

# *Audio Digitale*

*Prof. Sebastiano Battiato*

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## ➤ **Il punto di vista fisico**

Natura del suono

Descrizione del segnale sonoro nel dominio del tempo e della frequenza

Parametri fisici del suono

## ➤ **Il punto di vista psicoacustico e cognitivo**

Percezione del suono

Parametri percettivi

Relazione tra parametri fisici e percettivi

Multimedia A.A. 2011/2012 – Prof. S. Battiato



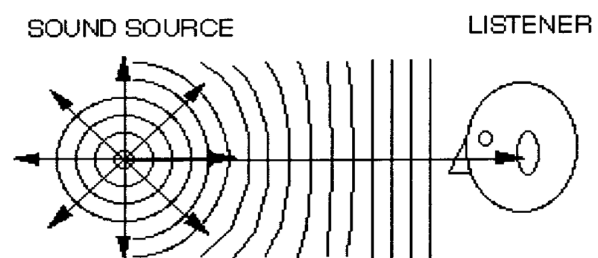
## ➤ La digitalizzazione del suono

Campionamento  
Quantizzazione  
Qualità del suono

## ➤ Elaborazione digitale e codifica

Strumenti SW per l'editing  
Operazioni di editing ed effetti: dominio del tempo e dominio della frequenza  
Compressione Audio

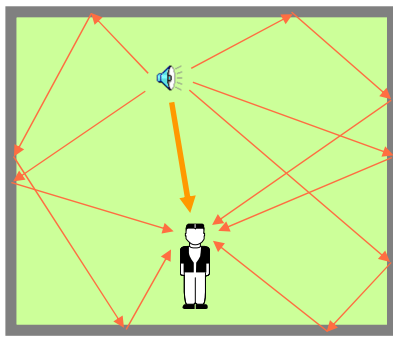
## Natura del suono



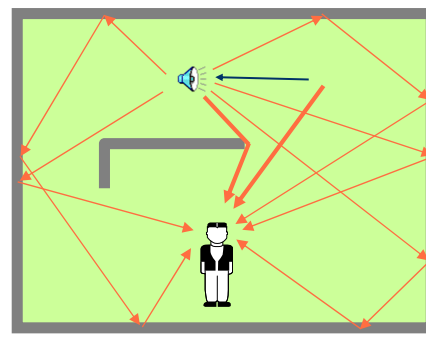
### Suono

Segnale di pressione - generato da una sorgente in vibrazione - che si propaga in un **mezzo elastico** (es. aria, acqua) fino ad un apparato sensoriale (es. orecchio umano dell'ascoltatore). Durante la propagazione si modifica – per assorbimento, riflessione, diffusione, ecc. - e si carica di “indizi” spaziali.

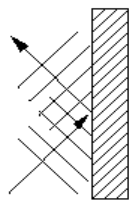
**Nel vuoto assoluto i suoni non possono propagarsi e quindi non si sentono i rumori**



Riverbero



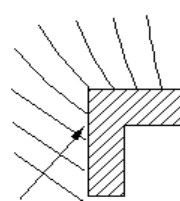
Occlusione



Riflessione



Diffusione



Diffrazione

Multimedia A.A. 2011/2012 – Prof. S. Battiato

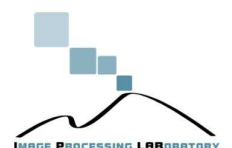


## La Natura del Suono

Come accennato la natura del suono è di tipo *ondulatorio*: si tratta di onde meccaniche che trasportano *energia* lontano dalla sorgente sonora. Viene quindi trasportato un *segnale*, cioè una variazione **continua** di qualche parametro legato all'ambiente in cui avviene la propagazione.

Il segnale sonoro è un'onda *longitudinale* poiché la sorgente sonora vibra nella stessa direzione di propagazione del suono.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Suono e Percezione

➤ La natura percettiva del suono è stata spiegata solo in parte. Alla fine dell'elaborazione effettuata dal nostro apparato uditivo è possibile percepire:

- ✓ **Musica**
- ✓ **Linguaggio**
- ✓ **Rumore**

## Il segnale sonoro: caratteristiche

Innanzitutto distinguiamo le tre caratteristiche fondamentali di ogni suono che sono l'**Altezza**, l'**Intensità** ed il **Timbro**. Poiché si tratta di grandezze percettive, per misurarle in modo esatto sono stati messi a punto altrettanti parametri di tipo fisico.

La “**vibrazione**” responsabile del suono può essere rappresentata come un'onda sinusoidale. E' possibile quindi mettere in relazione diretta i parametri percettivi con i relativi parametri fisici dell'onda.

<b>Parametro Percettivo</b>	<b>Parametro Fisico</b>	<b>Rappresenta</b>
<b>Altezza</b>	Frequenza	Tonalità Audio (bassi, medi, alti)
<b>Intensità</b>	Ampiezza	Volume
<b>Timbro</b>	Spettro	Tipologia di strumento

# Il segnale sonoro: caratteristiche (cont.)

Distinguiamo poi

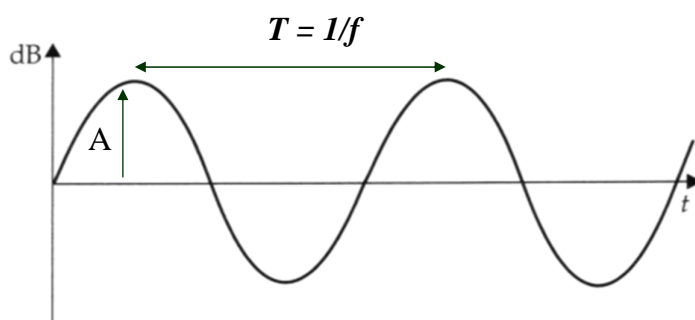
## ➤ Dominio del tempo

Come varia la pressione sonora nel tempo in corrispondenza di un determinato punto di ascolto;

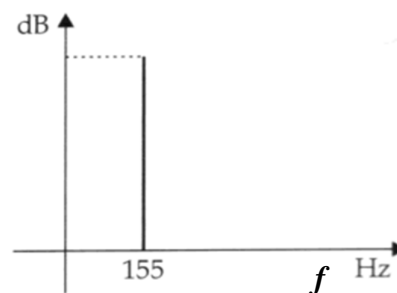
## ➤ Dominio della frequenza

Da quante e quali componenti elementari (**toni**) è composto il segnale sonoro;

## Suoni elementari: toni



*Dominio del tempo*



*Dominio della frequenza*

- Ampiezza (**A**) espressa in decibel **dB**;
- Periodo (**T**) espresso in secondi;
- Frequenza (**f**) numero di cicli (onde) al secondo; si misura in Hertz **Hz**

# La Frequenza

La frequenza si può definire come il numero di onde completate in un secondo.

La frequenza è il parametro che distingue tra loro le note musicali. A frequenze minori corrispondono i bassi, e in maniera crescente si hanno i toni medi e poi gli alti (o acuti), così come riportato in tabella:

Frequenza	Tipologia
<b>20-500 Hz</b>	Bassi
<b>500-8000 Hz</b>	Medi
<b>8000-20000 Hz</b>	Acuti

L'Hertz deriva il suo nome dal fisico tedesco Heinrich Rudolf Hertz (1857-1894)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## La Frequenza (cont.)

Le caratteristiche frequenziali inducono una differenziazione dei suoni in suoni *puri* e *complessi*.

Un suono *puro* (detto anche **tono**) è costituito da una sola frequenza ed è quindi descritto da un'onda sinusoidale semplice;

Un suono *complesso* consiste invece di più frequenze sommate in un'onda dall'andamento articolato; in un singolo periodo possono essere comprese più alternanze di compressioni e rarefazioni intermedie; l'ascolto rivela il timbro caratteristico di una sorgente. In generale in natura i suoni sono di tipo complesso, e lo specifico andamento deriva dal metodo di produzione del suono da parte della sorgente.

*Unica eccezione è il diaphason che riesce a produrre un suono sinusoidale quasi puro. I suoni puri, infatti, possono essere prodotti quasi esclusivamente in laboratorio mediante l'uso di oscillatori.*

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Frequenza e Note Musicali

I suoni sono segnali che hanno frequenze comprese all'incirca tra i 20 e i 20000 Hertz. Tali limiti derivano direttamente dal nostro sistema uditivo. Oltre tali valori si hanno gli *infrasuoni* e gli *ultrasuoni*.

Un suono complesso qualsiasi contiene molte frequenze. Affinchè in un suono si possa individuare una frequenza speciale, che caratterizza la sensazione globale di gravità/acutezza trasmessa dal suono, occorre che il segnale sia periodico. I suoni prodotti da strumenti musicali, hanno delle fasi di periodicità significative e per essi ha senso parlare della sensazione di altezza

## Note e Ottave

Si definisce ottava l'intervallo musicale tra due 'do' consecutivi. Dalla metà del settecento tutti gli strumenti ad intonazione fissa (pianoforte, organo,..) sono accordati secondo la scala temperata. Questa scala divide l'ottava in **12** intervalli ognuno dei quali vale un semitono. Ogni semitono corrisponde alla frequenza del semitono precedente moltiplicata per  $2^{1/12}$ , in modo tale che la prima nota dell'ottava seguente abbia frequenza doppia della nota corrispondente nell'ottava precedente.

# Note e Ottave

Note	Notazione Anglossassone	Frequenza (Hz)
la	A	$440.0 = 440 \times 2^{0/12}$
la#	A#	$466.2 = 440 \times 2^{1/12}$
si	B	$493.8 = 440 \times 2^{2/12}$
do	C	$523.2 = 440 \times 2^{3/12}$
do#	C#	$554.4 = 440 \times 2^{4/12}$
re	D	$587.3 = 440 \times 2^{5/12}$
re#	D#	$622.2 = 440 \times 2^{6/12}$
mi	E	$659.2 = 440 \times 2^{7/12}$
fa	F	$698.4 = 440 \times 2^{8/12}$
fa#	F#	$740.0 = 440 \times 2^{9/12}$
sol	G	$784.0 = 440 \times 2^{10/12}$
sol#	G#	$830.6 = 440 \times 2^{11/12}$
la	A	$880.0 = 440 \times 2^{12/12}$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Armoniche

La forma d'onda di un tono **non puro** dipende dalle “armoniche”. Le armoniche di un suono **puro** (ovvero di una segnale sinusoidale ad una certa frequenza che chiameremo **fondamentale**) sono i suoni (segnali) di frequenza multipla di quella fondamentale. Consideriamo ora un “do” e verifichiamo a cosa corrispondono le sue armoniche principali.

Do	Re	Mi	Fa	Sol	La	Si
261.6	294	329.6	349	392	440	494
523.2		659.2		784		
1046.4		1318.4		1568		

- seconda armonica  $261.6 \times 2 = 523.2$  **Do** dell'ottava successiva
- terza armonica  $261.6 \times 3 = 784.8$  **Sol** dell'ottava successiva
- quarta armonica  $261.6 \times 4 = 1046.4$  **Do** di due ottave sopra
- quinta armonica  $261.6 \times 5 = 1318$  **Mi** di due ottave sopra

Queste tre note “suonano bene assieme” e formano l'accordo di **Do-maggiore**.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# L'Ampiezza

L'ampiezza è il termine fisico con cui viene definita l'intensità di un suono, ossia l'energia trasportata da un'onda, ossia il suo volume.

Esistono due tipi di misura di solito utilizzati:

## Livello di pressione sonora (Sound Pressure Level)

$$\text{SPL} = 20 \log (p/p_0) \text{ db}$$

$p_0$ : pressione di riferimento. Rappresenta la *soglia minima di udibilità*. Corrisponde a 0,000025 Newton/m<sup>2</sup> ed è calcolata per una onda sinusoidale (tono puro) di 1000 Hz. La *soglia del dolore* corrisponde invece ad una pressione di circa 30 Newton/m<sup>2</sup>.

Nota: 0.5 kg sono circa 5 Newton.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# L'Ampiezza

- Poiché il rapporto tra un suono appena udibile e un suono alla soglia del dolore fisico è così elevato (circa un milione), conviene schiacciare in qualche modo la scala di riferimento.
- Si introduce quindi una scala logaritmica, basata sul rapporto (dovuta a Alexander Graham Bell) misurata in *decibel* ().
- Come noto, il logaritmo di un numero  $x$  in base  $b$  è l'esponente a cui deve essere elevata  $b$  per ottenere  $x$ , cioè  $\log_b x = a$  se  $x = b^a$ . Passare da  $a$  a  $a+1$  su una scala logaritmica significa passare da  $x = b^a$  a  $x = b^{a+1}$ . Se  $b = 10$  passare da  $x = 100$  a  $y = 1000$  corrisponde ad aumentare di 1 sulla scala logaritmica (lo stesso da 1000 a 10000).

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# L'Ampiezza

## Livello di Intensità sonora (Sound Intensity Level)

$$\text{SIL} = 10 \log (I/I_0) \text{ db}$$

Per intensità sonora si intende l'energia che passa attraverso una superficie unitaria nella unità di tempo.

$I_0$ : soglia minima di intensità sonora ( $10^{-12}$  Watt/m<sup>2</sup>). L'intensità  $I$  è proporzionale al quadrato della ampiezza  $A$  del suono.

---

**SPL** e **SIL** sono valori presi in relazione alle soglie di udibilità e quindi nella maggior parte dei suoni standard hanno valori identici. Da un punto di vista dell'interpretazione fisica i due valori sono diversi.

Le energie tipicamente in gioco nei fenomeni acustici sono irrilevanti rispetto alle energie in gioco nel fenomeno luminoso.

L'andamento percettivo del volume segue la scala logaritmica per cui, senza entrare nel dettaglio, basti sapere che ad un aumento di **10 dB** corrisponde un raddoppio dell'intensità sonora, cioè un raddoppio del volume.

# Livelli di intensità sonora

**Tabella 1.1** Livelli di intensità

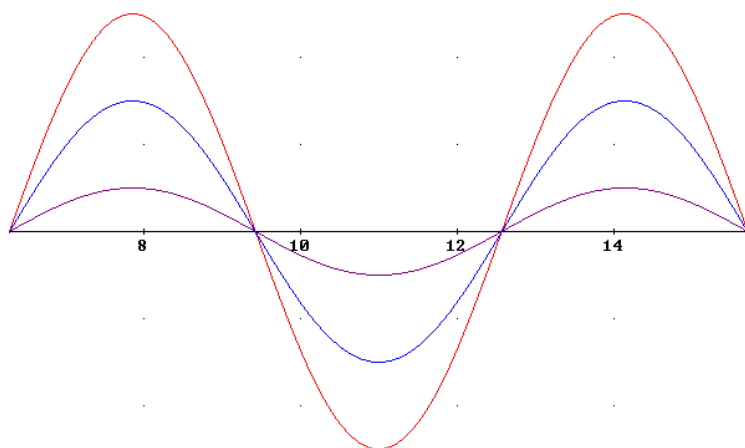
Suono	SPL (dB)	Reazione
Massimo rumore prodotto in laboratorio	210	Suono insopportabile
Lancio di un missile (a 50 m)	200	
Rottura del timpano	160	Dolore fisico
Jet al decollo (a 50 m)	130	
<b>Suono al limite del dolore</b>	<b>120</b>	
Complesso rock in locale chiuso	110	
Schianto di fulmine	110	
Urlo (a 1,5 m)	100	
Martello pneumatico (a 3 m)	90	Suoni utili
Traffico cittadino diurno	70-80	
Ufficio o ristorante (affollati)	60-65	
Conversazione (a 1 m)	50	
Teatro o chiesa (vuoti)	25-30	
Bisbiglio (a 1 m)	15	
Fruscio di foglie	10	
Zanzara vicino all'orecchio	10	Non udibile
Soglia dell'udito (a 1000 Hz)	0	

Tratto da: V. Lombardo e A. Valle. *Audio e Multimedia*. APOGEO (2005)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## L'Ampiezza (cont.)



Tre suoni aventi uguale frequenza ma ampiezza differente. In pratica, essi presentano la stessa altezza ma hanno un volume differente.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Il timbro (spettro)

Il timbro è il suono caratteristico di uno *strumento*. Una nota suonata da una pianoforte avrà un timbro differente rispetto alla stessa nota prodotta da un violino o da un flauto.

Il timbro è determinato dalle caratteristiche fisiche dello strumento, quali il mezzo utilizzato per produrre il suono (corde, pelle, ancia,...).

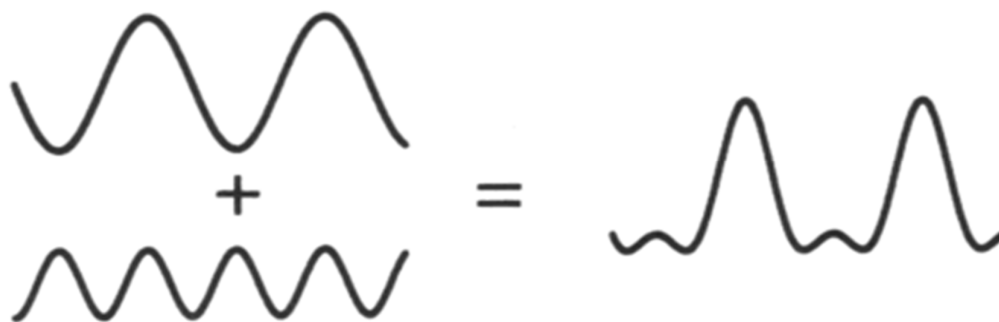
Generalmente il timbro viene riconosciuto più facilmente solo all'inizio della nota (*attacco*), mentre è più difficile da distinguere se il suono viene prolungato. Il timbro è legato alla forma d'onda. Gli strumenti musicali non emettono **toni puri** (ovvero sinusoidi perfette, che sarebbero anche sgradevoli) ma sono caratterizzati da forma d'onda molto differenti tra loro.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Suoni complessi

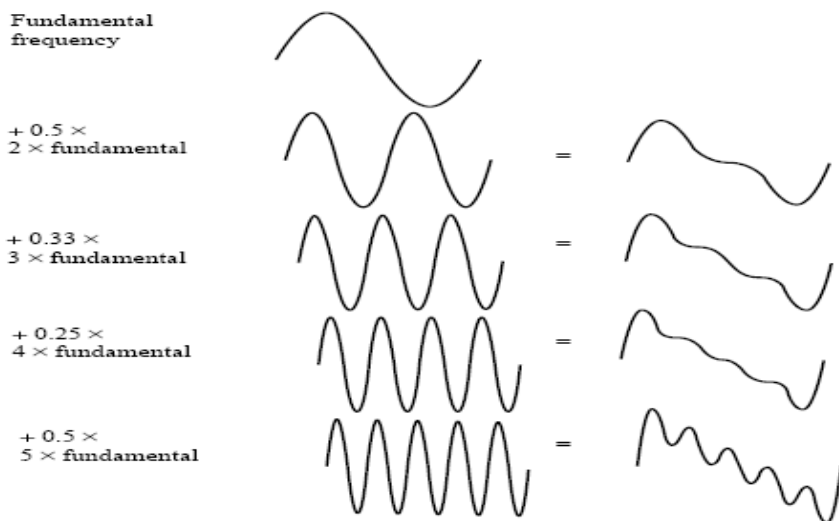
Un suono complesso è il risultato della sovrapposizione di più suoni elementari (**toni**) con frequenza, ampiezza e fase diverse



Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Suoni complessi

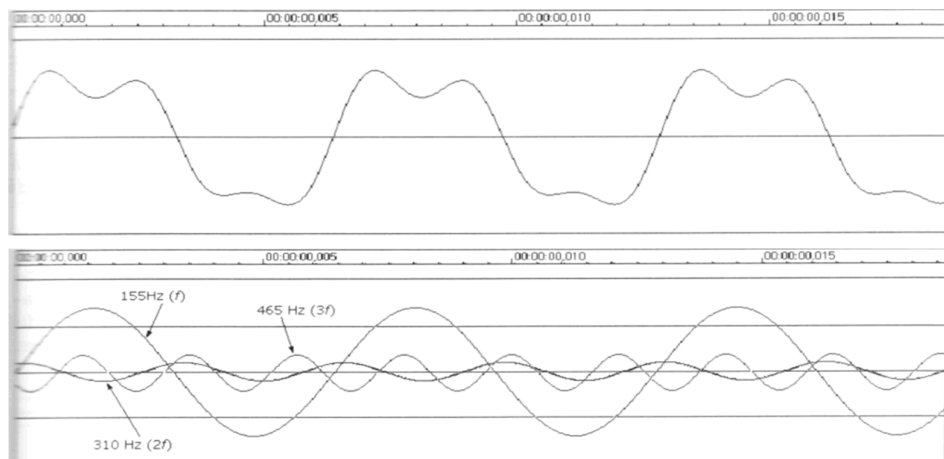


Un segnale complesso può essere costruito mediante un'opportuna somma pesata di "sinusoidi"

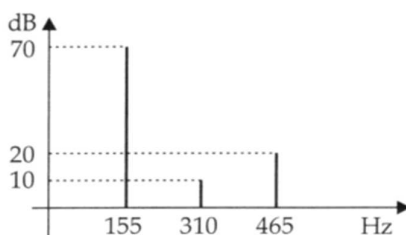
Multimedia A.A. 2011/2012 – Prof. S. Battiato



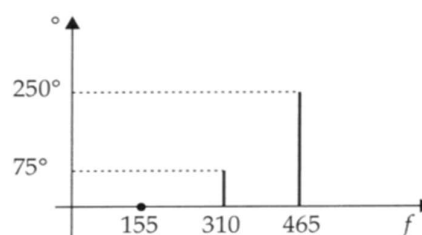
*Dominio del tempo*



*Dominio della frequenza*



a)



b)

Tratto da: V. Lombardo e A. Valle. *Audio e Multimedia*. APOGEO (2002)

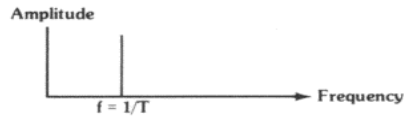
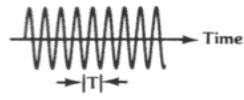
Multimedia A.A. 2011/2012 – Prof. S. Battiato



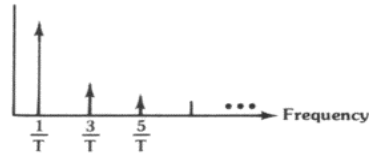
**Time Domain**

**Frequency Domain**

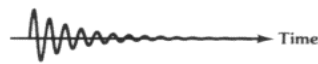
a) Sine Wave



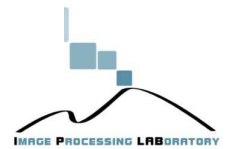
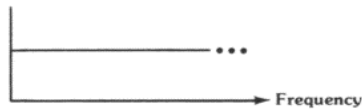
b) Square Wave



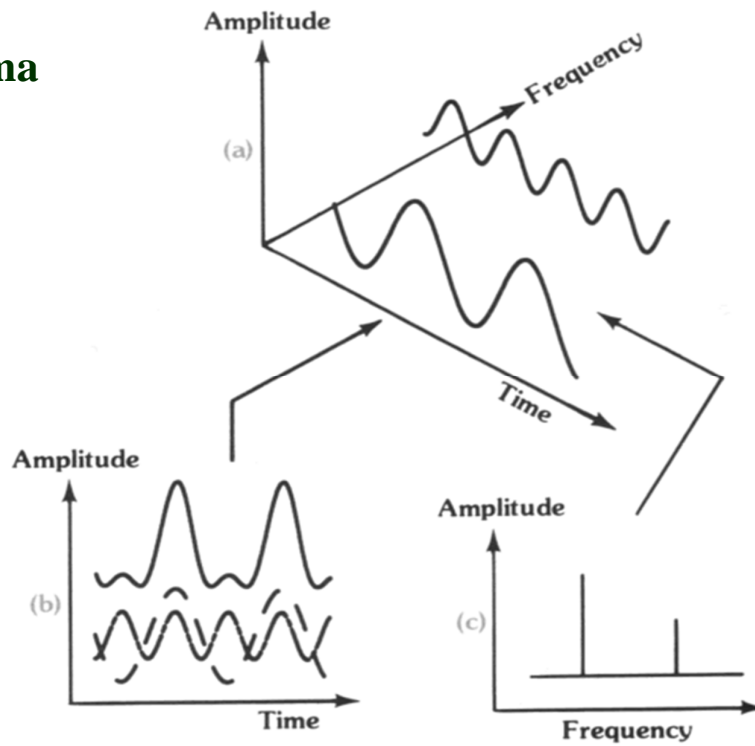
c) Transient



d) Impulse



**Sonogramma**



## Analisi di Fourier

- L'analisi di Fourier ed in particolare la decomposizione in serie di segnali periodici riveste un ruolo fondamentale nell'analisi e caratterizzazione dei suoni.
- Si “ascoltino” gli esempi a corredo del capitolo 1 del Testo Audio e Multimedia

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Analisi di Fourier

Una funzione **periodica** può essere espressa come somma di seni e/o coseni di differenti frequenze e ampiezze (*Serie di Fourier*).

Anche una funzione non periodica, (*sotto certe condizioni*) può essere espressa come integrale di seni e/o coseni, moltiplicati per opportune funzioni-peso (*Trasformata di Fourier*).

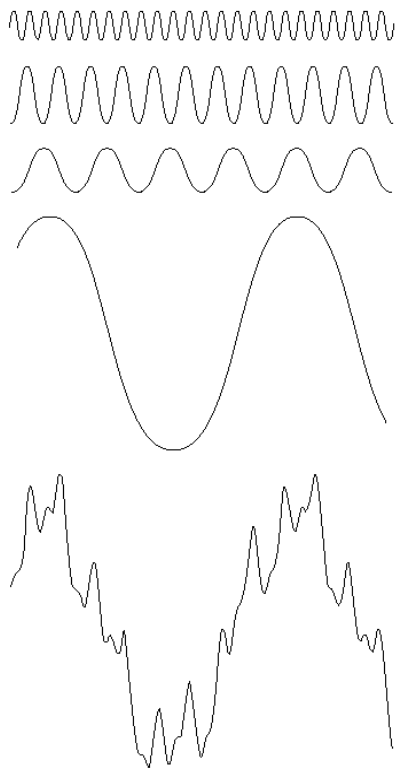


Jean Baptiste Joseph Fourier  
(Auxerre, 1768 – Paris, 1830)

Multimedia A.A. 2011/2012 – Prof. S. Battiato

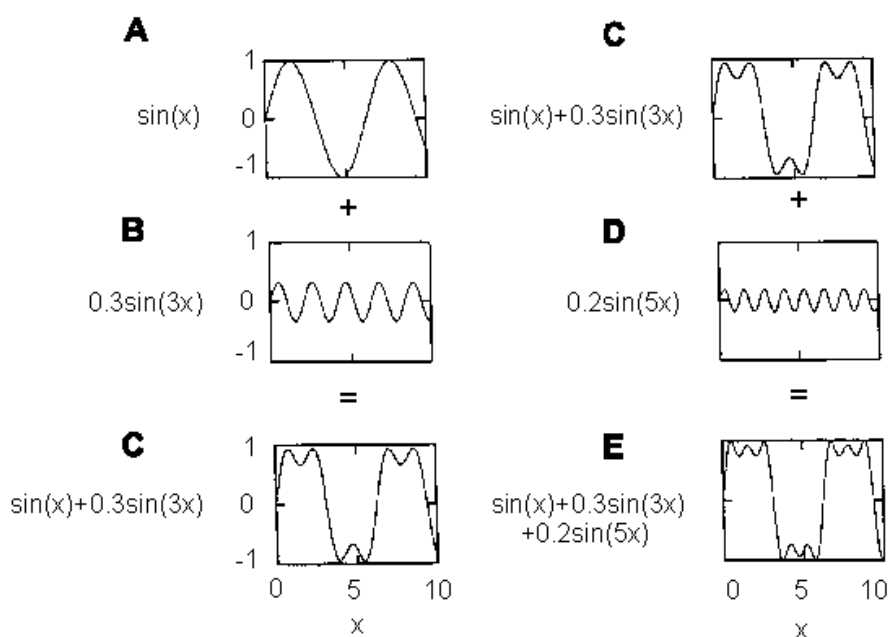


## Un primo esempio



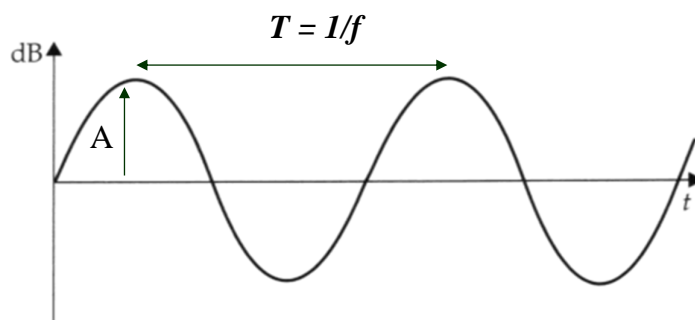
Questa funzione è la somma delle 4 funzioni periodiche di cui sopra.

## Ancora Esempi

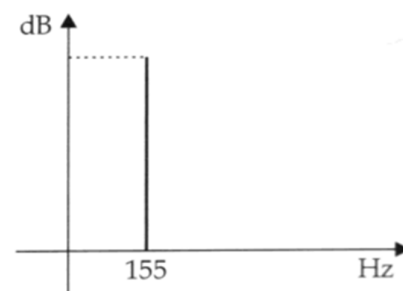


# Immagini e Segnali

- Un'immagine può essere vista come una funzione discreta in due dimensioni i cui valori rappresentano il livello di grigio di un determinato pixel.
- La funzione "immagine" può essere vista come un segnale, cioè una funzione variabile in un dominio con una propria frequenza (costante o variabile).

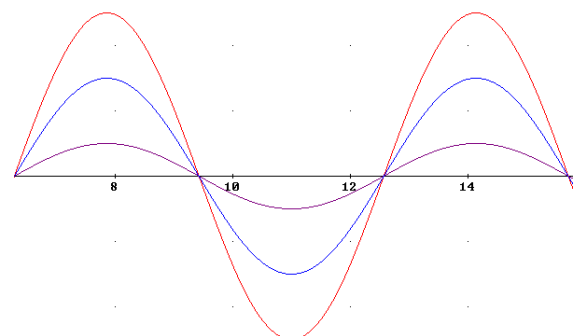


*Dominio del tempo*



*Dominio della frequenza*

- Ampiezza (**A**) espressa in decibel **dB**;
- Periodo (**T**) espresso in secondi;
- Frequenza (**f**) numero di cicli (onde) al secondo; si misura in Hertz **Hz**



## Preliminari

Cominciamo a studiare le funzioni periodiche da cui derivano le cosiddette “*armoniche*”, ossia le funzioni sinusoidali e cosinusoidali del tipo:

$$y = A \sin \omega x + \varphi$$

$$y = A \cos \omega x + \varphi$$

dove  $A$  indica l'ampiezza,  $\omega$  la pulsazione (definita come  $\omega = 2\pi/T$ , misurata in radianti al secondo) e  $\varphi$  la fase.

Sappiamo però che  $y = A \sin \omega x + \varphi = A \sin \varphi \cos \omega x + A \cos \varphi \sin \omega x$

quindi ponendo  $a = A \sin \varphi$  e  $b = A \cos \varphi$

otteniamo  $y = a \cos \omega x + b \sin \omega x$

Con tale sostituzione siamo riusciti a riscrivere la nostra funzione in termini di seno e coseno, riassumendo:

$$y = A \sin \omega x + \varphi = a \cos \omega x + b \sin \omega x$$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Preliminari

Analogamente

$$y = A \cos \omega x + \varphi = a \cos \omega x + b \sin \omega x$$

posto però  $a = A \cos \varphi$  e  $b = A \sin \varphi$

Quindi  $a$  e  $b$  sono le due nuove ampiezze e quel che è più interessante è l'assenza del coefficiente di fase nella nuova espressione.

Multimedia A.A. 2011/2012 – Prof. S. Battiato

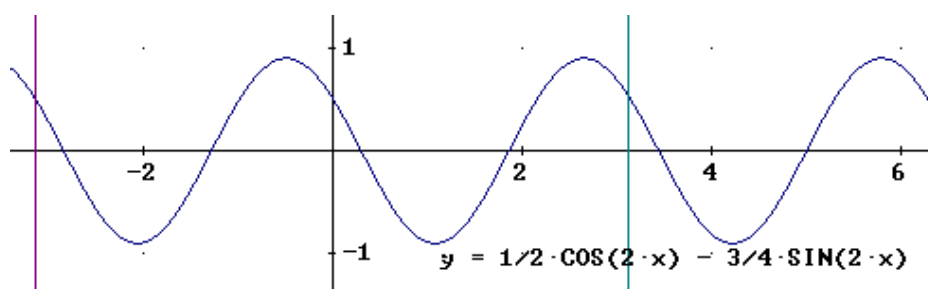


## Preliminari

Le funzioni sinusoidali di periodo  $2\pi$  possono essere espresse nella forma:

$$f(x) = a \cos \omega x + b \sin \omega x$$

Se la funzione sinusoidale è di periodo minimo  $T=2\pi$  abbiamo  $\omega=2\pi/2\pi=1$  cioè abbiamo un'oscillazione completa nell'intervallo  $2\pi$  mentre se  $T=\pi$  abbiamo esattamente 2 oscillazioni.



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Preliminari

In generale se la funzione  $f(x) = a \cos \omega x + b \sin \omega x$  ha come minimo periodo  $T = 2\pi/\omega$ , in un periodo  $2\pi$  si compiono  $\omega$  oscillazioni complete.

Se consideriamo la figura precedente si ha un periodo minimo  $T = \pi$  ma possiamo sempre considerarla come un'espressione di periodo  $2\pi$  giacchè compie in questo intervallo due oscillazioni esatte.

Tutto ciò vale in generale anche per la seguente espressione:

$$a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots + a_n \cos nx + b_n \sin nx$$

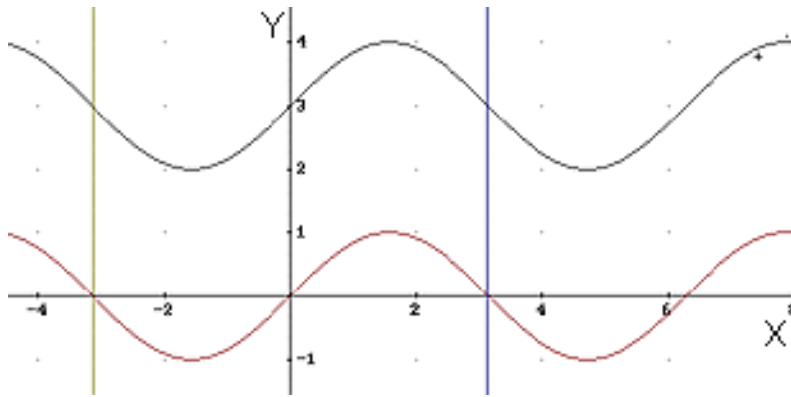
Infatti anche la somma di più funzioni sinusoidali con pulsazioni diverse da ancora una funzione periodica di periodo  $2\pi$  (**verificarlo sperimentalmente in Matlab per esercizio**)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Preliminari

La costante  $a_0 > 0$  ha il semplice effetto di spostare l'onda prodotta dalla sommatoria verso l'alto oppure verso il basso se  $a_0 < 0$  rispetto all'asse delle  $x$ . Nell'immagine consideriamo il contributo di una costante sommata a  $\sin x$ .



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Serie di Fourier

- Nel caso unidimensionale data una funzione variabile nel tempo  $f(x)$  e **periodica**, questa funzione può essere rappresentata come somma di infiniti termini sinusoidali (**armoniche elementari**), attraverso la **serie di Fourier**:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

- Lo sviluppo in serie di Fourier può avvenire purchè vengano soddisfatte le cosiddette **condizioni di Dirichlet**.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Le Armoniche

Prima di vedere le formule per trovare i coefficienti della serie di Fourier diamo la definizione di armonica.

La funzione  $a_1 \cos x + b_1 \sin x$  viene detta **prima armonica o armonica fondamentale** della funzione  $f(x)$ .

L'armonica fondamentale, come vedremo, ha frequenza minima rispetto alle armoniche di ordine superiore ed è quella che dà il maggiore contributo nella costruzione dell'onda risultante della serie.

Chiamiamo invece la funzione  $a_k \cos kx + b_k \sin kx$  la  **$k$ -esima armonica o armonica di ordine  $k$**  della  $f(x)$ .

**Esercizio:** Trovare la formula per il calcolo dell'ampiezza della  $k$ -armonica.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Ampiezza $k$ -esima armonica

Partendo da  $y = A \sin \omega x + \varphi$  con considerazioni analoghe a quelle fatte precedentemente otteniamo:

$$A_k \sin \varphi_k = a_k; \quad A_k \cos \varphi_k = b_k;$$

Elevando al quadrato e sommando entrambi i membri si ha:

$$A_k^2 \sin^2 \varphi_k + A_k^2 \cos^2 \varphi_k = a_k^2 + b_k^2$$

$$A_k^2 (\sin^2 \varphi_k + \cos^2 \varphi_k) = A_k^2 (1)$$

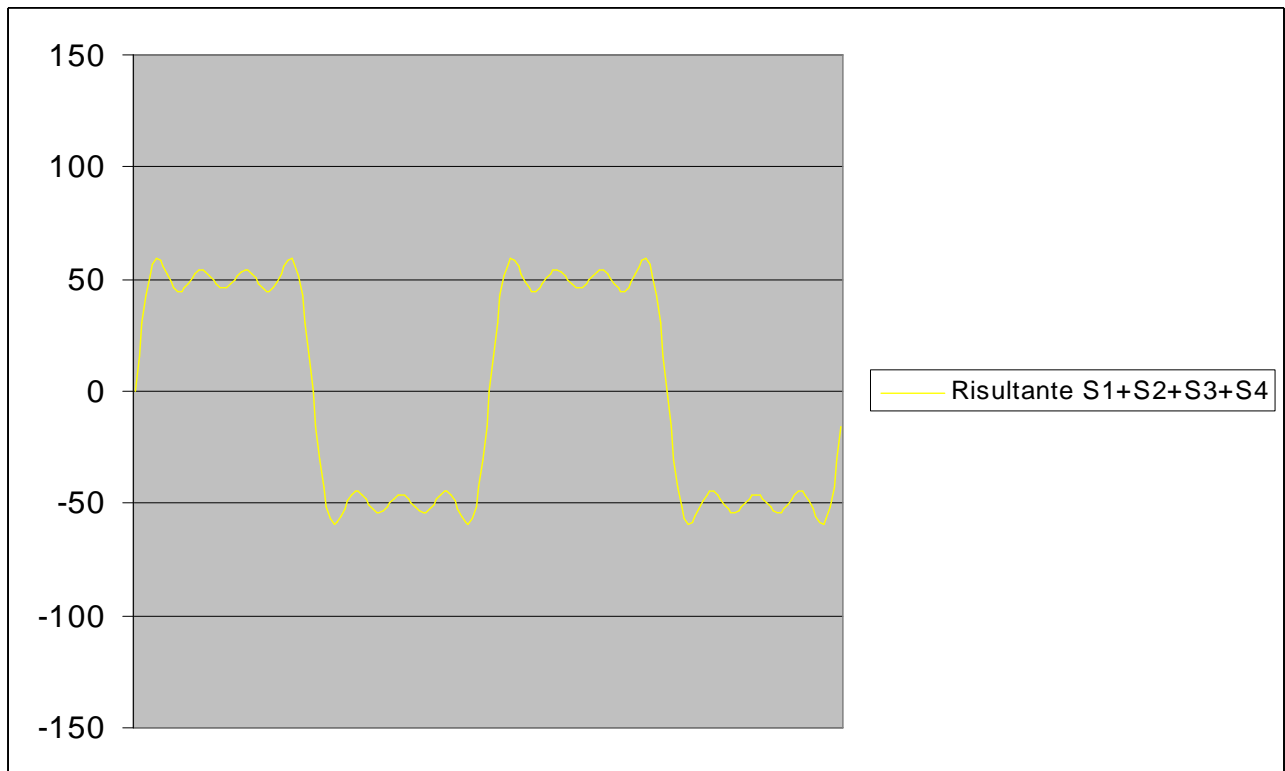
da cui:

$$A_k = \sqrt{a_k^2 + b_k^2}$$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Sintesi di forme d'onda – es. con onda quadra

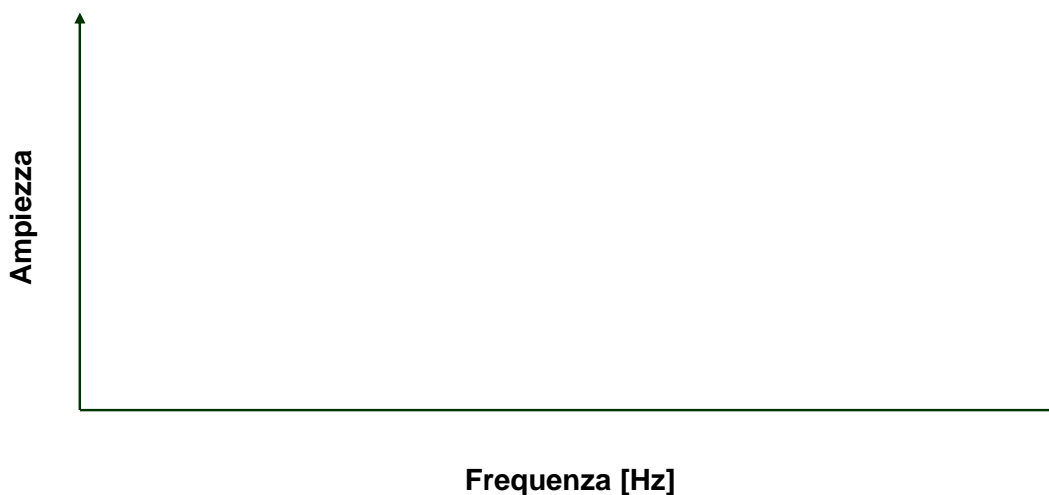


Multimedia A.A. 2011/2012 – Prof. S. Battiato

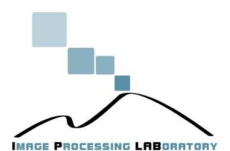


## Analisi di forme d'onda – lo spettro

Le componenti sinusoidali di un'onda possono essere rappresentate in un grafico, ciascuna come una barra di altezza pari *all'ampiezza* del senoide corrispondente e ascissa pari alla sua *frequenza*.

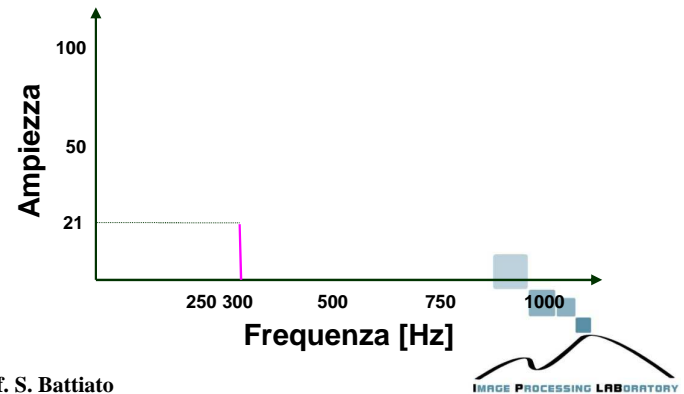
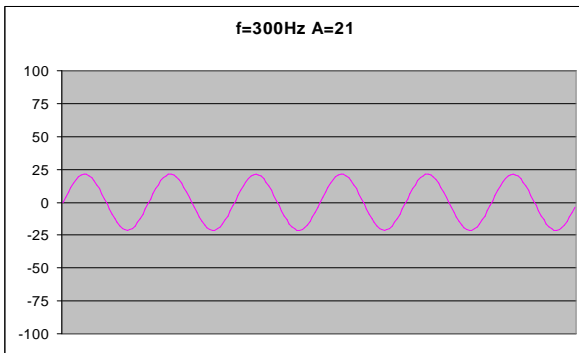
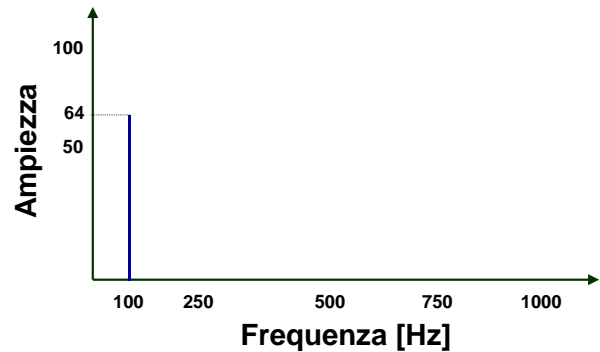
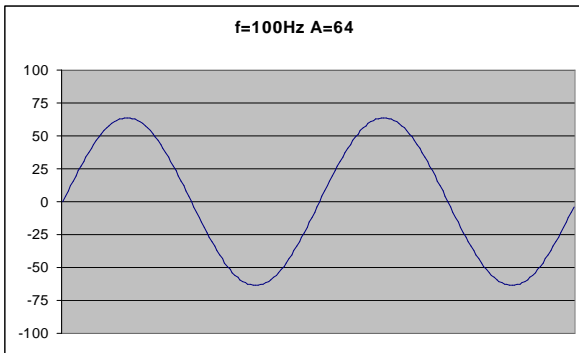


Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Analisi di forme d'onda – lo spettro

## Spettri di onde sinusoidali

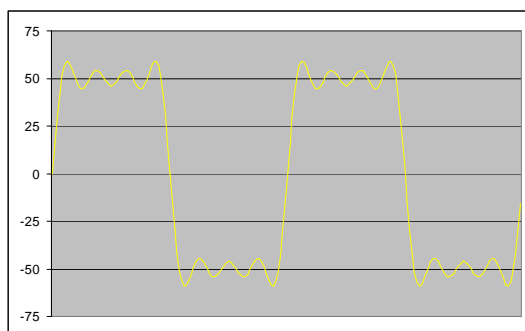


Multimedia A.A. 2011/2012 – Prof. S. Battiato

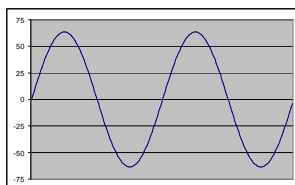
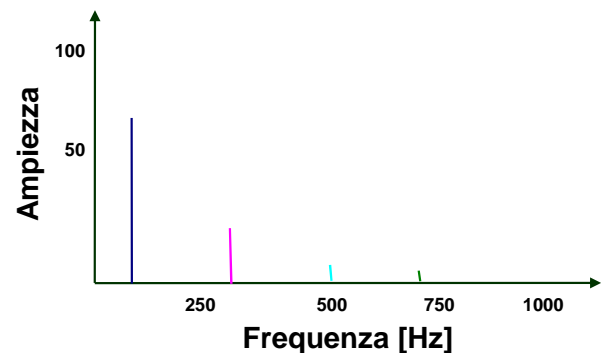


# Analisi di forme d'onda – lo spettro

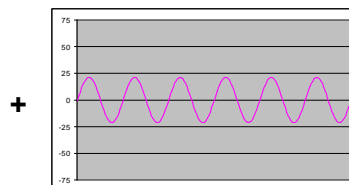
Forme d'onda complesse: lo spettro è la somma degli spettri dei sinusoidi che compongono l'onda. Il segnale è una somma di sinusoidi di frequenza multiple intere della frequenza del segnale ( $f_0$ ).



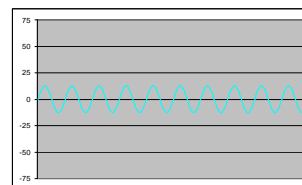
f = 100 Hz



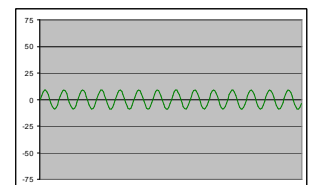
f=100Hz A=64



f=300Hz A=21



f=500Hz A=6

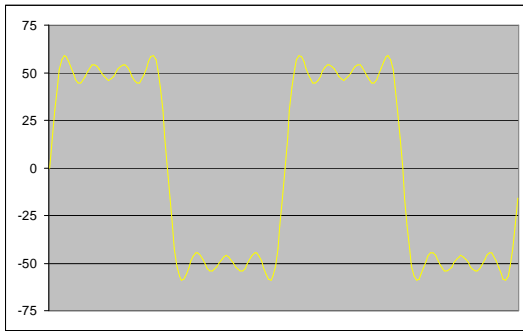


f=700Hz A=4

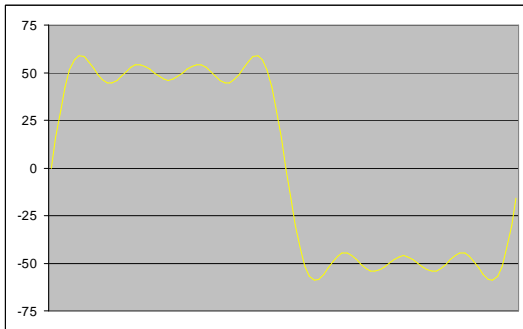
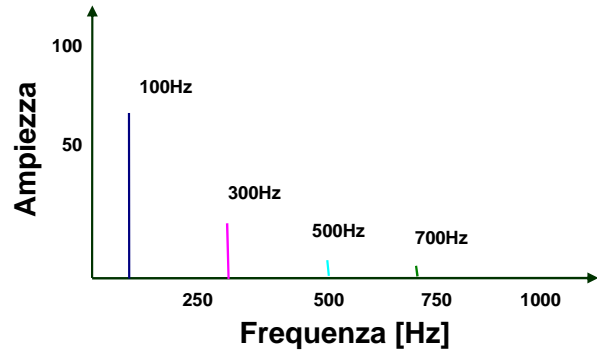
Multimedia A.A. 2011/2012 – Prof. S. Battiato



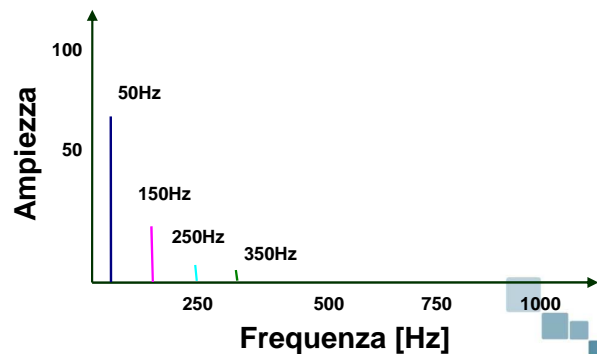
# Analisi di forme d'onda – lo spettro



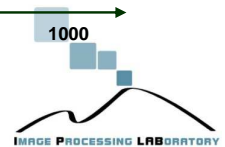
f=100Hz



f=50Hz



Multimedia A.A. 2011/2012 – Prof. S. Battiato

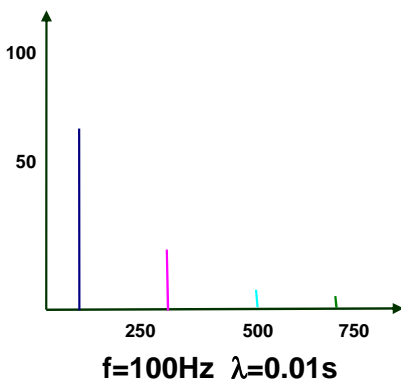


# Analisi di forme d'onda – lo spettro

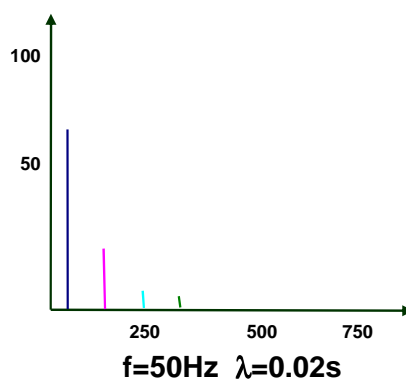
Aumentando la lunghezza d'onda di un segnale (il suo periodo), quindi diminuendo la sua frequenza, le barre dello spettro tendono a spostarsi verso l'origine degli assi ed ad avvicinarsi le une alle altre.

Intuitivamente, possiamo immaginare che se la lunghezza d'onda diventa infinita (ossia il segnale si ripete in un periodo infinito: è **non-periodico**), le barre dello spettro si fondono in una linea continua.

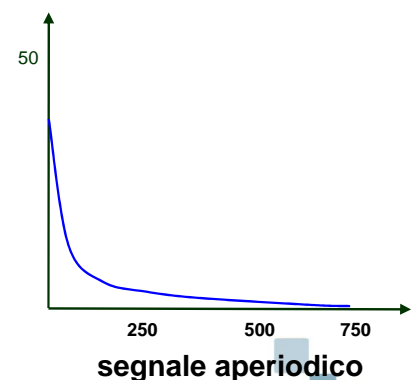
100



f=100Hz  $\lambda=0.01s$



f=50Hz  $\lambda=0.02s$

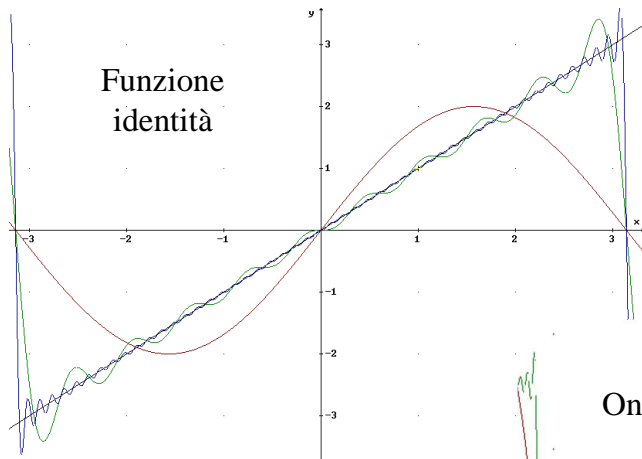


segnale aperiodico

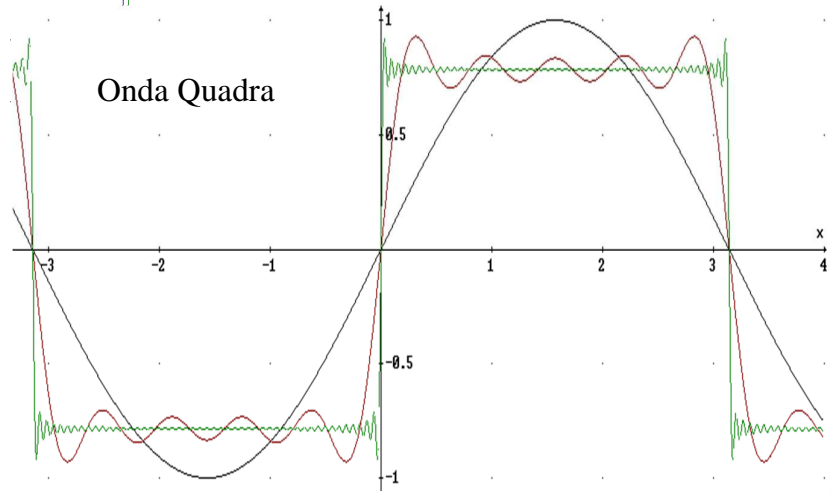
Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Esempi



<http://www.falstad.com/fourier/j2/>  
<http://www.falstad.com/dfilter/>



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Serie di Fourier: Calcolo dei Coefficienti

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(x) dx$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(x) \cos kx dx \text{ per } k = 1, 2, 3, \dots$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(x) \sin kx dx \text{ per } k = 1, 2, 3, \dots$$

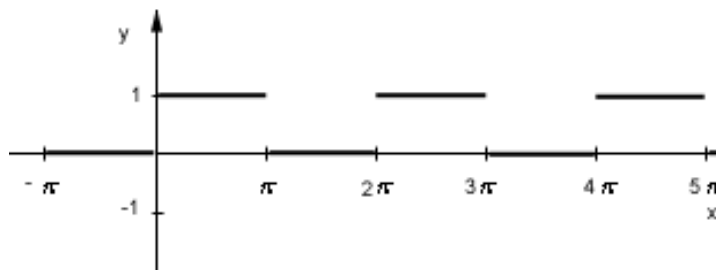
Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Esercizi

**Esercizio** : Trovare i coefficienti  $a_0$ ,  $a_k$ ,  $b_k$  per la seguente funzione:

$$f(x) = \begin{cases} 0 & \text{se } (2k-1)\pi \leq x < 2k\pi \\ 1 & \text{se } 2k\pi \leq x < (2k+1)\pi \end{cases}$$



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Esercizi

Esprimere mediante la serie di Fourier le seguenti funzioni elementari:

$$y=f(x)=\cos x \text{ (Utilizzare le formule di Werner)}$$

$$y=f(x)=x$$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



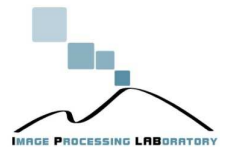
## Serie di Fourier: Forma Esponenziale

$$\begin{aligned} e^{ix} &= \cos x + i \sin x \\ \cos kx &= \frac{e^{ikx} + e^{-ikx}}{2} \\ \sin kx &= \frac{e^{ikx} - e^{-ikx}}{2i} = \frac{i e^{ikx} - e^{-ikx}}{2i} = -i \frac{e^{ikx} - e^{-ikx}}{2} \end{aligned}$$

Utilizzando le formule di **Eulero** (di cui sopra) nella serie di Fourier, otteniamo:

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \frac{e^{ikx} + e^{-ikx}}{2} - ib_k \frac{e^{ikx} - e^{-ikx}}{2} \right) = \\ &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( \frac{a_k e^{ikx}}{2} + \frac{a_k e^{-ikx}}{2} - \frac{ib_k e^{ikx}}{2} + \frac{ib_k e^{-ikx}}{2} \right) = \\ &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( \frac{a_k - ib_k}{2} e^{ikx} + \frac{a_k + ib_k}{2} e^{-ikx} \right) \end{aligned}$$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Serie di Fourier: Forma Esponenziale

$$f(x) = \frac{a_0}{2} \sum_{k=1}^{\infty} \left( \frac{a_k - ib_k}{2} e^{ikx} + \frac{a_k + ib_k}{2} e^{-ikx} \right)$$

ponendo:

$$\frac{a_0}{2} = c_0, \quad \frac{a_k - ib_k}{2} = c_k, \quad \frac{a_k + ib_k}{2} = c_{-k}$$

otteniamo:

$$f(x) = c_0 + \sum_{k=1}^{\infty} \left( c_k e^{ikx} + c_{-k} e^{-ikx} \right) = \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$$

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Suoni Complessi

- Si utilizza la DFT (Discrete Fourier Transform) sulla rappresentazione digitale del suono fissando di volta in volta una finestra temporale di  $N$  di campioni di un segnale digitale  $x[N]$ . La funzione in output  $X[k]$  ottenuta dalla DFT rappresenta  $N$  campioni di uno spettro di frequenze continuo.
- La risoluzione in frequenza è maggiore al crescere dei campioni utilizzati dalla DFT.



## Ancora DFT

Maggiore è il numero di punti, migliore è la risoluzione in frequenza, perchè, detto  $n$  il numero di punti su cui si effettua l'analisi e  $sr$  la frequenza di campionamento, la larghezza di banda  $bw$  di ogni canale di analisi è data dalla formula:

$$bw = sr/n$$



# Esempio Risoluzione in frequenza

Per una frequenza di campionamento di 44100 Hz ed un numero di punti pari a 256, la larghezza di banda sarà  $44100/256 = 172.27$  Hz

Questo equivale a dire che, se facciamo un'analisi di un suono campionato a 44100 Hz con un numero di punti pari a 256, lo spettro risultante conterrà le ampiezze delle seguenti bande di frequenza:

172.27, 344.5, 516.8, 689, 861.3, 1033.6, ...

## Rumore

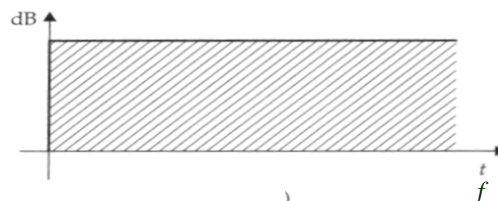
Esistono varie tipologie di rumore. Tra quelli a banda larga i più importanti sono:

**Rumore termico.** Questo rumore è generato dal calore insito in qualsiasi componente elettronico. Il calore fa sì che all'interno del componente si verifichino delle collisioni di elettroni in tutte le direzioni e a tutte le velocità generando delle correnti a tutte le frequenze. *Le ampiezze di queste frequenze ossia le intensità delle correnti sono mediamente costanti in quanto la direzione delle collisioni è assolutamente casuale.* Il rumore termico aumenta con la temperatura in quanto aumenta con essa l'energia cinetica associata alle particelle.

# Rumore



*Dominio del tempo*



*Dominio della frequenza*

**Rumore bianco.** Si intende con questa dicitura un rumore di ampiezza costante su tutto lo spettro di frequenza. In sostanza si tratta di un rumore termico solo che in questo caso si intende un rumore appositamente generato con finalità di test. Per vedere infatti il comportamento di un componente audio, per esempio di un canale di un mixer, si invia in ingresso un rumore bianco e si esamina il segnale di uscita. Generalmente in questo caso l'obiettivo sarà quello di ottenere un segnale in uscita mediamente costante a tutte le frequenze, questo significherà che il componente è affidabile a tutte le frequenze. In generale il rumore bianco viene usato per i test sui componenti elettronici.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Inviluppo di un suono

Il modo in cui un suono evolve nel tempo rispetto alla sua ampiezza è detto *inviluppo*.

Per introdurre questo concetto conviene considerare un esempio pratico. Il più eloquente è quello di uno strumento a corda, per esempio una chitarra. Quando il chitarrista esegue una nota, la sentiamo scoccare quasi, poi piano piano la nota si estingue. L'andamento dell'ampiezza della nota suonata viene chiamato *inviluppo ADRS* (acronimo delle parole: **Attack**, **Decay**, **Sustain**, **Release**) e ha un preciso schema che può essere applicato a qualsiasi suono e strumento.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



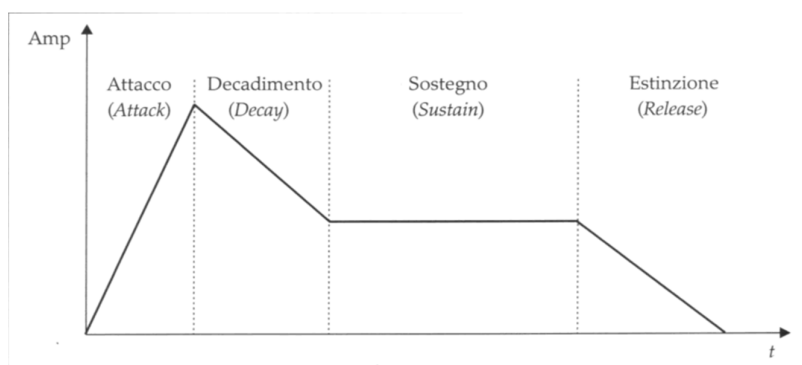
# Inviluppo di un suono (1)

**Attack:** L'ampiezza raggiunge molto rapidamente il massimo

**Decay:** Dopo l'attacco, parte dell'energia iniziale viene persa e l'ampiezza diminuisce.

**Sustain:** L'ampiezza mantiene un livello quasi costante per un certo tempo.

**Release:** L'ampiezza ricomincia a diminuire fino ad annullarsi.



Multimedia A.A. 2011/2012 – Prof. S. Battiato



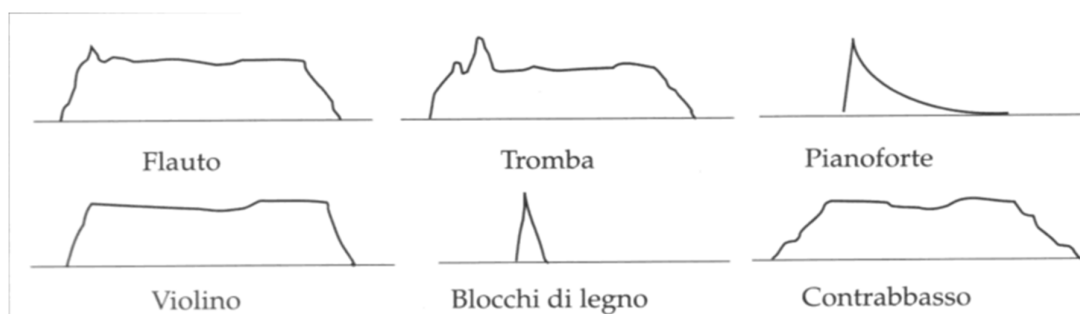
# Inviluppo di un suono (2)

**Attack:** L'ampiezza raggiunge molto rapidamente il massimo

**Decay:** Dopo l'attacco, parte dell'energia iniziale viene persa e l'ampiezza diminuisce.

**Sustain:** L'ampiezza mantiene un livello quasi costante per un certo tempo.

**Release:** L'ampiezza ricomincia a diminuire fino ad annullarsi.



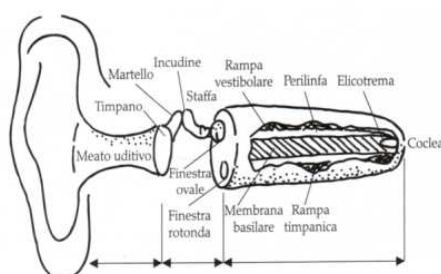
Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Percezione del suono

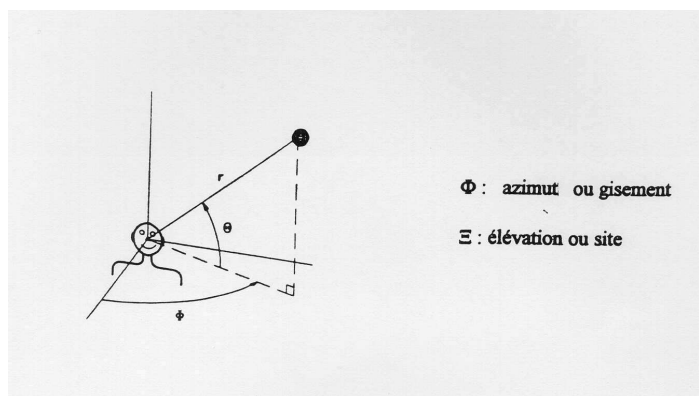
- In risposta al segnale di pressione il timpano dell'orecchio si muove avanti e indietro; l'energia meccanica del timpano si trasforma in impulsi nervosi (elettro-chimici) che vengono elaborati dal cervello e danno origine alla esperienza uditiva.
- Processi coinvolti
  - Sensazione;
  - Percezione;
  - Concezione/cognizione.

## L'orecchio



# Descrizione del percetto

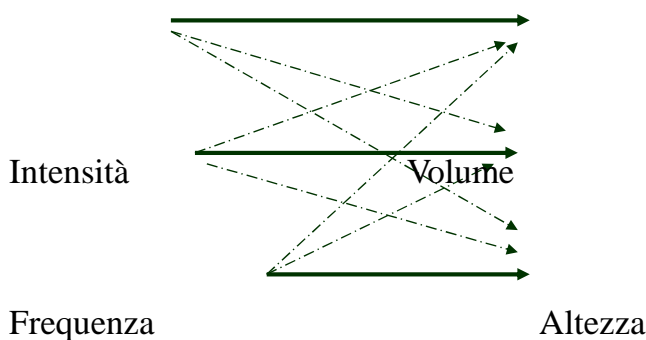
- *Altezza* (pitch): suono acuto, grave
- *Volume* (loudness): suono forte, debole
- *Timbro*: colore del suono (caratterizza la identità della sorgente)
- *Localizzazione* della sorgente: direzione (azimut ed elevazione) e distanza
- *Durata* (percepita)
- ...



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Relazioni tra parametri fisici e percettivi



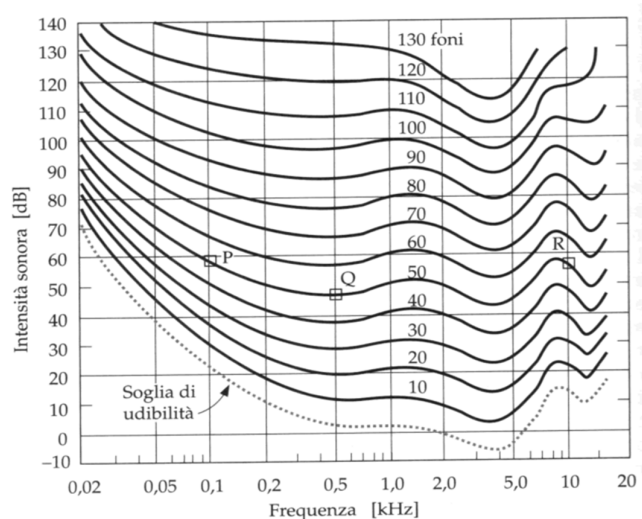
Le proprietà fisiche di ampiezza e frequenza corrispondono alle caratteristiche percettive di volume e tono. Tuttavia il legame tra queste grandezze non è costante. Inoltre, anche nei limiti di quelli che sono i suoni da noi percepiti, la relazione tra proprietà percepite e proprietà fisiche non è una relazione lineare: per esempio, aumentando l'ampiezza di una forma d'onda di una uguale grandezza, non si ottengono uguali incrementi di volume (il volume sembra aumentare di meno via via che diventa più elevato). Analoga caratteristica vale per la frequenza: ad aumenti uguali di frequenza non corrispondono uguali incrementi di tono (l'incremento di tono sembra via via più piccolo col crescere della frequenza).

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Diagramma di Fletcher e Munson

Le curve **isofone** rappresentano suoni percepiti con lo stesso volume. Il volume percepito (in **fon**) dipende dalla intensità e dalla frequenza



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## L'ascolto

### ➤ Tipi di suoni:

*Voce;*

*Effetti/rumore;*

*Musica;*

*Strati di suoni (Quanti? Di che tipo? Compromesso: chiarezza/densità).*

### ➤ Tipo (intento) di ascolto

*Causale:* il suono è un indice della causa che lo ha prodotto;

*Semantico:* il suono è portatore di un messaggio che comprendo;

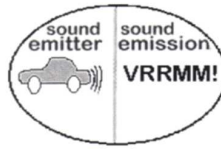
*Ridotto:* il suono in se (oggetto sonoro).

Multimedia A.A. 2011/2012 – Prof. S. Battiato

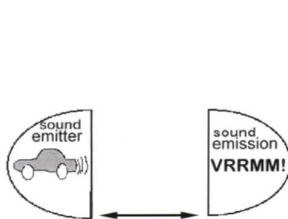


# L'illusione audiovisiva

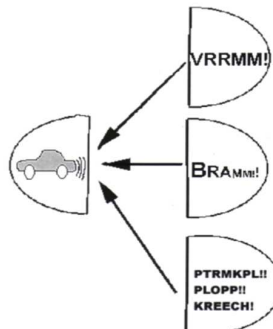
- Il suono nella realtà



- Il suono nelle produzioni multimediali



*Registrazione*



*Elaborazione*



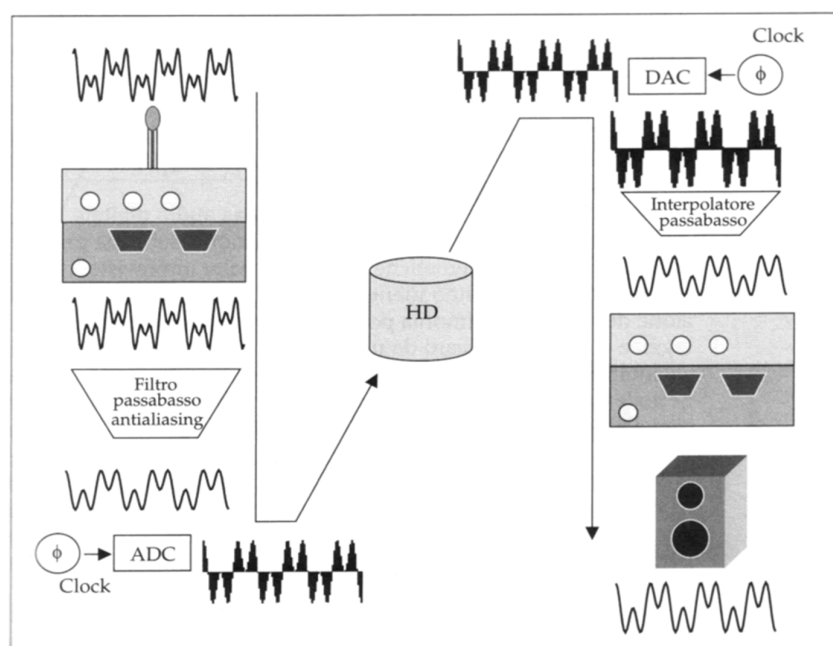
*Sincresi= Sincronismo+Sintesi*



# Il suono digitale



# Digitalizzazione del suono



Tratto da: V. Lombardo e A. Valle. *Audio e Multimedia*. APOGEO (2002)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Digitalizzazione del suono

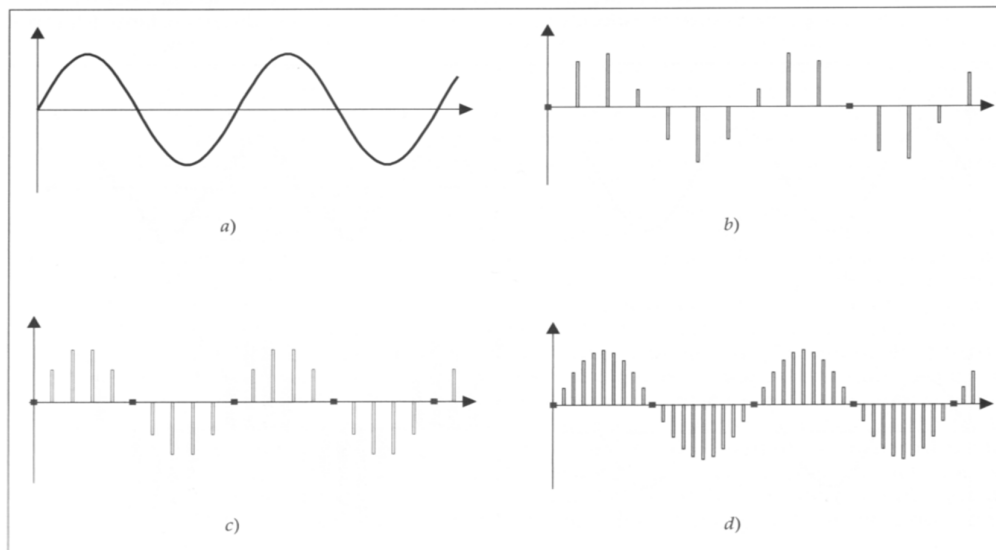
- I microfoni producono rappresentazioni analogiche del segnale audio. Questo è infatti rappresentato da un valore di tensione il cui andamento nel tempo riflette le oscillazioni di pressione nell'aria. Nel caso dei dischi in vinile o nel campo magnetico la curva continua nel tempo delle variazioni di ampiezza viene rappresentata da una curva continua nel tempo delle variazioni di tensioni elettriche ed è memorizzata nei solchi del disco o nel campo magnetico di un nastro.
- Per poter rappresentare il suono in un sistema digitale bisogna prima convertirlo in un flusso di numeri rappresentati in forma binaria. Una rappresentazione digitale assegna dei numeri

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Campionamento

E' la discretizzazione del segnale analogico nel tempo. La conversione del suono da formato analogico a digitale avviene per mezzo di una scheda di acquisizione (o digitalizzazione) che *campiona* il valore della forma d'onda ad intervalli regolari.



Multimedia A.A. 2011/2012 – Prof. S. Battiato



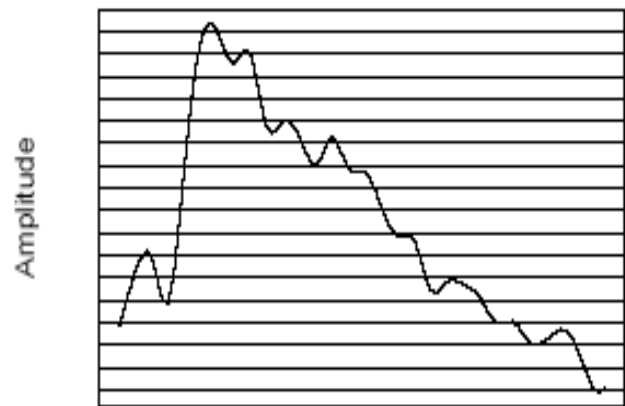
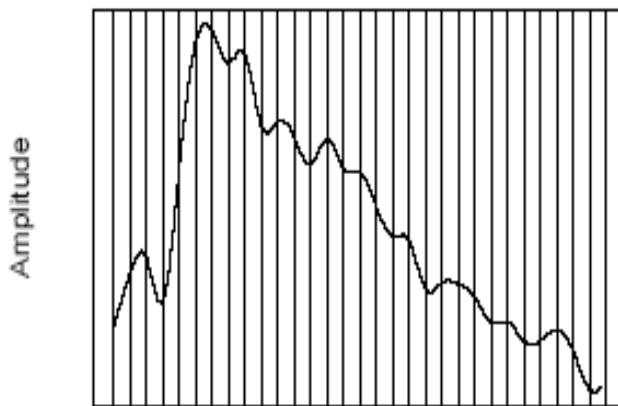
# Quantizzazione

- E' la discretizzazione della ampiezza. L'ampiezza di ogni campione, dovendo essere rappresentata digitalmente (cioè con una codifica binaria), non può assumere infiniti valori.
- Ogni singolo campione di ampiezza (tensione elettrica) viene quindi assegnato ad uno dei valori numerici che sono consentiti dalla codifica digitale (si commettono errori di **quantizzazione**).
- La conversione Analogico-Digitale richiede pertanto un processo di discretizzazione sia nel tempo (**campionamento**) che in ampiezza (**quantizzazione**).

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Campionamento e Quantizzazione



Per la conversione Analogico-Digitale è quindi necessaria la specificazione di 2 parametri relativi a:

- Quanto spesso campionare il segnale nel tempo (**Frequenza di campionamento**)
- Con quanti valori rappresentare ogni campione (**Precisione di quantizzazione**)

## Precisione di Quantizzazione

Con qualsiasi rappresentazione analogica una parte del segnale impiegato per rappresentare la grandezza è dovuta al rumore. Un tipico rumore che tutti sperimentiamo è quello causato dall'impressione magnetica sul nastro e viceversa dalla lettura del segnale registrato; tale rumore viene percepito come fruscio dal nostro apparato uditivo.

Per la riduzione del rumore sono stati sviluppati vari metodi: il noto sistema della Dolby, ad esempio, enfatizza in registrazione alcune regioni dello spettro nelle quali il rumore è maggiormente percepibile; in riproduzione, le stesse regioni vengono de-enfatizzate, con il risultato di riportare ai livelli corretti i rapporti fra le frequenze nel segnale e attenuare nel contempo il rumore.

# Precisione di Quantizzazione

Per stimare l'ammontare di rumore introdotto da un sistema analogico si utilizza il *signal-to-noise ratio* (**SNR**), cioè il rapporto tra la massima ampiezza utile del segnale e l'ampiezza del rumore presente (statico o bianco) sovrapposto al segnale. Viene anche definito come rapporto tra la potenza del segnale e quella del rumore. A valori alti di SNR, che si misura in **db** corrisponde una migliore qualità del suono.

I CD Audio hanno un valore teorico di SNR ratio di circa 96 decibels (dB) (circa 90 db in pratica). Valori inferiori a 70 db indicano un rumore di fondo udibile. Ridurre la quantizzazione a 8-bits, riducendo del 50% la quantità di dati farebbe diminuire la qualità di un CD audio di circa 50 db, producendo una qualità simile a quella della radio AM. **(Approssimativamente ad ogni bit corrispondono 16 db)**

## Dinamica

Molto semplicemente rappresenta la capacità di graduare in modo nitido l'intensità del suono (nel contesto complessivo) riproducendo nel giusto rapporto i picchi di intensità, i suoni di basso livello, e tutti i suoni la cui intensità è compresa tra i due estremi. Ad una maggiore profondità di bit corrisponde la possibilità di registrare e/o riprodurre una maggiore dinamica.

La gamma dinamica è il rapporto tra l'ampiezza massima e l'ampiezza minima presenti nel segnale.

**Esempio:** i brani di musica classica hanno un'alta dinamica, mentre la musica techno ha una dinamica molto limitata

# Teorema di Nyquist

Per avere una digitalizzazione senza perdita di informazione è necessario campionare con una frequenza almeno il doppio della massima frequenza che compare nello spettro della forma d'onda da acquisire:

**Il tasso di campionamento  $f_c$  deve essere almeno il doppio della frequenza massima  $f_{max}$  presente nel segnale**

$$f_c \geq 2 f_{max}$$

**Esempio:**

per la voce (al telefono)  $f_{max} = 3600-4000\text{Hz}$ , pertanto  $f_c = 8000\text{ Hz}$

# Teorema di Nyquist

- Se l'orecchio è in grado di captare suoni, fino a 20000Hz, occorre campionare ad almeno 40000Hz. In realtà la frequenza di campionamento standard attuale è pari a 44.100 Hz per almeno due ragioni:
  - il valore dei 20.000 Hz è un valore medio; fissando una frequenza di campionamento standard, leggermente superiore ci si è assicurata la massima fedeltà;
  - Nei primi anni '70, i supporti magnetici utilizzati impedivano comunque di oltrepassare questo limite.
- Al di sopra di tale soglia si ha il **sovracampionamento** che può portare solitamente ad uno spreco di banda
- Al di sotto si ha il **sottocampionamento** che spesso genera disturbi e distorsioni quali ad esempio **l'aliasing**

# Aliasing

- E' importante notare che ciò che avviene nella digitalizzazione rispetta il teorema del campionamento in senso inverso. Stabilito un tasso di campionamento SR, occorre eliminare dal segnale tutte le frequenze che sono maggiori di SR/2.
- Per fare ciò si usa un filtro *passabasso* in quanto fa passare solo frequenze sottouna certa soglia ed è detto di *antialiasing*, in quanto evita il problema dell'*aliasing*.

## Memoria in Kb dei files audio

Lo spazio di memoria (in KiloByte) occupato da un file audio si calcola con la seguente formula:

$$\text{Spazio (in KB): } (f_c * D * N_{bc} * N_c) / (8 * 1024)$$

dove

$f_c$ : tasso di campionamento (n. campioni al secondo)

$D$ : durata in secondi

$N_{bc}$ : numero di bit usati per rappresentare ciascun campione

$N_c$ : numero canali (1: mono; 2: stereo)

# Confronto qualità /data rate

Quality	Sample Rate (KHz)	Bits per Sample	Mono/ Stereo	Data Rate (uncompressed) (kB/sec)	Frequency Band (KHz)
Telephone	8	8	Mono	8	0.200-3.4
AM Radio	11.025	8	Mono	11.0	0.1-5.5
FM Radio	22.05	16	Stereo	88.2	0.02-11
CD	44.1	16	Stereo	176.4	0.005-20
DAT	48	16	Stereo	192.0	0.005-20
DVD Audio	192 (max)	24 (max)	6 channels	1,200.0 (max)	0-96 (max)

Il formato finora presentato è per certi versi uno standard e viene comunemente detto *linear PCM* (PCM – Pulse Code Modulation)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Formati Audio Digitale

I formati dei file audio sono stati sviluppati per standardizzare la riproduzione e la distribuzione di dati audio nei sistemi digitali. I parametri che determinano i dati audio sono tre:

- sampling rate, misurato in campioni/sec (Hz), per canale;
- lunghezza e tipo di codifica della parola binaria, ovvero il numero di bit per campione;
- numero di canali;

I formati si dividono in due tipi:

- con intestazione (header), autodescriventi;
- senza intestazione (headerless o raw);

L' intestazione contiene:

- Definizione codifica usata per i dati audio;
- descrizione brano e dati di copyright

L'intestazione inizia spesso con una parola chiave, e prosegue poi con i dati della codifica

Multimedia A.A. 2011/2012 – Prof. S. Battiato

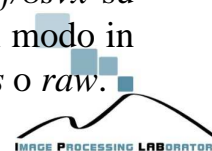


# Formati Audio

Estensione	Origine	Parametri	intestazione
au, snd	NeXT, Sun	Rate, canali, codifica	Y
aif(f), AIFF	Apple, SGI	Rate, canali, quantizz.	Y
aif(f), AIFC	Apple, SGI	AIFF con compressione	Y
Wav, WAVE	Microsoft	Rate, canali, quantizz.	Y
voc	Creative	rate	Y
Iff, IFF/8SVX	Amiga	Rate, canali, info strum.	Y
Mp3, mp4	MPEG	Rate, canali, qualità campioni	Y
sf	IRCAM	Rate, canali, codifica	Y
ra	Real	Rate, canali, qualità campioni	Y
Snd, fssd	Mac, PC	-	N
snd	Amiga	-	N

La larga diffusione del sistema PCM ha fatto sì che oggi siano disponibili decine di formati diversi, il *wav* nei sistemi Windows, *aiff* e *aifc* nei Macintosh, *iff/8svx* su Amiga, *au* nei sistemi Sun e Unix. Si noti che a cambiare è l'header, non il modo in cui sono rappresentati i dati. I file senza header vengono definiti *header-less* o *raw*.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Compressione Audio

Il formato CD Audio standard (PCM – Pulse Code Modulation) viene comunemente detto *linear PCM* in realtà non comprime il segnale.

La più semplice tecnica di compressione utilizza la differenza di campioni successivi realizzando la cosiddetta *DPCM* (Differential PCM).

Le differenze locali tendono ad essere molto piccole, richiedendo quindi un numero di bit sensibilmente inferiore.

Esistono inoltre versioni ottimizzate di questa tecnica come quella utilizzata dal formato Compact Disc-Interactive (CD-I) detta *ADPCM* (Adaptive DPCM), che ottengono risultati ancora migliori.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Compressione Audio: Codifica per modelli

Le codifiche per modelli sono tecniche legate ad una particolare sorgente sonora (in questo caso la voce) che si tenta di emulare tramite un modello più o meno semplificato. Le corde vocali e la gola hanno delle ben precise caratteristiche fisiche, il loro comportamento sarà quindi predicibile sulla base di un modello.

Queste codifiche rappresentano una scelta ottimale per la compressione della voce, tanto che vengono utilizzate nella telefonia mobile (GSM) e anche su Internet.

Le più famose sono LPC (Linear Predictive coding) e il CELP (Code Excited Linear Predictive).



## Compressione Audio: $\mu$ -law, A-law

Schemi che realizzano le specifiche contenute nella raccomandazione G.711 rilasciata dal CCITT, comitato standard per le TLC. La codifica  $\mu$ -law è utilizzata in Nord America e Giappone per i servizi di telefonia ISDN, A-law è usato invece in Europa e sul traffico internazionale ISDN;

- Frequenza di campionamento: 8 KHz;
- Quantizzazione logaritmica a 8 bit;
- Bitrate: 64 Kbps



# Compressione Audio

- Le più moderne tecniche di compressione audio (AC3, MP3) si basano sulle caratteristiche dell'apparato uditivo umano. La principale caratteristica su cui si basano le migliori tecniche di compressione audio e' basata sul cosiddetto effetto di *masking*.
- Così come nel dominio del visibile una forte sorgente luminosa tende ad abbagliare e nascondere la presenza di sorgenti luminose di minore intensità, la presenza di un forte suono ad certa frequenza tende a mascherare la presenza dei suoni a frequenze vicine.
- Queste tecniche di compressione si basano pertanto sulla suddivisione dello spettro audio in bande di frequenza di dimensione opportuna rispetto alla selettività del nostro sistema uditivo. Per ogni banda viene poi applicata una rappresentazione che tiene conto più che di tutte le componenti presenti nella banda, solo di quelle che sono effettivamente udibili (non mascherate).

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Compressione Audio Percettiva

Gli schemi di compressione “percettivi” comprimono il segnale eliminando quelle parti che il nostro apparato uditivo non percepirebbe

Sono codifiche di tipo lossy

Tali schemi si basano su considerazione di psicoacustica ed in particolare sull'effetto di mascheramento (**Masking**)

Multimedia A.A. 2011/2012 – Prof. S. Battiato

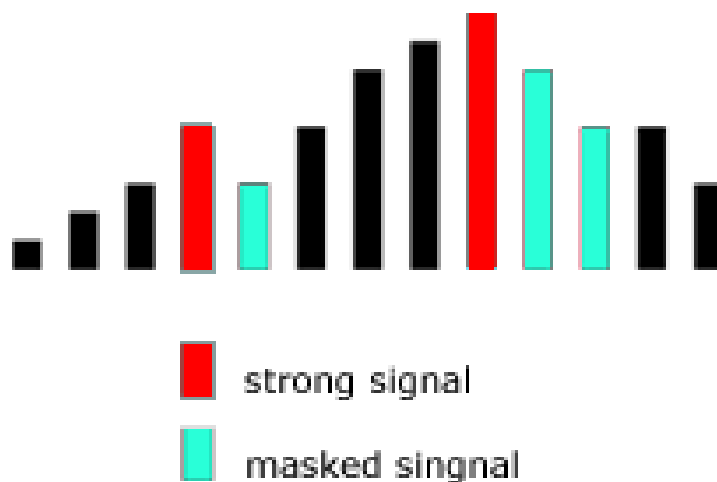


# Masking

Il mascheramento non è altro che la sovrapposizione di suoni deboli con suoni forti. Accade quasi sempre che suoni di strumenti differenti si sovrappongano tra loro. Nei casi in cui il suono più alto copre totalmente quello più basso, c'è il cosiddetto mascheramento.

Il nostro orecchio, non è in grado di percepire frequenze “deboli” adiacenti a frequenze “forti”, in quanto queste ultime “mascherano” le prime. Una forte componente di suono può quindi rendere non udibili i suoni a frequenze vicine: la presenza di una qualunque componente udibile in frequenza provoca una alterazione locale (e temporale) della soglia di udibilità attorno a quella frequenza.

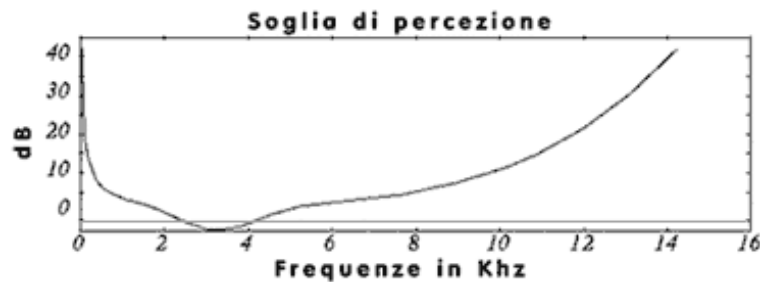
# Masking



**Il Segnale Rosso Maschera Il Segnale Azzurro Che Quindi Non Verra' Codificato**

# Sensibilità Uditiva

Il nostro orecchio è sensibile in misura diversa alle diverse frequenze:



Dal grafico emerge che l'orecchio umano è maggiormente sensibile alle frequenze comprese fra 2 e 4 KHz, che richiedono pochissimi **dB** per essere percepite. Per poter udire le frequenze successive ai 6 KHz bisogna incrementare il loro volume secondo l'andamento evidenziato dal grafico. Non è un caso che l'intervallo fra i 2 e i 4 KHz sia quello massimamente usato dalla nostra voce.

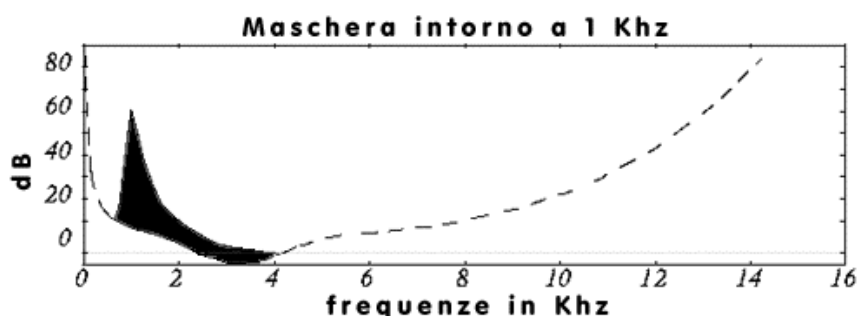
Ovviamente possiamo già usare questa caratteristica dell'orecchio a nostro vantaggio eliminando dallo spettro del segnale in analisi quelle componenti spettrali non udibili dall'orecchio medio. In sostanza si tagliano le alte frequenze e le bassissime frequenze. In generale, siccome l'orecchio a queste frequenze perde sensibilità e selettività, si può ridurre la quantità di informazione trasmessa in questa parte di spettro.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Masking Frequenziale

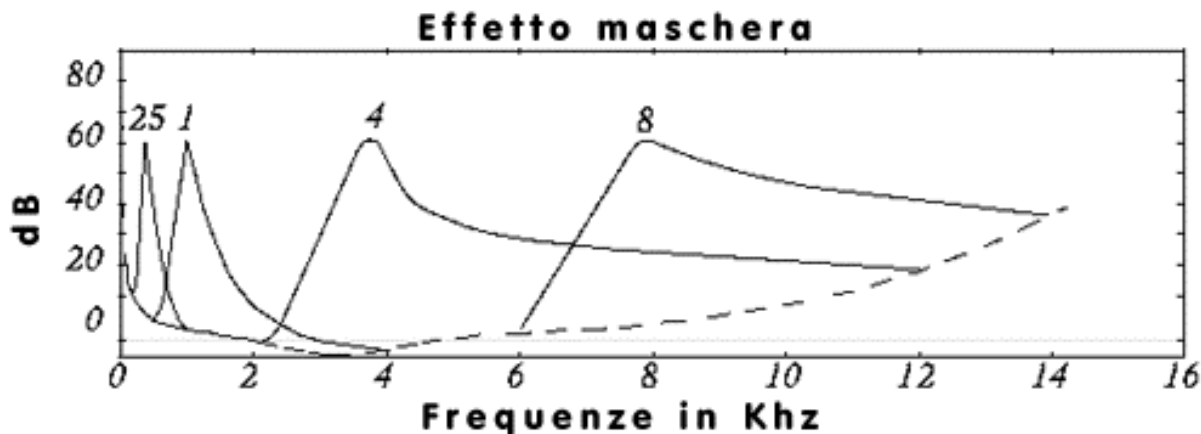
E' possibile mostrare come le componenti frequenziali di un segnale, sotto certe condizioni, interferiscano tra loro. Si emette un tono pari a 1 KHz (*tono maschera*) ad un volume fisso pari a 60 dB. Si emette un *tono test* ad un differente livello (es. 1.1 KHz) e se ne aumenta il volume finchè diventa appena distinguibile. Si varia la frequenza del tono test e si disegnano i valori risultanti in cui esso diviene udibile. Il tono fisso a 60 dB copre il tono test nelle frequenze immediatamente antecedenti e soprattutto nelle frequenze successive. l'inserimento del tono maschera ad un volume più alto, impone di aumentare il volume del tono test di una certa percentuale.



Multimedia A.A. 2011/2012 – Prof. S. Battiato



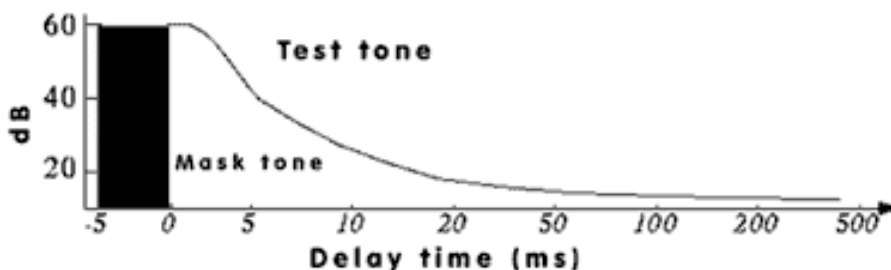
# Masking Frequenziale



Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Masking Temporale



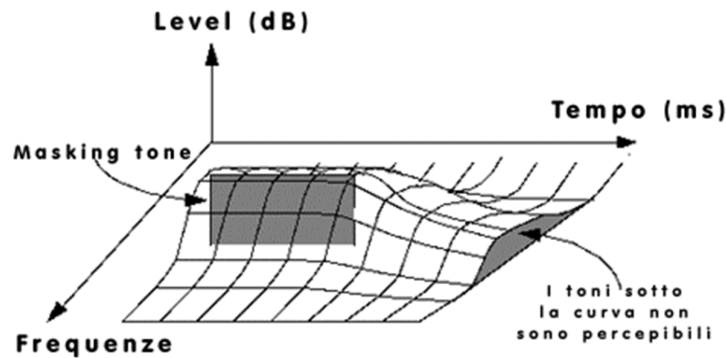
Ascoltando un suono forte, l'orecchio necessita di un certo tempo per poter percepire un suono più debole ad esso attiguo, quando il suono più forte viene interrotto. Ciò è dovuto al fatto che il sistema uditivo è composto da parti meccaniche in movimento che non si adattano istantaneamente alle variazioni di pressione acustica.

Il periodo di tempo necessario per poter percepire il suono più debole può essere quantificato. Emettendo un tono pari a 1 KHz (tono maschera) ad un volume fisso pari a 60 dB, insieme a un *test tone* di 1.1KHz a 40 dB, si può fare il seguente esperimento: si interrompe il tono maschera e subito dopo il test tone; si varia il tempo di ritardo con cui si interrompe il test tone, e si disegna la soglia di tempo in cui esso risulta udibile. Si osserva che il tono test diventa udibile se rimane attivo per almeno 5 ms oltre l'istante in cui il tono maschera viene interrotto.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Masking



L'effetto complessivo del mascheramento è che molti toni non saranno mai udibili perché collocati nel dominio della frequenza e del tempo troppo vicino a toni forti. Tenendo conto della sensibilità dell'orecchio e del fenomeno del Masking Audio (*effetto maschera*) è quindi possibile eliminare dallo spettro del segnale una quantità molto alta di informazioni inutili, perché non udibili dall'orecchio umano.

Questi sono i fenomeni Psico-Acustici su cui si basano i moderni algoritmi di compressione audio come MP3, MP3Pro, Atrac-3, AAC, etc.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## MPEG Audio Standard

- L'algoritmo di compressione audio MPEG è stato sviluppato dal *Motion Picture Experts Group* (MPEG), sotto la direzione dell'*organizzazione internazionale per gli standard* (ISO), che ha rilasciato lo standard nel 1992, e dell'*International Electro-Technical Commission* (IEC).
- Lo *standard di compressione audio MPEG* è solo una parte dello standard nato per la compressione video, audio e relativa sincronizzazione. Può essere usato in specifiche applicazioni, per comprimere audio digitale ad alta fedeltà con bit rate molto bassi. L'algoritmo MPEG/audio fa parte della classe di algoritmi di compressione *lossy* (con perdita), ma consente di ottenere notevoli fattori di compressione, con perdita contenuta a livello dei dati, comunque non percepita dal sistema uditivo umano.
- Il noto formato *Mp3* non è altro che l'applicazione dell'algoritmo di compressione *MPEG/Layer III*.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# MPEG Audio/Layer III: Prestazioni

<i>Qualità</i>	<i>Banda Coperta</i>	<i>Modalità</i>	<i>Bitrate</i>	<i>Fattore di riduzione</i>
Radio AM	7.5 kHz	mono	32 kbps	24:1
Radio FM	11 kHz	stereo	56..64 kbps	26..24:1
Simile a CD	15 kHz	stereo	96 kbps	16:1
CD	> 15 kHz	stereo	112..128 kbps	14..12:1

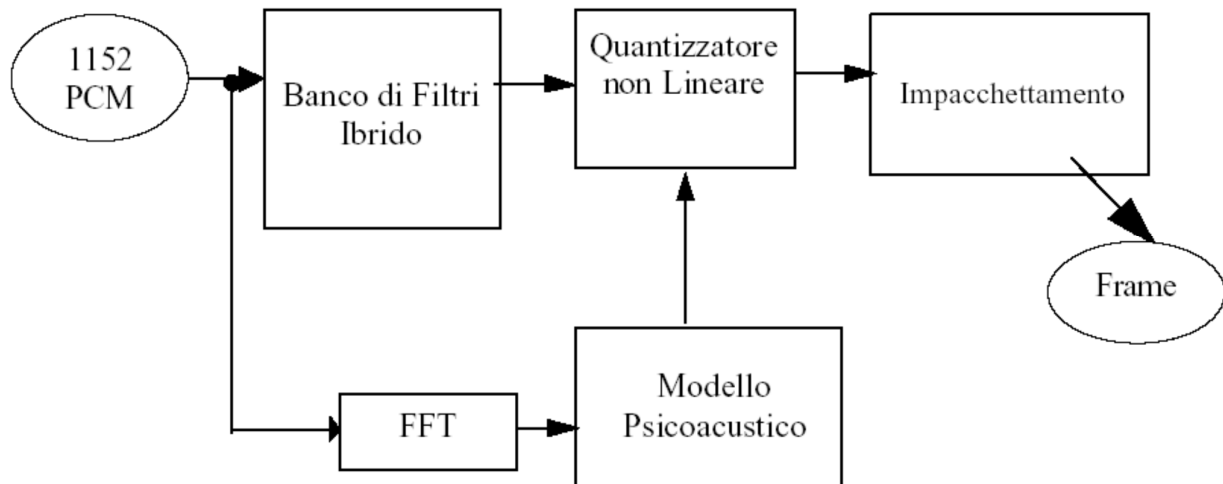
Questo significa che un segnale stereo in qualità CD, che normalmente occuperebbe circa 10 Mbyte/minuto, e richiederebbe un bit rate di 1.4 Mbit/sec, può essere gestito con un bitrate variabile da 112 a 128 Kbit/secondo ed un'occupazione di memoria dagli 840 Kbyte/minuto ( $112 \text{ kbit/sec} * 60 \text{ sec/min} = 6720 \text{ kbit} / 8\text{bit/byte} = 840 \text{ kbyte}$ ) ai 960 Kbyte/minuto (per bitrate di 128 kbit/sec). Quindi sarebbe possibile, in linea di principio, ascoltare un brano musicale in qualità CD, trasferendolo in tempo reale su due linee telefoniche ISDN da 64 kbit/sec. Si noti come i fattori di riduzione siano calcolati rispetto al bit rate dell'audio in qualità in CD (1.4 Mbit/sec).

## MPEG Audio: I Layers

Qualunque encoder MPEG/Audio è in grado di comprimere un segnale PCM con diversi algoritmi di compressione. Per quanto riguarda MPEG-1, gli algoritmi esistenti sono tre e vengono identificati da un "Layer" di appartenenza:

- Layer 1: è l'algoritmo più semplice dei tre e raggiunge buoni risultati con un bitrate pari a 384Kbit/sec per un segnale stereo. Esso associa ad un frame 384 campioni PCM per frame. Il formato di file associato è l'MP1.
- Layer 2: più complesso del primo in quanto associa ad un frame 1152 campioni PCM; è adatto per codifiche a bitrate intorno ai 192-256Kbit/sec per un segnale stereo.
- Layer 3: è il più complesso dei tre ed è anche quello che raggiunge le migliori prestazioni. Il formato MPEG-1 associa ad ogni frame 1152 campioni. Con bitrate tra 128-192kbit/sec si riesce ad ottenere un segnale stereo di qualità sufficientemente elevata. Il formato di file associato è MP3. I concetti che stanno alla base del layer 3 sono:
  - ✓ Dominio frequenziale suddiviso in funzione delle bande critiche
  - ✓ Utilizzo della codifica Huffman per l'impacchettamento finale dei dati audio

# MPEG Audio



Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Mpeg Audio

- **Banco di Filtri Ibrido:** questa fase ha il compito di convertire i campioni PCM nel corrispondente dominio frequenziale, utilizzando un "Banco di Filtri Polifasico seguito da una Trasformata Coseno Modificata (MDCT)". Questo blocco prende, quindi, il segnale rappresentato nel dominio del tempo e lo trasforma nella corrispondente rappresentazione nel dominio delle frequenze (spettro).
- **Modello Psicoacustico:** questo blocco rappresenta "il cuore" dell'encoder e di tutto il sistema MPEG/Audio. Il suo compito è di analizzare lo spettro del segnale (calcolato con la Trasformata di Fourier) e definire il livello di soglia di udibilità SMR (Signal to Mask Ratio) sfruttando i principi psicoacustici dell'apparato uditivo umano. In pratica, il modello psicoacustico determina quali sono le sole informazioni che il nostro orecchio è in grado di percepire e quali no, e fornisce questa informazione al blocco "**Quantizzatore non Lineare**" che la gestirà opportunamente.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Mpeg Audio

- **Quantizzatore non Lineare** : compito di questo blocco è di codificare numericamente lo spettro ricevuto dal blocco "**Banco di Filtri Ibrido**" in funzione dell'importanza di ogni banda di frequenze: se il blocco "**Modello Psicoacustico**" indica che una particolare banda di frequenze è percepita poco, essa verrà codificata con pochi bit; viceversa, se il blocco "**Modello Psicoacustico**", indica che una particolare banda di frequenze è percepita molto, essa verrà codificata con tanti bit. L'obiettivo finale è quello di ottenere una quantizzazione dello spettro per cui il rumore di quantizzazione introdotto si trovi al di sotto della soglia di udibilità (SMR) fornita dal modello psicoacustico.
- **Impacchettamento**: compito di questo blocco è prendere la codifica numerica dello spettro frequenziale generato dal blocco "**Quantizzatore non Lineare**" ed impacchettarla secondo la sintassi dello standard MPEG utilizzato. In questa fase, il layer 3 prevede un'ulteriore compressione con l'algoritmo di Huffman.

## MP3: Algoritmo di Codifica

L'algoritmo di codifica è composto da diversi steps che possono essere così riassunti:

1. Si usano dei filtri per dividere il segnale audio che è campionato con una certa frequenza, ad esempio di 44100 campioni al secondo, in 32 sottobande che coprono l'intera gamma di frequenze udibili dell'orecchio umano e per ognuna delle quali sono noti i parametri di mascheramento nel tempo e in frequenza.
2. Per ognuna delle sottobande, viene calcolata l'entità del mascheramento causata dalle bande adiacenti.
3. Se la potenza in una sottobanda è sotto la soglia di mascheramento, allora non viene codificata in uscita l'informazione che essa trasporta, poiché non sarebbe udibile.
4. Altrimenti, occorre calcolare il numero di bit necessari per rappresentare l'informazione della sottobanda facendo attenzione che in questo procedimento il rumore introdotto stia sotto la soglia.
5. Infine, formare il flusso di bit (bitstream) in uscita.

# MP3

- L'MP3 utilizza sempre il blocco dei filtri, però a differenza dei layers 1 e 2 le sottobande non sono tutte della stessa dimensione, poiché certe frequenze contengono molta più informazione e vanno trattate con maggiore dettaglio. Il layer 3, inoltre, fa uso di una MDCT, cioè di una trasformata discreta del coseno modificata.
- Si tratta, quindi, di effettuare una operazione che consenta di migliorare la risoluzione in frequenza per ognuna delle sottobande. Questa operazione consente di suddividere ognuna delle 32 sottobande in ulteriori 6 (short) o 18 (long) sottofrequenze, secondo un processo noto come filtraggio sottobanda (sub-band filtering).
- Il modello psico-acustico lavora ulteriormente su queste sottosottomaschere, in particolare sui coefficienti della MDCT che le rappresentano. Il modello psico-acustico deciderà quali coefficienti devono passare in uscita e quali no, sulla base del calcolo del mascheramento temporale e sul fatto che alcuni di questi sono ridondanti giacché provengono dai canali sinistro e destro che spesso portano la medesima informazione.

# MP3

- A questo punto il tutto è quasi pronto. I coefficienti "sopravvissuti" contengono le informazioni necessarie alle varie frequenze e devono ora essere organizzati in uscita. I coefficienti vengono ordinati passando dalla frequenza più bassa a quella più alta. Poiché la massima informazione è contenuta in bassa frequenza, i coefficienti di bassa frequenza sono più numerosi di quelli in alta frequenza (ed infatti i puristi lamentano la scarsa efficienza dell'MP3 per la riproduzione delle alte frequenze). L'intero intervallo viene diviso in tre parti (frequenze basse, medie e alte).
- Ognuno di questi intervalli viene codificato a parte secondo l'algoritmo di Huffman, che è uno degli algoritmi basilari nella teoria della compressione. L'algoritmo è ottimizzato per ognuno dei tre intervalli. A questo punto i dati vengono inviati in uscita sotto forma di pacchetti che contengono un CRC (codice per la correzione dell'errore) per rendere il sistema più robusto agli eventuali errori che si possono presentare durante il trattamento del file.

# Standard di codifica avanzati

- MPEG 2 Audio AAC (Advanced Audio Coding)
- MPEG 4 Audio AAC
- MP3 Pro
- Dolby AC-2, AC-3
- ....

## Formato MIDI

Il MIDI (**Musical Instrument Digital Interface**) è il protocollo standard per la comunicazione tra dispositivi musicali

Un protocollo è un insieme di regole che stabilisce attraverso quali modalità deve avvenire la comunicazione tra dispositivi hardware o processi software differenti.

Alla base del MIDI c'è la necessità di far comunicare sintetizzatori musicali diversi. Il MIDI si colloca sia a livello simbolico e sia a livello operativo. Attraverso il MIDI è possibile rappresentare come dati elettronici una performance musicale (n.b., non la forma d'onda del segnale audio)

L'informazione codificata dal protocollo consiste in un'insieme di istruzioni per un sintetizzatore che sulla base dei dati ricevuti, si occupa della sintesi audio

Il MIDI permette L'I/O tra dispositivi diversi.

# Comunicazione MIDI

- 1) Eseguiamo un brano sul un dispositivo di input MIDI (es. la tastiera).
- 2) I dati esecutivi (quali tasti sono stati premuti, con quanta forza, per quanto tempo) vengono convertiti in forma MIDI.
- 3) I dati MIDI vengono trasmessi, attraverso una connessione hardware dedicata.
- 4) I dati MIDI vengono decodificati da un altro dispositivo MIDI (es. un sintetizzatore) che genera il segnale audio.

- 
- Se i dati MIDI vengono memorizzati, possono essere modificati attraverso procedure di editing.
  - Si chiama **Sequencer** un sistema (dispositivo fisico o software per un computer generico) di registrazione e di esecuzione dotato di una memoria programmabile nella quale vengono memorizzati i dati di controllo operativo alla (ri-)generazione di eventi musicali. Il sequencer memorizza i dati da un dispositivo di input (un sintetizzatore ma anche la tastiera del computer), ne consente l'editing e (ri)-crea la performance inviando i dati al dispositivo di esecuzione (un sintetizzatore ma anche la scheda audio del computer)



## Elaborazione digitale del suono



# Editing del suono

- Un *Editor Audio* è un programma che consente di modificare un segnale audio digitale (mono o stereo) sia esso campionato o sintetizzato

Esempi: Audacity 1.0, Sound Forge 7.0, Wavelab 4.0

## Altri tipi di SW:

*Sequencer*: sono degli editor multitraccia in grado di gestire sia audio campionato sia MIDI (es. Cool Edit Pro 2.1, Cakewalk 9.0, Cubase SX 2.0)

*Groove Box*: sono dei sequencer in grado di memorizzare diversi pattern generati sfruttando la sintesi interna e/o suoni campionati e poi di missarli insieme applicando inoltre gli effetti disponibili (es. Fruity Loops 4, Acid Pro 4.0)

*SW per l'estrazione (ripping) e conversione audio*: sono strumenti per estrarre le tracce audio dai CD e per convertire i dati nel formato desiderato (es. Yamp)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Modalità di visualizzazione

- Rappresentazione del segnale nel *dominio del tempo*:  
descrive l'andamento nel tempo dell'ampiezza del segnale
  - ✓ Misure del tempo (asse orizzontale): cronometrica, musicale, di sincronizzazione audio-video (SMPTE)
  - ✓ Misure della ampiezza (asse verticale): valore assoluto, in dB, valore percentuale o normalizzato
- Rappresentazione del segnale nel dominio della frequenza:  
descrive il contenuto frequenziale del segnale
- Rappresentazione del segnale nel dominio tempo/frequenza/ampiezza (**sonogrammi**)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Statistiche

Vengono calcolate diverse statistiche:

- ✓ valore minimo
- ✓ valore massimo (di picco)
- ✓ posizione del minimo
- ✓ posizione del massimo
- ✓ valore efficace della ampiezza: **RMS** (Root Mean Square)
- ✓ valore medio del segnale (**DC Offset**)
- ✓ ...

## Operazioni di base

- *Selezione* di un segmento audio;
- *Eliminazione* del segmento selezionato (o di tutto il resto: trim/crop);
- *Copia* del segmento selezionato nella memoria temporanea;
- *Inserimento* di quanto presente in memoria in una posizione determinata;
- *Sostituzione* di un segmento selezionato con quanto presente in memoria temporanea;
- *Mix* di due segnali (con possibilità di regolare sia il volume sia la dissolvenza incrociata tra di loro);
- Inserimento di *marcatori* (indicatori di posizione);
- Definizione di *regioni* (tra due marcatori);
- *Editing* di singoli campioni;

...

## Modifica dei parametri di campionamento e salvataggio

- Ricampionamento
- Ri-quantizzazione
- Conversione di un file mono in stereo
- Salvataggio in vari formati:

WAVE (formato proprietario di Windows; .wav)

AIFF (standard di Apple)

AU (standard di Next/Sun)

RAW (senza intestazione)

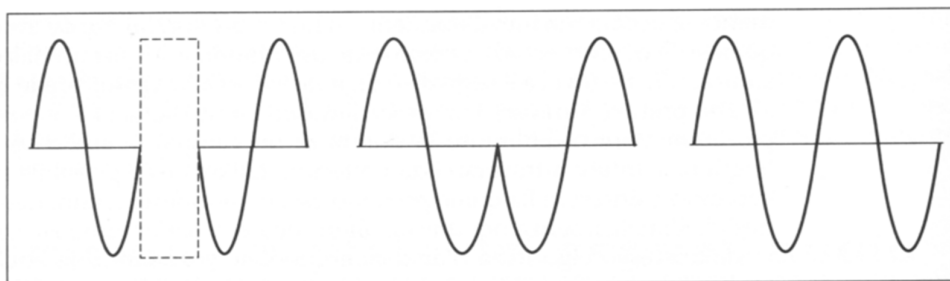
mp3, RealAudio (formati compressi)

Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Operazioni di “inversione”

- **Invert**: il segnale viene invertito rispetto l’asse delle ordinate. Uso nel restauro (eliminazione di click)



- **Reverse**: il segnale viene invertito rispetto l’asse delle ascisse (invertendo l’ordine dei campioni). Si ottiene effetto “aspirato” dovuto ad inversione di involuppo.

Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Operazioni sulla ampiezza del segnale

- **Amplificazione uniforme**
- **Normalizzazione**: massima amplificazione possibile senza che si produca distorsione (**clipping**). Si sfrutta tutta la gamma dinamica!
- **Variazione dell'involuppo** (es. dissolvenze: **fade in, out**)
- **Panning**: posizionamento e movimento laterale del suono (ripartizione) su due o più canali.
- **Distorsione**: si mappa il valore di ogni campione in un altro valore attraverso una funzione “distorcente”;



## Elaborazione della gamma dinamica

- **Compressione**: riduce l'escursione dinamica di un segnale
- **Limitazione**: “limita” la parte del segnale la cui ampiezza supera una data soglia
- **Espansione**: dilata l'escursione dinamica di un segnale
- **Noise gating**: “rimuove” la parte del segnale inferiore ad una data soglia (utile per tagliare il rumore di fondo)



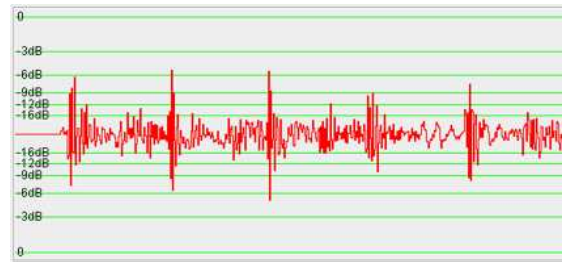
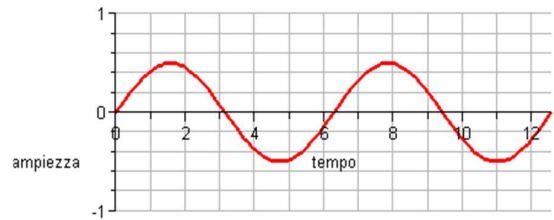
$$y_i = Kx_i \quad K \in \mathbb{R}_0^+$$

## Esempi: Effetti sull'ampiezza

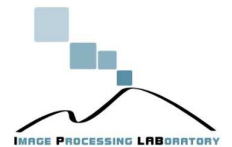
Gli effetti di questa classe possono essere riassunti nell'equazione:

$$y_i = K x_i$$

dove  $y_i$  è il generico  $i$ -esimo *frame* (o campione) in output,  $x_i$  è il generico  $i$ -esimo *frame* in input e  $K$  un valore reale. Questi effetti agiscono solo sull'ampiezza dell'onda e non modificano il timbro o l'altezza (fisicamente intesa come frequenza) dell'audio in input. A titolo esemplificativo i vari effetti sono stati applicati ad un'onda sinusoidale ed ad un'onda sonora.

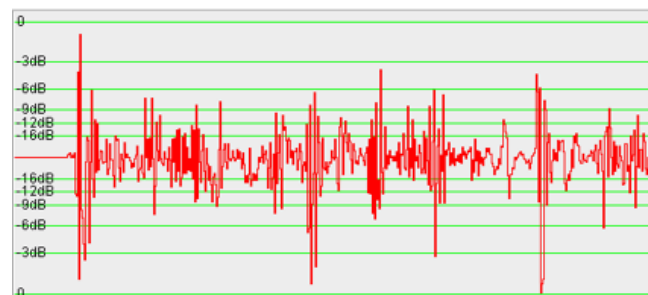
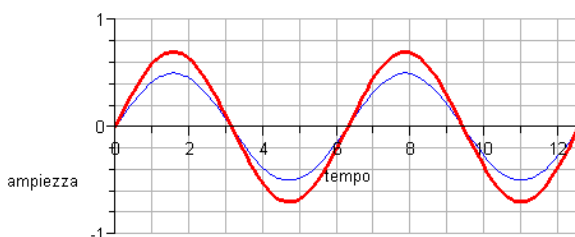


Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Volume

Variare l'ampiezza del segnale corrisponde a moltiplicare ogni singolo *frame* per una costante  $K$ , se  $0 < K < 1$  si avrà una diminuzione del volume (dunque dell'ampiezza), per  $K > 1$  si avrà un aumento del volume. Per  $K = 1$  non ci sarà variazione.



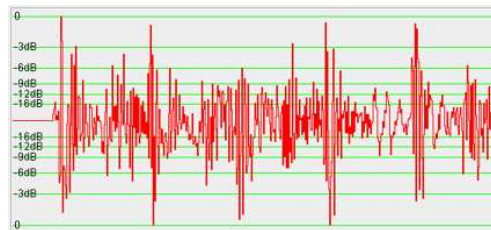
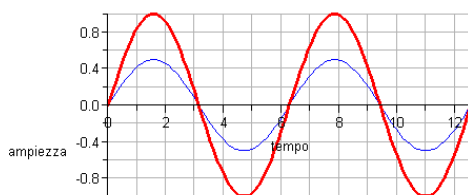
Multimedia A.A. 2011/2012 – Prof. S. Battiato



## Maximize

Questo effetto massimizza il volume senza che venga variato il rapporto tra due *frame* consecutivi qualunque, ovvero si aumenta al massimo il volume senza che si verifichino distorsioni o effetti di saturazione dell'onda. Si analizza ogni singolo *frame* e si prende il valore assoluto del massimo ( $K$ ) raggiunto. Se  $K < 2^{bps-1}$  dove  $bps$  è il numero di *bits per frame*, si moltiplicherà ogni *frame* per  $2^{bps-1}/K$ .

$$y_i = \frac{2^{bps-1}}{K} x_i \quad \text{con } K = \max\{|x|\}$$



## Minimize

Si analizza ogni singolo *frame* e si prende il valore assoluto del minimo raggiunto ( $K$ ). Si sottrae poi ad ogni *frame* il valore  $K$ . In questo caso si vuole diminuire al minimo il volume senza modificare la differenza tra due qualsiasi *frames*.

$$y_i = x_i - K \quad \text{con } K = \min\{|x|\}$$

## Fade In

Si agisce sull'ampiezza del segnale audio, facendone variare il volume, con  $K$  crescente nell'intervallo  $[0,1]$ . Considerando un generico intervallo di  $N$  frames  $[a,b]$ , il volume, partendo da 0, aumenterà dinamicamente in ogni frame. L'equazione del *Fade In* è espressa dalla seguente formula:

$$y_n = f(n-a)x_n \quad a \leq n \leq b$$

$$\text{si noti che } 0 \leq n-a \leq N$$

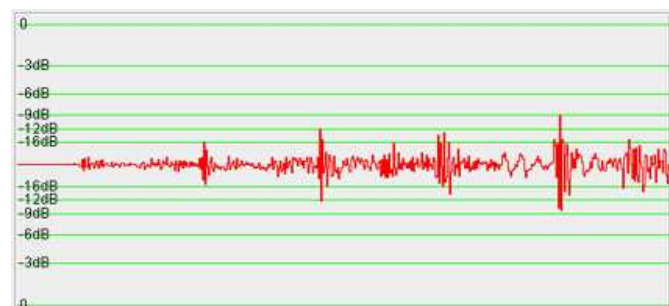
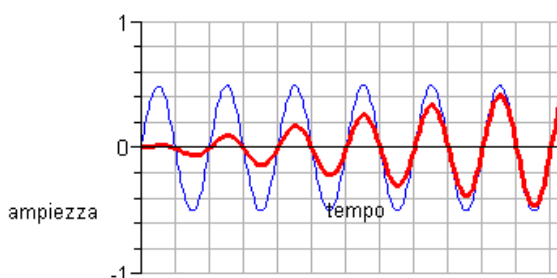
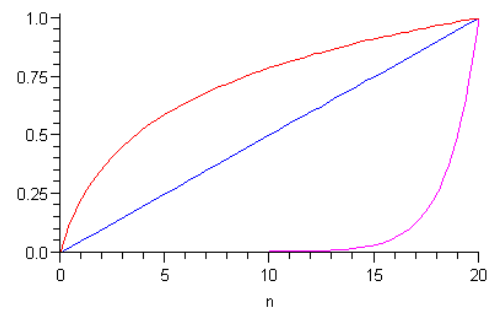
$$f : [0, N] \rightarrow [0, 1]$$

$f$  monotona crescente

$$f(0) = 0 \quad f(N) = 1$$

## Fade In

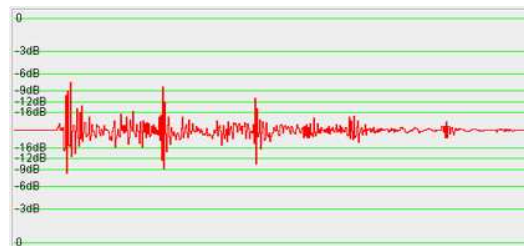
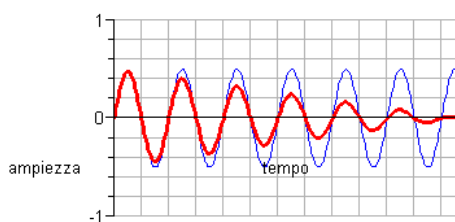
- lineare  $f(n) = \frac{n}{N}$
- logaritmica  $f(n) = \log_{1+N}(1+n)$
- esponenziale  $f(n) = \begin{cases} 0 & \text{per } n=0 \\ b^n & \text{altrimenti con } b > 1 \end{cases}$



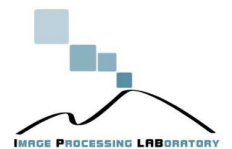
## Fade Out

Si agisce sull'ampiezza del segnale audio, facendone variare il volume con  $K$  decrescente nell'intervallo  $[0,1]$ . Considerando un generico intervallo di  $N$  frames  $[a,b]$ , il volume, partendo da 1, diminuirà dinamicamente in ogni frame. Considerando una generica funzione  $f$  che rispetti le stesse condizioni viste per il *Fade in*, l'equazione del *Fade Out* è espressa dalla seguente equazione:

$$y_n = (1 - f(n - a))x_n \quad a \leq n \leq b$$



Multimedia A.A. 2011/2012 – Prof. S. Battiato

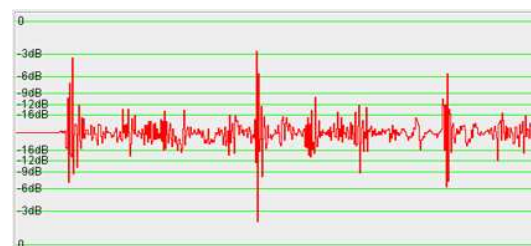
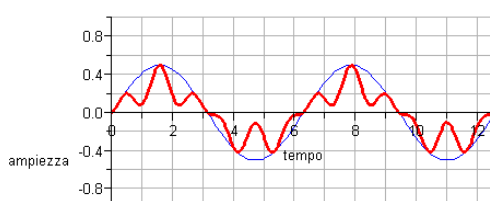


## Tremolo

Il volume varia in modo sinusoidale: sia  $factor$  un fattore costante e  $f_{wave}$  la frequenza di campionamento del file, l'effetto tremolo può essere espresso tramite l'equazione:

$$y_i = x_i \left( \alpha + \beta \sin \left( factor \frac{i\pi}{f_{wave}} \right) \right) \quad \text{con } 0 < \beta \leq \alpha < 1$$

solitamente  $\alpha=0.6$   $\beta=0.4$



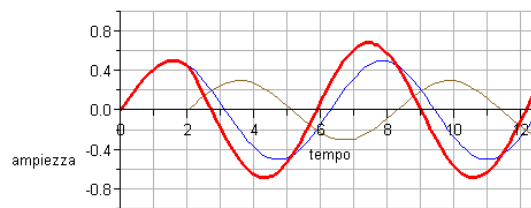
Multimedia A.A. 2011/2012 – Prof. S. Battiato



# Echo

Fissato un ritardo  $d$  si aggiunge al *frame*  $i$ -esimo il *frame*  $(i-d)$ -esimo. Il risultato finale è lo stesso che si prova in un ambiente con l'eco: dopo un certo istante si sente una ripetizione ma con un volume inferiore (selezionabile dall'utente come *decay*). Si può scegliere anche quante volte ripetere l'eco; ad ogni ripetizione il volume dell'eco si abbassa di un fattore *decay* (eco dell'eco):

$$y_i = x_i + \sum_{j=1}^{\text{ripetizioni}} (x_{i-d} \cdot \text{decay}^j) \quad \text{con } 0 < \text{decay} \leq 1$$



Echo (rosso) su onda sinusoidale (blu). L'onda marrone è l'originale con un ritardo  $d=2$  e  $\text{decay}=0.6$

# Filtraggi

- **Filtraggio:** un filtro (passa basso, passa alto, passa banda, elimina banda) agisce sullo spettro enfatizzando o attenuando determinate frequenze. Si considera come  $f_t$  (frequenza di taglio) quella a cui il filtro attenua di 3db il livello d'ampiezza massimo.

Si usano gli *equalizzatori grafici* o *parametrici*. Una delle operazioni più comuni e importanti; permette di modificare lo spettro e quindi anche il *timbro* del segnale

# Filtraggio

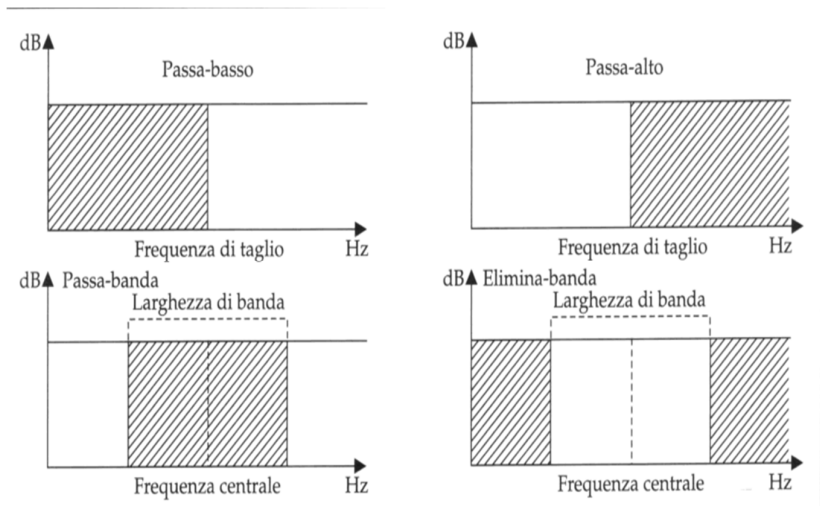
➤ Tipi di filtro:

*passa basso*  
*passa alto*  
*passa banda*  
*elimina banda*

➤ Parametri dei filtri:

*frequenza di taglio*  $f_t$   
*frequenza centrale*  $f_c$   
*larghezza di banda*  $\Delta f$   
*guadagno*  $Q=(f_c/\Delta f)$

*ordine*: specifica (in db per ottava) la pendenza della curva nella regione di transizione (es un filtro del 1° ordine presenta una pendenza di 6 db/ottava; del 2° ordine 12 db/ottava)



## Ritardi

- **Ritardi (delay)**: insieme di operazioni che utilizzano una o più copie ritardate (e talvolta modulate) del segnale originario che vengono sommate al primo per produrre effetti diversi.

Nota: i parametri principali del delay sono il tempo di ritardo ( $dt$ ) e quello di decadimento (specifica dopo quanto tempo gli echi non sono più udibili)

$1 < dt < 10$  ms spazializzazione del suono

$10 < dt < 50$  ms chorus/flanger

$dt > 50$  ms echi discreti

- **Modulazione**: alcuni aspetti di un segnale (detto **portante**) variano in relazione ad un altro (detto **modulante**). Esempio: nella modulazione di frequenza, la frequenza della portante viene fatta variare in funzione della modulante.

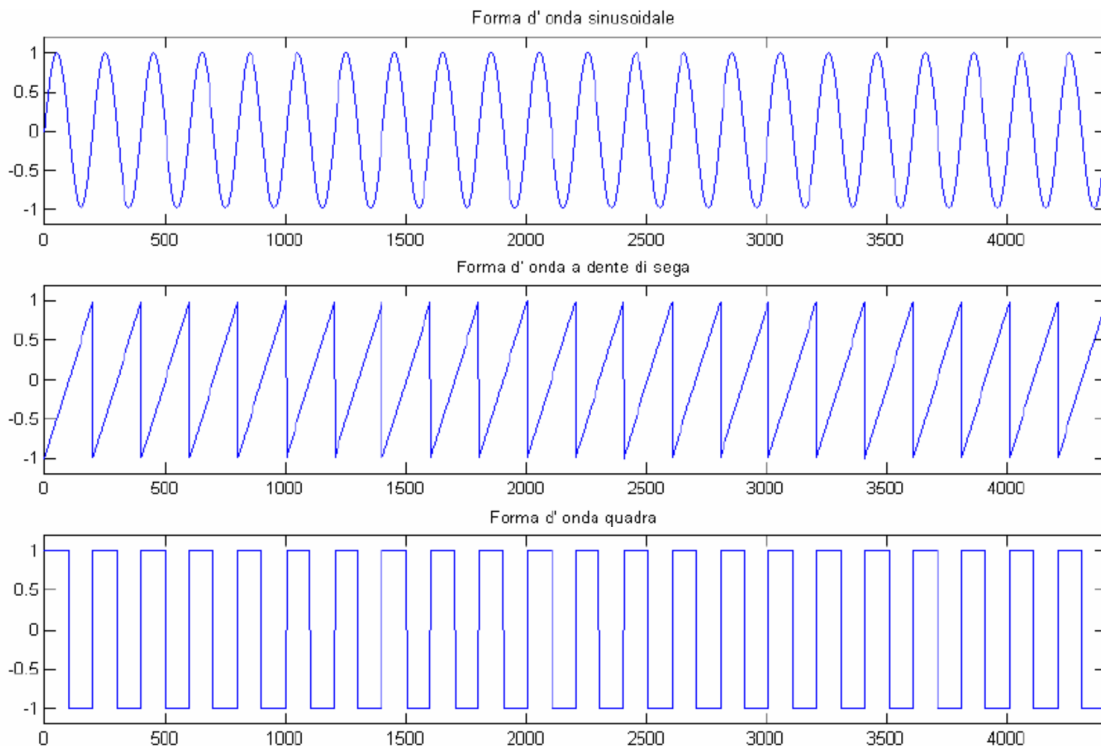
# Effetti

- *Chorus*
- *Flanger*
- *Phaser*
- *Delay-eco*
- *Riverbero*

## MATLAB e Audio

• Matlab fornisce diverse funzioni per generare forme d'onda. La maggior parte di queste richiede la dichiarazione preliminare di un vettore tempo.

```
t = linspace(0, 2, 2*22050) ;  
waveform = sin(2 * pi * freq * t);  
waveform = sawtooth(2 * pi * freq * t);  
waveform = square(2 * pi * freq * t);
```



## ***sound* e *soundsc*: funzioni di ascolto**

**Matlab fornisce due funzioni per l'ascolto dei suoni (vettori) creati.**

### **La funzione *sound***

#### **Sintassi:**

```
sound(y, Fs)
sound(y)
sound(y, Fs, bits)
```

#### **Descrizione:**

- *sound(y,Fs)* invia alla scheda audio il segnale nel vettore *y*, il quale ha frequenza di campionamento *Fs*. Si assume che i valori in *y* siano nel range  $-1.0 \leq y \leq 1.0$ . I valori al di fuori del range vengono tagliati.
- *sound(y)* esegue i suoni alla frequenza di campionamento di default di 8.192 Hz.
- *sound(y, Fs, bits)* esegue il suono usando un numero di bit per campione pari a *bits*. La maggior parte delle piattaforme supporta *bits* = 8 oppure *bits* = 16.

**Nota:** i suoni stereo vengono eseguiti solo su piattaforme che li supportano. *y* deve essere una matrice di dimensioni (n,2).

# ***sound* e *soundsc*: funzioni di ascolto**

## **La funzione *soundsc***

### **Sintassi:**

```
soundsc(y,Fs)  
soundsc(y)  
soundsc(y, Fs, bits)  
soundsc(y,...,slim)
```

### **Descrizione:**

- *soundsc*(*y*,*Fs*) invia alla scheda audio il segnale nel vettore *y*, il quale ha frequenza di campionamento *Fs*. Il segnale *y* viene scalato nel range  $-1.0 \leq y \leq 1.0$  prima di essere eseguito, in modo tale da evitare i tagli che la funzione *sound* apporta quando i valori in *y* sono al di fuori del range precedente.
- *soundsc*(*y*) esegue i suoni alla frequenza di campionamento di default di 8.192 Hz.
- *soundsc*(*y*, *Fs*, *bits*) esegue il suono usando *bits* bit per campione. La maggior parte delle piattaforme supporta *bits* = 8 oppure *bits* = 16.
- *soundsc*(*y*,...,*slim*), dove *slim* = [*slow shigh*] mappa i valori in *y* compresi tra *slow* e *shigh* all'intero range sonoro. Il valore di default è *slim* = [*min*(*y*) *max*(*y*)].

## **Lettura/Scrittura**

Matlab fornisce quattro funzioni per la lettura e la scrittura di soundfile. Le funzioni ***auread*** ed ***auwrite*** leggono e scrivono file in formato NeXT/SUN (.au), mentre le funzioni ***wavread*** e ***wavwrite*** leggono e scrivono file in formato Microsoft Wave (.wav).

# Esempio analisi mediante FFT

```
[y, fs, nbits] = wavread('triangle.wav');  
disp('Playing at the original sample  
rate.');
```

sound(y, fs);

% First try the specgram function on the  
sound.

```
specgram(y);  
Y = fft(y);  
plot(abs(Y));
```