

---

# *Informatica Personale*

*Contributi digitali per  
rappresentare immagini e suoni:*

*Parte II*

*Dante Picca*

*Anno Accademico 2009-2010*

---

## *Premessa*

Sino ad ora ci siamo focalizzati esclusivamente sulle immagini. Abbiamo quindi imparato come una qualsiasi immagine (ad esempio una fotografia) può essere trasformata – attraverso i processi di campionamento spaziale e di quantizzazione cromatica – in un insieme di bit, così da poter essere trattata da un computer.

Una cosa analoga ci aspettiamo dovrà avvenire nel caso si voglia trattare, attraverso un calcolatore elettronico, anche un contributo audio.

In questo caso, la prima domanda da porci è: qual'è lo stimolo fisico che vogliamo digitalizzare e da cosa è prodotto?

## *Premessa*

---

Nel caso della luce, la risposta è nota: lo stimolo fisico è prodotto da onde elettromagnetiche che incidono nei diversi punti della nostra retina (o della zona fotosensibile di una macchina fotografica).

I colori sono dati dalla diversa frequenza della radiazione elettromagnetica incidente.

Noto tutto questo, il processo di digitalizzazione delle immagini può essere svolto senza problemi nei modi visti precedentemente.

## *Che cosa è un suono?*

---

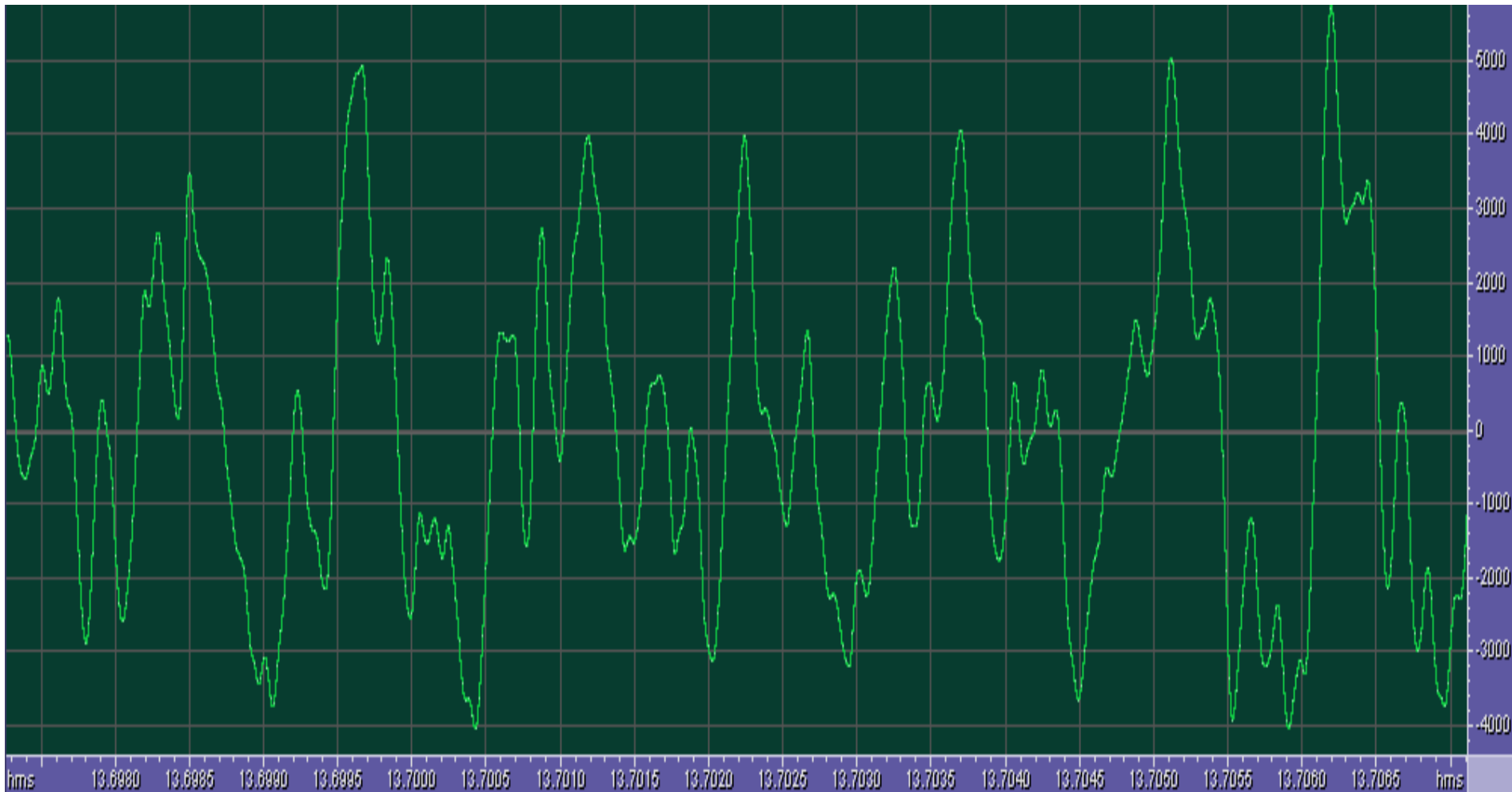
Il suono è quello stimolo prodotto sul nostro orecchio dalla vibrazione di un corpo in oscillazione che si propaga in un mezzo (ad esempio l'aria o l'acqua).

Anche il suono è, quindi, un'onda. Tale onda si propaga – attraverso l'aria – dalla sorgente al nostre orecchie.

L'aria fa vibrare il nostro timpano come un tamburo e tale vibrazione viene trasformata in un impulso nervoso per il cervello.

Ma come è fatta questa onda che arriva ai nostri timpani (oppure sulla membrana di un microfono)?

# *Che cosa è un suono?*



## *Digitalizzazione del segnale*

---

Che cosa significa digitalizzare questo segnale?

Significa trasformare la deformazione della membrana del microfono (o del nostro timpano) in un insieme di numeri, così da poterli trattare attraverso un computer.

Anche in questo caso, è necessario effettuare le attività di campionamento e di digitalizzazione, ma su un segnale completamente diverso rispetto a quello utilizzato per le immagini.

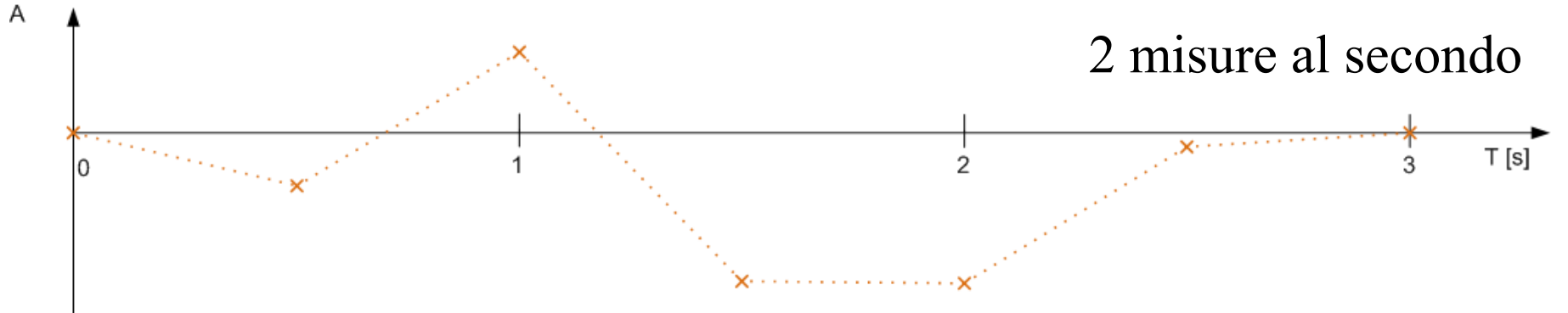
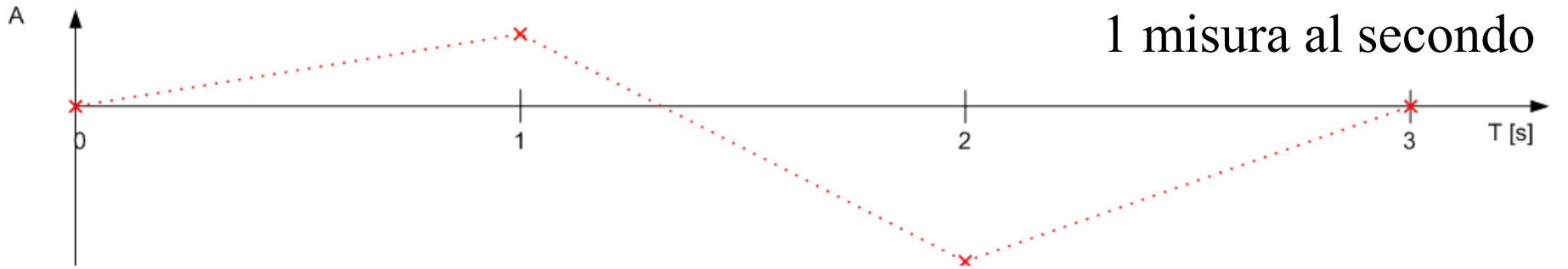
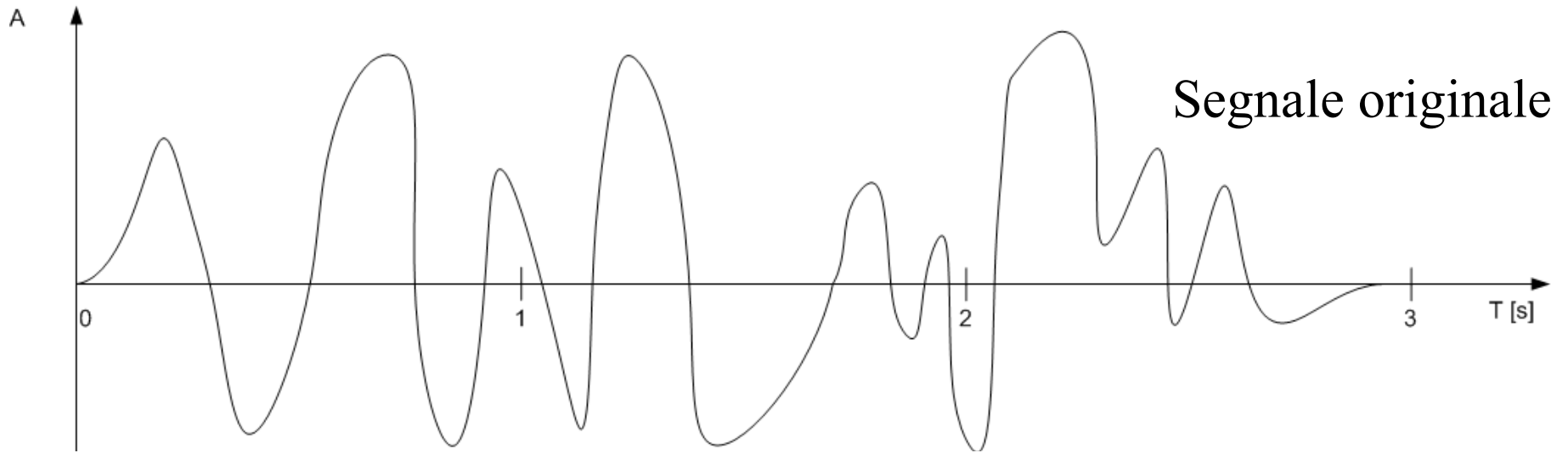
## *Campionamento di un'onda*

Con l'attività di **campionamento** si misura, ad intervalli di tempo regolari, l'ampiezza del segnale. Nel nostro caso, si misura la deformazione del timpano o del microfono.

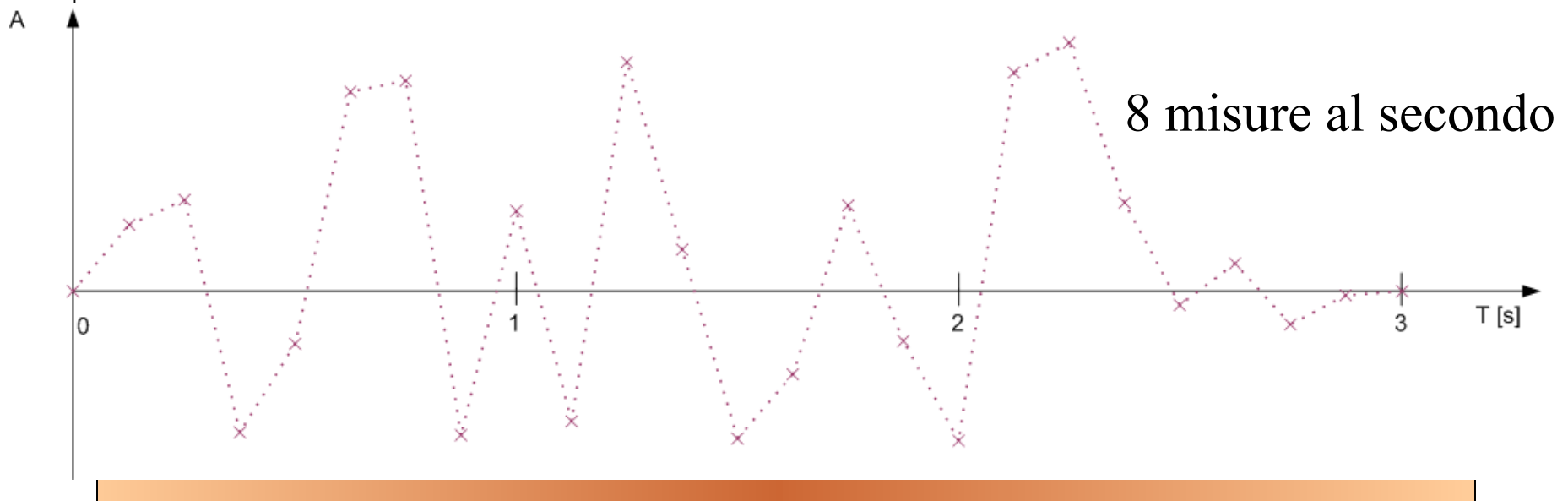
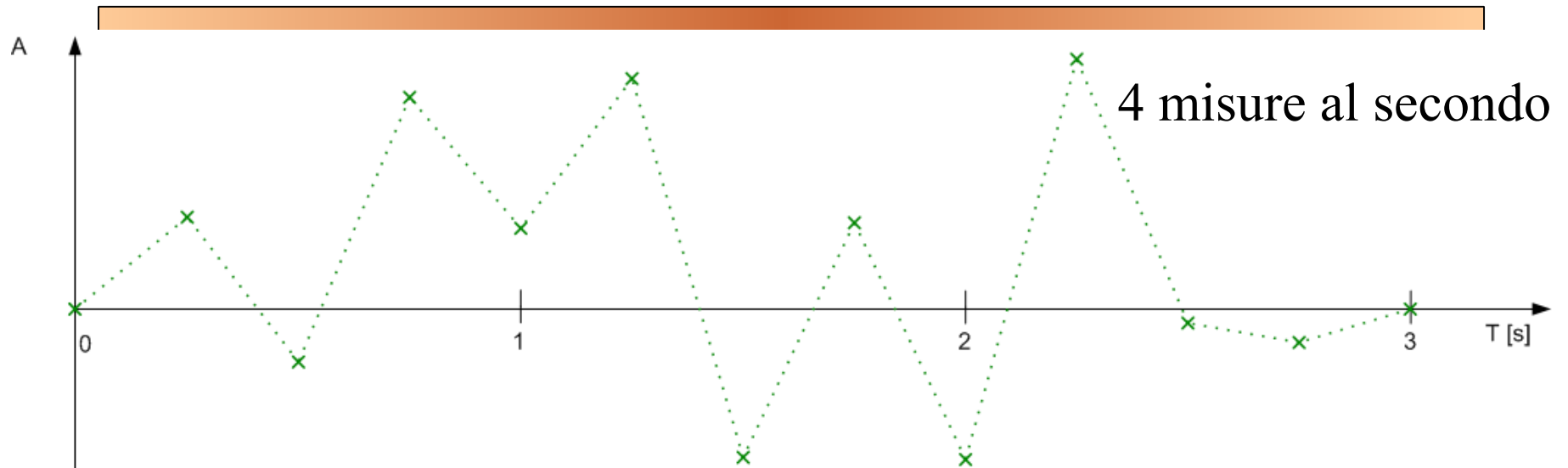
Come vedremo più avanti, il numero di misure da effettuare dipende dal tipo di segnale che si intende acquisire. Nel caso della musica conservata su un CD audio, il suono è stato campionato misurando l'ampiezza del segnale 44.100 volte al secondo (pari ad una misura ogni 23 milionesimi di secondo).

Maggiore è la frequenza di campionamento maggiori saranno i dati da conservare.

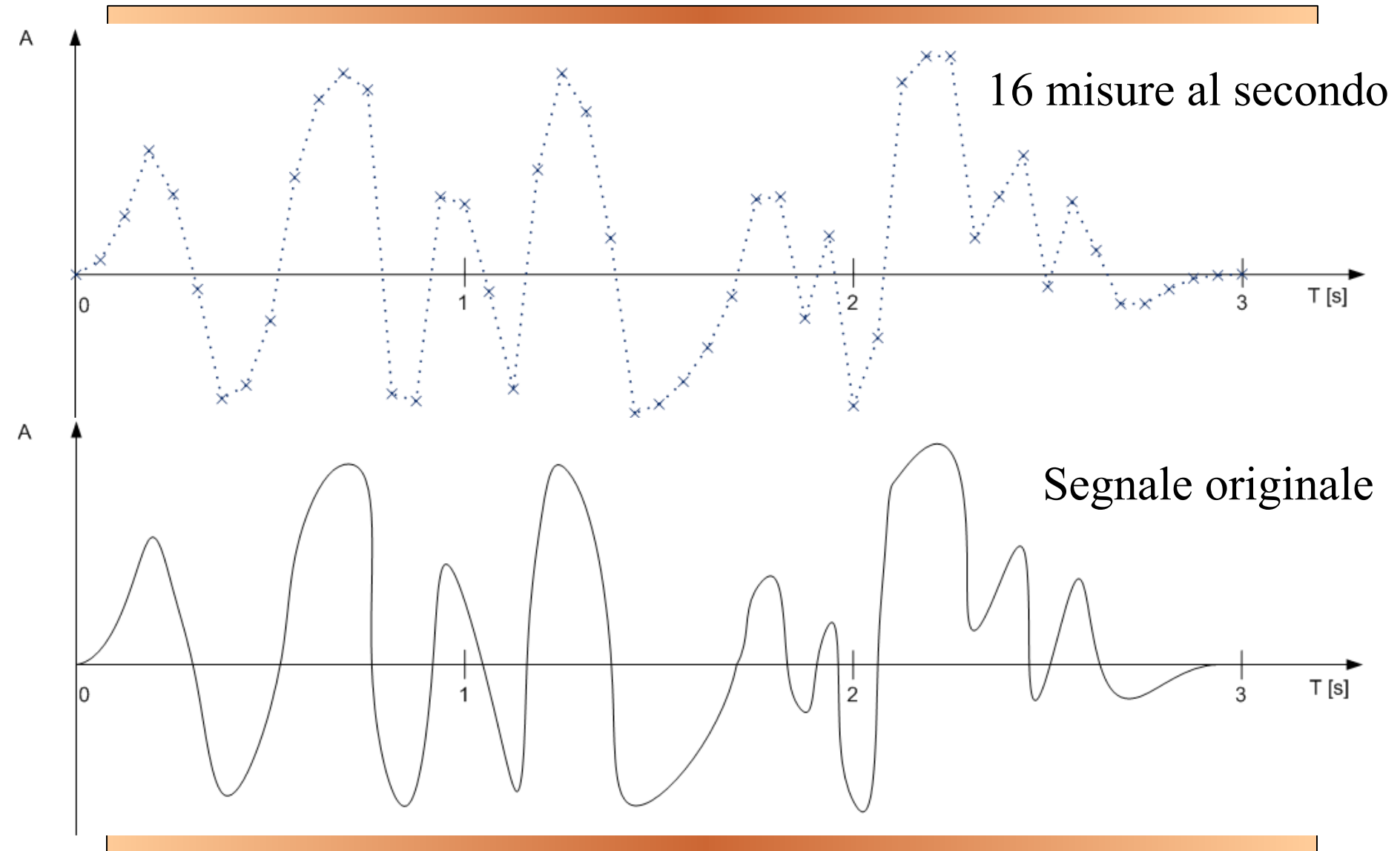
# *Campionamento di un'onda*



# *Campionamento di un'onda*



# *Campionamento di un'onda*



## *La quantizzazione*

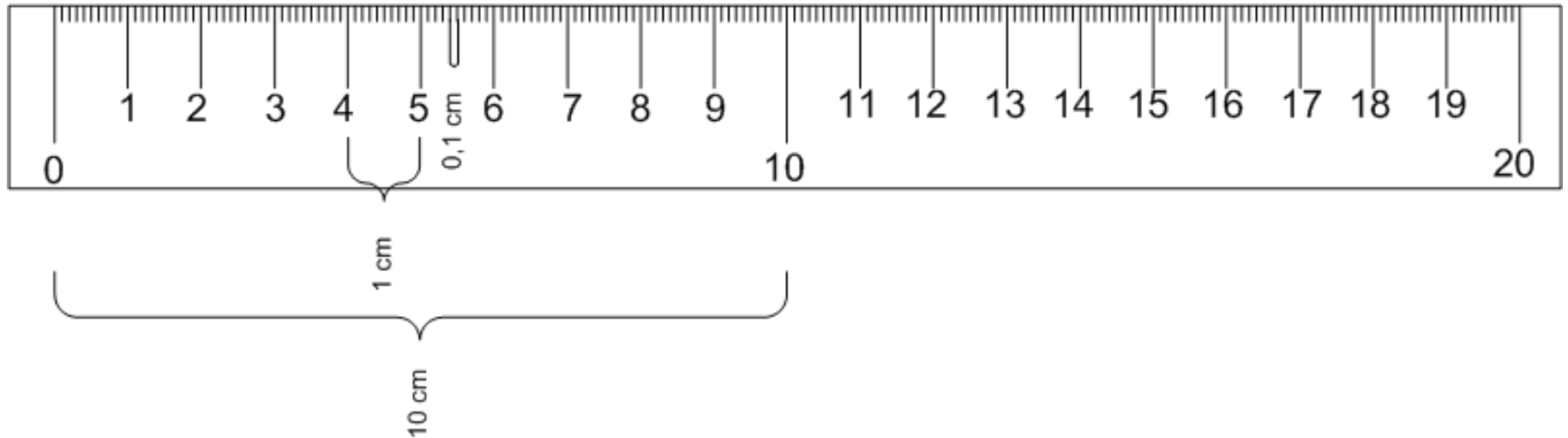
---

Una volta definito l'intervallo temporale tra le misure, è necessario specificare con quale precisione deve essere memorizzato il numero ricavato dalla misura.

L'operazione con cui viene assegnato un numero, compreso in un opportuno intervallo, all'ampiezza del segnale è detta **quantizzazione**.

In generale l'ampiezza di un segnale, per ogni campionamento, viene rappresentata da un valore all'interno di un insieme di 65.536 valori (16 bit). Tale valore è spesso indicato come “bit deep”.

# *La quantizzazione*



## *La digitalizzazione*

---

Al termine del processo di digitalizzazione, un segnale audio è trasformato in una sequenza di numeri che corrispondono all'ampiezza della deformazione del microfono misurata ad intervalli di tempo regolari.

Per riprodurre tale segnale è sufficiente inviare al dispositivo di riproduzione (ad esempio le casse audio) un impulso elettrico, con la stessa cadenza temporale con la quale è stato campionato il segnale, e di intensità proporzionale al valore dell'ampiezza del segnale.

Il fattore di proporzionalità tra ampiezza del segnale digitalizzato ed intensità dell'impulso elettrico inviato è definito utilizzando la manopola del volume.

---

## *La frequenza di campionamento*

In fase di acquisizione di un segnale audio, come già visto per le immagini, è necessario domandarsi quale sia la frequenza ideale di campionamento.

Come nel caso delle immagini anche per l'audio la domanda è mal posta, poiché non esiste un valore corretto *a priori*.

Poiché ad un minore intervallo tra un campionamento ed un altro (e questo vale sia per le immagini che per l'audio) corrisponde una maggiore quantità di informazioni acquisite tale valore dipende dalla tipologia di informazioni che si intende acquisire. Nel caso delle immagini, le informazioni d'interesse sono i dettagli e per l'audio?

## *La frequenza di campionamento*

---

Nel caso dell'audio, le informazioni a cui siamo interessati sono le frequenze che compongono il segnale. Ogni suono, infatti, può essere trattato come una opportuna combinazione di onde acustiche di frequenza diversa.

Questo è evidente pensando alla musica: ascoltando un'orchestra che suona, l'onda che arriva alle nostre orecchie è la somma delle onde prodotte da ogni singolo strumento: chi è ben allenato può anche distinguere i singoli strumenti nell'ensemble.

## *La frequenza di campionamento*

In realtà uno strumento non emette un segnale monocromatico (cioè composto da una sola frequenza), ma composto da molte frequenze diverse (che ne caratterizzano il timbro).

Per comprendere l'associazione tra informazioni e frequenza si può pensare anche al seguente caso: per comunicare con qualcuno, si potrebbe associare ad ogni tasto del pianoforte una lettera. Il mittente preme i tasti del pianoforte sulla base del messaggio, il destinatario riconosce quali tasti sono stati premuti ed è in grado di decifrare il messaggio inviato.

## *La frequenza di campionamento*

Una volta compreso che le informazioni da conservare sono legate alle frequenze che compongono il segnale audio, la domanda più sensata è: qual'è la massima frequenza che il nostro orecchio è in grado di percepire?

L'orecchio umano è in grado – in media – di percepire frequenze comprese tra i 20 Hz ed i 20.000 Hz.

Allora sarebbe sufficiente effettuare una misura ogni 1/20.000 secondi.

Sfortunatamente, il teorema di Shannon-Nyquist afferma che la frequenza di campionamento deve essere almeno il doppio di quella del segnale: almeno 40.000 Hz.

## *La frequenza di campionamento*

Nel caso dei suoni conservati su un CD audio, la frequenza di campionamento è pari a 44.100 Hz, così da acquisire anche segnali poco oltre la soglia di udibilità.

Ma sono veramente importanti tutte queste informazioni oppure è possibile utilizzare una frequenza di campionamento diversa?

Per saperlo è necessario ricavare il contributo delle diverse frequenze al segnale che stiamo campionando.

Il contributo delle diverse frequenze al segnale può essere ricavato calcolando la Trasformata di Fourier del segnale stesso. Non ci interessa capire come funziona ma vediamo come funziona nella realtà.

Una volta visto uno spettro di Fourier (e cioè quali frequenze compongono un segnale), è immediato comprendere il contributo del bit deep.

Quantizzando un segnale a 16 bit, si vede osservando lo spettro di Fourier che il livello del rumore è molto basso (circa -110 db: circa  $1/100.000$  del massimo contributo in frequenza!).

Quantizzando lo stesso segnale con 8 bit, il livello del rumore sale a circa -80 db (circa  $1/10.000$  del massimo contributo in frequenza). Un valore basso ma che può essere percepito.

## *Il CD audio*

Il CD audio è il sistema di qualità migliore per conservare i contributi audio: infatti il segnale è campionato a 44.100 Hz e digitalizzato a 16 bit. Inoltre il segnale è completamente stereo: i segnali per le due orecchie sono digitalizzati indipendentemente.

Per conservare un secondo di audio su CD sono quindi necessari:  $44.100$  (numero di campionamenti al secondo)  $\times 16$  (bit deep)  $\times 2$  (canali) =  $1.411.200$  bit. Una canzone di 3 minuti pesa, così:  $338.688.000$  bit

Come per le immagini, un file contenente un brano audio archiviato con una qualità molto elevata richiede molto spazio e non può essere inviato via internet.

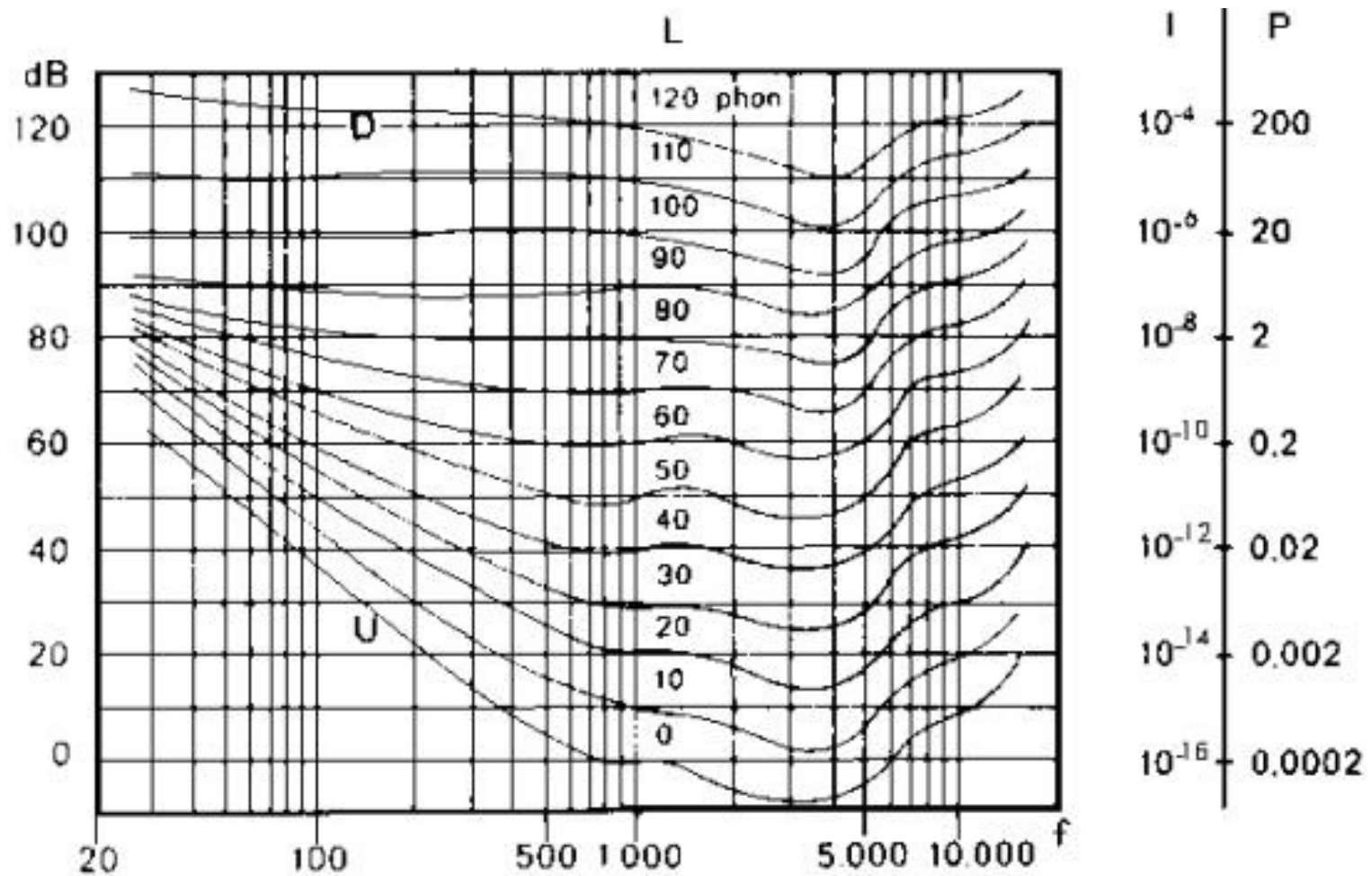
## *I formati psicoacustici.*

Per realizzare dei formati audio che consentano l'archiviazione di brani conservando una elevata qualità in fase di riproduzione ma riducendo il peso del file, sono state studiate le risposte del nostro orecchio agli stimoli acustici, così da comprendere quali informazioni sono essenziali e quali possono essere eliminate.

I formati che si basano su questi studi sono detti formati psicoacustici: ovviamente sono formati lossy.

La prima cosa da osservare è che la risposta in frequenza del nostro orecchio ad uno stimolo acustico dipende dalla frequenza dello stimolo stesso: uno stimolo appena percettibile ad una frequenza, può essere percepito come fastidioso se inviato ad una frequenza diversa.

# Audiogramma



## *I formati psicoacustici.*

Non ha quindi senso conservare informazioni su segnali di frequenza inferiore alla soglia di udibilità.

Inoltre, il nostro orecchio maschera un suono ad alta intensità con uno di frequenza vicina ma intensità più bassa. Per tale motivo, piuttosto che inviare le informazioni relative a tutte le frequenze che compongono il segnale, è sufficiente inviare informazioni relative ad alcuni segnali che mascherano gli altri.

Sempre gli studi psicoacustici hanno poi evidenziato che il nostro cervello non percepisce, per i suoni molto bassi, l'effetto stereo: è sufficiente quindi inviare alcuni segnali come stereo e gli altri come mono, risparmiando spazio.

## *Il formato MP3*

---

Il formato MP3 è stato il primo ad implementare un modello psicoacustico per valutare quali informazioni conservare e quali eliminare.

Oltre a quanto visto il formato MP3 codifica, sulla base del tipo di segnale, pezzi diversi dello stesso brano con qualità diverse.

Infine, una volta completata la codifica, i dati vengono compressi (in modo lossless) come farebbe winzip.