



AUDIO PROCESSING

*UNIVERSITA' DEGLI STUDI DI CATANIA
DIPARTIMENTO DI MATEMATICA E INFORMATICA
LAUREA TRIENNALE IN INFORMATICA
A.A. 2021/22
Prof. Filippo L.M. Milotta*

ID PROGETTO: 0C

TITOLO PROGETTO: Audio Engineer

AUTORE 1: Francesco Romano

AUTORE 2: Dario Rovito

AUTORE 3: Mauro Mancari

Indice

1. Obiettivi del progetto.....	2
2. Riferimenti Bibliografici	8
3. Argomenti Teorici Trattati	Errore. Il segnalibro non è definito.

1. Obiettivi del progetto

1.1 L'ingegnere del suono

L'ingegnere del suono (anche chiamato tecnico del suono o fonico) può lavorare in diversi ambiti quali quello dello spettacolo e della musica, ma anche teatro, cinema, televisione, concerti e conferenze. Si occupa di migliorare la qualità del suono attraverso algoritmi di elaborazione del segnale audio in modo da farlo arrivare alle orecchie del pubblico nella migliore forma possibile migliorando e bilanciando le sorgenti sonore tramite effetti audio; infatti può anche gestire e allestire gli strumenti necessari, come il posizionamento dei microfoni, per registrare ed elaborare i suoni.

Un buon tecnico del suono deve essere in grado di interpretare la visione del musicista o del regista per mettere in risalto, attraverso i suoni, quello che vuole comunicare al pubblico.

In Italia, attualmente, esistono pochi corsi di laurea dedicati alla formazione di questa figura professionale. Uno di questi è il corso di laurea magistrale (in inglese) in Music and acoustic engineering del Politecnico di Milano (sede di Cremona). Sono presenti anche alcuni master universitari, come quello in Ingegneria del Suono e dello spettacolo dell'Università di Roma "Tor Vergata", o il Corso di perfezionamento in Ingegneria del suono dell'università di Modena e Reggio Emilia.

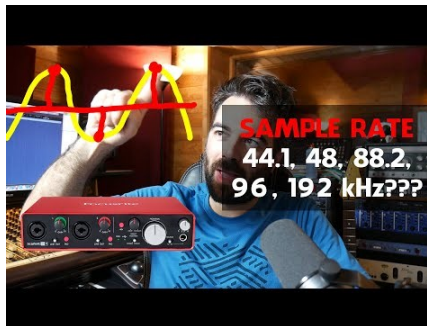


Di seguito verranno presentate le problematiche che l'audio engineer si ritrova a risolvere nel proprio lavoro, le strumentazioni più comuni e infine i formati audio utilizzati:

1.2 Frequenza di campionamento

Cos'è il sample rate? (0:02 – 3:34)

https://youtu.be/oaqCLMLx_9A



Il sample rate (la frequenza di campionamento) di un file audio rappresenta il numero di volte in cui un segnale analogico viene analizzato in un secondo, allo scopo di poterlo trasformare in un file audio vero e proprio.

Conoscere la frequenza di campionamento è importante non solo in fase di registrazione, ma anche sul lavoro di un file audio.

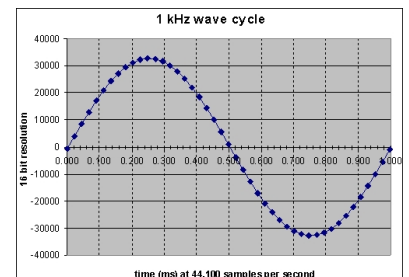
Qual è la Frequenza di Campionamento *Migliore*?

Intuitivamente la risposta sarebbe che la miglior frequenza è la più alta che si possa scegliere, così da avere un file composto da un maggior numero di misurazioni e quindi più definito. In realtà non è proprio così, infatti aumentando troppo le frequenze il computer viene sottoposto ad uno sforzo maggiore occupando più spazio, quindi basta superare un certo valore ed essa verrà considerata accettabile.

Questo valore è definito dal “teorema di Nyquist-Shannon”, esso definisce la minima frequenza, detta **frequenza di Nyquist** (o anche *cadenza di Nyquist*), necessaria per campionare un segnale analogico senza perdere informazioni, e per poter quindi ricostruire il segnale analogico tempo continuo originario, affermando che, data una funzione la cui trasformata di Fourier sia nulla al di fuori di un certo intervallo di frequenze (ovvero un segnale a banda limitata), nella sua conversione analogico-digitale la minima frequenza di campionamento necessaria per evitare aliasing e perdita di informazione nella ricostruzione del segnale analogico originario (ovvero nella riconversione digitale-analogica) deve essere maggiore del doppio della sua frequenza massima mediante questa formula:

$$s(t) = \sum_{k=0}^{+\infty} s(k\Delta t) \operatorname{sinc}\left(\frac{t}{\Delta t} - k\right) \quad \forall t \in \mathbb{R}$$

In definitiva quindi per codificare correttamente in digitale un segnale audio analogico, quest'ultimo dev'essere campionato con una frequenza di campionamento almeno doppia rispetto alla frequenza più alta contenuta nel segnale stesso; con frequenza intendiamo in termini pratici nella percezione che noi abbiamo dell'*altezza* di quel suono, infatti un segnale audio ad alta frequenza lo percepiamo come un *suono acuto*, un segnale audio a bassa frequenza lo percepiamo come un *suono grave*.



1.3 Formati Audio

Prima di parlare dei formati audio è necessario fare una piccola introduzione sull'audio digitale e sulla compressione.

L'audio digitale nasce per la prima volta all'inizio degli anni '80 con l'introduzione del Compact Disc. Esso con il passare del tempo, prese sempre più piede nell'ambiente della produzione e della post-produzione audio. Gli indiscutibili vantaggi per quanto riguarda la banda passante ed il rapporto segnale/rumore, hanno fatto sì che esso giungesse con una certa rapidità all'occupazione del vertice delle tecnologie audio. Attualmente l'audio digitale viene utilizzato in tutta la catena del segnale, sia da quando un suono entra all'interno del mixer, fino alla sua trasmissione.

Un segnale audio digitale viene ottenuto dal segnale analogico attraverso due operazioni che sono fondamentali, ovvero, il campionamento e la quantizzazione. Tale conversione viene effettuata tramite un convertitore analogico/digitale (A/D) che tiene anch'esso conto di due parametri fondamentali, **la frequenza di campionamento e i bit per campione**.

La quantizzazione, consiste nel determinare il valore di ogni campione. L'intervallo di variabilità del segnale, la cui ampiezza è determinata dalla frequenza di campionamento, viene diviso in intervalli uguali. Il numero di questi intervalli è determinato dal numero dei bit che si utilizzano per codificare ogni campione.

Il numero di bit utilizzati per la quantizzazione dipende essenzialmente da due fattori:

- La disponibilità e il costo di convertitori A/D adatti a trattare parole di questa lunghezza;
- La quantità di memoria occupata dopo il campionamento.

Per ridurre il rumore termico, tutta l'apparecchiatura (e anche gli ascoltatori) dovrebbero essere mantenuti a una temperatura prossima allo zero assoluto. È evidente, che sia il processo di campionamento che quello di quantizzazione introducono degli errori, rivestendo un'importanza fondamentale nella qualità del segnale finale. Questo concetto, si può sintetizzare nei seguenti punti:

- Più alta è la frequenza di quantizzazione, minore è la potenza totale del rumore digitale;
- Più alta è la frequenza di campionamento, più il rumore si distribuisce alle alte frequenze. Il rumore che si trova fuori dalla banda audio può essere eliminato in un secondo tempo ed è influente durante l'ascolto;
- In genere il rumore digitale è correlato con il segnale.

Con la tecnologia digitale è relativamente facile trovare soluzioni economiche a questi problemi.

Una volta ottenuto il segnale digitale questo potrà essere memorizzato ed eventualmente elaborato, ma sarà necessario trasformare nuovamente il segnale digitale in analogico prima della fase di amplificazione di potenza (questa procedura prende il nome di **conversione D/A**).

Prima di concludere questa piccola introduzione sull'audio digitale, possiamo sintetizzare quanto detto dicendo che, elevate frequenze di campionamento, anche se non servono a riprodurre segnali con frequenze non udibili dall'orecchio umano, consentono di ridurre le interferenze di tipo matematico nelle fasi di campionamento, quantizzazione e ricostruzione, nonché di semplificare i circuiti di conversione.

COMPRESSIONE

La compressione è la riduzione della gamma dinamica, la differenza tra le parti più rumorose e più silenziose di un segnale audio. Quando viene applicata la compressione, le parti più 'calme' del segnale vengono aumentate e quelle più alte vengono attenuate. Inoltre la compressione riduce la gamma dinamica di un segnale, ovvero riduce la differenza tra le parti di un segnale.

I compressori riducono il guadagno del segnale, abbassano il volume dei picchi rumorosi ed inoltre, pareggiano anche le note che escono dal mixer.

Idealmente la compressione suona più aggressiva e più serrata. Mentre analizzando la sovra-compressione, possiamo dire che essa suona distorta, noiosa e rumorosa.

Il pregio dei compressori è che ciascuno di essi è unico ed aggiungerà un colore diverso al suono. Quindi possiamo dire che, se utilizziamo i compressori giusti si ricava un suono più scintillante e pungente senza andare in clip. Consente infatti, di ottenere un maggiore volume generale tenendo sotto controllo i picchi, riducendo così, la gamma dinamica di alcuni suoni e consentendo di bilanciare i livelli tra gli strumenti in un mix, in maniera molto più agevole.

Esistono due tipi di compressione:

- Con perdita(**Lossy**): ovvero, quando l'informazione nel file compresso è minore di quella contenuta nel file di origine. Essa permette compressioni maggiori ma a scapito della qualità sonora;
- Senza perdita(**Lossless**): ovvero, quando l'informazione contenuta nel file complesso è identica a quella contenuta nel file di origine.

FORMATI AUDIO

In questo periodo storico in cui la musica digitale ha avuto una diffusione globale, sono stati creati numerosi formati audio. Ecco qui un elenco dei formati audio più diffusi in tutto il globo.

Formati audio della famiglia Lossy

AAC (Advanced Audio Coding o MPEG-4): È lo standard di default utilizzato da Apple per iTunes. Infatti quando si voleva importare un CD (in formato MP3) all'interno di un Mac, in automatico il formato MP3 veniva convertito in AAC. A parità di bitrate occupa lo stesso spazio di un MP3 ma, la conversione può risultare superiore.

MP3 (Acronimo di Moving Picture Expert Group- 1/2 Audio Layer 3): È il formato audio compresso più utilizzato al mondo ed è stato introdotto nel 1998. Il suo algoritmo di compressione va a rimuovere alcuni dettagli della traccia audio che possono essere difficilmente ascoltati dall'orecchio umano.

OGG Vorbis: È un formato Open Source di qualità paragonabile all'MP3, sconosciuto fino a poco tempo fa ma, utilizzato ora da servizi di musica streaming come Spotify. OGG è l'estensione del file mentre, Vorbis è l'algoritmo di compressione.

WMA (Windows Media Audio/Player): È il formato audio inventato da Microsoft in risposta allo standard MP3, la qualità è la medesima ma, ha lo svantaggio di essere supportato solamente dai dispositivi Microsoft.

Formati audio della famiglia Lossless

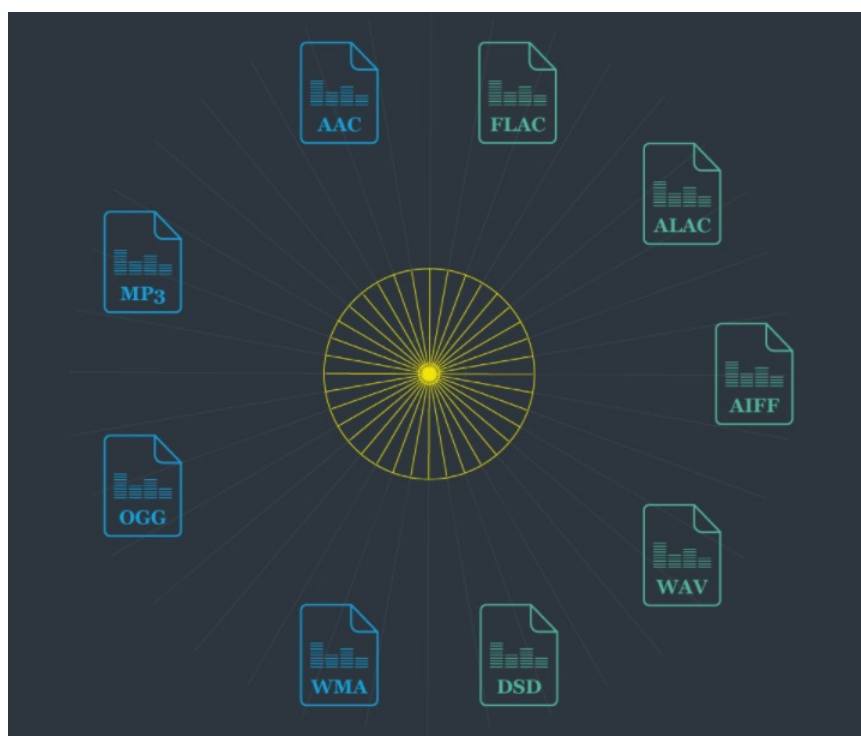
FLAC (Free Lossless Codec): È il formato open source più utilizzato per il music download senza perdite di qualità rispetto alla sorgente originale.

ALAC (Apple Lossless Audio Codec): Possiamo definirlo come il formato FLAC del mondo Apple, anche se in termini di peso non equipara il FLAC. È ideale solo se si utilizzano dispositivi Apple.

AIFF (Audio Interchange File Format): Anch'esso fa parte della famiglia Apple, ed è stato sviluppato basandosi sull'Interchange File Format della Electronic Arts ed è adatto particolarmente agli audio filati e a chi produce musica. Presenta essenzialmente le medesime caratteristiche del formato WAV ma senza compressione ed è utilizzato dagli artisti perché consente di embeddare all'interno del file anche dei metadati come testi, note o altre informazioni.

WAV (WAVE, Form Audio File Format): È stato introdotto dalla Microsoft e da IBM nel 1991 ed è ancora in uso. Rappresenta ciò che si ottiene quando si importa un CD musicale nel proprio computer Microsoft. Sono file molto pesanti che arrivano al massimo a 2GB e riproducono il suono in maniera fedele.

DSD (Direct Stream Digital): È un marchio registrato Philips e Sony sviluppato per il Super Audio CD. Utilizza un metodo di codifica che anziché prevedere profondità in bit maggiori, utilizza un solo bit ma a una frequenza di campionamento molto più elevata per registrare, archiviare e riprodurre una qualità del suono estremamente alta.



1.4 Il mixer

Il mixer avrebbe il compito di *miscelare* i segnali provenienti da varie fonti fino a farli confluire in un unico amplificatore. Quasi tutti i mixer sono dotati di un alimentatore *esterno*, per allontanare fisicamente tutti i circuiti che lavorano con corrente alternata dal mixer. I segnali provenienti dai vari ingressi confluiscono in vari modi nei bus di miscelazione, ovvero, dove avviene la miscelazione vera e propria, e in seguito si immettono nella sezione di uscita. Ogni circuito dei canali di ingresso fa capo a determinati controlli che possono apportare modifiche al segnale audio, oppure possono determinare le modalità con cui avviene la miscelazione.

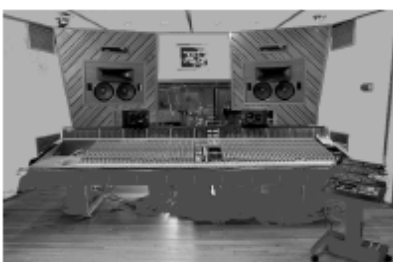
All'interno del mixer possiamo trovare i cosiddetti filtri di equalizzazione. In generale uno stadio di equalizzazione ha il compito di modificare la risposta in frequenza di un segnale audio. Possiamo in pratica far sì che una determinata gamma di frequenze di un programma sonoro risulti esaltata o attenuata. Tale operazione permette di variare la composizione delle armoniche e, entro certi limiti, il timbro del suono. Per comporre un programma musicale, risulta particolarmente importante poter intervenire sul timbro dei singoli suoni, prima di procedere al missaggio.

Lo stadio di equalizzazione di un canale di ingresso di un mixer è composto da un certo numero di filtri, di solito da due a quattro, che agiscono sul suono in maniera indipendente. Tali filtri possono essere di due tipi: grafici o parametrici:

- I filtri grafici agiscono su una frequenza fissa e produrranno un'esaltazione o un'attenuazione in una banda incentrata su quella frequenza.
- I filtri parametrici agiscono su una frequenza variabile, quindi la banda sulla quale si produrranno gli effetti di esaltazione e attenuazione è anche essa variabile. Nei filtri parametrici più completi possiamo variare anche la pendenza del filtro, ossia in un certo senso l'ampiezza della banda dove avviene l'esaltazione o l'attenuazione.

Esistono numerosi tipi di Mixer, ognuno dei quali viene impiegato in ruoli ed ambienti diversi in base a ciò che si vuole fare. Ecco qui alcuni dei seguenti tipi di mixer:

Mixer da studio



Gestisce sia i segnali provenienti dai microfoni, sia i *registratori multitraccia*, in entrata e in uscita. Per questo motivo gran parte dei mixer da studio sono del tipo *on line*. In questo tipo di mixer i canali possono essere di volta in volta configurati come canali di ingresso o di uscita, a seconda della fase della lavorazione.

Mixer per radio-tv (broadcast mixer)



Nelle regie audio delle radio e delle televisioni troviamo dei mixer abbastanza simili a quelli per impiego live, con l'aggiunta di dispositivi specifici, quali ingressi per linee telefoniche (telco), relais telecomandati per abilitare i canali da parte dello speaker, timer per la scansione dei tempi radiofonici.

Mixer per discoteca:



Nei mixer da discoteca è attivo solo un canale per volta, in quanto è necessario passare dalla musica riprodotta da un giradischi alla musica di un altro giradischi, senza creare vuoti di silenzio. In alcuni mixer da discoteca, fra cui quello raffigurato, esiste un particolare controllo denominato **crossfader**. Si tratta di un *fader* posto in posizione orizzontale che permette con un solo movimento di *passare* da un canale all'altro, effettuando una *dissolvenza incrociata*.

Mixer digitali:



In pratica il segnale analogico viene convertito in digitale, quindi in una serie di numeri, appena entrato nel mixer. Tutte le elaborazioni che il mixer compie, quali il controllo del livello, l'equalizzazione, ecc. avvengono tramite operazioni matematiche sulla serie dei numeri. I vantaggi sono quelli del digitale e sono facilmente intuibili: pulizia del suono, assenza di rumori indotti, interfacciamento diretto ad altre apparecchiature digitali.

2. Riferimenti Bibliografici

- <https://claudiomeloni.it/frequenza-di-campionamento-file-audio/>
- https://it.wikipedia.org/wiki/Teorema_del_campionamento_di_Nyquist-Shannon
- https://en.wikipedia.org/wiki/Audio_engineer
- <http://www.fispinacademy.it/blog/277-il-mixer-parte-i-descrizione-generale.html>
- https://it.wikipedia.org/wiki/Compressione_audio_digitale
- https://youtu.be/oaqCLMLx_9A

3. Argomenti Teorici Trattati

3.1 Analisi di Fourier

L'analisi di Fourier è uno strumento matematico che permette di descrivere un segnale complesso come somma di segnali più elementari, in particolare onde sinusoidali e cosinusoidali. Essa permette di valutare ampiezza, fase e andamento sia nel dominio del tempo che della frequenza.

L'analisi si basa sul teorema di Fourier:

Una funzione periodica, che soddisfi le condizioni di Dirichlet, di periodo T o di frequenza fondamentale $f_1=1/T$, viene rappresentata tramite somma di onde sinusoidali e/o cosinusoidali di opportuna ampiezza e di frequenza multipla delle frequenza fondamentale.

Una funzione periodica di periodo T può essere rappresentata come:

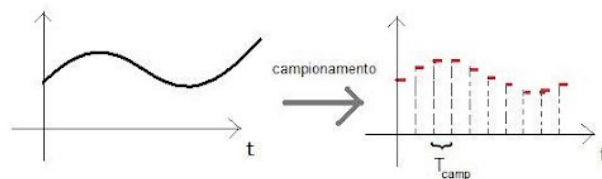
$$y(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n \frac{2\pi}{T} t) + b_n \sin(n \frac{2\pi}{T} t)$$

Dove:

- n è un numero naturale e $2\pi/T = 2\pi f = \omega$;
- l'espressione $a_n \cos(n\omega t) + b_n \sin(n\omega t)$ si chiama **n -esima armonica**;
- I termini a_n e b_n sono i coefficienti dell' n -esima armonica;
- L'armonica ottenuta per $n=1$ si chiama **armonica fondamentale** ed ha frequenza pari a quella dell'onda.

3.2.1 Frequenza di campionamento

- **Campionamento** — consiste nel convertire un **segnale** continuo nel tempo in un **segnale discreto**, valutandone l'ampiezza a intervalli di tempo regolari, si considerano solo alcuni valori del segnale tra loro equidistanti.



In particolare La frequenza di campionamento indica il numero di istantanee al secondo prelevate da un segnale audio. Questa frequenza determina l'intervallo di frequenza di un file audio. Maggiore è la frequenza di campionamento, più la forma della forma d'onda digitale sarà vicina a quella della forma d'onda analogica originale. Basse frequenze di campionamento limitano l'intervallo di frequenze che è possibile registrare, cosa che a sua volta può generare una registrazione che riproduce male il suono originale.

- Teorema del campionamento di Nyquist-Shannon

Definisce la minima frequenza, detta **frequenza di Nyquist** (o anche *cadenza di Nyquist*), necessaria per **campionare** un **segnale** analogico senza perdere informazioni, e per poter quindi ricostruire il segnale analogico tempo continuo originario. il teorema afferma che, data una funzione la cui **trasformata di**

Fourier sia nulla al di fuori di un certo intervallo di frequenze (ovvero un segnale a **banda limitata**), nella sua **conversione analogico-digitale** la minima **frequenza di campionamento** necessaria per evitare **aliasing** e perdita di informazione nella ricostruzione del segnale analogico originario (ovvero nella riconversione digitale-analogica) deve essere maggiore del doppio della sua frequenza massima .

- $f_c > 2f_n$ f_c : tasso di campionamento , ci riferiamo al numero di campioni per ogni secondo
 - L'aliasing deriva da una sovrapposizione delle più alte frequenze dello spettro originale , con le frequenze introdotte nel segnale campionato, diventando a volte indistinguibili, creando così un serio problema che si riflette direttamente sull'uscita del sistema.

3.2.2 COMPRESSIONE

La compressione audio è una tecnica di elaborazione dati, attuata a mezzo di un codec audio, che permette di ridurre le dimensioni (anche di molto) di un file audio o la banda passante richiesta per una trasmissione audio su un canale di comunicazione.

La **compressione dati lossy** (dall'inglese **loss**: perdita), individua una classe di algoritmi di compressione dati che porta alla perdita di parte dell'**informazione** originale durante la fase di compressione/decompressione.

Decomprimendo un file compresso con un metodo *lossy* la copia ottenuta sarà peggiore dell'originale per livello di precisione delle informazioni che codifica, ma in genere comunque abbastanza simile da non comportare perdita di informazioni irrinunciabili

La **compressione dati senza perdita** (o compressione dati **lossless**), in informatica e telecomunicazioni, è una classe di algoritmi di compressione dati che non porta alla perdita di alcuna parte dell'informazione originale durante la fase di compressione/decompressione dei dati stessi.

3.3 Formati audio

Uno degli standard più utilizzati è sicuramente il **MPEG** (Motion Picture-Coding Experts Group), inizializzato nel 1988 dalla **ISO/IEC** (**ISO**: International Organization for Standards; **IEC**: International Electrotechnical Commission). Il suo compito è quello di compressione, decompressione, elaborazione e codifica di video, audio e contenuti multimediali in generale.

Esistono diversi tipi di MPEG:

- **MPEG-1** (1992);
- **MPEG-2** (1994);
- **MPEG-3** (assorbito nel 2);
- **MPEG-4** (1999);
- **MPEG-7** (2001, standard per la ricerca, il filtraggio e la gestione delle informazioni, ma non della codifica come i precedenti. Insieme a MPEG-4 viene spesso denominato **MPEG-47** per codifica e descrizione);

- **MPEG-21** (2001, standard per la definizione di un framework per lo sviluppo di applicazioni multimediali);
- **MPEG-D** (2007);