

# **PROGETTAZIONE E PRODUZIONE MULTIMEDIALE**

**Prof. Alberto Del Bimbo  
Dip. Sistemi e Informatica  
Univ. degli Studi di Firenze**

# SUONO

## IL SUONO

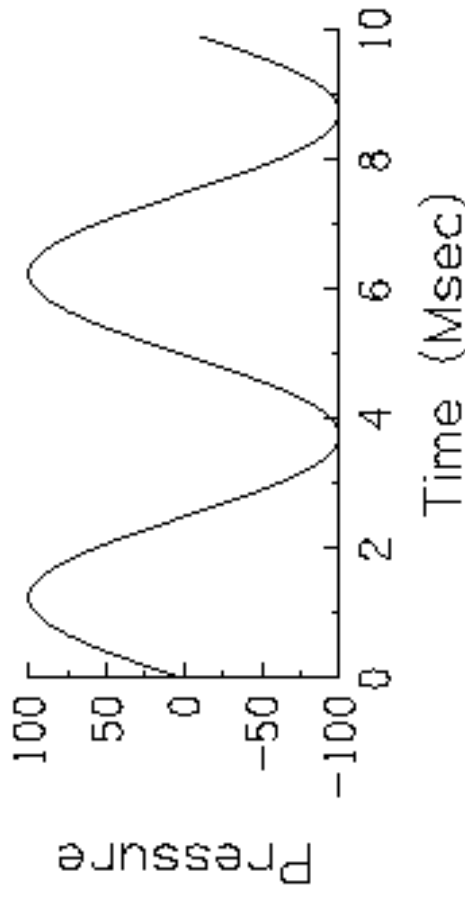
- ⌘ Il suono è un fenomeno fisico prodotto da vibrazioni della materia. Come la materia vibra si creano variazioni nell'aria circostante. La percezione del suono è legata alla vibrazione delle strutture interne del nostro orecchio, causata da variazioni di pressione dell'aria.
- ⌘ Tali variazioni di pressione possono essere grandi o piccole e possono aver luogo rapidamente o lentamente.
- ⌘ In ogni caso la natura della variazione di pressione è ciclica, nel senso che l'andamento del valore di pressione si ripete con valori analoghi nel tempo.




## FREQUENZA ED AMPIEZZA

- ⌘ La velocità con cui i valori di pressione fluttuano ciclicamente viene espressa dalla FREQUENZA della vibrazione.
- ⌘ La durata nel tempo di ogni ciclo della oscillazione viene detta PERIODO.
- ⌘ Il numero di cicli presenti in un secondo corrisponde alla frequenza.
- ⌘ Frequenza ( $f$ ) e periodo ( $T$ ) sono legati dalla relazione  $f=1/T$ . Tipicamente, la frequenza si misura in Hertz (Hz), equivalendo un Hertz ad una frequenza di un ciclo per secondo.
- ⌘ L'altro valore che caratterizza la fluttuazione di pressione è la sua AMPIEZZA, che si misura rispetto al valore medio della pressione dell'aria.

- ⌘ Il valore 0 indicato sull'asse della pressione corrisponde al valore medio della pressione dell'aria.
- ⌘ Il grafico dell'andamento dice quindi che il valore della pressione aumenta gradualmente sino al tempo 1.25 msec in cui raggiunge il valore massimo di 100.
- ⌘ Quindi decresce sino al tempo 3.75 msec in cui raggiunge il valore minimo -100 e ritorna a zero al tempo 5.0 msec prima di iniziare il secondo ciclo.



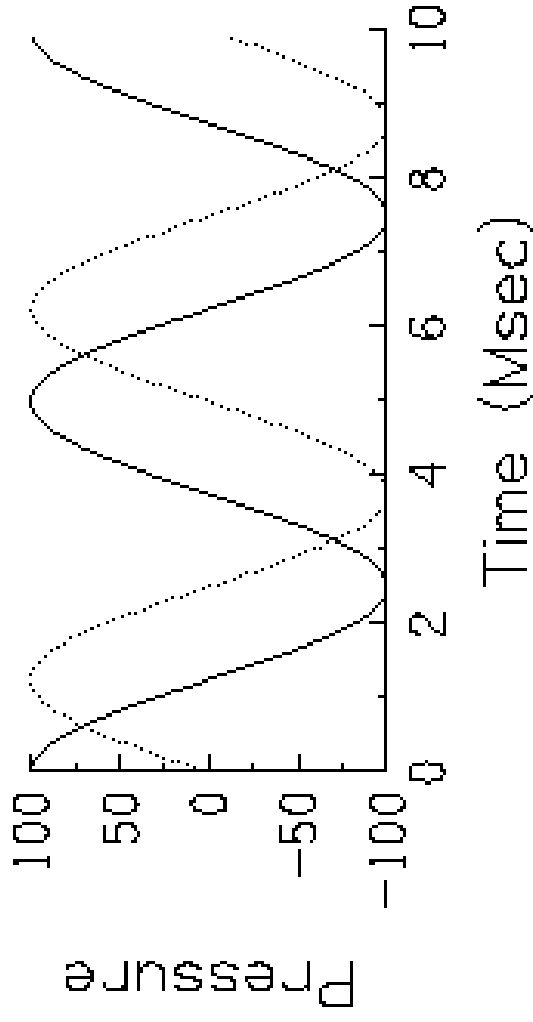


⌘ Ampiezza e frequenza della forma d'onda hanno effetto sul suono percepito

☒ Variazioni di piccola ampiezza producono suoni di bassa intensità. L'intensità del suono cresce col crescere della sua ampiezza.

☒ La frequenza della forma d'onda è invece legata alla tonalità del suono percepito: al crescere della frequenza cresce il tono.

**La differenza di fase ha unicamente a che vedere con il fatto che le due funzioni siano diversamente allineate rispetto al tempo.**




# Proprietà fisiche e caratteristiche percettive

⌘ Le proprietà fisiche di ampiezza e frequenza corrispondono alle caratteristiche percettive di volume e tono. Tuttavia il legame tra queste grandezze non è costante.

Da un punto di vista percettivo la nostra capacità di percepire suoni è limitata a quelli di frequenza compresa tra 20 Hz e 22.000 Hz circa.

⌘ Inoltre, anche nei limiti di quelli che sono i suoni da noi percepiti, la relazione tra proprietà percettive 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).



⌘ Esempio: si noti come, sebbene gli incrementi di frequenza siano costanti (25 Hz), via via che la frequenza cresce, gli incrementi di tono sembrano diminuire.

⌘ 200Hz 225Hz 250Hz 275Hz 300Hz  
325Hz 350Hz 375Hz 400Hz 425Hz  
450Hz 475Hz





⌘ Di solito vengono quindi usate delle scale di rappresentazione tali da riflettere uguali differenze percettive.

- ☒ Per la frequenza, una di queste scale è la scala di Mel. Uguali differenze sulla scala di Mel corrispondono ad uguali differenze percepite di tono, ma non uguali differenze di frequenza.
- ☒ Analogamente per il volume, la scala di rappresentazione impiegata è in Decibel (dB). In questa scala le ampiezze sono rappresentate in scala logaritmica. Nella scala in decibel, 0.0 dB corrisponde ad un suono alla soglia dell'udibilità, mentre 130 dB ad un suono alla soglia del dolore. Incrementi di 1 dB corrispondono a Just Noticeable Differences di volume (definite come le più piccole differenze di volume che vengono percepite in media il 50% delle volte).

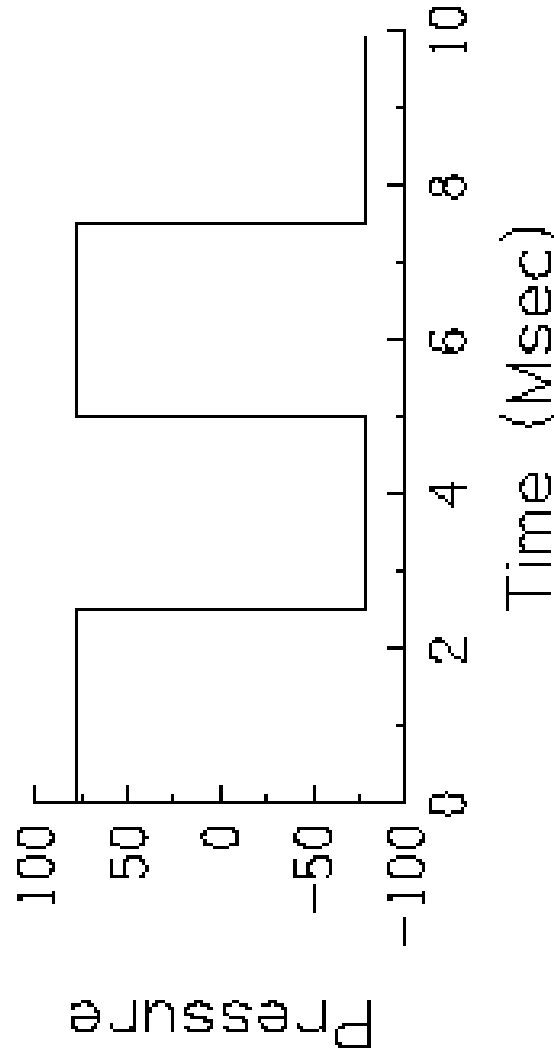


⌘ I suoni elementari hanno andamento sinusoidale, periodico, con estensione indefinita

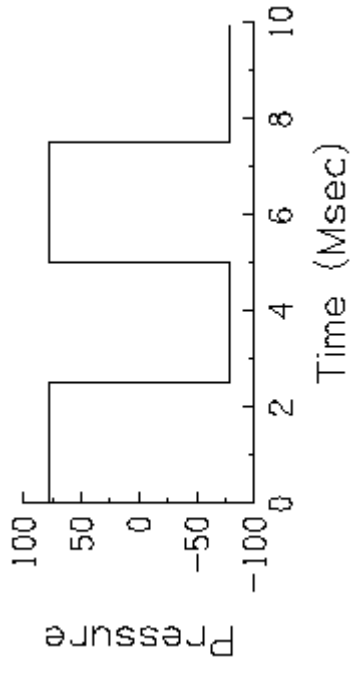
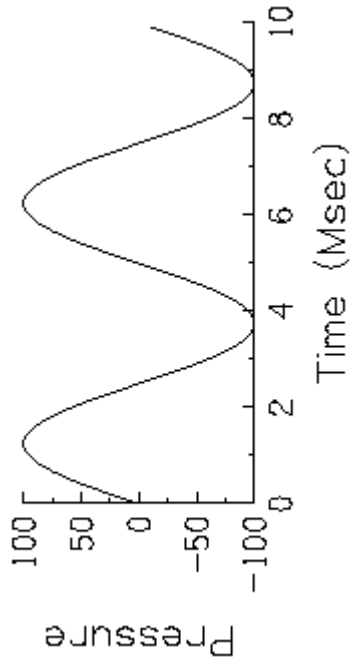
⌘ Tuttavia, la maggior parte dei suoni presenti in natura non possono essere descritti così semplicemente essendo caratterizzati da forme d'onda diverse da quelle sinusoidali.

⌘ Si può dimostrare che, fatte alcune ipotesi di regolarità sull'andamento della forma d'onda, un generico suono complesso può essere descritto come una combinazione di suoni elementari.

⌘ Si consideri per esempio il suono corrispondente alla forma d'onda qui si seguito riportata in figura. Tale suono e' caratterizzato dall'alternarsi di intervalli temporali in cui la pressione assume valori costanti `alti' e `bassi'.



Confronto tra due onde con la stessa frequenza: sebbene il tono delle due forme d'onda sia lo stesso, esse hanno un diverso timbro.






⌘ Ogni suono può essere descritto come la somma di una serie di suoni elementari.

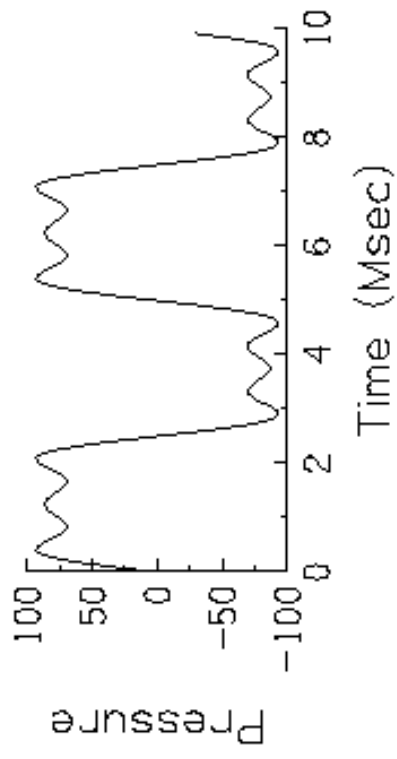
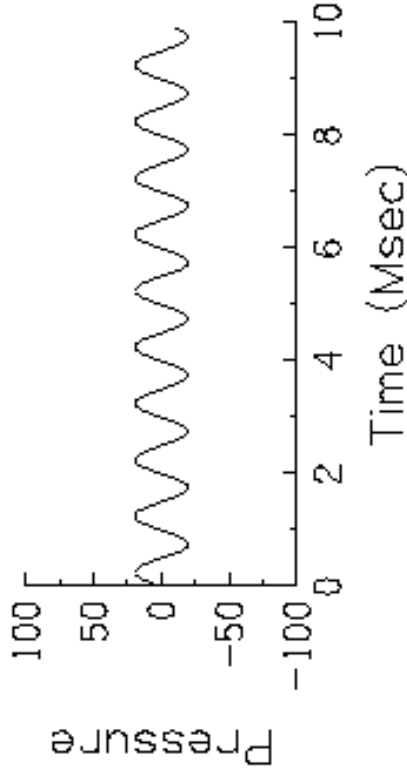
⌘ L'onda quadra può essere ottenuta sommando onde sinusoidali di opportuna frequenza, ampiezza e fase

⌘ Comprensibilmente, il primo e più rilevante contributo è dato dalla sinusoide di frequenza pari a quella dell'onda quadra: 200 Hz. Questa componente prende il nome di FREQUENZA FONDAMENTALE (F0) ed è quella che determina il tono del suono percepito. Quindi la componente F0 è quella che causa lo stesso tono percepito per la sinusoide a 200 Hz e per l'onda quadra a 200 Hz.

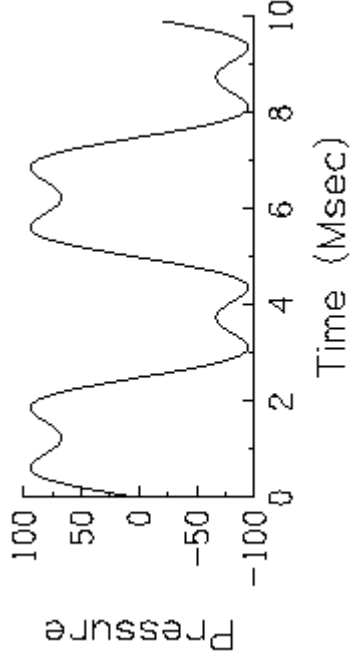
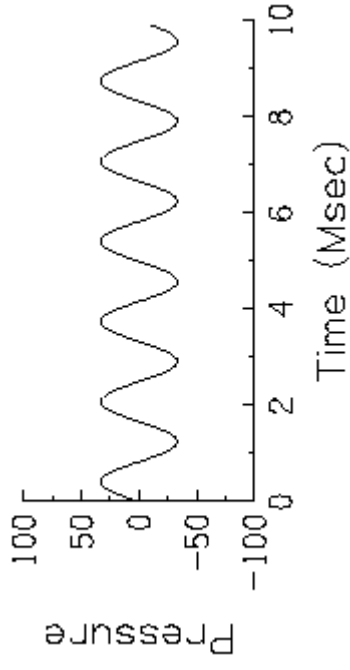
⌘ Tuttavia, per costruire l'onda quadra sono necessarie anche altre componenti elementari di frequenza maggiore. Queste altre componenti vengono dette ARMONICHE ed hanno tutte frequenza multipla di F0, quindi 400 Hz, 600 Hz, 800 Hz e così via.

- 
- ⌘ L'onda quadra e' un caso particolare in cui tutte le armoniche pari sono nulle. Qui l'onda quadra e' data dalla somma delle componenti  $F_0$ ,  $3 \cdot F_0$ ,  $5 \cdot F_0$  ecc. vale a dire 200Hz, 600Hz, 1KHz, 1.4KHz ecc.
  - ⌘ Ognuna di queste componenti ha una diversa ampiezza rispetto a quella dell'armonica fondamentale

Continuando nel processo di sommare contributi di armoniche di frequenze via via crescenti, si arriva alla terza componente (la 5a armonica che ha 1000Hz) la cui ampiezza deve essere  $1/5$  della fondamentale.

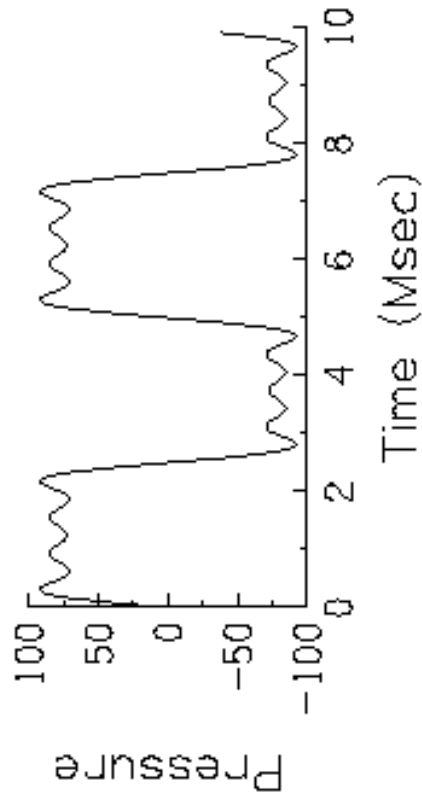
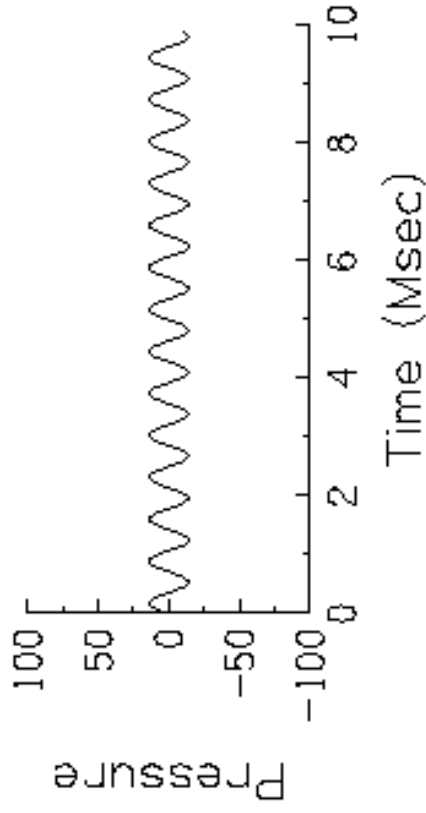



L'onda ottenuta sommando i contributi della sinusoide a 200Hz con quella a 600Hz ad  $1/3$  della sua ampiezza (a sx) e' la seguente (a dx):





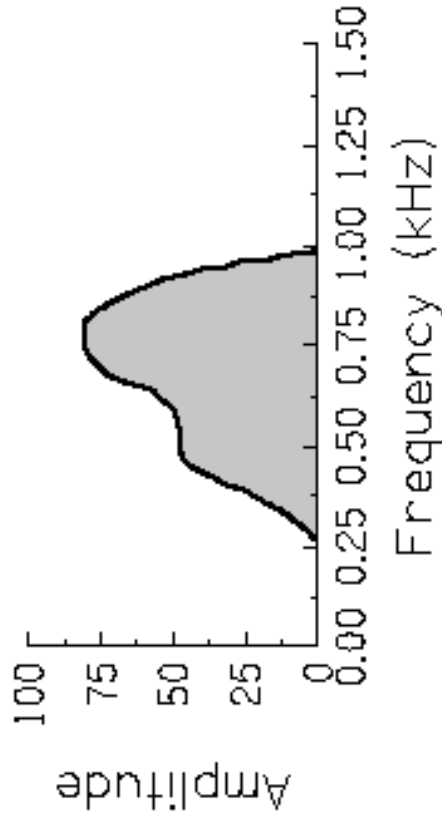
Quindi abbiamo la 7a armonica (1.4KHz) con ampiezza 1/7 della fondamentale



- 
- ⌘ La presentazione sino ad ora effettuata e' basata su una descrizione dell'effetto sonoro in base a variazioni di pressione nel tempo.
  - ⌘ La rappresentazione nel dominio del tempo e' sicuramente importante ma per molti versi non rende esplicite alcune caratteristiche che sono invece di basilare importanza nella determinazione dell'effetto sonoro.
  - ⌘ Una di queste caratteristiche e' la frequenza, il cui valore e' di difficile derivazione da una rappresentazione della forma d'onda nel dominio del tempo.

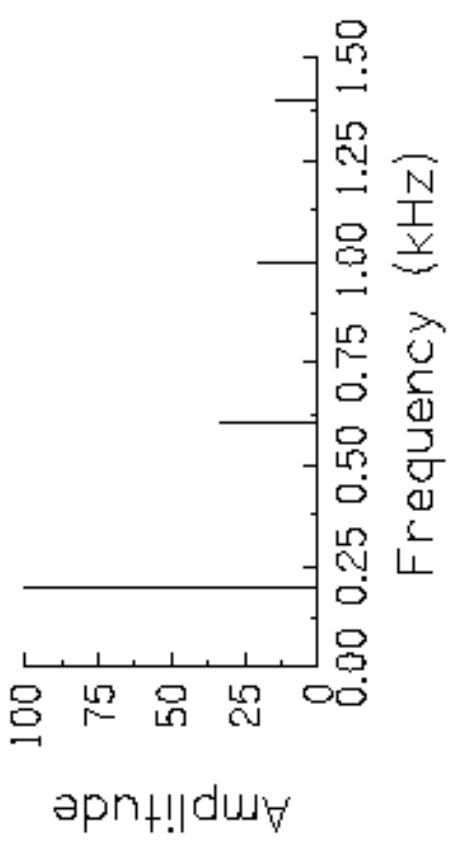
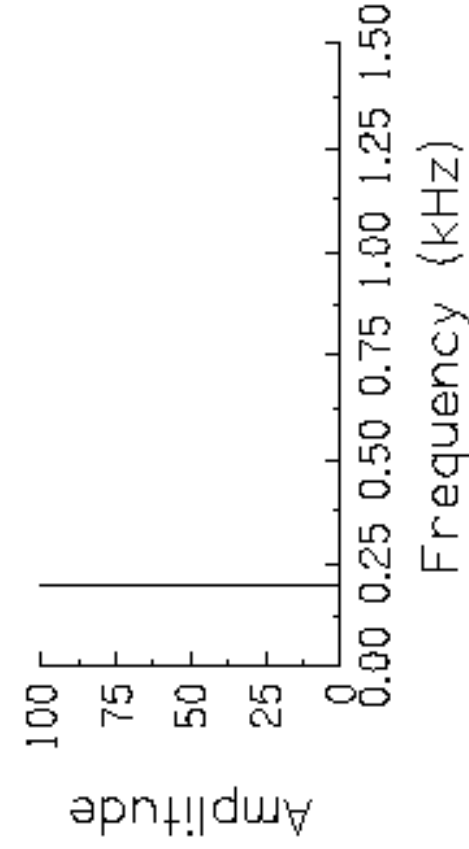
⌘ Per far fronte a queste difficoltà e' stata introdotta in una diversa rappresentazione del suono detta appunto rappresentazione nel dominio della frequenza.


⌘ In questa rappresentazione, l'asse dei tempi viene sostituito dall'asse delle frequenze e la rappresentazione di un generico suono e' effettuata con un grafico che riporta i valori di ampiezza di ciascun contributo in frequenza.



# Spettro Discreto e Spettro Continuo

⌘ Qui di seguito e' riportata la rappresentazione nel dominio della frequenza dell'ampiezza della forma d'onda sinusoidale a 200 Hz e dell'onda quadra.

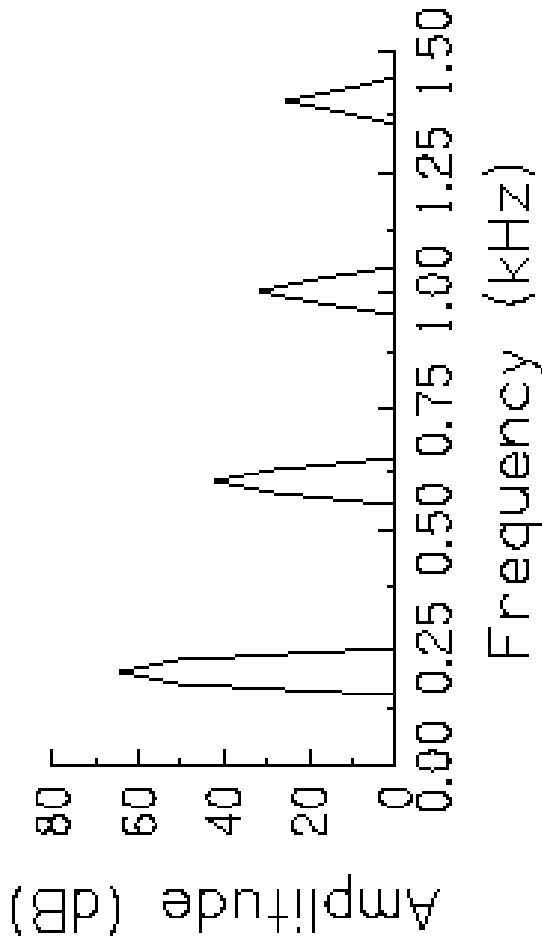





⌘ La presenza di una linea nello spettro in frequenza indica la presenza di un segnale esattamente sinusoidale, tuttavia, i suoni caratterizzati da uno spettro discreto sono pochi e la maggior parte dei suoni presenti in natura presentano un andamento non periodico.


⌘ Questo perché i suoni che normalmente si sentono hanno un inizio ed una fine (non sono quindi perfettamente periodici, dal momento che questa caratteristica richiede che il segnale si ripeta periodicamente all'infinito) vale a dire sono contenuti in un intervallo temporale finito.

⌘ Qui di seguito e' descritta la rappresentazione in frequenza della forma d'onda quadra troncata al di fuori di un intervallo temporale finito.





⌘ Un'altra causa di scostamento rispetto ad un perfetto andamento periodico e' rappresentata dal fatto che la maggior parte dei suoni, anche all'interno dell'intervallo temporale in cui sono definiti, non hanno un andamento periodico ma QUASIPERIODICO, vale a dire presentano delle lievi differenze tra un periodo e l'altro.

- 
- ⌘ La più importante differenza tra questa figura e la precedente è che quelle che prima erano delle linee armoniche, ora si sono allargate in una forma a campana.
  - ⌘ Rappresentano sempre delle armoniche ma non associate a suoni puri. Rappresentano un suono associato ed una energia distribuita attorno a quella dell'armonica pura.
  - ⌘ Il suono a cui questa rappresentazione si riferisce si caratterizza per avere un inizio ed una fine graduali invece che istantanei.
  - ⌘ Spettri di questo tipo vengono chiamati spettri armonici invece che linee spettrali.






⌘ La differenza tra le linee e le campane riflette una differenza fondamentale tra la rappresentazione nel dominio del tempo e quella nel dominio della frequenza.


⌘ I suoni che si estendono per tempi molto lunghi e con andamenti fortemente periodici nel dominio del tempo, sono caratterizzati da profili molto concentrati nel dominio della frequenza (una sinusoide di durata infinita e' caratterizzata da una unica linea spettrale). Tuttavia, suoni definiti su intervalli temporali ristretti manifestano un profilo molto allargato nel dominio della frequenza.

⌘ La situazione estrema e' quella rappresentata da un suono puramente impulsivo: con ampiezza nulla ovunque a parte in un istante temporale. Un suono di questo tipo e' caratterizzato da uno spettro in frequenza che ha valore costante non nullo (spettro completamente piatto).



⌘ Uno spettro di questo tipo non è detto spettro armonico ma spettro continuo.

⌘ Gli spettri continui sono associati a suoni con andamento non periodico: aperiodico.

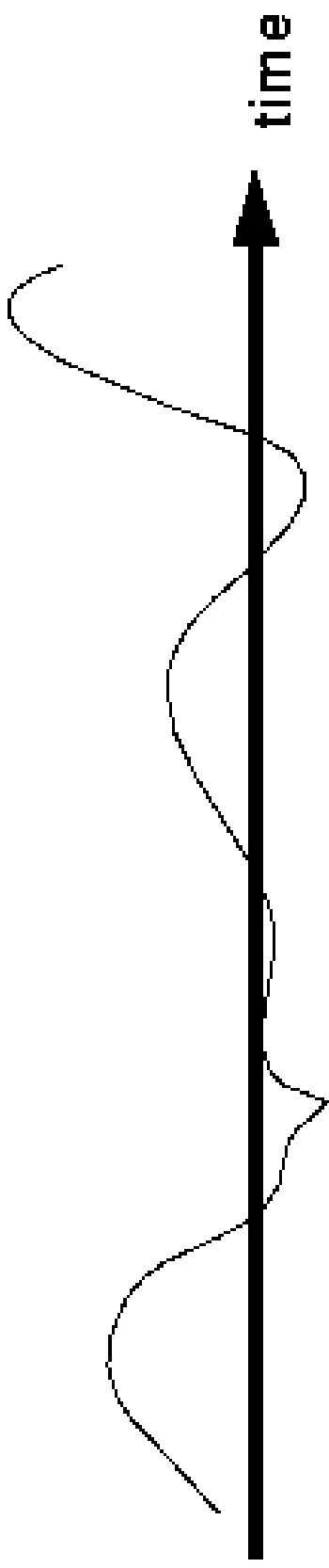



⌘ I suoni sono in generale caratterizzati da uno spettro continuo. Pertanto possono essere pensati come composti da tanti suoni elementari a frequenze ed ampiezze diverse. La presenza dei contributi a frequenze elevate rende il suono piu' brillante, mentre la loro assenza lo rende piu' cupo.

⌘ Le problematiche relative alla elaborazione del suono, sia in forma analogica che digitale sono connesse al fatto che tale elaborazione porta sempre (in misura variabile) ad un cambiamento dello spettro originale e quindi del suono ad esso associato.

# SUONO DIGITALE

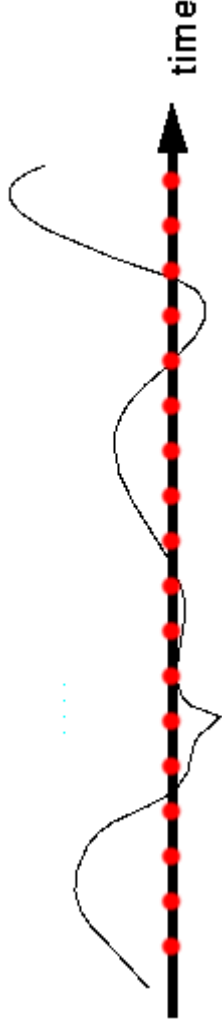
🔊 I microfoni producono rappresentazioni analogiche del segnale audio. Questo e' infatti rappresentato da un valore di tensione il cui andamento nel tempo riflette le oscillazioni di pressione nell'aria.



- 
- ⌘ Per poter rappresentare il suono in un sistema digitale bisogna prima convertirlo in un flusso di numeri rappresentati in forma binaria.
  - ⌘ 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.
  - ⌘ Inoltre, l'ampiezza di ogni campione, dovendo essere rappresentata digitalmente (cioè con una codifica binaria), non potrà assumere infiniti valori.
  - ⌘ La conversione Analogico-Digitale richiede pertanto un processo di discretizzazione sia nel tempo (campionamento) che in ampiezza (quantizzazione).

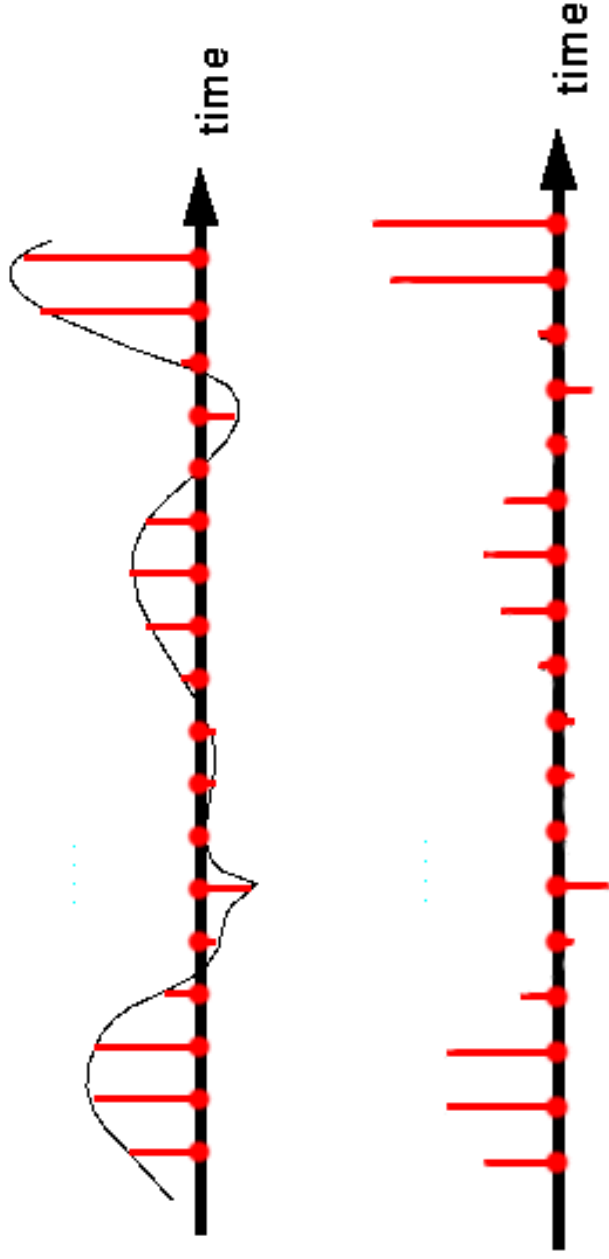
# Campionamento


✂ Divisione dell'asse orizzontale (quello temporale) in unità discrete. Il numero di campioni individuato in un secondo prende il nome di frequenza di campionamento.



# Quantizzazione

✂ divisione dell'asse verticale (quello delle ampiezze) in unita' di ampiezza costante





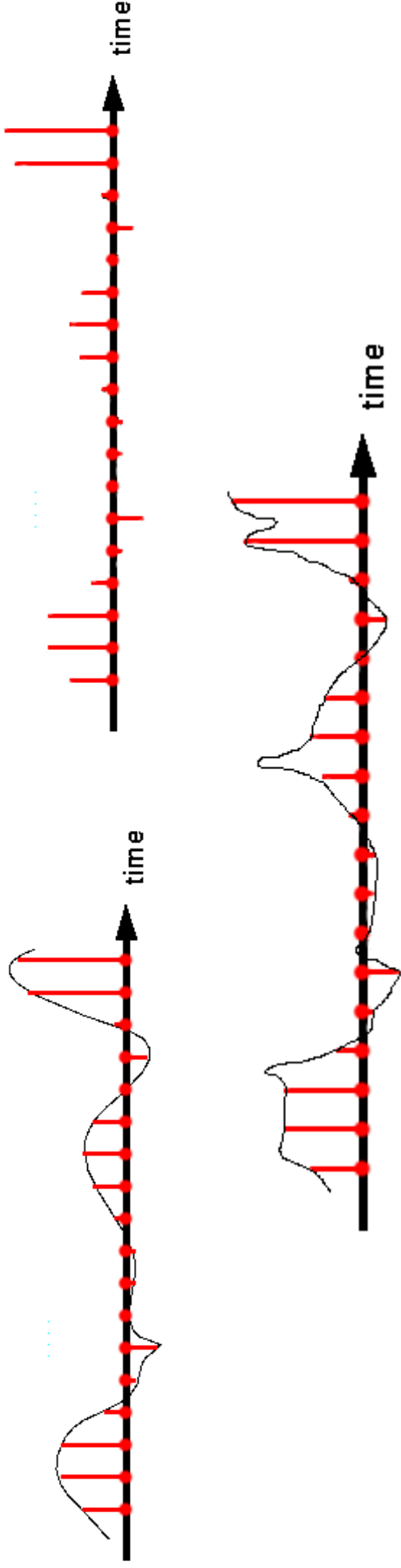
⌘ Per la conversione Analogico-Digitale e' quindi necessaria la specifica di 2 parametri relativi a:

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



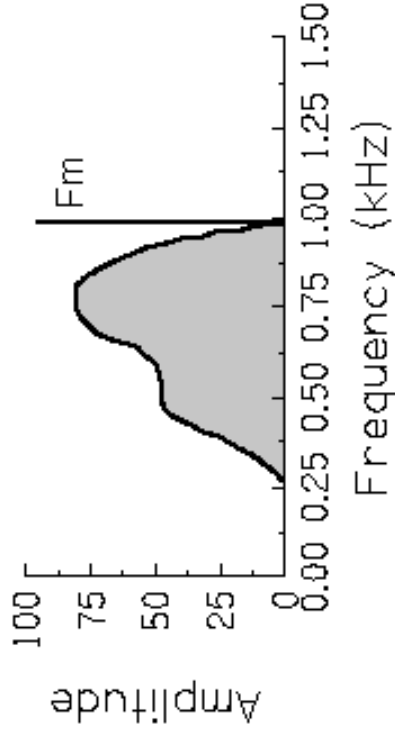
# RICOSTRUZIONE DEL SEGNALE

✂ Una volta estratti i campioni chi mi garantisce che questi individuino univocamente la forma d'onda di partenza?



# Teorema di Nyquist

- ⌘ Per avere una digitalizzazione senza perdita di informazione e' necessario campionare con una frequenza almeno il doppio della massima frequenza che compare nello spettro della forma d'onda da acquisire: La frequenza di campionamento di un segnale il cui spettro in frequenza abbia componenti non nulle sino a  $F_m$  dovra' essere pari almeno a  $2 \cdot F_m$ .





✂ Sistema uditivo in grado di percepire suoni di frequenza compresa tra 20Hz e 22KHz.

☑ Range del parlato: 500Hz - 2KHz

✂ un campionamento con frequenza maggiore di 44KHz dovrebbe garantire la perfetta riproducibilità del segnale audio.

# Rapporto Segnale Rumore (SNR)

- ⌘ Con qualsiasi rappresentazione analogica una parte del segnale impiegato per rappresentare la grandezza e' dovuta al rumore.
- ⌘ Il rapporto tra la potenza del segnale e quella del rumore e' chiamato *signal to noise ratio* (*SNR*).
- ⌘ SNR e' una misura della qualita' del segnale e si misura in decibel (dB):

$$SNR = 10 \log \frac{V_{signal}^2}{V_{noise}^2} = 20 \log \frac{V_{signal}}{V_{noise}}$$

# Precisione di quantizzazione

- ⌘ Tipicamente vengono usati 8 o 16 bit. Con 8 bits si hanno 256 distinti valori, con 16 bits si hanno 65536 distinti valori.
- ⌘ Ogni bit aggiunge approssimativamente 6 dB di risoluzione, per cui 16 bits consentono di avere un rapporto segnale rumore  $\text{SNR} = 96 \text{ dB}$ .  
(  $V_{\text{signal}} = 2 \exp 16$ ,  $V_{\text{noise}} = 1$ .  $\Rightarrow 20 \times \log(65536) = 96 \text{ dB}$ .)
- ⌘ I campioni possono essere rappresentati in scala lineare (*linear format*), o logaritmica ( *$\mu$ -law* (or *A-law* in Europe)) per ricondursi ad una scala di rappresentazione percettiva

# Audio Quality vs. Data Rate

Quality'	Freq. Camp.	Bit/sample	Mono/Stereo	Data Rate
Telephone	8 KHz	8	Mono	64Kb/s
AM Radio	11.025KHz	8	Mono	88.2Kb/s
FM Radio	22.050KHz	16	Stereo	705.6Kb/s
CD	44.1KHz	16	Stereo	1.411Mb/s
DAT	48KHz	16	Stereo	1.536Mb/s

# Dimensione del file Audio

⌘ La seguente formula da' una stima della dimensione del file rappresentante un minuto di traccia audio:

$$\text{(Frequenza in Hz)} * (\text{n. bit per campione}) / 8 = \text{Bytes/sec}$$

$$\text{Bytes/sec} * 60 = \text{Bytes/min}$$

# Dimensione del file Audio

⌘ Di seguito sono riportate le dimensioni del file audio rappresentante una singola traccia di 5 minuti, per diversi valori di frequenza di campionamento e numero di bit per campione

BITS PER SAMPLE		SAMPLING FREQUENCY RATE		
16 bits	8 bits	11.025 kHz	22.05 kHz	44.1 kHz
		3.30 MB	6.615 MB	13.2 MB
		6.615 MB	13.2 MB	26.46 MB



# Formati e supporti

⌘ 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;

# Formati

- ⌘ 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
  - dati di copyright

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

# 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

# Supporti


- ⌘ I campioni del segnale digitale possono essere registrati su una moltitudine di supporti, con caratteristiche fisiche e logiche diverse:
- fisica: supporti magnetici, ottici elettronici;
  - logica: ogni supporto è caratterizzato da uno standard proprio per la struttura dei dati

# Supporti per l'audio digitale

⌘ CD-AUDIO:  $F_c=44,1$  KHz, quantizzazione 16 bit, segnale stereo, velocità di trasferimento minima:  
 $44.100 * 2 * 16 = 1.411.200$  bit/sec;

⌘ CD-ROM: applicazioni multimediali, 650 MB di dati, diversi standard (Libro Giallo, Rosso, etc.);

- ⌘ DVD-ROM: sola lettura, da 4,7 Gb a 17 Gb di capacità, lettori più lenti di quelli per CD ma i dati sono più vicini e quindi velocità di trasferimento è maggiore;
- ⌘ DVD-Video: formato per i film in digitale, contiene sistema di protezione basato su CSS (content scrambling system), formato video MPEG-2;
- ⌘ DVD scrivibili: 5 formati diversi in funzione del supporto fisico e delle specifiche dei dati;
- ⌘ DVD-Audio: competitore del SuperCD (Sony-Philips), audio di qualità multicanale con protezione, contenuti extra (immagini, testo, menu, video, etc.), connessione a Internet, Fc fino a 192 KHz, quantizzazione fino a 24 bit;



⌘ DAT: nasce come formato audio per poi divenire nel 1998 un formato generale per l'archiviazione dei dati, nastro magnetico di 4 o 8 mm, formati variabili da 2 a 40 Gb;

⌘ Hard-Disk: la memorizzazione dei dati avviene tramite polarizzazione di sostanze ferromagnetiche, le performance dipendono dall'interfaccia verso il processore (IDE, ATA, SCSI, USB esterne più lente);

# Musica e MIDI

- ⌘ La relazione tra computer e musica si è rafforzata grazie allo sviluppo del MIDI (Music Instrument Digital Interface)
- ⌘ L'interfaccia è collegata direttamente alla porta seriale del computer e consente la trasmissione di segnali musicali
- ⌘ Il MIDI è uno standard riconosciuto dai costruttori di strumenti musicali





## ⌘ Un'interfaccia MIDI ha due diversi componenti:

- ☒ Hw: specifica il collegamento fisico tra gli strumenti musicali, specifica la porta MIDI, specifica il cavo MIDI e gestisce i segnali elettronici che vengono trasmessi tramite il cavo stesso.
- ☒ Sw: specifica il formato dei dati che codifica l'informazioni che viaggia attraverso l'hardware. Un dato in formato MIDI non codifica i singoli campioni, come fa invece il formato audio. La codifica contiene le specifiche degli strumenti, la definizione dell'inizio e della fine di una nota, la frequenza di base e il volume del suono.

# Formato dei dati MIDI

⌘ Il formato dei dati MIDI è digitale: i dati vengono raggruppati in messaggi MIDI.

☑ Ogni messaggio MIDI comunica un evento musicale tra i diversi dispositivi collegati. Questi eventi sono tipicamente azioni che provoca un musicista mentre suona uno strumento musicale. (esempio: - quando un musicista preme il tasto di una tastiera elettronica l'interfaccia MIDI crea un messaggio MIDI codificando l'inizio della nota e l'intensità della pressione; Il messaggio viene trasmesso; Quando il tasto viene rilasciato, viene creato un nuovo messaggio MIDI, contenente il segnale corrispondente 10' di musica sono circa 200KB di dati MIDI;

# Compressione Audio

- ⌘ Le piu' 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 intensita', 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 piu' che di tutte le componenti presenti nella banda, solo di quelle che sono effettivamente udibili (non mascherate).

# **Schemi di compressione semplici**

- ⌘ **Gli schemi di compressione semplici sono quelli con perdita di informazione che si basano su algoritmi non sofisticati**
- ⌘ **Non raggiungono grandi compressioni, ma sono di rapida esecuzione e semplici da implementare.**
- ⌘ **Esempi: standard  $\mu$ -law e A-law, schema ADPCM**

# **$\mu$ -law e 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

# Codifica ADPCM

- ⌘ La codifica ADPCM (adaptive differential pulse code modulation) è un metodo comune di compressione che raggiunge un buon compromesso tra velocità di elaborazione, qualità del segnale decodificato e rapporto di compressione;
- ⌘ Si basa su due concetti:
  - codifica dell differenza tra campioni (campioni adiacenti sono molto simili nel valore);
  - adattamento delle differenze allo specifico segnale di input
- ⌘ Invece di rappresentare ogni campione audio in modo indipendente, rappresenta la differenza tra un campione e un valore predetto a partire dal campione precedente e da un insieme di differenze quantizzate adattate sul segnale di input

# Schemi di compressione di tipo percettivo

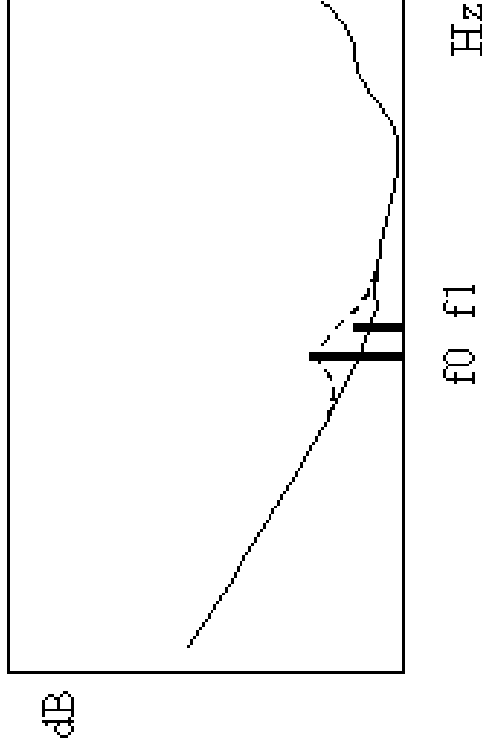


- ⌘ Gli schemi di compressione di questo tipo comprimono il segnale eliminando quelle parti che il nostro apparato uditivo non percepirebbe
- ⌘ Sono codifiche di tipo lossy
- ⌘ Tali schemi si basano sull'effetto di mascheramento

# Effetto di mascheramento

⌘ Una forte componente di suono può 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.



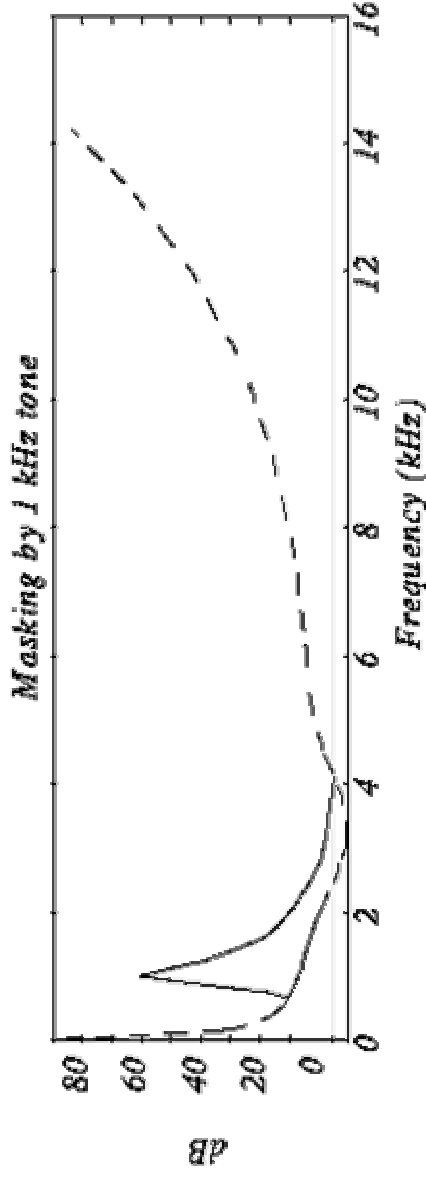


# Effetto di mascheramento

⌘ L'effetto di masking varia con la frequenza:

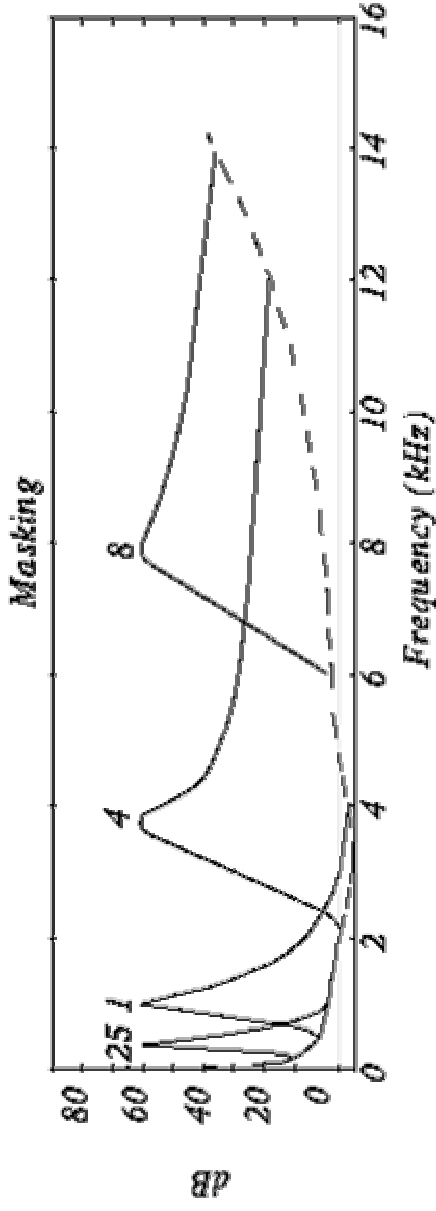
✗ 100Hz per  $f_0 = 20\text{Hz}$

✗ 4KHz per  $f_0 = 20\text{KHz}$ .



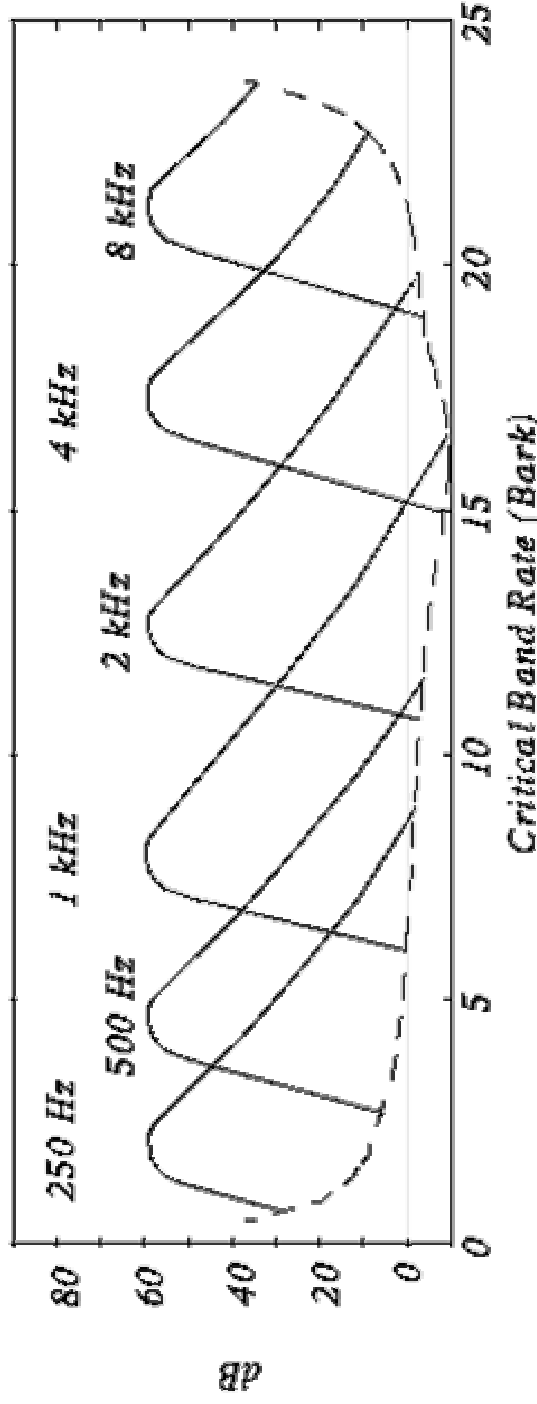
# Effetto di mascheramento

- ⌘ Circa 100Hz per frequenze di mascheramento < 500Hz, cresce per frequenze superiori.



# Banda critica

- ⌘ Il range di frequenza per cui si manifesta l'effetto di mascheramento è detto *banda critica*.
- ⌘ L'introduzione delle bande critiche permette di misurare la frequenza in un modo che sia uniforme dal punto di vista percettivo
- ⌘ Nuova unità di frequenza:
  - ⌘ 1 *Bark* = ampiezza della banda critica.
  - ⌘ Per frequenze  $< 500\text{Hz}$ ,  $1 \text{ Bark} \approx \text{freq}/100$
  - ⌘ per frequenze  $> 500\text{Hz}$ ,  $1 \text{ Bark} \approx 9 + 4\log(\text{freq}/1000)$



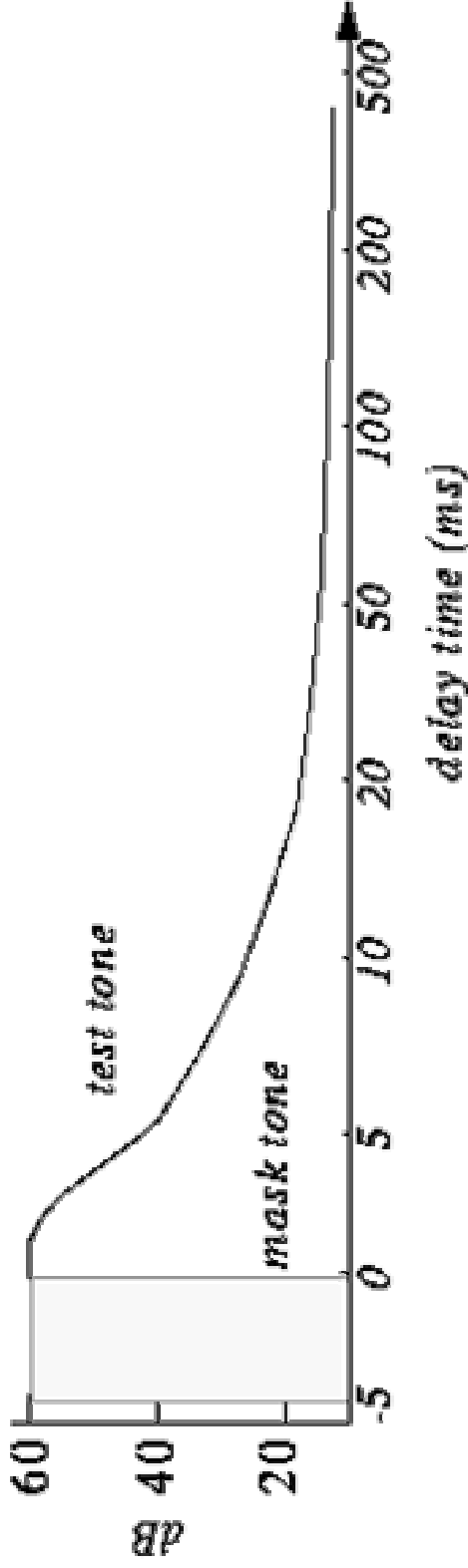
# Mascheramento temporale

⌘ Se ascoltiamo un suono di forte intensità, è necessario un certo tempo prima che un tono lieve possa essere percepito.

☒ tono di mascheramento di 60dB a 1KHz;

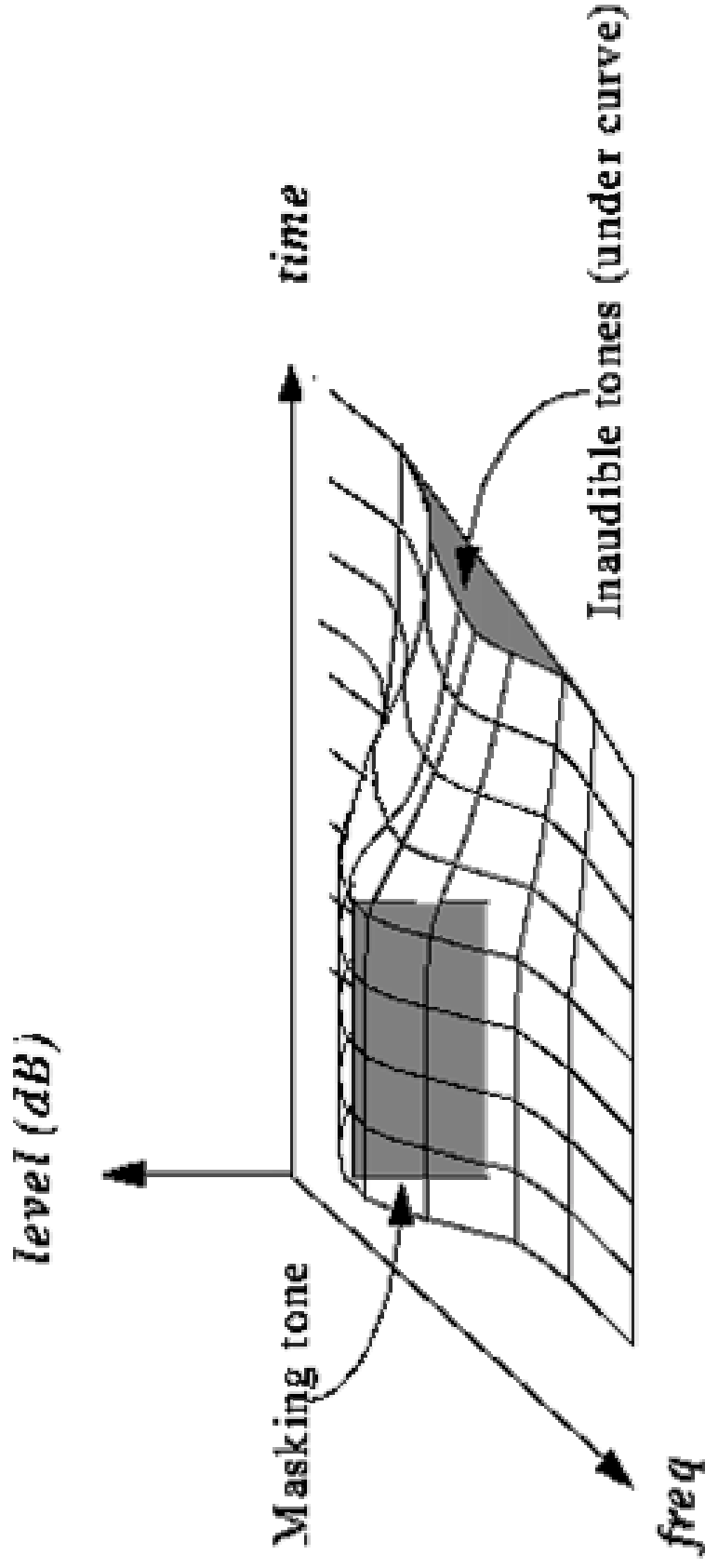
☒ tono di test a 1.1KHz a 40 dB;

⌘ Ritardo di ascolto al variare dell'intensità del tono di test.



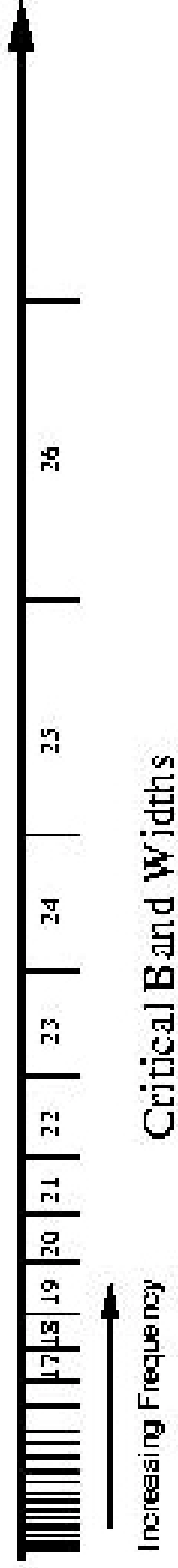
# Effetto complessivo di mascheramento

⌘ Variando la frequenza del tono di test, con tono di mascheramento di durata costante si ottiene il grafico seguente.



# Banda critica

⌘ Variazione della banda critica con la frequenza.



# MPEG-1 audio

- ⌘ MPEG-1: 1.5 Mbits/sec per audio e video.
  - ☒ Circa 1.2 Mbits/sec per la codifica video, 0.3 Mbits/sec per la codifica audio.
  - ☒ Come confronto un CD audio non compresso ha un bit-rate:  $44100 \text{ campioni/sec} * 16 \text{ bits/campione} * 2 \text{ canali} > 1.4 \text{ Mbits/sec}$ .
- ⌘ Fattore di compressione tra 2.7 e 24.
- ⌘ Con fattore di compressione 6:1 (16 bit stereo campionati a 48KHz ridotti a 256 Kbits/sec) e condizioni di ascolto ottimali, ascoltatori esperti non sono in grado di distinguere tra clip audio codificati e originali.

# MPEG-1 audio

- ⌘ Layer III dello standard Mpeg-1: è il formato comunemente indicato come mp3.
- ⌘ Prevede la compressione di uno stream con le seguenti frequenze di campionamento: 32KHz, 44.1KHz, 48KHz.
- ⌘ Il formato audio di ingresso può essere in uno di quattro possibili modi:
  - ☒ audio mono;
  - ☒ due canali per due tracce mono;
  - ☒ due canali per una traccia stereo;
  - ☒ due canali per una traccia joint-stereo (si avvantaggia della correlazione tra canali stereo).



# MPEG-1 audio

⌘ Sono previsti tre distinti livelli di compressione:

- ✒ *layer I*: è il più semplice e maggiormente indicato qualora non sia richiesto uno stream di uscita inferiore a 128Kbit/s;
- ✒ *layer II*: livello di complessità intermedio da impiegarsi qualora siano richiesti flussi di uscita di poco inferiore a 128Kbit/s. Le schede di acquisizione MPEG in real-time tipicamente supportano questo livello di compressione audio;
- ✒ *layer III*: è il più complesso e l'unico in grado di scendere sotto i 64Kbit/s mantenendo una elevata qualità audio.

⌘ Per tutti e tre i livelli esistono chip che effettuano la decodifica in real-time.

# Audio layer

Rapporto di compressione

Mpeg layer

1:4

384 kbps stereo

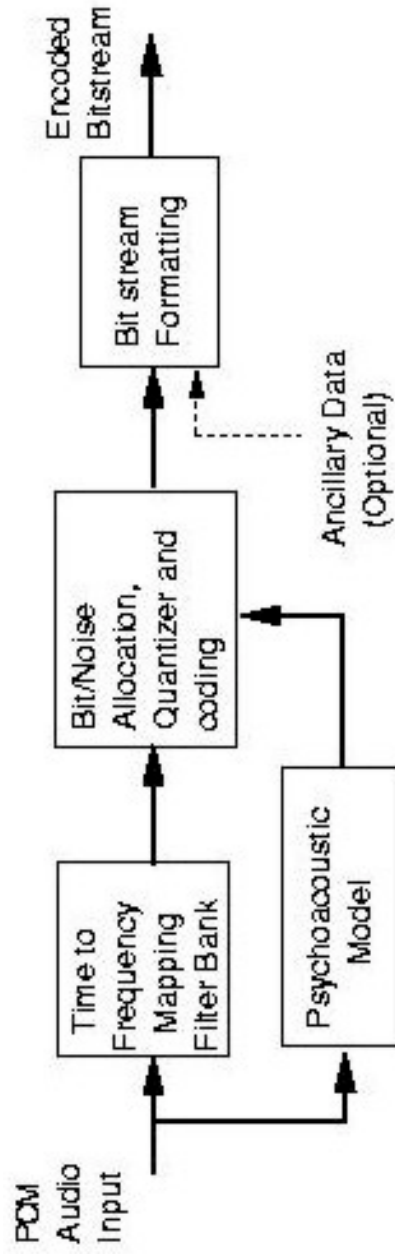
1:6, 1:8

256 - 192 kbps stereo

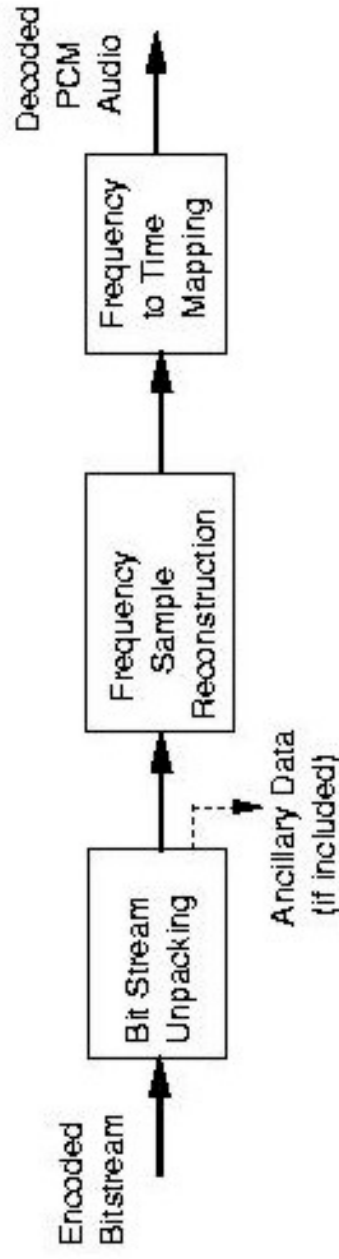
1:10, 1:12

128 - 112 kbps stereo

# MPEG-1 audio



## MPEG/Audio Encoder



# Algoritmo di codifica

⌘ Strategia di compressione basata sulla quantizzazione.

⌘ Obiettivo:

☒ rendere l'errore di quantizzazione non udibile.

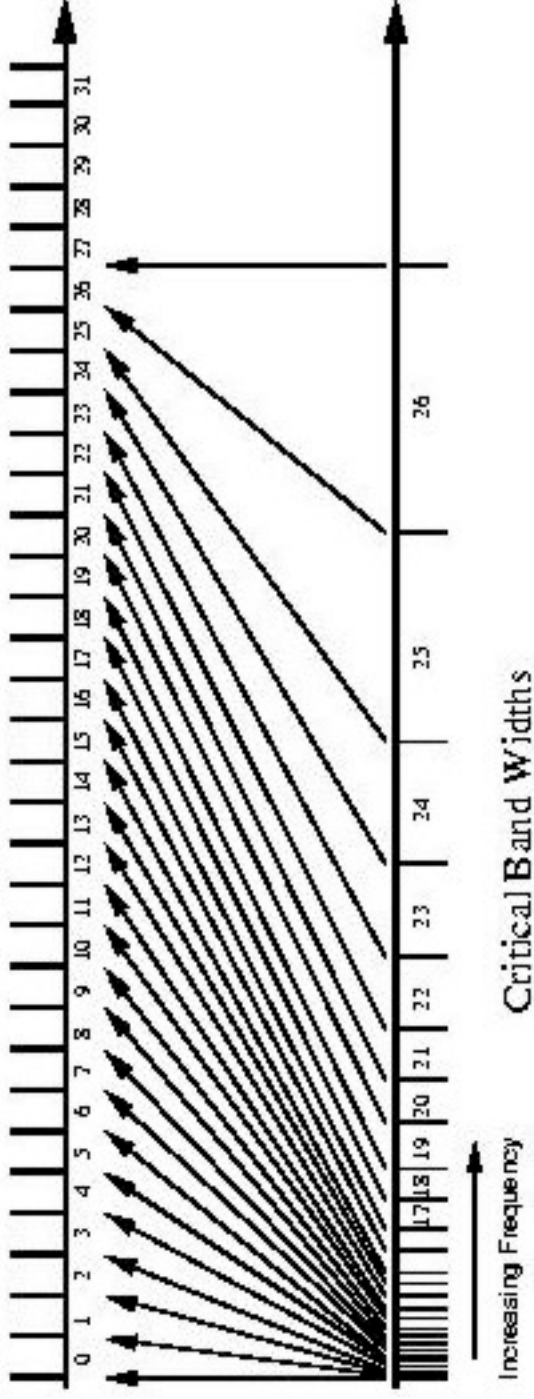
# Passi dell'algoritmo di codifica

- ⌘ 1. Filtri di convoluzione per dividere il segnale audio in sottobande di frequenza che approssimano le 32 bande critiche.
- ⌘ 2. Determinazione dell'ammontare di mascheramento per ciascuna banda causato da bande vicine usando il modello psicoacustico.
- ⌘ 3.1 Se la potenza di una banda è sotto la soglia di mascheramento non è codificata.
- ⌘ 3.2 Altrimenti determina il numero di bit necessari a rappresentare il coefficiente in modo tale che il rumore introdotto dalla quantizzazione sia sotto l'effetto di mascheramento (1 bit di quantizzazione introduce 6db di rumore:  $6\text{dB} = 20 \log_{10} 2$ ).
- ⌘ 4. Formattazione dello stream di bit.

# Banco di filtri

- ⌘ Scomponere il segnale di ingresso in 32 bande di frequenza di uguale ampiezza.

MPEG/Audio Filter Bank Bands



# Banco di filtri

⌘ Introduce tre approssimazioni:

✗ le 32 bande dovrebbero simulare le bande critiche. Tuttavia queste ultime non hanno ampiezza costante, mentre l'ampiezza delle 32 bande è costante.

✗ A bassa frequenza più bande critiche sono coperte dalla stessa banda: il numero di bit di quantizzazione calcolato per la banda non può adattarsi in modo efficiente alle diverse bande critiche. È la banda critica con il minimo effetto di mascheramento che vincola i bit di quantizzazione per l'intera banda.

✗ Ad alta frequenza una banda critica corrisponde a diverse bande: in questo caso si ha semplicemente uno spreco di risorse di calcolo (un'unica banda sarebbe sufficiente).

# Banco di filtri

- ✗ La precisione finita con cui viene implementato il filtraggio introduce un errore di approssimazione non recuperabile;
- ✗ il numero limitato di coefficienti di ciascun filtro provoca uno scostamento della risposta del filtro rispetto al caso ideale: filtri adiacenti hanno una sovrapposizione in frequenza così che il segnale ad una singola frequenza può provocare una risposta non nulla in uscita a più di un filtro.

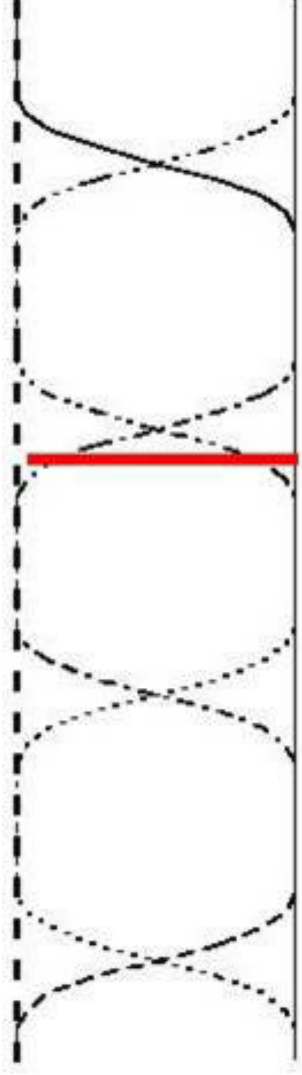


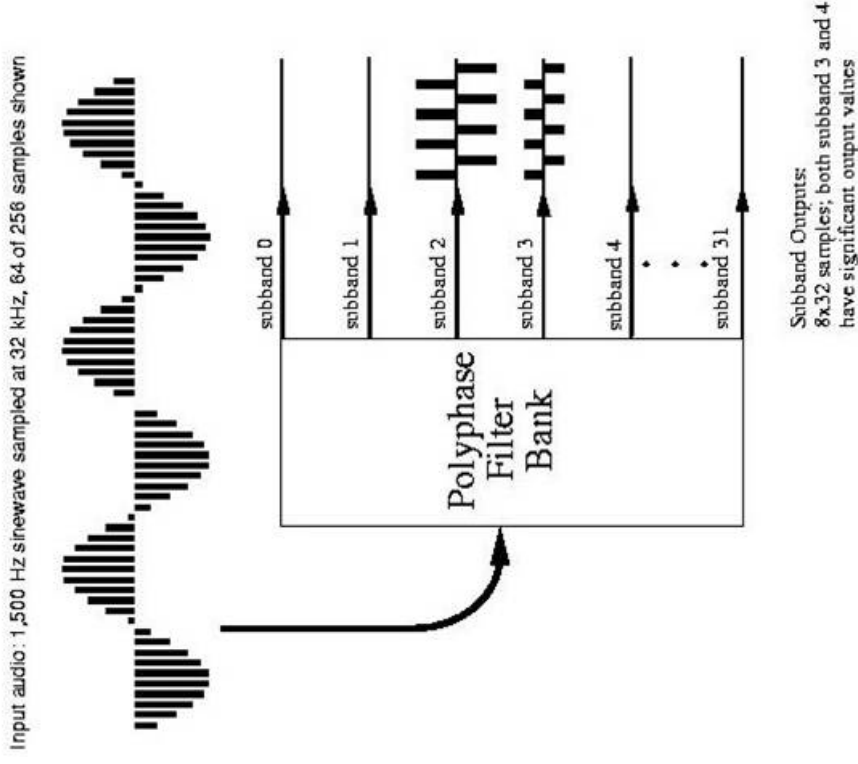
Figure 8 Composite Frequency Response of the Polyphase Filter Bank for 3 Subbands



# Banco di filtri

⌘ La risposta del filtro / del banco corrisponde ad un passa-banda traslato in frequenza.

⌘ Nell'esempio, un tono sinusoidale puro, che ha energia ad una sola frequenza, appare all'uscita di due filtri.



# Modello psicoacustico

- ⌘ Analizza il segnale audio e calcola il livello di masking in funzione della frequenza.
- ⌘ Lo standard prevede due possibili implementazioni del modello psicoacustico:
  - ✒ *Model I*, più semplice;
  - ✒ *Model II*, particolarmente indicato per la compressione Layer III.

# Modello psicoacustico

## ⌘ Esegue le seguenti operazioni:

- ☒ conversione del segnale audio nel dominio della frequenza;
- ☒ raggruppa le componenti spettrali sulla base delle bande critiche;
- ☒ applica una spreading function per calcolare l'effetto di masking di ogni componente;
- ☒ determina il valore minimo della soglia di udibilità;
- ☒ calcola la soglia di udibilità per ogni sottobanda:
  - ☒ *Model I* seleziona la minima soglia di mascheramento all'interno di ogni sottobanda;
  - ☒ *Model II* seleziona la minima soglia di mascheramento all'interno di ogni sottobanda solo se la sottobanda ha dimensione simile alla banda critica corrispondente. Altrimenti esegue una media delle soglie di mascheramento.

# Allocazione dei bit

⌘ È fatta a livello di banda:

- ☒ se la potenza in una banda è sotto la soglia di mascheramento non la codifica;
- ☒ altrimenti determina il numero di bit necessari a rappresentare il coefficiente tali che il rumore introdotto dalla quantizzazione sia sotto l'effetto di mascheramento (1 bit di quantizzazione introduce circa 6db di rumore).

# Algoritmo di allocazione dei bit

1. Calcola *Mask-to-Noise* ( $MNR$ ) ratio:

$$MNR_{db} = SNR_{db} - SMR_{db}$$

dove:

✘  $SNR$ : *signal-to-noise ratio*

✘  $SMR$ : *signal-to-mask ratio* dal modello psicoacustico.

Lo standard fornisce tabelle che danno stime per SNR risultanti dalla quantizzazione ad un dato numero di livelli di quantizzazione;


2. Ottiene MNR per ciascuna sottobanda;
3. Cerca la sottobanda con il più basso MNR;
4. Alloca bit di codice a questa sottobanda. Se la sottobanda ottiene più bit allocati di quanto è appropriato, cerca una nuova stima per SNR e ripete dal passo 1.

# Commenti

- ⌘ Il motivo per la scelta di 16 bit di precisione è per ottenere un buon rapporto segnale-rumore.
- ⌘ Il rumore che si ottiene è rumore di quantizzazione dal processo di digitalizzazione.
- ⌘ Per ogni bit aggiunto si ottiene un SNR migliore di 6db.
- ⌘ Audio CD realizzano circa 90db di SNR.
- ⌘ L'effetto di mascheramento consente di innalzare il livello del rumore intorno ad un suono forte in quanto il rumore sarà comunque mascherato.
- ⌘ Innalzare il livello di rumore è lo stesso che usare meno bit e usare meno bit significa comprimere.
- ⌘ L'effetto di mascheramento si verifica anche prima e dopo un suono di forte intensità (*pre/post masking*).  
*Premasking*: 2-5 ms, *Postmasking*: 100ms.

# QUALITA' DEL SUONO

- ⌘ Nel processo di digitalizzazione, la determinazione della qualità' con cui rappresentare il suono, avviene sulla base del tentativo di soddisfare due opposte esigenze: (1) fedeltà' nella riproduzione del suono, (2) dimensione del file audio.
- ⌘ In generale, la qualità' della riproduzione del suono e' legata alla possibilità' di riprodurre i suoni nel modo più' simile a come noi tali suoni percepiamo.
- ⌘ Pertanto, un altro fattore che interviene a determinare la qualità' del suono e' il numero di canali audio che si considerano: Le leggi che caratterizzano la propagazione delle onde acustiche ed il fatto che noi esseri umani (ma non solo noi) siamo dotati di due orecchie ci consentono non solo di udire i suoni, ma anche di farci una idea della direzione da cui questi provengono (pur senza vedere l'oggetto che genera il suono).



⌘ La qualita' del suono non e' pertanto legata solo a quanto la forma d'onda riprodotta sia fedele all'originale, ma anche alla capacita' del sistema di riproduzione di creare l'effetto di piu' suoni generati in punti distinti nello spazio. Per avere questo effetto bisogna ricorrere all'acquisizione di piu' canali audio, ognuno dei quali acquisisce il suono che proviene da una ben precisa direzione.

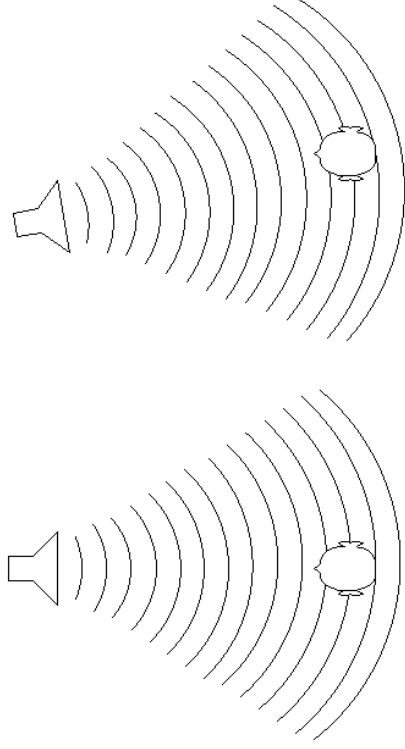




⌘ Il sistema uditivo fa uso di tre diversi meccanismi per ricostruire la provenienza nello spazio di un segnale sonoro.



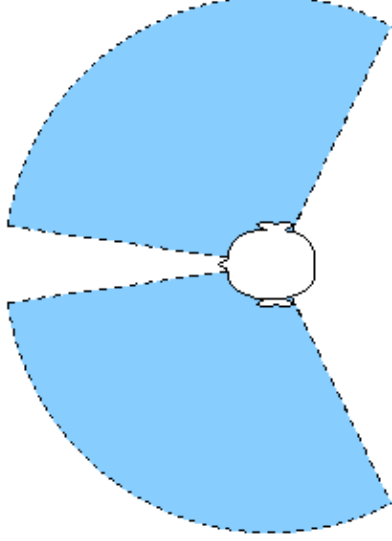
La differenza nel tempo di arrivo del fronte d'onda alle due orecchie. Un suono emesso da una sorgente posizionata sull'asse frontale dell'osservatore arriverà nello stesso istante alle due orecchie. Tuttavia, non appena la sorgente si discosta da questa posizione uno dei due orecchi percepirà l'arrivo del fronte d'onda prima dell'altro. A questa differenza temporale si dà il nome di Interaural Time Delay (ITD). Questo fenomeno dà risultati non ambigui solo se la lunghezza d'onda del suono è minore al doppio della distanza tra le orecchie.





⌘ Se la sorgente sonora e' posizionata da una parte rispetto all'asse frontale, l'orecchio posizionato dall'altra parte percepira' un suono piu' attenuato, per via dell'effetto di scudo fatto dalla testa. Questa differenza di livello sonoro e' nota come Interaural Level Difference (ILD).

⌘ A causa della forma della testa e del padiglione auricolare, la nostra sensibilita' alle diverse frequenze componenti lo spettro di un suono, varia con il variare della posizione della sorgente sonora. Questo effetto e' noto come Head Related Transfer Function (HRTF).





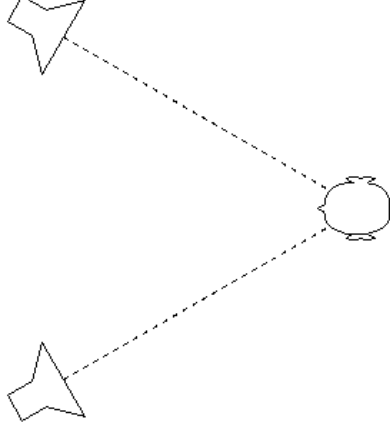
**Gli esseri umani hanno una naturale capacita' di muovere e ruotare la testa in modo da minimizzare ITD, ILD e la differenza di HRTF alle due orecchie.**



**Di questi tre fenomeni che stanno alla base della nostra capacita' di ricostruire la provenienza di un suono nello spazio, i tradizionali sistemi di riproduzione audio ne sfruttano soltanto uno: Interaural Level Difference. E' infatti sfruttando questo fenomeno che, con l'ausilio di due altoparlanti, si puo' generare una immagine sonora in cui la direzione percepita della sorgente sonora viene controllata con il livello del segnale che alimenta i due altoparlanti.**

## ⌘ L'efficacia di questo sistema e' vincolata da due fattori:

- ☐ l'ascoltatore deve essere posizionato in modo da formare un angolo di  $60^\circ$  con i due altoparlanti. In caso contrario si genererebbe un effetto di vuoto tra i due altoparlanti o un effetto di eccessiva concentrazione del suono.
- ☐ l'ascoltatore deve trovarsi sull'asse di separazione degli altoparlanti, altrimenti il suono sembrerà provenire dall'altoparlante più vicino.



# Formato mono e multicanale

- ⌘ Il formato audio multicanale ha origine con il cinema negli anni 1950. In questi anni, il suono stereofonico (che non necessariamente è sinonimo di due canali) viene propagandato, insieme al nuovo formato a schermo largo, quale strumento di rivalsa rispetto alla televisione, da cui il cinema di sentiva minacciato.
- ⌘ Tuttavia, a differenza di quello che in seguito sarà' il formato stereo televisivo caratterizzato da due soli canali, il formato stereo nel cinema prevede un minimo di quattro canali.

✂ **Traccia Audio Mono Ottica: la traccia audio è memorizzata contemporaneamente alla traccia video attraverso una traccia ottica in un'area adiacente a quella contenente i fotogrammi.**

✂ **Traccia Audio Stereo Magnetica: dopo la stampa della traccia video, la traccia audio è memorizzata attraverso l'applicazione di sottili tracce magnetiche, in tutto simili a quelle utilizzate nei registratori.**





**Traccia ottica: mentre il film e' proiettato un sottile fascio di luce attraversa la traccia ottica e finisce su una fotocella che converte la quantita' di luce incidente in un segnale elettrico. Tale segnale viene amplificato ed usato per pilotare gli altoparlanti.**

**⌘ Traccia magnetica: consente la registrazione e poi riproduzione in formati stereo multicanale. Tuttavia presenta una ridotta durata temporale del segnale magnetico, un maggior costo dei dispositivi per la riproduzione del segnale audio.**

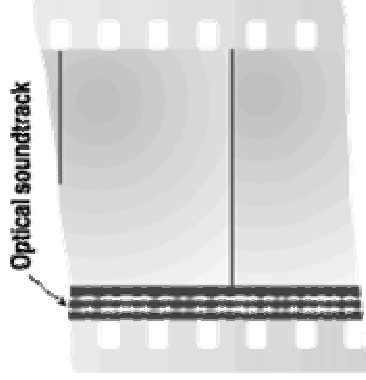
# Home Stereo


- ⌘ Nel 1958 comincia l'impiego della riproduzione stereofonica in campo discografico. Tuttavia, questa nasce come in grado di riprodurre due soli canali, non tanto a causa di precise preferenze di pubblico, quanto per la impossibilit  di registrare pi  di due tracce sonore sul supporto di memorizzazione in vinile.
- ⌘ Pochi anni dopo incomincia la diffusione radiofonica FM. Essendo anch'essa caratterizzata da stereo con due tracce audio contribuisce alla diffusione del connubio stereo = due tracce in ambiente di riproduzione home.



# Dolby surround

✂ Traccia Audio Stereo Ottica: Dolby Surround (1970): nello spazio riservato alla singola traccia ottica mono vengono introdotte due tracce ottiche in grado di codificare l'informazione su ben tre o quattro (Dolby Pro Logic Surround) canali audio: sinistro, destro, centrale e surround per effetti di ambiente.



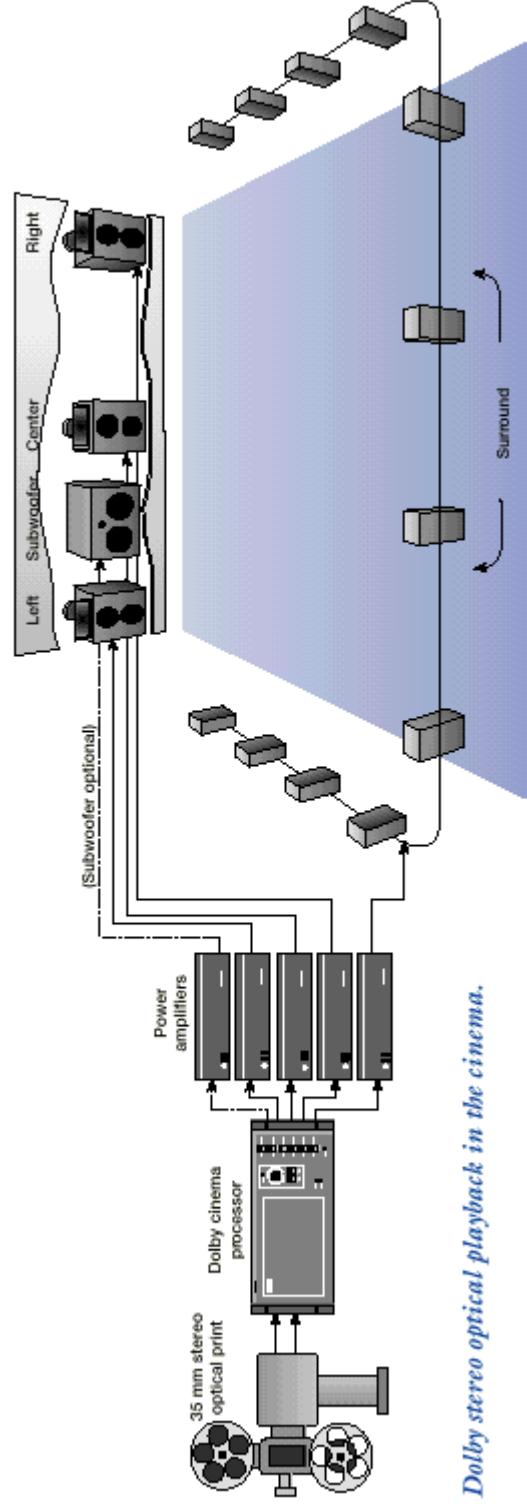


⌘ Il vantaggio di questa tecnologia e' rappresentato quindi dalla possibilita' di codificare e trasmettere quattro canali audio sulle stesse tracce necessarie a codificare e trasmettere due canali.

⌘ Un codificatore viene impiegato per fondere le quattro tracce in modo da codificarle su due sole tracce ed un decodificatore e' necessario per ripristinare, dati i due segnali in Dolby Surround, le quattro tracce corrispondenti ai canali sinistro, destro, centrale e surround.

⌘ La tecnologia Dolby Surround è anche impiegata per la riproduzione di home video, la trasmissione televisiva, i videogiochi ed applicazioni multimediali.

# Cinema: Dolby Surround

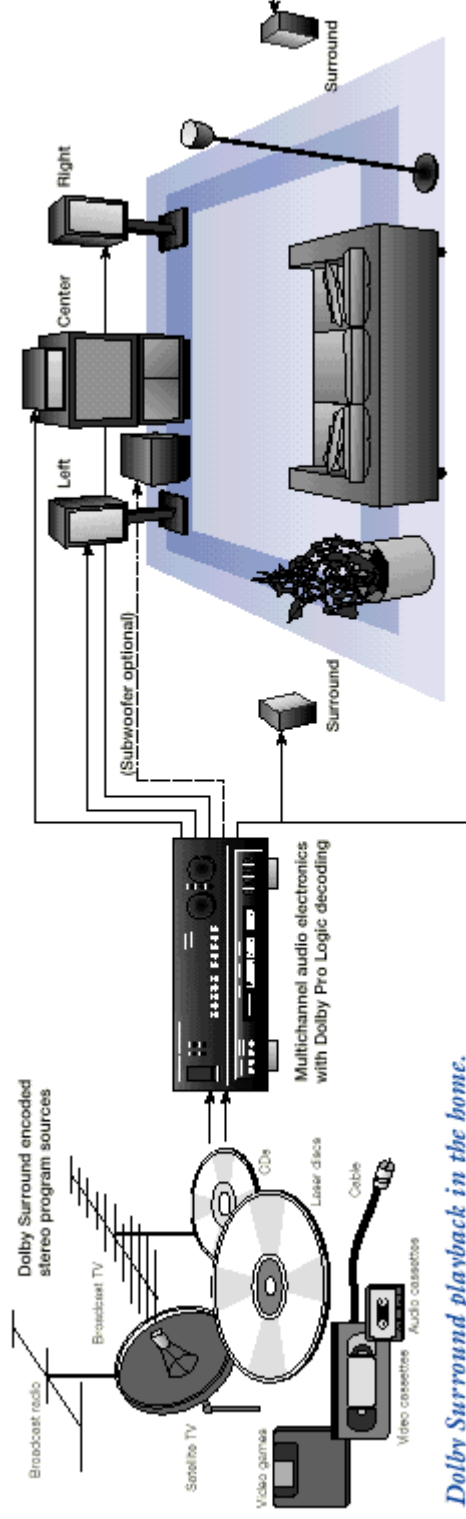


*Dolby stereo optical playback in the cinema.*

# Home Video: Dolby Surround

- ⌘ I primi videoregistratori furono introdotti nel 1970 con l'intento di fornire uno strumento per registrare i programmi televisivi e rivederli ad ore piu' convenienti. Successivamente, si prospettò la possibilità di impiegare questi dispositivi per la riproduzione home di film nati per il cinema.
- ⌘ I primi registratori supportavano solo una traccia audio. Tuttavia, sulla spinta della diffusione della musica HiFi in campo discografico, ben presto vennero introdotti i primi videoregistratori in grado di supportare la registrazione e riproduzione di due tracce audio: stereo. Parallelamente, anche alcune reti televisive (sia via cavo che via etere) cominciarono ad impiegare una trasmissione stereo a due canali del segnale audio.
- ⌘ In questo contesto, nel 1982 comincia a diffondersi l'impiego della tecnologia Dolby Surround per la riproduzione di home video. L'impiego di tale tecnologia continua tutt'oggi e si è anzi allargata al campo della trasmissione televisiva, videogiochi ed applicazioni multimediali.

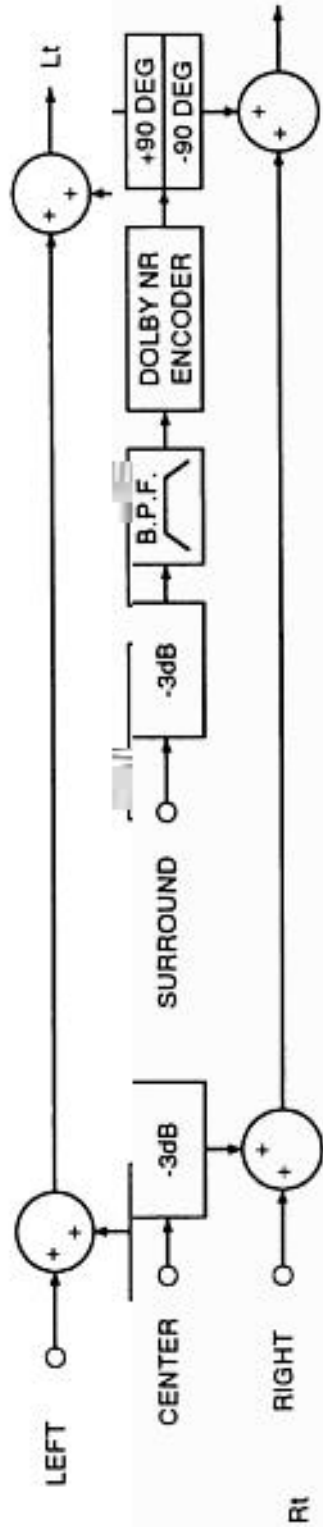
# Home Video: Dolby Surround



*Dolby Surround playback in the home.*

# Tecnologia Dolby MP (Motion Picture Matrix)

✂ Con la Tecnologia Dolby MP (Motion Picture Matrix) il codificatore accetta quattro distinti segnali di ingresso (L, R, C, S) e genera due segnali di uscita: left-total e right-total (Lt ed Rt).



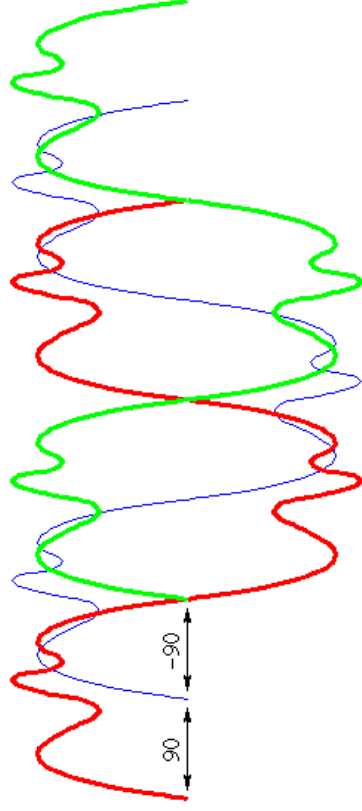
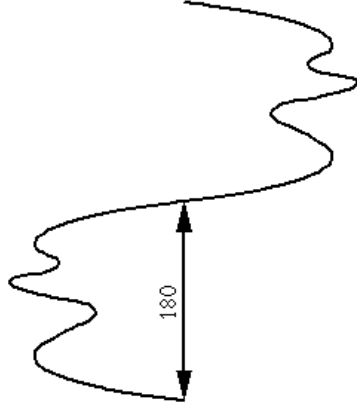


**I segnali L ed R passano inalterati allo stadio di uscita, il segnale C viene abbassato di livello e distribuito equamente tra Lt e Rt.**



**Il segnale S viene equamente distribuito su Lt ed Rt dopo essere stato sottoposto a:**

- ☐ 1) Riduzione del livello di suono;
- ☐ 2) Filtraggio passa basso 100~7000 Hz;
- ☐ 3) Filtraggio Dolby NR contro il rumore;
- ☐ 4) Shift in fase di +90 e -90 gradi, in modo da sommare ai due canali Lt ed Rt due segnali che sono in opposizione di fase.





**Il decodificatore accetta in ingresso i due segnali Lt ed Rt e ripristina in uscita i quattro segnali (L, R, C, S).**



**I segnali L ed R vengono direttamente ricavati da Lt ed Rt. Questo provoca in realta' la riproduzione in L ed R del segnale S. Tuttavia, comparendo il segnale S in opposizione di fase, questo genera un semplice effetto di diffusione del segnale S.**

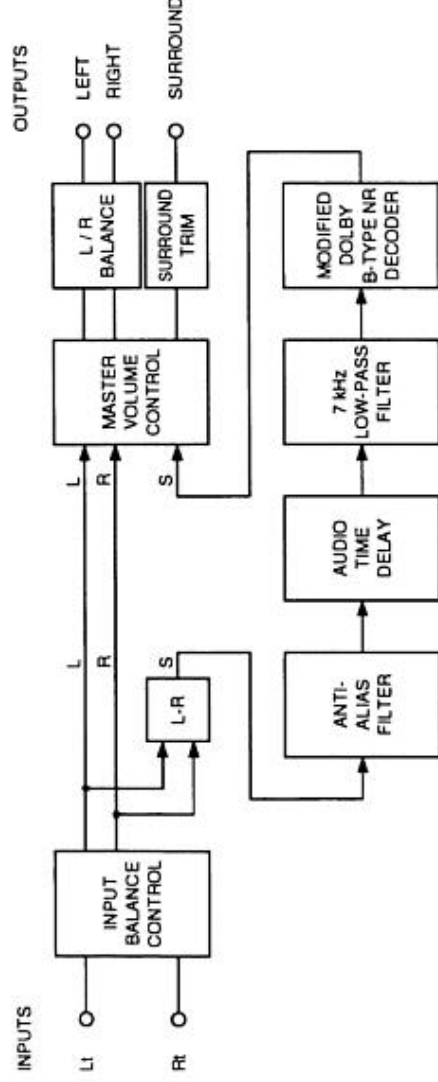


Fig. 2. Passive surround decoder block diagram.



# Tecnologia Dolby Pro Logic

- ⌘ Al fine di migliorare la separazione tra i vari canali e ridurre effetti di crosstalk, il decoder passivo viene sostituito con un decoder attivo, secondo la tecnologia Pro Logic.
- ⌘ Questa nuova tecnologia ha lo scopo di produrre una maggior localizzazione della provenienza del suono attraverso una piu' accurata decodifica tesa ad eliminare da ogni canale la presenza di segnali propri di altri canali. Per fare cio' e' necessario introdurre uno stadio di filtraggio attivo a valle dello stadio di decodifica passivo.

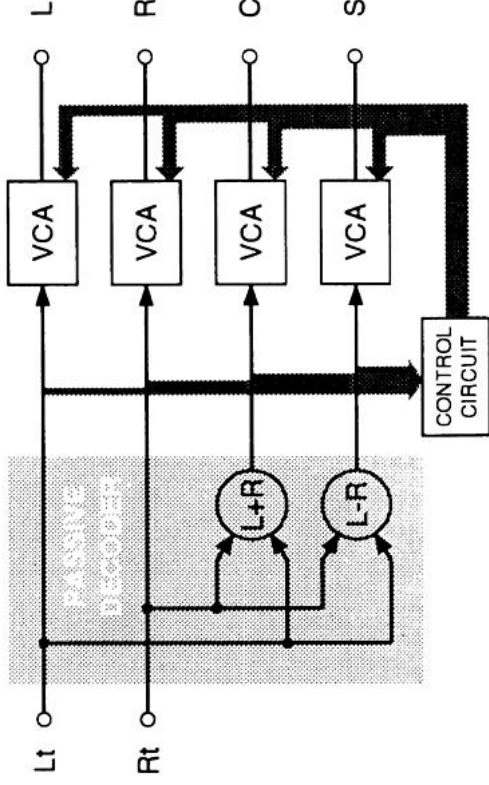


Fig. 6. Gain riding concept.

✂ Si consideri per esempio il caso in cui l'audio di un dialogo e' presente in entrambi i canali Lt ed Rt. In questo caso il decoder passivo consente a tale segnale audio di essere riprodotto sia dal canale C che da quelli L ed R. Invece, con l'impiego di un decoder attivo e' possibile usare un circuito di controllo che riduca opportunamente il livello dei segnali L ed R non appena questo si renda necessario.

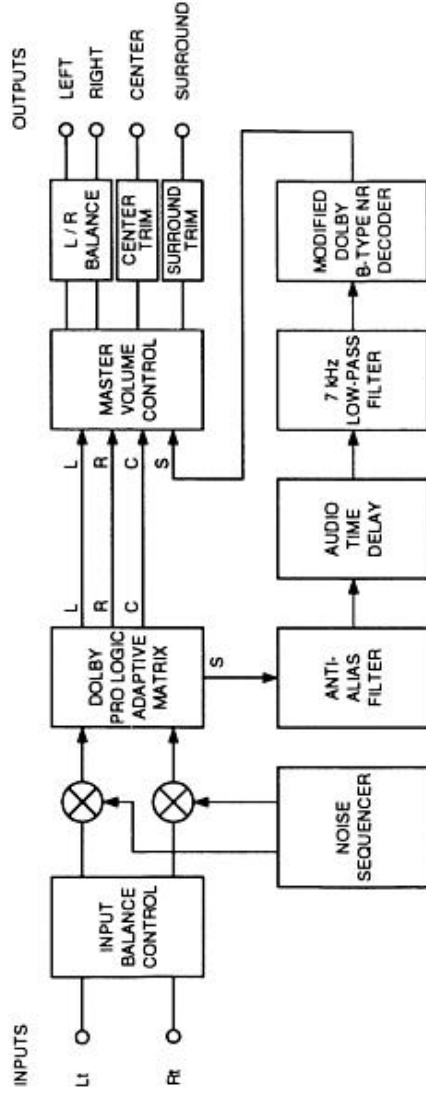


Fig. 9. Pro Logic decoder block diagram.

# Tecnologia Dolby Digital (1989)

⌘ Un limite della tecnologia Dolby Surround e' rappresentato dal fatto che in realta', in presenza di segnali in opposizione di fase sui due canali LeftT e RightT, i segnali Left, Right e Central vengono penalizzati a vantaggio del Surround. In presenza di segnali uguali sui due canali LeftT e RightT, i segnali Left, Right e Surround vengono penalizzati a vantaggio del Central. In sostanza non si ha una vera e propria quadrifonia in quanto i quattro canali non possono essere attivi contemporaneamente.

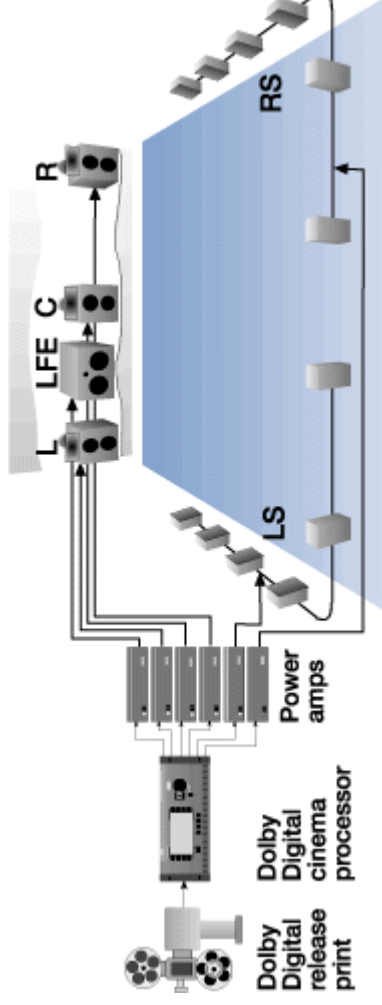
⌘ Verso la fine degli anni '80 l'industria cinematografica rivelo' un crescente interesse verso la possibilita' di sviluppare un formato di registrazione e riproduzione digitale dell'audio che superasse i limiti della tecnologia Dolby Surround. Al fine di preservare la tradizionale traccia audio ottica, si adottò come soluzione quella di registrare la traccia digitale (anche questa ottica) tra i fori presenti al bordo della pellicola.



Per quanto riguarda il numero di canali audio codificati, il nuovo formato Dolby Digital prevede un formato 5.1, riconosciuto da tutti i gruppi dell'industria cinematografica come quello in grado di soddisfare al meglio i requisiti di riproduzione nelle sale cinematografiche.



Il formato prevede l'impiego di sei canali separati: Left, Center e Right frontali, Left e Right Surround ed un canale subwoofer per la riproduzione di suoni in bassa frequenza (canale Low Frequency Effects).





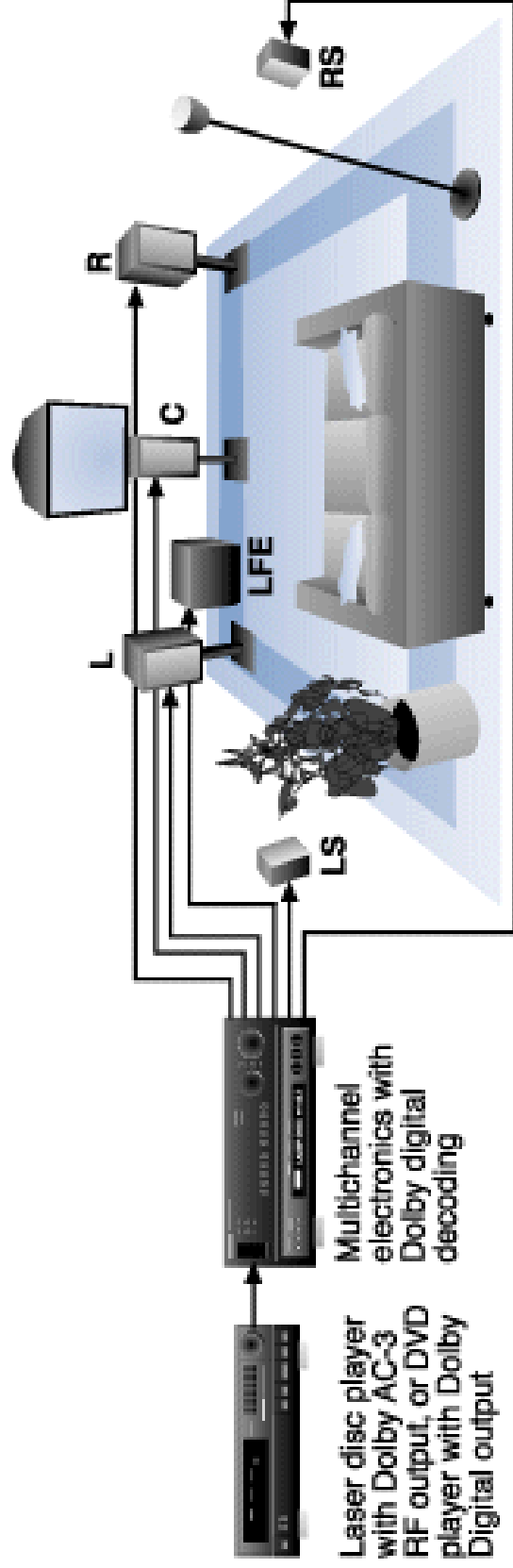
**Date le limitate dimensioni del supporto di memorizzazione, fu necessario introdurre una sorta di codifica in modo da permettere la memorizzazione di ben sei canali audio nello spazio compreso tra i fori a bordo pellicola. Fu così' sviluppato un formato di codifica che oggi consente di memorizzare e trasmettere sei tracce audio in meno dello spazio richiesto da un singolo canale nello standard Audio-CD.**



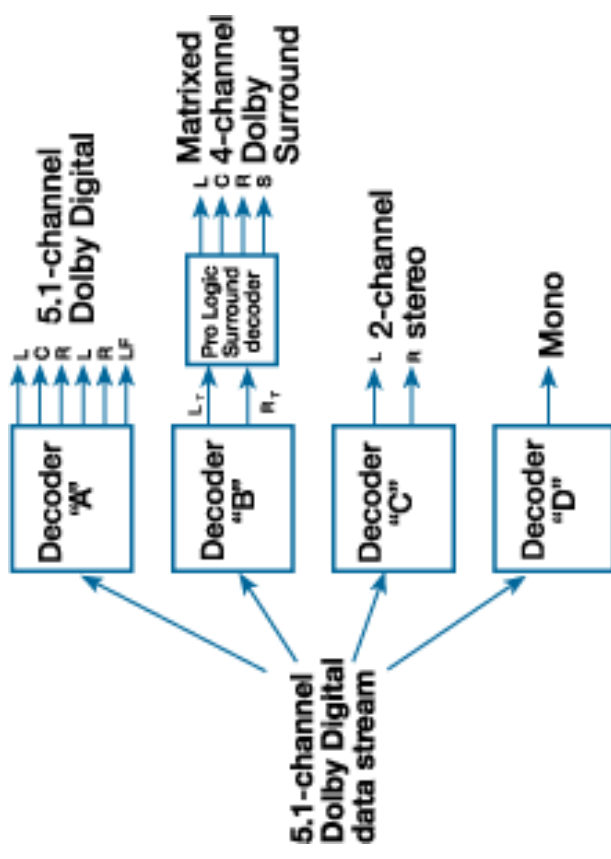
**L'impiego della tecnologia Dolby Digital in campo cinematografico si e' ben presto esteso al campo home video, cominciando nel '95 con la prima traccia audio su supporto Laser Disc e poi con il suo utilizzo nei sistemi DVD, TV digitale via cavo, TV digitale via etere nonche' applicazioni multimediali.**



**Rispetto al formato Dolby Surround analogico caratterizzato da un unico canale surround per audio in bassa frequenza (100-7000Hz), il nuovo formato digitale prevede l'impiego di due canali surround, ognuno con la stessa banda di riproduzione dei canali frontali, consentendo quindi un maggior effetto di profondità, localizzazione e realismo.**



⌘ Per poterne rendere effettivo l'impiego nelle piu' differenti configurazioni, il decoder Dolby Digital genera diversi tipi di uscite attraverso un processo di Downmixing, consentendo cosi' dalla traccia digitale di poter riprodurre l'audio compatibilmente con i formati 5.1 canali, 4 canali surround Pro Logic, 2 canali stereo, 1 canale mono.



- ⌘ Il formato digitale di rappresentazione audio impiegato nei Audio-CD prevede un formato PCM con una precisione di 16 bit/sample ed una frequenza di campionamento pari a 44.1 KHz. Questo porta ad un bit rate pari a  $1.411 \text{ Mb/s}$  ( $2 \text{ canali} * 16 * 44100$ ).
- ⌘ Tale data rate e' troppo elevato sia per trasmissione che memorizzazione (in un Audio-CD che ha capacita' di 650MB e' possibile al piu' memorizzare 3685 secondi di 2 tracce audio, equivalenti a 61 minuti).

- ⌘ I limiti imposti in termini di trasmissione e memorizzazione hanno portato allo sviluppo di una tecnica di codifica audio, nota come AC3, che consente con una precisione di 20 bit/sample e sample rate a scelta tra 32, 44.1, 48 KHz, di avere dei data rate pari a 32Kb/s (1 canale audio), 384 Kb/s (formato 5.1) sino a 640 Kb/s.





⌘ Il formato Dolby Digital rappresenta lo standard per la codifica delle tracce audio adottato per Laser Disc, standard TV digitale ATSC (Stati Uniti), TV via cavo digitale, TV satellitare digitale, DVD-Video in Stati Uniti, Europa e Giappone. I data rate massimi per questi formati sono di seguito mostrati:

Formato	Sample Rate	Data Rate
Laser Disc	48 KHz	384 Kb/s
Digital TV	48 KHz	384 Kb/s
Digital Cable	48 KHz	448 Kb/s
Digital Satellite	48 KHz	448 Kb/s
DVD-Video	48 KHz	448 Kb/s

# Certificazione THX Lucasfilm

- ⌘ Al vertice della tecnologia per la riproduzione audio sia cinematografica che home video e' la certificazione THX. Questa non rappresenta uno standard di codifica, quanto una certificazione, rilasciata su ambienti cinematografici e componentistica home video, per garantire la conformita' dell'effetto audio riprodotto con quello che il regista cinematografico intendeva generare.
- ⌘ Il sistema home video THX supporta sia la codifica analogica Dolby Surround Pro Logic che quella digitale Dolby AC3 e prevede l'impiego dei seguenti componenti: controller THX, altoparlanti THX, amplificatore THX. L'impiego di un sistema di riproduzione THX consente di avere i seguenti vantaggi: bilanciamento dei toni piu' naturale (sistemi non THX tendono a produrre suoni troppo brillanti), maggior chiarezza dei dialoghi che non devono essere coperti da suoni dovuti all'azione, effetto sonoro piu' uniforme, maggior localizzazione del suono che segue l'azione presente sullo schermo.

# Digital Theater System (DTS)

⌘ Alcuni film in DVD di ultima generazione utilizzano, oltre agli altri standard audio, anche un nuovo formato che prende il nome di Digital Theater System. Questo nuovo standard e' adatto agli audiofili piu' raffinati perche' utilizza per ogni singola traccia tutta la banda passante prevista, evitando qualsiasi penalizzazione di un canale rispetto agli altri a fronte di un data rate di 1.536 Mb/s (quasi 4 volte rispetto al Dolby Digital).

⌘ In altre parole, se il Dolby Digital restringe di volta in volta la banda dei vari canali per dare maggior risalto a quello con maggior informazione, il DTS offre una ampiezza di banda uguale per tutti i canali. Ovviamente lo spazio richiesto per la memorizzazione dei canali e' maggiore rispetto al Dolby Digital e ad oggi, sono pochi i film prodotti con questa codifica. Tale codifica inoltre richiede sistemi di decodifica molto piu' costosi rispetto a quelli per la decodifica Dolby Digital.