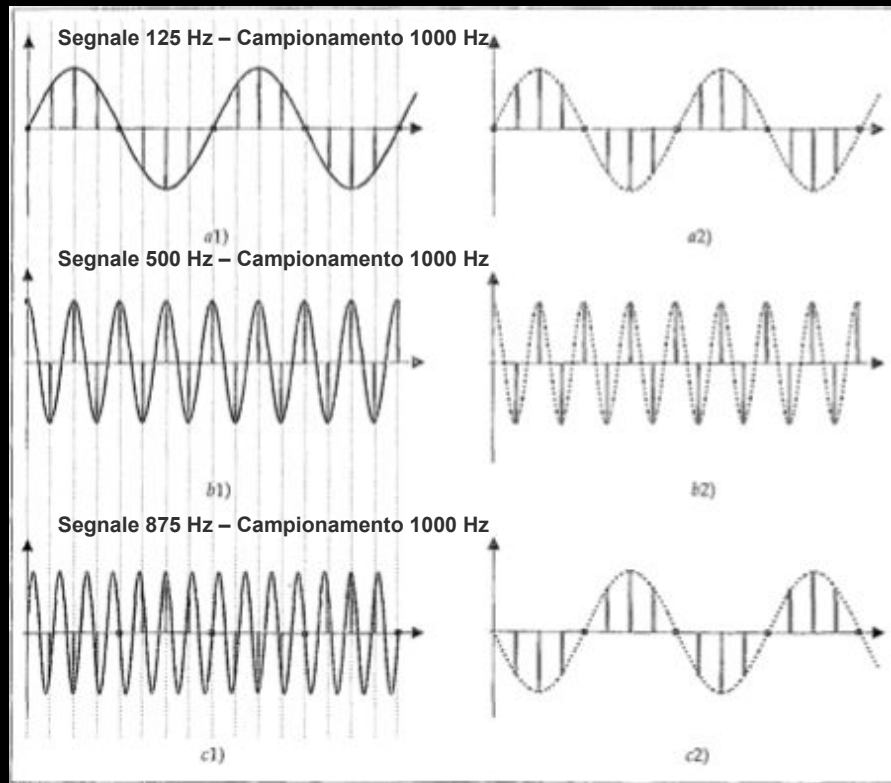


La frequenza di campionamento dipende dalla velocità di variazione del segnale che a sua volta dipende dalla parziale con frequenza più alta. Si considera pertanto come frequenza più alta da campionare il range frequenziale di udibilità dell'umano 20.000 Hz.

Pertanto è giusto campionare un segnale a 20.000 Hz ? Qui nasce il problema dell' **aliasing**...

II SUONO DIGITALE - CAMPIONAMENTO - *Aliasing*

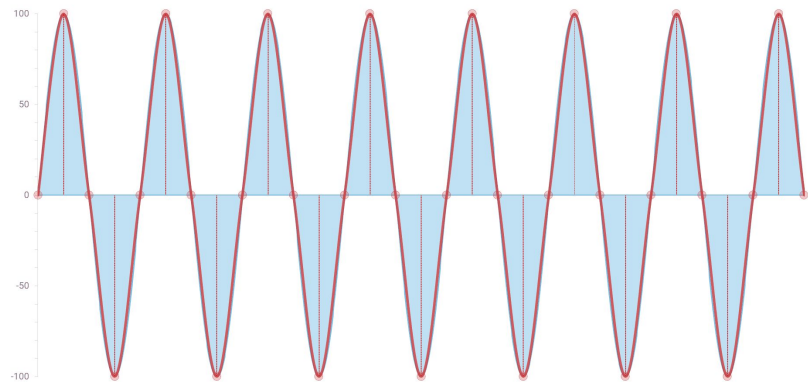
L'**aliasing** è l'artefatto che si produce come risultato del campionamento che, in determinate condizioni, può ricostruire un segnale differente dal segnale originale continuo nel tempo. Il risultato che si può ascoltare dell'aliasing è la produzione di frequenze non presenti sul segnale originale.



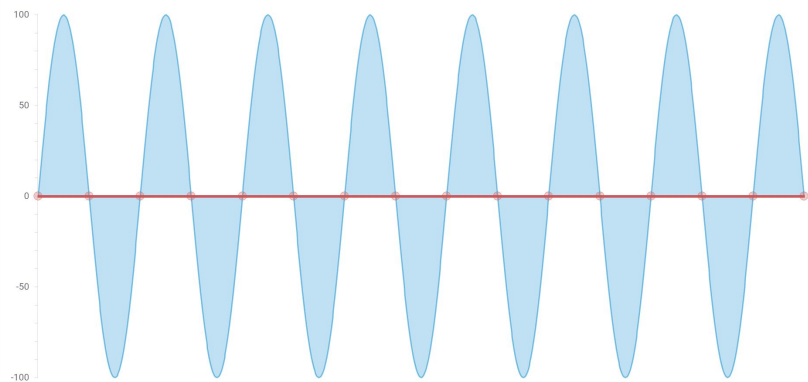
- a) *Oversampling* → la frequenza di campionamento è più che sufficiente per campionare il segnale (sovracampionamento). Condizione ottimale.
- b) *Campionamento critico* → La frequenza di campionamento è al limite per campionare il segnale. Se la fase del segnale originale fosse stata 0° invece di 90° il campionamento avrebbe ottenuto tutte ampiezze uguali a 0 e produrrebbe anche in questo caso l'effetto dell'aliasing.
- c) *Undersampling* → La frequenza di campionamento non è sufficiente per ricostruire il segnale (sottocampionamento). Il risultato prodotto è l'**aliasing**, ossia la creazione di un segnale completamente diverso dall'originale

II SUONO DIGITALE - CAMPIONAMENTO - *Aliasing*

60-sec period signal compared to 30-second Sampled Result

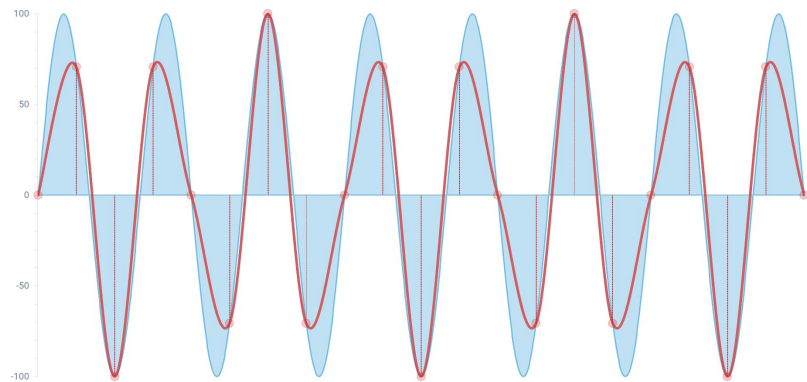


60-sec period signal compared to 60-second Sampled Result

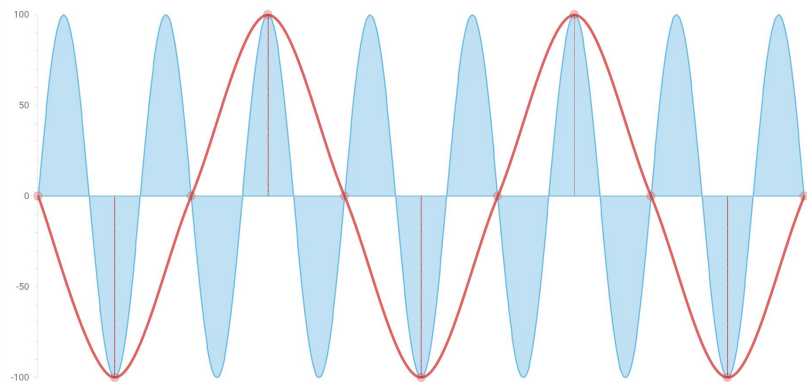


Source Signal Discrete Sample Sampled Result

60-sec period signal compared to 45-second Sampled Result

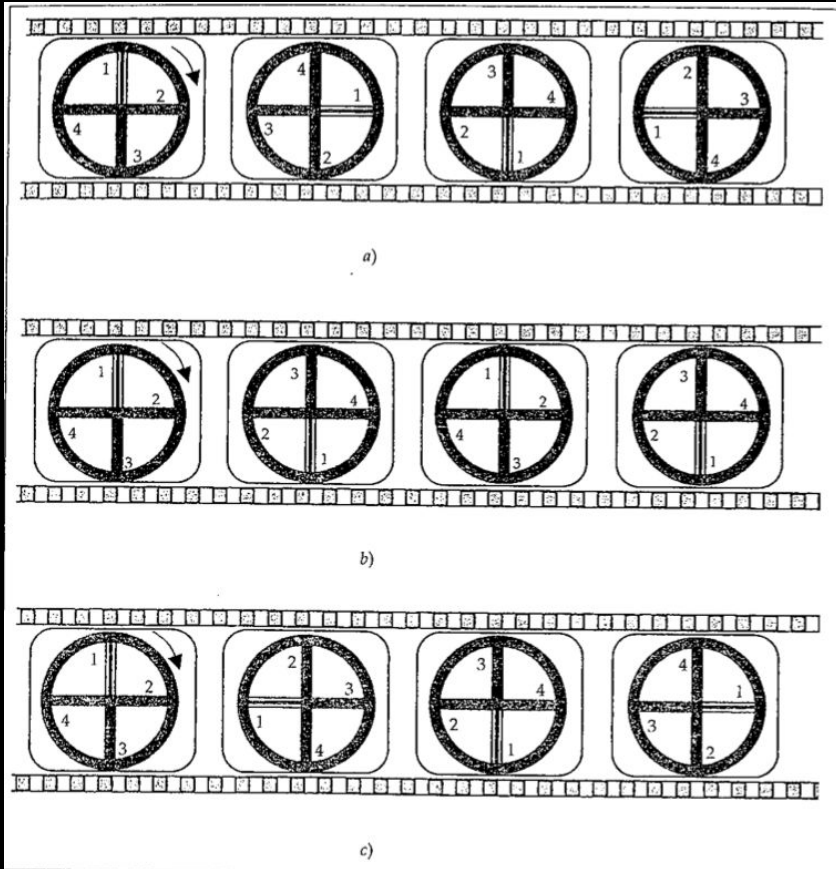


60-sec period signal compared to 90-second Sampled Result



Source Signal Discrete Sample Sampled Result

II SUONO DIGITALE - CAMPIONAMENTO - *Aliasing* in video



II SUONO DIGITALE - CAMPIONAMENTO - *Teorema di Nyquist-Shannon*

Il problema dell'aliasing può essere risolto sovracampionando il segnale analogico. Tuttavia per scegliere correttamente la frequenza di campionamento si può utilizzare il teorema di Nyquist-Shannon (spesso anche chiamato teorema del campionamento).

Il **teorema di Nyquist-Shannon** afferma che, per campionare correttamente (senza perdita di informazioni) un segnale a banda limitata, è sufficiente campionarlo con una frequenza di campionamento pari almeno al doppio della massima frequenza del segnale (tale frequenza viene anche detta **frequenza di Nyquist**).

$$f_{\max} < f_c / 2$$

Abbiamo bisogno di almeno due campioni per periodo poiché in un periodo il segnale cambia direzione due volte.

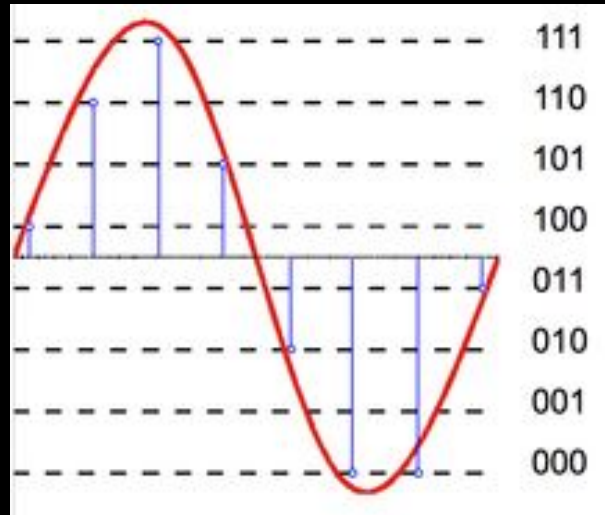
La frequenza di campionamento più tipica per l'audio è di 44100 Hz, e quindi con una frequenza di Nyquist (massima frequenza che si può campionare) a 22050 Hz. È la frequenza di campionamento che troviamo per i CD audio.

II SUONO DIGITALE - QUANTIZZAZIONE

Se il campionamento è la discretizzazione del segnale analogico nel tempo, la **quantizzazione** è la discretizzazione delle ampiezze del segnale.

Ciascun campione prelevato per ogni istante tempo (definito dalla frequenza di campionamento) rappresenta un valore discreto dell'ampiezza del segnale. Tale valore dipende dalla quantizzazione, e viene rappresentato da un numero finito di bits.

La quantizzazione lavora sulle ordinate (asse delle y) della rappresentazione nel dominio del tempo



II SUONO DIGITALE - QUANTIZZAZIONE - rappresentazione in bit

Il valore dei campione viene rappresentato in bits, quindi attraverso un sistema binario. Ogni campione è definito da una serie di 0 ed 1.

Il sistema binario funziona esattamente come il sistema decimale. Nel sistema decimale il valore di ogni cifra che compone il numero è calcolato a partire dalla posizione decimale che occupa (10^n dove n è la posizione decimale – la prima posizione a destra è 0):

$$8 \times 10^0 = 8 \times 1 = 8$$

$$(1 \times 10^1) + (8 \times 10^0) = 10 + 8 = 18$$

Pertanto il range di valori che possiamo ottenere aumenta con l'aumentare delle cifre utilizzate \Rightarrow valori = 10^N (dove N è il numero di cifre con cui rappresentiamo il valore). Vale la stessa regola per il sistema binario solo che il valore di ogni cifra che compone il numero sarà di 2^n e il numero massimo di valori che possiamo rappresentare sarà 2^N .

$$(1 \times 2^3) + (0 \times 2^2) + (0 \times 2^1) + (0 \times 2^0) = 8 + 0 + 0 + 0 = 8$$

$$(1 \times 2^4) + (0 \times 2^3) + (0 \times 2^2) + (1 \times 2^1) + (0 \times 2^0) = 16 + 0 + 0 + 2 + 0 = 18$$

Il valore 8 sarà rappresentato da 1000 ed il valore 18 da 10010.

II SUONO DIGITALE - QUANTIZZAZIONE - rappresentazione in bit

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
0	0	0	0	1	0	1	0	= 10

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
0	0	0	1	0	1	1	1	= 23

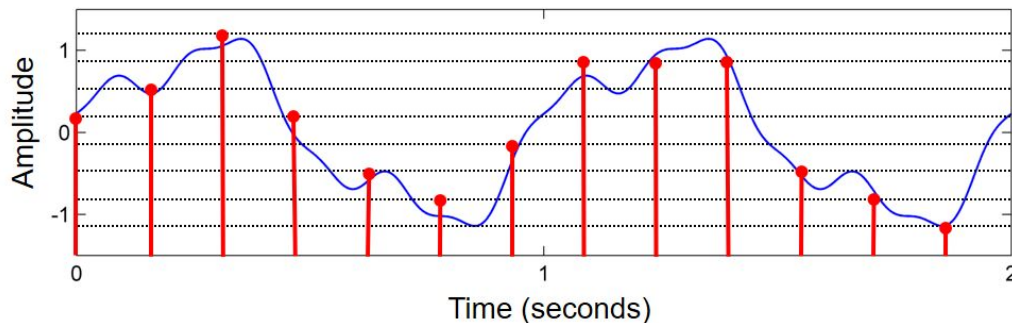
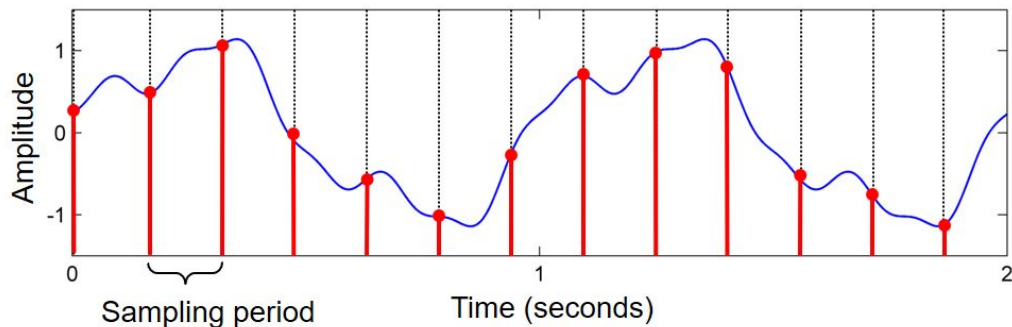
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
1	0	0	0	0	0	0	0	= 128

Numero più grande ad 8 bit

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
1	1	1	1	1	1	1	1	= 255

II SUONO DIGITALE - QUANTIZZAZIONE - errore di quantizzazione

Figure 2.13 from [Müller, FMP, Springer 2015]



Più grande è il numero di bit maggiore è l'accuratezza del segnale quantizzato e minore è l'*errore di quantizzazione* (anche chiamato rumore di quantizzazione).

Nell'approssimazione delle ampiezze a valori finiti viene per forza introdotto un errore.

Esempio: Approssimazione di numeri con virgola a numeri interi per mezzo della funzione *round* (funzione che arrotonda i numeri con virgola all'intero più vicino) otteniamo una perdita di informazione.

(valore originale)

0,1 - 0,2 - 0,3 - 0,4 - 0,5 - 0,6 - 0,7 - 0,8 - 0,9 - 1

(valore approssimato con round)

0 - 0 - 0 - 0 - 0 - 1 - 1 - 1 - 1 - 1

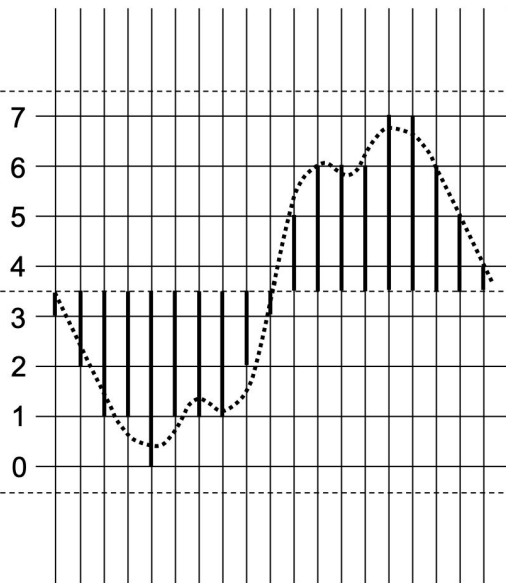
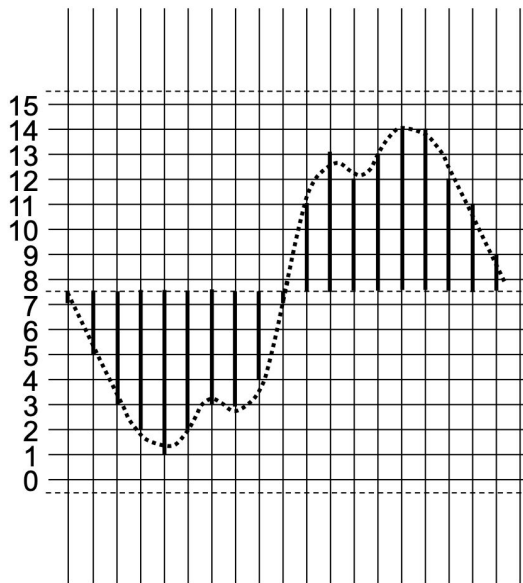
La stessa cosa può essere assimilata all'errore di quantizzazione. La perdita della continuità del segnale introduce una distorsione.

Diminuendo i bit si aumenta lo *step di quantizzazione* e di conseguenza l'errore e rumore.

II SUONO DIGITALE - QUANTIZZAZIONE - errore di quantizzazione

4 bit / 16 valori di quantizzazione

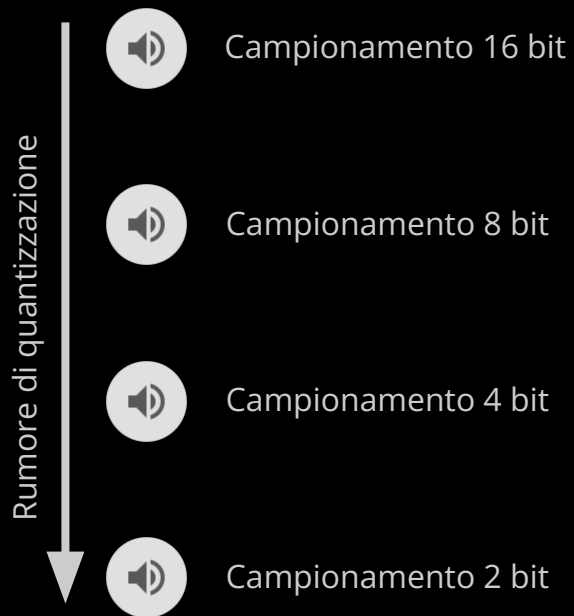
3 bit / 8 valori di quantizzazione



Valore di tensione	Parola binaria
$[-5.000, -4.961]$	00000000
$[-4.961, -4.922]$	00000001
$[-4.922, -4.883]$	00000010
$[-4.883, -4.844]$	00000011
...	00000100
...	...
...	...
...	11111011
$[+4,844, +4,883]$	11111100
$[+4,883, +4,922]$	11111101
$[+4,922, +4,961]$	11111110
$[+4,961, +5,000]$	11111111

Il segnale analogico varia da 5V a -5V

II SUONO DIGITALE - QUANTIZZAZIONE - errore di quantizzazione



II SUONO DIGITALE - FORMATI STANDARD - CD AUDIO

44100Hz/16bit. È il formato del CD Audio, dà una banda passante di circa **20KHz** con circa **96dB** di dinamica. Forse non tutti sanno perché è stato scelto proprio **44100**. Le prime apparecchiature di registrazione digitale erano basate su registratori video; per ovvi motivi di semplicità di progetto era bene avere un numero intero di campioni per ogni linea di quadro.

Il formato **NTSC** usa **30 frame** di **525** linee al secondo (di cui **490** utilizzabili); con **3** campioni per linea si ottiene una **Fc** di **$30 \times 490 \times 3 = 44100$** .

D'altra parte il formato **PAL/SECAM** usa **25 frame** di **625** linee al secondo (di cui **588** utilizzabili); con **3** campioni per linea si ottiene una **Fc** di **$25 \times 588 \times 3 = 44100$** , semplice no?

Purtroppo **$44100/2 = 22050$** e la frequenza di Nyquist è troppo vicina al limite della banda audio per premettere un'agevole conversione **D/A** e **A/D**, con tutti i problemi che ne derivano.

II SUONO DIGITALE - FORMATI STANDARD - DAT

48000Hz/16bit. Quando è stato progettato un sistema di registrazione dedicato all'audio (il **DAT**) la frequenza di campionamento massima è stata elevata a **48KHz**. In questo caso la frequenza di Nyquist sale a **24KHz** rendendo molto meno problematici i filtraggi pre- e post- digitalizzazione. Storicamente fu una sorpresa per i recensori scoprire la drammatica differenza di qualità tra i due formati.

Con il miglioramento della tecnologia elettronica sono divenuti disponibili convertitori **A/D** e **D/A** con **18**, **20** e **24** bit. Questo ha fatto ha creato un certo numero di formati intermedi, utilizzati specialmente nelle apparecchiature da studio che sfruttavano la maggiore risoluzione disponibile per ampliare la gamma dinamica almeno prima del riversamento su **CD**.

Contemporaneamente è sorto il problema di spremere il meglio da quanto disponibile in studio al momento della creazione del master da stampare.

Dopo molti anni è stato infine possibile definire il formato fisico del successore del **CD**. Il **DVD** (*Digital Versatile Disc*) è un contenitore grezzo che nelle sue varie versioni può contenere da **5** a **18 GigaByte** di dati. Il fatto che sia "versatile" non è stato proprio un bene visto che sono oggi disponibili troppi formati **DVD** alternativi e la maggior parte dei consumatori se ne sta alla finestra aspettando che si decidano.

II SUONO DIGITALE - FORMATI STANDARD - DVD

96000Hz/24bit tra le varie possibilità di audio su **DVD** questa è per la più diffusa, questo è anche il formato più ricco disponibile nelle apparecchiature commerciali a basso costo (perfino nelle moderne schede audio per PC). la frequenza di Nyquist è **48KHz** e questo risolve pressoché tutti i problemi legati al filtraggio. I soliti spiritosi hanno pensato di interpretare i vantaggi di questa banda passante allargata e sono cominciate a spuntare casse dotate di super-tweeter ad ultrasuoni (poco diffuse nel mercato occidentale). I **24** bit di profondità di quantizzazione forniscono una gamma dinamica teorica superiore al rumore termico delle normali apparecchiature. Alcuni recensori hanno criticato pesantemente questa scelta giudicandola una spreco. Bisogna notare che a fronte di un leggero incremento della occupazione in bit si riesce ad eliminare praticamente tutti i problemi legati alla quantizzazione: il fatto che **144dB** non siano raggiungibili significa solamente che i problemi veri saranno legati alla qualità delle elettroniche è in particolare dei convertitori, lasciando molto spazio al miglioramento.

RIFERIMENTI

- Lombardo, Vincenzo, and Andrea Valle. *Audio e multimedia*. Maggioli Editore, 2014. **CAPITOLO 2 – LA RAPPRESENTAZIONE DIGITALE DEL SUONO**
- Francesco Romani - *Formati e Conversioni* <http://pages.di.unipi.it/romani/DIDATTICA/IAD/IAD3.pdf>