

Il suono

Qualcuno pizzica una corda di una chitarra: la corda si sposta dalla sua posizione di riposo, si muove in una direzione, torna indietro, si sposta in direzione opposta, torna indietro. Ripete rapidamente questo movimento ciclico, è entrata in vibrazione. Ogni volta che si sposta in una direzione, esercita una pressione sullo strato di aria adiacente e lo comprime; quando torna indietro, con il ridursi della pressione si crea uno strato di aria più rarefatta (Figura 1.1).

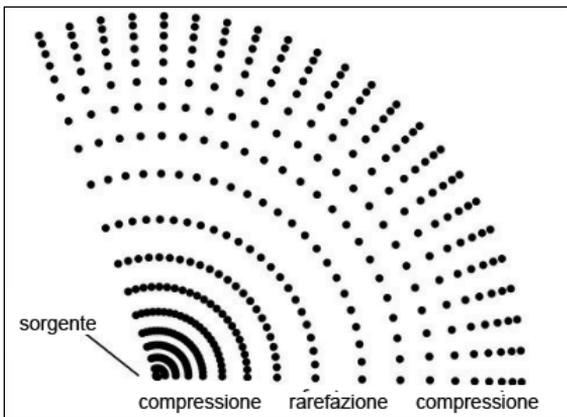


Figura 1.1 Una sorgente in vibrazione provoca la formazione di onde di compressione e rarefazione dell'aria circostante.

Nel frattempo, lo strato compresso è andato a esercitare una pressione sullo strato successivo e l'alternarsi di compressione e rarefazione si propaga come un'onda. Nell'aria, e a una temperatura costante di 21 °C, l'on-

In questo capitolo

- **Le caratteristiche dei suoni**
- **Onde e raggi**
- **Modi in ambienti chiusi**
- **Psicoacustica**
- **Da analogico a digitale (e ritorno)**

da si propaga velocemente: copre 344 metri al secondo. Raggiunge in una frazione di secondo una persona che si trova a pochi metri di distanza: l'alternarsi di compressione e rarefazione dell'aria giunge al suo padiglione auricolare, che funge un po' da imbuto e convoglia una parte di quelle vibrazioni dell'aria verso il canale uditivo dell'orecchio, una struttura all'incirca cilindrica. Le vibrazioni dell'aria in questo canale si comunicano alla membrana timpanica e da questa a un sistema di ossicini (detti incudine, martello e staffa) nell'orecchio medio (Figura 1.2).

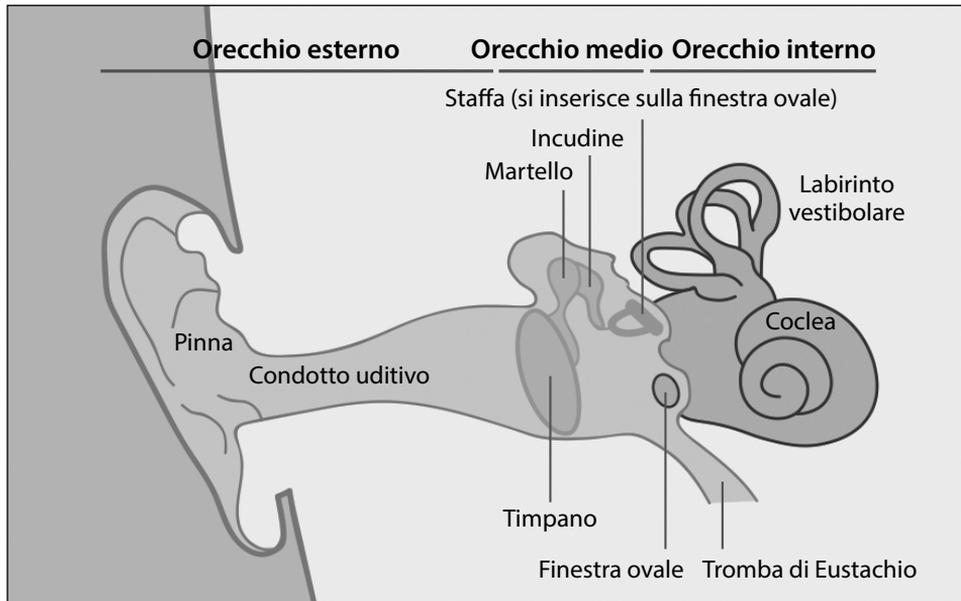


Figura 1.2 Anatomia dell'orecchio umano.

La staffa, l'ultimo ossicino della catena, è collegata a un'altra membrana, la finestra ovale, che trasmette le vibrazioni alla coclea. Quest'organo è una cavità piena di liquido, divisa in due dalla membrana basilare. Le vibrazioni dell'aria si comunicano al liquido e vengono raccolte dalle cellule ciliate, poste sulla membrana basilare, le quali effettuano la trasformazione delle vibrazioni meccaniche in segnali elettrici, i segnali neurali. Questi ultimi viaggiano verso il cervello, dove verranno interpretati – e il nostro ascoltatore potrà dire di aver sentito un suono prodotto dalla chitarra; se si tratta di un ascoltatore competente, potrà dire addirittura se si tratta di una nota *mi* o di un *do*.

Anche se sembra già complessa, questa è in realtà una descrizione molto semplificata di quello che succede realmente, quando un corpo elastico come una corda di chitarra entra in vibrazione e si innesca una serie di fenomeni fisici fino alla generazione di impulsi neurali che il cervello dell'ascoltatore interpreta come un suono.

Vediamo più da vicino tutta la catena e definiamo un po' di terminologia, che ci sarà utile nel seguito.

Le caratteristiche dei suoni

Il corpo che si mette in vibrazione (come la corda di chitarra nell'esempio, ma anche la colonna d'aria in uno strumento a fiato, o qualsiasi altra sorgente sonora) effettua una serie di oscillazioni, che si andranno smorzando nel tempo, poiché l'energia meccanica che gli è stata impartita dalla sollecitazione iniziale si disperde progressivamente per l'attrito con l'aria circostante. Se si riporta in un grafico lo spostamento della corda dalla sua posizione iniziale in funzione del tempo, si ottiene una forma come quella della Figura 1.3: la corda si allontana sempre di più dalla posizione iniziale in una direzione e il grafico sale, poi raggiunge il massimo dell'escursione e torna indietro (il grafico scende), supera la posizione di riposo (il grafico intercutta l'asse delle x), si allontana nella direzione opposta (il grafico scende sotto l'asse delle x), raggiunge la massima escursione in quella direzione e torna indietro verso la posizione di riposo (il grafico risale e intercutta di nuovo l'asse delle x). Quindi il ciclo ricomincia e il grafico si riproduce con un'identica forma.

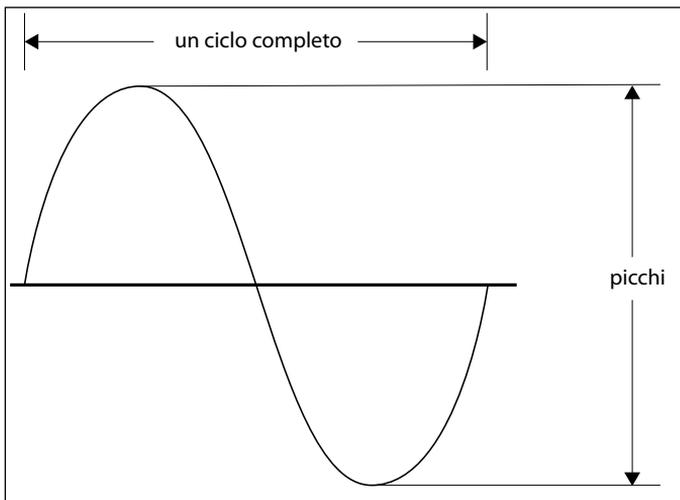


Figura 1.3 La rappresentazione grafica del moto ondulatorio.

Frequenza

Un ciclo completo (il tratto di grafico che corrisponde a un'oscillazione completa) si chiama *periodo* dell'oscillazione e la sua durata è costante (per quella onda sonora). Il numero di periodi nell'unità di tempo (un secondo) è la *frequenza* dell'oscillazione, che si misura in hertz (Hz in sigla). Una frequenza di 1 Hz significa una oscillazione completa in un secondo.

L'orecchio umano non percepisce suoni della frequenza di 1 Hz: la *soglia di udibilità inferiore* si colloca intorno ai 20 Hz (20 oscillazioni al secondo). Non percepisce neanche suoni di frequenza molto elevata: la *soglia di udibilità superiore* si colloca intorno ai 20.000 Hz. Diciamo "intorno" perché le capacità percettive sono una caratteristica personale: c'è chi percepisce qualche frequenza in più, chi ne percepisce qualcuna in meno; con l'età la capacità percettiva tende a ridursi, e lo stesso accade in condizioni di affaticamento

o di cattiva salute (si sente meno bene quando si ha il raffreddore, per esempio). Sotto i 20 Hz si parla di *infrasuoni*, sopra i 20.000 di *ultrasuoni*.

Lunghezza d'onda

Poiché nell'aria la velocità delle onde sonore è costante, esiste una relazione costante fra frequenza delle oscillazioni e lunghezza dell'onda: se in un secondo le onde coprono 344 metri e la frequenza è di 100 Hz, la lunghezza di ciascuna onda sarà di 3,44 metri. Di solito si indica con c la velocità del suono, con f la frequenza e con l la lunghezza d'onda: le tre grandezze sono legate dalla relazione $l = c/f$ (con l espressa in metri e f in hertz). La frequenza è legata a quella che noi percepiamo come *altezza* del suono: quanto più elevata la frequenza, tanto più acuto il suono che percepiamo.

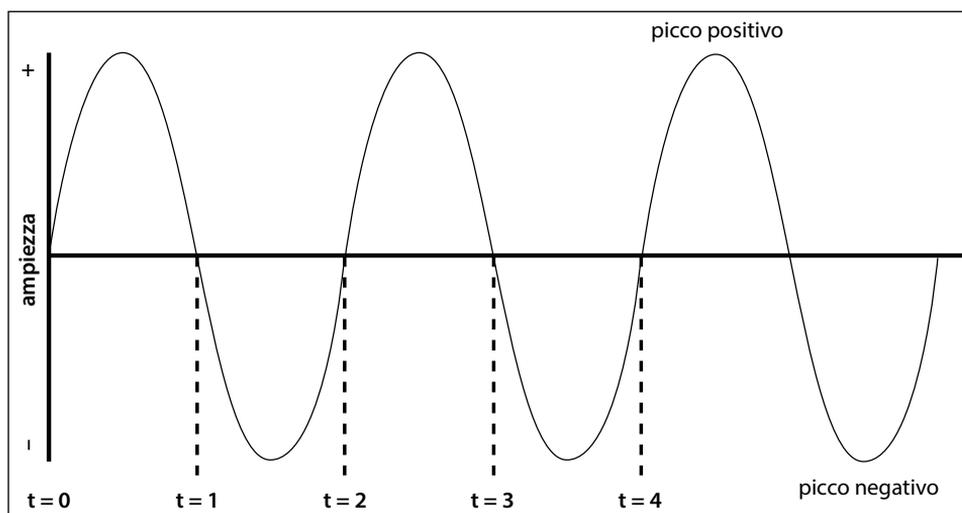


Figura 1.4 Andamento di un'onda nel tempo.

Ampiezza

Se riguardiamo il grafico, c'è un'altra caratteristica interessante delle onde sonore: la loro *ampiezza*, cioè la distanza fra il punto di massimo e l'asse delle x , che indica in che misura la nostra corda, una volta pizzicata, si allontana dalla posizione di riposo. L'ampiezza è legata a quella che noi percepiamo come *intensità* (o volume) del suono: maggiore l'ampiezza, più forte risulta per noi il suono. L'ampiezza delle onde non è costante nel tempo: le onde si smorzano progressivamente. L'andamento dell'ampiezza (o *inviluppo*) è una caratteristica peculiare di ciascun suono: nel suo grafico si distinguono di solito quattro fasi: una iniziale o attacco (*attack*), in cui l'ampiezza aumenta dallo zero a un massimo, un primo decadimento (*decay*), una fase stazionaria o di regime (*sustain*) in cui l'ampiezza rimane relativamente costante e una fase finale di estinzione (*release*). Un suono di clarinetto, per esempio, ha un attacco lento e una fase di sustain abbastanza lunga, con un'estinzione lenta; un suono di rullante di batteria, invece, ha un attacco molto rapido e un rapido decadimento.

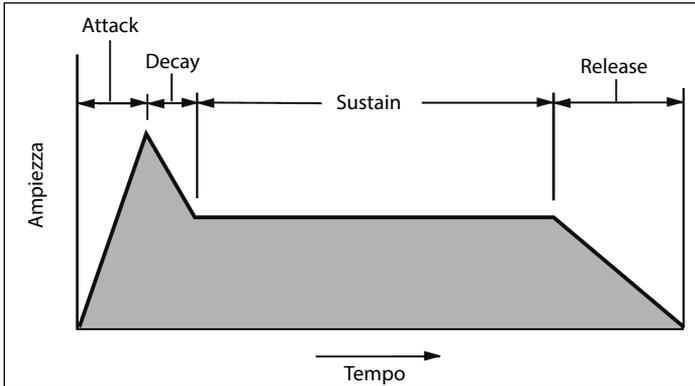


Figura 1.5 Rappresentazione delle quattro fasi dell'inviluppo di un suono o ADSR: Attack (attacco), Decay (primo decadimento), Sustain (fase di regime) e Release (fase finale di estinzione).

Fondamentale e armoniche

Fin qui abbiamo dato per scontato che la corda pizzicata producesse un'unica frequenza, ma il caso di un suono puro (frequenza unica), come quello rappresentato dal grafico della Figura 1.3, è molto raro; in realtà un'onda sinusoidale pura come quella del grafico si riesce a produrre solo con un generatore elettronico. La corda di una chitarra non oscilla solo nella sua interezza; oscillano anche le sue due metà, i suoi tre terzi e così via. Ciascuna di queste parti oscilla a una frequenza (detta *parziale*) che è un multiplo della frequenza *fondamentale* (quella della corda nella sua interezza), e il movimento risultante è la composizione dei moti delle singole parti vibranti. Solitamente quando le parziali hanno una frequenza che è un multiplo intero della fondamentale vengono chiamate *armoniche*: la distinzione ha un senso, perché esistono anche suoni con parziali la cui frequenza non è un multiplo della fondamentale.

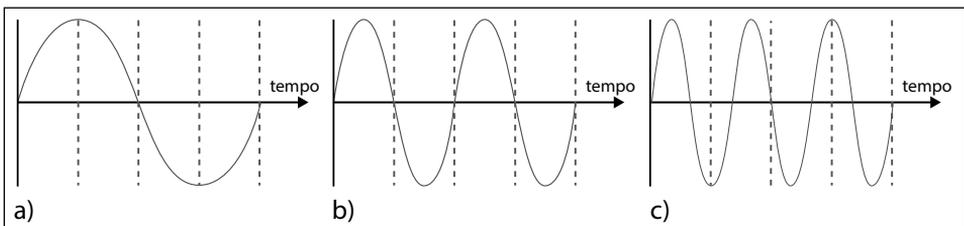


Figura 1.6 Un'onda fondamentale e le sue due prime parziali, di frequenza rispettivamente doppia e tripla.

Il risultato è un'onda dall'andamento molto più complesso e la sua forma dipende dalle caratteristiche della sorgente sonora: nella Figura 1.6 le due parziali hanno la stessa ampiezza della fondamentale, ma in genere non è così, le parziali hanno ampiezze diverse dalla fondamentale e l'una dall'altra. Quando ascoltiamo una nota emessa da un violino,

quella che arriva al nostro orecchio è un'onda complessa, costituita dalla somma di una fondamentale e di una serie di parziali di varia ampiezza; quando sentiamo una nota della stessa frequenza emessa da una tromba, l'onda che ci arriva ha una composizione differente e di conseguenza le due forme d'onda sono diverse (Figura 1.7). Questa forma complessa è correlata a quello che percepiamo come *timbro* o *qualità* del suono e che ci permette di distinguere un violino da una tromba, anche se emettono la stessa nota (ed è quello che ci permette di distinguere la voce di due persone diverse, anche se pronunciano le medesime parole).

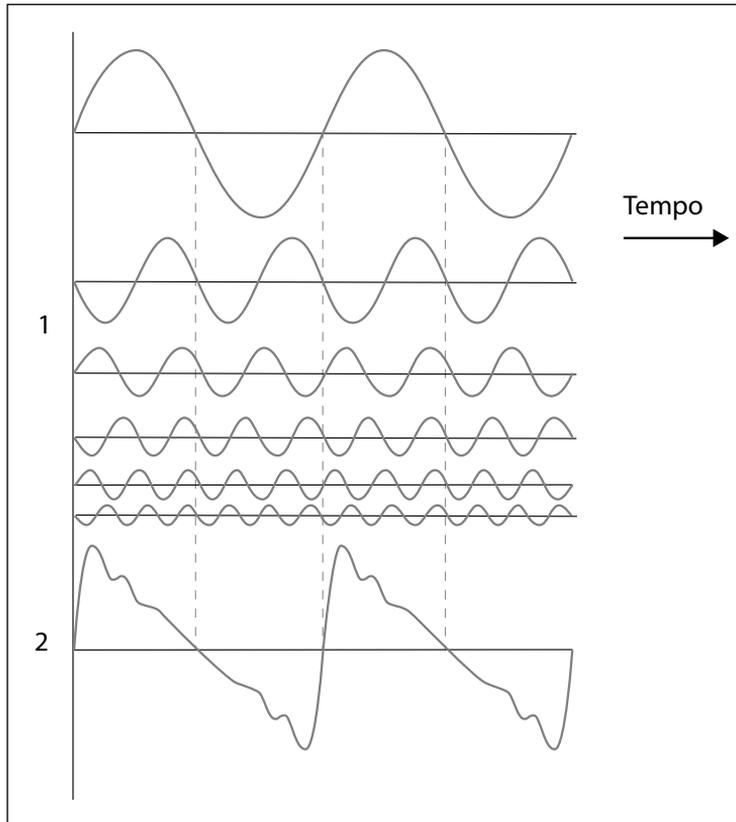


Figura 1.7 Un'onda complessa (2) e le parziali che la compongono (1).

Per quanto complessa possa essere una forma d'onda, la si può sempre scomporre nella somma di una serie di onde sinusoidali semplici: dal punto di vista matematico, lo strumento che permette di risalire dalla forma complessa alle onde componenti prende il nome di *analisi di Fourier*.

Una rappresentazione efficace della complessità di un suono è quella che ne dà lo *spettro* (Figura 1.8), ovvero la distribuzione della sua energia in funzione della frequenza, visualizzando con barre verticali di altezza proporzionale all'ampiezza (e quindi all'energia) il contributo della fondamentale e delle armoniche (Figura 1.8).

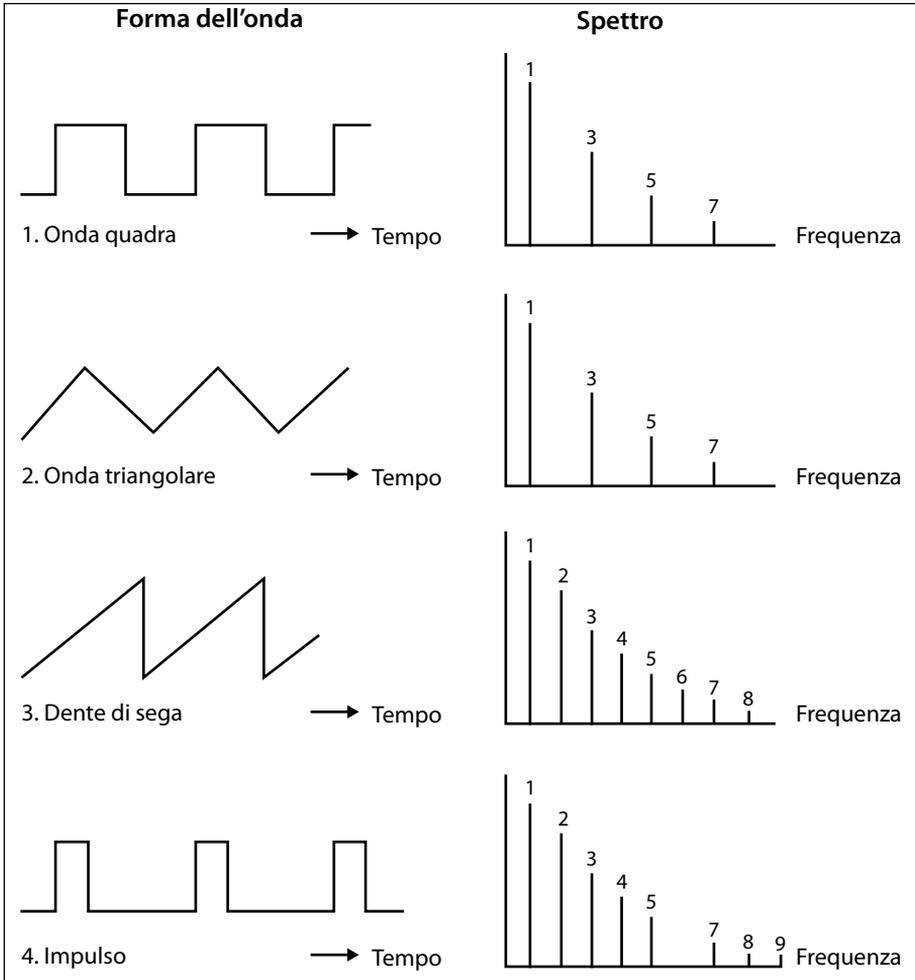


Figura 1.8 Alcune forme d'onda particolari e la loro rappresentazione spettrale (a destra).

Pressione sonora e decibel

Le onde sonore sono onde di *pressione* e in fisica la pressione si misura in pascal (simbolo Pa). Entro il campo di udibilità dell'orecchio umano, la pressione delle onde acustiche varia moltissimo: si va da pochi micropascal (μPa), cioè dai milionesimi di pascal, alle centinaia di pascal (intorno ai 100 pascal si colloca la soglia del dolore). Su una scala lineare, valori tanto diversi non sono facilmente rappresentabili e non risultano comunque molto intuitivi. Per questo si ricorre normalmente a una scala logaritmica: si considerano i logaritmi dei valori, anziché i valori stessi.

Il logaritmo di un valore x è l'esponente che bisogna assegnare a una base (qui si usa di solito la base 10) per ottenere quel valore. Così, per esempio, $1000 = 10^3$ e il logaritmo di 1000 in base 10 è 3, in formula: $\log_{10} 1000 = \log_{10} 10^3 = 3$. Con l'uso dei logaritmi, si ottiene una rappresentazione più "compatta", ma anche il vantaggio di trasformare i

prodotti in somme: $1000 \times 1000 = 10^3 \times 10^3 = 1000000 = 10^6$. Con i logaritmi: $\log_{10}(1000 \times 1000) = \log_{10}(10^3) + \log_{10}(10^3) = 3 + 3 = 6$. I calcoli con le potenze di dieci sono ovviamente molto facili, se si usa il 10 come base dei logaritmi; per tutti gli altri valori si ricorre per comodità alle tavole logaritmiche.

Per comodità, nel mondo dell'audio si usa come misura il *livello di pressione sonora* L_p (SPL o *Sound Pressure Level* in inglese), che è il rapporto fra la pressione sonora che interessa e una pressione sonora di riferimento, che si assume pari a $20 \mu\text{Pa}$, all'incirca la pressione che corrisponde alla soglia di udibilità per un suono a 1 kHz; il rapporto è espresso in forma logaritmica e l'unità di misura è il decibel (dB). In formula:

$$\text{dB} = 10 \log_{10}(p/p_0)^2 = 20 \log_{10}(p/p_0)$$

Utilizzando i decibel, la soglia di udibilità corrisponde a 0 dB; il livello di una comune conversazione a circa 30 cm di distanza è di 70 dB; la soglia del dolore si colloca intorno ai 125-130 dB. La differenza di un dB può essere considerata la differenza minima di intensità percepibile: in altre parole, percepiamo di "volume" differente due suoni se differiscono per almeno 1 dB. (Si tratta di un valore "medio": c'è chi percepisce anche variazioni minori e chi invece non percepisce che differenze maggiori; il valore dipende anche dalla frequenza del suono e da altri fattori.) Un aumento di 6 dB viene percepito come un raddoppio del volume.

Si usa il decibel (e quindi i logaritmi) anche per la misura di altre grandezze e per distinguere i vari usi si ricorre a indici sottoscritti: dB_{SPL} è l'unità riferita al livello di pressione sonora. Quando ci si riferisce invece al livello del segnale (potenza misurata in Watt), la formula precedente cambia un po' e diventa:

$$\text{dB} = 10 \log P/P_{\text{ref}}$$

dove P è la potenza e P_{ref} è una potenza di riferimento misurata in Watt. Di solito si indica come dBm il decibel rispetto alla potenza di riferimento di 1 milliwatt.

Quando si usa il decibel per rapporti fra tensioni (voltaggi), si torna alla formula precedente ($20 \log V/V_{\text{ref}}$) dove V è la tensione e la tensione di riferimento è 0,775 volt (nel qual caso si usa la forma dBu o dBv), oppure 1 volt (e allora si scrive dBV).

Onde e raggi

Le onde sonore non sono visibili; per averne un modello visuale, si può pensare alle onde che si producono in uno stagno quando si getta nell'acqua un sasso. Dal punto in cui cade il sasso, le onde si propagano circolarmente, allontanandosi dal punto di impatto. Quelle che si propagano sono solo le onde: l'acqua non si muove, se non verticalmente. Lo stesso accade con le onde sonore: gli spostamenti dell'aria sono solo locali e le onde si propagano circolarmente.

Interferenza e fase

Che cosa succede se si lanciano due sassi nell'acqua, in posizioni diverse? Da ciascun punto di impatto si propagano delle onde, che finiranno per incontrarsi: si ha un'*interferenza*,

se si incontrano due creste, il risultato è un'onda più alta; se si incontrano una cresta e un ventre, le due onde si annullano. In casi intermedi, anche il risultato è intermedio. Quando due onde raggiungono negli stessi istanti i loro punti di massimo e di minimo, si dicono *in fase* (o *in coincidenza di fase*) quando l'una raggiunge un punto di massimo mentre l'altra raggiunge un punto di minimo, sono *in opposizione di fase*. Le onde che si incontrano si sommano, si può dire, algebricamente: due onde in fase danno come risultato un'onda di maggiore ampiezza; due onde in opposizione di fase si cancellano. (Figura 1.9). Quando le due onde sono fuori fase (in tutti gli altri casi) il risultato dipenderà dalle fasi del ciclo in cui si incontrano e può essere un'amplificazione o un'attenuazione.

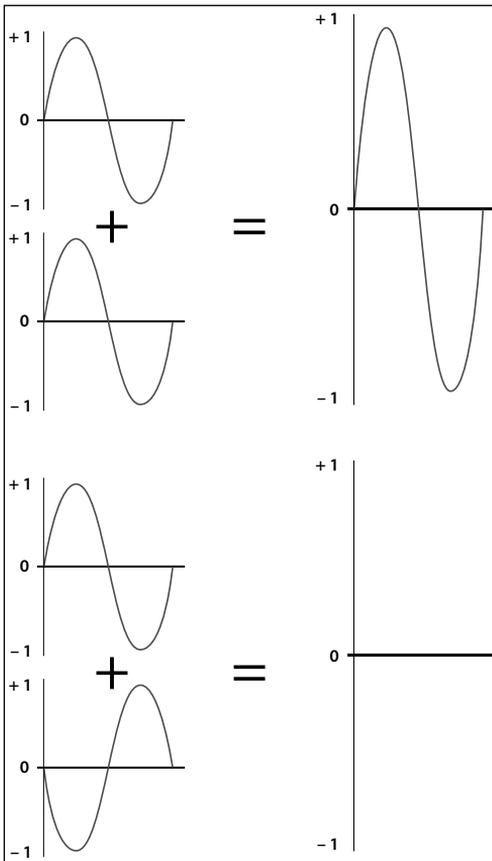
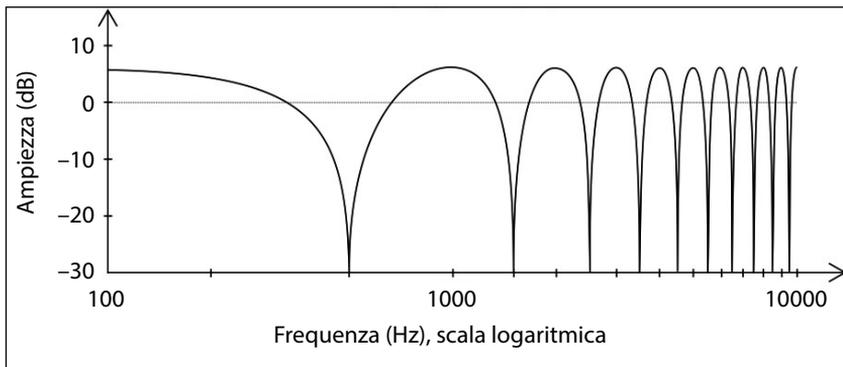


Figura 1.9 Due onde che si incontrano, se sono in coincidenza di fase (in alto) si sommano; se sono in opposizione di fase (in basso) si cancellano a vicenda. È il fenomeno dell'interferenza, costruttiva nel primo caso, distruttiva nel secondo.

In campo libero, senza alcun ostacolo, le onde sonore si propagano circolarmente fino a che l'energia sonora non si annulla. In presenza di ostacoli o in un ambiente chiuso (che è il caso che ci interessa in modo particolare), si verificano invece fenomeni che si intuiscono meglio adottando la rappresentazione per raggi: un raggio indica la direzione in cui l'onda sonora si propaga.

Filtraggio a pettine

Quando si incontrano due onde sonore, come abbiamo visto, le relazioni di fase possono produrre, a seconda dei casi, amplificazioni o cancellazioni, totali o parziali. Quando le due onde sono complesse, le relazioni di fase delle loro varie componenti sono in genere molto varie: mentre alcune componenti verranno amplificate, altre subiranno cancellazioni (al solito, totali o parziali). Il risultato non sarà una semplice amplificazione o cancellazione delle onde complesse, ma un cambiamento timbrico. Un caso particolarmente importante nel campo della registrazione è quello di due onde identiche, una delle quali viene ritardata rispetto all'altra, cosa che può essere provocata intenzionalmente applicando un effetto di ritardo, oppure avvenire naturalmente perché lo stesso segnale arriva a due microfoni, posti a una certa distanza l'uno dall'altro, in tempi leggermente diversi, semplicemente perché deve percorrere un tratto più lungo per arrivare a uno dei due, rispetto all'altro. Con il ritardo, alcune frequenze componenti restano in fase, altre finiscono fuori fase, altre ancora in opposizione di fase. La distribuzione dell'energia nello spettro del segnale risultante assume una forma particolare, che ricorda un po' i denti di un pettine:



L'effetto prende il nome di *filtro a pettine* o *effetto pettine* e il risultato è una alterazione del timbro, che influisce sulla qualità della registrazione o dell'ascolto.

Assorbimento e riflessione

Quando l'onda sonora incontra una parete, in parte la sua energia viene assorbita dal materiale di cui è costituita la parete, in parte viene invece riflessa e, come nel caso dei raggi di luce, l'angolo di incidenza e l'angolo di riflessione sono uguali ma opposti (Figure 1.10 e 1.11).

La capacità di assorbimento della parete dipende dal materiale: se il materiale avesse un assorbimento nullo, avremmo una situazione analoga a quello che è uno specchio per la luce, cioè l'onda verrebbe totalmente riflessa.

Nella maggior parte dei casi l'assorbimento da parte del materiale non è nullo, ma non è nemmeno totale (caso di un assorbitore perfetto, come potrebbe essere una finestra aperta, che lascia fuoriuscire completamente il suono) e l'onda riflessa perciò è in qualche misura attenuata rispetto a quella incidente.

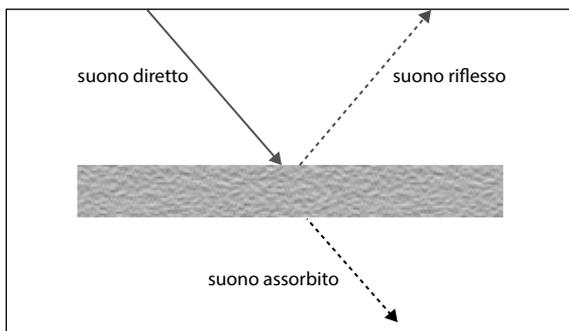


Figura 1.10 Quando un suono (qui rappresentato come un raggio) incontra un ostacolo di grandi dimensioni, come una parete, viene riflesso, ma una parte della sua energia può essere assorbita dall'ostacolo, a seconda del tipo di materiale di cui è costituito. L'angolo di incidenza e quello di riflessione sono identici. Il comportamento del suono è analogo a quello della luce.

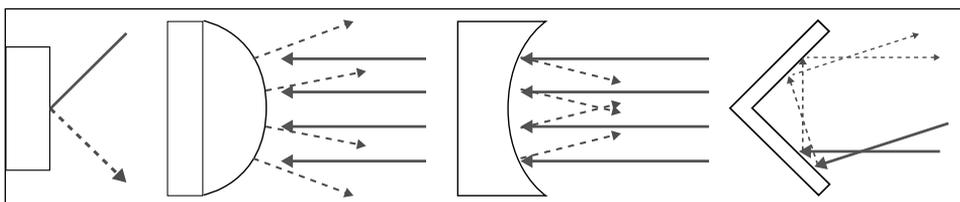


Figura 1.11 A seconda della forma delle pareti, i suoni riflessi generano situazioni diverse. Nel caso di una superficie convessa, i suoni riflessi vengono dispersi; nel caso di una superficie concava, vengono invece concentrati. In prossimità degli angoli, il percorso delle riflessioni diventa più complicato.

In una stanza chiusa, le riflessioni sono molteplici: dalla parete di fronte alla sorgente sonora e da quella alle sue spalle, dalle pareti laterali, dal soffitto e dal pavimento. Poiché i tragitti che le onde percorrono prima di raggiungere le varie pareti sono di lunghezza diversa, le onde riflesse si incontrano fra loro e con le onde dirette in genere in condizioni di fase diverse: in qualche punto, dove saranno in fase, si amplificheranno, in altri si attenueranno, in altri ancora, dove si incontrano in opposizione di fase, si annulleranno completamente. I comportamenti varieranno in dipendenza della lunghezza d'onda (e quindi della frequenza): le lunghezze d'onda maggiori, che corrispondono alle frequenze più basse e quindi alle note più profonde, creano più problemi.

Lo si nota in particolare nel caso delle onde stazionarie, che si possono formare quando la lunghezza delle onde è pari alla distanza fra pareti opposte, o è un multiplo o sottomultiplo di quella distanza (Figura 1.12). Le onde che arrivano perpendicolari alle pareti, in tali casi, vengono riflesse nella stessa direzione (ma in senso opposto) e formano *onde stazionarie*, che hanno i *ventri* (punti di massima pressione) e i *nodi* (punti di pressione nulla) in posizioni che rimangono fisse. Così, chi si trova all'interno di un ventre sente quella frequenza amplificata, chi si trova in un nodo non la percepisce affatto. Se la frequenza è bassa, i nodi sono distanziati e il fenomeno è ben avvertibile; se la frequenza è alta i nodi sono ravvicinati, la distribuzione dell'energia più regolare e le differenze meno apprezzabili.

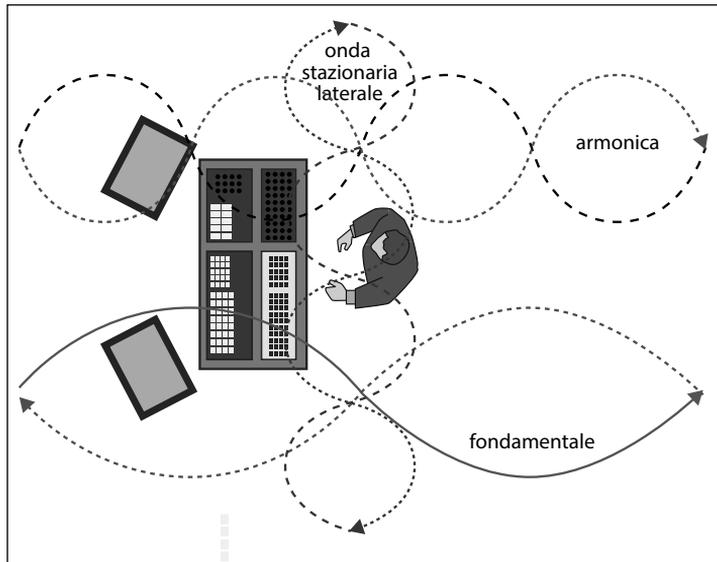


Figura 1.12 Onde stazionarie in una sala di regia non trattata.

Un'onda stazionaria genera una *risonanza*: capita, in sale amatoriali non ben trattate, che qualche nota, tipicamente del basso, si senta molto più forte delle altre, al punto da risultare fastidiosa. Il fenomeno è più accentuato in un ambiente cubico e in quelli in cui una dimensione è un multiplo dell'altra; è meno evidente in spazi grandi. Ove il problema si presenti, bisogna correggerlo con opportuni accorgimenti (come vedremo nei Capitoli 2 e 3).

Rifrazione

Quando incontrano una parete, abbiamo detto, le onde sonore cedono in genere parte della loro energia; quanta energia venga assorbita dipende dal tipo di materiale che incontrano: i *coefficienti di assorbimento* variano da materiale a materiale, ma anche in funzione della frequenza del suono. Il coefficiente di assorbimento del normale vetro di una finestra, per esempio, è molto più elevato alle basse frequenze che alle alte; un tendaggio invece assorbe le alte frequenze più delle basse.

Nel passaggio da un mezzo a un altro, le onde sonore, come la luce, subiscono una *rifrazione*: la loro direzione di propagazione viene modificata, di un angolo che dipende dal materiale. Il fenomeno si manifesta anche all'interno di uno stesso mezzo, fra strati con caratteristiche diverse: fra strati d'aria con temperature diverse, per esempio. Il cambiamento di direzione può essere immediato, nel passaggio da un mezzo all'altro, oppure graduale, durante la propagazione nel nuovo mezzo: così, per esempio, nel passaggio da uno strato d'aria più freddo a uno più caldo la direzione di propagazione di un'onda sonora può variare in modo più continuo, magari creando un arco con effetti curiosi a distanza dalla sorgente.

Certi materiali, assorbendo energia dalle onde sonore incidenti, possono entrare a loro volta in vibrazione in misura significativa: il vetro di una finestra entra in vibrazione

quando viene colpito dalle onde sonore, ad alcune frequenze anche in misura fastidiosa, e le vibrazioni si propagano da entrambe le facce del vetro. Per questo la finestra che permette la visibilità fra una stanza di regia e una sala di registrazione richiede sempre una progettazione attenta, in modo da eliminare il più possibile questo tipo di fenomeno, che influisce negativamente sia sulla registrazione che sull'ascolto.

Diffrazione

Quando le onde sonore incontrano sul loro cammino un ostacolo, per esempio un pilastro in mezzo a una stanza, un mobile, una persona, lo aggirano: il fenomeno prende il nome di *diffrazione* (Figura 1.13).

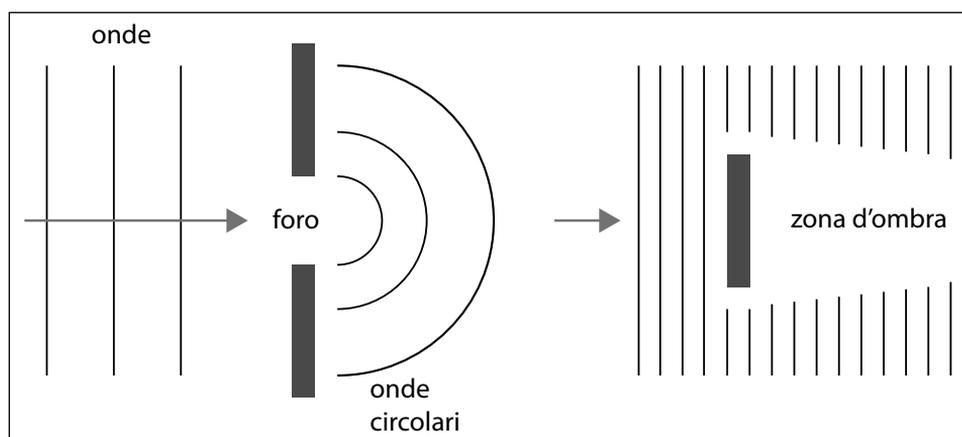


Figura 1.13 Il fenomeno della diffrazione: le onde si propagano oltre gli ostacoli.

Se l'ostacolo è di piccole dimensioni rispetto alla lunghezza dell'onda sonora, questa si propaga immediatamente oltre l'oggetto; se l'ostacolo è invece di dimensioni relativamente grandi rispetto alla lunghezza dell'onda, dietro l'ostacolo si forma una "zona d'ombra" in cui le onde dirette non arrivano, ma arrivano le onde diffratte dagli spigoli, che agiscono come nuove sorgenti sonore; lo stesso fenomeno si verifica quando un'onda incontra un ostacolo con un'apertura di piccole dimensioni.

Riverbero

In uno spazio chiuso come una sala di registrazione, dunque, in presenza di vari oggetti e persone, in genere con più sorgenti sonore, si produce un gran numero di onde dirette, riflesse dalle pareti, diffratte dagli ostacoli, con comportamenti difficilmente prevedibili sulla carta. Quando viene prodotto un suono, un ascoltatore percepirà non solo il suono diretto, ma anche le onde riflesse e diffratte. Poiché le onde riflesse (magari non solo una, ma due o più volte) percorrono un tragitto più lungo, arrivano alle orecchie dell'ascoltatore con un certo ritardo rispetto al suono diretto. Se il ritardo è piccolo, il suono riflesso non viene percepito come separato, ma contribuisce a modificare il suono

diretto, possibilmente arricchendolo (è il motivo per cui ci sembra di avere una voce particolarmente bella, quando cantiamo dentro la cabina della doccia) e in generale contribuendo a dare una sensazione di spazio.

Quanto sia significativo questo fenomeno lo si può apprezzare con un esperimento semplicissimo: basta mettersi a cantare liberamente e poi prendere in mano un libro aperto e metterlo davanti alla bocca, variando la distanza mentre si continua a cantare: il libro è una superficie riflettente e la variazione del suono della voce è immediatamente percepibile.

Oltre un certo punto, se il ritardo diventa apprezzabile, la riflessione diventa un disturbo (raramente è un effetto voluto sotto il profilo artistico), fino a diventare una eco quando le distanze sono sufficientemente grandi, nel qual caso il suono diretto e quello di ritorno vengono percepiti come nettamente distinti.

Se si produce in una stanza un suono, per effetto delle riflessioni non cesserà immediatamente quando cessa l'emissione dalla sorgente: l'effetto è quello che si definisce *riverbero*. Ambienti chiusi diversi, in funzione delle dimensioni, dei materiali delle pareti e dei rivestimenti, e di altri fattori ancora, hanno *tempi di riverbero* differenti: con questo termine si indica il tempo che impiega il riverbero per decadere di 60 dB (in sigla RT60). Una grande chiesa di costruzione tradizionale ha in genere tempi di riverbero elevati, anche intorno ai 3 secondi: è ciò che in molti casi contribuisce a rendere affascinante l'esecuzione di un coro e il suono di un organo a canne. Lo stesso tempo di riverbero in una sala di registrazione per la musica pop risulterebbe invece fastidioso: il suono si impasterebbe e sarebbe difficile distinguere le varie frequenze. In una sala per musica pop/rock in genere si preferiscono tempi di riverbero intorno ai 4 decimi di secondo (o meno); per la musica classica si preferiscono tempi di riverbero più elevati, oltre il secondo.

A un ascoltatore posto di fronte a una sorgente sonora in una stanza chiusa giunge innanzitutto il *suono diretto*, quello che viaggia direttamente dalla sorgente; poi le *prime riflessioni*, che sono le prime onde riflesse dalle pareti della stanza; infine il *riverbero*, l'insieme delle riflessioni successive, dall'andamento sempre più casuale, che sono il risultato dei progressivi rimbalzi da una parete all'altra, con tutte le interferenze che si vengono a produrre. Le prime riflessioni sono quelle che danno all'orecchio dell'ascoltatore un'indicazione sulle dimensioni dello spazio in cui si trova; il riverbero successivo è costituito da così tante onde riflesse a breve distanza l'una dall'altra da essere percepito come un segnale unico, che va progressivamente estinguendosi (Figura 1.14).

Il tempo di riverbero dipende dal coefficiente di assorbimento delle superfici che delimitano lo spazio della stanza e dal volume dello spazio: al crescere del volume, il tempo di riverbero aumenta, mentre diminuisce al crescere del coefficiente medio di assorbimento delle pareti. Se si devono registrare strumenti diversi o formazioni strumentali diverse, la qualità delle registrazioni trarrebbe beneficio da tempi di riverbero differenti: per questo si possono utilizzare volta a volta accorgimenti come l'inserimento di tappeti o di pannelli mobili, che contribuiscono a variare l'assorbimento o a modificare i percorsi delle riflessioni.

Gli studi professionali in cui è frequente la necessità di registrare strumenti differenti realizzano spazi separati per gli strumenti che richiedono maggiore attenzione o che possono contribuire negativamente alla qualità della registrazione: per le registrazioni di musica pop/rock, per esempio, alla batteria può essere dedicato uno spazio separato (con una finestra per la visibilità da parte degli altri membri del gruppo musicale); lo stesso

può essere fatto anche per la voce. Queste soluzioni permettono allo stesso tempo di mantenere ben separati gli strumenti e di arricchire il suono di ciascuno con condizioni di riverbero differenti.

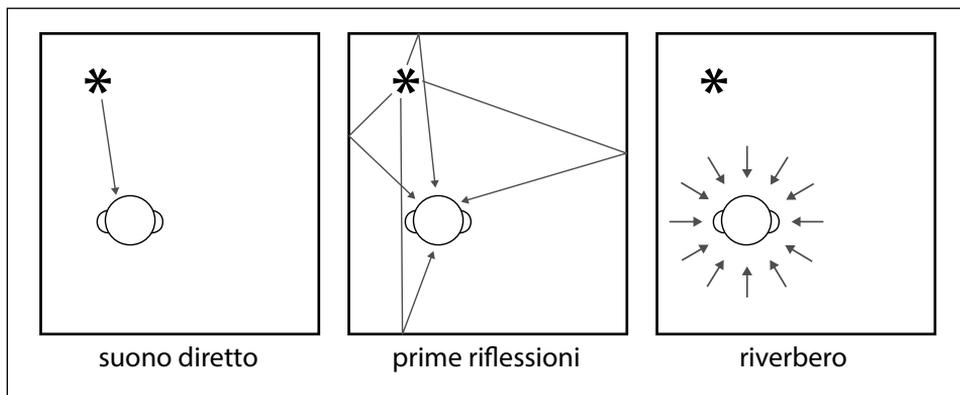


Figura 1.14 Una sorgente emette un suono: quello che giunge all'ascoltatore senza subire alcuna riflessione è il suono diretto. Propagandosi in tutte le direzioni, le onde incontrano le pareti e vengono riflesse: quelle che giungono all'orecchio dell'ascoltatore dopo essere state riflesse una sola volta sono le prime riflessioni, e poiché il percorso che hanno seguito è più lungo, arriveranno all'ascoltatore con un certo ritardo rispetto al suono diretto. Con il proseguire delle riflessioni, che si fanno sempre più numerose e progressivamente più attenuate, si ha il fenomeno del riverbero.

Gli studi in cui, oltre alle registrazioni musicali, si effettuano registrazioni puramente vocali (per esempio: interviste per la radio o per podcast) riservano a questo scopo una stanza di dimensioni relativamente contenute con caratteristiche di riverbero controllate in modo da garantire qualità ma anche intelligibilità delle voci (quando il riverbero è eccessivo, diventa difficile distinguere bene le parole pronunciate).

Modi in ambienti chiusi

Ogni ambiente chiuso ha una peculiare distribuzione di onde stazionarie a bassa frequenza, che sono chiamate *modi*. Sono le risonanze che dipendono direttamente dalle dimensioni dell'ambiente chiuso e possono essere modi *assiali* quando la risonanza coinvolge una sola dimensione della stanza (le onde vengono riflesse da due pareti opposte), *tangenziali* quando la risonanza coinvolge due dimensioni (per esempio, le quattro pareti della stanza) e *obliqui* quando coinvolgono tutte le tre dimensioni (tutte le pareti, il soffitto e il pavimento della stanza), come è rappresentato nella Figura 1.15. Abbiamo fatto un'ipotesi semplificatrice: che la stanza, il nostro ambiente chiuso, sia un parallelepipedo, quindi con i lati opposti paralleli – come sono del resto la maggior parte delle stanze di costruzione moderna. Le cose si fanno più complicate se la geometria dell'ambiente è diversa, come nel caso di stanze di forma poligonale, magari non simmetriche, o con pareti curve, o con il soffitto inclinato o a volta e così via.

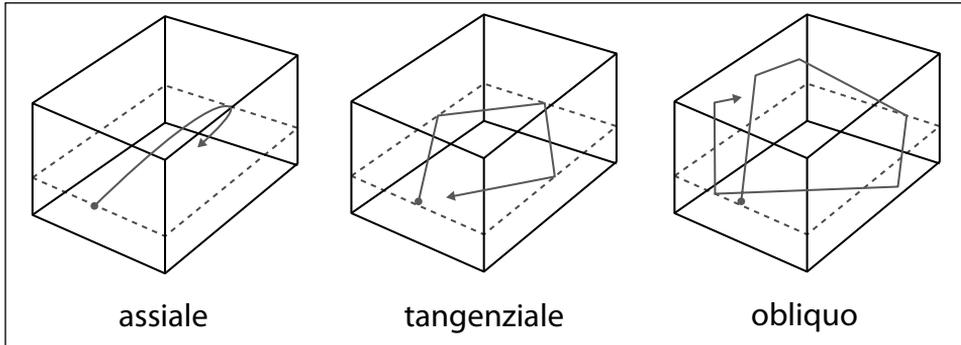


Figura 1.15 I modi assiali, tangenziali e obliqui coinvolgono rispettivamente due, quattro e tutte le sei pareti di un locale chiuso.

Va tenuto presente che, se una stanza ha un modo, poniamo, a 50 Hz, a causa di una delle sue dimensioni, ha anche i modi corrispondenti a multipli di quella frequenza: quindi anche a 100, 150, 200 Hz, e così via. Le frequenze che ci interessano, in vista dell'utilizzo della stanza come sala di registrazione o di controllo, sono quelle più basse – dai 300 o dai 200 Hz (gli esperti hanno pareri un po' diversi in proposito) in giù: per contestualizzare questi valori, il Re4 corrisponde a 293,66 Hz, il Sol3 a 196 Hz; la nota più grave di un contrabbasso è un Mi1 a 41,2 Hz. Stiamo parlando quindi di un intervallo di frequenze molto significativo, che copre praticamente tutta l'estensione di un basso elettrico e quasi metà dell'estensione di un pianoforte.

Se chiamiamo L (per *length*, lunghezza) la dimensione maggiore di una stanza, W (*width*, larghezza) la seconda dimensione in ordine di grandezza, e infine H (*height*, altezza) la terza, espresse in metri; indichiamo con c la velocità del suono nell'aria (344 m/s), con p , q , r tre numeri interi maggiori o uguali a 0, e con f la frequenza, la formula matematica per calcolare i modi di quell'ambiente è:

$$f = \frac{c}{2} \sqrt{\frac{p^2}{L^2} + \frac{q^2}{W^2} + \frac{r^2}{H^2}}$$

I tre interi p , q e r permettono di identificare i modi e di stabilire, per ciascuno, di che tipo di modo si tratti: se, per esempio, p e q sono 0 e r è 1, il modo è (0, 0, 1) ed è un modo assiale, perché è coinvolta solo una coppia di superfici (pavimento e soffitto, in questo caso); se p è 1, mentre q e r sono 0, si ha il modo (1, 0, 0), primo modo assiale delle due pareti opposte in lunghezza. Quando c 'è un solo zero, come in (1, 1, 0) o (1, 0, 1) si tratta di un modo tangenziale (sono coinvolte due dimensioni), quando non ci sono zeri, come in (1, 1, 1) o (2, 1, 1), il modo è obliquo (sono coinvolte tutte le tre dimensioni).

Nel caso del modo (0, 0, 1), per esempio, sotto il segno di radice resterebbe $1/H^2$, e la radice sarebbe $1/H$; $c/2$ è $344/2 = 172$; quindi il risultato sarebbe $172/H$. Se l'altezza della stanza fosse di 2,5 metri, quindi, il primo modo assiale per l'altezza darebbe una frequenza di 68,8 Hz.

Per ottenere i vari modi, si fanno variare i tre interi p , q , r , partendo da 1 e salendo. Per ogni terna si otterrà una frequenza. Si procede finché si arriva alla soglia di frequenza che interessa. Le frequenze ottenute possono poi essere disposte in ordine crescente per avere una panoramica.

NOTA

Solitamente si valuta il numero di modi presenti in ogni ottava o, meglio ancora, in ogni terzo d'ottava: se le proporzioni di un locale sono adeguate, il numero dei modi presenti in ciascuna banda larga un terzo di ottava dovrebbe essere maggiore o al più uguale a quello dei modi presenti nella banda precedente.

I modi sono fra i fenomeni più fastidiosi per la registrazione e l'ascolto: per ottenere una risposta più equilibrata, l'ambiente deve essere trattato con opportuni interventi (lo vedremo nei prossimi capitoli).

Psicoacustica

Anatomia e fisiologia dell'apparato uditivo sono molto più complesse di quanto possa sembrare dalle poche righe dedicate all'orecchio umano all'inizio di questo capitolo; i dettagli non ci interessano qui (ma saperne di più certo non fa male, e chi fosse curioso trova ottimi testi specifici). Dedichiamo invece qualche osservazione agli aspetti più propriamente psicologici della percezione dei suoni, in rapporto con le caratteristiche fisiche dei suoni stessi, che è il campo di ricerca della *psicoacustica*. Anche questo è un ambito molto vasto; vediamo solamente alcuni aspetti che hanno una particolare rilevanza per la registrazione e la riproduzione.

Volume percepito

Per cominciare, abbiamo collegato l'*intensità*, ovvero l'ampiezza delle onde sonore, al *volume* (o *loudness*, in inglese) percepito, ma si tratta di una semplificazione. Il volume percepito è sicuramente legato al livello di pressione sonora, misurato, come abbiamo visto, in decibel, ma dipende anche dalla frequenza: un suono di bassa frequenza deve avere un livello di pressione sonora più elevato rispetto a un suono di media frequenza (intorno ai 4 kHz), per essere percepito di pari intensità.

Parametri fisici e parametri percettivi

Abbiamo detto inizialmente che i tre parametri fisici dei suoni, intensità, frequenza e forma d'onda, corrispondono ai tre parametri volume, altezza e timbro, della nostra percezione. Come abbiamo appena visto, però, il volume percepito è influenzato anche dalla frequenza. In realtà, tutti i parametri fisici influenzano in qualche misura tutti i parametri percettivi. Così, anche se i rapporti più forti sono quelli che abbiamo visto all'inizio, l'intensità di un suono contribuisce anche alla percezione dell'altezza e influisce su quella del timbro e la forma d'onda influenza in parte anche la percezione del volume e dell'altezza.

È esperienza comune che, quando si abbassa il volume di un impianto audio, i bassi si sentano meno; il volume di ascolto è quindi rilevante, in particolare in sala di regia durante le fasi di mixaggio, ed è importante che non sia troppo basso né troppo alto, oltre che per ragioni di salute del sistema uditivo, per non rischiare di effettuare correzioni che abbiano un impatto negativo sull'ascolto in altre condizioni.

Localizzazione dei suoni

L'aver due orecchie ci permette di localizzare i suoni: le onde sonore arrivano ai due padiglioni seguendo tragitti diversi, in tempi anche minimamente diversi perché diversa è la distanza che devono percorrere, e anche lo spettro del suono è diverso (per il diverso mix di suono diretto e riflessioni). Per quanto piccole siano queste differenze, permettono al nostro sistema uditivo di identificare bene la posizione delle sorgenti sonore. Nei missaggi stereo o surround questa capacità viene utilizzata, distribuendo in modo diverso i diversi suoni per ottenere effetti di distribuzione spaziale.

Effetto Haas

Se alle nostre orecchie giungono suoni simili in rapida successione, il primo viene percepito come suono diretto, mentre i successivi sono percepiti come riflessioni, anche nel caso in cui fossero di intensità maggiore. È questo il cosiddetto *effetto Haas* o *effetto precedenza*, che va anche sotto il nome di *legge del primo fronte d'onda*. Questo effetto è importante per la localizzazione dei suoni, perché è sul suono diretto che il nostro sistema uditivo applica la sua analisi delle differenze fra ciò che arriva a un orecchio e ciò che arriva all'altro.

Grazie all'effetto precedenza, siamo in grado di individuare facilmente da quale direzione provenga un suono anche se siamo in un ambiente riverberante. Tutti i suoni che ci arrivano nell'arco di circa 50 millisecondi dal suono diretto sono convogliati in un'unica percezione; oltre i 50 millisecondi vengono percepiti come qualcosa di diverso, di separato, e a quel punto si ha il fenomeno dell'eco.

Mascheramento

La nostra capacità di separare i diversi suoni in base alle loro caratteristiche è notevole, e la diamo per scontata: se suonano un pianoforte e un violino, non abbiamo difficoltà a stabilire quale sia la parte dell'uno o dell'altro. Esistono però casi in cui il nostro orecchio può essere tratto in inganno. Quando due strumenti producono frequenze molto vicine, è possibile che uno dei due *mascheri* l'altro: i suoni di un basso elettrico possono mascherare, per esempio, quelli della grancassa, il che impoverisce l'effetto complessivo. In un mixaggio, il mascheramento può essere ridotto intervenendo sui due suoni per differenziarli maggiormente, per esempio mediante filtri o equalizzazione.

Adattamento

Bisogna fare attenzione anche al fenomeno dell'*adattamento*: suoni che si ripetono con costanza, e che inizialmente producono fastidio, dopo un po' vengono ignorati e ritornano percepibili solo in momenti particolari (il ticchettio dell'orologio alla parete); particolari colorazioni dei suoni, dovute alle caratteristiche degli altoparlanti o dell'impianto di riproduzione, dopo un po' diventano "normali" e non sono più apprezzate. Per questo è importante, nelle fasi di monitoraggio delle registrazioni e di mixaggio avere monitor il più possibile equilibrati: se i monitor esaltano i bassi, quello che in un mix sembrerà molto evidente si appiattirà quando verrà ascoltato con un sistema di riproduzione di-

verso. Per lo stesso motivo, oltre a scegliere monitor equilibrati e a tenere sotto controllo le risonanze della sala di controllo, è sempre opportuno ascoltare i mix su (almeno) un secondo impianto di riproduzione.

Da analogico a digitale (e ritorno)

Fin qui abbiamo parlato di suono come fenomeno “naturale”, ma per l’elaborazione informatica il suono deve essere convertito in formato digitale. Come avviene la trasformazione di un fenomeno analogico, continuo, in una rappresentazione digitale, discreta? Sostanzialmente, attraverso una coppia di processi, che prendono il nome di *campionamento* e *quantizzazione*.

Campionamento

Opportuni apparecchi elettronici permettono di misurare l’ampiezza di un’onda sonora in un dato istante: quella misura è un *campione* dell’onda. Se si misura l’ampiezza dell’onda in più istanti, si ottiene una sua rappresentazione per punti: quanto più ravvicinati nel tempo i campionamenti, tanto più numerosi i punti registrati e tanto migliore sarà l’approssimazione che si otterrà (Figura 1.16).

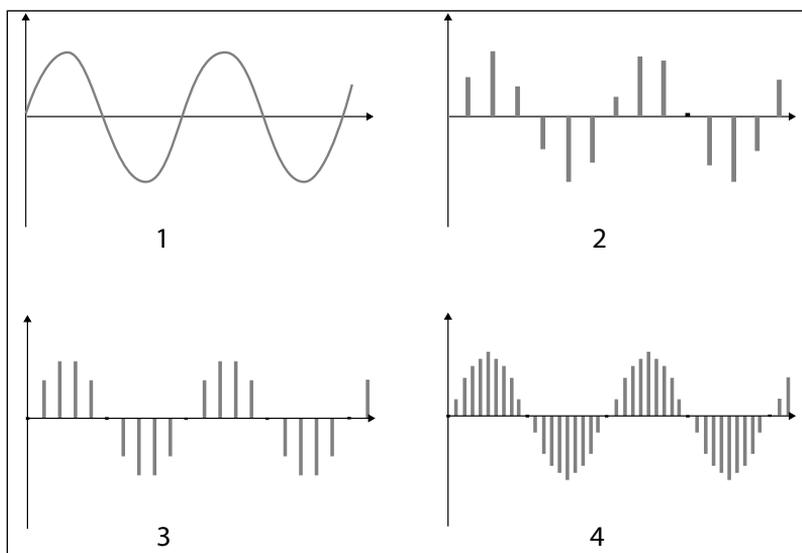


Figura 1.16 Il campionamento di un segnale analogico (1). Al crescere del numero dei campionamenti (2, 3, 4), la rappresentazione digitale dell’onda diventa sempre più fedele.

Il numero dei campionamenti effettuati nell’unità di tempo è la *frequenza di campionamento* e, come la frequenza delle onde sonore, si misura in hertz: 1 hertz equivale a 1 campionamento al secondo. Le frequenze di campionamento utilizzate nella digitalizzazione dei suoni sono normalmente però molto più elevate: un teorema matematico (*teorema di*

Nyquist) dice che, per poter ricostruire correttamente le onde di partenza, la frequenza di campionamento deve essere (almeno) il doppio della frequenza del suono più acuto che si vuole digitalizzare. Ricordando che la soglia di udibilità massima per un essere umano è intorno ai 20 kHz, questo significa che una buona frequenza di campionamento non può essere inferiore ai 40 kHz circa; quella adottata per i CD audio, infatti, è di 44,1 kHz; nelle fasi di acquisizione ed elaborazione, si usano oggi anche frequenze di campionamento più elevate (96, 192 kHz).

Per avere un'analogia visiva: la frequenza di campionamento per un suono corrisponde al numero dei pixel per unità di superficie in un'immagine digitalizzata. Se i pixel sono pochi, l'immagine appare sgranata, se non addirittura irricognoscibile; all'aumentare dei pixel, l'immagine è sempre più definita.

Quantizzazione

Sulla qualità della digitalizzazione influisce non solo la frequenza di campionamento, ma anche (almeno) un altro fattore, cioè il grado di "finezza" della misura ottenuta a ogni singolo campionamento. Anche in questo caso, quella che nel mondo analogico sarebbe una misura continua deve essere trasformata in valori discreti (*quantizzazione*) e diventa importante il numero di valori diversi che il sistema di misurazione è in grado di distinguere. Quanto maggiore quel numero, tanto più fine sarà la misurazione.

Nel mondo digitale, tutto alla fine si riduce a *bit*, cifre digitali che possono assumere solo il valore 0 o 1: 1 bit permette di distinguere quindi solo 2 valori, 2 bit permettono di distinguerne 4 e via continuando, sulla base delle potenze di due. Quindi, n bit permettono di codificare n^2 valori distinti. Con 8 bit si codificano al più 256 valori, con 16 bit se ne codificano 65.536, con 24 bit 16.777.216; 24 bit è il livello di quantizzazione tipico delle apparecchiature di digitalizzazione moderne.

Per riprendere l'analogia visiva, il livello di quantizzazione corrisponde alla profondità di colore di un'immagine: in quel caso, all'aumentare del numero dei bit aumenta il numero dei colori che possono essere differenziati.

Se si digitalizzano suoni a una frequenza di campionamento di 44,1 kHz, con quantizzazione a 24 bit, ogni secondo di suono digitalizzato produce $24 \times 44.100 = 1.058.400$ bit, ovvero 132.300 byte (1 byte = 8 bit), il che corrisponde a circa 8 megabyte per un minuto. Possono non sembrare molti, ma quando si registra non bisogna dimenticare che questo è il "peso" di una sola traccia: se si lavora con molte tracce, la quantità di megabyte che deve essere memorizzata (e poi elaborata) cresce molto rapidamente. Questo influisce sui requisiti di memoria (di lavoro e di massa) della stazione di lavoro che si utilizza per la registrazione digitale.

Da digitale ad analogico

Anche se tutta la catena di registrazione ed elaborazione dei suoni sta nel mondo digitale, alla fine è necessario tornare al dominio analogico: tutti quei bit devono essere riconvertiti in onde sonore continue, emesse da qualche tipo di altoparlante. Il processo di conversione da digitale ad analogico in sostanza prende i singoli punti dei campioni digitali e ricostruisce un'onda continua interpolando le "parti" mancanti.