

Capitolo 2

Elementi di Acustica e Psicoacustica

Carlo Drioli Nicola Orio

Copyright ©1999 by Carlo Drioli and Nicola Orio. All rights reserved.

2.1 Acustica: introduzione

La percezione sonora è normalmente legata alle vibrazioni del timpano nell'orecchio. Queste vibrazioni sono provocate da piccole variazioni di pressione nell'aria. La variazione di pressione dell'aria è quindi l'equivalente fisico del suono. Questo fenomeno può essere visualizzato appoggiando un foglio di carta sopra il cono di un altoparlante: quando viene emesso un suono, il foglio inizia a vibrare. Infatti il movimento verso l'esterno della membrana dell'altoparlante determina un aumento di pressione e quindi spinge in fuori il foglio di carta. Inversamente il movimento verso l'interno della membrana determina una diminuzione di pressione ed attrae il foglio verso l'altoparlante. La membrana del timpano ha un comportamento analogo a quello del foglio di carta: un incremento di pressione spinge la membrana del timpano verso l'interno, mentre una diminuzione di pressione la attrae verso l'esterno. I movimenti del timpano sono quindi trasmessi alla coclea che li trasforma in impulsi elettrici che vengono inviati al cervello attraverso le terminazioni nervose.

2.2 Oscillazioni e onde

Dato che il suono corrisponde a variazioni di pressione nell'aria, è naturale che le proprietà di queste variazioni determinino le proprietà del suono percepito. Molti suoni musicali presentano variazioni regolari di pressione. In particolare la regolarità implica che un determinato andamento della pressione si ripeta nel tempo. Viene definita *forma d'onda* la ripetizione di tale andamento. In questo caso il suono è detto *periodico* e la durata della singola forma d'onda è detta *periodo*, indicato con il simbolo T e misurato in secondi. Se la funzione $p(t)$ indica l'andamento della pressione nel tempo in un punto dello spazio, per un suono periodico si ha la relazione:

$$p(t) = p(t + T)$$

Nel caso opposto, in cui l'andamento della pressione è privo di qualsiasi regolarità, il segnale associato viene percepito come *rumore*. Il rumore può essere diviso di due classi principali:

- *rumore impulsivo*: è determinato da rapide variazioni di pressione circoscritte nell'arco di pochi millisecondi. Un tipico esempio di rumore impulsivo si ha quando un corpo rigido viene percosso. Va notato che il rumore impulsivo viene regolarmente generato durante la produzione di suoni musicali, si pensi ad esempio al suono di chitarra nel quale è chiaramente percepibile il rumore prodotto dal plettro sulla corda; oppure al suono di pianoforte dove è fondamentale per il riconoscimento del timbro il rumore prodotto dal martelletto sulla corda.
- *rumore stazionario*: ha generalmente una elevata estensione temporale ma è comunque privo di regolarità. Tipici esempi di rumore stazionario sono il rumore prodotto dal vento o quello proveniente da uno schermo televisivo in assenza di segnale (*effetto neve*). Per questo genere di segnali audio si ricorre generalmente ad una descrizione statistica dell'andamento della pressione.

I suoni periodici¹ sono alla base della musica occidentale e di molti altri repertori, per cui a questi verrà posta particolare attenzione.

2.2.1 Suoni periodici

E' noto, dal teorema di scomposizione in serie di Fourier, che ogni funzione periodica può essere suddivisa nella somma di funzioni elementari. Per segnali reali, come nel caso dell'andamento della pressione in un mezzo trasmissivo, vale la relazione:

$$p(t) = a_0 + \sum_{n=1}^{\infty} a_n \cdot \sin\left(\frac{2\pi n t}{T} + \phi_n\right)$$

Dalla formula risulta quindi che un suono periodico, di periodo T , è scomponibile nella somma, pesata dai termini a_n , di sinusoidi di periodo $T, T/2, T/3, T/4, \dots$. Nella formula è inoltre presente il termine ϕ_n che tiene conto della fase iniziale di ognuna delle sinusoidi, che in generale può essere diversa per ogni funzione elementare. Considerando che il suono solitamente si propaga in aria, ove è già presente un termine costante dato dalla pressione atmosferica e tenendo conto che l'orecchio è sensibile alle variazioni di pressione, il termine a_0 viene usualmente trascurato. In Figura 2.1 vengono riportati gli andamenti delle forme d'onda rispettivamente associate a una sinusoide, ad un segnale periodico costituito da una somma di 16 sinusoidi in rapporto armonico tra loro e ad un segnale rumoroso stazionario.

In campo musicale si è soliti descrivere un suono periodico in termini di *frequenza*, usualmente indicata con il simbolo f e misurata in Hertz (Hz). Il legame tra periodo T e frequenza f è descritto dalla formula

$$f = \frac{1}{T}$$

La scomposizione di un suono periodico di frequenza f in forme d'onda elementari, indica che queste avranno rispettivamente frequenze $f, 2f, 3f, 4f, \dots$. La sinusoide di frequenza f , pari alla frequenza del suono periodico di partenza, è detta *fondamentale* mentre le sinusoidi di frequenza multipla intera di f vengono dette *parziali*. Si fa riferimento alle forme d'onda elementari che costituiscono un suono con il termine *armoniche*.

La frequenza è associata alla sensazione di altezza (*pitch*) di un suono: maggiore è la frequenza, maggiore risulta l'altezza del suono, in altre parole il suono risulta più acuto. Gli esseri umani sono

¹Spesso viene compiuta dagli autori la distinzione tra suono e rumore in base alla presenza o meno di periodicità. In questo contesto si è preferito usare la terminologia alternativa *suono periodico* e *suono aperiodico*.

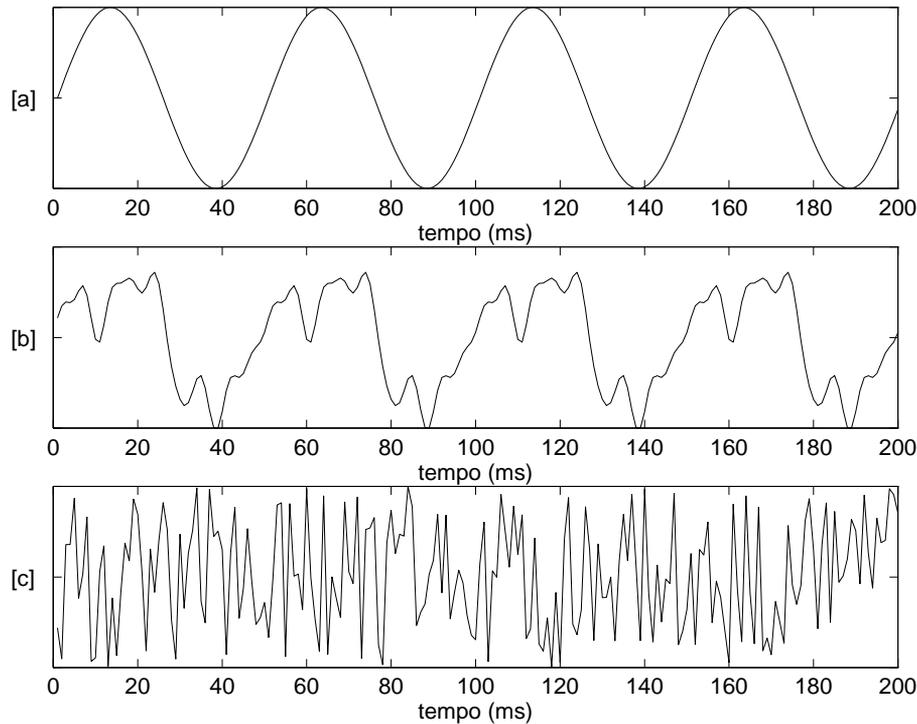


Figura 2.1: Andamento nel tempo di tre segnali rispettivamente con andamento: [a] sinusoidale, [b] periodico (somma di 15 sinusoidi), [c] aperiodico

in grado di percepire suoni nell'intervallo di frequenze da circa 20 Hz a circa 16 kHz, anche se alcuni soggetti sono in grado di percepire suoni in intervalli più ampi, ma comunque contenuti tra i 16 Hz e i 20 kHz. L'estensione di un pianoforte, così come l'estensione di un'orchestra sinfonica, va da 27.5 Hz a 3729.3 Hz. Al di sotto di 15 Hz, le variazioni di pressione non vengono più percepite come un singolo suono ma come una rapida successione di impulsi. Frequenze al di sopra della soglia di udibilità (*ultrasuoni*) non vengono percepite, quindi il filtraggio del segnale audio al di sopra dei 20 kHz non ne altera la qualità percepita. E' per questa ragione che, ricordando il teorema del campionamento di Shannon, la frequenza di campionamento dei Compact Disc (44.1 kHz) e' sufficiente per una perfetta ricostruzione del segnale analogico originario dal punto di vista percettivo.

2.2.2 I suoni reali

Come si è visto, la sinusoide è la più semplice forma d'onda perché non è ulteriormente scomponibile. I suoni prodotti dagli strumenti musicali acustici non hanno però mai un andamento così semplice. I suoni naturali infatti sono sempre costituiti da serie di armoniche, che contribuiscono a dare ricchezza ai suoni musicali. Si prendano come esempio i modi di vibrazione di una corda, fissata ai due estremi, a sezione costante e lunghezza L . Si può dimostrare, ricordando nozioni di fisica sul moto oscillatorio, che la corda vibra alle frequenze:

$$f_i = \frac{i\pi\sqrt{S/\rho_L}}{L}$$

dove S è la tensione della corda e ρ_L è la massa per unità di lunghezza. Quindi la forma d'onda di un suono prodotto da una corda è la risultante di un insieme di sinusoidi a frequenza multipla di una

frequenza fondamentale. Analogamente, considerando i modi di vibrazione di un tubo acustico, si può dimostrare che un tubo aperto da entrambe le estremità può vibrare alle frequenze:

$$f_i = \frac{i\pi c}{L}$$

dove c è la velocità del suono in aria e L è la lunghezza del tubo. Da entrambi gli esempi proposti risulta chiaro inoltre che una delle tecniche più efficaci per l'analisi dei suoni è l'analisi di Fourier, ovvero la scomposizione dei suoni prodotti dagli strumenti musicali acustici in elementi fondamentali (le sinusoidi).

In realtà, i suoni prodotti dagli strumenti musicali non hanno mai un comportamento così regolare. Innanzitutto le armoniche hanno un rapporto che solo approssimativamente può essere espresso come rapporto tra interi. Ad esempio la corda reale si differenzia dalla corda ideale principalmente a causa di questa inarmonicità: le parziali risultano avere dei rapporti leggermente maggiori dei numeri interi previsti dalla teoria (si dice in questo caso, usando il lessico musicale, che le parziali sono *crescenti* rispetto alla fondamentale). Alcuni suoni reali inoltre sono caratterizzati proprio a causa dell'assenza di armonicità; è il caso delle campane, nelle quali non è nemmeno presente il termine relativo alla fondamentale e le armoniche hanno rapporti solo approssimativamente armonici. Una seconda caratteristica dei suoni reali è che questi non sono mai esattamente periodici: le forme d'onda si ripetono nel tempo assumendo degli andamenti simili, ma non del tutto uguali. L'orecchio percepisce quindi un andamento approssimativamente periodico, ma percepisce anche le variazioni nella forma d'onda, che contribuiscono a dare dinamicità al suono prodotto. Infatti una delle caratteristiche dei suoni di sintesi è appunto l'eccessiva regolarità del loro sviluppo temporale. Questo spesso si traduce nella percezione di un suono che rapidamente diventa poco interessante per l'ascoltatore.

2.3 Inviluppo dei suoni

Considerando la musica come una forma di comunicazione ottenuta attraverso l'organizzazione dei suoni, risulta evidente che un suono perfettamente periodico in senso matematico (ovvero un suono che si ripete indefinitamente) non consente alcuna forma di comunicazione. I suoni musicali vengono quindi organizzati nella scala dei tempi; in particolare ogni suono ha un inizio e una fine che vengono percepiti dall'ascoltatore. Grossa importanza nella comunicazione musicale ha l'evoluzione nel tempo della forma d'onda. Come si è visto un suono è caratterizzato principalmente da una frequenza, legata alla percezione del pitch, e da una ampiezza delle oscillazioni della pressione, legata alla percezione di intensità. Una forma d'onda elementare che evolve nel tempo può quindi essere espressa dalla formula:

$$s(t) = A(t) \cdot \sin(2\pi ft)$$

dove f è la frequenza del suono e $A(t)$ è l'*inviluppo di ampiezza* del segnale. Per meglio chiarire il concetto di inviluppo di ampiezza, prendiamo come esempio la generazione di un suono da una corda di violino eccitata con l'archetto. In condizioni di riposo la corda ha ovviamente vibrazione nulla, e quindi non produce alcun suono. Quando il violinista inizia a sfregare l'archetto sulla corda, questa inizia a vibrare abbandonando la situazione di riposo. Esiste un periodo di tempo nel quale le oscillazioni della corda, da nulle, si fanno sempre più ampie. Questa viene definita fase di attacco e solitamente indicata con il corrispondente termine inglese *attack*. Questa fase dura solitamente pochi centesimi di secondo, in relazione al tipo di strumento musicale. La fase successiva a quella di *attack* è definita con il termine inglese *decay*: corrisponde ad un rapido assestarsi della ampiezza ad un valore stabile dopo una sovralongazione a cui è stata portata dalla fase di *attack*. Anche il *decay* è

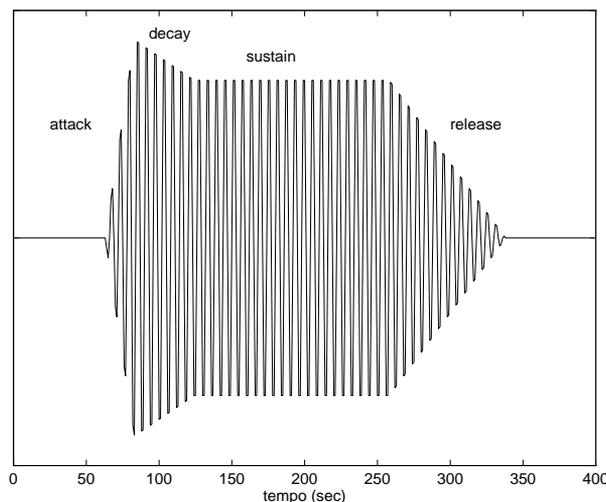


Figura 2.2: Evoluzione del segnale musicale nel tempo: sono evidenziate le parti di *attack*, *decay*, *sustain* e *release*

molto rapido. A questo punto, esaurito il transitorio di attacco, si è realizzato un accoppiamento tra lo sfregamento dell'archetto e le oscillazioni della corda. Questo corrisponde alla fase di *sustain*, che può durare anche parecchi secondi, nella quale il suono viene appunto sostenuto dal musicista, che continua a fornire l'energia necessaria per mantenere le vibrazioni. L'ultima fase, che ha inizio nel momento in cui il musicista smette di mantenere eccitato il sistema di vibrazione, viene denominata *release* (ovvero rilascio) e corrisponde al tempo in cui il corpo vibrante (nel nostro esempio la corda di violino) smorza l'entità delle vibrazioni, fino a portarsi nuovamente nello stato di quiete. In Figura 2.2 sono illustrate le quattro diverse fasi descritte. In questo caso la funzione involuppo $A(t)$ è stata approssimata con la successione di quattro segmenti, ma in generale può assumere degli andamenti molto più complessi, solitamente seguendo una curva esponenziale. Va peraltro sottolineato che spesso, in sede di sintesi, si preferisce approssimare l'involuppo con delle spezzate, poiché si è visto che la qualità sonora dei risultati non viene compromessa da questa approssimazione.

2.4 Propagazione del suono

Come detto, il suono ha natura oscillatoria: lo studio dell'acustica musicale fa quindi riferimento alla teoria delle onde. Il periodo e la frequenza sono quindi le caratteristiche principali a cui si farà riferimento per l'analisi del comportamento di un'onda acustica. Torniamo a fare riferimento al movimento della membrana di un altoparlante, tenendo conto che considerazioni analoghe possono essere compiute, ad esempio, sul movimento oscillatorio di una colonna d'aria all'interno di un tubo acustico o sulla membrana di uno strumento a percussione. Il movimento della membrana dell'altoparlante causa compressione e rarefazione dell'aria, che appunto corrispondono a variazioni di pressione acustica. Quando la membrana si sta muovendo verso l'esterno, le molecole presenti nell'aria vengono compresse, determinando quindi un incremento della pressione locale. Questa incremento di pressione si propaga agli strati d'aria adiacenti. Inversamente quando la membrana si muove verso l'interno, si crea una diminuzione di pressione che si propaga agli strati adiacenti. Ne risulta che le particelle d'aria sono spinte leggermente in avanti e indietro nella direzione di propagazione del suono. Oscillazioni di questo tipo vengono definite *longitudinali*, e sono tipiche della trasmissione del suono nell'aria. Le

oscillazioni possono anche essere perpendicolari alla direzione di propagazione del suono. Ad esempio in una corda percossa la deformazione, che è perpendicolare alla corda, si propaga lungo la corda stessa.

Se andiamo a misurare la pressione dell'aria lungo la direzione di propagazione di un suono periodico, notiamo quindi una successione di aumenti e diminuzioni di pressione. Questa distribuzione viene definita onda sonora. La distanza più piccola tra due punti corrispondenti dell'onda sonora (ad esempio tra due massimi consecutivi) è detta *lunghezza d'onda*. Essa è comunemente indicata con il simbolo λ . La lunghezza d'onda dipende dal periodo e dalla velocità di propagazione del suono. La velocità del suono in aria viene solitamente indicata con la lettera c e, a temperatura ambiente (20°C), è di circa 344 m/sec, ovvero 1238 km/h. La relazione tra periodo T , velocità del suono c e lunghezza d'onda λ è

$$\lambda = c \cdot T$$

Utilizzando la definizione di frequenza data in precedenza, si ottiene la seguente relazione:

$$f = \frac{c}{\lambda}$$

che pone in evidenza come la frequenza sia inversamente proporzionale alla lunghezza d'onda. Si riconosce questa relazione in molti strumenti musicali. Ad esempio le corde corrispondenti alle note gravi del piano sono lunghe, mentre quelle corrispondenti alle note acute sono corte. Negli organi le note basse sono generate dalle canne più lunghe. Ricordando che le frequenze udibili sono contenute nell'intervallo da 20 Hz a 16 kHz, sostituendo i valori numerici nella formula che lega frequenza e lunghezza d'onda, otteniamo che le lunghezze d'onda dei suoni udibili vanno da circa 17 m (suono grave) a 21 cm (suono acuto). La velocità del suono nell'aria dipende dalla temperatura, aumenta di circa 0.6 m/sec per grado centigrado, mentre è indipendente dalla pressione atmosferica e dalla frequenza del suono.

La velocità dipende inoltre anche dal mezzo in cui si propaga. In Tabella 2.1 vengono riportate le velocità di propagazione, a 0°C , per alcuni mezzi trasmissivi. Queste differenze implicano che la lunghezza d'onda di un suono ad una data frequenza vari a seconda del mezzo in cui si propaga. Ad esempio, poiché la velocità di propagazione in acqua è circa 4.35 volte maggiore di quella in aria, le due lunghezze d'onda in acqua e in aria manterranno la stessa proporzionalità.

Mezzo trasmissivo	Velocità (m/sec)
Gomma	70
Ossigeno	317
Aria	331
Azoto	337
Idrogeno	1270
Acqua marina	1440
Acciaio	5050

Tabella 2.1: Velocità di propagazione del suono a 0°C in alcuni mezzi trasmissivi

2.4.1 Onde sferiche e onde piane

La sorgente sonora più semplice da analizzare è la sfera pulsante. Si tratta evidentemente di una situazione ideale in cui una sfera si contrae e si espande radialmente attorno ad una posizione media.

La variazione di pressione causata dalle pulsazioni della sfera si espande con la stessa efficienza in tutte le direzioni, dando luogo ad una *onda sferica*. Un'altra semplice sorgente sonora può essere considerato un pistone che si muove all'interno di un tubo. Se vengono trascurati gli effetti ai bordi del tubo, il movimento del pistone causerà una variazione di pressione solamente lungo la direzione del movimento. Avremo in questo caso una *onda piana*, che si propaga in un'unica direzione. A distanze sufficientemente elevate, il raggio di curvatura di un'onda sferica può essere considerato trascurabile, e anche in questo caso l'onda si considera piana.

In situazioni non ideali, la propagazione del suono non ha un andamento così semplice. In particolare, la propagazione in un mezzo non omogeneo dà adito ad alcuni fenomeni analoghi a quelli riscontrati nella propagazione della luce. Tra questi i principali sono la diffrazione e la riflessione.

2.4.2 Diffrazione

Nel caso di una sorgente reale, come ad esempio il cono di un altoparlante o la campana di una tromba, l'efficienza di irradiazione dipende dalla lunghezza d'onda. Questo effetto viene definito *diffrazione*. Se la dimensione della sorgente (ad esempio il raggio dell'altoparlante) è piccola rispetto alla lunghezza d'onda, la sorgente può essere considerata puntiforme e irradiante in tutte le direzioni con la stessa efficienza, generando quindi onde sferiche. Nel caso la lunghezza d'onda sia confrontabile con le dimensioni della sorgente, il suono viene irradiato con efficienza diversa a seconda della direzione. In particolare se la lunghezza d'onda è minore della dimensione della sorgente vi è un angolo al di sopra del quale non vi è praticamente irradiazione. In Figura 2.3 viene illustrata la diversa diffrazione nel caso la dimensione dell'apertura D sia, rispettivamente, minore o maggiore della lunghezza d'onda λ .

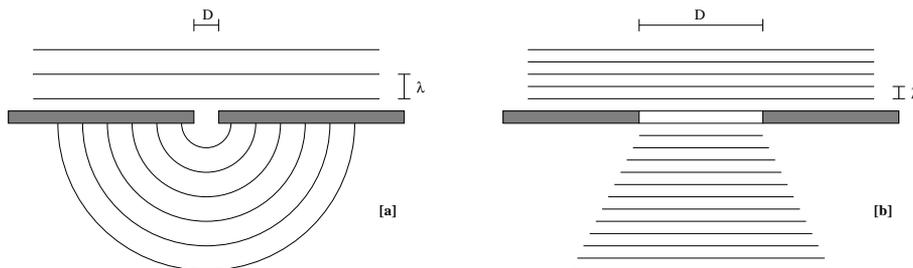


Figura 2.3: Effetto della diffrazione nei casi $\lambda > D$ in [a] e $\lambda < D$ in [b]

Un primo effetto della diffrazione è la direzionalità dei suoni acuti rispetto ai suoni bassi: è per questo motivo che è molto più semplice identificare la posizione di una sorgente se questa emette frequenze acute. Di questo fenomeno tengono conto gli apparecchi HiFi, nei quali le basse frequenze non necessitano di diffusione stereofonica. Inoltre è a causa della diffrazione che le frequenze basse possono essere più facilmente percepite anche in presenza di ostacoli che non consentono la propagazione diretta del suono, come ad esempio nel caso di stanze comunicanti. L'effetto della diffrazione è inoltre responsabile del tipo di propagazione della voce: essendo l'apertura della bocca sufficientemente piccola rispetto alle frequenze di emissione sonora, le onde sonore hanno propagazione sferica.

Inoltre la direzionalità della voce è rinforzata dall'effetto di un secondo fenomeno, legato al comportamento delle onde in presenza di ostacoli (nel caso della voce l'ostacolo è la testa dello stesso parlante). Infatti quando un'onda incontra un ostacolo di dimensioni piccole rispetto alla lunghezza d'onda (suono grave, ostacolo piccolo), il suono viene diffratto e riesce a superare l'ostacolo. Nel caso contrario (suono acuto, ostacolo grande) il suono non riesce a superare l'ostacolo e si crea una

zona d'ombra. Questo avviene perché le onde sonore sono maggiormente riflesse che diffratte. Tornando al caso della voce, questo fenomeno spiega perché è difficile capire il parlato ponendosi dietro ad una persona, nonostante si percepisca comunque il suono: sono le basse frequenze quelle che maggiormente riscono ad aggirare l'ostacolo, ma queste non sono sufficienti per rendere intelleggibile il parlato (in particolare, come si vedrà, non vengono riconosciuti i formanti).

2.4.3 Riflessione

In generale avviene una riflessione ogni volta che cambiano le caratteristiche del mezzo trasmissivo. La causa più comune della riflessione è la presenza di una discontinuità, ad esempio quando un'onda che si propaga in aria incontra un ostacolo. Ponendosi nel caso più semplice, si può supporre l'ostacolo come una parete liscia di dimensioni sufficientemente grandi rispetto alla lunghezza d'onda. In questo caso si ha che l'onda viene parzialmente riflessa e parzialmente assorbita, e che l'angolo di riflessione è uguale all'angolo di incidenza. La percentuale di onda assorbita dipende dal tipo di materiale.

Nel caso l'onda si rifletta su di una superficie irregolare, ove le irregolarità abbiano dimensioni paragonabili alla lunghezza d'onda, si ha un tipo di riflessione detta *eco diffuso*, dove la direzione di propagazione varia a seconda della lunghezza d'onda e della forma dell'ostacolo. L'effetto della riflessione è estremamente importante nella progettazione di sale da concerto e di teatri, e l'*acustica architettonica* è divenuto un ramo molto importante dell'acustica. A causa della riflessione alle pareti, all'ascoltatore infatti non giunge solamente il suono proveniente dagli strumenti musicali (o dalla voce degli attori), ma anche una successione di onde riflesse che, a causa della maggiore distanza percorsa, giungono all'ascoltatore con un dato ritardo. Questo fenomeno è noto con il nome di *riverberazione*, il cui controllo è uno dei principali obiettivi dell'acustica architettonica.

Sono state progettate inoltre particolari camere nelle quali la riflessione delle pareti è resa massima (camere ecoiche) o minima (camere anecoiche). In una camera ecoica, l'elevata riflessione delle pareti, fa sì che il suono in un punto giunga con la stessa intensità da tutte le direzioni: le camere ecoiche sono quindi utilizzate per effettuare misure della potenza acustica di un sistema. Al contrario, in una camera anecoica, la riflessione alle pareti è pressoché nulla e quindi il suono che giunge in un determinato punto proviene esclusivamente dalla sorgente: le camere anecoiche sono appunto utilizzate per studiare le caratteristiche delle sorgenti sonore.

2.5 Intensità del suono

Si è detto che l'equivalente fisico del suono è la variazione di pressione nell'aria (la pressione si misura in pascal, simbolo Pa). L'entità delle variazioni di pressione è legata alla percezione di volume sonoro (*loudness*): maggiore è la variazione di pressione, maggiore è il volume sonoro percepito. Spesso, più che non ai picchi nella variazione di pressione, si fa riferimento alla *pressione efficace*, simbolo p_{eff} , che è la media quadratica delle variazioni di pressione. Nel caso di un andamento sinusoidale della pressione del tipo

$$p = P_0 \sin\left(\frac{2\pi t}{T}\right)$$

si ha $p_{eff} = P_0/\sqrt{2}$. La minima pressione efficace che può essere percepita è di 0.00002 Pa, mentre la soglia del dolore varia intorno ai 20 Pa, in relazione alla frequenza del suono come vedremo in seguito parlando di psicoacustica.

Si consideri di dover determinare il volume sonoro prodotto da una sorgente. L'esperienza comune ci dice che la pressione efficace di un suono varia in relazione alla distanza della sorgente; inoltre,

come si è visto parlando della diffrazione, una sorgente sonora può irradiare in maniera diversa in differenti direzioni. Infine il fenomeno della riflessione può ulteriormente complicare la misurazione, rendendola sensibile, non solo alla distanza e alla posizione rispetto alla sorgente, ma anche alla presenza di ostacoli o elementi riflettenti. E' per questa ragione che una sorgente sonora viene caratterizzata in base alla propria *potenza acustica*, ovvero in base al lavoro prodotto nell'unità di tempo. Come ogni potenza, anche la potenza acustica si misura in watt (W). In Tabella 2.2 viene riportata la potenza acustica del parlato e di alcuni strumenti musicali. Gli strumenti musicali sono comunque caratterizzati da una bassissima efficienza, ovvero il rapporto tra i watt acustici e i watt spesi si aggira intorno all'1%.

Sorgente sonora	Potenza (W)
Parlato (normale)	10^{-5}
Parlato (litigio)	10^{-3}
Cantante lirico	0.03
Clarinetto	0.05
Tromba	0.3
Pianoforte	0.4
Trombone	6
Orchestra	60

Tabella 2.2: Potenza massima prodotta da alcune sorgenti sonore

Si definisce *intensità acustica* (simbolo I) la potenza media trasmessa per unità di superficie nella direzione di propagazione dell'onda. Si può dimostrare che, per onde piane e onde sferiche, vale la relazione:

$$I = \frac{p_{eff}^2}{\rho c}$$

dove ρ è la densità del mezzo trasmissivo (in aria, a temperatura ambiente e a pressione atmosferica standard $\rho = 1.21 \text{ kg/m}^3$) e, al solito, p_{eff} è la pressione efficace e c è la velocità del suono nel mezzo. Considerando l'intervallo di valori assunti dalla pressione efficace, si nota che l'intensità acustica assume valori in un range molto elevato, andando da circa 10^{-12} W/m^2 per la soglia di udibilità a circa 1 W/m^2 per la soglia del dolore.

2.5.1 Decibel e misure del suono

I valori di pressione, potenza e intensità acustica dei suoni si distribuiscono in un intervallo di valori molto esteso. Per questa ragione queste grandezze sono comunemente espresse in scala logaritmica. Va inoltre osservato che la scala logaritmica ha un andamento più vicino a quello delle scale percettive che verranno illustrate nel capitolo 2.11. Viene definito come *livello di pressione acustica* (in inglese *pressure level*, con simbolo PL) il logaritmo del rapporto tra la pressione misurata e una pressione di riferimento. In formule:

$$PL = 20 \cdot \log \frac{P}{P_{ref}}$$

dove si fa implicitamente riferimento alla pressione efficace. Il valore di PL è adimensionale e viene espresso in *decibel* (dB). In generale non è necessario utilizzare una pressione di riferimento standard. Può risultare comunque conveniente utilizzare come riferimento la minima pressione efficace udibile;

in questo caso si parla di *Sound Pressure Level* (SPL) che viene quindi definito come:

$$SPL = 20 \cdot \log \frac{p}{0.00002} \quad \Rightarrow \quad SPL = 20 \cdot \log p + 94$$

Analogamente, anche la potenza e l'intensità acustica vengono espresse in decibel utilizzando un valore di riferimento. Il *livello di intensità acustica* (in inglese intensity level, simbolo IL) è definito dalla formula:

$$IL = 10 \cdot \log \frac{I}{I_{ref}}$$

Anche in questo caso non è necessario scegliere un riferimento standard. Dalla formula si ricava agevolmente il raddoppiamento dell'intensità corrisponde ad un aumento di $10 \cdot \log 2 = 3$ dB.

La scelta di moltiplicare il logaritmo per un coefficiente 10 è dovuta alla semplicità di notazione che ne consegue: utilizzando come riferimento la minima intensità udibile, la scala in decibel assume valori da 0 (soglia di udibilità) a 120 (soglia del dolore) e risulta quindi più pratica della scala in Bel. Si pone in evidenza che il fattore moltiplicativo è diverso per la misura di pressione e di intensità e rispecchia la relazione di proporzionalità tra l'intensità e il quadrato della pressione precedentemente espressa (si ricorda che l'elevamento al quadrato nei logaritmi corrisponde alla moltiplicazione per 2). In Tabella 2.3 vengono riportati i valori in dB prodotti da un'orchestra a seconda delle indicazioni di volume nella partitura e da diverse sorgenti sonore.

Indicazione	Sorgente sonora	Intensità (dB)
	Silenzio	0
	Spillo che cade	10
	Sussurro a 1m	20
	Sala vuota	30
<i>ppp</i>	Libreria	40
<i>pp</i>	Interno auto silenziosa	50
<i>p</i>	Conversazione pacata	60
<i>mp</i>	Traffico	70
<i>mf</i>	Fabbrica	80
<i>f</i>	Metropolitana	90
<i>ff</i>	Discoteca	100
<i>fff</i>	Concerto rock	110
	Jet in partenza a 500m	120

Tabella 2.3: Livello di intensità associato alle indicazioni di partitura (prima colonna) e prodotto da alcune sorgenti sonore (seconda colonna)

Anche se la soglia del dolore è intorno a 120 dB, una prolungata esposizione a sorgenti sonore di elevata intensità può causare danni permanenti all'orecchio. In particolare è considerata a rischio l'esposizione a 100 dB, mentre le leggi sulla sicurezza obbligano l'uso di apposite cuffie negli ambienti di lavoro nel caso di prolungata esposizione ad un livello di intensità superiore a 85 dB.

I valori riportati in Tabella 2.3 vanno presi come puramente indicativi. In particolare le indicazioni di partitura hanno solo una corrispondenza approssimativa con i valori in dB, poiché dipendono dalla rumorosità della sala, dalla dinamica degli strumenti e dalle diverse scelte esecutive. Inoltre si è già accennato al fatto che il valore dell'intensità e della pressione variano con la distanza. Si prenda ad esempio una sorgente approssimativamente puntiforme, a cui corrisponde la propagazione di un

onda sferica. Dalla definizione di intensità sappiamo che, a parità di potenza della sorgente, questa risulta proporzionale all'inverso della superficie attraversata dall'onda. La dimensione della superficie sferica S dipende dal raggio r e quindi dalla distanza dalla sorgente, secondo la relazione $S = 4\pi r^2$. Nel caso si raddoppi la distanza la superficie risulta quadruplicata. Applicando la formula per il calcolo del livello di intensità si ottiene una variazione di -6 dB. Nel caso la sorgente sia, invece, una colonna di traffico la propagazione non è sferica, ma cilindrica: il raddoppiamento della distanza porta al raddoppiamento della superficie, a cui corrisponde una attenuazione di soli -3 dB. E' per questo motivo che il rumore causato dal traffico viene percepito a distanza maggiori che non il rumore di sorgenti singole (ad esempio macchinari industriali).

Il livello di intensità acustica dipende evidentemente anche dal numero di sorgenti presenti. E' possibile dimostrare che, nel caso di sorgenti tra loro scorrelate, l'incremento massimo del livello di intensità dato dalla somma di due sorgenti è di 3 dB e che questo si verifica quando le due sorgenti hanno pari intensità. In Figura 2.4 viene illustrato l'andamento dell'incremento del livello di intensità sonora nel caso di due sorgenti scorrelate, in funzione dell'intensità di una delle due sorgenti (l'altra è fissa a 60 dB). Risulta evidente che nel caso di sorgenti con intensità molto diverse, l'effetto della sorgente con maggiore intensità risulta preponderante; ad esempio nel caso di due sorgenti a 60 dB e 70 dB la risultante è a 70.4 dB.

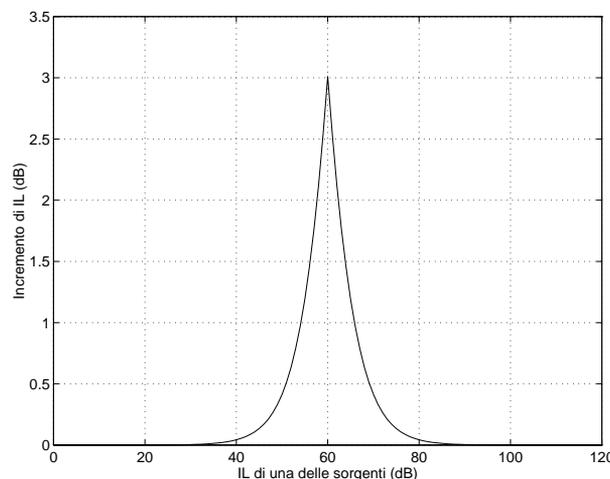


Figura 2.4: Incremento del livello di intensità nel caso di due sorgenti, la prima fissa a 60 dB e la seconda variabile da 0 dB a 120 dB

2.6 Psicoacustica: introduzione

Con il termine *musica* si fa generalmente riferimento a un complesso di processi che vanno dalla generazione di pattern sonori fino alla percezione e alla elaborazione del messaggio musicale da parte di un ascoltatore. Il processo di produzione e percezione del suono si può schematizzare come una catena di tre sistemi connessi detti *sorgente*, *mezzo* e *ricevitore*, le cui funzioni sono riportate in Figura 2.5.

Table 1.1

	System	Function
Source	Excitation mechanism	Energy supply
	Vibrating element	Determination of fundamental tone characteristics
	Resonator	Conversion into air pressure oscillations (sound waves), final determination of tone characteristics
Medium	Medium proper	Sound propagation
	Boundaries	Reflection, absorption, reverberation
Receptor	Eardrum	Conversion into mechanical oscillations
	Inner ear	Primary frequency sorting, conversion into nerve impulses
	Nervous system	Processing, display, identification, storage and transfer to other brain centers

Figura 2.5: Le funzioni dei sistemi sorgente, mezzo e ricevitore

In queste note si analizza il sistema ricevitore, cioè come i suoni vengono elaborati ed interpretati dal sistema uditivo e dal cervello.

2.7 Caratteristiche fisiche del suono e sensazioni uditive

Sentiamo un suono quando il timpano dell'orecchio viene eccitato da un'onda di pressione avente caratteristiche fisiche ben definite (come intensità o periodicità). La conseguenza per noi è una percezione del fenomeno costituita da diverse sensazioni che ci permettono di distinguere quel suono rispetto ad altri. Le tre sensazioni primarie che accompagnano l'ascolto di un suono sono: *altezza*, *intensità* e *timbro* (in inglese, rispettivamente *pitch*, *loudness* e *timbre*). Queste sensazioni sono il risultato della elaborazione dell'orecchio e del cervello e non sono grandezze misurabili direttamente (sono, invece, misurabili le quantità fisiche del suono che sono principali cause di altezza, intensità e timbro). Per suoni periodici (o quasi periodici), il pitch è determinato principalmente dalla frequenza fondamentale (si ricorda che la frequenza fondamentale di un suono è il numero di ripetizioni in un secondo del pattern di vibrazione). Tra i suoni periodici si distinguono i suoni puri, formati cioè da una sola componente sinusoidale. Visto nel dominio della frequenza, un suono puro è rappresentato con una riga in corrispondenza della frequenza della sinusoide. Per un suono composto da più armoniche (sinusoidi), la frequenza fondamentale è il massimo comun divisore della serie di frequenze che costituiscono lo spettro.

L'intensità percepita dipende dal flusso di energia che accompagna la vibrazione. Essa è tuttavia dipendente anche da altri fattori quali pitch, durata e presenza di altri suoni.

Il principale elemento che caratterizza il timbro è lo spettro di energia. Come si avrà modo di vedere, l'evoluzione temporale dello spettro è un elemento fondamentale per il riconoscimento e la caratterizzazione dei suoni strumentali. Se un suono viene privato del proprio attacco, viene persa, nella gran parte dei casi, la capacità dell'ascoltatore di riconoscere chiaramente lo strumento acustico.

Oltre alle sensazioni primarie appena viste, ne esistono altre che rivestono notevole importanza nel processo percettivo. La *direzionalità* è una sensazione legata alla percezione dei suoni, ed è legata alla differenza di fase con cui il suono giunge alle orecchie (differenza che è dovuta alla direzione di incidenza del suono). *Consonanza* e *dissonanza* sono infine due sensazioni (indotte, legate cioè al condizionamento culturale) provocate dalla sovrapposizione di due o più suoni e dai battimenti dovuti alla sovrapposizione (si parlerà nel seguito del fenomeno dei battimenti). Il rapporto fra le frequenze e la frequenza di battimento determinano il grado di consonanza e dissonanza percepito.

I messaggi musicali sono composti da suoni che si avvicinano nel tempo e da pattern ritmici che si ripetono nel tempo. Il tempo ha dunque un ruolo determinante nella costruzione di elementi essenziali del messaggio musicale come la melodia ed il ritmo. Nella Tabella 2.7 si mettono a confronto le sensazioni dell'udito con le scale dei tempi relative alla loro elaborazione e con lo stadio del processo uditivo in cui tali sensazioni sono elaborate.

SCALA DEI TEMPI (s)	LUOGO DI ELABORAZIONE	SENSAZIONE	INFLUENZA (Cultura, Ambiente, Stato)
$6 * 10^{-5} \div 6 * 10^{-2}$	orecchio interno	altezza, intensità, timbro	debole
~ 0.1	collegamento nervoso tra orecchio e corteccia	transitori, timbro, direzionalità, identificazione, discriminazione	
> 0.1	corteccia cerebrale	ritmo, messaggio musicale	
	emisfero sinistro	breve termine (sequenzializzazione, parlato)	
	emisfero destro	lungo termine (integrazione spaziale (visiva) e temporale (uditiva))	forte

Tabella 2.4: Scala dei tempi e luogo di elaborazione per le sensazioni

L'indagine quantitativa sulle sensazioni appena viste è condotta attraverso misure psicofisiche su soggetti umani. Oggetto delle misure (i cui dati sono trattati solitamente con metodi statistici) sono soglie, soglie differenziali, eguaglianza e scale di valori. Misure di soglia e soglia differenziale mirano a stabilire ad esempio a quale intensità un suono inizia ad essere percepito o qual'è la minima variazione di frequenza per cui la sensazione di pitch cambia; misure di uguaglianza servono a stabilire ad esempio quando due suoni a frequenza diversa hanno uguale intensità; misure su scale di valori mirano a stabilire ad esempio quando due suoni sono percepiti con altezza o intensità doppia o tripla.

2.8 L'organo dell'udito

Il sistema uditivo umano ha una struttura complessa e svolge funzioni notevolmente avanzate. Non solo è in grado di elaborare un ampio insieme di stimoli, ma può identificare precisamente l'altezza o il timbro di un suono, o la direzione da cui esso proviene. Molte funzioni del sistema uditivo vengono svolte dall'organo che chiamiamo *orecchio*, ma grande enfasi di recente viene attribuita alla elaborazione che ha luogo nel sistema nervoso centrale. Per semplificarne la descrizione, l'orecchio è spesso diviso in tre parti principali: orecchio *esterno*, orecchio *medio* ed orecchio *interno* (Figura 2.6). L'orecchio esterno è formato dalla *pinna* esterna e dal *canale uditivo* (*meato*). L'orecchio medio inizia con la membrana del *timpano*, alla quale è attaccato il primo dei tre ossicini (chiamati *martello*, *incudine* e *staffa*) che compongono questo stadio. Il compito di questi ultimi è quello di amplificare il moto del timpano (essi formano un sistema di leve) e di trasferirlo ad un'altra membrana, la *finestra ovale*. Con la finestra ovale inizia l'orecchio interno, formato principalmente dalla *coclea*. La coclea contiene i meccanismi per trasformare le variazioni di pressione in corrispondenza del timpano in impulsi nervosi che vengono interpretati dal cervello come suono.

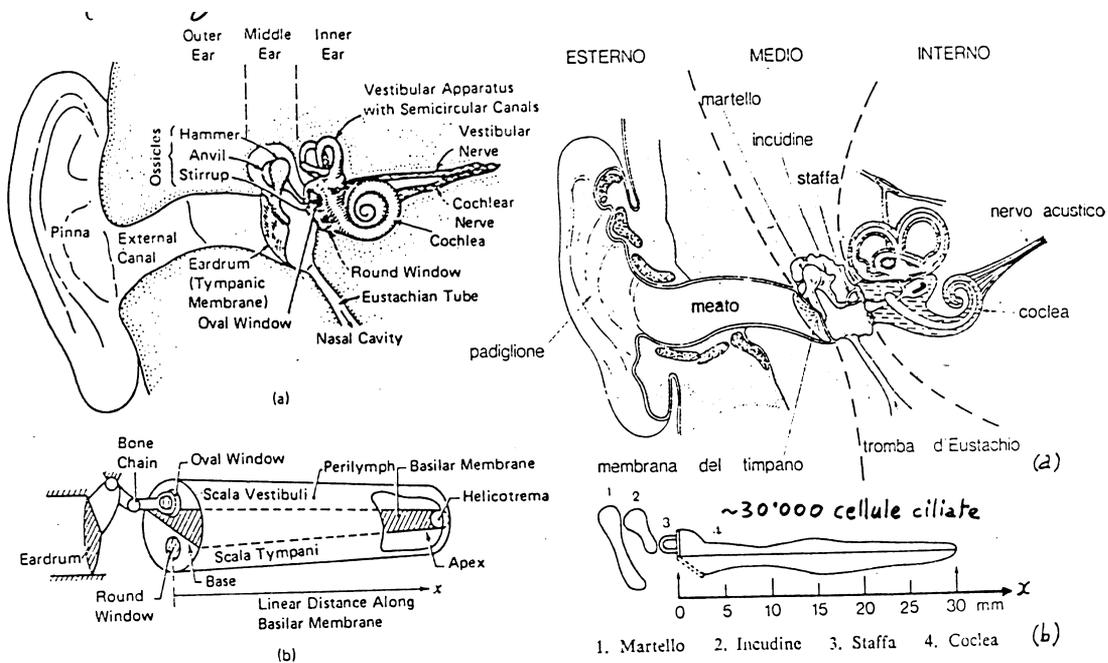


Figura 2.6: Le funzioni dei sistemi sorgente, mezzo e ricevitore

Nel dettaglio, vediamo come la coclea ci aiuta a percepire l'altezza (pitch) di un suono puro. Questo organo è costituito da una cavità divisa in due da una membrana (*membrana basilare*) e contenente un liquido incompressibile (*perilinfo*) che passa da una parte all'altra della membrana attraverso un'apertura (*elicotrema*). Le vibrazioni trasmesse dall'orecchio medio alla finestra ovale si trasmettono al fluido del dotto cocleare, che a sua volta provoca un moto della membrana basilare simile a quello di una bandiera. Lungo la membrana basilare, circa 30000 recettori nervosi (*cellule ciliate*) convertono il moto della membrana in segnali che sono a loro volta trasmessi ai neuroni del nervo acustico. E' importante sottolineare che, a fronte di un suono puro di una data frequenza, il massimo della ampiezza

di oscillazione della membrana basilare è localizzato in una regione ben delimitata della membrana. La posizione di questa regione dipende dalla frequenza del suono. Per ogni frequenza c'è dunque una regione di massima sensibilità della membrana (regione di risonanza). Più bassa è la frequenza e più la regione di risonanza è prossima all'*Apex* (elicotrema). L'estensione delle frequenze udibili va da 16 Hz a 20 kHz.

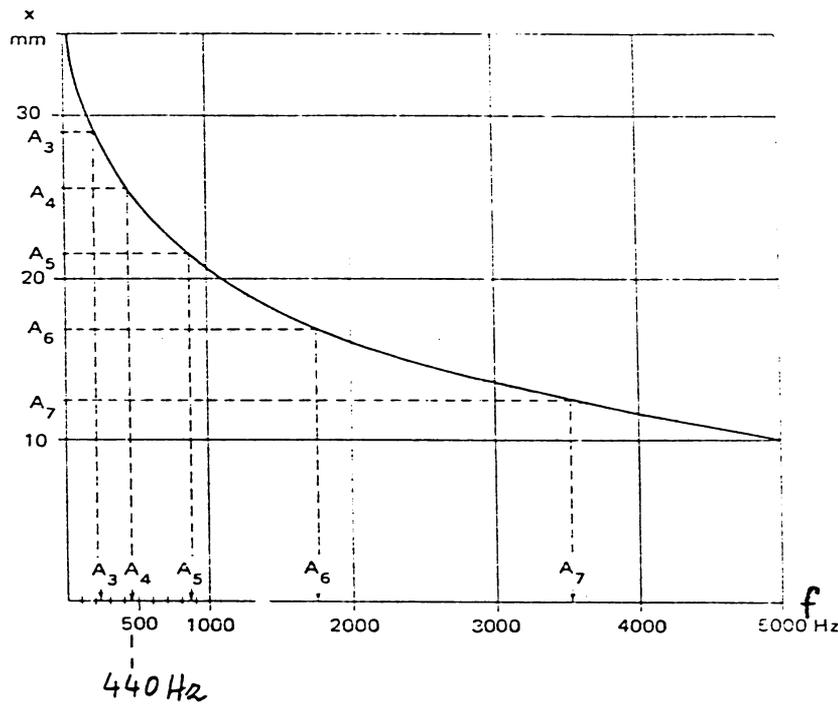


Figura 2.7: Posizione della zona di risonanza sulla membrana basilare

In Figura 2.7 si può osservare come la posizione x (misurata dalla base, Figura 2.6) della regione di massima risonanza varia al variare della frequenza f di un suono puro. Dalla figura è possibile trarre alcune considerazioni fondamentali:

- L'estensione di frequenze che va approssimativamente da 20 Hz fino a 4000 Hz copre circa i due terzi dell'estensione della membrana basilare (dai 12 ai 35 mm dalla base). La rimanente porzione della scala di frequenze (4000 - 16000 Hz) è compressa nel rimanente terzo. Il range di frequenze visto corrisponde alle prime 7 ottave musicali, riconosciute come le più importanti in musica.
- A fronte di un raddoppio della frequenza (salto di ottava) dello stimolo sinusoidale la regione di risonanza subisce uno spostamento costante di 3.5-4 mm, indipendentemente dalla frequenza di partenza. In altre parole, quando la frequenza f è moltiplicata per un dato valore, la posizione del massimo di risonanza viene traslata di una certa quantità seguendo una legge di tipo logaritmico.

2.8.1 Percezione di altezza dei suoni puri e discriminazione del pitch (jnd di frequenza)

Il pitch è una sensazione soggettiva. In molte scale musicali si tende a considerare l'*ottava* come unità fondamentale: note giudicate essere l'una l'ottava dell'altra hanno frequenze l'una il doppio dell'altra, anche se il rapporto di 2:1 non sempre è esatto. Tuttavia esiste una discordanza, particolarmente evidente al di sopra dei 1000 Hz, fra la frequenza reale del suono puro e l'altezza mediamente percepita dall'ascoltatore (Figura 2.8).

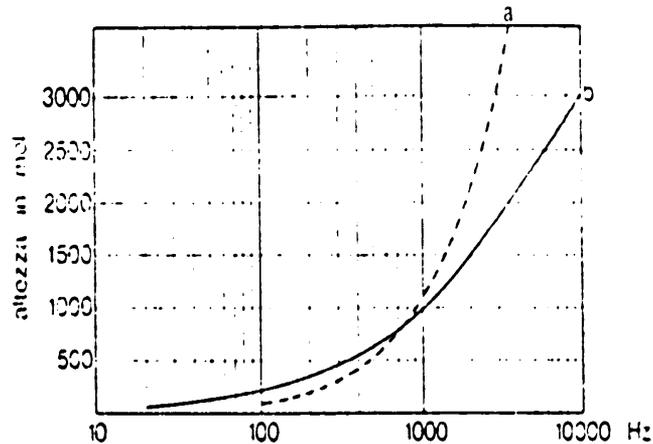


Figura 2.8: Altezza in frequenza (linea tratteggiata) e giudizio medio (linea continua)

Allo scopo di avere per il pitch una scala coerente con la curva di percezione dell'altezza, è stata introdotta la scala *mel* (Figura 2.9). Per definizione, infatti, a 1000 Hz corrispondono 1000 mel (con pressione sonora 60 db sopra la soglia di udibilità a 1000 Hz) e ad ogni ottava i mel si raddoppiano (o si dimezzano). La scala mel è una scala psicofisica del pitch.

La capacità di distinguere tra due stimoli pressoché uguali è spesso caratterizzata, negli studi psicofisici, da una misura di minima differenza apprezzabile (*just noticeable difference, jnd*). Due stimoli sono giudicati uguali se differiscono per meno del jnd. In psicoacustica si incontrano misure di jnd per molte delle sensazioni uditive. Gli studi sulla percezione del pitch hanno mostrato che il jnd di frequenza dipende, oltre che dal valore di frequenza iniziale dello stimolo, anche dall'intensità sonora, dalla durata e dalla velocità di variazione della frequenza (per cambiamenti improvvisi le soglie si abbassano anche di 30 volte). La Figura 2.10 mostra il jnd medio per suoni puri al variare della frequenza. Si può notare che la risoluzione di frequenza (definita come $\frac{jnd}{f_c}$, con f_c frequenza centrale) è massima intorno ai 2000 Hz e raggiunge il suo minimo alle basse frequenze.

La descrizione del meccanismo di discriminazione delle frequenze (noto come *teoria della localizzazione*) cerca di spiegare il fenomeno della percezione del pitch attraverso la conversione di una vibrazione temporale in una vibrazione nello spazio ad opera della coclea. Questa teoria spiega alcuni fenomeni, ma non fornisce una spiegazione completa del funzionamento della percezione dell'altezza. Non spiega, ad esempio, perché percepiamo suoni complessi come una sola entità avente una altezza ben definita, pur risuonando la membrana basilare in corrispondenza di ogni componente del suono. Né spiega perché percepiamo l'altezza corretta anche in suoni complessi nei quali la componente fondamentale sia stata eliminata (fenomeno noto come *ricostruzione della fondamentale*). Si

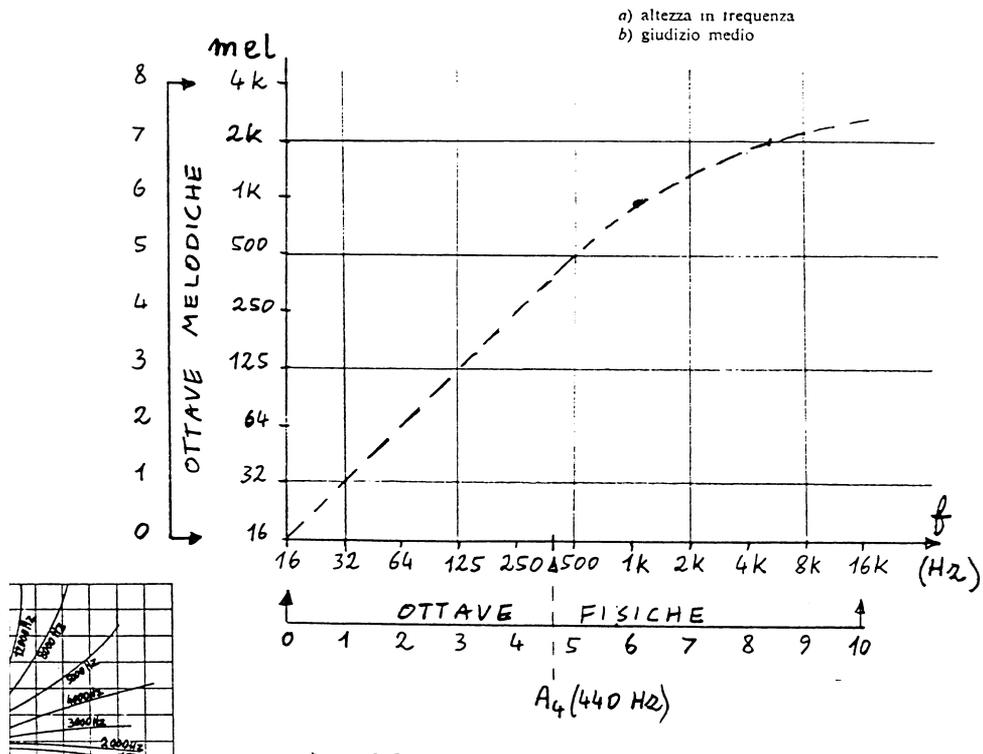


Figura 2.9: Scala mel

vedrà in seguito come alla teoria spaziale ne venga affiancata un'altra, detta teoria temporale o della periodicità, per cercare di spiegare questi fenomeni.

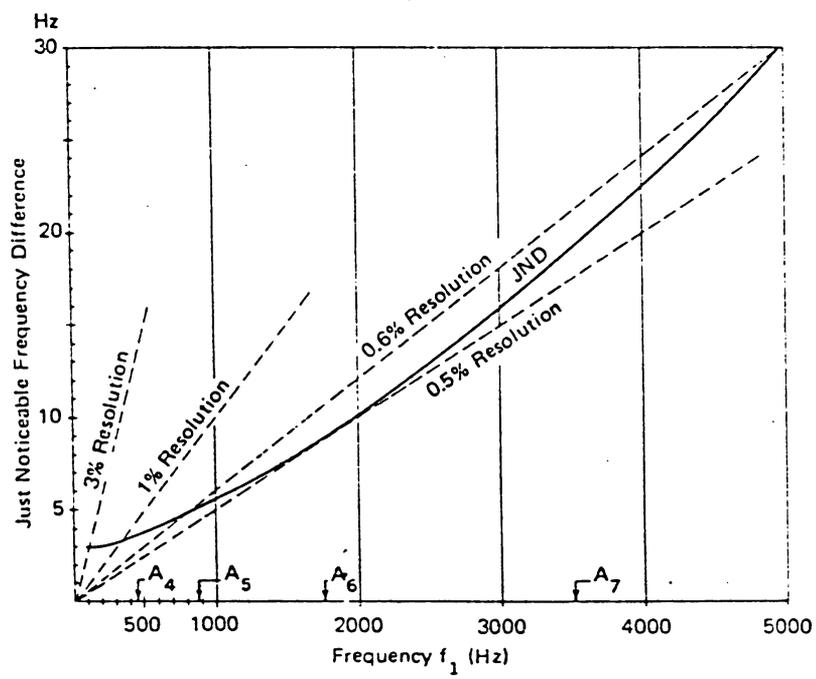


Figura 2.10: Soglia differenziale (jnd) per la frequenza

2.8.2 Pitch e intensità dei suoni puri

La frequenza dello stimolo sinusoidale risulta essere il parametro fondamentale nella determinazione della sensazione di altezza. Tuttavia, esso non è il solo. Gli esperimenti di Stevens (1937) hanno mostrato che al crescere dell'intensità dello stimolo da 40 a 90 db, per frequenze al di sopra dei 1000 Hz il pitch percepito subisce un incremento, mentre per frequenze al di sotto dei 1000 Hz esso subisce un decremento rispetto all'intensità iniziale. Per frequenze intorno ai 1000 Hz, la variazione è quasi nulla. Questi dati sono riassunti nel diagramma di Stevens e Wolkmann (Figura 2.11).

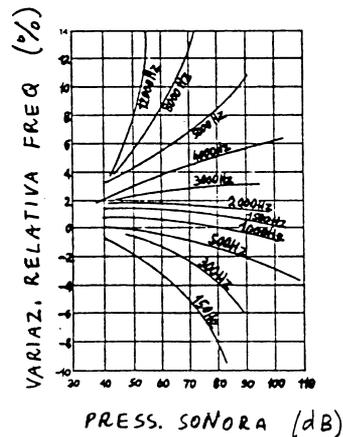


Figura 2.11: Diagramma di Stevens e Wolkmann (1937)

2.9 Sovrapposizione di suoni puri

In tutti gli esperimenti di cui si è trattato fin ora, si è fatto implicitamente riferimento a stimoli uditivi composti da una sola componente sinusoidale. Si vuole vedere adesso quali sono gli effetti prodotti dalla sovrapposizione di due suoni puri. Classificheremo gli effetti risultanti in due famiglie: gli effetti del primo ordine e quelli del secondo ordine. Gli effetti del primo ordine sono caratterizzati dall'essere elaborati nell'orecchio interno (elaborazione meccanica), mentre quelli del secondo ordine sono caratterizzati dall'essere elaborati in uno stadio successivo (elaborazione neurale).

2.9.1 Gli effetti del I ordine

Effetti del primo ordine emergono quando lo stimolo uditivo è costituito da due suoni puri con stessa frequenza e fase, e la frequenza di uno dei due stimoli viene fatta crescere (o decrescere) gradualmente. Sia f_1 la frequenza del primo suono e f_2 quella del secondo e sia inizialmente $f_1 = f_2$. Finché i due suoni presentano stessa frequenza e stessa fase iniziale, la membrana basilare presenta una eccitazione in corrispondenza della posizione relativa alla frequenza comune, di ampiezza pari alla somma delle ampiezze dei due stimoli. Quando la frequenza di uno dei due stimoli comincia a crescere ($f_2 = f_1 + \Delta f$) e fino a quando Δf non supera un certo valore Δf_d , l'ascoltatore percepisce un unico suono a frequenza $f = (f_1 + f_2)/2$ e modulato in ampiezza. Questa modulazione di ampiezza è chiamata

battimento del I ordine. La frequenza di battimento del primo ordine è pari a $fb = (f_2 - f_1)$. La spiegazione è nella formula della somma di due suoni puri:

$$\sin \omega_1 t + \sin \omega_2 t = 2 \sin \frac{(\omega_1 + \omega_2)t}{2} \cos \frac{(\omega_1 - \omega_2)t}{2} \quad (2.1)$$

con $\omega t = 2\pi f t$. Si vede che l'argomento del seno determina l'altezza del suono ($f = \frac{f_1+f_2}{2} = f_1 + \frac{\Delta f}{2}$), mentre l'argomento del coseno determina la frequenza di battimento $f_b = 2(\frac{f_2-f_1}{2}) = f_2 - f_1 = \Delta f$ (figg. 2.12 e 2.13).

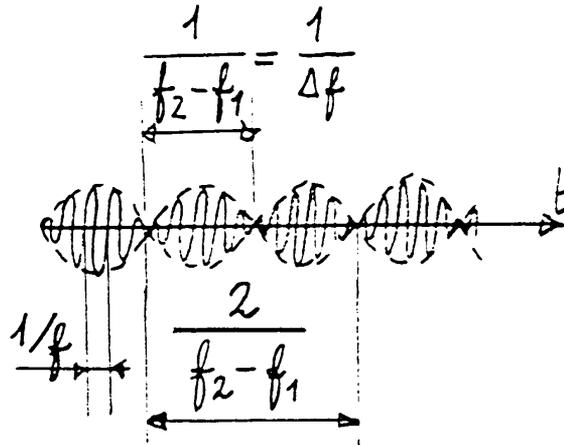


Figura 2.12: Battimenti del I ordine: frequenza di battimento

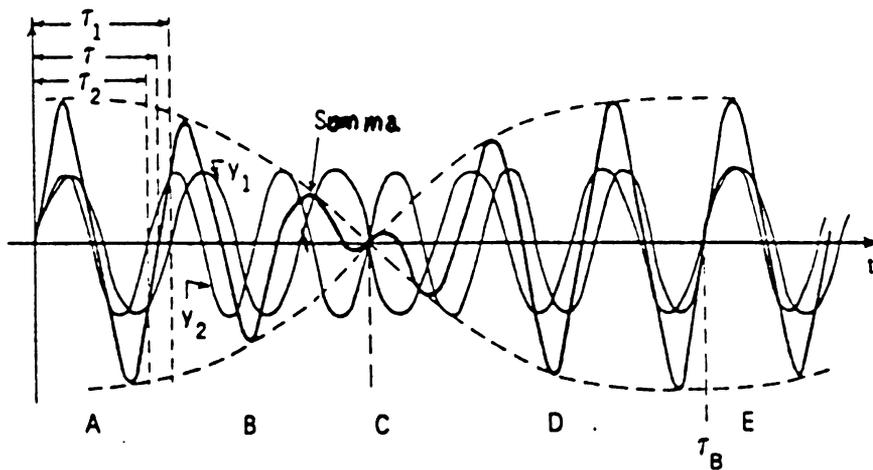


Figura 2.13: Battimenti del I ordine: pitch risultante ($\tau_1 = \frac{1}{f_1}, \tau_2 = \frac{1}{f_2}, \tau = \frac{1}{f}$)

All'aumentare del modulo di Δf oltre i 15Hz, la sensazione di battimento scompare, lasciando il posto a una sensazione sgradevole di ruvidità (*roughness*). Quando Δf supera in modulo la grandezza

Δf_D (soglia di discriminazione) i due suoni risultano distinguibili, pur esistendo ancora la sensazione di ruvidità del suono. Solo quando il modulo di Δf supera una seconda soglia Δf_{CB} , la sensazione dei due suoni distinti risulta netta e piacevole. La grandezza $2\Delta f_{CB}$ è chiamata banda critica (Figura 2.14).

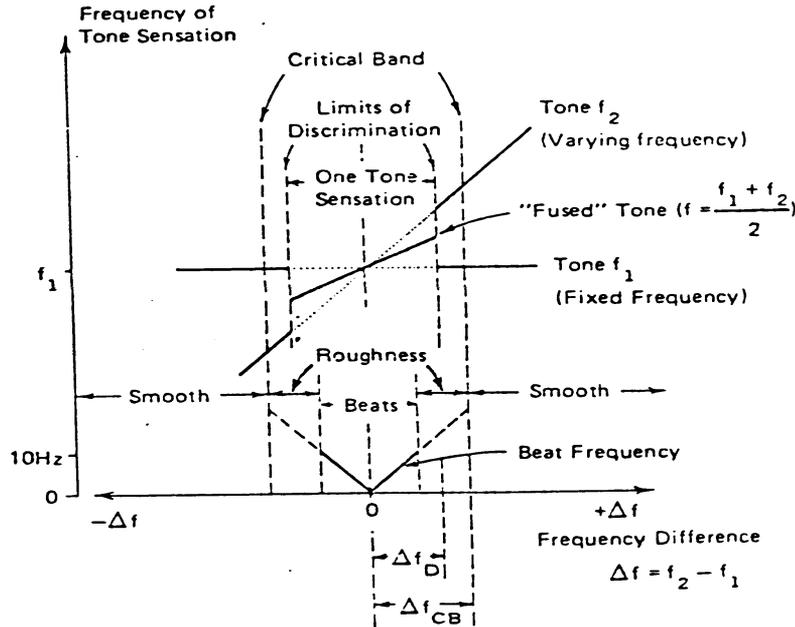


Figura 2.14: Sensazioni uditive al variare della differenza di frequenza

La Figura 2.15 mostra come la discriminazione del pitch Δf_D e la banda critica Δf_{CB} dipendono dalla frequenza centrale. Dalla figura si può vedere anche come alcuni intervalli musicali siano consonanti o dissonanti a seconda della frequenza centrale (semitono, tono e terza minore hanno rapporti di frequenza rispettivamente pari a 16/15, 9/8 e 6/5). Si noti come l'intervallo di mezzo tono è dissonante su tutta l'estensione delle frequenze, mentre l'intervallo di terza minore risulta dissonante solo al di sotto dei 600 Hz circa.

Si può notare, infine, che fra jnd e bande critiche vale la relazione approssimata:

$$\Delta f_{CB} = 30 jnd \quad (2.2)$$

Una banda critica corrisponde a circa 1.3 mm di membrana basilare e a circa 1300 ricettori.

Tra gli effetti del primo ordine vi è poi quello dei *suoni di combinazione*. Questi sono suoni che vengono percepiti anche se non sono originariamente presenti nello stimolo, e sono il risultato di distorsioni non lineari presenti al passaggio del segnale da orecchio esterno a orecchio interno. L'esperimento descritto in precedenza viene condotto in questo caso aumentando l'intensità dei due suoni e facendo variare la frequenza f_2 da f_1 a $2f_1$. Sia x lo stimolo costituito da due suoni puri e sia y il segnale distorto a causa delle non linearità del passaggio fra orecchio esterno ed interno. Se esplicitiamo la funzione nonlineare fermandoci al termine del secondo ordine otteniamo le formule:

$$x = k(\sin \omega_1 t + \sin \omega_2 t) \quad (2.3)$$

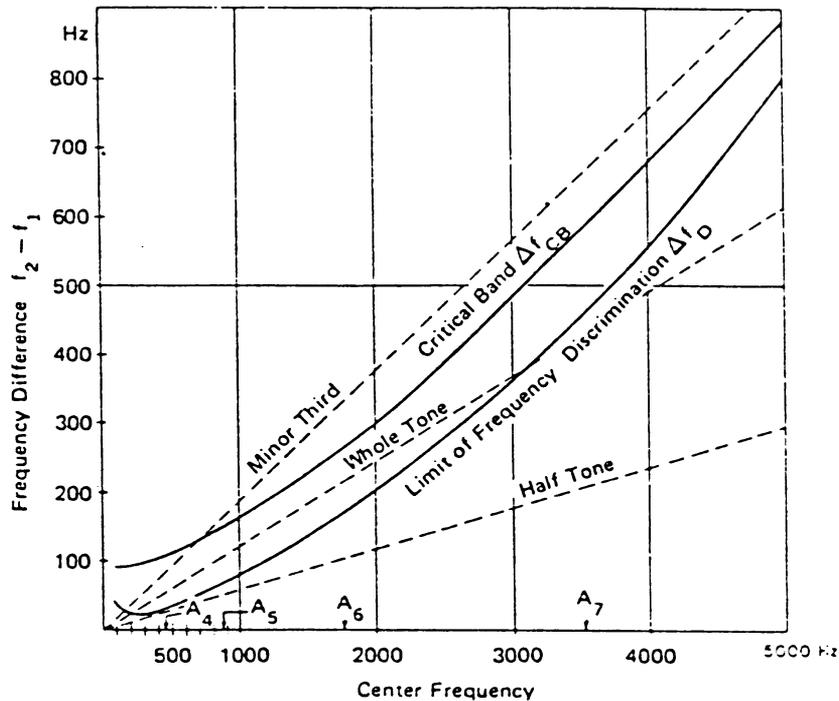


Figura 2.15: Banda critica e discriminazione del pitch al variare della frequenza centrale

$$y = a_0 + a_1x + a_2x^2 \quad (2.4)$$

$$\begin{aligned} x^2 &= k^2(\sin^2 \omega_1 t + \sin^2 \omega_2 t + 2 \sin \omega_1 t \sin \omega_2 t) \\ &= k^2 \frac{1}{2} [(1 - \cos 2\omega_1 t) + (1 - \cos 2\omega_2 t) + \cos(\omega_2 - \omega_1)t - \cos(\omega_2 + \omega_1)t] \end{aligned} \quad (2.5)$$

I suoni di combinazione percepiti avranno frequenza $f_1, f_2, f_2 - f_1, f_1 + f_2, 2f_1, 2f_2$. Se nello sviluppo in serie di potenze della nonlinearity consideriamo termini superiori a quello del secondo ordine, ulteriori suoni di combinazione sono calcolabili. In Figura 2.16 è mostrato l'andamento dei suoni a frequenza $f_{c1} = f_2 - f_1, f_{c2} = 2f_1 - f_2, f_{c3} = 3f_1 - 2f_2$, al variare di f_2 da f_1 a $2f_1$.

Gli effetti dovuti alla nonlinearity vengono percepiti anche a fronte dell'ascolto di un solo suono di intensità estremamente elevata. In questo caso saranno percepiti suoni a frequenze $2f_1, 3f_1, 4f_1, \dots$ (*armoniche auricolari*).

2.9.2 Gli effetti del II ordine

Gli effetti del II ordine sono il risultato dell'elaborazione degli stimoli da parte del sistema nervoso. Se gli effetti del primo ordine avevano origine già sulla membrana basilare a causa della sovrapposizione di due regioni di risonanza, quelli del secondo ordine hanno la caratteristica di non derivare da simili cause fisiche. Analisi sperimentali hanno dimostrato che non vi sono tracce di essi nei pattern di vibrazione del liquido cocleare e che quindi essi devono avere origine in uno stadio di elaborazione successivo.

Il fenomeno dei *battimenti del secondo ordine* consiste nella sensazione di modulazione di ampiezza che si avverte quando lo stimolo è composto da due suoni puri che eccitano la membrana

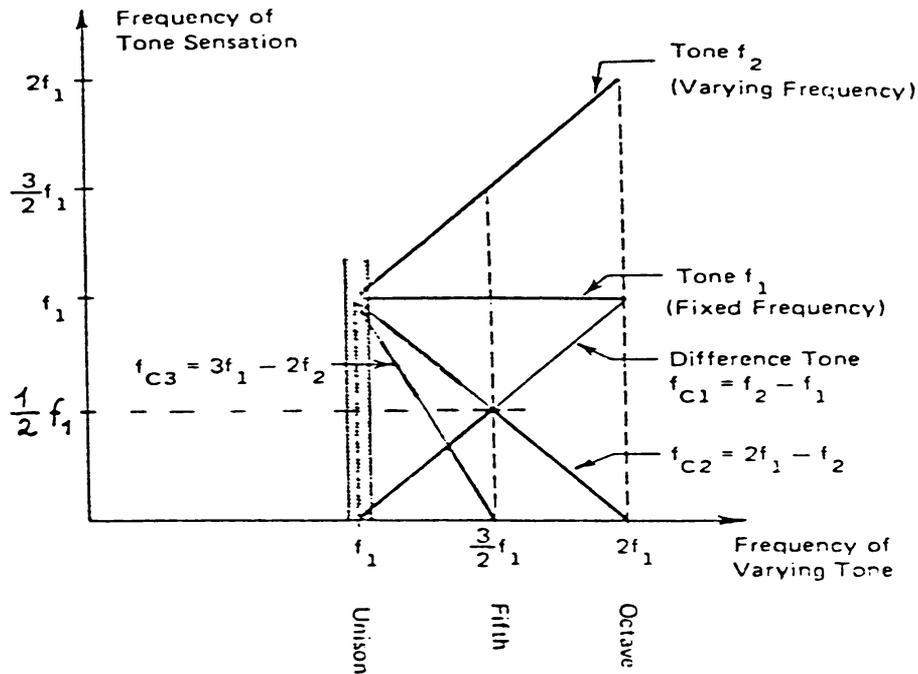


Figura 2.16: Frequenza dei suoni di combinazione f_{c1} , f_{c2} e f_{c3} evocati dalla sovrapposizione di due suoni a frequenza f_1 e f_2

basilare in regioni che non si sovrappongono. La frequenza f_2 sia posta inizialmente ad un valore $f_2 = 2f_1$. Si può notare che per diverse differenze di fase iniziale fra le due componenti la forma d'onda cambia notevolmente (Figura 2.17). Se la differenza di fase si mantiene perfettamente costante, tuttavia, l'ascoltatore non percepirà alcuna differenza. Quando la frequenza f_2 viene leggermente stonata rispetto all'ottava ($f_2 = 2f_1 + \epsilon$), la differenza di fase non rimane più costante. Il sistema uditivo percepisce in queste condizioni battimenti a frequenza $f_b = \epsilon$. Battimenti del secondo ordine si ottengono anche stonando leggermente intervalli di quarta ($f_2 = 4/3f_1$) e di quinta ($3/2f_1$), con frequenze di battimento rispettivamente di $f_b = 3\epsilon$ e $f_b = 2\epsilon$ Hz. Questo fenomeno mette bene in risalto come il senso dell'udito sia insensibile a differenze di fase costanti nel tempo, ma sensibile alla variazione nel tempo dello sfasamento.

Una importante differenza tra battimenti del primo ordine e battimenti del secondo ordine è la seguente: i primi presentano una modulazione di ampiezza senza mutazione della forma d'onda nel tempo; i secondi, al contrario, presentano un cambiamento ciclico del pattern di vibrazione senza che vi sia mutazione nella ampiezza totale.

Tra gli effetti del II ordine vi è ancora quello del *rintracciamento della fondamentale* (o *virtual pitch*). Consideriamo in questo caso un suono composto da numerose componenti, esattamente armoniche, con frequenze $f_1, 2f_1, 3f_1, \dots, nf_1$. Il pitch percepito è naturalmente il massimo comun divisore della serie di armoniche, cioè f_1 . Se ora viene generato uno stimolo simile al primo, ma da cui sia stata tolta la prima armonica con frequenza f_1 (o le prime l armoniche), la sensazione di pitch risultante dall'ascolto risulta uguale a quella del suono originario. Il sistema uditivo ha dunque ricostruito la componente mancante. Questa sensazione è più netta se sono presenti le armoniche più prossime alla

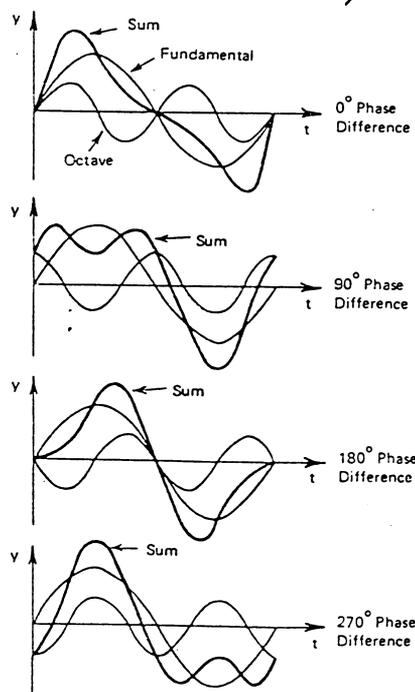


Figura 2.17: Sovrapposizione di due suoni puri con diverso sfasamento costante

fondamentale ($2f_1, 3f_1, 4f_1, \dots$), meno netta se sono presenti quelle a partire da un numero d'armonica superiore al quarto.

Il rintracciamento della fondamentale è un fenomeno che si può osservare anche con stimoli formati da due suoni puri. Ad esempio, se la coppia di suoni puri presenta rapporti di frequenze quali $f_2 = (3/2)f_1$ o $f_2 = (4/3)f_1$, il pattern di vibrazione risultante è caratterizzato da periodo maggiore. In particolare, a livello neurale viene generata la sensazione di un suono a frequenza $f_0 = (1/2)f_1$ e $f_0 = (1/3)f_1$ rispettivamente. In Figura 2.18 è mostrato il pattern di vibrazione risultante nel caso $f_2 = \frac{3}{2}f_1$.

Gli effetti del II ordine sono il risultato di una elaborazione da parte del sistema nervoso centrale. Ciò vuol dire che, al contrario degli effetti del I ordine, essi vengono percepiti anche con ascolto binaurale (quando, cioè, ciascuno dei due stimoli (contemporanei) è presentato ad un solo orecchio).

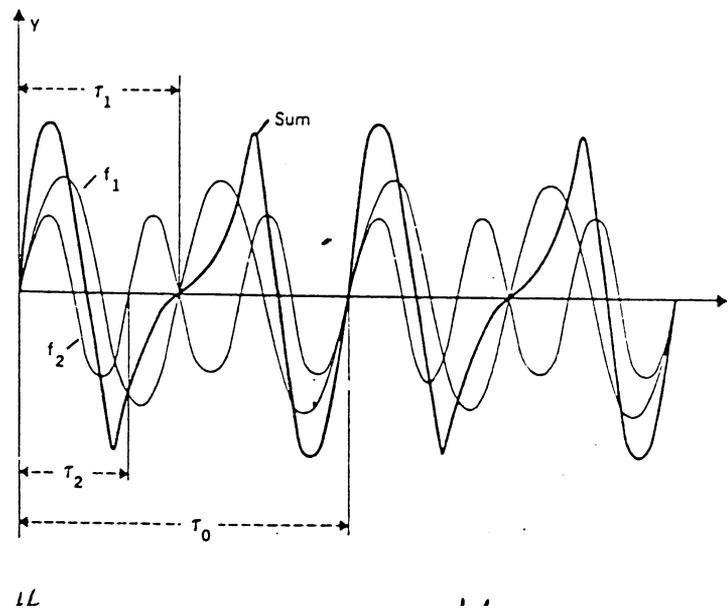


Figura 2.18: Ricostruzione della fondamentale ($\tau_1 = \frac{1}{f_1}$, $\tau_2 = \frac{1}{f_2}$, $\tau_0 = \frac{1}{f_0}$)

2.10 Elaborazione dello stimolo uditivo nel sistema nervoso

Gli effetti del secondo ordine mettono in evidenza l'incompletezza della teoria della localizzazione per spiegare molti fenomeni uditivi. La percezione dei battimenti può essere spiegata se si ipotizza un sistema di analisi dei pattern temporali della vibrazione. Il meccanismo di ricostruzione della fondamentale, d'altra parte, può essere spiegato anche con un sistema di analisi dell'informazione neurale generata dalla particolare configurazione spaziale che l'eccitazione della membrana basilare assume a fronte di un suono complesso. Al fine di capire meglio le teorie proposte negli ultimi anni sul funzionamento del processo uditivo, è necessario introdurre alcune nozioni sul funzionamento del sistema nervoso uditivo.

La prima funzione per il trasferimento dell'informazione localizzata sulla membrana basilare verso il sistema nervoso centrale è espletata dalle *cellule ciliate*. Questi ricettori coprono la membrana basilare per tutta la sua estensione e fanno da ponte verso le terminazioni nervose. Essi sono soggetti a una sollecitazione meccanica ogni volta che la membrana è eccitata nella zona corrispondente e provocano un impulso elettrico nelle terminazioni nervose ogni volta che tale sollecitazione supera una certa soglia. Le cellule che formano le terminazioni nervose, e che sono l'elemento fondamentale di elaborazione e trasmissione nel sistema nervoso, sono chiamate *neuroni* (fig 2.19). Nel neurone si distinguono il *corpo*, i *dentriti* e gli *assoni*. I dentriti ed il corpo della cellula sono i ricettori dei segnali neurali provenienti da altre cellule, mentre attraverso l'assone il neurone passa l'impulso elettrico ad altri neuroni nei punti di contatto con essi (*sinapsi*). I segnali neurali consistono in impulsi elettrici (*potenziali di azione*) dell'ordine delle decine di millivolt e della durata di pochi millisecondi. Vi sono sinapsi di tipo eccitatorio e di tipo inibitorio. Quando, in un certo intervallo di tempo, un neurone riceve un numero di stimolazioni eccitatorie che supera di un certo valore di soglia il numero di stimolazioni inibitorie, esso sarà indotto a produrre uno stimolo a sua volta. Questo potenziale (eccitatorio o inibitorio a seconda della natura del neurone) verrà trasmesso ad altre cellule nervose mediante le sinapsi presenti sull'assone del neurone. È importante notare che un neurone emette un

impulso elettrico in funzione della distribuzione temporale e spaziale dei segnali presinaptici.

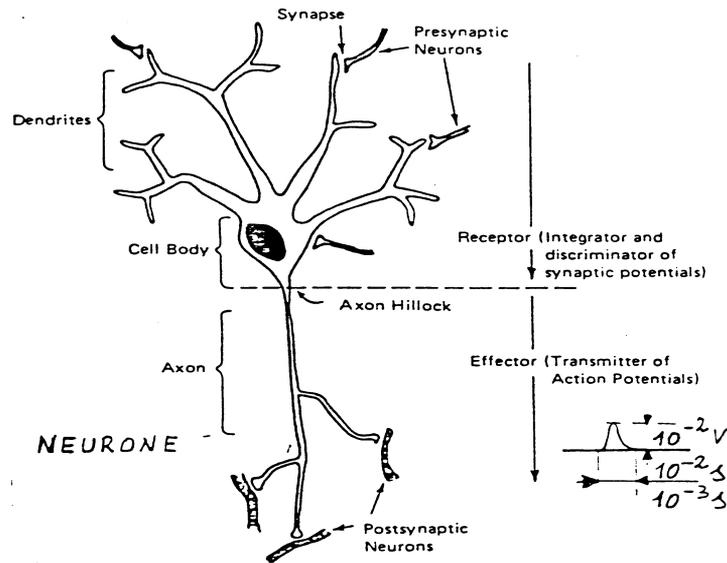


Figura 2.19: Neurone

Possiamo ora descrivere come il sistema nervoso riceve le informazioni dall'organo della coclea. Quando un suono puro provoca il moto in corrispondenza di una zona di risonanza della membrana basilare le cellule ciliate vengono sollecitate e provocano un treno di impulsi elettrici nelle fibre nervose ad esse collegate. Avviene che la densità temporale degli impulsi nervosi dipende dalla velocità con cui la posizione della membrana passa dalla *Scala Vestibuli* alla *Scala Timpani* (Figura 2.6). L'inibizione della trasmissione si manifesta nel passaggio inverso e attività minore si osserva in situazione di velocità minima. La Figura 2.20 mostra l'andamento temporale del treno di impulsi provvati da un'onda periodica a bassa frequenza. Da questa analisi si evince che una fibra nervosa del nervo uditivo è capace di trasmettere i seguenti tipi di informazioni: a) posizione della risonanza sulla membrana (ogni fibra nervosa è associata a una zona); b) distribuzione temporale degli impulsi e dunque periodicità e configurazione della forma d'onda.

Possiamo ora chiederci come le informazioni sulla distribuzione temporale degli impulsi nervosi possa venire usata dal sistema nervoso per elaborare un effetto come i battimenti del secondo tipo. Se due stimoli con frequenze a distanza di un'ottava eccitano la coclea, due fibre nervose in corrispondenza delle due zone di massima risonanza saranno attivate (informazione spaziale, che non spiega la sensazione del battimento). Tuttavia, alla particolare forma d'onda corrisponde un particolare pattern periodico di impulsi nervosi. La periodicità di tale pattern dà informazioni sulla frequenza di ripetizione, mentre la particolare conformazione dà informazioni sul pattern di vibrazione. La figura 2.21 (un esempio di istogramma delle occorrenze di intervalli temporali tra spike un una fibra nervosa) dà un'informazione statistica legata al pattern di vibrazione. Questo tipo di analisi si ha verosimilmente per suoni a bassa frequenza, mentre alle alte frequenze l'informazione dovuta al pattern diventa confusa. L'analisi dettagliata della distribuzione temporale degli impulsi richiede dunque un ulteriore meccanismo, detto comunemente di *autocorrelazione temporale*, che mette in rilievo le caratteristiche periodiche del pattern impulsivo e che sopprime le altre, sulla base della comparazione del treno di impulsi attuale con treni di impulsi precedenti.

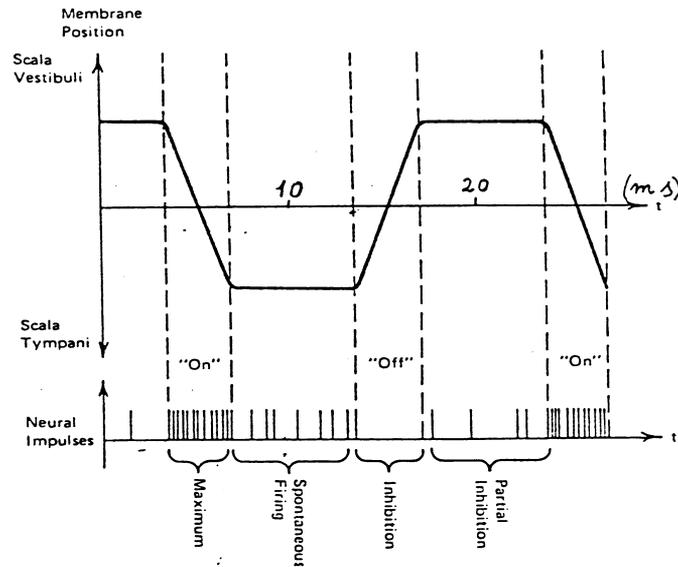


Figura 2.20: Treno di impulsi generato da un'onda periodica

Un simile meccanismo di correlazione temporale è responsabile della sensazione di spazializzazione del suono. Allo scopo di stabilire dei ritardi temporali e quindi per elaborare informazioni di localizzazione della sorgente sonora, il sistema nervoso si serve della *crosscorrelazione* fra segnali neurali provenienti dalle due orecchie. Un modello di crosscorrelazione neurale, in cui un neurone è attivato solo quando sia eccitato simultaneamente dalle due fibre nervose provenienti dalle due orecchie, è illustrato in figura 2.22.

I sistemi di analisi temporale e spaziale ora descritti sono modelli a cui si è fatto ricorso per spiegare molti fenomeni percettivi, spesso legati a suoni puri o composti da componenti armoniche. Un ulteriore fenomeno uditivo fondamentale è quello per cui l'ascolto di un suono spettralmente ricco con componenti in relazione armonica fra loro produce un percepito unico avente pitch determinato dalla componente fondamentale, anziché dare luogo alla percezione di tanti suoni separati, ognuno con pitch legato alla frequenza della singola componente. Nessuna delle due teorie viste (temporale e spaziale) è in realtà in grado di fornire una spiegazione esauriente. Per spiegare come sia possibile assegnare ad un complesso spettrale un pitch unico, si fa allora l'ipotesi di esistenza di un *sistema centrale di elaborazione dell'altezza*. Questo sistema agisce ad alto livello riconoscendo che strutture sonore di natura armonica riproducono sulla membrana basilare pattern simili per proprietà (ad esempio, al crescere dell'ordine della parziale le distanze fra zone di risonanza mantengono una certa relazione definita, si veda figura 2.23). La funzione dell'elaboratore di altezza è quella di assegnare un pitch unico a eventi di questo tipo. Attraverso criteri di matching con templates di base (sorta di database di pattern spaziali), l'elaboratore è anche in grado di assegnare un pitch a un suono senza fondamentale, elaborando il matching parziale.

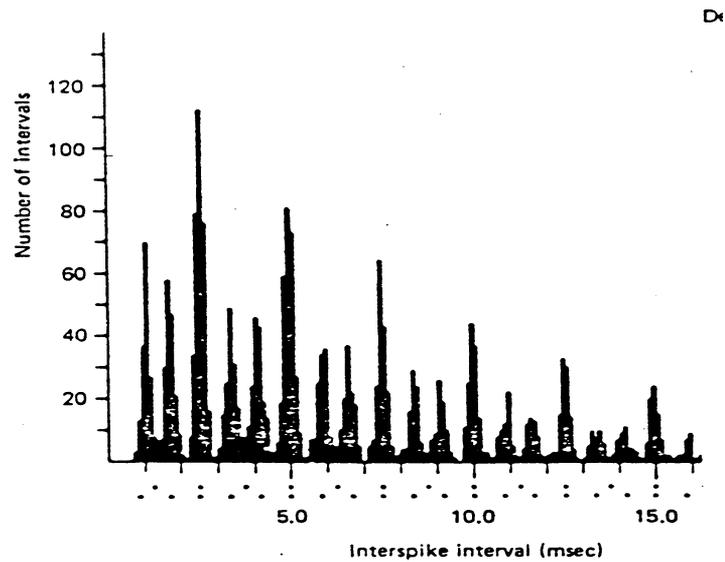


Figura 2.21: Istogramma del numero di occorrenze di dato tempo di intercorrenza fra spike successivi nel nervo uditivo a fronte di un'eccitazione composta da due suoni puri

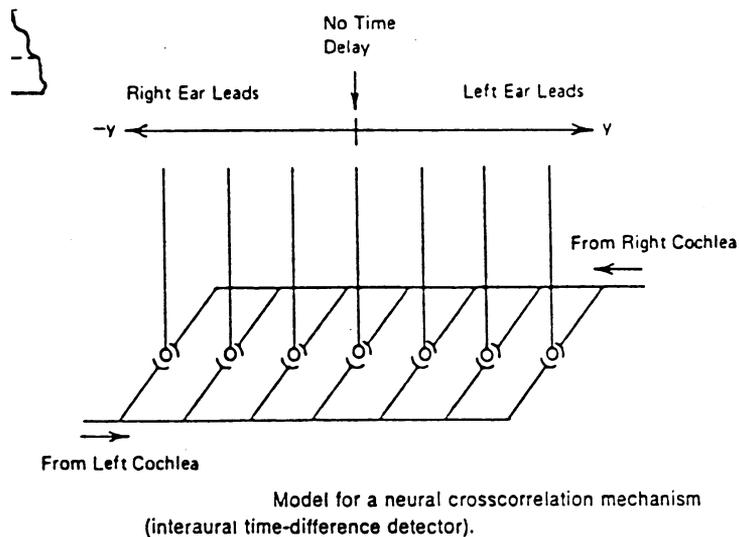


Figura 2.22: Modello di individuatore di differenze temporali interaurali

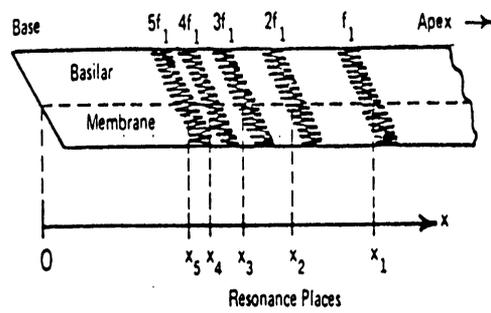


Figura 2.23: Pattern spaziale sulla membrana basilare dovuto a un suono armonici

2.11 Misure di intensità acustica

Si è visto, quando si sono descritti orecchio medio e orecchio interno, che l'ampiezza dell'oscillazione del timpano dovuta a un suono puro determina l'ampiezza dello scostamento della membrana basilare. Questa, a sua volta, determina la sensazione di intensità del suono. L'ampiezza dell'oscillazione del timpano è in stretta relazione con i parametri fisici di variazione di pressione Δp e di flusso di potenza acustica, o intensità, I . L'orecchio è sensibile a un range dinamico estremamente ampio ($10^{-12} \div 1 \frac{W}{m^2}$) e l'impraticità di una scala lineare ha portato all'uso della misura di livello di intensità (o IL), espresso in dB. Vi sono degli evidenti vantaggi legati all'uso questa scala: una scala logaritmica offre innanzitutto una notevole compressione dei valori; è una scala relativa, i valori sono riferiti a un valore di riferimento (soglia dell'udibile); l'unità della scala ha un valore che rappresenta approssimativamente la minima variazione percepibile di intensità, definita *jnd di intensità*. Questo valore è funzione della frequenza del suono puro, e dell'intensità di partenza. Si può comunque dire che il jnd di intensità assume valori massimi dell'ordine di 1.5 dB e valori minimi intorno ai 0.3 dB (figura 2.24).

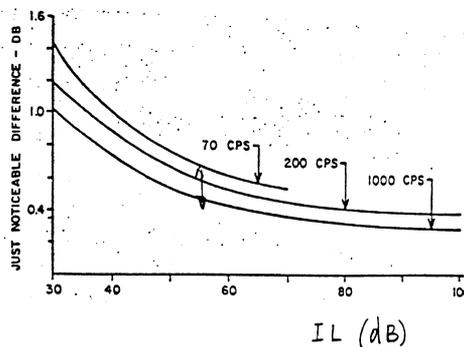
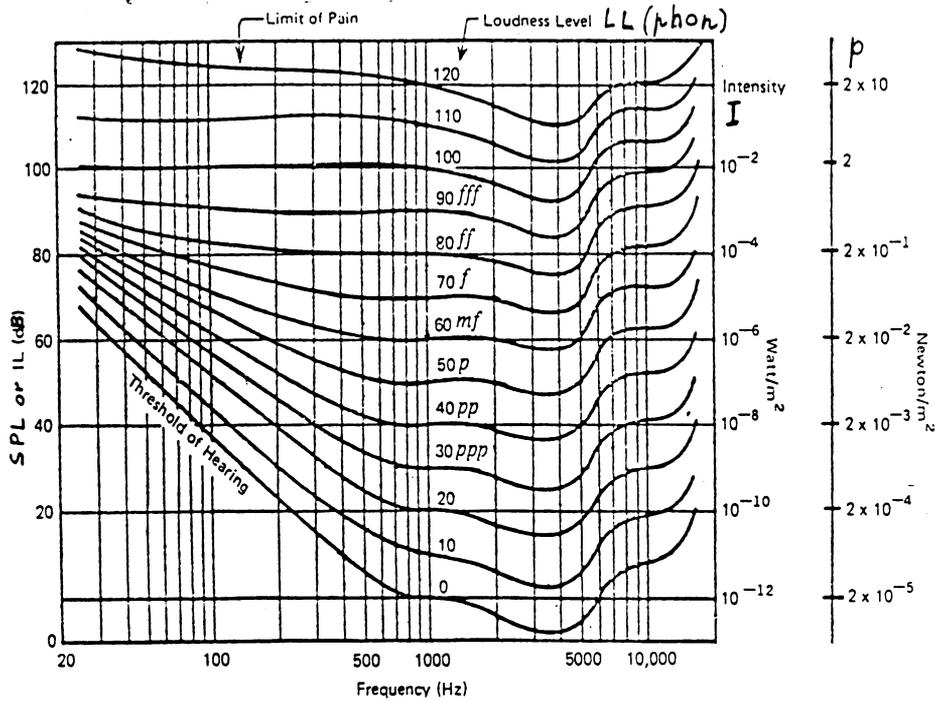


Figura 2.24: Valori di jnd di intensità per valori diversi di IL e frequenza del suono puro

Nel paragrafo 2.5.1 è stata introdotta una seconda scala logaritmica che definisce il livello di intensità acustica in funzione della pressione sonora (SPL). Sperimentalmente si osserva che suoni puri continui, caratterizzati da stesso SPL ma a frequenze diverse, producono sensazioni diverse di intensità. Questo indica che l'SPL non è una buona misura dell'intensità percepita se confrontiamo suoni puri a frequenza diversa. È stato dunque necessario trovare sperimentalmente i valori di eguale intensità percepita (*curve isofoniche* o *curves of equal loudness*) al variare della frequenza considerando come riferimento l'SPL a 1000 Hz. Il risultato è riassunto nel diagramma di figura 2.25, dovuto a Fletcher e Munson. Si noti, ad esempio, come un suono puro con SPL di 50 dB a 1000 Hz è considerato *piano* mentre è appena udibile a 60 Hz. In altre parole, per produrre la stessa sensazione di intensità alle basse frequenze è necessaria molta più energia di quella necessaria per produrre la stessa sensazione alla frequenza di riferimento di 1000 Hz. Si può notare ancora, osservando la curva che rappresenta la soglia dell'udibile, che la sensibilità dell'orecchio diminuisce notevolmente alle basse frequenze e alle frequenze alte al di sopra dei 6000 Hz.

In stretta relazione con le curve isofoniche è la definizione di *Phon* (o *Loudness Level, LL*): il LL di un suono a frequenza f è dato dal SPL di un suono a 1000 Hz che determina la stessa percezione di intensità. Dire dunque che un suono è a 80 Phon vuol dire fornire i valori della curva di isofonia relativa a 80 SPL per tutte le frequenze. Si noti che la scala dei Phon non è ancora una scala soggettiva (un raddoppio dei phon non determina un raddoppio dell'intensità percepita). Gli studi per la determinazione di una scala soggettiva hanno portato alla definizione della scala dei *Son* (o *subjective*



Curves of equal loudness (Fletcher and Munson 1933) in a sound

Figura 2.25: Curve di eguale intensità percepita (Fletcher and Munson, 1933)

loudness, L). In questa nuova scala la sonorità soggettiva raddoppia ogni 10 Phon. La relazione fra scala dei Phon e scala dei Son è illustrata in figura 2.26.

La legge che lega i Son e l'intensità del suono (o la variazione di pressione Δp) si esprime con la formula approssimata $L = C_1 \sqrt[3]{I} = C_2 \sqrt[3]{\Delta p^2}$, dove C_1 e C_2 sono parametri che dipendono dalla frequenza. E' possibile valutare questa relazione anche per suoni composti dalla sovrapposizione di più componenti sinusoidali. In questo caso è importante distinguere i seguenti casi: per suoni le cui frequenze cascano nella stessa banda critica, l'intensità percepita risultante è in relazione alla somma delle intensità individuali: $L = C_1 \sqrt[3]{I_1 + I_2 + I_3 + \dots}$. Quando le frequenze del suono complesso superano i limiti della stessa banda critica, la relazione diventa $L = C_1 \sqrt[3]{I_1} + C_2 \sqrt[3]{I_2} + C_3 \sqrt[3]{I_3} + \dots$. Quando, infine, le componenti sono molto diverse per frequenza e intensità l'ascoltatore tende a focalizzare l'attenzione su una sola componente (quella a frequenza più elevata o quella a intensità maggiore), attribuendo al suono complesso altezza e intensità di quella singola componente.

I valori di sonorità soggettiva visti fanno riferimento a suoni puri di durata superiore al mezzo secondo. Questa assunzione è necessaria perchè, per tempi inferiori, la durata del suono influenza la sonorità soggettiva: più breve è il suono, più basso risulta il livello percepito se l'intensità del suono è mantenuta costante. La figura 2.27 illustra la relazione al variare della durata fra l'effettiva sonorità soggettiva (r) e la sonorità soggettiva (L) di un suono stazionario di eguale frequenza e ampiezza.

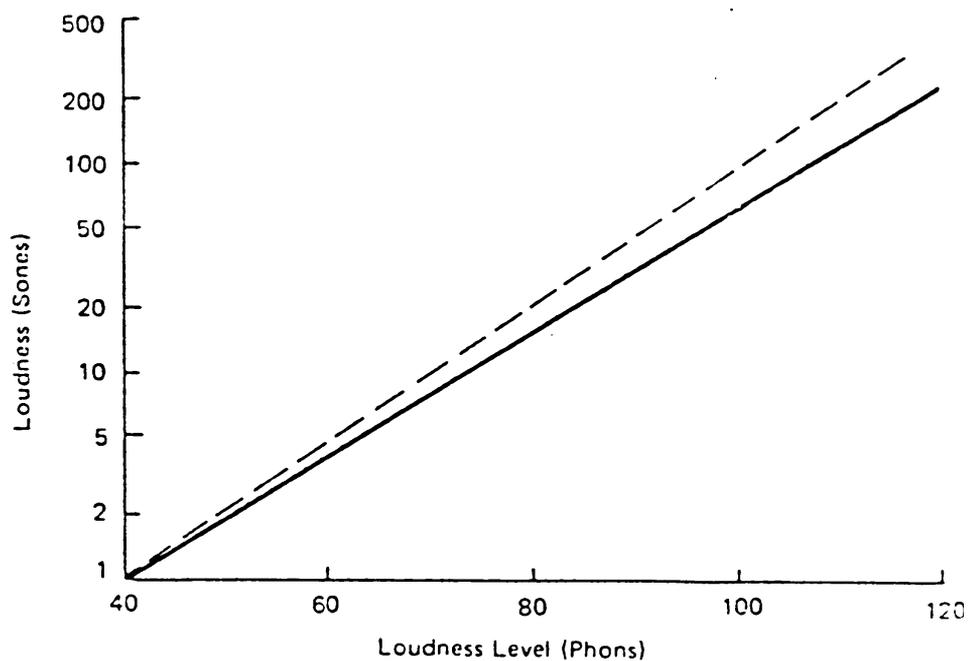
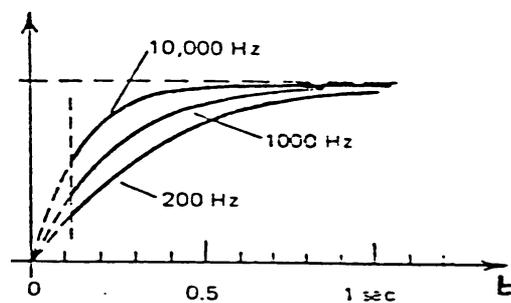
Figura 2.26: Confronto fra le scale dei *Phon* e quella dei *Son*

Figura 2.27: Influenza della durata sulla sonorità soggettiva

2.12 Mascheramento dei suoni

Ascoltando un suono composto da due suoni puri, non sempre l'ascoltatore percepisce le componenti distintamente. Quando, ad esempio, uno dei due suoni è caratterizzato da un'intensità molto maggiore dell'altro, quest'ultimo risulta inudibile o *mascherato*. Si definisce *livello di mascheramento, ML* l'intensità del suono mascherato alla soglia del mascheramento. In figura 2.28 sono illustrate le curve di mascheramento per diverse coppie di suoni puri. Si noti come in corrispondenza del punto in cui i due suoni assumono stessa frequenza (o l'uno frequenza doppia dell'altro), l'insorgere di battimenti del primo e del secondo ordine abbassa notevolmente la soglia. In alcuni grafici questo effetto è rimosso estrapolando i dati in corrispondenza dei punti critici (curve tratteggiate).

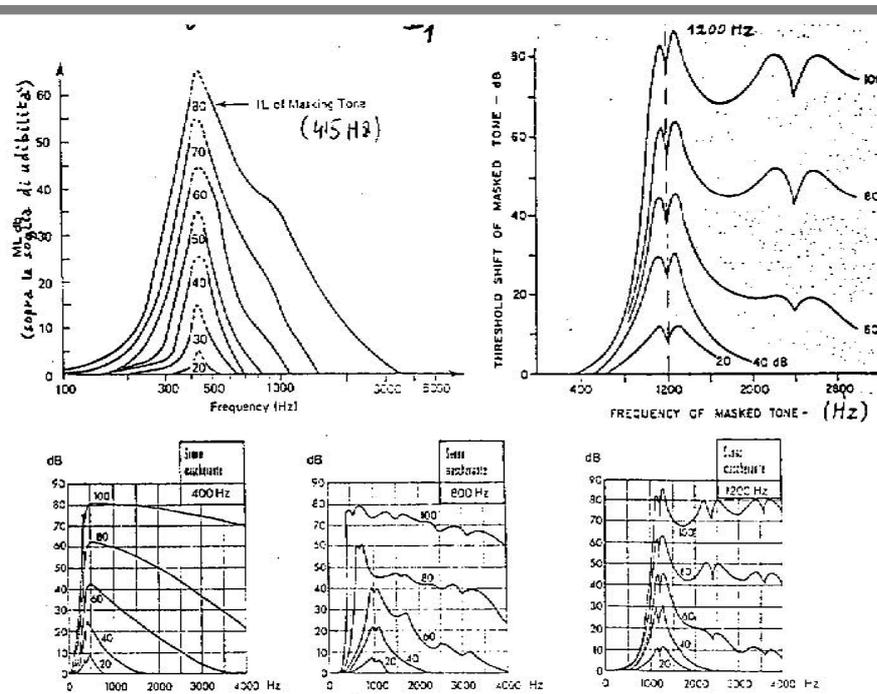


Figura 2.28: Curve del Livello di Mascheramento per coppie di suoni puri

Se I_1 è l'intensità del suono mascherante e I_{2m} è l'intensità del suono mascherato alla soglia del mascheramento, le relazioni seguenti definiscono il valore di *jnd* per il mascheramento:

$$\begin{aligned}
 I_t &= I_1 + I_{2m} = I_1 \left(1 + \frac{I_{2m}}{I_1} \right) \\
 &= I_1 \left(1 + \frac{I_{2m} I_0}{I_0 I_1} \right) = I_1 \left(1 + 10^{\frac{ML - IL_1}{10}} \right) \\
 jnd &= 10 \log \frac{I_t}{I_1} = 10 \log \left(1 + 10^{\frac{ML - IL_1}{10}} \right)
 \end{aligned}$$

dove I_t è l'intensità totale del suono risultante, I_0 è la soglia di udibilità e $ML = 10 \log \frac{I_{2m}}{I_0}$ è il livello di mascheramento.

2.13 La percezione del timbro

La parola *timbro* è usata per denotare la "qualità" o il "colore" del suono. Il timbro è quell'attributo che ci permette di giudicare diversi due suoni che abbiano stessa intensità, stessa altezza (e stessa durata). Il timbro è determinato in primo luogo dallo spettro di potenza dello stimolo, in secondo luogo dalla forma d'onda (fase), dall'intensità e dalle caratteristiche temporali (durata).

I principali parametri associati allo spettro di un suono sono l'altezza (frequenza fondamentale), l'intensità (integrale delle ampiezze delle parziali) e timbro (configurazione spettrale). Studi sperimentali hanno mostrato che il timbro è determinato dalla distribuzione della potenza acustica nelle bande critiche, non dai rapporti di intensità delle varie armoniche con la fondamentale. A questo scopo il range di frequenze udibili è stato diviso in 24 bande di circa un terzo di ottava (corrispondenti alle bande critiche)(Figura 2.29) ed è stato poi misurato il grado di variazione timbrica percepita in funzione del cambiamento di potenza distribuita nelle bande. Un esempio comune di deformazione timbrica dovuto a spostamento della potenza nelle bande critiche è l'effetto prodotto dall'accelerazione di un nastro audio.

Banda critica n.	1	2	3	4	5	6	7	8
Frequenza centrale in Hz	50	150	250	350	450	570	700	840
Larg. della banda in Hz	100	100	100	100	110	120	140	150
Banda critica n.	9	10	11	12	13	14	15	16
Frequenza centrale in Hz	1000	1170	1370	1600	1850	2150	2500	2900
Larg. della banda in Hz	160	190	210	240	280	320	380	450
Banda critica n.	17	18	19	20	21	22	23	24
Frequenza centrale in Hz	3400	4000	4800	5800	7000	8500	10500	13500
Larg. della banda in Hz	550	700	900	1100	1300	1800	2500	3500

Figura 2.29: Bande critiche

Il timbro di un suono non è tuttavia determinato esclusivamente dalla configurazione statica dello spettro. Qualsiasi suono reale è caratterizzato da una continua evoluzione temporale dello spettro. La dinamicità dello spettro è particolarmente accentuata nelle fasi di *transitorio* del suono (*attack*, *decay* e *release*), particolarmente importanti per il riconoscimento timbrico e per l'identificazione dello strumento. La fase sostenuta dei suoni strumentali è comunque caratterizzata da variazioni spettrali minori che conferiscono naturalezza al suono. La necessità di rappresentare l'evoluzione temporale dello spettro di un suono, giustifica l'uso delle rappresentazione *tempo-frequenza-intensità*, di cui un esempio è dato in figura 2.30.

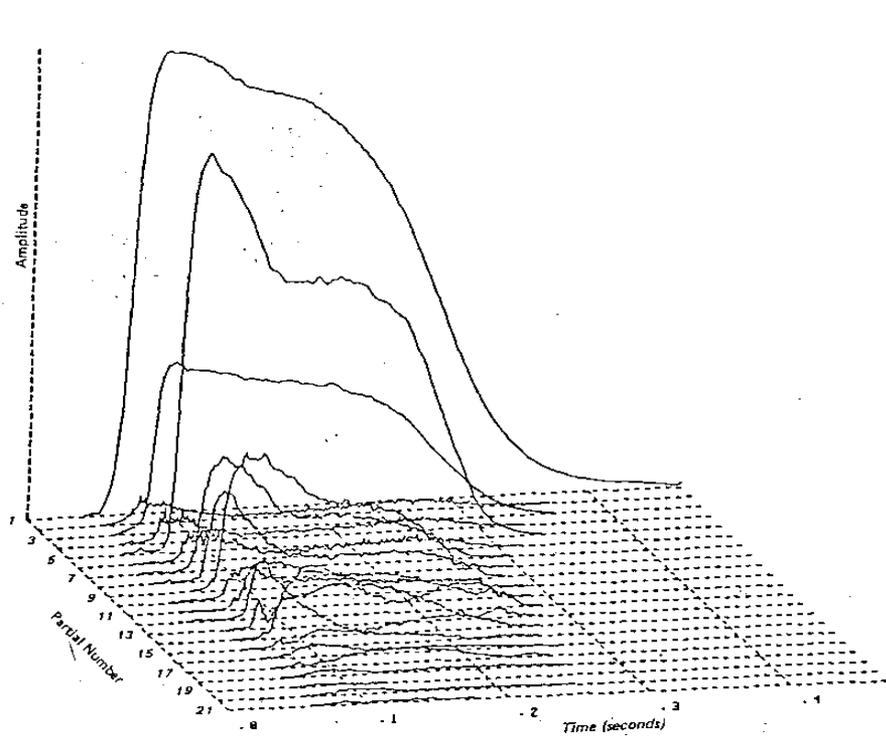


Figura 2.30: Rappresentazione tempo-frequenza di un suono

Bibliografia

[1] *J. G. Roederer, The Physics and Psychophysics of Music, Springer Verlag, 1980.*

[2] *J. R. Pierce, La Scienza del Suono, Zanichelli, 1988.*

NOTA: *La trattazione degli argomenti relativi alla Psicoacustica costituisce un ampliamento delle originarie dispense del corso relative a questo argomento, di cui è autore il Prof. G. B. Debiasi.*