AudioLand 2015

(Appunti su nozioni basilari del suono analogico e digitale)

a cura di Efrem Castellano

Indice :

Il suono e le sue caratteristiche …….. pag. 1

Esempi matematici su frequenza e lunghezza d'onda ……... pag. 4

Fase e tipi di fase…….…… pag. 4

Applicazione pratica della controfase …… pag. 7

Distinzione tra suono e rumore …….. pag. 8

Suoni semplici e suono composti ……... pag. 8

Tipi di forme d'onda …….. pag. 9

Aspetti della frequenza ……... pag. 11

Inviluppo …….……... pag. 13

Il "movimento del suono" …….. pag. 14

Suono analogico e suono digitale …….. pag. 14

Convertitori ADC - DCA e frequenza di campionamento …… pag. 17

Dimensione di un file audio digitale (calcolo matematico) …..…..…….. pag. 18

Formati e compressioni audio …… pag. 19

Tecniche di esportazione e modifica dell'audio ……..……. pag. 24

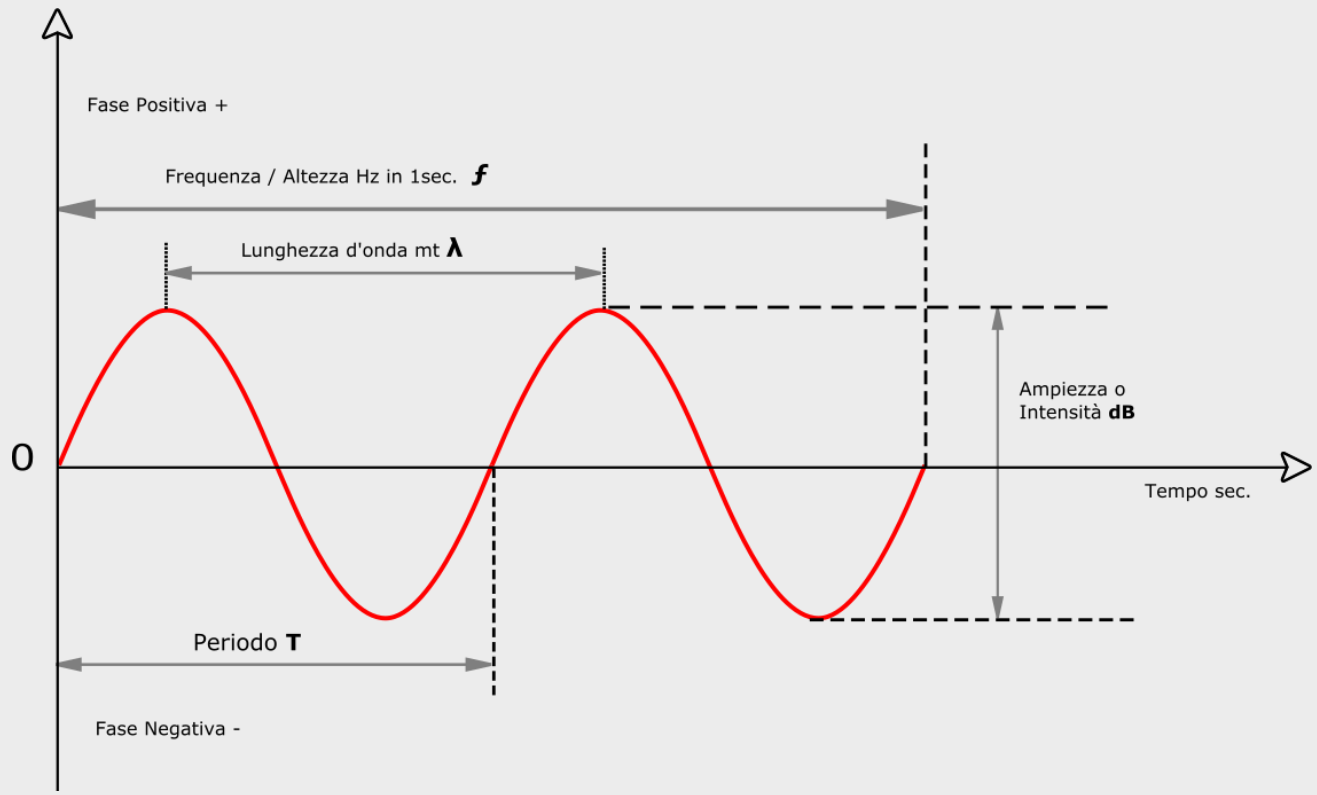
Tabella di riferimento file audio (.wav) …….…… pag. 26

Tabella note e frequenze 440 & 432 Hz ….……pag. 27

Il suono e le sue caratteristiche

L'audio che percepiamo (sotto forma di musica e rumori) è definito: "suono". Il suono è la vibrazione (regolare o costante) di un corpo che oscilla nell'aria che, propagandosi, è percepita appunto dall'apparato uditivo di esseri umani e animali. Nel "vuoto" non c'è né suono, né rumore.
L'oscillazione costante di tale corpo, è rappresentata come un'onda sinusoidale che è poi la base (o punto di partenza) per rappresentare graficamente tutte le onde sonore.
Anche il suono è matematica: ha le sue grandezze e caratteristiche che sono rappresentate in lettere e numeri:

- **Ampiezza** o **Intensità** o detto anche volume si misura in **dB** (decibel)
- **Altezza** o **Frequenza** si misura in **Hz** f
- **Periodo** rappresentato con la lettera **T** e si misura in **sec.**
- **Lunghezza d'onda** si misura in metri **(mt)** λ (Lambda)
- **Fase** φ (Phi)
- **Velocità** si misura con **c** = velocità del suono nell'aria (**344 m/sec** alla temperatura di 20 °C ed alla pressione atmosferica del livello del mare).
- **Timbro** qualità che, a parità di frequenza, distingue un suono da un altro.

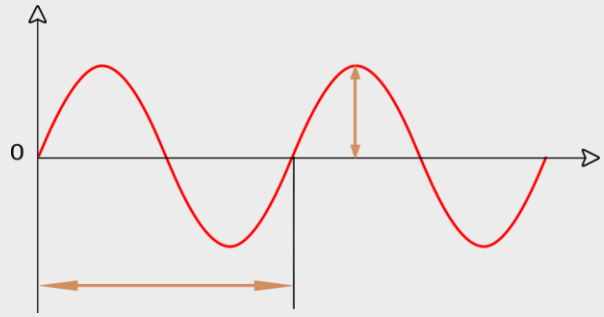


(Fig. 1)

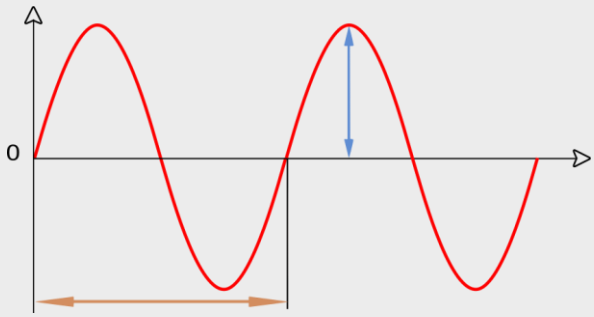
In questa figura è rappresentata un'onda sinusoidale semplice. Sull'asse delle ascisse (x) è rappresentato il tempo. Su quello delle ordinate (y) lo spazio. Sulla linea dello 0, avremo una condizione di silenzio.

Ampiezza / Intensità:

L'intensità è quella caratteristica (grandezza) che permette di distinguere i suoni deboli da quelli forti.



suono meno intenso



suono più intenso

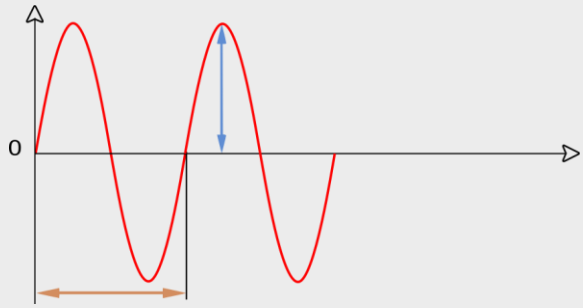
Qui varia l'ampiezza ma non il periodo.

Tra le caratteristiche dell'ampiezza/intensità del suono abbiamo:

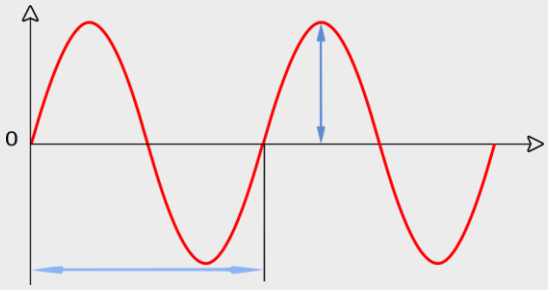
- dB (*Decibel*, decimo di bel (simbolo **B**): 10 dB = 1 B)
- RMS (*Root Mean Square*, cioè la radice della media dei quadrati o meglio detto *valore efficace*)
- SIL (*Sound Intensity Level*, cioè il livello di Intensità Sonora) SIL= 10 log (*I/I*o)
- SPL (*Sound Pressure Level*, cioè Livello di Pressione Sonora) SPL= 20 log (*p/p*o)
- SOL (*Standard Operating Level*, cioè Livello Standard di Operatività)
- SNR (*Signal To Noise Ratio*, cioè rapporto segnale-rumore)

Frequenza / Altezza

L'altezza o frequenza di un suono corrisponde al numero di vibrazioni complete che la sorgente compie in un secondo. Quindi, se abbiamo una frequenza di 30 Hz vuole dire che la nostra onda sonora compirà 30 periodi completi in un secondo.
Grazie alla frequenza possiamo distinguere i suoni acuti da quelli gravi.



Suono acuto



Suono grave

Qui varia il periodo ma non l'ampiezza.

Un periodo maggiore comporta una frequenza minore, un periodo minore comporta una frequenza maggiore.

$$f = 1/T$$

Periodo

Il periodo è il tempo che serve a un'onda sonora per compiere un ciclo completo. Per ciclo completo si intende un'oscillazione che parte da zero, fino a raggiungere i valori massimi di oscillazione, (sia in fase positiva, che negativa) per poi tornare nuovamente a zero.
Un periodo maggiore comporta una frequenza minore, un periodo minore comporta una frequenza maggiore.

$$T = 1 / f \quad \text{oppure} \quad T = \lambda / c$$

Lunghezza d'onda

Caratteristica che serve per calcolare in metri la distanza tra due punti massimi di una forma d'onda. Di norma si misura in metri.

$$\lambda = c / f \quad \text{oppure} \quad \lambda = c \times T$$

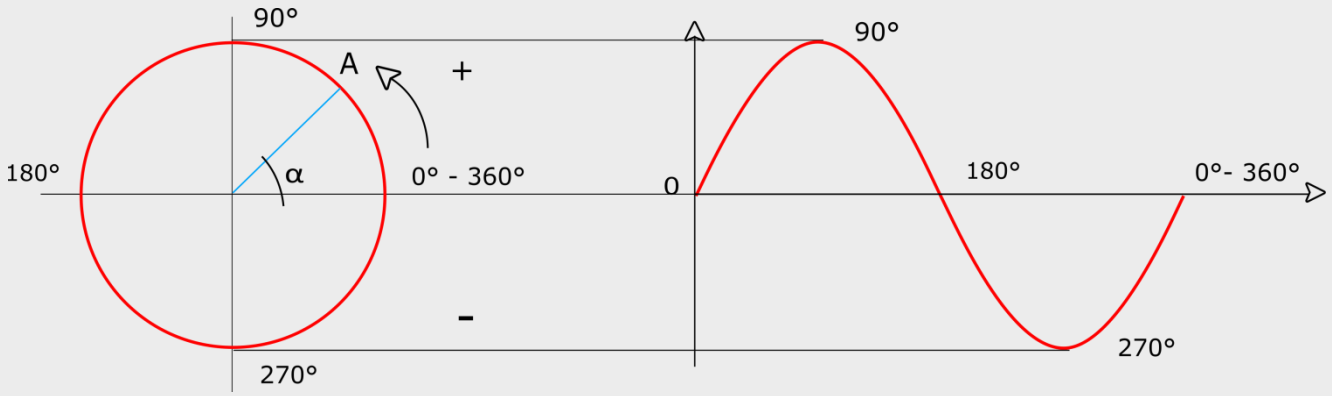
Esempi matematici su frequenza e lunghezza d'onda

Abbiamo un suono di 35Hz e vogliamo conoscere la sua lunghezza d'onda. Il calcolo da fare sarà il seguente:

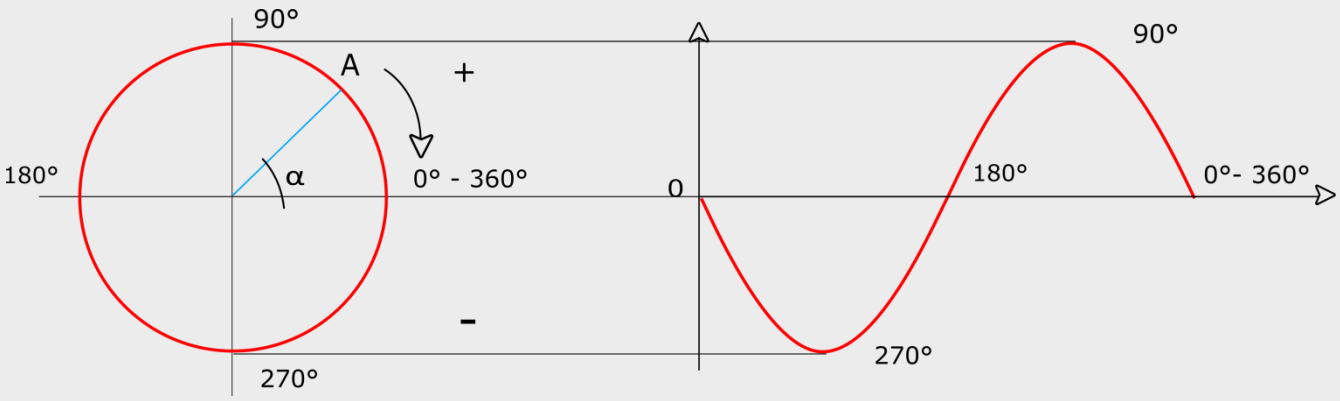
Dalla formula → $\lambda = c / f$ oppure $\lambda = c \times T$ abbiamo $c = 344$m/s ed $f = 35$Hz
Dunque: $\lambda = 344 / 35 =$ **9,828 mt** e il tempo? $T = 1 / f$ → $1 / 35 =$ **0,02857 sec.**
Per avere un'ulteriore conferma applichiamo ora la seconda formula per ottenere λ :
$\lambda = c \times T$ → $\lambda = 344 \times 0{,}0285 =$ **9,828 mt**

Fase e tipi di fase

La fase si misura in gradi sessagesimali e serve per individuare il punto esatto di lettura in un ciclo della forma d'onda. 0° è il punto iniziale, 90° è l'ampiezza massima, 180° è il punto dove la forma d'onda si interseca con l'asse delle ascisse, 270° è il picco negativo, 360° è la conclusione del ciclo.



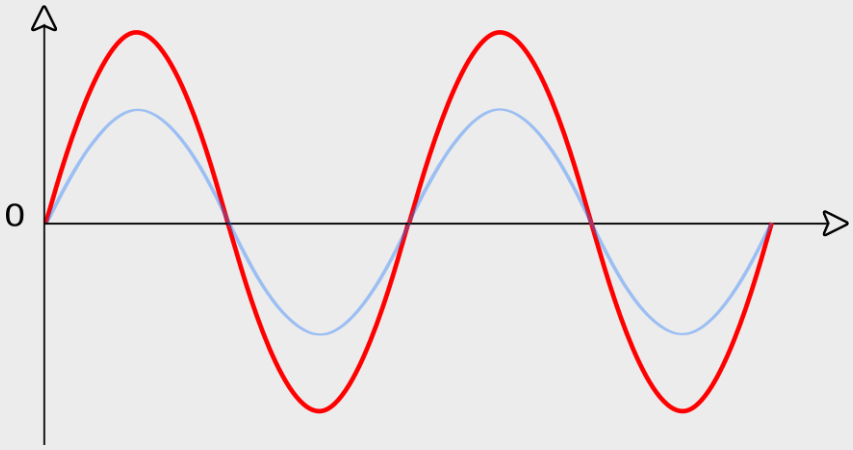
Nel caso invece che la proiezione del punto A (raggio blu) girasse in senso orario, ecco che avremo questo caso:



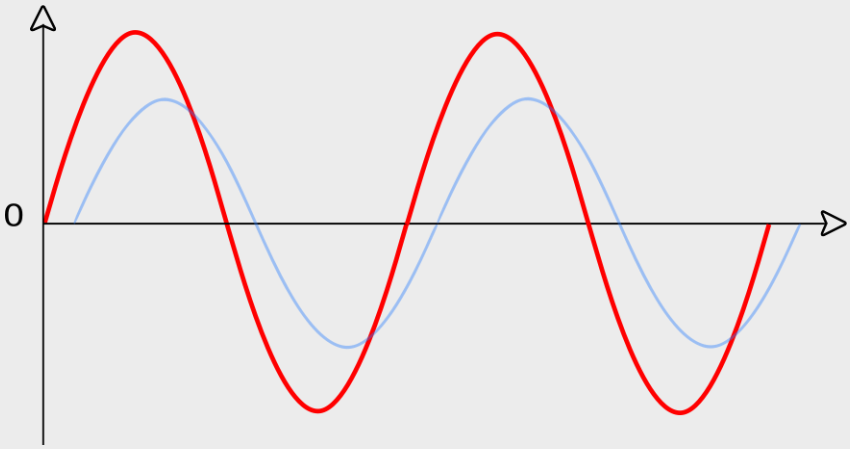
Ecco che da questa proiezione, avremo un'onda sinusoidale perfettamente in controfase. Fase e tempo sono legate tra di loro da un'equazione :

$$\varphi = 2\pi \ rad \times f \times \Delta t$$

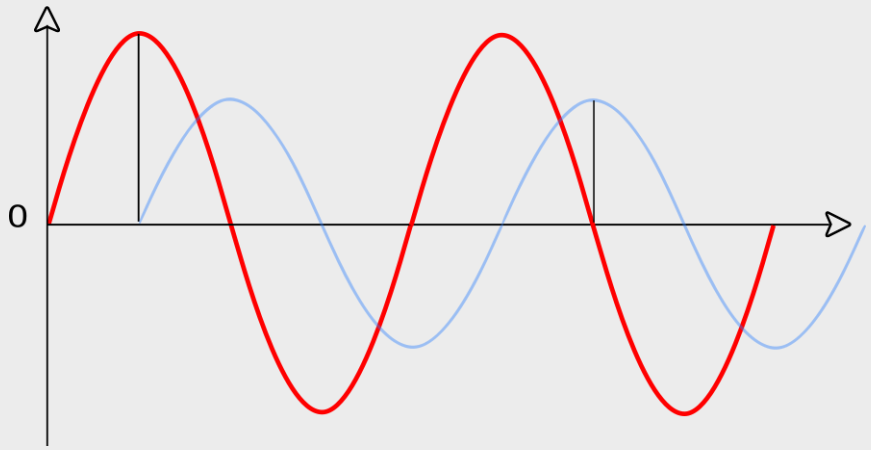
Essere in fase:
Due segnali sinusoidali con stessa frequenza ma diversa ampiezza si dicono in fase quando lo sfasamento $\Delta\varphi$ è pari a 0° e si dice che due segnali sinusoidali sono **in fase**: si noti come "creste" e "gole" delle onde siano allineate verticalmente (quindi sincrone).



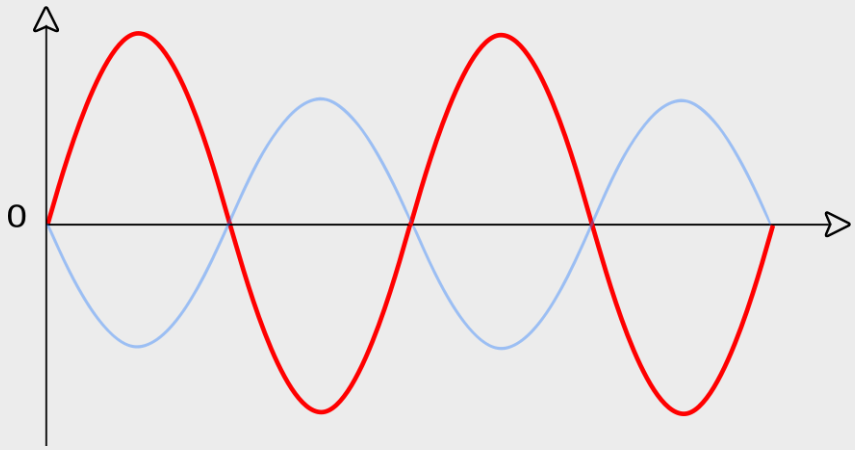
Sfasamento:
Quando si considerano due segnali sinusoidali aventi la stessa frequenza, ma diversa ampiezza, si può poi parlare di **differenza di fase** tra loro **$\Delta\varphi$** o **sfasamento**, intendendo con ciò, da un punto di vista matematico, la differenza tra le due costanti di fase.



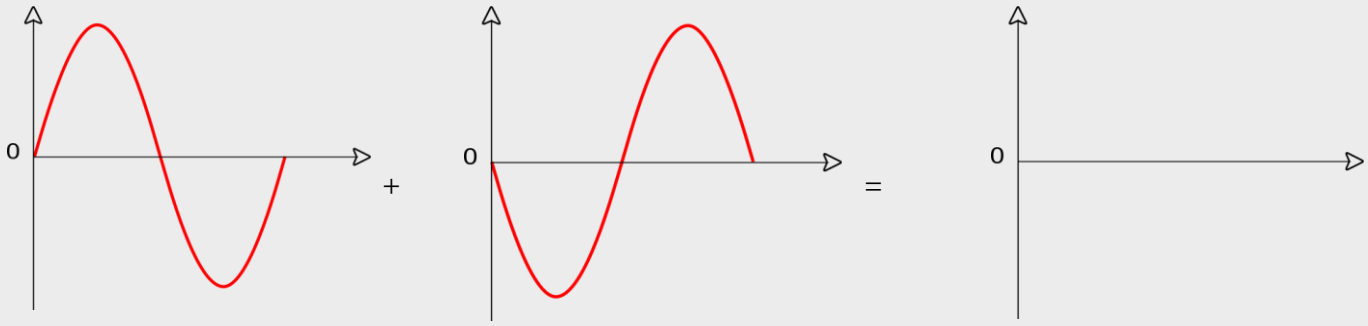
In quadratura:
lo sfasamento è pari a ± 90° e si dice che i segnali sono **in quadratura**: punti corrispondenti, ad es. le creste dei due segnali, sono spostati di un quarto di periodo.



Controfase:
lo sfasamento è pari a 180° e si dice che i segnali sono **in controfase**: le creste di un segnale sono allineate verticalmente con le gole dell'altro e viceversa.



Cancellazione di fase:
Due onde sinusoidali di frequenza e ampiezza identiche ma con fase invertita, se sommate, tenderanno ad annullarsi con conseguente ottenimento del silenzio.



Un'applicazione pratica della controfase:

A volte mettere un suono in controfase può servire a ricavarne un'altro. Come? Facciamo un esempio teorico. Ho un brano strumentale e vocale. Posso ottenere solo la parte vocale? Si se possiedo e metto in controfase la sua esatta riproduzione strumentale:

Brano strumentale e vocale + *Brano solo strumentale* = *Parte Vocale*
(FASE) (CONTROFASE)

Non è detto che si ottenga una parte vocale perfettamente pulita perché, certe frequenze di determinati strumenti musicali, potrebbero andare a miscelarsi molto bene con le frequenze della voce umana..

Velocità

Abbiamo detto che la velocità è una costante e più precisamente, quella del suono, cioè 344m/s alla temperatura di ± 20°.
La velocità è un elemento fondamentale che dà al suono svariati "comportamenti" e propagazioni..

Timbro

Il timbro è quella caratteristica sonora che ci permette di distinguere un suono da un altro pur se questi hanno stessa intensità e altezza.

Distinzione tra suono e rumore

Nella vita quotidiana di tutti i giorni, noi abbiamo a che fare con suoni e rumori. Negli ultimi decenni anche i rumori sono entrati a far parte della musica, tanto da definirsi "musicali". Pensando a certi generi musicali, molti rumori (opportunamente modificati) entrano far parte di un brano musicale..

Il suono è:

- Concreto o determinato
- Emette vibrazioni regolari / periodiche
- E' facilmente percepibile / distinguibile
- E' facilmente reperibile
- E' classificabile
- Spesso è gradevole all'ascolto

Il rumore è:

- Astratto o indeterminato
- Emette vibrazioni irregolari
- A volte è difficilmente percepibile / distinguibile
- A volte è difficilmente registrabile / reperibile / riproducibile
- E' inclassificabile
- A volte sgradevole all'ascolto

Grazie alla tecnologia digitale (campionatori e librerie sonore), oggi si possono reperire "rumori" con maggiore facilità.
Anche attraverso le varie tecniche di sintesi del suono è possibile ottenere "rumori strani".

Suoni semplici e suoni composti

L'onda sinusoidale oltre ad essere rappresentata graficamente può essere riprodotta in formato audio. Il diapason è l'unico strumento musicale "naturale" che riproduce un'onda sinusoidale pura. Si possono poi riprodurre onde o segnali sinusoidali tramite generatori sonori appositi o sintetizzatori analogici e digitali. Con una sola onda sinusoidale si fa ben poco di musicale, magari si ottiene un suono di basso o di pad e poco di più.. Ma cosa succede se prendiamo più onde sinusoidali e le suoniamo contemporaneamente ? Avremo un suono composto cioè formato da più suoni semplici. Dietro queste poche righe in realtà ci sarebbero da spiegare tutti gli studi di Jean Baptiste Fourier, che possono essere sintetizzati cosi:

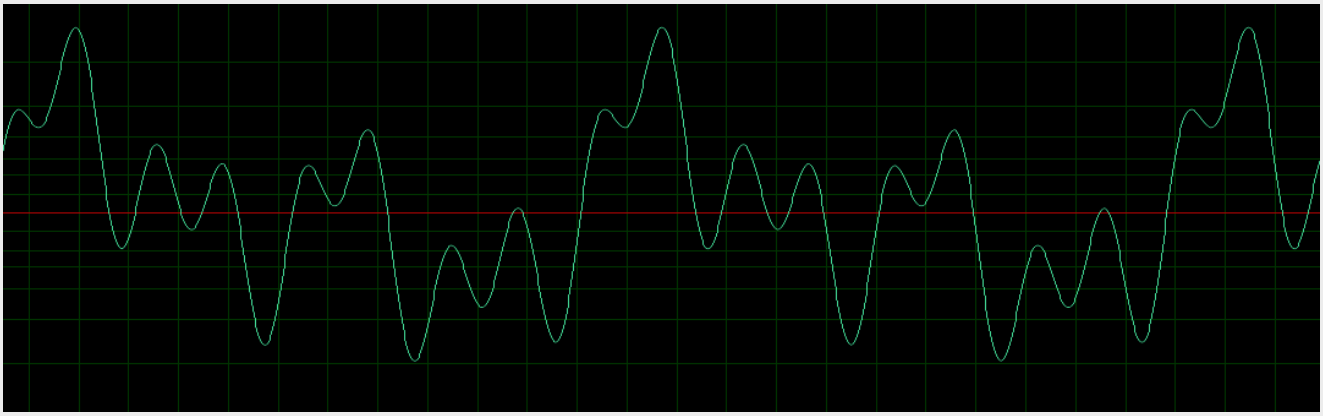
- *Un suono composto è costituito da una frequenza fondamentale più le sue armoniche.*

- *Un suono composto periodico è costituito da una somma di segnali sinusoidali semplici, ognuno con una sua propria ampiezza e fase.*

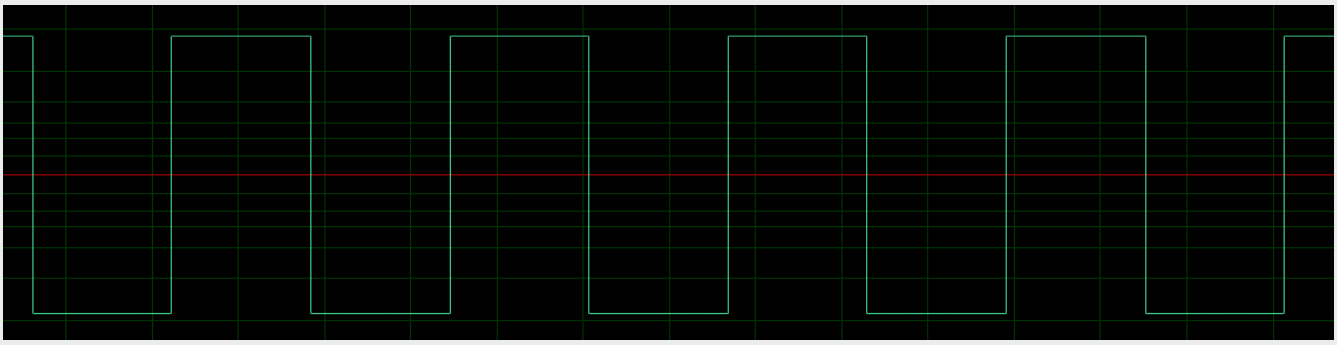
Grazie ai calcoli di Fourier, possiamo rappresentare graficamente (nei software digitali) le forme d'onda di suoni, segnali e rumori sull'asse delle ascisse, sia nel dominio del tempo (audio editor), sia in quello della frequenza (analizzatore di spettro). Sull'asse delle ordinate viene sempre rappresentata l'ampiezza del suono espressa in $dB_{fs.}$

Tipi di forme d'onda

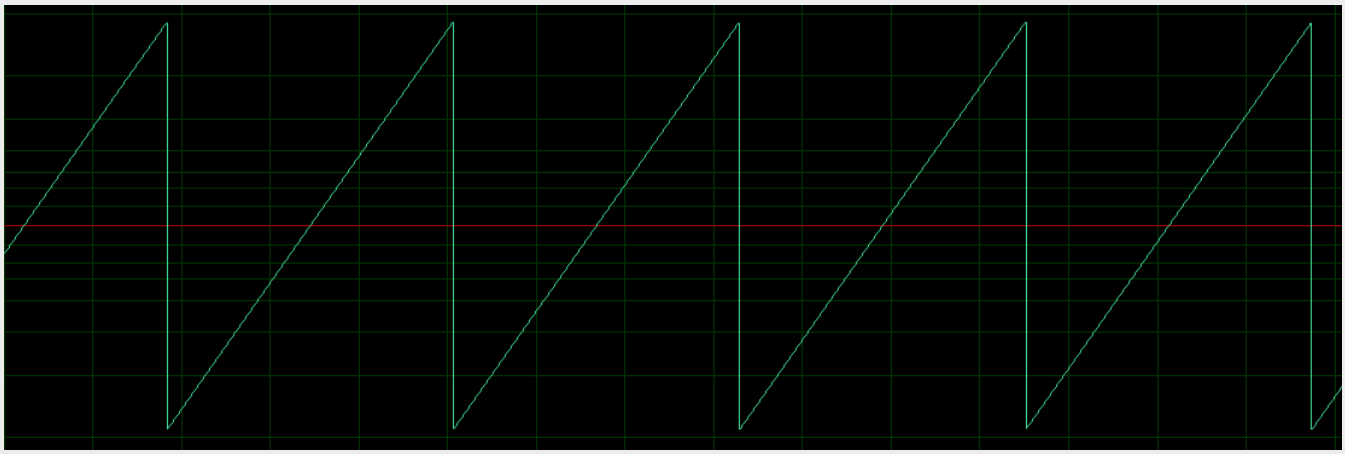
Onda sinusoidale complessa:
Come citato sopra, un suono o segnale audio complesso deve essere composto dalla prima frequenza denominata fondamentale e dalle armoniche. Nel caso della sinusoidale sono calcolate e prese armoniche multiple della fondamentale, sia pari, sia dispari.
Esempio :
f (fondamentale) = 85Hz → $2f$ = 170Hz → $3f$ = 340Hz → $4f$ = 680Hz



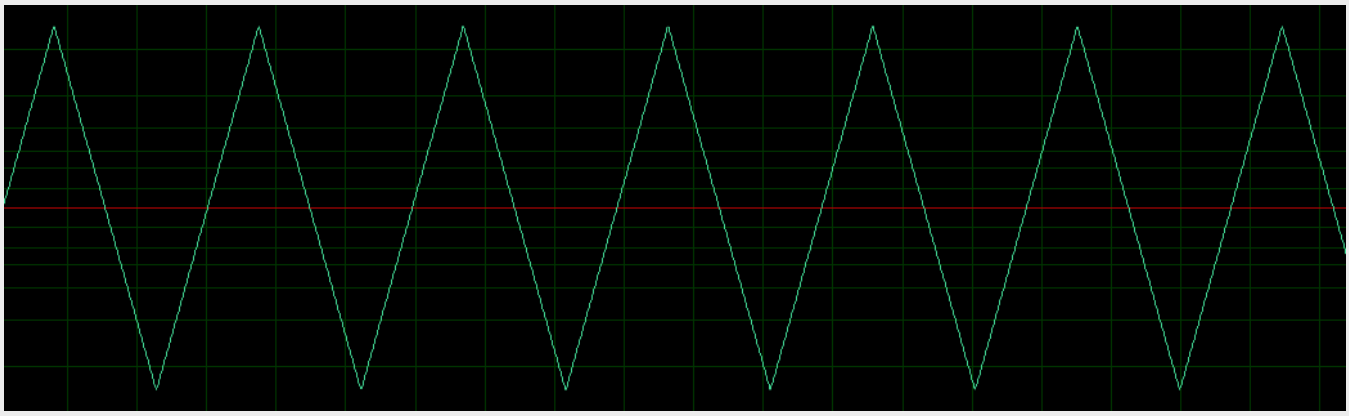
Onda quadra:
Sempre partendo dalla sinusoidale e prendendo armoniche dispari con ampiezza pari a 1/3, 1/5, 1/7 della fondamentale, ecco che si otterrà l'onda quadra. Per avere un'ottima onda quadra c'è bisogno di un numero infinito di armoniche. Si possono ottenere onde quadre anche con un numero finito di armoniche però, più sono e più l'onda quadra sarà precisa, sia nella grafica, sia nel suono. Naturalmente il suono dell'onda quadra è diverso da quello sinusoidale.
Esempio :
f (fondamentale) = 85Hz → $3f$ = 255Hz → $5f$ = 425Hz → $7f$ = 595Hz → ..



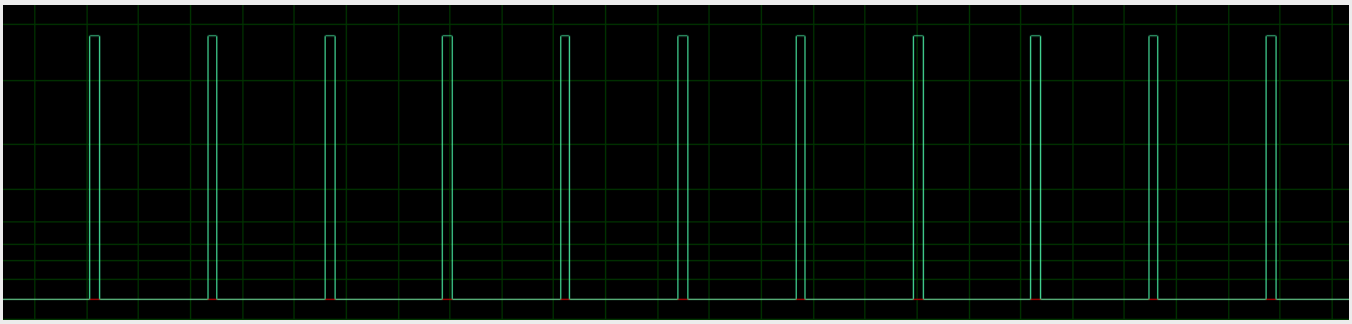
Onda a dente di sega:
Nell'onda a dente di sega sono presenti tutte le armoniche (sia pari che dispari). L'ampiezza di ogni armonica è pari alla metà dell'armonica precedente. Anche qui, suono differente.



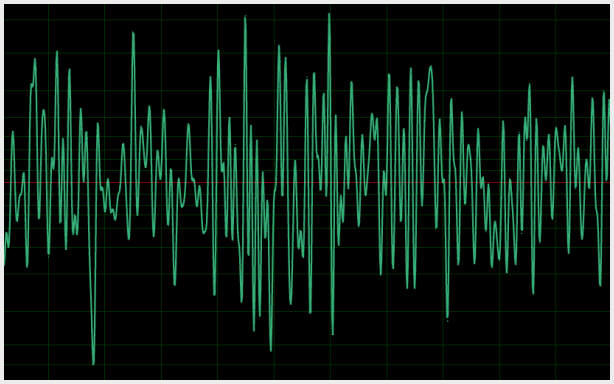
Onda triangolare:
Ha un contenuto armonico molto simile a quello dell'onda quadra. La differenza è che le ampiezze decrescono con un andamento del tipo $1/f^2$. Anche qui, suono differente dalle altre onde viste finora.



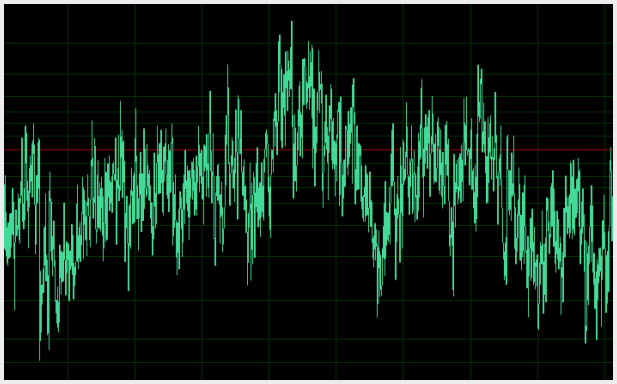
Onda a impulsi:
L'onda a impulsi si ottiene limitando il tempo in cui l'onda quadra rimane nella massima ampiezza.
Lo spettro è simile all'onda quadra, ma in questa a impulsi, sono presenti tutte le armoniche tranne il reciproco del tempo dell'impulso e i suoi multipli. Anche questa forma d'onda ha un suono differente.



Rumore bianco, rosa e marrone:
In questi tre elementi non è presente alcuna forma di periodicità. Il loro spettro si dice inarmonico e contiene tute le frequenze con la stessa ampiezza.



Rumore bianco



Rumore Rosa

Il *rumore marrone* è simile al rumore rosa, ma più ricco di frequenze basse e può essere presente in alcuni sintetizzatori virtuali..

Alcuni aspetti della frequenza

Sulla frequenza del suono, è bene citare e ricordare alcuni aspetti che possono servire in futuro:

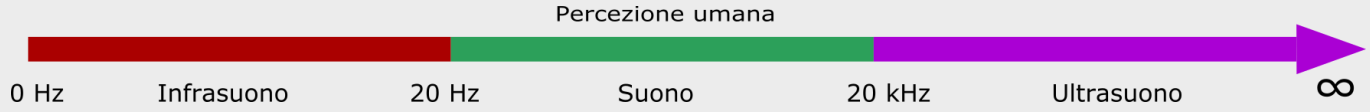
- Battimenti
- Accordatura aurea
- Ultrasuoni
- Infrasuoni o frequenze subsoniche

Battimenti:
I battimenti si verificano quando si fanno suonare due onde sinusoidali aventi una frequenza poco differente l'una dall'atra (in questo caso c'è una tolleranza di 2 - 8 Hz max. di differenza). Facendo un esempio, abbiamo una frequenza di 120 Hz e una di 124 Hz. Otterremo un segnale che oscilla a un'ampiezza di 4Hz, cioè quattro volte al secondo. Quest'oscillazione è chiamata onda modulante mentre, la media tra le due frequenze, cioè 122 Hz, è chiamata onda portante. A livello uditivo, sentiremo appunto che il suono "pulsa", "oscilla" o "batte". In musica, ci sono spesso fenomeni di battimenti (per esempio suonando due note musicali molto vicine tra di loro) ma sono meno percepibili quando i suoni non sono puri, come nel caso delle onde sonore sinusoidali, triangolare, ETC.
Ci sono anche casi di battimenti anche tra le armoniche dei suoni complessi, ma qui il discorso si fa lungo e complicato. Anche in Fisica (scienza), si continuano a studiare i fenomeni dei battimenti..

Accordatura aurea:
L'accordatura standard degli strumenti musicali è impostata a 440 Hz corrispondente alla nota La di IV [a] ottava. Nell'accordatura aurea, tale nota è impostata a 432 Hz. Per questo tipo d'intonazione, c'è tutta una teoria scientifica e filosofica che è spesso fonte di dibattiti e discussioni.. Si fa riferimento alla Risonanza di Schumann e alla musicoterapia. Secondo alcuni studi, multipli e sottomultipli della frequenza aurea, unita alla tecnica dei battimenti, porterebbero giovamento alla salute dell'uomo.

Ultrasuoni:
L'essere umano riesce a percepire un range di frequenze che va all'incirca da un minimo di 20 Hz a un massimo 20 kHz (20'000 Hz). Con l'avanzare dell'età, naturalmente, questo range va sempre più a restringersi, specie sulle frequenze alte.. Oltre i 20 kHz, abbiamo gli ultrasuoni. Sono frequenze inudibili dall'essere umano, ma comunque utilizzate e riprodotte artificialmente per scopi medici e industriali. Solitamente i diffusori acustici per Hi-Fi, nella riproduzione delle frequenze alte, non superano i 20 kHz tuttavia, i monitor da studio, possono superare questo limite, toccando anche i 50 kHz.
In natura ci sono alcuni animali che possono percepire ed utilizzare gli ultrasuoni per spostarsi e comunicare: cani, gatti (solo percezione), pipistrelli, topi, balene, delfini (movimento e comunicazione).

Infrasuoni:
Al contrario degli ultrasuoni, gli infrasuoni sono frequenze sotto i 20 Hz. Possono propagarsi a lunghe distanze e con poca dissipazione. Gli infrasuoni sono presenti in natura nei tuoni, terremoti, venti, eruzioni vulcaniche, maremoti e sono percepibili da alcuni animali (gli stessi citati sopra). A livello musicale sono definite anche come frequenze subsoniche. In un brano musicale possono essere presenti, ma raramente si riescono a produrre frequenze sotto i 20 Hz, almeno che non si disponga d'impianti molto grandi e molto potenti..
Si dice che determinate frequenze infrasoniche (se emesse ad alto volume) possono recare danno all'uomo provocando vertigini, nausee, emicranie e alterazioni dell'umore.

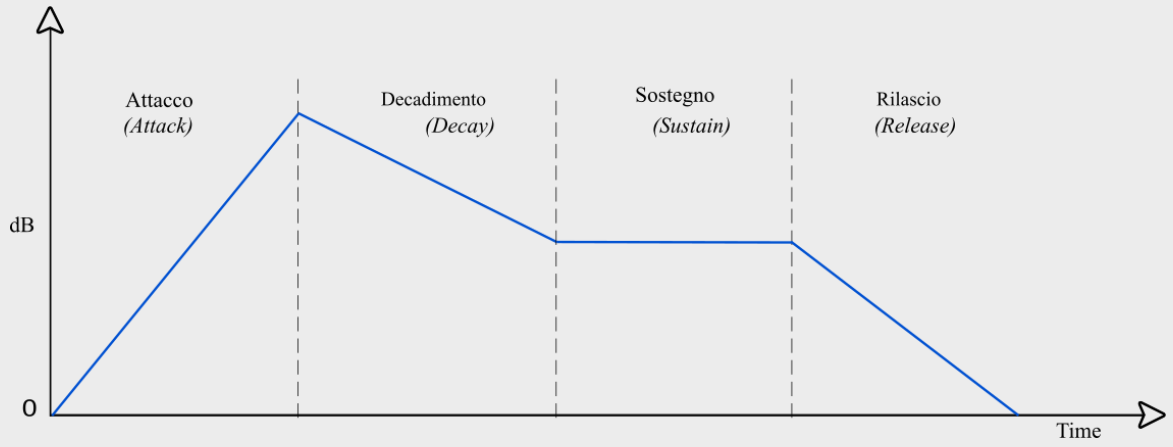


Inviluppo

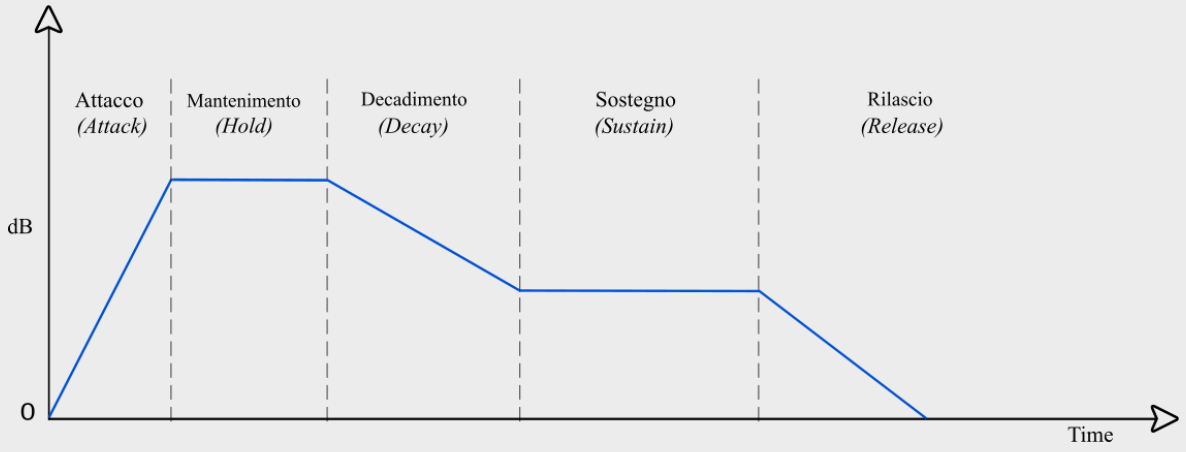
L'inviluppo è il modo in cui il suono evolve nel tempo rispetto alla sua ampiezza ed è composto da quattro transitori, che sono:

- **Attacco** (*Attack*). L'ampiezza parte da zero alla massima ampiezza.
- **Decadimento** (*Decay*). Dopo l'attacco l'ampiezza tende a diminuire fino ad un certo livello.
- **Sostegno** (*Sustain*). L'ampiezza mantiene costante per un certo tempo.
- **Rilascio** (*Release*). L'ampiezza diminuisce fino ad andare a zero.

L'inviluppo è presente nei sintetizzatori analogici, digitali e softsynth. Viene abbreviato con il termine ADSR.



Ci sono casi particolari di ADSR che aggiungono anche la funzione "Hold Time". Questo sistema permette di tenere le note di sustain per un certo periodo di tempo prima di decadere. Viene abbreviato con il termine AHDSR.



Ci sono ulteriori forme d'inviluppo come nel Prophet '08, sintetizzatore di Dave Smith, dove viene inserito davanti all'attacco un ritardo (Delay). DADSR (Delay, Attack, Decay, Sustain, Release).

Il “movimento del suono”

Come scritto nelle primissime pagine, Il suono è la vibrazione di un corpo che oscilla nell’aria. Queste nano particelle si propagano in maniera irregolare, dando forma a “problemi acustici”. Quelli più comuni si formano all’interno di locali, dove si fa o semplicemente si ascolta musica.

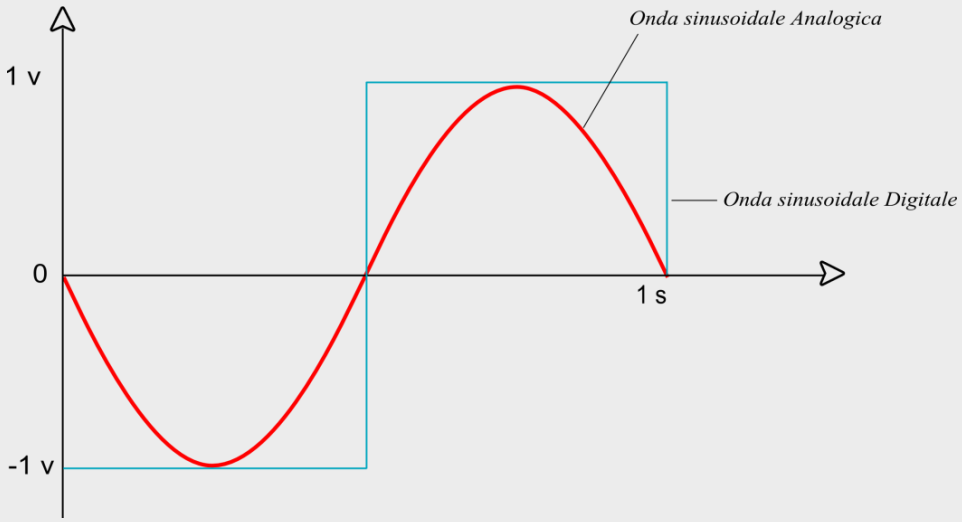
- **Riflessione** Il suono “rimbalza” con lo stesso angolo di incidenza dell’ostacolo, cambiando direzione.
- **Rifrazione** Il suono cambia direzione quando attraversa due corpi di densità diversa.
- **Diffrazione** Il suono (soprattutto le frequenze basse) aggirano un ostacolo.
- **Assorbimento** Il suono “scontrandosi” con un materiale , viene in parte assorbito e in parte disperso.
- **Effetto Doppler** Il suono si sposta rispetto all’ascoltatore. (Es. sirena 118).

Questi fenomeni andrebbero trattati molto più approfonditamente, perché sono legati allo studio dell’acustica, e per la quale se ne occupano fisici e ingegneri del suono. E’ noto che in tutti gli ambienti dove viene riprodotta, ascoltata o eseguita musica, c’è l’opportuno trattamento acustico al fine di evitare o ridurre al minimo i problemi citati sopra.
Il trattamento acustico si effettua prima a livello strutturale e poi con l’impiego di opportuni pannelli fonoassorbenti di vario modello e materiale, all’interno del locale.

Suono analogico e suono digitale

Con l’avvento del digitale, la musica è diventata sempre più “virtuale”. Le macchine interpretano il digitale sotto forma di bit, ovvero una combinazione di due numeri: l’1 e lo 0.
Semplificando, il suono analogico è la rappresentazione più fedele possibile di un segnale elettrico o audio.
Per digitale, invece, si intende un sistema che utilizza segnali discreti per interpretare un segnale continuo analogico.
Il processo di conversione dall’analogico al digitale, viene detto anche: digitalizzazione.
Supponiamo di voler “digitalizzare” un’onda sinusoidale, la quale avrà valori massimi d’ampiezza -1Volt e +1Volt.
Utilizzando un solo bit, cioè un solo numero non possiamo altro che avere a disposizione lo 0 e l’1. Quindi assegniamo:

- 0 per -1 Volt
- 1 per +1 Volt



Nell'immagine sopra, vediamo la digitalizzazione di una sinusoidale di 1Hz utilizzando 1bit. Quindi cosa cerca di fare il sistema digitale: in termini semplici, cerca di imitare il più possibile il suono analogico e per farlo ha bisogno di molti bit. Più sono i bit, più la linea azzurra si "spezzetterà" in parti sempre più piccole fino ad assomigliare alla "curva perfetta" del segnale analogico. Ovviamente bisogna dividere il valore (nel nostra caso) di 1 volt con un numero sempre maggiore di bit.
Infatti, cosa succede se abbiamo a disposizione 2 bit ? Avremo quattro possibilità di combinare lo 0 e l'1.

00	01	10	11

Quattro combinazioni da dividere in due, (perché bisogna considerare l'oscillazione nella fase positiva e in quella negativa), quindi l'operazione da fare sarà:
4/2 = 2 e il valore di 1 Volt andrà diviso per 2 → 1/2 = 0,5

00 per -1Volt	01 per -0,5Volt	10 per 0,5 Volt	11 per +1 Volt
-1 + 0,5= -0,5	-0,5 + 0,5 = 0	0 + 0,5 = 0,5	0,5 + 0,5 = +1

E se si utilizzano 3 bit ?

000	001	010	011	100	101	110	111

Otto combinazioni. Si farà dunque la stessa operazione fatta sopra:
8/2 = 4 → 1Volt /4 = 0,25

000 per -1Volt	-1 + 0,25 = - 0,75
001 per -0,25Volt	-0,75 + 0,25 = -0,5
010 per -0,5Volt	-0,5 + 0,25 = -0,25
011 per -0,75Volt	-0,25+0,25 = 0
100 per +0,25Volt	0 + 0,25 = 0,25
101 per +0,5Volt	0,25 + 0,25 = 0,5
110 per +0,75 Volt	0,5 + 0,25 = 0,75
111 per +1Volt	0,75 + 0,25 = +1

Infine proviamo ad utilizzare 4bit: Si otterranno sedici combinazioni: 16/2 = 8 → 1/8 = 0,125

0000	0000 per -1 Volt	-1 + 0,125 = -0,875
0001	0001 per -0,125 Volt	-0,875 + 0,125 = -0,75
0010	0010 per -0,25 Volt	-0,75 + 0,125 = -0,625
0011	0011 per -0.375 Volt	-0,625 + 0,125 = -0,5
0100	0100 per -0,5 Volt	-0,5 + 0,125 = -0,375
0101	0101 per -0,625 Volt	-0,375 + 0,125 = -0,25
0110	0110 per -0,75 Volt	-0,25 + 0,125 = -0,125
0111	0111 per -0,875 Volt	-0,125 +0,125 = 0
1000	1000 per +0,125 Volt	0 + 0,125 = 0,125
1001	1001 per +0,25 Volt	0,125 + 0,125 = 0,25
1010	1010 per +0,375 Volt	0,25 + 0,125 = 0,375
1011	1011 per +0,5 Volt	0,375 + 0,125 = 0,5
1100	1100 per +0,625 Volt	0,5 + 0,125 = 0,625
1101	1101 per +0,75 Volt	0,625 + 0,125 = 0,75
1110	1110 per +0,875 Volt	0,75 + 0,125 = 0,875
1111	1111 per +1 Volt	0,875 + 0,125 = +1

Dunque, avremo nell'ordine:

1bit – 2 livelli
2 – 4
3 – 8
4 – 16
5 – 32
6 – 64
7 – 128
8 – 256
9 – 512
10 – 1.024
11 – 2.048
12 – 4.096
13 – 8.192
14 – 16.384
15 – 32.768
16 – 65.536
17 – 131.072
18 – 262.144
19 – 524.288
20 – 1.048.576
21 – 2.097.152
22 – 4.194.304
23 – 8.388.608
24 – 16.777.216
25 – 33.554.432
26 – 67.108.864
27 – 134.217.728
28 – 268.435.456
29 – 536.870.912
30 – 1.073.741.824
31 – 2.147.483.648
32 – 4.294.967.296

Come si può vedere dai numeri, se si aggiunge un bit, "il livello di digitalizzazione" diventa il multiplo del precedente.
Più aumentano i bit, più precisi saranno: la rappresentazione digitale della nostra curva e la campionatura dei livelli di ampiezza del segnale. Quest'operazione di calcolo è chiamata **Quantizzazione lineare.**
In termine tecnico, il numero di bit prende il nome di **risoluzione.**

Convertitori ADC - DCA e la frequenza di campionamento

La conversione Analogico / Digitale (o digitalizzazione del suono analogico) è un termine che nella musica si trova e se ne discute molto spesso. Nell'audio ad alta fedeltà, ci sono apparecchi appositi che convertono appunto il suono analogico in digitale e viceversa. I convertitori Analogico-Digitali sono abbreviati con il termine A/D (o ADC) mentre, quelli che da digitale convertono in analogico, sono denominati DAC (Digital to Analog Converter).
Nel settore della musica digitale, ci sono schede audio di fascia medio - alta in cui la conversione A/D e DAC è affidata a un DSP ovvero un chip o processore. E' anche possibile acquistare separatamente appositi hardware DSP (schede acceleratrici) che aiutano l'elaborazione del suono, alleggerendo la fatica di calcolo al processore principale del computer, o a quello della scheda audio stessa. La frequenza di campionamento è strettamente correlata alla conversione ADC – DAC. Più è alto questo valore più il suono sarà "fedele".
Per frequenza di campionamento s'intende la misura espressa in Hertz del numero di volte al secondo in cui un segnale analogico è misurato e memorizzato in forma digitale.
La frequenza di campionamento non va confusa con la frequenza di un suono o un segnale audio. La frequenza di campionamento stabilisce la qualità del suono.
Lo standard internazionale (Audio CD Red Book) ha fissato la risoluzione a **16 bit e la frequenza di campionamento in 44.100 Hz stereo digitale.**
Schede audio di qualità professionale, possono arrivare a frequenze di campionamento di 192 kHz a 24 bit.

Qui sotto viene riportata la tabella dei bit:

- 1 Byte = 8 bit (bit)
- 1 Kilobyte = 1.024 byte (KB)
- 1 Megabyte = 1.024 KB (MB)
- 1 Gigabyte = 1.024 MB (GB)
- 1Terabyte = 1.024 GB (TB)

Altri valori relativi al byte:

- 1 nibble = 4 bit
- 1 word = 2 byte (16 bit)
- 1double word = 4byte (32 bit)
- 1quad word = 8byte (64 bit)

- kibibyte 1024 byte, indicato con KiB
- mebibyte 1024 kibibyte, indicato con MiB
- gibibyte 1024 mebibyte, indicato con GiB
- tebibyte 1024 gibibyte, indicato con TiB
- pebibyte 1024 tebibyte, indicato con PiB
- exbibyte 1024 pebibyte, indicato con EiB
- zebibyte 1024 exbibyte, indicato con ZiB
- yobibyte 1024 zebibyte, indicato con YiB

Calcolo matematico che verifica quanto spazio può occupare un file audio

Ora che abbiamo "un'infarinatura" generale di bit, Byte e campionamento, proviamo a calcolare quanto "spazio" verrà occupato nella realizzazione o registrazione di materiale audio digitale.
Se vogliamo digitalizzare o campionare un suono analogico ad una frequenza di campionamento di 44.100Hz e 16bit di risoluzione, avremo che, per ogni secondo di campionamento, il sistema audio digitale fotograferà il segnale analogico in ingresso al convertitore A/D per ben 44.100 volte.

Per quanto riguarda la risoluzione in bit avremo che ogni nostro campione potrebbe assumere una quantità di valori d'ampiezza (livelli di volume sonoro) pari a 2^{16} = 65'536 valori possibili organizzati da –32768 a +32768 sull'asse verticale dell'ampiezza.
Mettiamo ora il caso di voler registrare una traccia audio stereo della durata di 60 secondi; il calcolo che andremo a fare sarà il seguente:

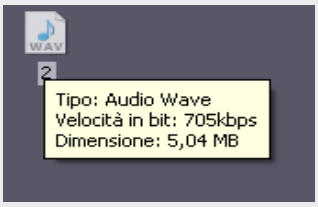
16 bit x 44.100 x 2 (canali stereo) = 1.411.200 bit (un secondo in stereo)
1.411.200 / 8bit = 176.400 byte
176.400 / 1.024 = 172,265 Kbyte (KB)
172,256 / 1.024 = 0,168 MB
0,168 x 60secondi = **10 MB** → questo è lo spazio effettivo che occuperà la nostra traccia audio

Nel caso volessimo registrare una traccia mono, basterà non moltiplicare per due; e quindi:
16 x 44.100 = 705.600 bit
705.600 / 8 = 88.200 byte
88.200 / 1.024 = 86,132 Kbyte
86,132 / 1.024 = 0,084 MB
0,084 x 60 = **5,04 MB** → traccia audio mono di 60 secondi o 1 minuto

Per un'ulteriore verifica, utilizzando Windows come sistema operativo, basterà mettere la freccia del mouse sull'icona del nostro file audio, e appariranno in un riquadro giallo o dati del nostro progetto: la dimensione in MB e la velocità in kbps (kilo bit per secondo) detta **bitrate** facilmente ottenibile dividendo per 1.000 la prima operazione matematica che abbiamo fatto sopra nel nostro calcolo:

1.411.200 / 1.000 = 1.141,2 kbps (bitrate)

Traccia Stereo

705.600 / 1.000 = 705,6 kbps (bitrate) 

Traccia Mono

Es. proviamo a veder quanto spazio occupa un file stereo di 48.000Hz a 24bit della durata sempre di 60 secondi:

24bit x 48.000 x 2(canali stereo) = 2.304.000 bit (un secondo in stereo)
2.304.000 / 8bit = 288.000 byte
28.000 / 1.024 = 281,25 Kbyte (KB)
281,25 / 1.024 = 0,274 MB
0,274 x 60secondi = **16,47 MB →** questo è lo spazio effettivo che occuperà la nostra traccia audio.

Un'ulteriore verifica, per avere un quadro più completo dei dati si può eseguire cliccando col tasto destro del mouse sull'icona del file e andando poi a cliccare sulla voce "Proprietà".

Come possiamo notare, più alta sarà la frequenza di campionamento e più spazio occuperà il nostro file.

Formati e compressioni audio

Le frequenze di campionamento sono gestite da diversi tipi di formati audio, che nel corso degli anni si sono sviluppati nell'ambito dell'audio digitale. La conversione in diversi formati audio, viene realizzata da programmi chiamati ***Audio Converter***, come ad esempio AVS Audio Converter, mentre per la manipolazione dell'audio ci sono programmi chiamati ***Audio Editors.*** Oggi giorno, abbiamo a disposizione molti programmi per lavorare con l'audio, alcuni free, altri invece a pagamento e più professionali:

- Adobe Audition
- Sony Sound Forge
- Steinberg WaveLab

Possiamo per comodità dividere i formati audio in ***lossy*** (con perdita di qualità audio) e ***no lossy*** (o **lossless**, senza perdita di qualità). I formati *lossy* (i più diffusi tra gli utenti comuni) nascono dall'idea che non tutti i suoni presenti nei 44.100 campioni al secondo contenuti in un file WAVE standard vengono correttamente percepiti dall'orecchio umano. In questo modo (così come avviene per il jpg delle immagini) si vanno a tagliare le alte frequenze, che si ritiene siano meno distinte dal nostro orecchio. Più si tagliano queste frequenze, più lo spazio occupato dalla nostra traccia diminuisce (per esempio 3 MB al posto di 6 MB per un brano di alcuni minuti, o molto di più), ma diminuisce anche la qualità del risultato, dato che le frequenze che vengono tagliate non sono più tanto 'inaudibili'.
La riconversione in WAVE, partendo dai formati lossy, **non porta alcun beneficio** (a meno che si debba elaborare la traccia e salvarla più volte, in questo caso la perdita di qualità si limiterà al primo passaggio); esistono inoltre, programmi che riescono a restituire ai nostri brani alcune di queste alte frequenze, come Steinberg Clean Plus o per esempio, le prime versioni di Easy CD Creator, o altri ancora più professionali, tuttavia, pur facendo un discreto lavoro, più di tanto non possono fare.
I formati senza perdita (no lossy), invece, cercano di diminuire lo spazio occupato dalla traccia, senza andare a toccare il suono; la percentuale di compressione sarà decisamente inferiore rispetto ai lossy, ma non ci sarà perdita di qualità. Se riconvertiamo in WAVE, un formato "*lossless*" (magari per essere elaborato), il suono sarà identico a prima.

No Lossy o Losless

WAVE (**.wav**) sta per ***WAVE form audio format*** è un formato di file audio usato soprattutto in ambiente Windows. La sua risoluzione standard è a 16bit e la frequenza di campionamento è a 44,1 kHz. I formati di codifica possono essere:

- PCM
- LPCM
- DPCM
- ADPCM
- Onda A/mu-Law
- Forma d'onda ACM
- DVI/IMA ADPCM

PCM: Pulse-code modulation (letteralmente "modulazione codificata di impulsi")
E' un formato Microsoft Windows PCM supporta file sia mono che stereo a diverse risoluzioni e frequenze di campionamento. Il PCM non richiede compressione ed è considerato un formato senza perdita di informazioni. (Questa è la codifica maggiormente usata nell'ambito dell'audio professionale).

LPCM: Linear pulse code modulation. E' un metodo particolare che rappresenta una forma d'onda audio come una sequenza di valori di ampiezza registrati in una sequenza di volte. Precisa che i valori memorizzati sono proporzionali alle ampiezze, piuttosto che dire che rappresenta il logaritmo dell'ampiezza, o sono legati in qualche altro modo. In pratica questi valori saranno quantizzati. Le risoluzioni comuni per il formato LPCM sono 8, 16, 20 o 24 bit per campione.

DPCM: La differential pulse-code modulation codifica i valori della PCM tramite le differenze tra il valore successivo e quello precedente. Per i segnali audio questo tipo di codifica riduce il numero di bit richiesti per campione di circa il 25% rispetto al PCM.
La DPCM è una tecnica di predizione dei campioni utilizzata per la trasmissione di segnali analogici in formato numerico (digitale). La tecnica prevede la quantizzazione e codifica binaria dell'errore tra la predizione generata e il campione analogico in ingresso. Il valore di predizione viene ottenuto mediante un predittore più complicato rispetto a quelli utilizzati in tecniche più semplici ma meno precise e tiene conto della funzione di autocorrelazione di un segnale analogico.

Microsoft ADPCM: L'adaptive DPCM è una variante del DPCM che rende arbitraria la dimensione dei passi di campionamento, così da consentire ulteriori riduzioni della larghezza di banda richiesta per un dato rapporto segnale-rumore. Il formato Microsoft ADPCM fornisce una compressione 4:1. I file salvati in questo formato si espandono automaticamente a 16 bit all'apertura, indipendentemente dalla risoluzione originale. Per questo motivo, utilizzate questo formato con file a 16 bit e non a 8 bit.

Onda A/mu-Law: I formati A-Law e mu-Law WAV (CCITT standard G.711) sono comuni nelle applicazioni di telefonia. Questi formati di codifica comprimono l'audio originale a 16 bit in audio a 8 bit (per un rapporto di compressione 2:1) con un intervallo dinamico di circa 13 bit (78 dB). Oltre ad avere un rapporto segnale/rumore maggiore rispetto a PCM a 8 bit, le forme d'onda codificate A- Law e mu-Law presentano una maggiore distorsione rispetto all'audio originale a 16 bit. Inoltre, la qualità è superiore a quella che otterreste con alcuni formati ADPCM a 4 bit.

Forma d'onda ACM: Microsoft ACM (Audio Compression Manager) fa parte di tutte le versioni di Windows a 32 bit. Alcuni di questi formati vengono forniti insieme a Windows, mentre altri sono forniti da terzi. Potete acquisire i formati ACM quando installate altri software.

DVI/IMA ADPCM: La caratteristica International Multimedia Association (IMA) di ADPCM comprime i dati a 16 bit in 4 bit/campione (4:1) utilizzando un metodo diverso (più veloce) da Microsoft ADPCM. Esso presenta diverse caratteristiche di distorsione che possono produrre risultati migliori o peggiori a seconda del campione compresso. Analogamente a Microsoft ADPCM, utilizzate questo formato con file a 16 bit e non a 8 bit. Questo schema di compressione può costituire un'ottima alternativa a MPEG; esso fornisce una decodifica ragionevolmente veloce della compressione 4:1 e diminuisce solo leggermente la qualità del campione.

AIFF (**.aif**). E' il formato audio Apple utilizzato dalla Apple per i Mac. E' in pratica l'equivalente del WAV che viene utilizzato da Windows.

APE (Monkey Audio; **.ape**): Formato no lossy che permette di ridurre di circa il 50% lo spazio occupato dalla nostra musica (in certi casi anche di più) senza alcuna perdita in qualità. In questo modo, un album che in formato wav occupa circa 600 MB, viene ad occupare in media 300 MB (molto di più dei 100MB circa di un mpc a bitrate alti, e dei 60 di un mp3, tuttavia la qualità è identica all'originale); dico in media perché ci sono certi tipi di musica in cui il livello di compressione è ancora più elevato. Per ascoltare brani in questo formato, si possono usare dei plugin per WinAmp o -meglio ancora- player che lo incorporano in modo nativo come Foobar 2000. In questo momento è probabilmente il *miglior codec lossless*, considerando un equilibrio tra velocità e compressione (clicca qui per la tabella comparativa tra i formati lossless).

FLAC (.flac): Formato Open Source, di grande popolarità. Una volta era estremamente diffuso (tra puristi della musica), ma sta perdendo piano piano terreno con l'affermarsi del formato ape di MonkeyAudio, che a parità di qualità, permette una migliore compressione. Il nostro album da 600Mb, andrà ad occupare in FLAC circa 330-340 MB.

LA (Lossless Audio: **.la**): Il formato no-lossy che comprime di più. La conversione in questo formato è lentissima, però ottiene la migliore compressione in assoluto (riprendendo l'esempio di MonkeyAudio, quell'album occuperebbe circa 290 MB, contro i 300 MB dell'APE). E' poco diffuso, dato che qualche MB guadagnato sull'ape non vale le lunghe attese in più (e un minor supporto da programmi terzi).

OPTIM FROG (**.ofr**)**:** ottimo codec lossless, il migliore dopo LA come comprimibilità. La compressione, se impostata a livelli alquanto elevati, è circa 3 volte più lenta rispetto a Monkey Audio, e il guadagno in MB è di circa il 2%. Il nostro album di riferimento andrà ad occupare 295 MB circa. Come tutti gli altri formati di questa pagina, può essere ascoltato senza plugin con Foobar 2000.

WAVPACK (.wv, wvc**)**: E' un valido formato di compressione open source, che permette sia risultati lossless, lossy e ibridi. La velocità di compressione è altissima, e il grado di compressione buono (l'album di riferimento occuperà circa 320 MB). Interessante anche il formato ibrido, in cui al file lossy si affianca un altro file, che permetterà di ricostruire interamente il file originale in maniera lossless nel caso lo volessimo ripristinare.

Lossy

WMA (**.wma**): formato windows media audio, compresso e molto simile a un mp3. Formato di compressione audio della Microsoft. I file compressi in questo formato sono di circa il 20 % più piccoli degli Mp3

MP3 (**.mp3**) è l'abbreviazione di MPEG 1 layer III e identifica i file audio che utilizzano questo algoritmo. E' uno standard che elimina i suoni non udibili dall'orecchio umano. In questo modo un brano mp3 a 128 kbps occupa uno spazio pari a 1/11 rispetto allo stesso in formato .wav o cd audio. Alcuni stimano un bitrate a 128 kbps come 'Cd Quality'; in realtà, le frequenze che vengono tagliate per risparmiare spazio sul disco rigido non sono tanto 'inaudibili'. A 192 i risultati cominciano ad essere buoni, e solo a 320 kbps si può parlare di CD Quality.. Fino a qualche anno fa la qualità dell'mp3 era considerata elevata, ma ora -con l'affermarsi di schede audio più avanzate, a 24 bit- si continua ad usare l'mp3 rispetto ad altri formati come ogg o mpc più per la sua diffusione e compatibilità che non per la qualità del risultato. Un album di 600 MB in formato wav, andrà ad occupare circa 50 MB in mp3 128. Il **bit rate** è il numero di unità binarie che fluiscono al secondo ed è variabile per i file MP3. La regola generale è che maggiore è il bitrate, più informazione è possibile includere dall'originale, maggiore è la qualità del file audio compresso. Attualmente per le codifiche dei file MP3 fissano un tasso di compressione equivalente per tutto il file audio.
Per l'MPEG-1 layer 3 i bitrate disponibili sono: 32, 40, 48, 64, 80, 96, 112, 128, 160, 192, 224, 256 e 320 kbit/s (103 bits per secondo), e le frequenze campionate disponibili sono 32, 44.1 e 48 Khz. La frequenza di campionamento a 44.1 kHz è quasi sempre utilizzata per i CD audio, mentre i 128 Kbit/s come una sorta di bitrate standard "abbastanza buono". L'MPEG-2 e l'MPEG-2.5 (non- ufficiale) contemplano un numero maggiore di bitrate: 8, 16, 24, 32, 40, 48, 56, 64, 80, 96, 112, 128, 144 e 160 kbit/s.
I file MP3 audio sono suddivisi in settori ("chunks" in inglese) chiamati frames, ("fotogrammi" in italiano). Ognuno di questi settori è provvisto di un marcatore del tasso di compressione; in questo modo, mentre il file audio viene riprodotto, è possibile modificarlo dinamicamente. Questa tecnica rende possibile utilizzare più bit per la parte del suono ad alta dinamica (suoni più complessi) e meno bit per la parte a bassa dinamica (suoni meno complessi).
Nei coder più evoluti (per es. LAME) l 'MP3 è gestito in tre modi, che a seconda dei programmi che si stanno utilizzando, possono essere scelti dall'utente:

CBR = bitrate costante
Può assumere valori che vanno da 320 a 8 e viene mantenuto costante su tutti i frames (vedi Codifica).
Vantaggio: permette di prevedere in modo preciso le dimensioni finali del file e la banda necessaria in trasmissione.
Svantaggio: spreco di banda nei passaggi con poca dinamica o silenzio.

ABR = bitrate medio
Il coder può cambiare il bitrate sui vari frames, diminuendolo quando è possibile risparmiare (passaggi con poca dinamica o silenzio) e aumentandolo quando necessario, cercando di mantenere il valore medio richiesto.
Vantaggio: qualità migliore rispetto al CBR nei passaggi difficili e spreco ridotto in quelli semplici.
Svantaggio: una certa incertezza sulle dimensioni finali del file, ma soprattutto difficoltà di valutazione della banda.

VBR = bitrate variabile
Ritenuto più flessibile del precedente.
Il coder può cambiare il bitrate sui vari frames con maggiore libertà in base alle specifiche di min e max dell'utente che esegue la codifica.
Vantaggio: il bitrate si adatta meglio alle problematiche di codifica e alle richieste dell'utente.
Svantaggio: maggiore incertezza sulle dimensioni finali del file.

Dalle specifiche tecniche, il VBR appare la scelta migliore e in effetti i file in VBR suonano praticamente identici ai CBR con bitrate superiore. Resta il fatto che parecchi audiofili hanno i brividi all'idea che la qualità della codifica possa cambiare da un istante all'altro.

AC3 (**ac3**). E' il formato audio usato dai DVD. In genere lo troviamo a 384 kbps (e 6 canali), ma è possibile anche averlo in soli 2 canali e a bitrate inferiori. Uno dei programmi gratuiti che lo supporta (e che consente per esempio di ridurre il bitrate) è BeLight (Besweet).

OGG VORBIS (**.ogg**): Ottimo codec open source. Riesce a dare risultati migliori dell'mp3 soprattutto a bitrate bassi (maggior qualità in minor spazio), cioè inferiori a 128 kbps. E' supportato praticamente da ogni player audio (sia software che audio).

MusePack (.mpc): Formato di altissimi risultati, soprattutto a bitrate alti (oltre 192 kbps e oltre). I risultati sono decisamente migliori dell'mp3: basta ascoltare una traccia codificata in mp3 a 192 kbps e una in mpc allo stesso bitrate per sentire immediatamente la differenza (in certi casi si resta senza parole...). Erano in molti a considerarlo il miglior formato audio lossy, almeno fino all'uscita dell'mp4. E' supportato da WinAmp tramite plugin, e in modo nativo da player più evoluti come Foobar 2000. Un album di 600 MB in formato WAVE andrà ad occupare circa 100 MB se codificato in mpc con bitrate intorno ai 192-256 kbps. Da notare che dà il meglio di sé ad alti bitrate, mentre a bassi bitrate sono probabilmente preferibili ogg e acc.

Nero MP4 AAC (.mp4, .aac**)**: Il più recente, è in questo momento il miglior formato lossy, avendo superato anche la pur elevatissima qualità dell'mpc. La qualità è estremamente elevata, seppure i tempi di conversione sono un po' più lenti rispetto agli altri formati.

AAC (di iTunes - **.aac**): Formato lossy che consente una qualità elevata (sempre parlando di formati lossy...), almeno un gradino superiore a quella dell'mp3. Caratteristica di questo formato è la possibilità di essere protetto tramite DRM, in modo da non poter essere copiato liberamente da una piattaforma all'altra (a meno di utilizzare programmi che tolgano questa protezione..

Real Player (**.ra**, **.rm** e altre estensioni simili): codec molto usato in certi video e in certi audio. RealPlayer è un programma gratuito che consente di eseguire tracce audio e video create con il suo codec, tuttavia è alquanto pieno di pubblicità. Per questa ragione, suggerisco Real Alternative.

Tecniche di esportazione audio il “Dithering”

Abbiamo scritto nelle pagine precedenti che maggiore è la frequenza di campionamento e migliore sarà la qualità del nostro file audio. Ma lo spazio? Un file di 5 minuti alla frequenza di 48.000Hz e 24 bit ci tiene la bellezza di 80Mb circa. E allora come possiamo mantenere una qualità più alta risparmiando spazio? Ecco che una mano ci arriva dalla tecnica chiamata **Dithering.**
Il **dithering**, nell’elaborazione numerica di segnali, è una forma di rumore con un’opportuna distribuzione aggiunta ai campioni, con l'obiettivo di minimizzare la distorsione introdotta dal troncamento del numero di bit in eccesso, nel caso in cui si riquantizzino i campioni stessi. Il dithering viene usato abitualmente nell'elaborazione di segnali video e audio campionati e quantizzati.

Ovviamente, vi sono vari algoritmi di dithering, più o meno efficaci e più o meno adatti al tipo di materiale cui devono essere applicati. I software di editing audio, (Adobe Audition) solitamente, permettono di scegliere il tipo di dithering da utilizzare in un ricampionamento. In generale, quando si applica il dithering si deve scegliere il tipo di rumore introdotto, descritto da una funzione specifica (di seguito le più diffuse):

- **TPDF** (Triangular Probability Density Function, ovvero Funzione triangolare di densità della probabilità) è una delle più utilizzate per il materiale audio, equivalente al lancio di due dadi. La somma dei dadi ha differenti probabilità d'uscita:

1/1 = 2
1/2 2/1 = 3
1/3 2/2 3/1 = 4
1/4 2/3 3/2 4/1 = 5
1/5 2/4 3/3 4/2 5/1 = 6
1/6 2/5 3/4 4/3 5/2 6/1 = 7
2/6 3/5 4/4 5/3 6/2 = 8
3/6 4/5 5/4 6/3 = 9
4/6 5/5 6/4 = 10
5/6 6/5 = 11
6/6 = 12

il 7 uscirà molto più spesso del 2 o del 12, e la relazione tra queste probabilità viene detta *triangolare*.

- **RPDF** (Rectangular Probability Density Function, Funzione rettangolare di densità della probabilità) ed è equivalente al lancio di un dado. Ogni numero ha la stessa probabilità di uscire.

- **Gaussian PDF** (Gaussian Probability Density Function, Funzione Gaussiana di densità della probabilità) è equivalente al lancio di infiniti dadi. La relazione tra le probabilità dei risultati ha una forma a campana detta curva Gaussiana. Il dithering prodotto dalla PDF Gaussiana è quasi coincidente al rumore atmosferico naturale, il fruscio delle musicassette, ecc.

-Il dithering colorato viene a volte descritto come quel dithering che è stato filtrato per differenziarsi dal rumore bianco. Alcuni algoritmi di dithering usano un rumore che possiede più energia nelle alte frequenze, e meno nella banda audio critica.

-Il **noise shaping** non è in realtà un dithering, poiché consiste nel "modellare" il rumore, come dice il suo nome, spostandolo in zone dello spettro audio in cui è meno udibile. Normalmente viene associato al dithering, e quindi sposta appunto questo rumore. È tuttavia possibile anche usarlo senza dithering. In tal caso il rumore di ri-quantizzazione rimane distorto, ma viene spostato in zone dello spettro audio in cui è meno udibile.

Se il segnale da processare deve subire un ulteriore lavoro, dovrebbe essere usato il TPDF che ha l'ampiezza di due passi di quantizzazione (in modo che i valori calcolati varino tra -1 e +1 o, se vogliamo, tra 0 e 2). Se il dithering colorato viene usato come processo intermedio, la gamma di frequenze può sfumarsi in un'altra e diventare vagamente udibile.
Se il segnale da lavorare non deve essere ulteriormente processato (ovvero accettiamo direttamente il risultato del dithering) è conveniente usare il dithering colorato o il noise shaping, in grado di diminuire enormemente il livello del rumore udibile spostandone buona parte nell'area non critica.

Downsampling

Un’altra tecnica di esportazione audio, o per meglio dire “riduzione della frequenza di campionamento” è il ***Downsampling.***
Il fattore di downsampling (comunemente denotato con la lettera *M*) è nella maggior parte dei casi un intero o una frazione razionale maggiore dell'unità. Questo fattore moltiplica il periodo di campionamento o, in maniera equivalente, ne divide la frequenza. Ad esempio, se ad un CD audio viene applicato un fattore M uguale a 5/4, la risultante frequenza di campionamento scenderà da 44.100 a 35.280 Hz, e la velocità di trasmissione dei dati passerà da 1.411.200 a 1.128.960 bit al secondo:

44.100Hz / 5/4 = 35.280 x 16 x 2 = 1.128.960 bit/s /1.000 = 1.129 kbps.

In teoria, durante il donwsampling, non dovrebbe essere infranto il teorema di Nyquist-Shannon. Se però questo avviene, si può presentare aliasing. Al fine di assicurare che il teorema sia soddisfatto, viene utilizzato un filtro passa basso assieme ad un filtro anti-aliasing per ridurre l' ampiezza di banda del segnale prima che esso venga compresso.
L'**aliasing** (letteralmente *impersonazione*), o *distorsione da campionamento lento*, da *sottocampionamento*, è il fenomeno che introduce in un segnale digitale, delle frequenze spurie, che non erano presenti nel segnale analogico in ingresso.

Resampling

Con il *resampling* è possibile convertire o modificare la frequenza di campionamento di un file audio, modificandone anche la quantizzazione e il numero dei canali (ad esempio da Stereo a Mono o viceversa).

Tabella di riferimento File Audio (.wav)

Tabella di riferimento File Audio .wav Mono				
File	Codec	Freq. di Campionamento	Bitrate	Risoluzione
Wave (.wav)	PCM	44,100 kHz	705 kbps	16 bit
Wave (.wav)	PCM	48,000 kHz	768 kbps	16 bit
Wave (.wav)	PCM	96,000 kHz	1’536 kbps	16 bit
Wave (.wav)	PCM	192,000 kHz	3’072 kbps	16 bit
Wave (.wav)	PCM	44,100 kHz	1'058 kbps	24 bit
Wave (.wav)	PCM	48,000 kHz	1'152 kbps	24 bit
Wave (.wav)	PCM	96,000 kHz	2'304 kbps	24 bit
Wave (.wav)	PCM	192,000 kHz	4'608 kbps	24 bit
Wave (.wav)	PCM	44,100 kHz	1'411 kbps	32 bit
Wave (.wav)	PCM	48,000 kHz	1'536 kbps	32 bit
Wave (.wav)	PCM	96,000 kHz	3'072 kbps	32 bit
Wave (.wav)	PCM	192,000 kHz	6'144 kbps	32 bit

Tabella di riferimento File Audio .wav Stereo				
File	Codec	Freq. di Campionamento	Bitrate	Risoluzione
Wave (.wav)	PCM	44,100 kHz	1’411 kbps	16 bit
Wave (.wav)	PCM	48,000 kHz	1’536 kbps	16 bit
Wave (.wav)	PCM	96,000 kHz	3’072 kbps	16 bit
Wave (.wav)	PCM	192,000 kHz	6’144 kbps	16 bit
Wave (.wav)	PCM	44,100 kHz	2’116 kbps	24 bit
Wave (.wav)	PCM	48,000 kHz	2’304 kbps	24 bit
Wave (.wav)	PCM	96,000 kHz	4’608 kbps	24 bit
Wave (.wav)	PCM	192,000 kHz	9’216 kbps	24 bit
Wave (.wav)	PCM	44,100 kHz	2’822 kbps	32 bit
Wave (.wav)	PCM	48,000 kHz	3’072 kbps	32 bit
Wave (.wav)	PCM	96,000 kHz	6’144 kbps	32 bit
Wave (.wav)	PCM	192,000 kHz	12’288 kbps	32 bit

Per trovare il bitrate bisogna moltiplicare la frequenza di campionamento per la risoluzione in bit (mono) e raddoppiare la cifra ottenuta, nel caso il file sia stereo.

Tabella note e frequenze

Note It.	Note Ang.	Ottave									
		0	1	2	3	4	5	6	7	8	9
DO	**C**	16,35	32,70	65,40	130,81	261,62	523,25	1046,50	2093,00	4186,01	8372,02
DO#	**C is**	17,32	34,64	69,30	138,60	277,18	554,36	1108,73	2217,46	4435,00	8870,00
RE	**D**	18,35	36,70	73,41	146,83	293,66	587,32	1174,66	2349,31	4698,63	9397,27
RE#	**D is**	19,44	38,89	77,78	155,56	311,12	622,25	1244,50	2489,01	4978,03	9956,06
MI	**E**	20,60	41,20	82,40	164,81	329,62	659,25	1318,51	2637,02	5274,04	10548,08
FA	**F**	21,82	43,65	87,30	174,61	349,23	698,45	1397,00	2793,82	5587,65	11175,30
FA#	**F is**	23,12	46,25	92,50	185,00	370,00	740,00	1480,00	2960,00	5920,00	11840,00
SOL	**G**	24,50	49.00	98,00	196,00	392,00	784,00	1568,00	3136,00	6272,00	12543,85
SOL#	**G is**	25,96	51,91	103,82	207,65	415,30	830,61	1661,21	3322,43	6644,87	13290,00
LA	**A**	27,50	55	110	220	**440**	880	1760	3520	7040	14080
LA#	**A is**	29,13	58,27	116,54	233,08	466,16	932,32	1864,65	3729,31	7458,62	14917,24
SI	**B**	30,86	61,73	123,47	246,94	493,88	987,76	1975,53	3951,06	7902,13	15804,26

Note It.	Note Ang.	Ottave									
		0	1	2	3	4	5	6	7	8	9
DO	**C**	16,05	32,10	64,21	128,43	256,86	513,73	1027,47	2054,94	4109,89	8219,79
DO#	**C is**	17,00	34,01	68,03	136,07	272,14	544,28	1088,57	2177,14	4354,28	8708,57
RE	**D**	18,02	36,04	72,08	144,16	288,32	576,65	1153,30	2306,60	4613,20	9226,41
RE#	**D is**	19,09	38,18	76,36	152,73	305,47	610,94	1221,88	2443,76	4887,52	9775,04
MI	**E**	20,22	40,45	80,90	161,81	323,63	647,26	1294,53	2589,07	5178,14	10356,29
FA	**F**	21,42	42,85	85,71	171,43	342,87	685,75	1371,51	2743,02	5486,05	10972,11
FA#	**F is**	22,70	45,440	90,81	181,63	363,26	726,53	1453,06	2906,13	5812,27	11624,55
SOL	**G**	24,05	48,10	96,21	192,43	384,86	769,73	1539,47	3078,94	6157,89	12315,78
SOL#	**G is**	25,48	50,96	101,93	203,87	407,75	815,50	1631,01	3262,02	6524,05	13048,11
LA	**A**	27	54	108	216	**432**	864	1728	3456	6912	13824
LA#	**A is**	28,60	57,21	114,42	228,84	457,68	915,37	1830,75	3661,50	7323,00	14646,01
SI	**B**	30,30	60,61	121,22	242,45	484,90	969,80	1939,61	3879,22	7758,45	15516,91

F sup. = f rif x $2^{n/12}$ (Per trovare frequenze superiori)
F inf.= f rif / $2^{n/12}$ (Per trovare frequenze inferiori)

- Fr = 440 Hz o 432 Hz
- 2 = n° fisso o costante (1Tono = 2 Semitoni)
- n = n° semitoni
- /12 = n° tot. dei semitoni (Temp. Equabile). L'ottava è divisa in 12 semitoni (7 diatonici e 5 cromatici).

$Fs_{La\#}$ = 440 x $2^{1/12}$ = 440 x 1,0594… = **466,16 Hz**

Fs_{Si} = 466,16 x $2^{1/12}$ = 466,16 x 1,0594… = **493,88 Hz** oppure 440 x $2^{2/12}$ = 440 x 1,1224… = **493,88 Hz**

$Fi_{Sol\#}$ = 440 / $2^{1/12}$ = 440 / 1,0594… = **415,30 Hz**