

**Disclaimer:** il seguente documento contiene degli appunti rielaborati, pertanto questo non sostituisce né il testo consigliato dal docente né tantomeno il materiale didattico fornito.

## **LA DIGITALIZZAZIONE**

Segnale analogico: rappresentabile da una funzione continua nella variabile temporale.

Segnale digitale: rappresentabile da una funzione che figura graficamente solo per certi istanti di tempo (una funzione discreta).

Cronologicamente, prima dell'audio si ha a che fare col suono, il problema è riprodurre il suono dato l'audio.

**Trasduzione:** è il fenomeno per il quale è possibile convertire le onde sonore in informazione digitale, ossia un'onda elettrica.

Un canale audio è un dispositivo fisico il cui utilizzo è trasmettere una certa armonica apparente al suono originale (infatti i suoni complessi sono dati dalla somma di diverse armoniche). Sfruttando diversi flussi informativi, ossia servendosi di più canali capaci di produrre armoniche diverse, è possibile mescolare queste informazioni al fine di ottenere suoni complessi. Un **audio**

**multicanale** è un segnale audio costituito da diversi flussi informativi, ognuno dei quali viene considerato come un segnale audio a se stante.

Un audio che per essere riprodotto necessita di un solo canale è detto "mono", mentre uno più complesso che richiede più canali è detto "stereo".

Il suono genera variazioni di pressione nell'aria, una digitalizzazione, per potere poi riconvertire l'audio nel suono di origine, deve in qualche maniera essere in grado di mantenere questa informazione circa le variazioni di pressione. L'obiettivo è associare ad una certa variazione di pressione una certa variazione di tensione elettrica.

Come anticipato prima, grazie al fenomeno della trasduzione è possibile produrre un'informazione digitale a partire da una analogica, banalmente questa operazione è possibile grazie ad un trasduttore che in seguito all'ottenimento dell'informazione digitale si occupa di memorizzarla su un qualsiasi apparecchio di memoria (disco in vinile, audiocassetta...).

Per realizzare questi apparecchi viene impiegato il quarzo, esso è una gemma che ogni volta che subisce una pressione produce una certa differenza potenziale proporzionale alla pressione ricevuta. Il quarzo ha delle molecole esagonali, ognuna delle quali possiede una propria carica (positiva e negativa); quando questo materiale è soggetto ad una forza di pressione, le sue molecole si avvicinano fra loro e come risultato la differenza di carica fra di esse produce, nell'area di spazio che le separa, un campo elettrico che è causa differenza di potenziale prima citata.

Capito come si riesce a creare un'associazione fra pressione prodotta dal suono e differenza di potenziale che dà vita all'audio, serve ora risolvere la problematica di salvare da qualche parte questo segnale.

Gli esempi più occhi di memorizzazione sono:

-Disco in vinile: questo materiale veniva scavato; i solchi prodotti in esso venivano poi letti dalla puntina del giradischi (la cui architettura è omessa data la sua complessità), essa era in grado di riprodurre le differenze di potenziale associate all'informazione dell'audio impresso nel disco. Il resto dell'architettura del giradischi si occupava di riprodurre il suono dell'informazione audio letta.

-Nastri di videocassette: in essi sono presenti in abbondanza dei dipoli magnetici, questi producono un campo magnetico che induce un campo elettrico il quale genera le differenze di potenziale che tengono memoria della traccia audio incisa sulla cassetta. Quando poi la cassetta verrà inserita all'interno del lettore, esso sarà in grado di ricostruire l'informazione sonora a partire da quella audio memorizzata.

Fasi di un processo di registrazione:

Trasduzione —> pre-amplificazione —> amplificazione —> registrazione —> incisione su supporto di memoria

Fasi di un processo di registrazione/ riproduzione:

lettura di un supporto di memoria → pre-amplificazione → amplificazione → riproduzione del suono

Tutte le trasformazioni fisiche alle quali il suono è soggetto per ottenere l'audio, introducono un errore che influirà sulla riproduzione sonora della traccia incisa. Queste distorsioni circa l'informazione originale vengono genericamente identificate come rumori.

L'SNR (signal-noise rateo) è un indice che definisce il rapporto fra segnale e rumore e che quindi informa sulla servilità dell'audio per il quale esso è misurato:

$$SNR = S/N$$

S: ampiezza massima del segnale originale

N: ampiezza massima della distorsione introdotta dal rumore

Più è piccolo l'SRN, più è inquinato (e quindi meno utilizzabile) il segnale audio oggetto di studi. L'SNR si può anche misurare in dB:

$$SNR_{dB} = 10 \log(S^2/N^2)$$

Si fa il quadrato del rapporto perché può capitare che le ampiezze siano negative. Ciò non introduce errori calcolativi perché essendo presente un logaritmo in questa relazione, è bene prendere precauzioni circa il segno del suo argomento.

Pro e Contro delle apparecchiature per la riproduzione sonora:

Vantaggi: la strumentazione che consente di riprodurre segnali digitali nella loro forma analogica è di semplice realizzazione, economica e le diverse frequenze che consente di riprodurre sono moltissime.

Svantaggi: i suoni originali possono essere soggetti a distorsioni da parte del rumore elettromagnetico, ciò ne inquina la qualità in riproduzione; i macchinari che svolgono queste funzioni non sono duraturi nel tempo, infatti invecchiando hanno prestazioni peggiori; macchinari costruiti in maniera diversa, rappresentano il suono in maniera diversa.

La rappresentazione digitale sfrutta dei numeri (infatti i segnali audio sono discreti) per tenere conto delle variazioni che il segnale subisce dal suo inizio alla sua fine. Ci si distacca dalla funzione continua che caratterizza i segnali analogici poiché essi sono assai più complicati da trattare dai macchinari.

Ci si concentri adesso sul processo che produce un segnale digitale (audio) a partire da uno analogico (suono).

Il suono originale è soggetto ad una trasduzione, viene quindi generata, a partire dalla sua pressione, una serie di differenze di potenziale che ne contengono l'informazione. Studiando mediante una nuova funzione la variazione delle differenze di potenziale prodotte dalle variazioni di pressione si ottiene ancora una volta una funzione continua (ciò è logico dato che questa funzione studia una variazione nel tempo); la nuova relazione ottenuta viene poi: campionata, quantizzata e codificata.

ADC (Analog to Digital Converter): apparecchio che si occupa di eseguire la trasduzione, il campionamento, la quantizzazione e la codifica

DAC: apparecchio che inverte le operazioni del primo dispositivo restituendo il suono associato all'audio.

Alla fine delle operazioni compiute dall'ADC si ottiene la versione digitalizzata del segnale (ossia un segnale discreto) che viene rappresentata in un formato specifico e salvata su una qualsiasi memoria di massa.

quantizzazione: è un particolare campionamento che avviene lungo sull'asse delle y (il campionamento avviene invece sull'asse delle x).

È durante la fase di codifica che viene scelto il formato nel quale salvare l'audio; migliore è il formato, migliore sarà la qualità dell'audio (resistenza ai rumori, qualità del suono...).

Per esempio si può scegliere una codifica che preservi valori di frequenza specifici perché l'audio, nel contesto in cui sarà impiegato, è migliore se dotato di queste frequenze scelte.

Nota: in un calcolatore, il DAC corrisponde con la scheda audio del dispositivo. Le casse audio sono poi in grado di captare le variazioni della grandezza elettrica individuate dal DAC, producendo le variazioni di pressione nell'aria ad esse associate.

ADC e DAC assieme formano il DAQ (Data Acquisition), questo è un dispositivo che sia digitalizza i suoni, sia li riproduce a partire dagli audio.

I dispositivi DAQ contengono: memorie temporanee (utili in fase di registrazione per mantenere alcuni valori effimeri); un decodificatore; un quantizzatore; un clock; un campionatore.

Catena dell'audio digitale:

suono → trasduzione → campionamento + quantizzazione + codifica → copia digitale  
audio salvato in digitale → decodifica + interpolazione → riproduzione sonora

### **il campionamento**

Operazione che consente di ottenere a partire da una funzione continua, una discreta.

Per eseguire un campionamento il primo passo è definire un "sample rate", questo è un intervallo di una certa ampiezza; ogni volta che si percorre un intervallo ampio quanto il sample rate, l'operazione di campionamento prevede di prelevare dallo spettro continuo della funzione sulla quale sta operando, un campione. Com'è intuibile, l'output di questa operazione è un segnale a tempo discreto ottenuto a partire da uno continuo.

Quanto appena descritto è ciò che avviene durante un'operazione di campionamento, tuttavia questa deve essere disciplinata da delle regole per produrre una discretizzazione che una volta riconvertita in forma continua, coincida in maniera più fedele possibile al segnale analogico originale.

Il sample rate deve essere scelto con criterio, non in maniera casuale; il teorema di Nyquist consente di definire la frequenza migliore di campionamento (ossia l'ampiezza del sample rate). Il valore identificato da questo teorema è detto "frequenza di Nyquist".

La definizione di frequenza di Nyquist in un segnale periodico a banda limitata, coincide con la frequenza di valore maggiore individuata lungo l'asse delle ordinate dello spettro del segnale considerato (quindi questo valore di frequenza può banalmente essere ottenuto con strumenti matematici quali la serie di Fourier).

ATTENZIONE → il fatto che il segnale sia a banda limitata è un'ipotesi del teorema, il che significa che qualora il segnale che si volesse campionare fosse a banda illimitata, esso necessiterebbe di essere opportunamente sezionato da un filtro passa-basso; grazie a questo strumento verrà prelevata una porzione di spettro finita la quale, se campionata, consentirà di ricostruire il segnale di partenza.

Un segnale digitalizzato è considerabile fedele se si prende un campione entro un intervallo ampio almeno il doppio della frequenza di Nyquist.

esempio: se la  $f$  di Nyquist vale 100Hz, si potrebbe prendere un campione almeno ogni 200Hz (in verità si prende sempre poco più del doppio).

Potrebbe accadere (raramente) che l'operazione di campionamento, se eseguita al valore preciso del doppio della frequenza Nyquist, prelevi dalla funzione continua dei punti in cui essa coincide esattamente con i punti in cui lo spettro d'onda (funzione periodica) tocca l'asse delle  $x$ , ivi la funzione è nulla, quindi il campionamento preleverebbe campioni tutti nulli. Lo scenario appena descritto è problematico poiché una serie di campioni nulli, quando viene convertita nella sua forma analogica, restituisce una funzione identicamente nulla (ossia un silenzio) → questo campionamento è detto **campionamento critico** perché l'operazione di campionamento di un suono non nullo ha prodotto un audio al quale è associato un suono nullo.

Prendendo una frequenza di poco maggiore al doppio di quella di Nyquist, il campionamento critico non avviene mai.

Essendo la massima soglia di frequenza udibile dall'uomo di 20000 Hz, due buone frequenze alle quali eseguire campionamenti sono: 44100 Hz e 40000 Hz; il primo valore fornito è il migliore fra i due dato che per i motivi già detti la seconda potrebbe portare ad un campionamento critico.

Shannon fu un fisico Americano che si ingegnò nel risolvere il problema della trasmissione dei segnali per tutto il continente, egli intuì la rappresentazione dei segnali in 0 e 1.

Alcuni difetti dei dati registrati sono più facili da risolvere quando il segnale è già salvato su un supporto fisico piuttosto che nella sua forma digitalizzata (infatti il segnale salvato su supporto fisico è visualizzabile come funziona continuo).

La rappresentazione di suoni molto complessi è migliore nella versione analogica piuttosto che in quella digitale.

L'operazione di "**inverse**" inverte asse x e y di uno spettro di fase, in questa maniera si possono sommare suoni che annullano certe componenti del suono (l'obiettivo sono le componenti di disturbo) e annullarle; in questa maniera, ripristinando asse x e y originale dello spettro, si otterrà il segnale più pulito.

Nyquist fu un ingegnere vissuto nel periodo della seconda guerra mondiale, si occupò di studiare sistemi per la rilevazione balistica.

Il tasso di campionamento di uno spettro va scelto individuando nello spettro la sua frequenza più alta (frequenza di Nyquist), raddoppiandola e prendendo un valore poco sopra questo. Ma perché il tasso di campionamento viene scelto così?

Si prenda la funzione treno di impulsi unitario **nel dominio del tempo** (questa è una funzione periodica che presenta un impulso di altezza unitaria ogni volta che si ripete il periodo). La funzione scelta, campiona nel tempo i diagrammi d'onda dei suoni.

Il periodo del segnale treno di impulsi utilizzato per svolgere il campionamento è detto "periodo di campionamento" ( $t_c$ ); l'inverso di questo tempo di campionamento è appunto la frequenza di campionamento (quella frequenza che è poco più del doppio della frequenza di Nyquist).

Modulando il periodo del segnale treno di impulsi, si cambia il tasso di campionamento che si vuole applicare ai diversi segnali.

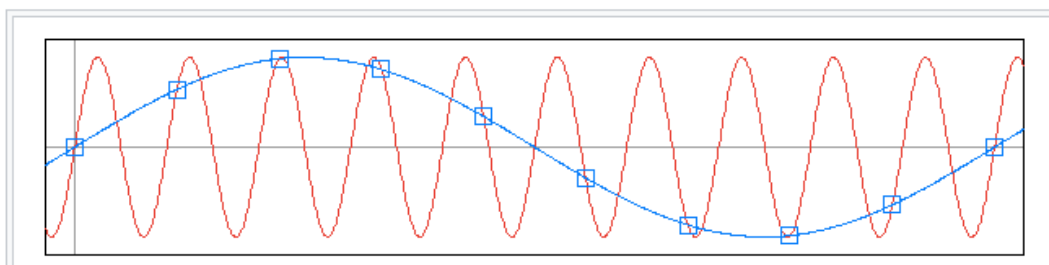
Dato un suono generico, applicando al suo diagramma d'onda il treno di impulsi si ottiene un campione (un impulso) in corrispondenza di ogni multiplo di  $t_c$  (quindi: primo campione a 0

secondi; secondo campione a  $t_c$  secondi; terzo campione a  $2t_c$  secondi; quarto campione a  $3t_c$  secondi...). I campioni individuati saranno tanto alti quanto lo era la funzione originale in corrispondenza dei punti in cui è avvenuto il campionamento, algebricamente e geometricamente altro non si fa che il prodotto fra il diagramma d'onda del suono e la funzione treno di impulsi. Nel diagramma dello spettro, a seconda del tempo di campionamento scelto, si otterrà un grafico che dipende molto dalla frequenza di campionamento ( $f_c$ ) data da:  $1/t_c$ .

Sia dato per scontato, senza enunciati, che lo spettro del segnale campionato nel tempo sia uno spettro righe.

L'**aliasing** è un fenomeno che ha luogo quando si esegue un campionamento al di sotto del doppio della frequenza di Nyquist.

Di seguito è descritto questo fenomeno: data una certa forma d'onda (grafico della funzione del suono nel dominio del tempo) se il campionamento avviene al di sotto della frequenza consigliata, può capitare che i campioni prelevati in fase di ricostruzione del segnale continuo, individuino una funzione diversa da quella campionata originariamente, ciò avviene a causa della distanza fra un campione e l'altro che suggerisce al sintetizzatore una ricostruzione diversa da quella di partenza. Il fenomeno descritto ha questa etimologia poiché quando si avviene individua un "alias" di quel segnale campionato, diverso da quello originale.



In figura è stato eseguito un campionamento al di sotto della frequenza suggerita. attenzione! -> il grafico riporta le funzioni nel dominio del tempo, la frequenza di cui si sta parlando è quella data dalla relazione:  $t_c = 1/f_c$ .

- Funzione rossa: forma d'onda del suono che si vuole campionare
- Quadrati blu: campioni prelevati dalla funzione rossa, questi campioni sono stati presi ad una frequenza inferiore a quella suggerita
- Funzione blu: forma d'onda ottenuta a partire dai campioni prelevati (alias della funzione rossa)

Come si vede chiaramente, la funzione ottenuta a partire dal campionamento eseguito è diversa da quella originale, è appunto venuto un aliasing.

Ultime considerazioni sul campionamento:

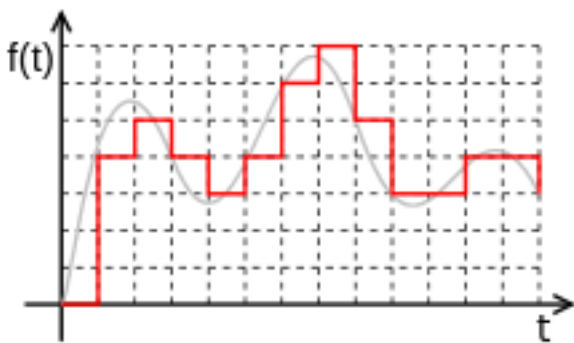
È stato più volte detto che un segnale elettrico, durante i suoi spostamenti, può essere contaminato da dei rumori. Grazie a dei filtri passa-basso è possibile eliminare i campioni di frequenza che trasportano l'informazione relativa al disturbo (in realtà non si parla mai di cancellazione, ma di attenuazione tale che le componenti del rumore risultino il più trascurabili possibile).

Un campionamento che avviene al di sotto della frequenza consigliata è detto sottocampionamento; una conseguenza di questo fenomeno è già stato detto essere l'aliasing. Il motivo dietro la scelta di 44100 Hz come massimo valore della frequenza di campionamento è legato al fatto che non ha senso campionare quei suoni che l'essere umano non è in grado di ascoltare. Ponendo questo limite in fase di lavoro/progettazione si risparmia parecchio perché di lavorerà su una banda di frequenze limitata.

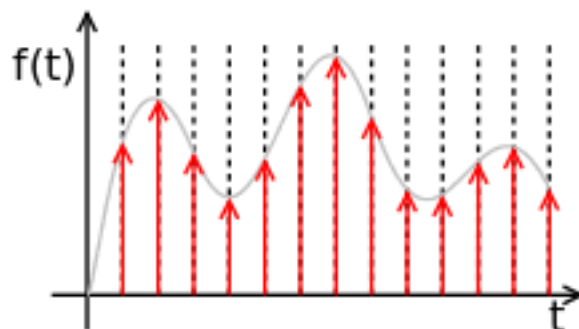
### la quantizzazione

Processo che discretizza un segnale continuo, approssimandolo in seguito a delle operazioni svolte lungo l'asse delle ordinate. Ovviamente, visto che la quantizzazione è un processo di approssimazione, ci si aspetta che essa introduca un certo errore sulla funzione d'origine.

ATTENZIONE -> segnale a valori discreti/discreto e segnale a tempo discreto sono due cose diverse!! Il primo è un segnale che assume valori in ogni istante di tempo, tuttavia questi valori non sono lineari, variano da intervallo a intervallo; il secondo è invece un segnale che vale 0 ovunque, fatta eccezione per alcuni istanti di tempo in cui assume un valore finito.



SEGNALE DISCRETO



SEGNALE A TEMPO DISCRETO

La qualità di un processo di quantizzazione si misura con un valore detto **precisione**, maggiore è la precisione di una quantizzazione, minore è l'approssimazione fatta sulla funzioni e di conseguenza minore sarà l'errore introdotto dall'operazione.

La quantizzazione prevede la suddivisione del grafico, in cui è rappresentata una forma d'onda, in aree parallele all'asse delle x (prese quindi lungo l'asse delle y) dette quanti.

La scelta dei quanti distingue 2 diversi tipi di quantizzazione:

**uniforme/ lineare:** l'ampiezza dei quanti è costante

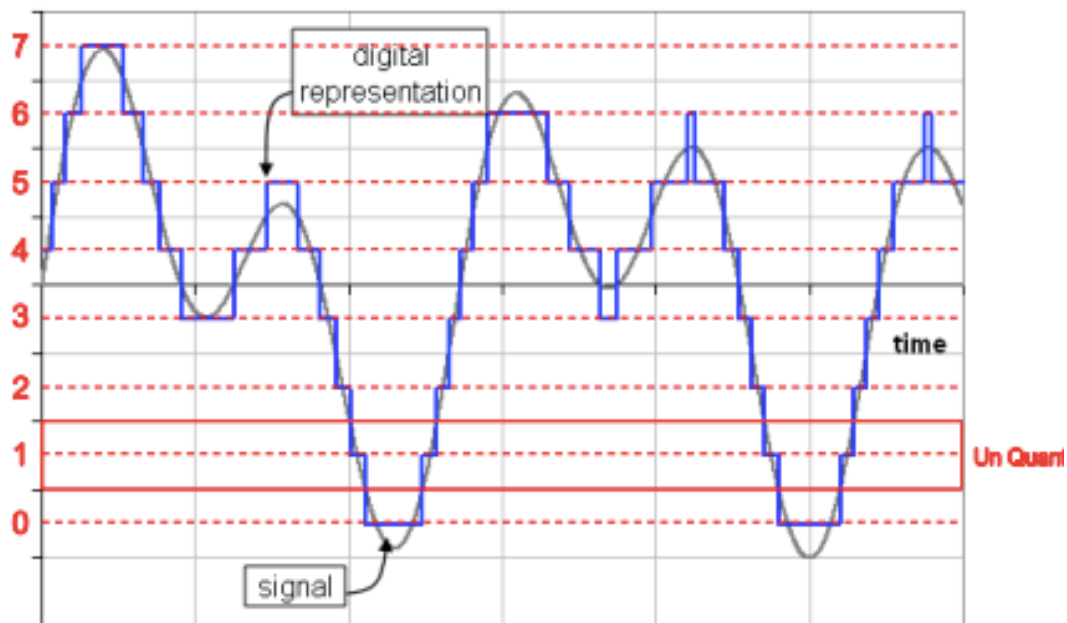
**non uniforme/ non lineare:** l'ampiezza dei quanti non è costante

Si definisce come **valore di fondo scala** (VFS, anche noto come “gamma dinamica”) il valore che identifica la dimensione massima che può essere impiegata per realizzare i quanti di una quantizzazione di un segnale. Il VFS si calcola come la sottrazione fra il massimo e il minimo valore che può assumere la funzione soggetta alla quantizzazione:

$$VFS = MAX_{funz} - MIN_{funz}$$

Secondo questa definizione, se  $Max=10$ ,  $Min=2$ , allora  $VFS=Max-Min=8$ , in questo esempio si hanno quindi a disposizione 8 livelli; li si prenda (sempre per esempio) di dimensione di 3 bit ognuno, in questa maniera si ottengono 8 quanti uniformi distribuiti come sotto-intervalli  $[2,3)$ ,  $[3,4)$ , ...,  $[7,8)$ , altrimenti si possono scegliere 8 quanti non uniformi dove la grandezza del quanto cresca in maniera logaritmica.

Riassumendo quanto detto finora: la quantizzazione è un'operazione che avviene sull'asse delle y discretizzandola in un certo numero di intervalli (i quanti), il cui risultato è un'approssimazione della funzione bersaglio; grazie a questa si evita di visualizzare infinite armoniche nello spettro di un segnale. Questo risultato è utile perché è impossibile studiare funzioni i cui spettri hanno infinite componenti. Graficamente, l'approssimazione della funzione originale è una funzione formata da tanti “gradini” di altezza scostante.



In questo esempio in figura:

la funzione è stata divisa in 8 quanti, ognuno con la stessa ampiezza (quindi si tratta di quantizzazione uniforme). Come si vede ogni quanto è caratterizzato da: un inizio, un valore di centro esatto e una fine. Quando la funzione cade all'interno di un certo quanto, la quantizzazione approssima il valore della funzione a quello di mezzo del quanto in cui essa è caduta.

Esempio: si supponga di avere un quanto ampio 1 da 3 a 4, perciò il valore di mezzo di questo (essendo asse ampio 1) è 3,5. Ognuno dei punti della funzione originale che cade in un intervallo lungo la direzione di y compreso fra 3 e 4, verrà approssimato a 3,5; tutti i punti al di fuori di questo range verranno approssimati al valore medio del quanto nel quale essi cadono.

Per fare sì che il segnale quantizzato sia trasmissibile, bisogna numerare i quanti (nell'immagine sono numerati da 0 a 7 per esempio) e dare loro una **codifica binaria**. Procedendo in questa maniera, quando verrà letta una certa sequenza di bit, sarà noto il numero del quanto al quale ci si sta riferendo e di conseguenza sarà anche noto il valore al quale quel quanto approssima la funzione originale (esempio: dati 8 valori di intervalli di quantizzazione, essi verranno numerati su 3 bit da 000 a 111 -> se il quanto 101 approssima la funzione a +26, quando il ricevente riceve una sequenza di 3 bit corrispondente a 101, capisce che la funzione vale +26).

Come già detto, essendo l'operazione di quantizzazione un'operazione di approssimazione, essa introduce un certo errore. Esiste una relazione per valutare il massimo errore introducibile da un'operazione di quantizzazione:

$$E_{max} = VFS/2Q$$

Dove:

$E_{max}$ : errore massimo che l'operazione di quantizzazione considerata è capace di introdurre

VFS: valore di fondo scala

Q: numero di quanti della quantizzazione considerata

$Q = 2^N \rightarrow N$ : numero di bit scelti per la codifica

**Profondità(N)**: numero di bit scelto per la codifica di una quantizzazione (3 nel caso dell'immagine scelta come esempio).

Maggiore è la profondità, maggiore è la qualità del segnale quantizzato (ossia che trasporta una maggiore quantità di informazioni del suono originale) ciò comporta anche che l'informazione trasportata sarà soggetta a meno distorsioni.

La qualità del segnale in uscita è direttamente proporzionale al numero di bit scelti per la codifica. Allora perché non scegliere N molto grande? Perché più N è elevato, più sarà corpulenta la dimensione del file audio che si otterrà alla fine del processo di sintetizzazione.

Oltre a rilevare l'errore che una certa quantizzazione introduce, si può anche stimare la distorsione che in fase di ricostruzione del suono analogico avrà comportato l'operazione di quantizzazione.

SQNR  $\rightarrow$  indice (misurabile anche in decibel) che informa circa la distorsione sonora alla quale sarà soggetta il suono quando verrà riconvertito a partire dal suo formato digitale.

"Signal to Quantization Noise Ratio"

Nel caso di una quantizzazione uniforme, la formula di quest'indice è:

$$SQNR = 2^N \times \sqrt{3/2} \rightarrow \text{graficando questa funzione, si vedrebbe che essa è un dente di sega!}$$

Più l'SQNR è elevato, maggiore è il numero di bit usati per la codifica, migliore è la qualità del segnale in uscita.

Quello appena calcolato è un numero puro adimensionale, se si vuole invece calcolare quanto vale la distorsione a livello di decibel il calcolo viene:

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

Da questa relazione si capisce che la distorsione aumenta di circa 6 dB per ogni nuovo bit che si aggiunge ad un'operazione di quantizzazione.

I canali di trasmissione supportano il trasporto di byte (8 bit) ciò significa che più sarà ampia la quantizzazione, maggiore sarà la quantità di canali necessari alla trasmissione.

esempio:

se ho un tasso di campionamento di 22050 campioni al secondo, codificati in 16 bit, necessiterò di 2 canali (16 bit sono 2 byte ovvero 2 canali da un byte ciascuno); il tasso di trasferimento su questo canale sarà di:

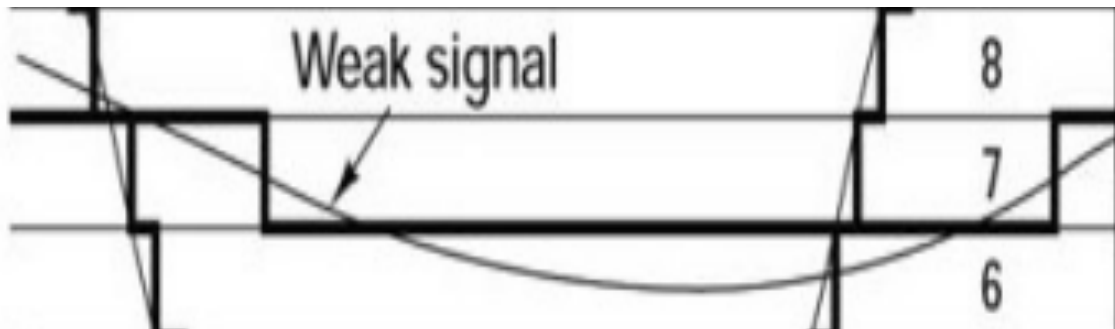
$$[(\text{numero dei canali}) \times (\text{numero di byte di codifica}) \times (\text{tasso di campionamento})] / 1000$$

quindi, usando i dati dell'esempio, ogni campione viaggia alla velocità di:

$$[2 \times 2 \times 22050] / 1000 \text{ kbyte al secondo (88,2 kbyte/sec)}$$

Segnali analogici con intensità bassa sono più soggetti alla distorsione rispetto a quelli la cui intensità è maggiore.

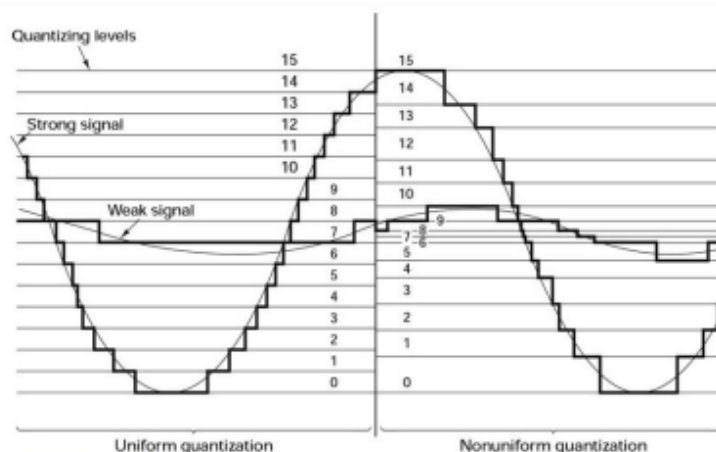
Questo fenomeno avviene perché essendo queste funzioni molto basse (ossia che oscillando molto vicino all'asse delle ascisse), esse finiscono con l'incontrare un numero di quanti molto inferiore a quello che la quantizzazione mette a disposizione.



Quello in immagine è il caso più estremo in cui una funzione di ampiezza molto bassa (quella meno calcata), a causa della sua altezza è in grado di incontrare 2 soli quanti della quantizzazione scelta, di conseguenza verrà approssimata ad una funzione onda-quadra (ossia una funzione che oscilla fra 2 soli valori), ma la funzione di partenza oscilla fra più valori (molto piccoli) che non è stato possibile discretizzare a causa della ampiezza troppo elevata dei quanti, pertanto se si provasse a convertire la funzione approssimata con questa quantizzazione ad una funzione continua, se ne otterrebbe una profondamente diversa da quella originale (appunto perché la funzione approssimata considerata oscilla fra 2 soli valori, quindi dà un'informazione praticamente nulla su tutti gli altri valori assunti da quella di partenza).

Una soluzione che viene in mente è quella di ridurre la dimensione dei quanti, così da accrescerli in numero e ottenere una precisione maggiore, tuttavia quando si progetta un dispositivo per la quantizzazione si deve tenere a mente che per esso passeranno suoni di intensità sempre differenti, un suono di alte intensità che incontra tutti questi quanti, oltre ad essere molto pesante da veicolare sarà anche soggetto ad un grandissimo numero di distorsioni (ciò è noto grazie alla formula del SQNR).

La soluzione a questo problema è introdurre la **quantizzazione non uniforme**. Questa quantizzazione si ottiene, data una scala di quantizzazione uniforme preesistente, andando a quantizzare ulteriormente i quanti in corrispondenza dei quali passano funzioni di bassa intensità, in questa maniera le funzioni di bassa intensità passanti per quelle aree di grafico saranno quantizzate in maniera più affidabile considerando valori più piccoli. Questa soluzione consente di realizzare dispositivi di quantizzazione più elastici.





Esiste in realtà un altro modo per risolvere questo problema (anche se è preferibile l'impiego della quantizzazione non uniforme), ossia il **dithering**. Questa tecnica consiste nell'aggiungere alla funzione di bassa intensità un rumore blu, questo la distorce lievemente facendola accedere a quanti più elevati, consentendo una quantizzazione più affidabile; ovviamente quando si riconvertirà in analogico la funzione il contributo del rumore andrà tolto (altrimenti averi il suono compromesso!).

La quantizzazione non uniforme introduce una scala di valori non costante, quindi i segnali di alta intensità soggetti ad essa passeranno per le aree più ricche di quanti destinate alle funzioni di bassa intensità, ciò quindi aumenta il numero di quanti che queste funzioni attraversano, ma ad un numero di quanti maggiore è stato visto corrispondere una distorsione maggiore, quindi ivi ci sarà un rumore di intensità maggiore. Anche se il discorso appena concluso è corretto, questo rumore è tollerabile perché essendo l'intensità del segnale molto alta per ipotesi, mentre il rumore nasce in zone di grafico in cui l'intensità è bassa per costruzione, l'intensità del suono originale soverchierà quella del rumore rendendolo inefficace. Quindi questi dispositivi sono in tutto e per tutto funzionanti.

Interpolazione: è il processo inverso a quello di quantizzazione.

L'interpolazione riproduce il segnale quantizzato con un certo margine di errore a causa delle diverse approssimazioni che vengono introdotte in fase di quantizzazione, è stato appunto detto più volte che la quantizzazione è un'operazione che introduce parecchi errori.

Una volta capito come attuare la conversione da analogico a digitale resta da rispondere all'ultimo interrogativo: cosa serve conoscere per salvare il file audio?

Le grandezze necessarie a determinare le dimensioni di un file audio sono: numero dei canali (C), tasso di campionamento ( $T_c$ ), ampiezza/profondità in bit dei quanti (N).

La dimensione (S: size) dell'audio si misura come:

$$S = T_c \times N \times C \times D$$

Dove "D" indica la durata della traccia audio in secondi.

Togliendo la durata, il bitrate (BR) della traccia, ossia il numero di bit della traccia audio trasmessi in un secondo è:

$$BR = T_c \times N \times C$$

L'obiettivo è garantire un buon rapporto fra qualità dell'audio e bitrate di esso, del resto un bitrate più piccolo è sinonimo di un audio più piccolo perciò sarà sia più facile da trasferire sia meno invasivo per il calcolatore che lo ospita.

ES//

Sia dato un tasso di campionamento pari a 44.1 kHz e una PCM (pulse code modulation) a 8 bit, calcolare quanti byte servono per memorizzare un audio **stereo** della durata di 2 secondi:

attenzione! -> essendo la PCM a 8 bit (ossia un byte) si potrebbe ingenuamente pensare che il canale di cui necessita l'audio per essere propagato sia uno solo, tuttavia esso deve essere memorizzato in formato stereo e questo genere di audio è stato detto necessitare di più canali (2 nello specifico). Quindi il numero di canali considerato è di 2!!

$$(44100 \times 8 \times 2 \times 2) / 8 = 176400 = 176,4 \text{ kbyte}$$

Si divide per 8 perché altrimenti il numeratore da solo darebbe le dimensioni del file in "bit" non in byte.

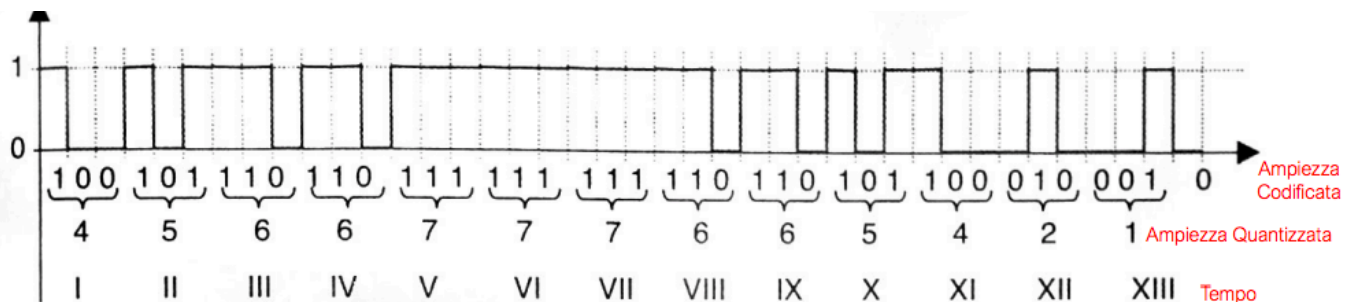
//FINE ES

La Pulse Code Modulation (PCM) è una tecnica di codifica dell'audio digitale.

"parola binaria" = sequenza di bit

Step della PCM:

1. Individuare la profondità dei quanti del processo di quantizzazione subito dal suono in fase di digitalizzazione
2. Prendere un grafico, suddividere l'asse delle X in intervalli ampi quanto lo sono i quanti della quantizzazione scelta per il segnale, mentre lungo l'asse Y disporre i valori 1 e 0
3. Prendere il segnale campionato, quantizzarlo e osservare una per una la codifica del quanto in cui cade ciascun campione
4. Nel grafico preso al punto "2" individuare lungo l'intervallo dell'asse X, associato a ciascuno dei campioni osservati al punto "3", i trattini 0 e 1 in corrispondenza del quanto in cui esso era contenuto



Si spieghi di seguito come è stato ottenuto questo grafico:

Gli intervalli di questa asse X sono di ampiezza 3 bit, quindi i quanti dell'operazione di quantizzazione subito da questo audio hanno una profondità di 3 bit.

Il primo campione della funzione campionata si deduce essere caduto nel quarto quanto della quantizzazione, quindi tradotto in binario esso corrisponde alla parola binaria 100, infatti nel grafico per il primo bit il valore della funzione è 1, per i restanti 2 è 0.

Il secondo campione della funzione campionata si intuisce che cade nel quanto 5 la cui codifica binaria in 3 bit è 101, perciò si osservano 1 nel primo bit, 0 nel secondo e 1 nel terzo.  
etc.

Quindi dal grafico ottenuto mediante PCM si riesce a risalire al grafico dell'audio in seguito alle operazioni di campionamento e quantizzazione, poi a partire da questo sarà possibile riottenere l'audio originale.

Grazie alla rappresentazione PCM è stato possibile trovare una rappresentazione per memorizzare su supporto digitale il suono originale.

Esistono diversi metodi per definire una scala binaria attraverso la quale identificare i valori, le più impiegate sono: offset binario, complemento a 2, segno e magnitudo.

L'offset binario è quella tradizionale alla quale si è abituati, rappresenta i numeri decimali in forma binaria partendo dallo zero fino ad un numero N.

Il complemento a 2 sfrutta l'ultimo bit della parola binaria (quello più a sinistra) per esprimere il segno del numero decimale al quale essa si riferisce, quindi la scala va da -N a N a seconda del range numerico che si vuole rappresentare; il problema di questa scala è che necessita di un bit in più rispetto al normale per rappresentare i numeri rispetto all'offset (esempio: se voglio rappresentare i numeri negativi da -8 a 8 dovrò rappresentarli su 4 bit invece che su 3).

L'ultima scala è quella meno impiegata perché introduce due problematiche: la prima è che essa consente di rappresentare pochi valori (infatti su 3 bit si possono rappresentare solo i numeri da 1 a 3 positivi e negativi), la seconda è legata al fatto che essa possiede DUE ZERI (ossia +0 e -0).

Definito come gli audio vengono salvati sui dispositivi, diventa necessario introdurre delle tecniche di controllo sugli errori dei bit quando questi viaggiano da un dispositivo ad un altro.

Il più semplice è il metodo di controllo di parità, ossia il parity check.

Data una sequenza di bit, si conta il numero di bit 1 contenuti in questa e si tiene conto se questi siano pari o dispari -> se sono PARI il bit per il controllo, detto **bit di parità**, viene impostato a 0; se sono dispari viene invece impostato ad 1. Questo metodo serve a individuare il controllo solo su bit singoli, quindi non è molto efficace se devono essere effettuate correzioni a seguire.

Questo metodo è estremamente elementare, tuttavia può diventare problematico qualora il bit sia affetto da errore sia giusto che quello di controllo. In situazioni di inconsistenza derivanti da questa situazione ci si accorge solo dopo aver processato i dati e scovato il problema viene chiesta una ritrasmissione. Questo espediente funziona ma allunga i tempi di attesa quindi è stata escogitata una versione più avanzata del perito check che impiega più bit di parità:

Una soluzione per fare fronte agli errori che potrebbe introdurre il bit di parità è di introdurre per uno stesso payload (che è un certo carico di dati in uscita) più bit di parità che si occupano di controllare determinate sezioni del codice (esempio: se ho una sequenza di 16 bit posso aggiungere al payload altri 4 bit di parità, ognuno controlla un quartetto di bit del payload). Vantaggi di questo metodo: i controlli sono non solo più immediati, ma se un bit di parità viene affetto da errore solo una piccola parte del payload viene coinvolta, infatti visto che il controllo è così localizzato, la macchina se si accorge dell'errore può correggerlo senza richiedere nuovamente l'invio di tutti i dati. Svantaggi di questo metodo: ad un payload piccolo vengono aggiunti più dati per i diversi bit, il che aumenta la dimensione dei dati.

Se ho un suono che digitalizzato impiega 2 canali, se voglio ridurre ad 1 il numero di canali da esso impiegati devo sommare l'ampiezza dei segnali sui due diversi canali e dividere per 2 il risultato ottenuto, si fa insomma una media fra i 2 canali!

### **cenni sulla rappresentazione digitale**

Esistono 3 diversi modi per rappresentare l'ampiezza digitale di una forma d'onda: valore assoluto, decibel, normalizzata.

- Valore assoluto: Scelta una codifica in bit dell'ampiezza, i valori digitali di essa vanno dal valore negativo al valore positivo decimale dell'ampiezza di bit scelta (esempio: se scelgo una codifica 16 bit, i valori della rappresentazione in valore assoluto andranno da -32.768 a 32.768)
- dB: Una normalissima scala in dB, più volte incontrata nel corso
- Normalizzata: La più utilizzata è quella che va da -1 a 1.

Queste sono scale, ossia rappresentazione di valori, perciò se il sistema utilizza una certa scala, esso potrà essere visualizzato secondo un'altra.

Promemoria: a causa di come è stato definito il dB a partire dalla grandezza fisica della potenza (la cui misura si esprime in Watt), il massimo valore della scala in decibel preso in valore assoluto è lo 0, il che significa che questa scala di rappresentazione individuerà quantità tutte negative.

### **gli equalizzatori**

Un equalizzatore è un dispositivo costituito da tanti filtri passa-banda, questi filtri possono: attenuare, amplificare o annullare le frequenze degli audio che passano attraverso essi. Grazie agli equalizzatori è possibile bilanciare le frequenze del segnale audio attenuandole (privandole di energia) o amplificandole (aggiungendo loro energia).

Parametrici e grafici -> sono i 2 tipi esistenti di equalizzatori

Gli equalizzatori di tipo parametrico sono più precisi rispetto quelli grafici, d'altro canto il suo rivale è più semplice da utilizzare dato che dispone di interfacce grafiche che ne agevolano il funzionamento, oltre che renderlo più intuitivo per gli utilizzatori.

Gli equalizzatori hanno diverse applicazioni pratiche, se ne elenchino alcune:

- correzione del timbro sonoro
- eliminazione dei rumori (tipicamente suoni disturbatori la cui frequenza è nota)
- creazione di nuovi suoni
- amplificazione sonora
- creazione di effetti sonori particolari (squillo della telefonata, suono del walkie-talkie...)

Gli equalizzatori lavorano su diversi canali di frequenza, solitamente questi sono 10 perché 10 è il numero delle ottave musicali.

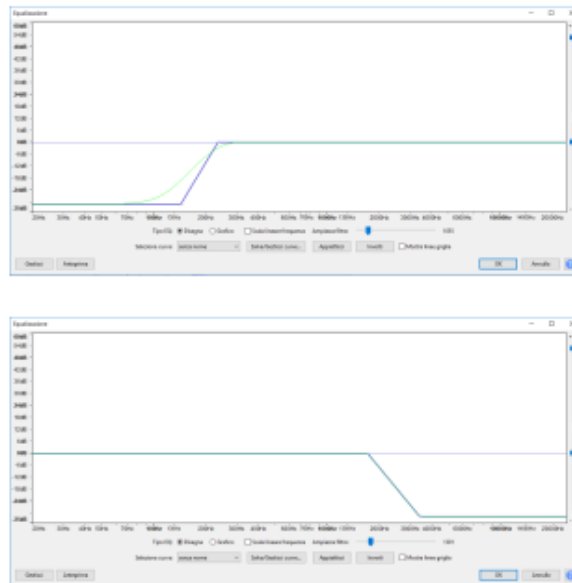
Si osservino ora più nel dettaglio i due equalizzatori citati:

#### equalizzatore parametrico

Questo dispositivo impiega sia dei filtri passa-alto (HPF - High Pass Filter) sia passa-basso (LPF - Low Pass Filter). I filtri sono organizzati in gruppi detti “banchi”, all’interno di uno stesso equalizzatore ci sono più filtri dello stesso tipo perché ciascuno di essi si occupa di una specifica banda di frequenza.

Il banco degli HPF si occupa di annullare le frequenze assai elevate; viceversa il banco degli LPF si occupa delle frequenze di intensità minore.

Elenco delle possibili pendenze della funzione utilizzata dal filtro: 6, 12, 24, 36, 48 dB/ottava  
(Nota: in un diagramma di Bode con scala semi-logaritmica, si apprezzerrebbe meglio l’osservazione di tali grafici di funzione)



**HPF & LPF**

#### equalizzatore grafico

Questo dispositivo, invece di utilizzare separatamente le due tipologie di filtro illustrate nel caso precedente, impiega dei filtri passa-banda che combinano l’effetto di un HPF e un LPF; come risultato, a livello grafico, si osserverà una funzione che vale 0 ovunque fatta eccezione in un certo range di frequenze. Ovviamente, è possibile modulare la grandezza di questo range di frequenze per consentire di ottenere sezioni sempre diverse del segnale audio che sta passando per l’equalizzatore.

Dello spettro del segnale che passa per questo dispositivo, usciranno solo le frequenze contenute nell’intervallo di sezione del filtro passa-banda. A causa del risultato che esso produce, il range di frequenze non cancellate dal filtro sono dette “frequenze di cut-off” e analogamente, la pendenza che ha il filtro (che ne definisce appunto l’ampiezza di azione) è detta “pendenza di cut-off”.



**FILTRO PASSA-BANDA OTTENUTO DALLA  
COMBINAZIONE FRA UN HPF E UN LPF**

In verità, quanto detto riguardo i filtri parametrici, è molto riassuntivo. Su questi dispositivi ci sono altri appunti da fare e di seguito verranno discussi:

Quando è stato introdotto generalmente l'equalizzatore, è stato detto che esso era anche in grado di esaltare le frequenze, non solo eliderle. Infatti ci sono 2 particolari filtri di cui si servono gli EQ parametrici ossia i Low-Shelving(LS) e gli High-Shelving(HS), essi sono rispettivamente un HPF e un LPF particolari. La particolarità sta nel fatto che la funzione che regola il funzionamento di questi filtri vale zero per un intervallo maggiore rispetto a quella dei normali filtri, in questa maniera la banda non nulla sulla quale essi si concentrano è più ristretta, ossia si concentra su frequenze molto specifiche. Il risultato di quanto descritto è logico essere un'esaltazione maggiore delle frequenze passanti per questi intervalli di frequenza, rispetto tutte le altre dell'audio. Il sistema di esaltazione appena descritto necessita di 3 parametri: la frequenza che si vuole esaltare; la pendenza che deve avere il filtro; il valore in dB che esso deve coprire (ossia la sua ampiezza).

Ingenuamente si potrebbe pensare che gli EQ parametrici siano più limitanti di quelli grafici dato che non dispongono di un filtro per passare le bande di frequenze; in realtà così non è perché anche questa tipologia di equalizzatori possiede il suo filtro passa-banda detto **filtro di peaking** (dall'inglese: peak -> cima/picco).

Il filtro di peaking è caratterizzato dall'avere una funzione a campana (ossia una curva gaussiana), infatti il "picco" dal quale esso prende il nome è proprio la cima di questa. Per caratterizzare il filtraggio eseguito da questo dispositivo, vengono passati 3 parametri all'equalizzatore: una frequenza ( $f_0$ ) in corrispondenza della quale si individua il picco della funzione, questa frequenza è infatti al centro perfetto della banda di frequenza sulla quale questo filtro opera; un'ampiezza in dB che definisce la banda di azione del filtro; un coefficiente (o parametro come si vedrà fra poco) detto "bandwidth" (che si indicherà con Q).

Q è l'**indice di curtosì**, è esso che definisce quanto è appuntito il picco della campana -> per valori di Q molto alti la curva è maggiormente appuntita, il che significa che generalmente la funzione, più sale verso il suo picco più va restringendosi; per un Q piccolo invece si hanno gli effetti opposti. Ovviamente, più la curva è stretta e a punta, più sarà possibile focalizzarsi su valori di frequenze precise.

Curva alta, stretta e a punta ( $Q \gg 1$ ): curva lepicurtica  
Curva bassa, larga e piatta ( $Q \ll 1$ ): curva platocurtica

Come detto prima, Q può essere o costante (ossia si dà un valore di Q e quello rimane fino alla fine del processo) o variabile (ossia che Q varia in seguito agli avvenimenti del processo). Quando Q varia esso è detto "semi-parametrico", si osservino ora 2 casi di filtri di peaking particolari che seguono alla scelta di Q semi-parametrico:

-Q costante: Quando la banda di azione in dB del filtro varia, Q resta sempre costante -> quindi graficamente si avrà per qualsiasi ampiezza di banda una curva la cui forma della punta non varia mai, semplicemente la curva sarà più o meno alta a seconda dell'ampiezza di banda.

-Q proporzionale: Q varia a secondo una legge di proporzionalità con le variazioni dell'ampiezza in dB del filtro.

Nota conclusiva: è anche possibile ottenere filtri di peaking in cui le funzioni a campana non sono simmetriche, quindi il picco della campana sarà o più a destra o più a sinistra rispetto  $f_0$ .

I filtri di peaking la cui funzione è al di sopra dell'asse delle ascisse (funzione geometrica positiva) amplificano le frequenze che sono contenute nella loro banda di azione; i filtri di peaking la cui funzione è al di sotto dell'asse delle ascisse (funzione geometrica negativa) attenuano le frequenze che sono contenute nella loro banda di azione.

Nota: un equalizzatore parametrico non utilizza tutte le tipologie di filtri illustrate, si possono realizzare diversi tipi di equalizzatori parametrici la cui funzione fra a seconda dei filtri che si decide di implementare per il loro funzionamento (HPF & LPF; LS & HS; filtri di peaking).

### **Operazioni sul range dinamico**

Sono operazioni sulla modulazione dell'onda sonora eseguite dagli equalizzatori. Alcune delle operazioni possibili sono: normalizzazione, compressione, limitazione, espansione, noise gate.

La normalizzazione è un'operazione che amplifica una frequenza al massimo valore possibile che non introduce una distorsione digitale.

Queste operazioni sono fra loro diverse, ma hanno un modo di agire che le accomuna: esse variano il livello di uscita di un segnale sulla base di quello che è il livello di entrata. La legge che regola le trasformazioni che avvengono dall'ingresso all'uscita del segnale nel e dal dispositivo, è detta funzione di trasferimento, essa è una legge che sulle ascisse riporta le ampiezze del segnale in entrata e sulle ordinate quelle del segnale in uscita.

Il compressore è uno strumento che riduce la differenza di range dinamico nella traccia audio. esempio: in una traccia ci sono musica (un sottofondo) e un cantato (voce del cantante solitamente in sovrapposizione); il range dinamico del cantato deve essere maggiore di quello della musica di sottofondo affinché essa sia più distinguibile/pronunciata rispetto la musica di accompagnamento, il compressore si occupa infatti di produrre questo risultato.

Attenzione! -> in seguito ad un'operazione di compressione, è buona norma eseguirne una di normalizzazione, questa infatti aggiusta meglio le ampiezze dell'audio in uscita dal compressore.

Il valore di fondo scala è il range dinamico al quale ci si è fino ad ora implicitamente riferiti.

Il compressore è uno strumento che definisce una soglia, questa è individuata da un valore preso sull'asse delle ordinate che, non a caso, è quella sulla quale sono rappresentati i valori in uscita del segnale (è stato detto prima quando si è parlato della funzione di trasferimento). In pratica, la logica dietro questo dispositivo, è che tutti i valori in un uscita da esso che sono al di sotto di questa soglia (rappresentata graficamente da una retta costante che parte da un punto dell'asse delle ordinate e che è parallela a quella delle ascisse), vengano amplificati così da risultare più marcati, mentre i valori che invece sono al di sopra della soglia vengono attenuati.

Di quanto i valori target vengano amplificati/attenuati, è una specifica che dipende dal compressore in utilizzo. Il risultato di questa operazione è quello già anticipato, ossia di rendere i suoni di sottofondo, il cui range dinamico è più basso, più in risalto nella traccia audio in uscita rispetto a quelli che prima dell'operazione di compressione erano i suoni principali che sono ora stati attenuati per lasciare spazio a quelli amplificati.

Generalmente il compressore viene infatti utilizzato per dare risalto a quei suoni che sono di sottofondo nella traccia audio.

In audacity il compressore prima porta tutti i valori alla soglia, poi ivi attua le amplificazioni di ampiezza.

Il limitatore è un'operatore dinamico che svolge la limitazione; questa operazione si serve, come la compressione, di una soglia. A differenza dell'operazione vista in precedenza, solo le ampiezze del suono in entrata che superano questa soglia vengono modificate, mentre quelle che sono al di sotto di essa restano invariate. Le ampiezze che superano la soglia imposta dal limitatore vengono abbassate al di sotto di essa.

Solitamente, ma non è obbligatorio, prima di un'operazione di limitazione se ne esegue una di amplificazione/ normalizzazione.

Un compressore che ha una pendenza pari a "10 a 1" si comporta come un limitatore.

L'espansore è l'operatore opposto al compressore, esso fa sì che tutti i valori al di sopra della sua soglia vengano aumentati mentre tutti i valori al di sotto di essa vengono diminuiti (sempre in termini di ampiezza dell'onda). In questa maniera viene reso più chiaro il suono principale mentre quelli di sottofondo vengono ulteriormente attenuati.

Il limita rumore (noise gate) è l'opposto del limitatore, esso pone a zero i valori di ampiezza che superano la soglia. Questo operatore si può realizzare con un compressore di pendenza tendente ad infinito (ossia la retta è praticamente verticale).

Di seguito si mostri un immagine che racchiude tutte le possibili operazioni eseguite dai diversi operatori di range dinamico:

