

MANUALE TEORICO

1.	CAMPO SONORO	3
2.	MODELLI PREVISIONALI	13
3.	RAMSETE	18
3.1	Generazione delle piramidi	18
3.2	Ricerca degli impatti con pareti e ricevitori	21
3.3	Estinzione delle piramidi	29
3.4	Attenuazione dell'energia	29
3.5	Correzione della coda sonora	31
3.6	Taratura dell'algoritmo in ambienti chiusi	35
3.7	Taratura dell'algoritmo all'aperto	44
4.	TECNICHE DI MISURA	48
4.1	La sorgente sonora per la misura della risposta all'impulso	48
4.2	Tecnica dei segnali	49
4.3	La convoluzione	51
4.4	Auralizzazione	61
5.	CONVOLVER PER RAMSETE	68
6.	POTERE FONOISOLANTE, FONOASSORBENTE E FONODIFFONDENTE	78
7.	EFFETTI DIFFRATTIVI	83
8.	SORGENTE DIRETTIVA	85
9.	PARAMETRI ACUSTICI PER GLI AMBIENTI CHIUSI	88

1. CAMPO SONORO

Se si vuole affrontare lo studio del comportamento di un campo sonoro reale generato da una o più sorgenti poste all'interno di un qualsiasi ambiente chiuso, descrivendo matematicamente i parametri che lo caratterizzano (ad esempio la densità di energia) in funzione delle coordinate spaziali e temporali, occorre considerare un numero così elevato di fattori che, senza l'introduzione di opportune semplificazioni, l'impresa può risultare ardua o addirittura proibitiva.

Ad esempio si consideri un ambiente chiuso e regolare, un parallelepipedo.

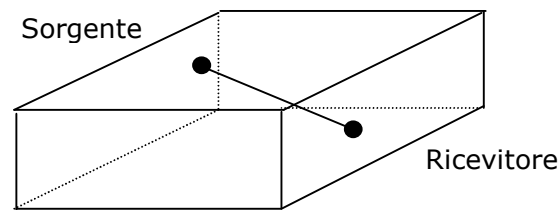


Figura 1.1 - Ambiente chiuso

Si pone al suo interno una sorgente puntiforme e impulsiva, cioè una sfera pulsante che emette dei fronti d'onda perfettamente sferici con uguale intensità in tutte le direzioni per brevi istanti di tempo. All'istante $t_0=0$ la nostra sorgente emette un impulso. Prima però che il livello sonoro arrivi sul ricevitore deve percorrere la distanza d fra il ricevitore stesso e la sorgente.

Quindi il primo fronte d'onda emesso arriverà dopo un tempo:

$$t_1 = \frac{d}{c}$$

dove $c = 343$ m/s è la velocità del suono nell'aria al 50% di umidità. Quando il suono arriva all'ascoltatore, esplica in un istante un livello piuttosto elevato e poi ridecade rapidamente a zero.

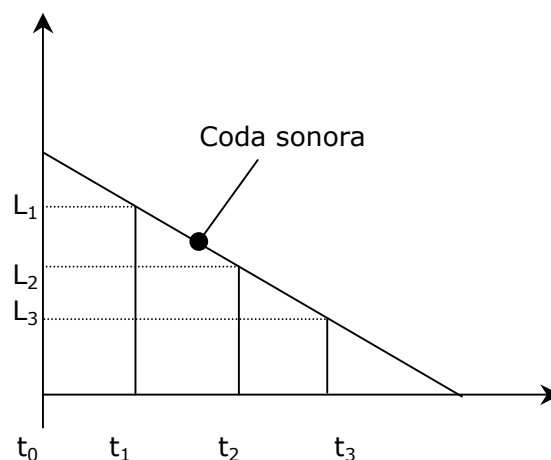


Figura 1.2 - Livelli sonori esplicitati da una sorgente impulsiva

Questo livello chiamato L_1 è il livello del suono diretto, cioè del suono che ha percorso la minima distanza possibile tra sorgente e ricevitore senza aver subito alcun fenomeno di riflessione con le pareti. Nella

sua espansione però il fronte d'onda sferico incontrerà prima o poi le pareti del parallelepipedo e a quel punto verrà riflesso. Il suono riflesso raggiungerà a sua volta il ricevitore, ma avrà percorso una distanza maggiore, come se provenisse da una sorgente immaginaria S' posta al di fuori della stanza e a distanza $d_2 > d_1$ (figura 1.3).

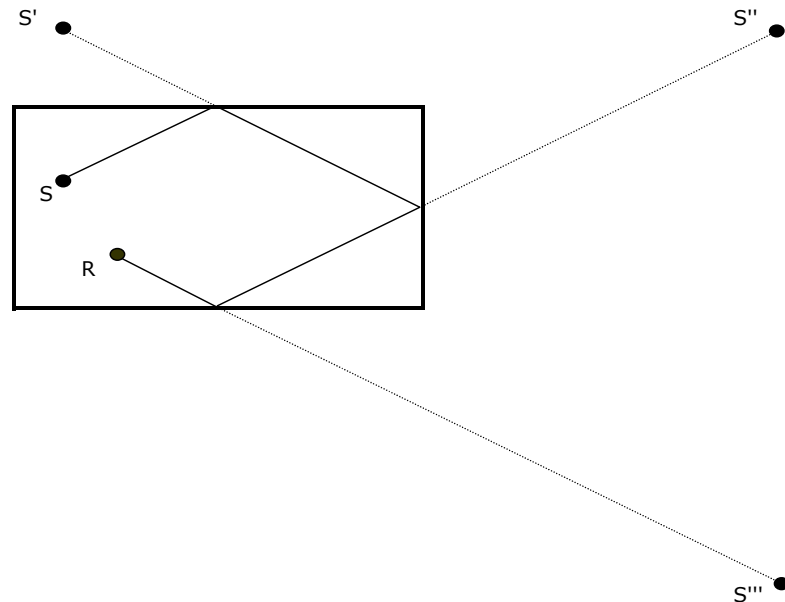


Figura 1.3 - Costruzione delle sorgenti immagine

Il suono riflesso una volta arriverà quindi dopo un tempo $t_2 > t_1$ essendo $d_2 > d_1$. Se l'onda arriva dopo essere stato riflesso più volte si avranno sorgenti immaginarie del secondo ordine, del terzo ordine e così via.

Il numero dei possibili cammini quindi si sviluppa con progressione geometrica. In un parallelepipedo (6 facce) chiamando ad esempio NS' il numero dei possibili cammini dei raggi riflessi una volta si ha:

$$\begin{aligned} NS' &= 6 \\ NS'' &= 6 \cdot 5 = 30 \\ NS''' &= 30 \cdot 5 = 150 \\ NS'''' &= 150 \cdot 5 = 750 \end{aligned}$$

Se si visualizza in un diagramma il livello sonoro in funzione del tempo (figura 1.2) ci si aspetta che accadano due fenomeni contemporaneamente: i tempi ai quali arriva una nuova onda diventano sempre più fitti a causa della crescita veloce del numero di raggi riflessi; ma la loro intensità si riduce fino a decadere a zero a causa dell'assorbimento. Si ha allora il fenomeno noto come coda sonora o riverberazione.

Con una sorgente impulsiva il fenomeno è molto chiaro. Le sorgenti impulsive non sono però in uso da molto. Tradizionalmente il campo sonoro veniva studiato con sorgenti di tipo stazionario (costanti nel tempo) che venivano bruscamente interrotte. Interrompendo l'emissione di energia da parte della sorgente il livello sonoro dopo un certo periodo di tempo decade a zero. In tale periodo il suono non è stazionario ma siamo in presenza di quello che si dice un transiente.

Supponiamo di essere in un ambiente silenzioso accendiamo in un certo istante $t=0$ una sorgente di rumore stazionario.

Costruiamo un nuovo diagramma nel quale andiamo a scrivere i livelli sonori che si sviluppano per effetto di tale sorgente. Come prima il

fronte diretto arriva al tempo t_1 e questo assume il valore L_1 , uguale a quello della sorgente impulsiva a patto che la potenza sia la stessa. In questo caso però la sorgente non cessa di emettere dopo un millisecondo ma continua ad emettere, il livello rimane quindi costante fino al tempo t_2 al quale arriva anche il suono prodotto dalla prima riflessione che va a sommarsi al suono della sorgente diretta. Fatta l'ipotesi che il nostro suono sia rumore incoerente, i due segnali si sommano senza fenomeni di interferenza quindi il livello che si raggiunge è la somma in dB di L_1 e L_2 dove L_1 è il suono dell'onda diretta e L_2 è il suono della prima onda riflessa. Dopo di che il livello rimane costante finché non arriva la seconda riflessione la quale fa salire l'intensità ancora di un gradino, questi gradini diventano però sempre più piccoli perché quando ad un suono forte si somma un suono molto più debole il contributo finale tende a diventare trascurabile. Quindi con una serie di piccoli gradini corrispondenti ai singoli arrivi di energia delle successive riflessioni si giunge asintoticamente ad un livello costante L_0 che rappresenta la condizione di equilibrio energetico di un ambiente.

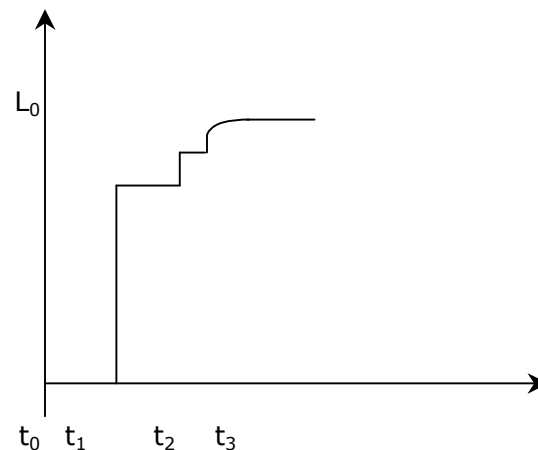


Figura 1.4 - Accensione di una sorgente stazionaria

La potenza erogata dalla sorgente, a regime, viene incessantemente assorbita dalle pareti. Quindi il livello costante L_0 corrisponde al bilancio tra la potenza che continuamente la sorgente sta immettendo e quella che continuamente le riflessioni sulle pareti stanno togliendo all'ambiente. E' possibile ricavare matematicamente la situazione di regime ragionando su base energetica ovvero utilizzando come grandezza la densità di energia D .

Poiché in generale D è una funzione del tempo ma anche delle coordinate spaziali lo studio del suo andamento diventerebbe troppo complicato. Si ipotizza quindi che D sia esclusivamente funzione del tempo $D(t)$.

I possibili percorsi che il raggio sonoro può intraprendere per andare da una parete all'altra del parallelepipedo sono infiniti, si considera allora il cammino libero medio la cui lunghezza L_m , che rappresenta la lunghezza media fra due successive incidenze del raggio sonoro sulle pareti, viene calcolata statisticamente e vale:

$$L_m = \frac{4V}{S}$$

dove S è la superficie totale del parallelepipedo e V il suo volume. La stima viene fatta come media aritmetica ponderale delle lunghezze di tutti i possibili percorsi, assumendo come pesi le probabilità che hanno tali cammini di essere effettivamente percorsi. Tali probabilità dipendono dal tipo di diffusione delle pareti e dal fatto che la sorgente abbia o meno una direzione privilegiata di emissione. In ogni caso i risultati saranno tanto più veritieri quanto meno la situazione reale si discosta dalle ipotesi fatte.

Tornando al raggio sonoro il tempo medio che intercorre tra due incidenze successive deve essere:

$$t_m = \frac{4V}{cS}$$

All'interno dell'ambiente le uniche trasformazioni energetiche che avvengono sono l'immissione di energia E_s da parte della sorgente e l'assorbimento dell'energia E_a da parte delle pareti. Quindi l'espressione della conservazione di energia può essere scritta come:

$$dE = E_s - E_a$$

Si ipotizza che in ogni intervallo di tempo infinitesimo (dt) esiste una frazione di energia circolante dE , anch'essa infinitesima, che incontra le pareti della sala e viene in parte assorbita e in parte riflessa; si ammette che il processo avvenga senza discontinuità. Allora se l'energia circolante nel parallelepipedo è per definizione: $V \cdot D(t)$, la variazione di energia sarà data da:

$$dE = V \cdot \frac{dD(t)}{dt}$$

Nel tempo t_m l'energia incide una sola volta sulle pareti quindi nel tempo dt l'energia incidente vale:

$$E_{\text{incidente}} = \frac{V \cdot D(t)}{t_m} dt$$

Si vuole stabilire ora quanta di questa energia incidente viene assorbita e quanta riflessa. Se α_i è il coefficiente di assorbimento dell' i -esima parete, la quantità di energia assorbita vale:

$$\bar{\alpha} \cdot D(t) \cdot V$$

con

$$\bar{\alpha} = \frac{\sum_i \alpha_i \cdot S_i}{\sum_i S_i}$$

coefficiente di assorbimento medio, cioè la media dei coefficienti di assorbimento di tutte le pareti.

Per l'ipotesi di continuità l'assorbimento risulta distribuito in tutto l'intervallo t_m e quindi l'energia assorbita nel tempo dt vale:

$$E_a = \bar{\alpha} \cdot \frac{V \cdot D(t)}{t_m} dt$$

da cui:

$$E_a = \frac{\bar{\alpha} \cdot D(t) c S}{4} dt$$

Essendo poi: $E_s = dE + E_a$

si ha:

$$W(t) \cdot dt = V \cdot \frac{dD(t)}{dt} dt + \frac{c}{4} D(t) \bar{\alpha} \cdot S dt$$

($W(t)$: potenza della sorgente)

La relazione è generale ed è applicabile sia al caso di sorgente di suono stazionario che di suono variabile nel tempo. Nel caso di suono stazionario dopo un certo tempo dall'accensione della sorgente, la densità di energia nell'ambiente raggiungerà un valore di regime. Più precisamente l'energia tenderà ad assumere un valore massimo costante. Dopo un certo tempo dall'accensione della sorgente si può considerare

$$\frac{dD(t)}{dt} = 0$$

quindi, dopo un certo periodo di tempo

$$D(t) = D_{reg} = \frac{4W}{c\alpha S}$$

La densità di regime è, nel caso di sorgente stazionaria, direttamente proporzionale alla potenza emessa e inversamente proporzionale all'assorbimento totale (o potere fonoassorbente)

$$A = \bar{\alpha} S$$

Nell'ipotesi fatta il livello sonoro, costante in ogni punto, è pari a

$$L_D = L_w + 10 \log \frac{4}{\alpha S}$$

Questa è nota come formula del campo riverberante e esprime il livello sonoro prodotto da una sorgente di rumore stazionario in un campo perfettamente diffuso e riverberante. Tale situazione descrive solo in maniera molto approssimata quello che accade in realtà in quanto avvicinandosi alla sorgente sonora è esperienza comune che i livelli sonori aumentano. Effettivamente quello che nelle approssimazioni si può considerare costante in tutti i punti della stanza è il livello del campo riverberante, cioè del suono che ha già subito almeno una riflessione. Per il suono diretto questa semplificazione non è affatto ragionevole. Quindi la formula del campo riverberante è in generale falsa. Ne esiste una versione modificata che tiene conto di entrambe le componenti, energia del suono diretto ed energia del suono diffuso, semplicemente sommandole, nell'ipotesi che la sorgente emetta ru-

more incoerente che non dia origine a fenomeni di interferenza. Il livello di densità totale allora si può scrivere:

$$L_D = L_W + 10 \log \left[\frac{1}{4\pi d^2} + \frac{4}{\alpha S} \right]$$

dove si vede che il contributo diffuso è lo stesso ovunque, mentre quello diretto varia con l'inverso della distanza al quadrato, quindi è trascurabile se ci si allontana e diventa viceversa predominante approssimandosi alla sorgente. Questa formula è nota come formula del campo semi-riverberante e risulta essere una relazione molto semplice che è da tutti comunemente impiegata per valutare il campo sonoro dentro gli ambienti chiusi.

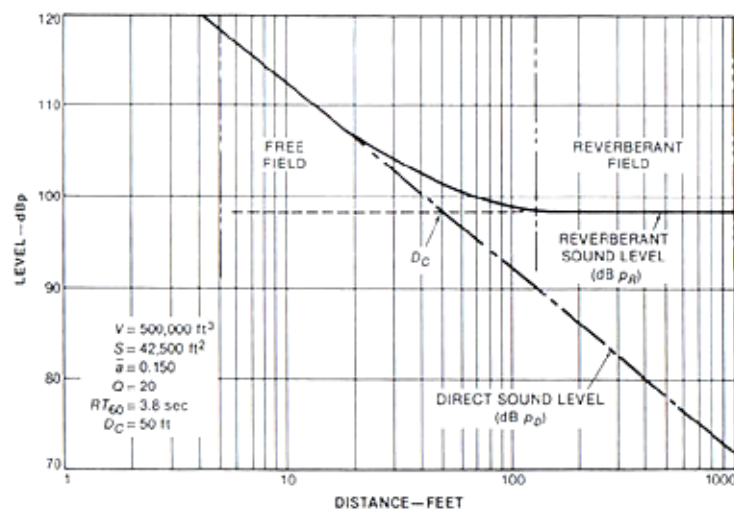


Figura 1.5 - Campo riverberante

Il grafico riporta l'andamento del livello sonoro in dB in funzione della distanza dalla sorgente. La linea tratteggiata mostra il decadimento perfettamente rettilineo che si avrebbe considerando solo il suono diretto. La somma dei contributi, diretto e riverberante, tende invece a stabilizzarsi ad un livello che dipende dalla superficie complessiva dell'ambiente e dal suo coefficiente di assorbimento medio.

Nel caso che la sorgente non sia omnidirezionale, ma abbia delle direzioni privilegiate di emissione, l'equazione del livello sonoro in campo stazionario può essere scritta come:

$$L_D = L_W + 10 \log \left[\frac{Q}{4\pi d^2} + \frac{4}{\alpha S} \right]$$

Ovviamente Q (coefficiente di direttività) può essere maggiore o minore di uno in funzione che ci si trovi entro o fuori dall'angolo di massima emissione. E' chiaro che essendo l'energia totale fissa se questa è maggiore in una direzione deve essere minore nell'altra. Quindi nella maggior parte delle zone il coefficiente Q è minore di 1.

Distanza critica Si definisce distanza critica quella particolare distanza dalla sorgente alla quale il livello del campo sonoro diretto e del campo sonoro riverberante assumono lo stesso valore, si indica con D_c e vale:

$$D_c = \sqrt{\frac{QS\bar{\alpha}}{16\pi}}$$

Equazione che si desume facilmente eguagliando i due termini relativi al campo diretto e diffuso

$$\frac{Q}{4\pi D^2} = \frac{4}{\alpha S}$$

Dal punto di vista grafico la D_c rappresenta il punto in cui la curva complessiva del campo riverberante più campo diretto è sollevata di 3 dB rispetto alla retta orizzontale del campo puramente riverberante. In quella posizione dell'ascissa si legge la distanza critica (figura 1.5).

La distanza critica è un parametro estremamente importante. In termini di qualità e comprensione del messaggio che viene riprodotto da un altoparlante; infatti se ci si trova entro tale distanza dalla sorgente il suono diretto è predominante sul campo riverberante. Il suono diretto è chiaro, è nitido e porta un'informazione perfettamente intellegibile, viceversa il suono riverberante è confuso. Quindi soprattutto per quanto riguarda la comprensione della parola è importante che l'ascoltatore venga a trovarsi sempre entro la distanza critica dalla sorgente. Questo significa che non è possibile in ambienti molto vasti posizionare un unico altoparlante per esempio al centro e sperare che il suono arrivi chiaro in ogni punto dello spazio. Le possibili soluzioni alla perdita di qualità da parte del suono sono due. Si può puntare un altoparlante molto direttivo direttamente sul pubblico come mostrato nella figura sotto. Così facendo, poiché le persone si possono pensare come un materiale molto assorbente, l'energia viene quasi completamente assorbita e non va ad alimentare il campo riverberante.

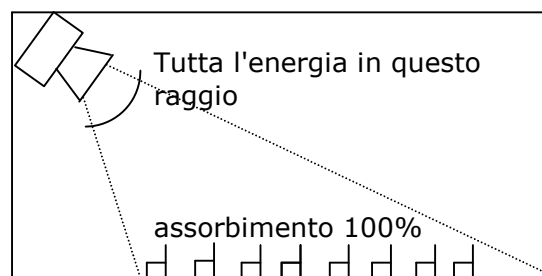


Figura 1.6- Altoparlante direttivo sul pubblico

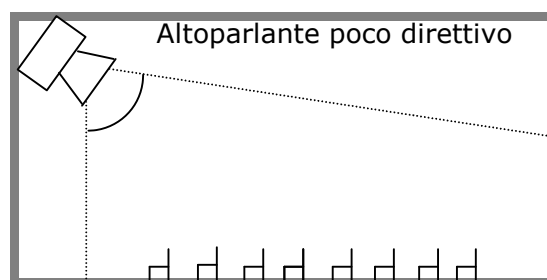


Figura 1.7- Ambiente fonoassorbente

Oppure si può usare una sorgente poco direttiva, a patto però che l'ambiente presenti un alto coefficiente di assorbimento.

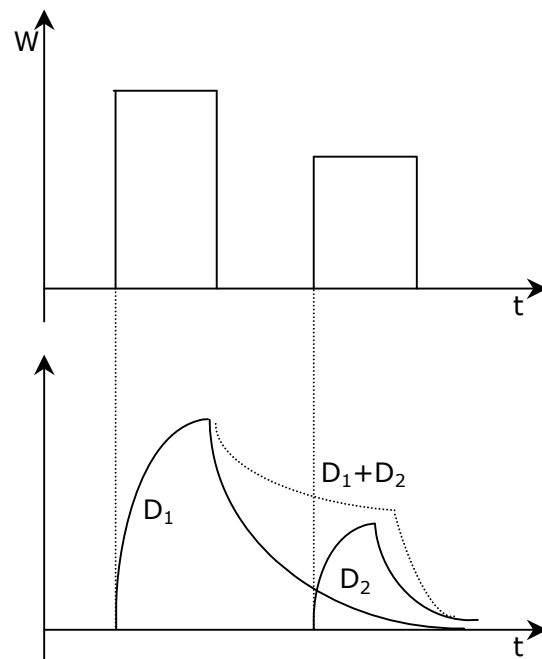


Figura 1.8- Potenza e densità di energia di due suoni

Nel caso che una sorgente emetta due suoni separati da un breve silenzio (grafici di figura 1.8) il transitorio di estinzione del primo segnale interferisce con il transitorio iniziale del suono seguente. Questo crea evidentemente un disturbo nel secondo segnale, tanto maggiore quanto maggiore è il transitorio di spegnimento ovvero la riverberazione del segnale che lo precede. Quindi affinché il suono non subisca un'alterazione che lo renda irriconoscibile occorre che la densità sonora del transitorio di estinzione di ciascuno dei segnali acustici discenda velocemente a valori sufficientemente bassi. Tale rapidità in acustica ambientale viene valutata mediante una grandezza denominata tempo di riverberazione, proposta per la prima volta da W. Sabine.

Tempo di riverbero o di riverberazione Si dice tempo di riverberazione (T_{60})

- il tempo necessario affinché la densità di energia sonora discenda ad un milionesimo del valore che aveva nell'istante di spegnimento della sorgente
- il tempo necessario affinché la pressione acustica nell'ambiente diminuisca di 60 dB

In una situazione a regime, si spenga la sorgente. La prima energia che viene a mancare è quella del suono diretto. Mentre in accensione però, rispetto al silenzio, quest'ultima aveva causato un brusco aumento del livello sonoro, ora rispetto al livello di regime alimentato dalla moltitudine di raggi riflessi, l'abbassamento che ne consegue è di entità molto più modesta. Ancora meno rilevante è l'abbassamento dovuto al venir meno del primo raggio riflesso. Il livello tende dunque a decadere a zero con un andamento come in figura 1.9.

Dopo un primo tratto scalinato la curva di decadimento diviene "liscia" e rettilinea.

Per valutare il tempo di riverberazione seguendo le disposizioni della

normativa ISO 3382 del 1997 si fa partire il cronometro dopo che il livello è sceso di 5 dB dal livello di regime in modo da evitare il primo tratto "scalinato" e si ferma dunque quando quest'ultimo è sceso complessivamente di 65 dB.

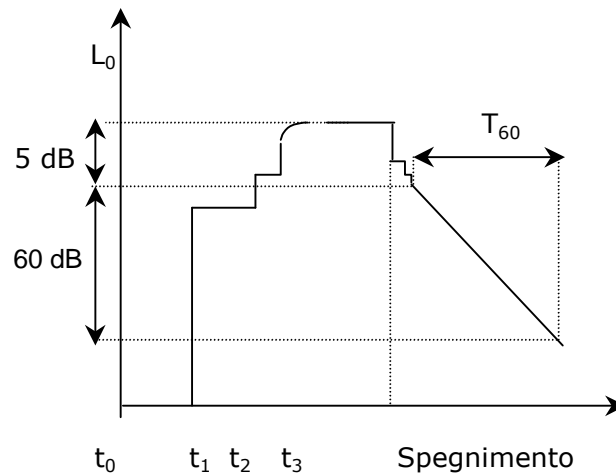


Figura 1.9- Spegnimento di una sorgente stazionaria

La relazione che lega il T_{60} alle caratteristiche dell'ambiente è stata trovata sperimentalmente da Sabine e prende quindi il nome di formula di Sabine:

$$T_{60} = 0,16 \frac{V}{\alpha S}$$

La relazione è facile da ricavare anche per via teorica. Nel momento in cui si spegne la sorgente, $W(t)=0$ quindi, tornando all'equazione generale di equilibrio energetico si ha

$$V \frac{dD(t)}{dt} dt = -\frac{c}{4} D(t) A dt \quad \text{da cui} \quad \frac{dD(t)}{D(t)} = -\frac{cA}{4V} dt$$

e integrando dall'istante t_0 di spegnimento al generico t si ha:

$$\int_{t_0}^t \frac{dD(t)}{D(t)} = -\frac{cA}{4V} \int_{t_0}^t dt \quad \text{cioè} \quad \ln \frac{D(t_0)}{D(t)} = \frac{cA}{4V} (t-t_0)$$

$(t-t_0) = T_{60}$, dalla definizione 1) $D(t)=D(t_0)10^{-6}$ quindi sostituendo

$$T_{60} = 0.16(V/A)$$

Il T_{60} indica la possibilità per l'uomo di percepire o meno la coda sonora. Il suono è "molto riverberante" quando $T_{60} > 2$ sec e "molto asciutto" quando $T_{60} < 1$ sec. Quindi la regolazione del tempo di riverberazione di un ambiente è uno dei principali parametri di progettazione acustica di un ambiente. Il suo valore preferenziale dipende dall'utilizzo della sala:

Aula piccola	0.5
Aula grande	1

Cinema	0.7-0.8
Teatro dell'opera	1.3-1.5
Concert hall	1.7-2.3

Risposta all'impulso ('Misure di isolamento al rumore aereo, del rumore di calpestio e del tempo di riverberazione', A. Farina, P. Fausti).
L'andamento nel tempo di un segnale sonoro per effetto della riverberazione risulta ben evidente nei diagrammi in cui viene riprodotta graficamente la successione delle onde che in seguito all'emissione di un segnale impulsivo raggiungono un determinato punto dell'ambiente.

Sotto è visibile un diagramma di questo tipo, in cui l'onda diretta e le successive riflessioni sono rappresentate da segmenti verticali positivi, di altezza corrispondente alle rispettive intensità e posizionati, rispetto all'origine, in funzione dei tempi impiegati a coprire le distanze percorse.

Questi diagrammi, che rappresentano per definizione la "risposta" dell'ambiente ad un segnale di tipo impulsivo, contengono tutte le informazioni legate alla trasmissione del suono fra sorgente e punto di ascolto necessarie per valutare la qualità del segnale ricevuto.

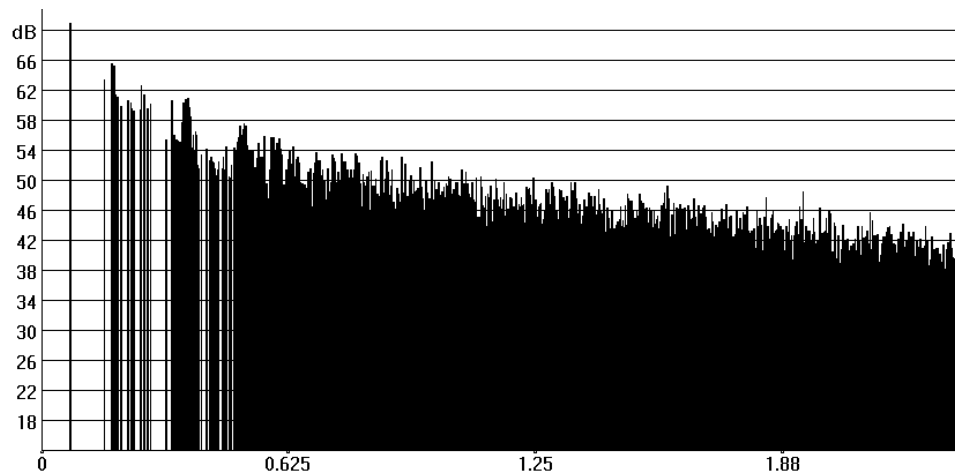


Figura 1.10 – Risposta all'impulso

Già visivamente è possibile apprezzare la presenza di onde riflesse confrontabili con l'onda diretta e il relativo ritardo.

Da questi dati è possibile risalire al tempo di riverbero tramite un procedimento matematico noto come integrazione all'indietro di Schroder, che ha il merito di aver inventato un procedimento molto veloce che permette in maniera semplice il calcolo del decadimento di un suono stazionario interrotto (tempo di riverbero) in una stanza mediante la misura della risposta all'impulso.

E' solo da dopo questa invenzione (1975) che si è incominciato a misurare in maniera sistematica il tempo di riverbero mediante la risposta all'impulso.

2. MODELLI PREVISIONALI

Esistono diversi modelli matematici che consentono di ottenere questo risultato; i più utilizzati:

- il metodo delle sorgenti immagini;
- il Ray Tracing;
- il Pyramid Tracing.

Metodo delle sorgenti immagine

Il primo ad essere utilizzato fu il metodo delle sorgenti immagine, nome che deriva dal fatto di ipotizzare riflessioni speculari sulle pareti di una stanza da analizzare. Esso è particolarmente efficace in ambienti parallelepipedi, ma può essere usato anche in ambienti a geometria diversa: data una sorgente la si ribalta simmetricamente rispetto alle superfici, quindi si possono creare sorgenti del primo ordine, che specularizzate, a loro volta, producono sorgenti del secondo ordine e così via. Un determinato ricevitore riceve il suono dalla sorgente reale e il suono riflesso delle altre sorgenti immagine.

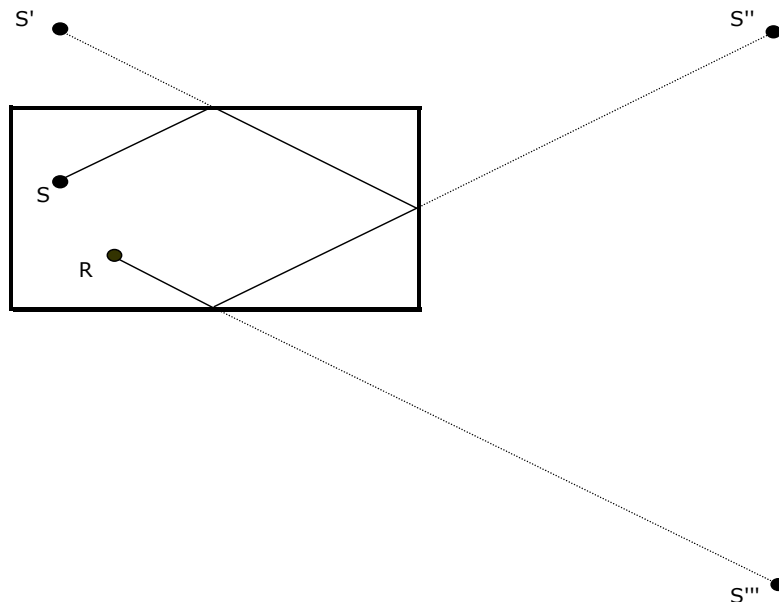


Figura 2.1 – Sorgenti immagine

L'estrema semplicità di questo metodo negli ambienti aventi una geometria parallelepipeda, come ad esempio nel campo industriale, fa sì che questa tecnica funzioni egregiamente e consenta di prevedere la risposta all'impulso fino ad ordini di riflessione molto elevati. Sussistono invece dei problemi quando dobbiamo considerare sorgenti immagini d'ordine elevato in presenza di geometrie complicate; infatti, questo metodo, per geometrie generiche, richiede che, per ogni sorgente immagine e per ogni ricevitore considerato, sia fatto un controllo di visibilità tra la sorgente e il ricevitore e questo rende il calcolo molto lento. Inoltre, poiché il numero delle sorgenti immagine cresce esponenzialmente con l'ordine delle stesse, nel caso di geometrie complicate si può arrivare al massimo alla determinazione dei raggi del quinto ordine e quindi si trascura, nella previsione, la restante coda sonora.

Metodo Ray Tracing

Altro metodo è il Ray Tracing, nel quale s'inizia da una sorgente puntiforme, dalla quale si emettono a caso raggi lineari in cui l'energia sonora è immagazzinata, a tali raggi si applica la legge della riflessione (legge di Snell). Tale procedimento è significativo solo se dalla

sorgente si emettono molti raggi (decine di migliaia o anche milioni); inoltre è necessario avere un ricevitore di dimensione finita, perché la probabilità che un raggio colpisca un ricevitore puntiforme è quasi nulla: per questo solitamente si usa un ricevitore sferico.

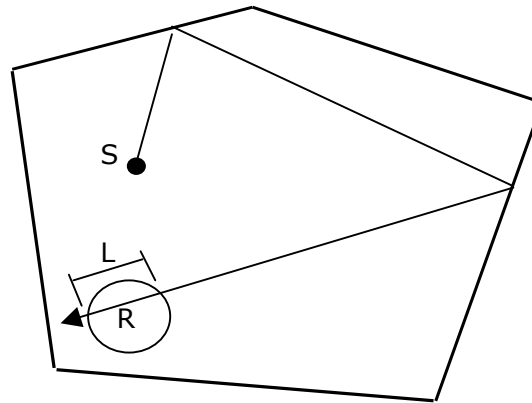


Figura 2.2 - Schema di generazione dei raggi e d'impatto su un ricevitore sferico.

Il parametro che interessa, avendo un ricevitore volumetrico, è la densità d'energia che arriva all'interno del volume:

$$W' = \frac{P_{wr} \cdot Q_{\delta} \cdot L \cdot \prod_i (1 - \alpha_i)}{N_{raggi} \cdot c \cdot V_{sfera}} \cdot e^{-\gamma \cdot x}$$

dove P_{wr} è la potenza sonora emessa dalla sorgente, Q_{δ} rappresenta la direttività nella particolare direzione in cui il raggio iniziale era stato generato, N_{raggi} è il numero di raggi, c è la velocità del suono, V_{sfera} è il volume della sfera, L la lunghezza del segmento con cui il raggio ha intersecato la sfera e $e^{-\gamma x}$ è un coefficiente di riduzione con la distanza che rappresenta l'assorbimento da parte dell'aria. Questo metodo è molto brillante nelle previsioni.

Metodo Pyramid Tracing - Il terzo metodo è quello del tracciamento dei fasci divergenti di cui fanno parte due diverse metodologie: il Cone Tracing (subito abbandonato a favore di una sua diversa implementazione) ed il Pyramid Tracing.

Nel primo, il ricevitore è puntiforme e, quando lo stesso è dentro al cono, questo gli comunica una certa intensità di energia sonora; il primo problema che si pone con questo metodo è il fatto che i coni non coprono completamente una superficie sferica: se sono adiacenti restano delle parti non coperte, se sono sovrapposti, vi sono parti coperte per due volte.

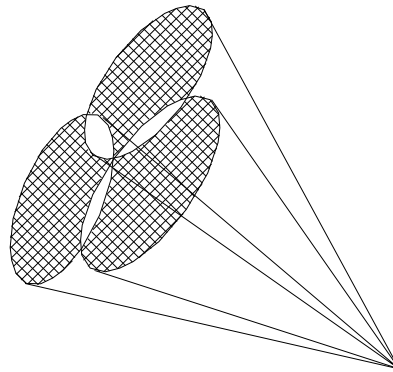


Figura 2.3 - Risultato della sovrapposizione di coni.

Per questo motivo il Cone Tracing è stato trascurato o comunque se lo si utilizza è necessario sovrapporre i coni e creare un algoritmo per evitare le rilevazioni multiple o che "pesi" l'energia (in media) in modo che i contributi multipli producano il livello sonoro corretto. Questo problema non sussiste con il Pyramid Tracing, siccome piramidi adiacenti coprono perfettamente la superficie sferica, come mostrato in figura.

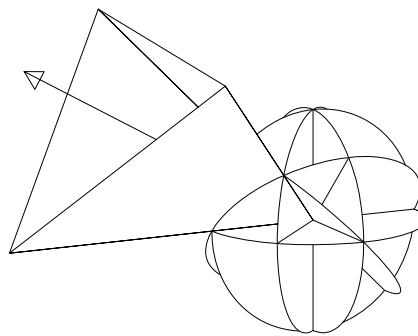


Figura 2.4 - Pyramid Tracing

La suddivisione della superficie in triangoli, è fatta utilizzando una versione modificata dell'algoritmo di Tenebaum, procedendo alla suddivisione in 8 ottanti della sfera: in questo modo il numero delle piramidi generate può essere di un numero che sia una potenza di 2 ($8 \cdot 2^N$), inoltre tutte queste hanno la stessa area di base, generando così una sorgente sonora isotropica.

I vantaggi di questa tecnica sono il ricevitore puntiforme e il numero contenuto di fasci da emettere (si fanno simulazioni discrete già con 2048 piramidi) il tutto a vantaggio della velocità di elaborazione dei dati da parte di un calcolatore. Chiaramente il tempo di calcolo cresce con diretta proporzionalità al numero di piramidi tracciate.

Il tracciamento del raggio centrale di ciascuna piramide avviene seguendo le usuali ipotesi dell'acustica geometrica: riflessione speculare all'impatto con una superficie, ubbidendo alla legge di Snell. Dopo ogni riflessione si provvede a costruire la posizione della sorgente immagine rispetto alla superficie impattata, e si prosegue il tracciamento del raggio a partire da tale nuova sorgente virtuale.

La verifica dell'impatto sui ricevitori avviene quando uno di essi si trova all'interno di una piramide che si sta tracciando. Questo fatto viene verificato tracciando indietro sulla superficie sferica della sorgente immagine la congiungente sorgente-ricevitore, e verificando che tale punto sia interno al triangolo costituito dai tre spigoli della piramide. Se si verifica la condizione di arrivo di energia sul ricevitore-

re, il contributo ricevuto viene memorizzato in una opportuna matrice, costituita da 10 colonne (le 10 bande di frequenza) e da alcune centinaia o migliaia di righe, corrispondenti ciascuna un certo bucket temporale rispetto all'istante di emissione. Chiaramente la matrice dei risultati sarà tanto più lunga (e dunque occuperà più memoria) quanto più lunga è la risposta all'impulso da calcolare, e tanto maggiore è la risoluzione temporale voluta.

Il valore memorizzato è una "intensità sonora scalare" cioè non è il vettore intensità che si misura con gli intensimetri, ma piuttosto un valore fittizio, legato al quadrato della pressione sonora, ottenuto dalla relazione:

$$I = \frac{Q}{4\pi \cdot d^2} \cdot \prod_i (1 - \alpha_i)$$

dove:

Q = direttività nella direzione considerata

d = distanza fra la sorgente virtuale ed il ricevitore

α_i = coefficienti di assorbimento delle superfici su cui il raggio ha impattato.

La presenza della direttività significa che è possibile impiegare sorgenti dotate di "balloon" di direttività molto dettagliati, definiti con una risoluzione di 10^0 sia in longitudine, sia in latitudine, per ciascuna banda di frequenza. Anche nel caso si adottino poche piramidi, con apertura angolare maggiore di 10^0 , il dato di direttività impiegato sarà sempre quello relativo alla effettiva congiungente sorgente-ricevitore, e pertanto il calcolo dell'onda diretta è sempre esatto.

Il tracciamento delle piramidi viene proseguito fino ad ordini elevatissimi (anche se è possibile limitarlo ad un ordine prefissato), in modo da ricostruire l'intera coda sonora, in ciascun punto ricevitore. Questo modello, pertanto, al contrario degli altri tracciatori di fasci divergenti sinora implementati, non impiega alcuna relazione dell'acustica statistica per ricostruire la parte tardiva della coda riverberante, rimanendo sempre del tutto deterministico.

Al crescere del percorso, la "base" della piramide si allarga progressivamente, e finisce per diventare più grande dell'intero ambiente. In queste condizioni tutti i ricevitori si trovano sempre entro la piramide, e dunque vengono sempre colpiti dai raggi. Questo fatto è decisivo, perché in tal modo la coda riverberante mantiene un sufficiente dettaglio, e non compaiono i frastagliamenti tipici di altri modelli di calcolo, che danno luogo ad una probabilità di impatto su un ricevitore decrescente con il percorso compiuto dallo stesso. Infatti, sia con i tradizionali modelli di Ray Tracing, sia con i più recenti modelli ibridi (Ray Tracing + Sorgenti Immagine), è necessario tracciare un numero di raggi crescente con il quadrato della lunghezza della risposta all'impulso.

Ad esempio, Lehenert indica il numero minimo di raggi da lanciare per un Ray Tracing tradizionale:

$$N_{min} = 2\pi \cdot \left(\frac{l_{max}}{r} \right)^2$$

dove:

l_{max} = lunghezza massima del tragitto di un raggio

r = raggio del ricevitore.

Ad esempio, con un ricevitore avente 1 m di raggio, ed una risposta all'impulso lunga 1 s ($l_{max} = 340$ m), si ottiene $N_{min} = 726336!$

Maercke e Martin, invece, indicano il numero minimo di raggi necessari per un tracciatore di coni (metodo ibrido):

$$N_{min} = 50 \cdot \left(\frac{l_{max}}{l_{cm}} \right)^2$$

Se si ipotizza una risposta all'impulso lunga 1 s ed un libero cammino medio l_{cm} di 15 m, si ottiene $N_{min} = 25689$. Il metodo ibrido è dunque decisamente più veloce, ma richiede comunque calcoli di durata non trascurabile.

Con il tracciatore di piramidi, invece, la precisione del risultato non dipende così drasticamente dal numero di raggi lanciati. Il numero di piramidi da tracciare dipende infatti soltanto dalla risoluzione temporale richiesta alla risposta all'impulso e non dalla sua lunghezza. Ciò significa che, anche con un numero ridotto di piramidi, ad esempio 256, è possibile prevedere correttamente risposte all'impulso di parecchi secondi con una risoluzione temporale discreta (10 ms).

Volendo viceversa una risoluzione temporale molto spinta, onde apprezzare i singoli echi e valutare con precisione il loro tempo di arrivo, si può utilizzare una risoluzione di 1 ms ed un numero di piramidi pari a 2048.

Come in tutti i tracciatori di fasci divergenti (Beam Tracing, Cone Tracing, etc.), la coda sonora tende ad essere sottostimata in campo riverberante, poiché, al crescere delle dimensioni della base di ciascun fascio, aumenta la probabilità di "perdere" una sorgente immagine. E' pertanto necessario effettuare una correzione della coda sonora, che viceversa non era necessaria con algoritmi di tipo Ray Tracing, a patto però di tracciare effettivamente il numero di raggi richiesto dalla relazione.

Ramsete impiega una correzione della coda sonora di tipo moltiplicativo, come originariamente suggerito da Naylor. L'intensità sonora ricevuta viene pertanto così corretta:

$$I_{corr}(t) = \frac{I(t)}{\left(1 - e^{-\frac{l_{cm}^2 \cdot N}{4\beta \cdot c_0^2 t^2}} \right)}$$

Nella precedente relazione compaiono due parametri dipendenti dalla natura del campo sonoro: il libero cammino medio l_{cm} , che viene calcolato da Ramsete in base ad una statistica sull'effettiva percorrenza dei raggi sparati in asse alle piramidi, ed il coefficiente adimensionale β , che dipende dalla natura più o meno sabiniana del campo sonoro (in campo perfettamente diffuso risulta $\beta = 0.1$).

(Propagazione sonora e previsione del rumore negli ambienti di lavoro, A. Farina)

3. RAMSETE

3.1 Generazione delle piramidi

Al primo passo (a livello 0), la sfera verrà suddivisa in otto parti, al secondo in sedici e così via secondo la serie 8×2^N . Questo calcolo viene fatto una volta per tutte all'inizio dell'elaborazione e quindi non incide in modo pesante sui tempi di calcolo.

La generazione delle piramidi è perfettamente isotropa, grazie all'algoritmo di Tenenbaum, costituito da una progressiva bisezione degli 8 spicchi di partenza.

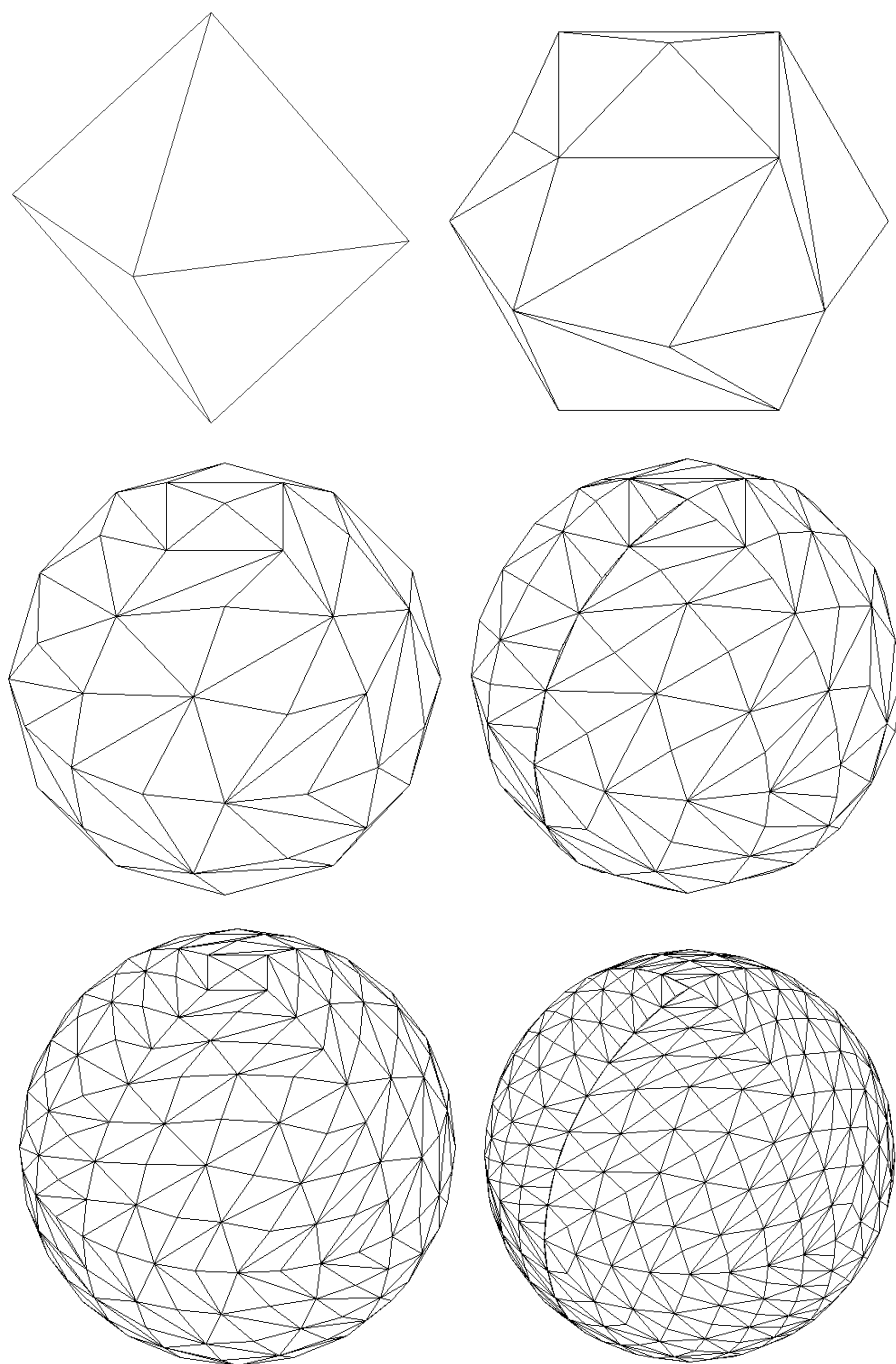


Figura 3.1 - Emissione della piramide e suddivisione della sfera in 8, 32, 128, 256, 512 e 1024 triangoli.

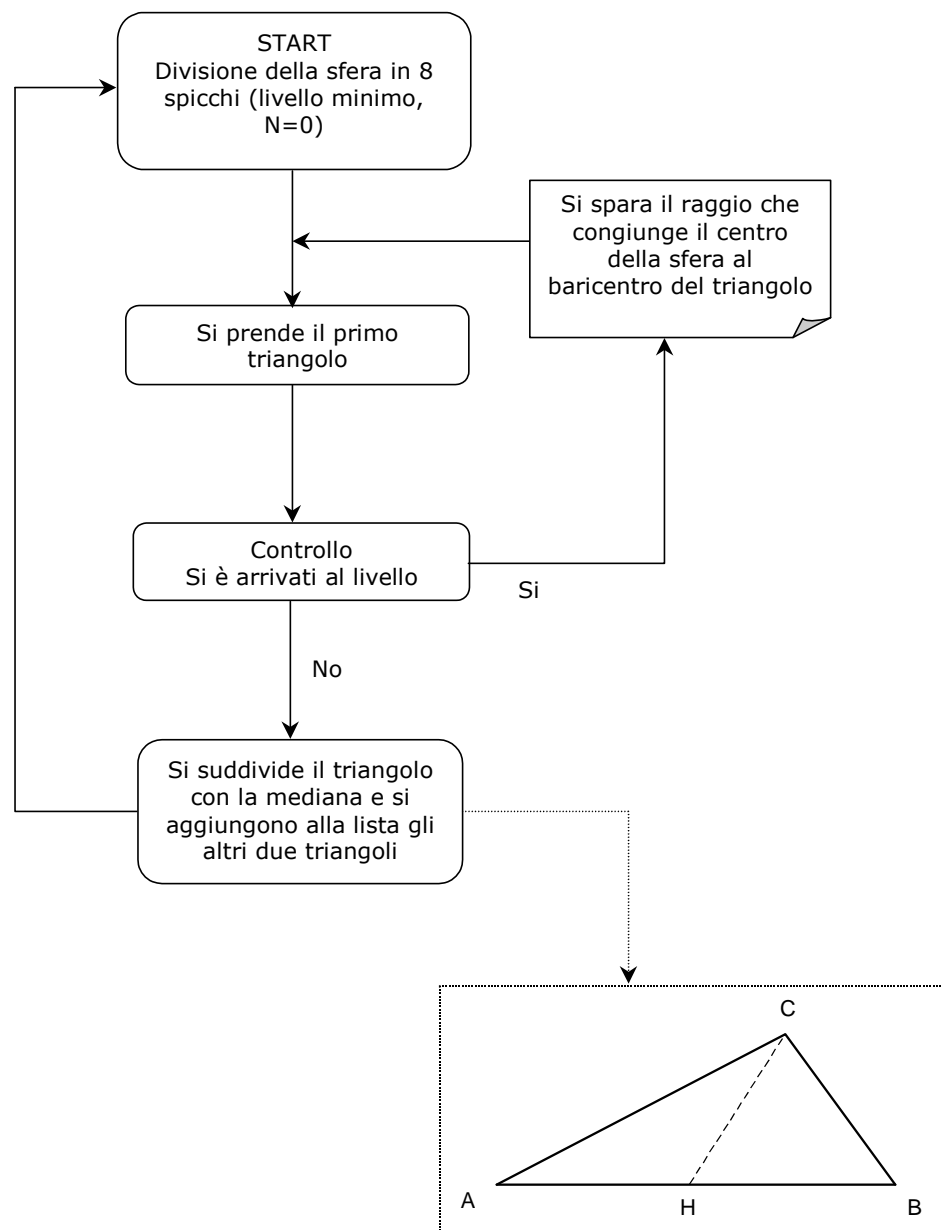


Figura 3.2 – Generazione delle piramidi

A questo punto viene riefettuato il controllo sul raggiungimento del livello di suddivisione desiderato; si ripete poi il ciclo fino a quando non sarà raggiunto il livello di suddivisione impostato.

Il numero di piramidi sparate P dipende quindi dal livello di suddivisione N secondo la seguente relazione:

$$P = 8 \cdot 2^N$$

N	P	N	P	N	P
0	8	6	512	12	32768
1	16	7	1024	13	65536
2	32	8	2048	14	131072
3	64	9	4096	15	262144
4	128	10	8192	16	524288
5	256	11	16384	17	1048576

Si ottiene quindi un Pyramid Tracing perché la sfera di partenza è esattamente divisibile in un qualsiasi numero, appartenente alla serie 8×2^N , di triangoli (curvi) di ugual area, mentre ciò non è esattamente vero nel caso del Cone Tracing. Si otterrebbero infatti delle sovrapposizioni fra i cerchi intersezione fra il cono con centro di generazione il centro della sfera e la sfera stessa. Poiché ogni sezione del cono (o della piramide) identifica una certa quantità di energia, le eventuali sovrapposizioni comporterebbero la creazione di zone in cui la quota di energia iniziale ha molteplicità due.

Questo non accade con la suddivisione della superficie sferica per triangoli che risolve anche il problema della scelta fra l'approccio pseudocasuale e quello deterministico a favore di quest'ultimo. Nel caso della sorgente omnidirezionale, per esempio, cioè con il rapporto energia emessa/superficie costante su tutta la superficie sferica, si ha la certezza di avere sparato omogeneamente sulla completezza della superficie sferica anche con un numero relativamente basso di raggi. E tanti più raggi si spareranno, in tanti più triangolini di ugual area, e quindi ugual contributo energetico sarà suddivisa la sfera, ottenendo una reale miglior accuratezza nei risultati, piuttosto che una maggior convergenza statistica degli stessi.

In realtà è stata prevista una leggera sovrapposizione fra piramidi adiacenti, quantificabile in strisce di larghezza pari a 2ε :

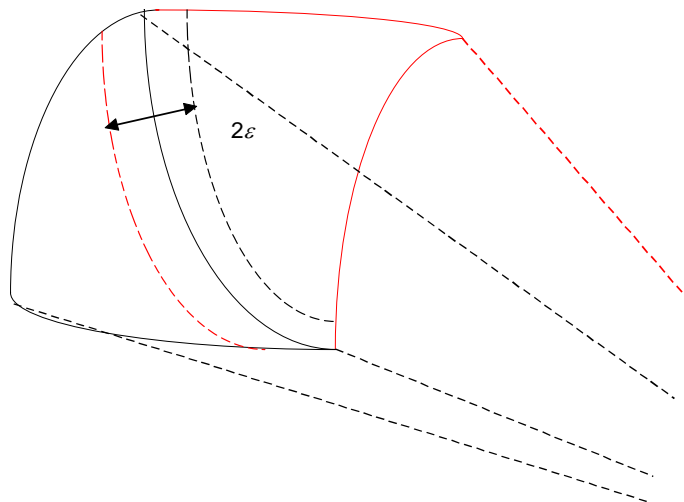


Fig. 3.3 - Sovrapposizione fra piramidi adiacenti.

questo per evitare eventuali buchi causati da errori numerici. In tal modo, però, se un ricevitore si trova proprio nell'ombra della proiezione di tale sovrapposizione, può capitare che riceva due contributi invece di uno. Ciò può accadere abbastanza spesso sull'onda diretta, ove l'errore è più grave.

Come unico inconveniente di tale controllo si è verificato che, nel caso di mappature molto fitte ed utilizzando slice temporali di grandezza notevole (0.1 s), può capitare che un ricevitore non riceva neppure il contributo che gli è dovuto. D'altronde tale errore è facilmente identificabile da una mappatura dell'SPL (ove si verificherà un minimo inspiegabile) e correggibile (basterà spostare anche di poco il ricevitore in questione).

In più, per evitare dannose situazioni di eccessiva simmetria è previsto un piccolissimo spostamento automatico della posizione della sorgente ottenuto incrementando o decrementando di qualche millimetro le sue coordinate.

3.2 Ricerca degli impatti con pareti e ricevitori

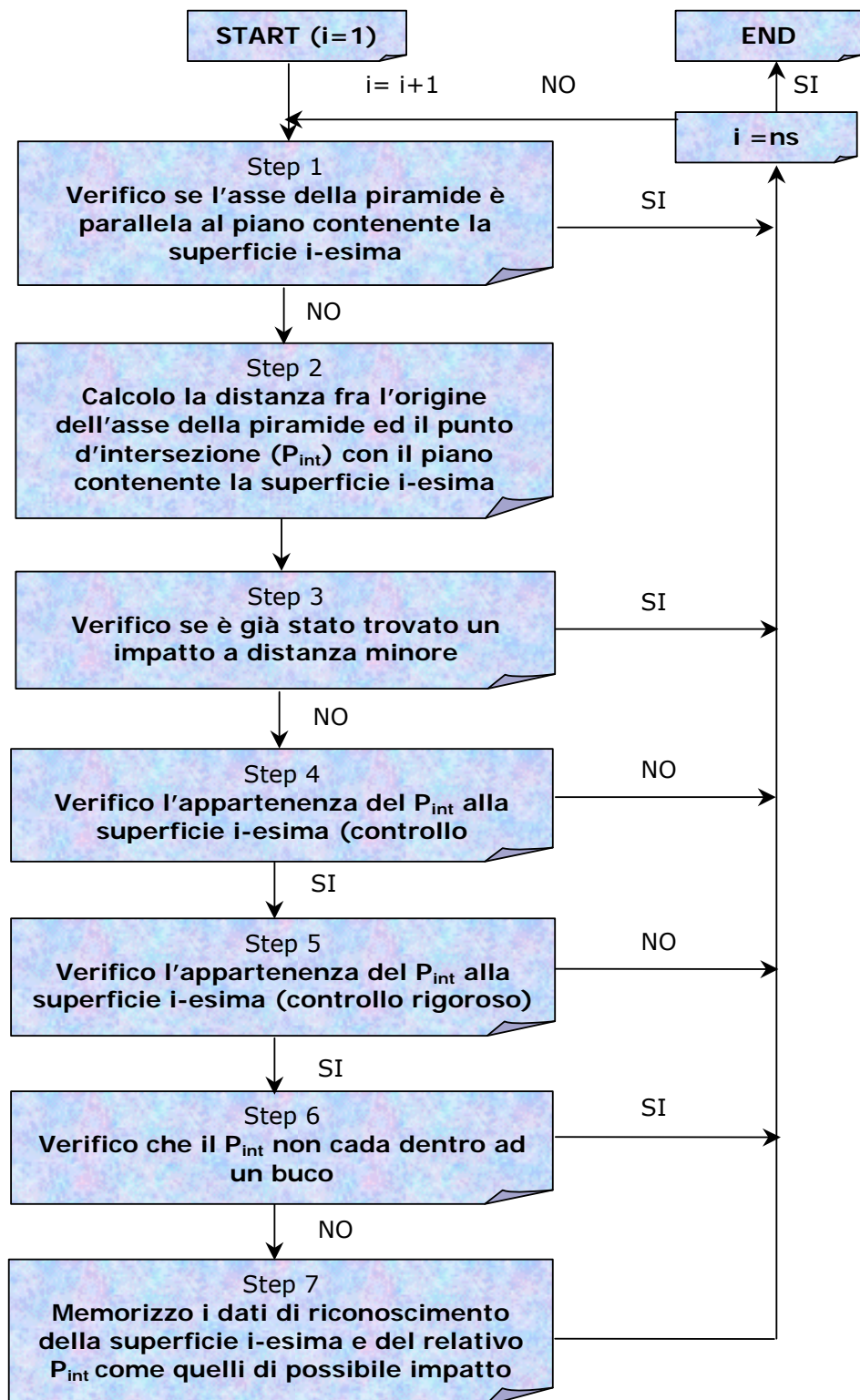


Fig. 3.4 – Ricerca degli impatti

Per ogni triangolo in cui è stata suddivisa la sfera, dalla sorgente viene sparata una piramide con vertice coincidente col centro della sorgente e con asse passante per il baricentro di ciascun triangolo, i vertici dei quali sono tutti punti appartenenti alla superficie sferica. Il vettore che rappresenta l'asse della piramide viene fatto viaggiare

alla velocità del suono nella direzione iniziale, quindi, al primo impatto con un ostacolo, verrà riflesso secondo le leggi dell'ottica geometrica (ipotesi di lavoro verificata nel campo delle frequenze audio) e viaggerà per tutto l'intervallo di tempo che l'utente vuole studiare.

La ricerca dell'impatto dei raggi sonori, fra tutte le routine che intervengono nel Pyramid Tracing, è indubbiamente quella più complessa e onerosa in termini di tempo.

Onerosa in termini di tempo nel senso che viene "eseguita" un gran numero di volte: si tratta infatti di quella routine, che nella fase di ricerca di un impatto viene percorsa tante volte quante sono le superfici in gioco.

Complessa, invece, non per il tipo di determinazione matematica che viene richiesta, (si tratta infatti di trovare l'intersezione di una retta con un piano), ma per il sistema di controlli "logici" che devono segnalare al calcolatore se l'intersezione del raggio con il piano di una generica superficie può corrispondere all'impatto cercato oppure no.

Al termine, ripetuta questa serie di controlli per ogni superficie in gioco, la routine restituisce la superficie cercata e le coordinate del punto di impatto.

Nel momento in cui si entra nella routine, del raggio si conoscono l'origine (x_0, y_0, z_0) che può essere l'ultimo impatto subito, o la sorgente, la direzione e il verso, attraverso i coseni direttori $\alpha, \beta,$ e γ . L'intersezione del raggio con il piano di una generica superficie si ottiene sfruttando l'equazione di una retta in forma parametrica:

$$x = x_0 + \rho\alpha \quad y = y_0 + \rho\beta \quad z = z_0 + \rho\gamma$$

dove $x_0, y_0, z_0,$ e α, β, γ assumono i valori già noti, mentre il parametro $\rho,$ da calcolare, si deduce dall'equazione:

$$a(x_0 + \rho\alpha) + b(y_0 + \rho\beta) + c(z_0 + \rho\gamma) + d = 0$$

che si ottiene andando a sostituire x, y e z nel piano espresso in forma cartesiana di cui si conoscono tutti i parametri.

Se la retta, che rappresenta il raggio, non è parallela al piano cui appartiene la superficie, ovvero se:

$$a\alpha + b\beta + c\gamma \neq 0$$

il valore di ρ è dato dall'espressione:

$$\rho = \frac{(ax_0 + by_0 + cz_0 + d)}{(a\alpha + b\beta + c\gamma)}$$

La relazione rappresenta il primo controllo previsto dalla routine per accertare se la superficie data può essere effettivamente quella su cui impatta il raggio oppure no; in caso negativo la superficie è scartata e il calcolatore riparte dall'inizio della routine con una nuova superficie, in caso affermativo si prosegue con le successive determinazioni e verifiche.

Questo meccanismo si ritrova al termine di ogni controllo; se una superficie riesce a soddisfare anche l'ultimo di questi controlli, viene memorizzata come la possibile superficie di impatto del raggio, il ciclo interno della routine si chiude, ed il calcolatore passa a considerare un'altra superficie.

Riprendendo l'analisi dettagliata della routine, si osserva che il parametro ρ corrisponde alla distanza percorsa dal raggio per giungere ad intersecare il piano della superficie data dal momento che α , β , e γ individuano un vettore unitario (per definizione infatti

$$\alpha^2 + \beta^2 + \gamma^2 = 1$$

Scatta allora un secondo controllo: se questa distanza risulta negativa (per cui il raggio si allontana dalla superficie) o se è già stato memorizzato un possibile impatto ad una distanza inferiore, la superficie in questione viene scartata.

Proseguendo ancora, si incontrano uno dopo l'altro tre controlli aventi la funzione di verificare se l'intersezione trovata è esterna o interna al contorno della superficie. Il primo, sommario ma molto più veloce del terzo ed ultimo, permette di scartare quasi sempre la superficie se non viene soddisfatta questa condizione di appartenenza; in particolare si verifica se le coordinate x , y e z del punto di intersezione rimangono comprese entro i valori minimo e massimo che rispettivamente caratterizzano le coordinate dei vertici della superficie.

Viene quindi verificato il soddisfacimento delle seguenti condizioni:

$$x_{\min} \leq x \leq x_{\max}$$

$$y_{\min} \leq y \leq y_{\max}$$

$$z_{\min} \leq z \leq z_{\max}$$

con:

$$x_{\max} = \max(x_1, x_2, x_3, x_4) + \varepsilon \quad x_{\min} = \min(x_1, x_2, x_3, x_4) - \varepsilon$$

$$y_{\max} = \max(y_1, y_2, y_3, y_4) + \varepsilon \quad y_{\min} = \min(y_1, y_2, y_3, y_4) - \varepsilon$$

$$z_{\max} = \max(z_1, z_2, z_3, z_4) + \varepsilon \quad z_{\min} = \min(z_1, z_2, z_3, z_4) - \varepsilon$$

Il termine ε è una costante, posta pari a 10^{-9} m, che copre gli inevitabili errori di troncamento sui numeri con molti decimali; si tratta di un accorgimento indispensabile tenendo conto che quasi sempre una delle coordinate rimane costante per tutta la superficie. Proprio per questo sono sempre poche le superfici che pur non soddisfacendo la condizione di appartenenza, riescono ad evadere questo controllo.

La definizione delle coordinate max e min per ogni superficie in gioco, avviene prima di entrare nella routine di tracciamento dei raggi sonori, in modo da non pesare sui tempi di calcolo.

A questo punto si verifica se la superficie trovata è un rettangolo con tutti i lati paralleli o perpendicolari ai piani ortogonali: se si passa al test delle finestre, altrimenti passa al test di controllo delle aree che, solo se soddisfatto, invierà la superficie al test delle finestre.

Se il punto non appartiene alla bounding box certamente non può appartenere neppure alla superficie che in essa è completamente contenuta e il test termina con esito negativo (non c'è impatto, e quindi riflessione, con questa superficie).

Se la bounding box degenera nella superficie stessa, so già che il punto di impatto appartiene alla superficie, altrimenti devo effettuare un ulteriore controllo che comunque vale solo per superfici definite da poligoni convessi.

Il terzo controllo di appartenenza, rigoroso ma decisamente più lento

del precedente, per arrivare ad esprimere un giudizio definitivo, si basa su di un confronto di aree. Vengono infatti calcolate le aree dei triangoli A_{1-2} , A_{2-3} , A_{3-4} e A_{4-1} che il punto di intersezione individua rispetto ai lati della superficie, ed è ammessa l'appartenenza se si verifica che:

$$A_{1-2} + A_{2-3} + A_{3-4} + A_{4-1} \leq A_{sup} + \varepsilon$$

dove A_{sup} è l'area della superficie e ε , posto pari a 10^{-6} m^2 , assume lo stesso significato visto in precedenza.

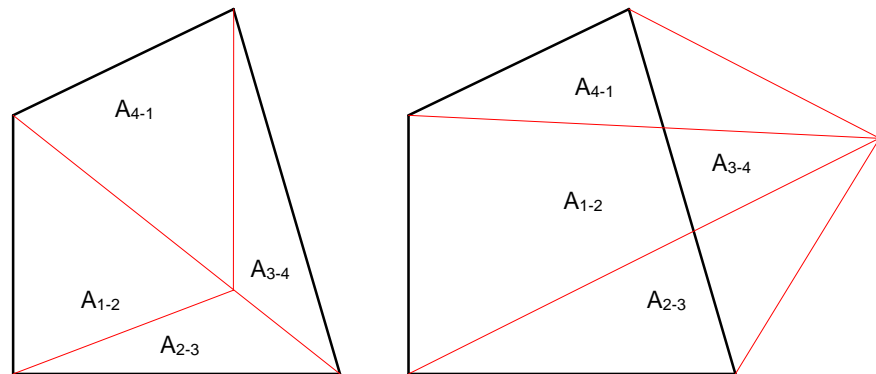


Figura 3.5 - Controllo di appartenenza alla superficie.

Se (x, y, z) è il punto di impatto del raggio sul piano della superficie, e se (x_m, y_m, z_m) e (x_n, y_n, z_n) sono due vertici generici di quest'ultima, l'area A_{m-n} del triangolo individuato è calcolata con l'espressione:

$$A_{m-n} = \frac{l}{2} \cdot \sqrt{v_x^2 + v_y^2 + v_z^2}$$

dove v_x , v_y e v_z rappresentano le componenti del prodotto vettoriale dei vettori che collegano il punto di intersezione con i due vertici considerati, per cui risulta:

$$v_x = (y_m - y)(z_n - z) - (y_n - y)(z_m - z)$$

$$v_y = (x_n - x)(z_m - z) - (x_m - x)(z_n - z)$$

$$v_z = (x_m - x)(y_n - y) - (x_n - x)(y_m - y)$$

L'area delle superfici è ovviamente calcolata una volta per tutte all'inizio del programma, procedendo sempre con lo stesso metodo, dopo aver scomposto la superficie in triangoli. Affinchè il tutto funzioni correttamente devono essere sempre utilizzate superfici convesse.

Constatata l'appartenenza alla superficie della intersezione data, interviene un ultimo controllo con il compito di verificare se l'impatto avviene in corrispondenza di eventuali buchi, porte o finestre; nel caso venga incontrato un buco la superficie viene immediatamente scartata.

La presenza di eventuali buchi (quanti e quali) è rapidamente verificata "interpellando" un indice predisposto in fase di caricamento dei dati; il controllo di appartenenza avviene esattamente nello stesso modo visto nel caso di superfici senza buchi.

Se è stata incontrata una porta o una finestra, questa viene memorizzata con le sue caratteristiche di assorbimento.

La superficie che giunge a questo punto, viene memorizzata dal calcolatore, con tutte le informazioni del caso, come possibile superficie di impatto del raggio (sono quindi scartati i dati di qualsiasi altra superficie precedentemente memorizzata).

La ricerca dell'impatto si conclude dopo aver preso in considerazione anche l'ultima superficie in gioco; quella che, a questo punto, è "ricordata" dal calcolatore come possibile elemento di impatto per il raggio, rappresenta la superficie effettivamente cercata.

Si possono presentare due situazioni anomale:

- non è stato trovato nessun impatto valido;
- l'impatto trovato è condiviso da più superfici.

Il primo caso si verifica, ovviamente, se lo spazio descritto presenta delle aperture anche di piccola entità che vengono mantenute tali nel modello numerico sottoposto ad elaborazione.

In particolare, il valore della costante ε che corregge l'area delle superfici nel controllo rigoroso di appartenenza, è definito per tentativi in modo da rendere minimo il numero dei non impatti. Nel caso di sale disegnate correttamente con Ramsete CAD non si hanno errori di questo tipo.

La seconda delle anomalie contemplate si verifica con una certa frequenza solo se ci sono delle superfici anche solo parzialmente sovrapposte, altrimenti diventa estremamente improbabile, perché in tal caso il raggio dovrebbe centrare esattamente uno spigolo (si parla infatti di "errori di spigolo").

In realtà questi errori di spigolo si manifestano con una certa facilità nel caso di impatti molto prossimi allo spigolo in comune a due superfici adiacenti in quanto il test di appartenenza è approssimato a meno di ε .

Ci possono essere due tipi di errori di spigolo: quando il raggio impatta il punto intersezione di due oppure di tre superfici.

Nel primo caso lo si fa riflettere specularmente su entrambe le pareti.

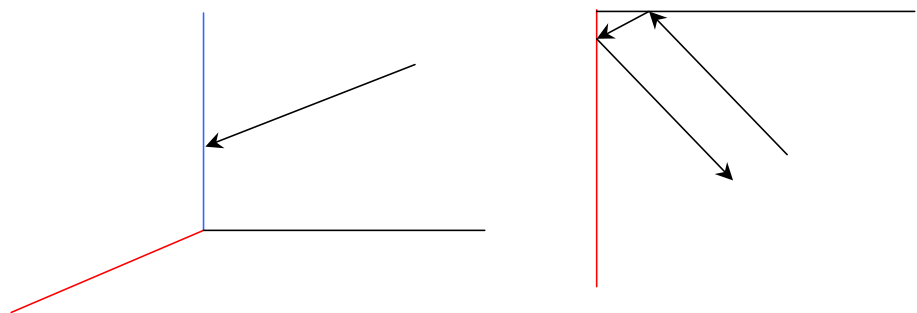


Figura 3.6 - Errore di spigolo (1° tipo: intersezione fra due piani).

Nel secondo caso è sensato riflettere il raggio cambiando semplicemente il verso e non la direzione del vettore che lo rappresenta.

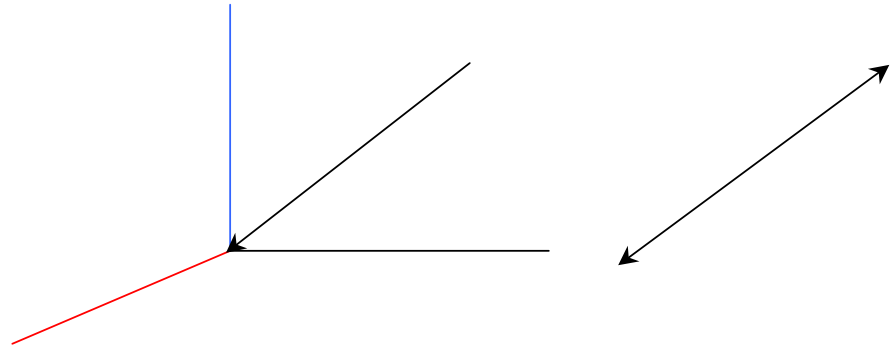


Figura 3.7 - Errore di spigolo (2° tipo: intersezione fra tre piani).

Durante il suo percorso, oltre ad ostacoli, la piramide di cui il vettore è l'altezza, potrà incontrare anche ricevitori, che sono schematizzati come dei punti nello spazio predisposti dall'utente. Per sapere quali ricevitori ogni piramide abbia impattato e per ciascuno di essi quante volte, occorre operare ad ogni rimbalzo un test di impatto sui ricevitori.

Si congiunge il ricevitore alla sorgente immagine corrispondente all'ultima parete impattata, si calcola il piano perpendicolare a tale retta e passante per il ricevitore e si costruisce il cerchio circoscritto al triangolo avente come vertici i punti appartenenti a tale piano e alla piramide. Se il punto intersezione fra la retta ed il piano è esterno al cerchio allora l'impatto non è avvenuto; se invece è interno, per determinare se l'impatto è avvenuto o no, si ricorre al test delle aree dei triangoli come prima.

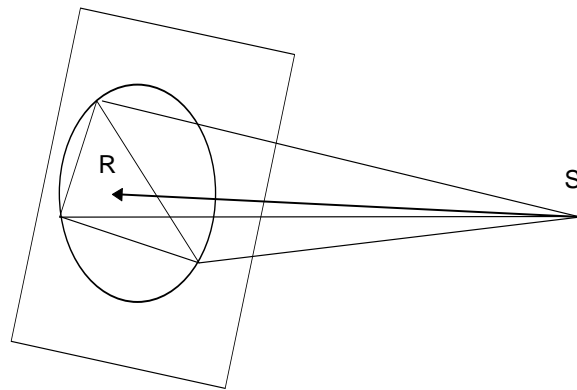


Figura 3.8 - Costruzione del cerchio circoscritto.

L'altro aspetto che differenzia il tracciatore piramidale di Ramsete dagli altri metodi di Beam Tracing è la possibilità di dichiarare "obstructing" alcune superfici (o tutte). Quando una piramide colpisce una di queste superfici, vengono attivati ulteriori controlli, onde verificare se dietro di essa si trova un ricevitore. In tale caso, si calcola il contributo sullo stesso fornito dall'onda sonora che ha attraversato la superficie (in base al potere fonoisolante della stessa). Si verifica poi se la superficie presenta bordi liberi, ed in caso affermativo viene portato un ulteriore contributo di energia al ricevitore a partire da ciascun bordo libero, calcolato con la relazione di Kurze-Anderson:

$$L_{diff} = L_{dir} - 5 - 10 \log \left(\frac{\sqrt{2\pi|N|}}{\tanh \sqrt{2\pi|N|}} \right)$$

in cui L_{dir} è il livello diretto, che giungerebbe al ricevitore se la parete "obstructing" non ci fosse, ed N è il n° di Fresnel, dato da:

$$N = \frac{2}{\lambda} \cdot \delta = \frac{2 \cdot f}{c_0} \cdot \delta$$

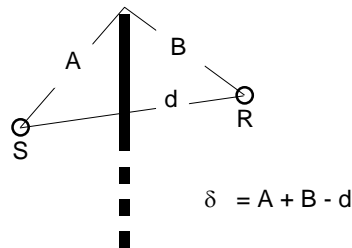


Figura 3.9 - Traiettoria dell'onda diretta e del raggio diffratto.

La seguente figura mostra le traiettorie dei raggi che vengono tracciati in questi casi (diffratti ed attraversante). L'algoritmo di ricerca dei bordi liberi non è ricorsivo, pertanto non viene considerata l'onda diffratta dal bordo sinistro del pannello laterale.

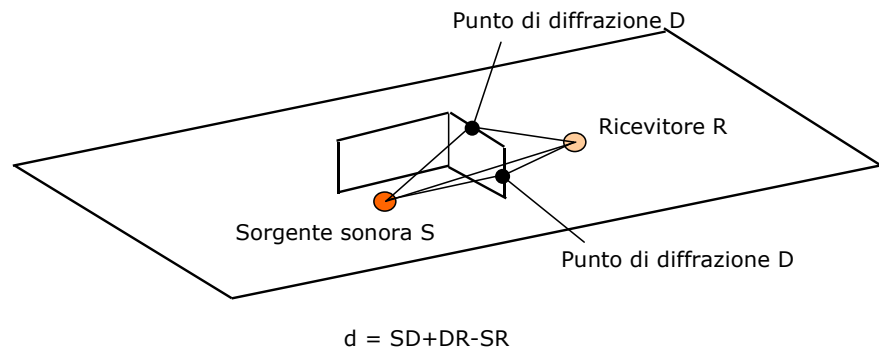


Figura 3.10- Diffrazione dei bordi liberi più energia passante attraverso la barriera.

Per verificare automaticamente che un lato di una parete obstructing sia diffrangente si procede così:

- si posizionano, sul prolungamento delle mediane, quattro punti di riferimento detto M_1 , M_2 , M_3 ed M_4 ad 1 cm di distanza dai lati della superficie dichiarata obstructing;
- si traccia la congiungente fra la sorgente S ed il primo sensore (SM_1) e quella fra il sensore stesso ed il ricevitore R ($M1_R$) e si verifica, con i controlli già visti, che non ci siano ostacoli;
- se la risposta è no, allora quel lato viene considerato diffrangente e viene applicata la relazione di Kurze-Anderson; nel caso opposto tale contributo non viene computato;
- si ripete lo stesso ciclo anche per gli altri tre sensori.

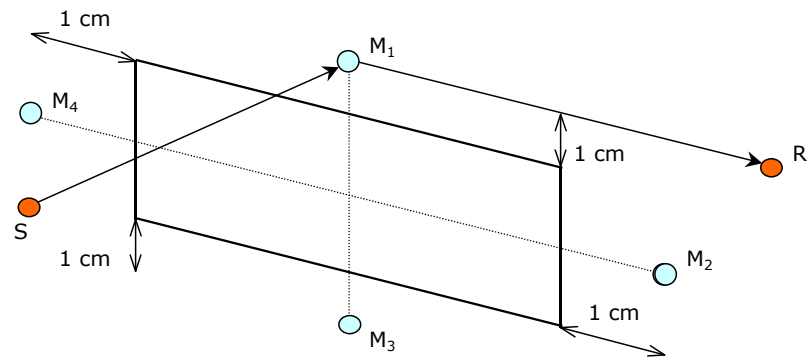


Figura 3.11 - Localizzazione dei sensori per la ricerca del bordo libero nel calcolo della diffrazione.

La distanza $A+B$ viene calcolata proiettando sul bordo libero (diffrangente) il punto intersezione della congiungente la sorgente ed il ricevitore con lo schermo. Dopodichè viene calcolato il cammino effettivamente percorso in due tratti: il primo che va dalla sorgente al bordo dello schermo più il secondo che parte da quest'ultimo ed arriva al ricevitore.

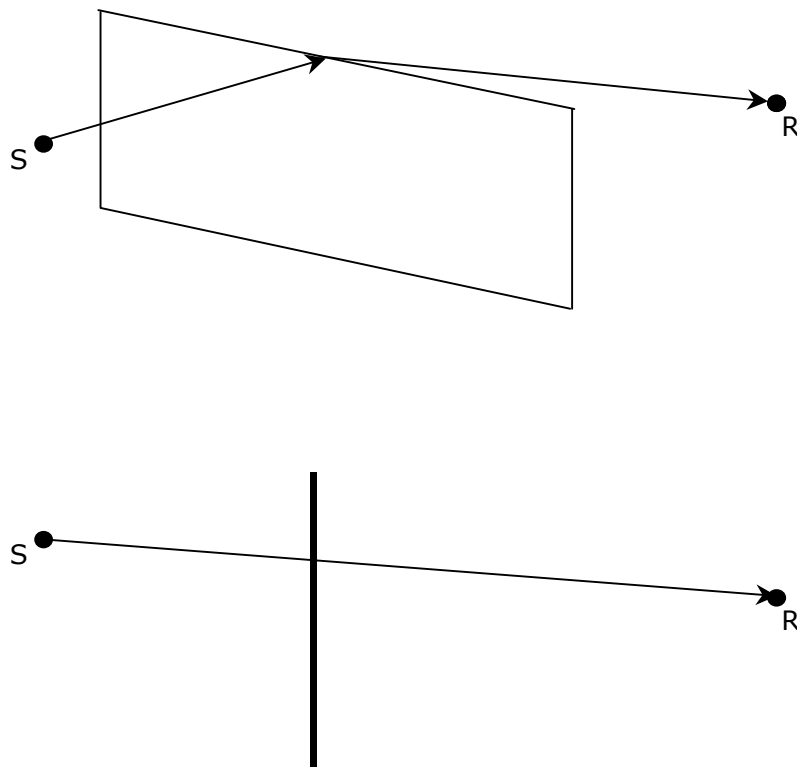


Figura 3.12 - Calcolo del minimo cammino del raggio diffratto.

Così facendo si calcola veramente il cammino percorso dall'onda, che risulta essere anche il minimo possibile.

Grazie a queste possibilità, l'algoritmo di tracciamento è in grado di affrontare lo studio della propagazione sonora anche in spazi geometricamente complessi, in presenza di schermature parziali o totali. E' pure possibile valutare l'efficacia di cabine per la sosta degli operai, oppure prevedere l'emissione di rumore verso ricevitori esterni all'edificio.

Inoltre il modello è utilizzabile anche nell'ambiente esterno, ad

esempio per la verifica di schermature attorno a sorgenti fisse, purchè la distanza dei ricevitori non sia così grande da dover considerare, oltre all'attenuazione dell'aria, anche effetti atmosferici che portino alla curvatura dei raggi sonori.

3.3 Estinzione delle piramidi

Esistono due modi per decidere quando un raggio debba essere considerato "estinto", ovvero non si debba più seguire la sua storia. Si può decidere di seguirlo per tutto il tempo dell'intervallo impostato dall'utente, oppure fino a quando la sua energia fosse nulla. Dal punto di vista computazionale c'è qualche differenza di costo che rende favorevole un approccio o l'altro a seconda dei casi. Se si imposta un intervallo di tempo molto lungo (4-5 s), usando il primo approccio, può verificarsi il caso che si debba continuare a seguire, e quindi a tracciare, un raggio con energia nulla, spendendo quindi inutilmente tempo di calcolo. Viceversa se si vogliono studiare tempi relativamente brevi (1-2 s), seguendo il secondo metodo, ci si potrà trovare a dover continuare inutilmente il tracciamento di raggi anche oltre l'intervallo di tempo di interesse, poiché capiterà che la maggior parte dei raggi non si sarà estinta entro l'intervallo da considerare.

Il programma segue l'approccio più trasparente per l'utente, che si è ritenuto debba poter scegliere l'intervallo di tempo da studiare e pagare in termini di tempo di calcolo solo in relazione ad esso; si evita così di dover studiare sempre ed in ogni situazione il caso peggiore, seguire il raggio fino alla sua estinzione, anche quando, fin dall'inizio, si sa che l'interesse sarà concentrato sui primi istanti di tempo dall'emissione dello stesso.

Il controllo in realtà viene fatto sullo spazio s percorso dal raggio dal momento della sua emissione. Conoscendo infatti il tempo a sua disposizione t e la velocità del suono c_0 , considerata costante ed uguale a 340 m/s alla temperatura di riferimento di 20° C, è immediato calcolare quanto potrà essere il cammino complessivo del raggio tramite la semplice formula:

$$s = c_0 t$$

Esiste anche un altro modo per estinguere un raggio; infatti è possibile scegliere, mediante il valore assegnato al parametro "history", per quante riflessioni il programma dovrà seguire il percorso di un raggio. Il valore "0" consentirà di calcolare solo l'onda diretta, mentre "-1" permetterà di seguire il raggio fino a quando gli altri controlli non ne decreteranno la fine.

3.4 Attenuazione dell'energia

Una sorgente sonora irradia una certa potenza P che si propaga nel mezzo con l'avanzare delle onde sonore. In un campo sonoro se $P[W]$ è la potenza trasportata da un'onda con fronte di superficie $S[m^2]$, la quantità di energia che fluisce nell'unità di tempo attraverso l'unità di area del fronte d'onda si definisce intensità I ed è data dall'espressione:

$$I = \frac{P}{S} \left[\frac{W}{m^2} \right]$$

Per un'onda piana o per un fronte di un'onda sferica sufficientemente lontano dalla sorgente da poter essere considerato piano, si dimostra che l'intensità è proporzionale al quadrato della pressione sonora. Per una sorgente sonora, che irradia uniformemente in tutte le direzioni l'energia, viene diffusa su superfici sferiche il cui centro è la sorgente sonora ed il cui raggio cresce con il propagarsi del suono; se la sorgente ha una potenza sonora P e il mezzo non è dissipativo, l'intensità ad una distanza di r metri dalla sorgente è data dalla:

$$I = \frac{P}{4\pi r^2} \left[\frac{W}{m^2} \right]$$

In un'onda sferica pertanto l'intensità diminuisce con il quadrato della distanza.

Dalla precedente relazione fra potenza ed intensità, dividendo entrambi i membri per l'intensità di riferimento I_0 e ricordando che $I_0 = P_0 / 1 \text{ m}^2$ si ha:

$$\frac{I}{I_0} = \frac{P}{P_0} \cdot \frac{1 \text{ m}^2}{4\pi \cdot r^2}$$

e passando ai livelli in dB:

$$10 \log \frac{I}{I_0} = 10 \log \frac{P}{P_0} - 10 \log 4\pi \cdot r^2$$

per cui si ha:

$$L_I = L_w - 10 \log 4\pi \cdot r^2$$

E' stata così considerata la divergenza sferica.

Il livello di intensità sonora è anche influenzato dalle condizioni ambientali e dalla direttività della sorgente, per cui l'equazione precedente assume la seguente forma:

$$L_I = L_w + DI_\theta - A_e - 10 \log 4\pi \cdot r^2$$

dove DI_θ è l'Indice di direttività nella direzione θ , espresso da $10 \log Q_\theta$, essendo Q_θ il fattore di direttività, rapporto fra l'intensità sonora alla distanza r nella direzione θ e l'intensità sonora media, alla medesima distanza, su tutte le direzioni.

A_e considera l'assorbimento di tutte le pareti (A_p) su cui impatta ogni singolo raggio, dell'eventuale potere fonoisolante di qualche schermo A_o , che va dichiarato obstructing, e dell'assorbimento dell'aria A_β :

$$A_e = A_p + A_o + A_\beta$$

A_p viene calcolato sottraendo, ad ogni rimbalzo, dall'energia attuale di ciascun raggio una quota di energia proporzionale al coefficiente di fonoassorbimento del materiale α di cui è costituita la parete su cui è avvenuto il rimbalzo stesso. Sarà quindi:

$$A_p = 10 \log \prod_i (1 - \alpha_i)$$

Tutte le volte che il percorso del raggio sarà ostacolato da uno schermo dichiarato obstructing costituito da un materiale di cui nel Material Manager siano inseriti anche i coefficienti di fonoisolamento, che altrimenti vengono considerati infiniti, non tutta l'energia rimasta al raggio dopo l'attenuazione subita a causa del potere fonoassorbente dello schermo (α) viene riflessa; una parte di essa, infatti, diminuita del potere di fonoisolamento dello schermo stesso (R), viene associata ad un nuovo raggio che continua oltre lo schermo la traiettoria del raggio incidente, come in figura.

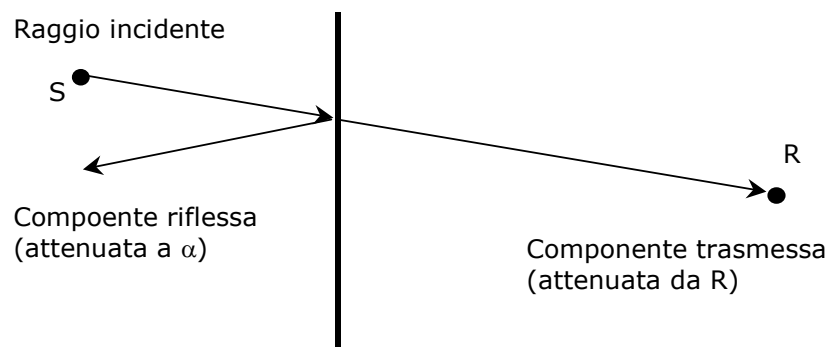


Figura 3.13 - Attenuazione subita da un raggio nell'impatto con una parete di materiale fonoassorbente e fonoisolante.

L'assorbimento dell'aria, viene calcolato alla temperatura di riferimento, considerata costante, di 20 °C. Essa varia con l'umidità e con la frequenza secondo la seguente relazione:

$$A_\beta = 1.7 \cdot \frac{f^2 r}{\Phi} \cdot 10^{-8}$$

dove:

- f = frequenza centrale di banda [Hz]
- Φ = umidità relativa %
- r = distanza in m.

I gradienti di temperatura e di velocità del vento, nonché l'effetto radente, nel caso di problemi in campo aperto possono dar luogo a variazioni anche notevoli sui risultati; questo programma non ne tiene conto proprio perché è stato progettato principalmente per prevedere il campo sonoro in ambienti chiusi o in spazi aperti di dimensioni non troppo elevate (intorno al centinaio di metri, dove appunto, tali fenomeni si possono trascurare).

3.5 Correzione della coda sonora

La correzione della coda sonora è di tipo moltiplicativo. La trattazione che segue si basa dapprima sull'ipotesi di campo strettamente Sabiniano, per poi essere estesa anche a campi non Sabiniani.

Ambiente Sabiniano

Il numero di impatti su un ricevitore nell'unità di tempo $n(t)$ può essere descritto matematicamente dalla seguente espressione:

$$n(t) = \frac{4\pi \cdot c_0^3 t^2}{V}$$

dove:

$n(t)$ = numero di impatti per unità di tempo
 c_0 = velocità del suono
 V = volume.

Nel caso di un tracciatore di fasci divergenti, il numero di impatti su di un ricevitore nell'unità di tempo segue la formula, solo quando l'apertura del fascio è ancora minore delle dimensioni delle superfici della sala; quando la base del fascio diventa, per effetto della divergenza, più grande della sala, tutti i ricevitori vengono impattati dopo ogni riflessione del fronte d'onda, in modo che il numero di impatti per unità di tempo tende ad un valore costante:

$$n(t \rightarrow \infty) = \frac{c_0 N}{l_{cm}}$$

dove:

c_0 = velocità del suono
 N = numero totale di piramidi tracciate
 l_{cm} = libero cammino medio.

A questo punto si può introdurre il concetto di Tempo Critico t_c : che rappresenta il punto di ideale separazione fra la prima parte della coda sonora, in cui vengono rilevate correttamente tutte le sorgenti immagine, e la parte tardiva, in cui il numero di arrivi di energia nell'unità di tempo è costante. Svolgendo i calcoli e ricordando che, per un ambiente perfettamente Sabiniano, vale la relazione:

$$l_{cm} = \frac{4V}{S_{tot}}$$

dove:

V = volume
 S_{tot} = superficie totale

si ottiene:

$$t_c^2 = \frac{NS_{tot}}{16\pi \cdot c_0^2}$$

introducendo, infine, il parametro adimensionale β :

$$\beta = \frac{4\pi \cdot l_{cm}^2}{S_{tot}}$$

si ottiene:

$$t_c^2 = \frac{Nl_{cm}^2}{4\beta \cdot c_0}$$

L'utilizzo di quest'ultima espressione implica l'utilizzo di due parametri ignoti invece di uno (l_{cm} e β , al posto della S_{tot}) ma entrambi più facilmente calcolabili della superficie totale "visibile" all'interno di una sala. Infatti il libero cammino medio l_{cm} è calcolato da Ramsete in base ad una statistica sull'effettiva percorrenza dei raggi sparati in asse alle piramidi, ed il coefficiente adimensionale β , dopo la taratura dell'algoritmo, è stato fissato ad un valore medio pari a 0.1. Tale valore varia con la forma della geometria della sala, e non con le sue dimensioni (nel caso di una stessa sala disegnata in due differenti scale, il valore ideale di β rimane invariato).

Il numero effettivo di impatti su un ricevitore nell'unità di tempo $n(t)$, ottenuto da un tracciatore di fasci divergenti, può essere descritto matematicamente, secondo Maercke-Martin, dalla relazione:

$$n(t) = \frac{4\pi \cdot c_0^3 t^2}{V} \cdot \left(1 - e^{-\frac{l_{cm}^2 N}{4\beta \cdot c_0^2 t^2}} \right) = \frac{4\pi \cdot c_0^3 t^2}{V} \cdot \left(1 - e^{-\frac{t_c^2}{t^2}} \right)$$

Quindi la coda sonora andrà corretta moltiplicando l'espressione per il reciproco del termine fra parentesi, e cioè:

$$\left(1 - e^{-\frac{l_{cm}^2 N}{4\beta \cdot c_0^2 t^2}} \right)^{-1} = \left(1 - e^{-\frac{t_c^2}{t^2}} \right)^{-1}$$

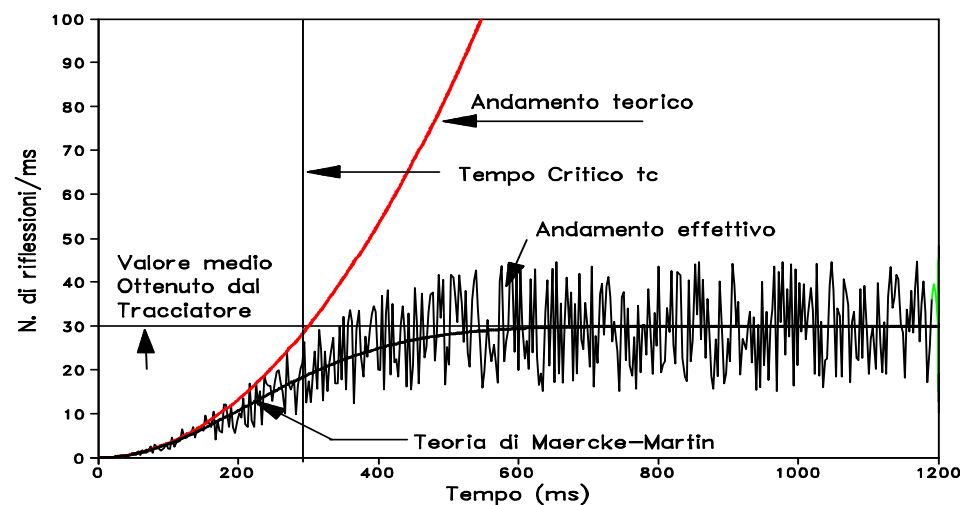


Figura 3.14 - Andamento teorico ed andamento reale di $n(t)$.

La figura sopra illustra una situazione reale, a confronto con le stime teoriche di cui sopra.

Vari autori hanno tentato di correggere i loro tracciatori di fasci basati sulle precedenti formule: Maercke e Martin usando una correzione additiva, sovrapponendo una coda statistica addizionale alla risposta all'impulso calcolata con il tracciatore di coni, mentre Naylor ha proposto di moltiplicare la coda sottostimata per un fattore di correzione tempo-variante. In Ramsete la teoria è stata ulteriormente perfezionata, e viene utilizzata una correzione di tipo moltiplicativo, in tal modo si evitano occasionali discontinuità nella densità di impatti rilevati.

In più, Ramsete corregge separatamente ciascun ricevitore; in altre parole la coda riverberante (diversa da ricevitore a ricevitore) viene

corretta dello stesso valore, ma mantiene la propa unicità. In questo modo possono essere analizzate situazioni aventi ricevitori in condizioni acustiche diverse (camere accoppiate, all'esterno, etc...) senza incorrere in errori gravi.

Per tali ragioni Ramsete non può considerarsi un modello di calcolo "ibrido"; infatti non viene effettuata alcuna distinzione fra la parte "primaria" deterministica della risposta all'Impulso e quella "tardiva" statistica, sebbene per tempi maggiori del tempo critico l'andamento del campo acustico solo in parte dal numero totale "reale" di arrivi.

In altre parole, Ramsete può lavorare anche come modello puramente deterministico: se il numero di piramidi è abbastanza elevato da far diventare maggiore il valore del tempo critico rispetto alla lunghezza della risposta all'impulso, in pratica non viene effettuata alcuna correzione ed è così possibile ottenere "esattamente" i valori dei parametri acustici più sensibili. Comunque, nei casi pratici, è sufficiente un numero molto piccolo di piramidi per ottenere una stima ragionevolmente precisa del livello di pressione sonora e del tempo di riverbero (che è ciò che si richiede per gli studi di bonifica del rumore), avendo cura che la risoluzione temporale scelta per la risposta all'impulso non sia troppo bassa.

Per queste ragioni Ramsete è incredibilmente veloce nel calcolo delle mappe dell'*SPL* anche per geometrie complesse.

La figura sotto mostra, ad esempio, il confronto fra le risposte all'impulso energetiche ottenute con un numero di piramidi molto elevato (128000) e quelle stimate con sole 256 piramidi sia in presenza di correzione della coda ($\beta = 0.1$), sia in assenza della stessa ($\beta = 0$).

Si nota come, pur trattandosi in questo caso di un ambiente fortemente non Sabiniano, e poco riverberante, la correzione della coda sia in grado di riprodurre con sole 256 piramidi la situazione che, con altre tecniche previsionali, avrebbe richiesto un numero enormemente superiore di raggi tracciati.

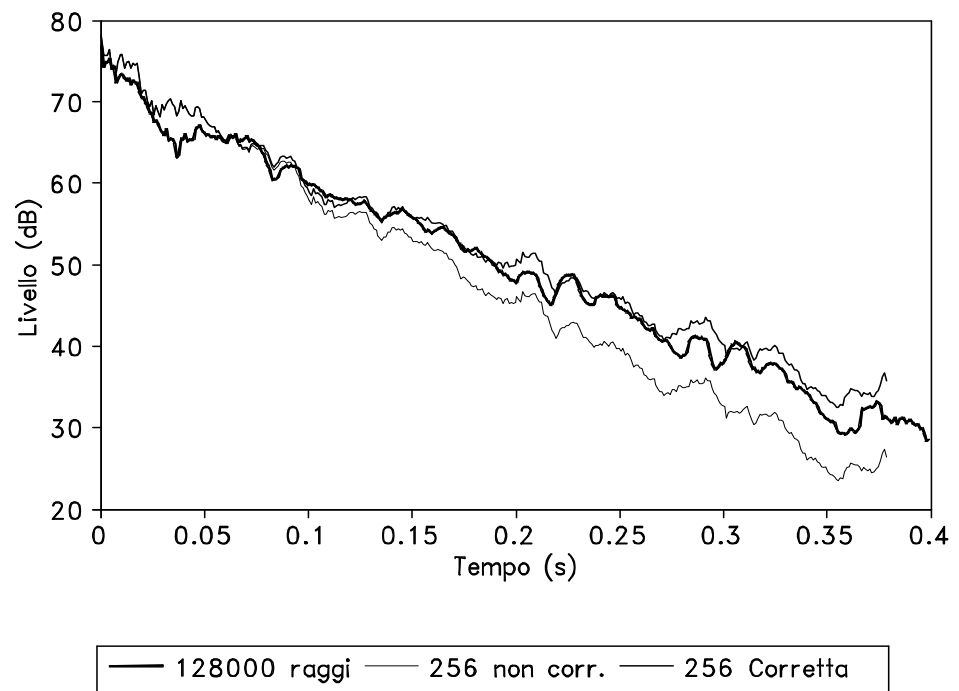


Figura 3.15 - Confronto dell'andamento della coda sonora con e senza correzione.

Fino ad ora è stata considerata l'ipotesi di Sabinianità del campo sonoro e, nonostante ciò, si è dimostrato come il modello possa funzionare correttamente anche in casi in cui questa ipotesi non sia verificata anche se, non essendo valida la teoria precedentemente illustrata, si deve usare un numero di piramidi più alto di quello che ci eravamo proposti.

Ambienti non sabiniani

Supponendo ora di essere in un ambiente non Sabiniano, si valuta come migliorare ulteriormente il modello.

Se si utilizza lo stesso tipo di correzione della coda anche in questo caso, si possono notare principalmente due effetti:

- il "vero" numero di riflessioni per unità di tempo non aumenta con t_c elevato al quadrato, ma ad un esponente $\alpha < 2$;
- il tempo critico è normalmente più basso di quello calcolato con le precedenti equazioni.

Comunque se dimensionando opportunamente il fattore di forma β si riesce ad ottenere un corretto valore di t_c e si corregge con α il grado di non Sabinianità della sala, si riescono ad ottenere risultati accurati anche con un numero di piramidi molto piccolo.

Il fattore di correzione della relazione

$$\left(1 - e^{-\frac{l_{cm}^2 N}{4\beta \cdot c_0^2 t^2}}\right)^{-1} = \left(1 - e^{-\frac{t_c^2}{t^2}}\right)^{-1}$$

nella sua forma più generale sarà quindi:

$$\left(1 - e^{-\frac{t_c^\alpha}{t^\alpha}}\right)^{-1}$$

3.6 Taratura dell'algoritmo in ambienti chiusi

Per verificare l'affidabilità del tracciatore, lo stesso è stato provato con varie tipologie geometriche che comportassero particolari difficoltà al modello numerico ed ai test di controllo. Quindi geometrie di forma marcatamente non sabiniana, strette e lunghe, a doppia cavità, rientranti su loro stesse ecc., e geometrie sabiniane per forma ma non sabiniane come distribuzione dei materiali, contenenti cioè materiali con coefficienti di fonoassorbimento anche molto diversi fra loro.

Nel primo caso un parametro molto significativo e correlato al grado di sabinianità della geometria è il libero cammino medio che viene calcolato dal programma poiché necessario per la correzione della coda riverberante. Il suo calcolo è relativamente semplice e si ottiene incrementando ad ogni rimbalzo il percorso fatto dal raggio e memorizzando il numero di rimbalzi. Questo per ciascun raggio sparato. Quindi si avrà:

$$l_{cm} = \frac{\sum_{n=1}^m l_n + \sum_{n=1}^p l_n + \dots + \sum_{n=1}^r l_n}{m + p + \dots + r}$$

dove:

l_{cm} = libero cammino medio della sala
 l_n = lunghezza del cammino del raggio fra due rimbalzi successivi
 m = numero rimbalzi del primo raggio
 p = numero rimbalzi del secondo raggio
 r = numero rimbalzi dell'ultimo raggio.

Un'ipotesi migliorativa potrebbe essere quella di calcolare separatamente il libero cammino medio dei raggi di ciascun ricevitore ed applicare così una diversa correzione della coda a seconda della posizione di ogni ascoltatore.

Allo stato attuale, può essere utile un secondo parametro statistico che, unito al libero cammino medio, può dare importanti informazioni sul grado di sabinianità geometrica della sala: la deviazione standard del libero cammino medio. Possono esistere, infatti, due sale con identico libero cammino medio, ma di forma diversa che si possono comportare acusticamente anche in modo assai differente. La deviazione standard del libero cammino medio può dare all'esperto in acustica la possibilità di discriminare i due casi e riconoscere anche il caso di equivalente sabinianità di due sale di forme diverse ma con uguale libero cammino medio e uguale deviazione standard del libero cammino medio. La deviazione standard del libero cammino medio $\sigma_{l_{cm}}$ si ottiene dalla seguente relazione:

$$\sigma_{l_{cm}}^2 = \frac{\sum_{i=1}^s d_i^2}{R} - l_{cm}^2$$

dove:

d = percorso totale di ogni singolo raggio
 $R = m + p + \dots + r =$ numero di rimbalzi complessivo
 l_{cm} = libero cammino della sala
 s = numero totale di raggi.

Tutte queste informazioni sono direttamente accessibili dato che il file risultati è di tipo ASCII:

```
[Version]
Program = Ramsete 2.0
Version = 2
Author  = Genesis
Date    = Feb 18 2001
Time    = 12:18:54

[File]
Name     = generate_file.____
Revision = 1
Author   = Genesis
Date     = Jan 27 2002
Time     = 15:09:20

[Model]
Name           = c:\programmi\Ramsete2\file creation\generate_file.ray
Microphones   = 2
Sources       = 1
```

Walls = 1
 Obstructing = 0

[Source]

Name = omni.spk
 Power = 110 110 110 110 110 110 110 110 110 110
 Position = -13 -1 5.5
 Direction = 0.00181818 0 -0.999998
 Roll angle = 0
 First band = 0 (31.5Hz)
 Last band = 9 (16KHz)
 Number = 1 of 1

[Statistics]

Pyramids traced = 2048
 Mean free path = 4.41185 m
 Standard deviation = 1.69513 m
 Mean history = 0.25293 reflections
 Pyramids escaped = 2048
 Edges hit = 0
 Processing time = 00:00:01

[Parameters]

Level = 8
 Time = 2 s
 Precision = 0.001 s
 History = 50 Reflections
 Humidity = 50 %
 Temperature = 20 C
 SoundSpeed = 343.294 m/s
 Alfa = 2
 Beta = 0.1
 Diffraction = 1
 Diffusion = On
 DiffusionMode = 7
 Format = 1.2
 ErrorLog = On
 HitLevel = 8
 RandomizeAfter = 4
 SaveAll = Off

[Microphone_1]

Position = -13 4 5.5 0.707107 0.707107 0
 Direct Wave = 0.0145642 s
 LFC = 0.332943 0.296974 0.287016 0.284456 0.280944 0.27797
 0.277476 0.278742 0.278467 0.269724
 LE = 0.00279615 0.00283372 0.00261981 0.0020792
 0.00141508 0.00115015 0.000905353 0.000511539 0.000312744
 0.000273069
 LF = 0.317909 0.289125 0.282548 0.280839 0.27768 0.274875
 0.274455 0.275756 0.275523 0.266956
 Impulse response =
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0

```

0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
3.18335e+008 3.18335e+008 3.18333e+008 3.18328e+008
3.18304e+008 3.18211e+008 3.17837e+008 3.16348e+008
3.1046e+008 2.87983e+008
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
.....(continua)

```

Con questi parametri statistici sono stati ricercati i valori ottimali di α e β contenuti nelle relazioni responsabili della correzione della coda. Si è proceduto facendo calcolare la risposta all'impulso di una sala a vari livelli di suddivisione. Notando che per livelli di suddivisione molto alti (>10) il modello di calcolo diventava in pratica un Ray Tracing deterministico, vista la selettività di ciascun fascio è stato assunto come riferimento "vero" il livello 17 e si è poi cercato di riportare la risposta all'impulso della stessa sala calcolata con livelli di suddivisione inferiori a 17 al tracciato.

Analizzando l'espressione

$$\left(1 - e^{-\frac{l_{cm}^2 N}{4\beta \cdot c_0^2 l^2}}\right)^{-1} = \left(1 - e^{-\frac{l_c^2}{l^2}}\right)^{-1}$$

si può notare come all'aumentare del livello di suddivisione la correzione applicata tenda a zero. Questo è giustificato dal fatto che tale correzione è indispensabile per bilanciare la perdita di sorgenti immagine dovuta alla situazione in cui la divergenza del fascio gli fa raggiungere sezioni paragonabili alla pianta della sala; nel caso in cui il livello di suddivisione aumenti, diminuirà la divergenza dei fasci e, di conseguenza, saranno meno le sorgenti immagine che andranno perse e quindi la correzione dovrà essere inferiore.

A livello 17 la correzione è praticamente nulla, si è quindi ottenuto un risultato "vero" nel senso che non necessita di correzione. Un'elaborazione con tale grado di precisione esige tempi troppo lunghi nel normale utilizzo del programma, quando è auspicabile un tempo di calcolo dell'ordine della decina di minuti nelle situazioni più complesse (tipico del livello 8).

Si procede quindi a caricare la risposta all'impulso a livello 17 corretta e senza alcuna correzione per verificare che praticamente siano sovrapponibili. Dopodichè sono caricate le risposte all'impulso calcolate con livelli di suddivisione inferiori con α e β come variabili. Si opera una ottimizzazione ai minimi quadrati cercando i valori di α e di β per i quali tutte le curve coincidano con quella calcolata con livello 17. Sono così trovati quei valori di α e di β che minimizzano lo scarto quadratico medio fra ciascuna risposta all'impulso e quella di riferimento.

Dopo vari tentativi è possibile scoprire come i due parametri influenzano la pendenza media della curva e la sua traslazione verso l'alto; si trova il valore 2 per α e 0.1 per β , che corrispondono bene ai valori teorici del campo sabiniano.

Per validare tali valori si è ripetuto tutto il procedimento con diverse geometrie.

Tali geometrie sono state molto utili anche per controllare l'affidabilità numerica dei controlli, affinare le tolleranze (gli ε) e seguire in generale l'operato del tracciatore, studiandone gli errori ed i tempi di calcolo in relazione alle diverse geometrie per poterne trarre utili informazioni per ottimizzazioni finalizzate alla riduzione dei tempi stessi.

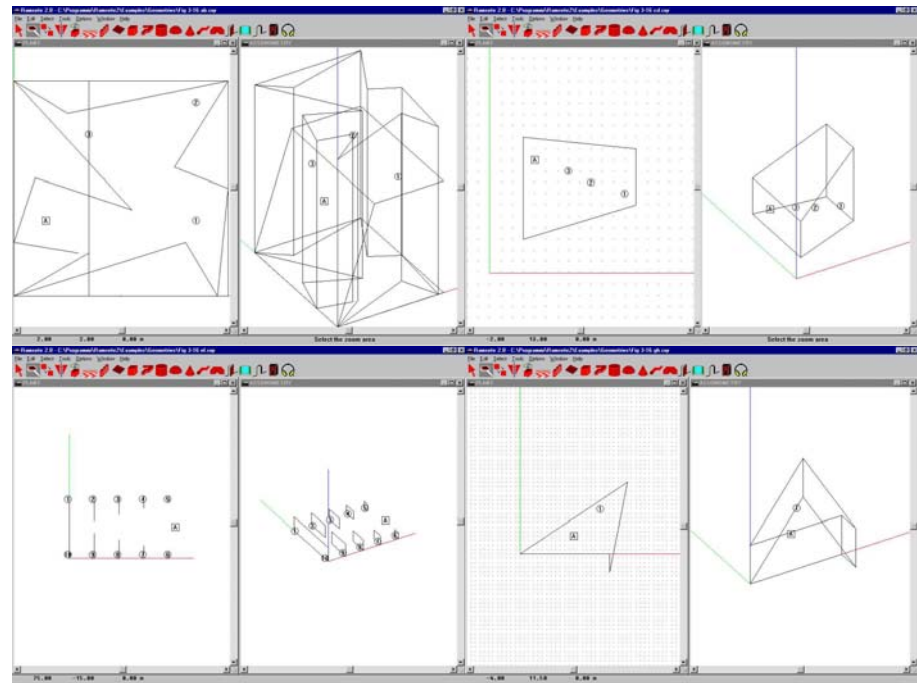


Figura 3.16 - Geometrie usate per la taratura ed i test di affidabilità del modello.

Si consideri una sala perfettamente Sabiniana. La Figura. 3.17 mostra la geometria di una classica camera riverberante; in questo caso i coefficienti di assorbimento sono gli stessi per ogni parete, in modo tale da avere un campo sicuramente Sabiniano e da poter considerare un unico ricevitore.

Il confronto viene fatto plottando sullo stesso grafico la risposta all'Impulso integrata all'indietro (Schroeder plot) in dB per la banda d'ottava centrata ad 1 kHz, calcolata con un algoritmo di Ray Tracing (128000 raggi) e con Ramsete (con tre differenti quantità di piramidi emesse).

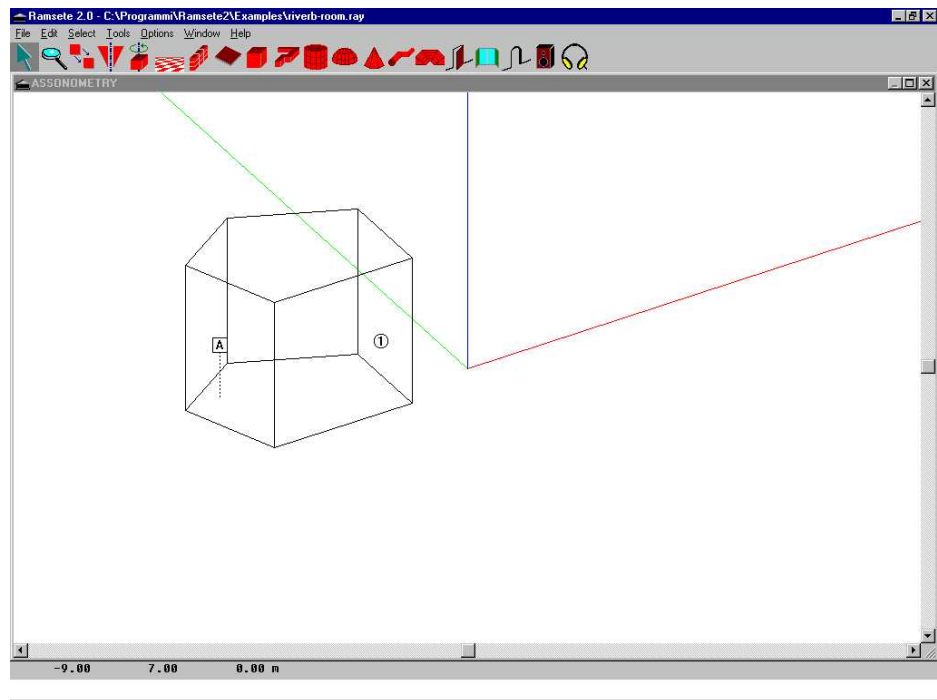


Figura 3.17 - Geometria di una camera riverberante.

In figura 3.18 il confronto viene fatto due volte: a sinistra sono riportati i risultati ottenuti con Ramsete senza correzione della coda, a destra, invece, con i valori di α e β teoricamente corretti ($\alpha = 2$, $\beta = 0.1$).

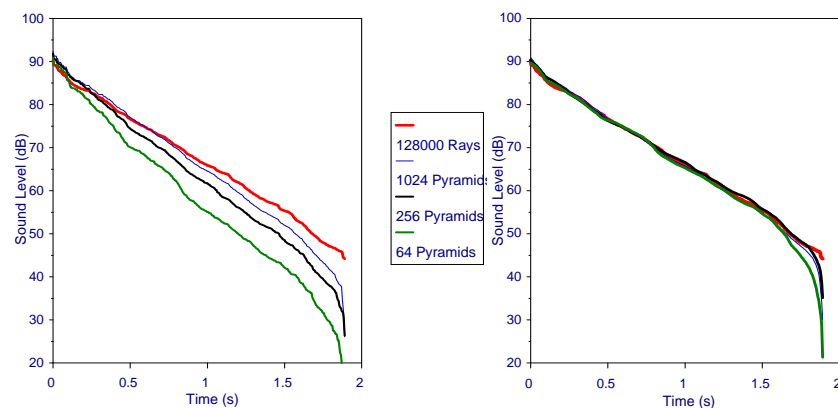


Figura 3.18 - Confronto delle Curve di Decadimento di una camera riverberante.

L'accuratezza dei risultati può essere verificata confrontando i valori numerici del Tempo di Riverbero T_{30} ottenuti con Ramsete e con il Ray Tracing (2.768 s).

n° piramidi	T_{30} senza correzione	T_{30} con correzione
1024	2.368	2.614
256	2.136	2.828
64	1.784	2.691

Valori di T_{30} calcolati con Ramsete.

Un altro caso interessante è rappresentato da due volumi accoppiati con coefficienti di assorbimento diversi.

Nella figura sotto sono riportati la geometria e i risultati di un caso con queste caratteristiche: il Teatro Buero Vallejo recentemente costruito in Spagna, ad Alcorcon (Madrid), su progetto dell'architetto Isicio Ruiz. La simulazione si riferisce al teatro terminato, con la sala completamente arredata ed il palcoscenico completamente vuoto (e quindi riverberante).

Il grafico mostra la risposta all'impulso (non integrata) ad 1 kHz ottenuta nel ricevitore n° 19 con il Ray Tracing e con Ramsete (con vari numeri di piramidi sparate) con la coda corretta con $\alpha = 1.9$ e $\beta = 0.22$.

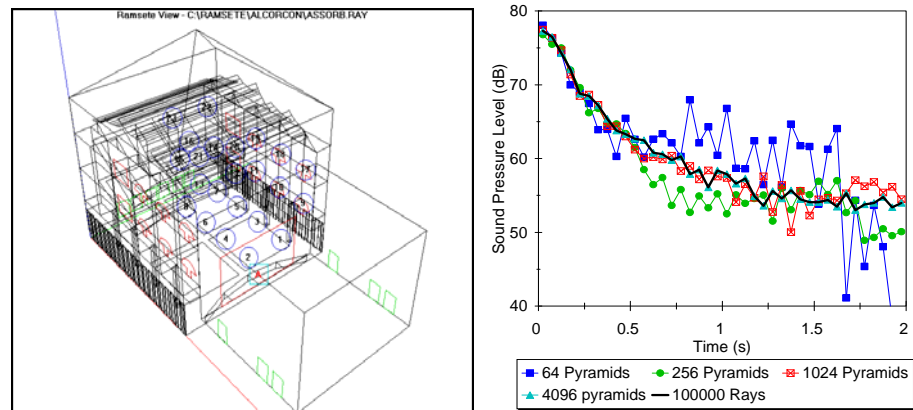


Figura 3.19 - Geometria e Risposta all'impulso di volumi accoppiati.

In questo caso i risultati mostrano una notevole discrepanza fra il riferimento (Ray Tracing 100000 raggi) e 64 piramidi, ed anche il caso con 256 non dà risultati sufficientemente buoni. Comunque, aumentando il numero di piramidi fino a 1024 o più, il risultato diventa pressochè indistinguibile dal riferimento, mentre il tempo di calcolo rimane ragionevole (4 minuti per 1024 piramidi su un Pentium 90 MHz).

Un altro caso tipico è costituito dalle chiese, ambienti tipicamente molto riverberanti e con caratteristiche anche di forte discontinuità del campo acustico e per questo motivo buon test per il programma. La Basilica di S. Domenico a Foligno è una antica chiesa non più usata per cerimonie religiose: dopo molti anni durante i quali è rimasta inutilizzata, è stata ora restaurata ed adattata a sala polivalente. Viene quindi usata come auditorium per importanti conferenze, ma a volte anche come sala da concerto.

Sono state condotte misure sperimentali con un altoparlante non omnidirezionale, una scheda di acquisizione *MLS* ed una testa artificiale in 24 punti uniformemente distribuiti in tutta l'area.

La stessa configurazione è stata ricostruita anche con Ramsete, usando i dati di direttività e potenza dell'altoparlante ricostruiti con Ramsete partendo da misure precedentemente effettuate secondo le norme ISO 3744.

Per questo confronto sono state sparate 2048 piramidi, con un tempo di calcolo di pochi secondi.

I grafici mostrano il tempo di riverbero e la mappatura del *RASTI*.

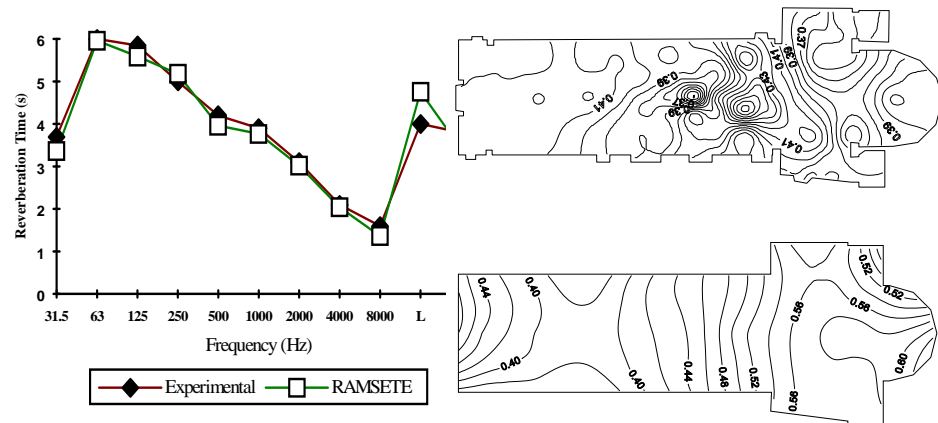


Figura 3.20 - Tempi di riverbero (sinistra) e *RASTI* (destra): la mappa in alto è sperimentale, quella in basso numerica.

Si può notare come i valori dei tempi di riverbero siano circa uguali mentre le mappe del *RASTI* sembrano un po' diverse, poiché quella sperimentale mostra un andamento complesso che risulta molto più semplice in quella numerica. Ma i valori numerici nei singoli punti non sono molto diversi: la differenza pesata *RMS* fra i valori numerici e sperimentali del *RASTI* è 0.114, che è dello stesso ordine di grandezza dell'errore nella misura sperimentale del *RASTI*. Ovviamente si può ottenere un migliore accordo fra dati misurati e calcolati se si è in possesso di dati più dettagliati sulla direttività della sorgente.

Capita sempre più frequentemente che i palazzetti dello sport vengano utilizzati per manifestazioni musicali di grande richiamo, ove occorra una grande capienza di pubblico, e sempre più spesso si interviene per correggerne l'acustica e renderli così adatti allo scopo. E' quindi stato interessante sottoporre Ramsete anche a questo test. Il Palazzetto dello sport di Modena è un grande edificio, con un volume interno di 70000 m³ ed una capacità di 5000 posti a sedere. Nel caso di concerti rock, vengono ricavati altri 1000 posti nel parterre. Questo ambiente ha una pessima acustica, a causa di tempi di riverbero eccessivamente alti, echi e focalizzazioni.

Sono state effettuate misurazioni in 44 punti usando come sorgente un colpo di pistola che veniva registrato da un DAT con un microfono omnidirezionale a condensatore. Le risposte all'impulso sono state poi riversate su PC attraverso una scheda audio di 16 bit a basso costo, convertiti in formato *MLSSA* (.TIM) e quindi analizzati come se fossero state ottenute da una misura di tipo *MLS*.

Il modello numerico del palazzetto dello sport introdotto con AutoCAD, ed importato in Ramsete, risulta composto da più di 1800 facce di 14 materiali diversi. Per il calcolo sono state sparate 256 piramidi, con un tempo di calcolo di 65 s. Qui sono riportati solo i risultati del tempo baricentrico, nella banda di 1 kHz, t_s sotto forma di mappe isolivello.

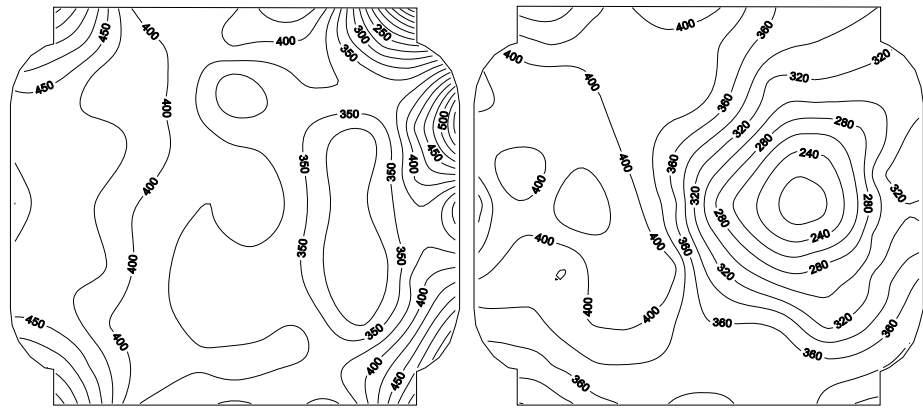


Figura 3.21 - Tempo Baricentrico nel Palazzetto dello Sport di Modena: sperimentale (sinistra) e numerico (destra).

I valori sono molto simili in punti situati lontano dalla sorgente, ma i risultati sperimentali ottenuti in prossimità della sorgente sono più alti di quelli numerici (che sembrano più ragionevoli). Probabilmente in queste misure si è avuto un overload della sezione analogica, e ciò causa una sottostima dell'onda diretta rispetto alla coda riverberante. Tale problema si è verificato solo nei punti molto vicini alla sorgente. Il valore medio pesato *RMS* fra i valori numerici e sperimentali è 51.2 ms, mentre scartando i sei punti intorno alla sorgente si riduce a 43.8.

Un altro caso interessante è quello di una industria: una geometria molto bassa ed allungata con coefficienti di assorbimento non uniformi e quindi con caratteristiche marcatamente non Sabiniane.

La figura sotto mostra sia la geometria, sia i risultati ottenuti: essendo un edificio ad uso industriale, il parametro acustico più importante è il livello di pressione sonora (*SPL*) e la sua attenuazione con la distanza.

Il risultato di riferimento (Ray Tracing, 160000 raggi) è messo a confronto con i risultati ottenuti con Ramsete (1024 piramidi), presentati con due diverse coppie di valori per i coefficienti α e β . La prima coppia (Ramsete1, $\alpha = 1.6$, $\beta = 0.2$) produce risultati molto simili al Ray Tracing.

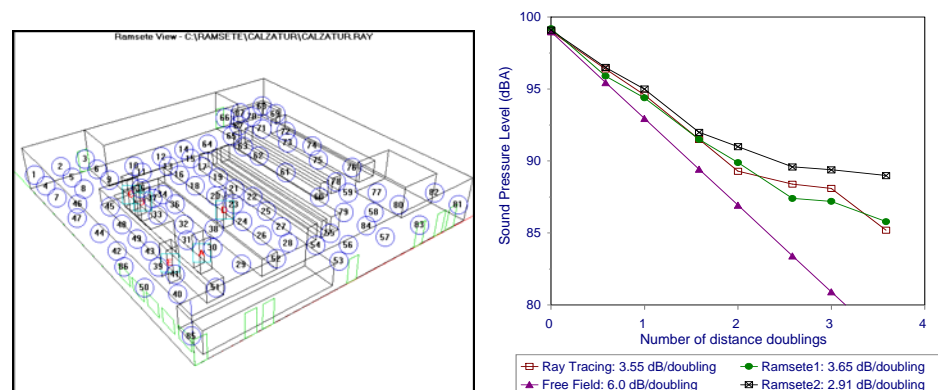


Figura 3.22 - Edificio industriale molto basso (con soffitto assorbente).

Va notato come, adottando valori errati di α e β si ottengano per l'*SPL* valori errati solo in punti molto lontani dalla sorgente: la linea tratteggiata in nella figura sopra è relativa al caso dei valori teorici (Sabiniani) di α and β (2.0 e 0.1 rispettivamente), e si può verificare

come il massimo scostamento consista in una sovrastima di 3.8 dBA.

Per un approfondimento delle situazioni non sabiniane in cui una delle tre dimensioni del locale è significativamente inferiore alle altre due, si rimanda all'articolo 'Studio della propagazione del rumore in ambienti industriali bassi e vasti', A. Farina, G. Fornari). Tale studio mostra come negli ambienti bassi e vasti esista un contributo rilevante dovuto alle riflessioni sonore con un conseguente innalzamento del livello sonoro complessivo, maggiore di quello ipotizzato da Sabine. Questo effetto è tanto più forte in quei casi in cui ci sono aree con sorgenti sonore distribuite con operatori ed esse prossime. L'articolo propone una relazione empirica da inserire nella legge di Sabine per correggerla; per applicarla è sufficiente conoscere l'altezza media del locale e la distanza del punto dalla sorgente sonora.

3.7 Taratura dell'algoritmo all'aperto

Ramsete, limitatamente ad un raggio di qualche centinaio di metri, può essere utilizzato anche per l'ambiente esterno. E' necessario quindi verificarne l'efficacia confrontando i suoi risultati con quelli sperimentali e quelli di un modello di calcolo appositamente studiato per l'esterno.

L'ambiente di studio è riportato nella figura sotto; è possibile vedere la posizione della sorgente (etichettata "A") e delle 16 posizioni microfoniche, poste su una linea retta e distanziate di 10 m. La sorgente era orientata verso gli edifici, parallelamente alla linea di misurazione.

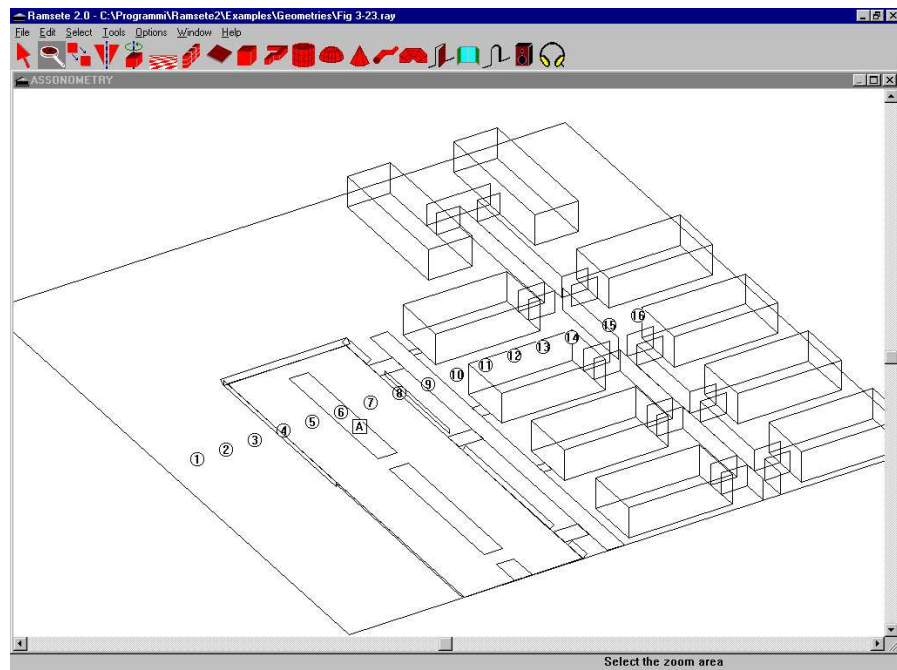


Figura 3.23 - Vista 3D della geometria studiata.

La sorgente utilizzata è un altoparlante coassiale a due vie montato su una cassa chiusa di forma sferoidale, in modo da minimizzare i fenomeni diffrattivi ed approssimare il più possibile ad una sorgente puntiforme. Lo spettro e la direttività della sorgente sono le seguenti.

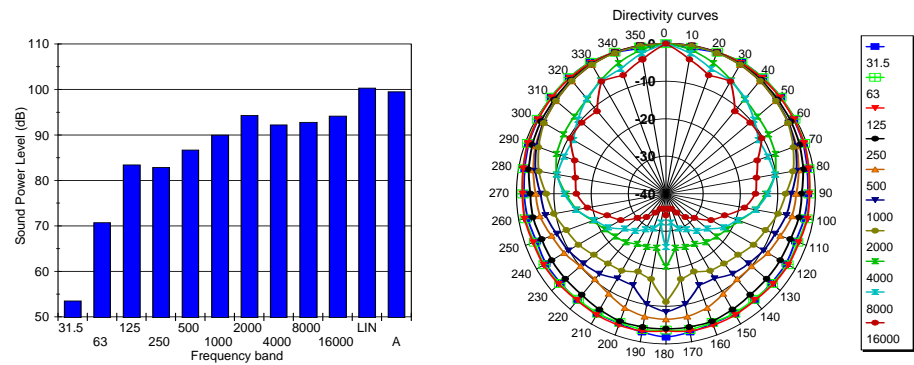


Figura 3.24 - Livelli di potenza e direttività della sorgente sonora.

Il test implica la propagazione del suono su due tipi diversi di terreno (asfalto ed erba), con terrapieni alti 1 m ed edifici alti 7 m o 10 m; le facciate di tali edifici sono totalmente composte da finestrate in cristallo. La sorgente posta ad 1 m dal suolo, e i microfoni posti a quota 1.3 m.

In ciascuna posizione di misura è effettuata una registrazione digitale (usando un DAT) di 60 s di segnale *MLS* ed una seconda, sempre di 60 s, a sorgente *MLS* spenta, per verificare il livello del rumore di fondo. Poiché il livello totale di potenza dell'altoparlante è limitato (100 dB), in molti punti il livello del segnale è sceso sotto la soglia del rumore di fondo, ma è ugualmente possibile misurarlo grazie alle proprietà del segnale *MLS*.

I risultati sperimentali sono presentati insieme a quelli delle simulazioni numeriche, per facilitare i confronti e sottolineare le differenze. Guardando nel grafico successivo risulta chiaro come Ramsete dia migliori risultati nelle vicinanze della sorgente, mentre il modello di calcolo per esterno, basato sulla ISO/DIS 9613, sia più accurato nei punti più distanti e nascosti fra gli edifici. Tale risultato era prevedibile, poiché Ramsete considera le riflessioni multiple e considera la direttività della sorgente, ma non include la valutazione di attenuazioni che all'esterno andrebbero considerate (effetto suolo, etc.). In più sembra che il tracciante per esterni sia più accurato nella valutazione dell'effetto di schermatura dovuto agli edifici, che Ramsete sembra sovrastimare.

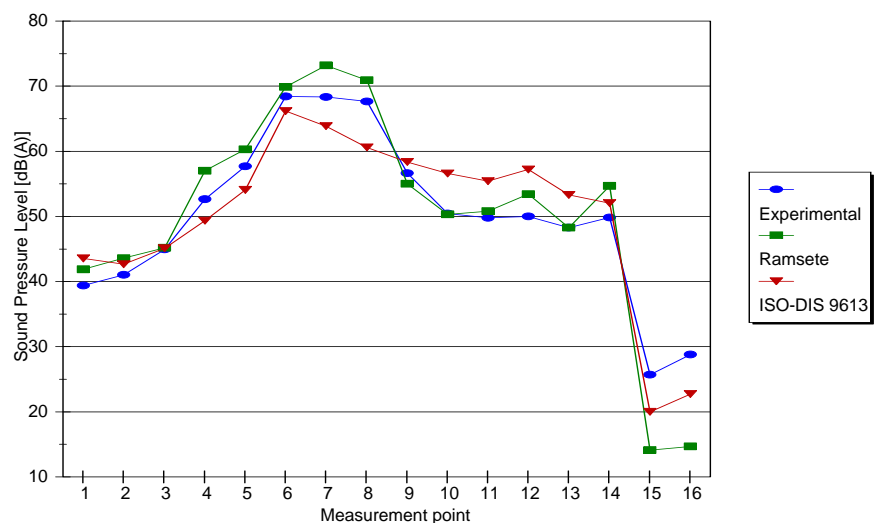


Figura 3.25 - Confronto dei risultati dB(A).

Un'altra interessante considerazione appare osservando lo spettro in qualche particolare punto di misura:

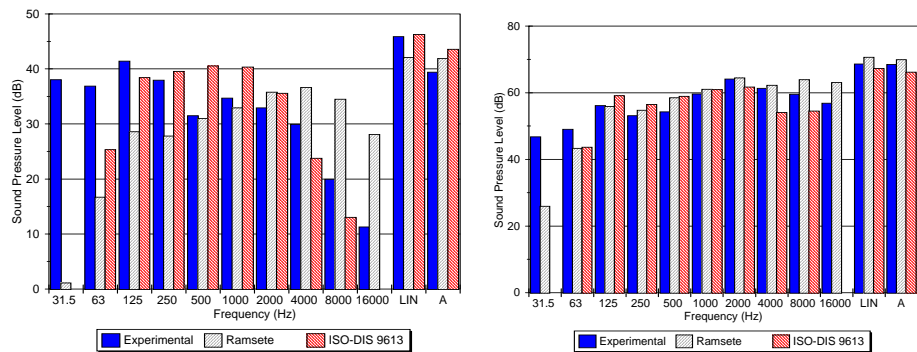


Figura 3.26 - Spettro dei punti n° 1 (sinistra) e n° 6 (destra).

Si può osservare come nel punto n° 1 (dietro la sorgente, e parzialmente schermato), i tre spettri siano abbastanza dissimili: quello sperimentale mostra una forte contaminazione dal rumore di fondo nelle bande a frequenza più bassa, mentre Ramsete sovrastima le alte frequenze di circa 10 dB. È evidente, a questo punto, che la formula dell'assorbimento dell'aria non sia realistica per un uso all'esterno, in presenza di un manto erboso, tale effetto sembra invece essere simulato correttamente dal modello per esterno.

Osservando cosa accade nel punto n° 6 (il più vicino alla sorgente sonora), invece, si può notare come i tre spettri siano molto simili: ovviamente a piccole distanze, i termini di assorbimento da considerarsi solo all'esterno non sono molto importanti, e l'andamento dello spettro dipende, per la maggior parte, dalla direttività della sorgente.

Poiché Ramsete non si limita al calcolo dei livelli di *SPL*, ma fornisce anche la risposta all'impulso fra sorgente e ricevitore, è possibile il confronto fra quest'ultima e quella sperimentale: si può notare come le principali riflessioni speculari siano state calcolate con precisione.

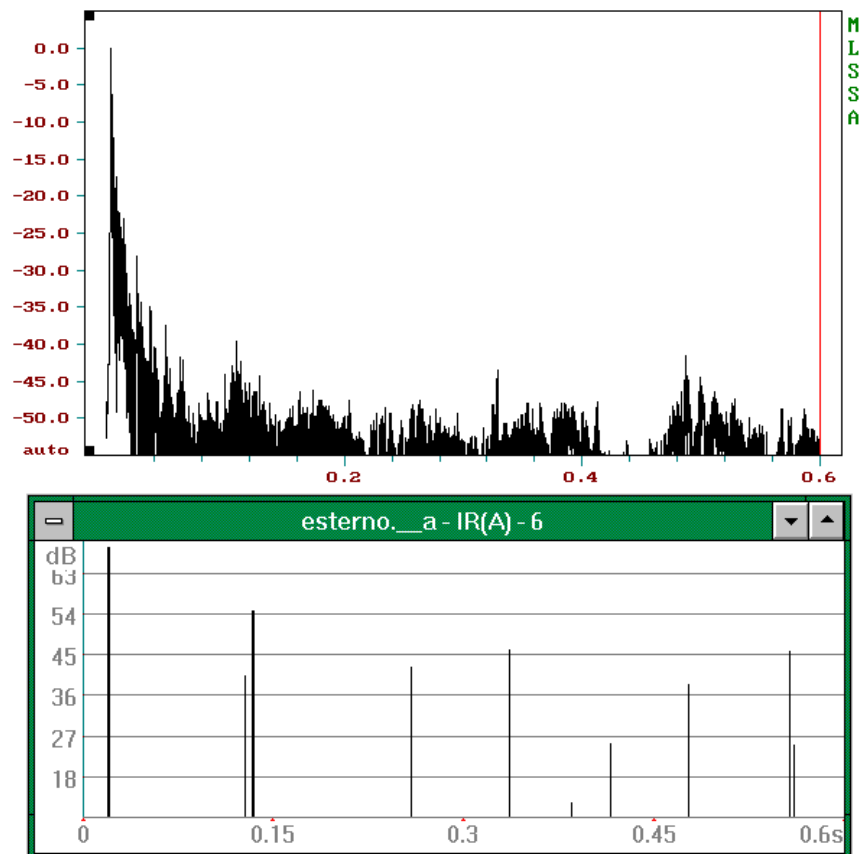


Figura 3.27 - Confronto fra la risposta all'impulso sperimentale e calcolata da Ramsete del punto n° 6.

Per un approfondimento sull'uso di Ramsete in ambienti aperti e chiusi si consigliano la lettura dei seguenti articoli:

'Ramsete - A new Pyramid tracer for medium and large acoustic problems', A. Farina

'Ramsete - Un nuovo software per la previsione del campo sonoro in teatri, ambienti industriali ed in ambienti esterni', A. Farina, P. Galaverna, G. Truffelli

'Pyramid Tracing vs Ray Tracing for the simulation of sound propagation in large rooms', A. Farina

'Validation of the pyramid tracing algorithm for sound propagation outdoors: comparison with experimental measurements and with the ISO/DIS 9613 standards', A. Farina

'Propagazione sonora e previsione del rumore negli ambienti di lavoro', A. Farina

4. TECNICHE DI MISURA

4.1 La sorgente sonora per la misura della risposta all'impulso

Nella realtà operativa, fino ad oggi, si è cercato di conciliare ipotesi teoriche con esigenze di carattere pratico. In particolare, per evitare l'onere del trasporto di strumentazione eccessivamente ingombrante, si sono ricercate sorgenti sonore che potessero approssimare l'impulso unitario positivo di pressione, di durata temporale nulla, della funzione Δ di Dirac.

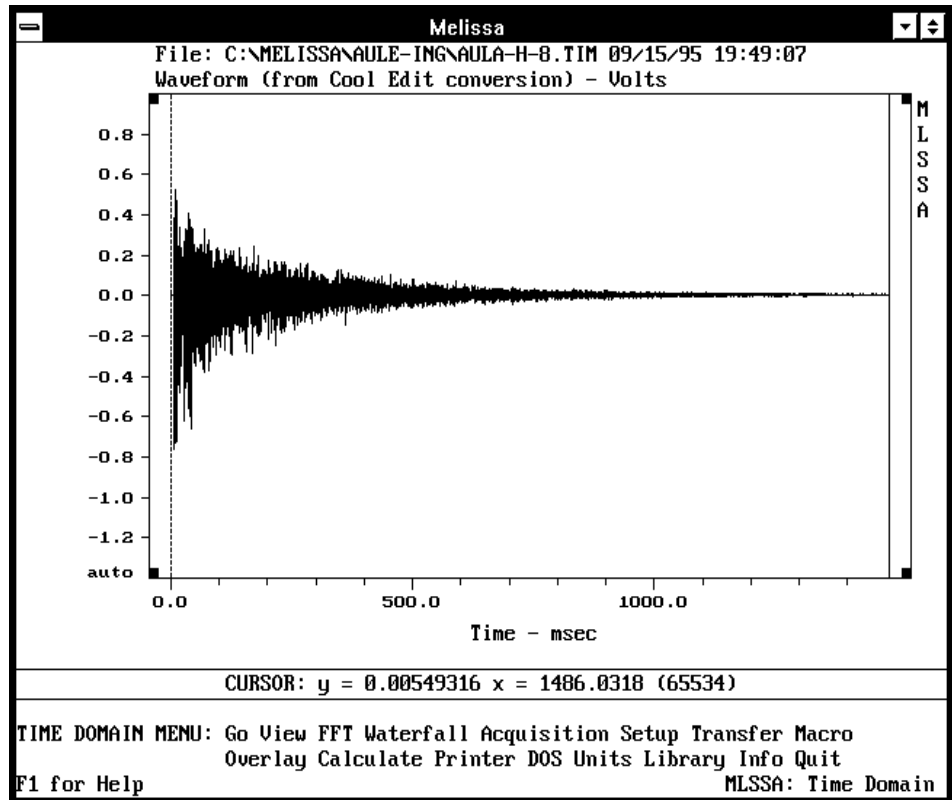


Figura 4.1 - Esempio di risposta all'impulso sperimentale.

La scelta della sorgente sonora rappresenta indubbiamente una componente rilevante all'atto di misurare le caratteristiche acustiche di un ambiente chiuso.

Le sorgenti normalmente impiegate sono:

- l'altoparlante
- il pallone
- la scintilla elettrica
- il colpo di pistola.

La scelta del tipo di sorgente con le peculiarità più adatte viene compiuta a seconda della ricerca acustica che si deve eseguire.

Analizzando, infatti, le caratteristiche temporali mediate su un numero notevole di eventi, si è potuto constatare che, mentre la scintilla elettrica è più adatta alle misure nei modelli in scala, il colpo di pistola va preferito nello studio di qualificazione acustica delle sale, e lo scoppio dei palloncini risulta efficace nello studio delle basse frequenze.

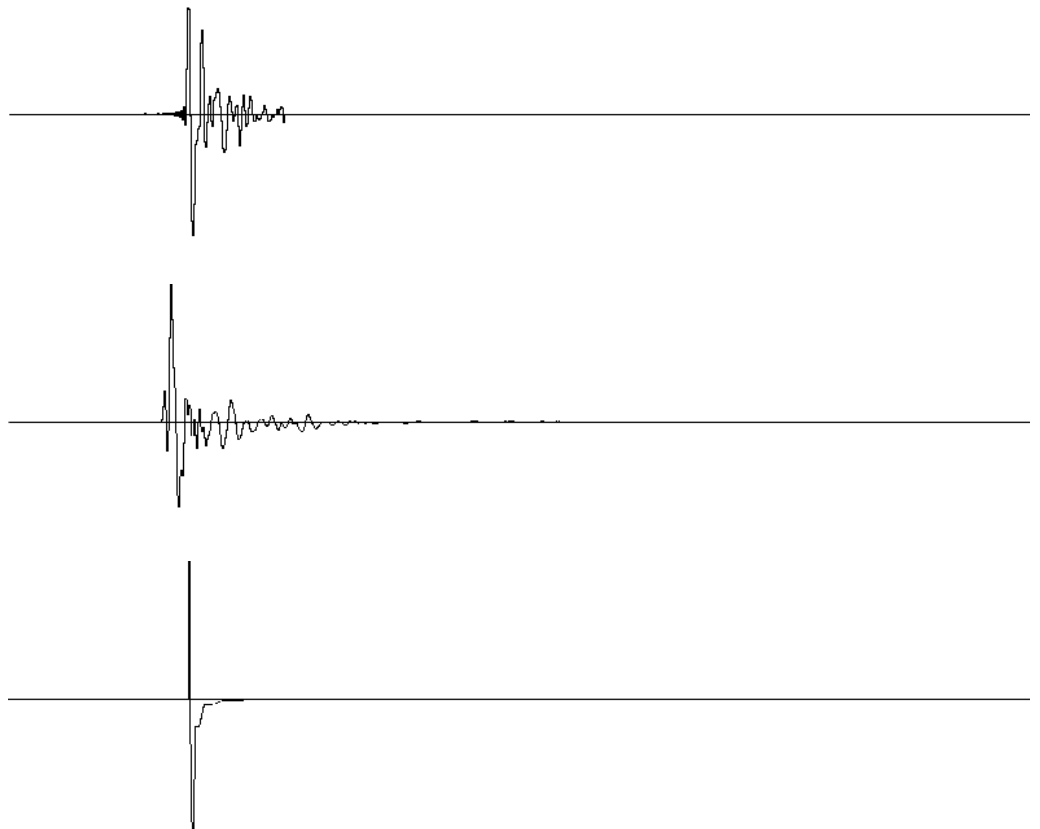


Figura 4.2 - Caratteristiche temporali delle sorgenti impulsive (dall'alto verso il basso: pistola, altoparlante e scintilla).

Per la misura del campo acustico di sale di ampie dimensioni tramite tecniche impulsive, è risultato preferibile l'impiego del colpo di pistola. Gli spari esplosivi risultano caratterizzati da un transitorio sufficientemente breve (dell'ordine di grandezza compreso fra 1 e 3 ms), mentre nel dominio delle frequenze risultano essere leggermente deboli al di sotto dei 100 Hz ed al di sopra dei 5000 Hz.

Per gli studi di qualificazione acustica, tuttavia, le frequenze minori di 100 Hz rivestono importanza minore rispetto alle altre, pertanto anche se con tali approssimazioni, risultano accettabili.

Montando sulla bocca della pistola un diffusore a forma sferica, poi, si ottiene una sorgente pressochè omnidirezionale.

Nel caso fosse richiesta una analisi più accurata alle basse frequenze, si può impiegare, in ausilio alla tecnica del colpo di pistola, lo scoppio di palloncini, la cui risposta in frequenza risulta essere particolarmente efficace attorno ai 100 Hz.

4.2 Tecnica dei segnali

Ipotesi di campo

Nello studio dell'acustica l'ambiente viene schematizzato con il concetto di scatola nera: un oggetto avente un ingresso ed una uscita ad esso correlata ma il cui funzionamento interno non è noto.

Si presuppone che il campo acustico in un ambiente sia lineare.

Un sistema si dice lineare se l'uscita è una funzione lineare dell'entrata ovvero se vale il principio di sovrapponibilità degli effetti.

Il principio di sovrapposibilità degli effetti si può esprimere così: dati due ingressi generici A e B e le loro relative uscite C e D, ottenute applicando singolarmente gli ingressi, applicando nel sistema l'ingresso A+B si ottiene l'uscita C+D.

Inoltre si presuppone che i sistemi studiati siano tempo invarianti, cioè indipendenti dal tempo.

Nell'ambito di queste ipotesi si applica il concetto matematico di risposta all'impulso. La formulazione di questa teoria è particolarmente semplice nel dominio digitale dei segnali campionati.

Il segnale in questo dominio è rappresentato così: l'intervallo di variabilità della tensione viene diviso in 2^n (dove n è il numero di bits usati nella rappresentazione) sottointervalli rappresentati ciascuno dal proprio valore medio; periodicamente il segnale analogico viene misurato (campionato) ed a seconda dell'intervallo in cui cade la tensione il campione assume il valore medio dell'intervallo.

Quindi la rappresentazione di un segnale nel dominio digitale è semplicemente un insieme ordinato di numeri interi. Si definiscono inoltre alcune grandezze che influenzano la rappresentazione digitale: frequenza di campionamento ν , periodo (è l'inverso della frequenza di campionamento) $\Delta\tau$.

```

10010000
01001010
10001001
01110011
00011000
10100000
00000011
00111111
01010101
10000111
00010000

```

Esempio di segnale nella rappresentazione digitale (e binaria) ad 8 bit.

Come si è visto il segnale entra nel sistema come un vettore di numeri (il nome dei vettori sarà scritto minuscolo in neretto) ed esce dal sistema come un altro vettore di numeri con la stessa frequenza di campionamento. Chiamando \mathbf{x} il vettore di numeri in entrata e \mathbf{y} quello in uscita essi possono essere rappresentati così:

```

x1  y1
x2  y2
x3  y3
x4  y4
x5  y5
x6  y6
x7  y7
x8  y8

```

4.3 La convoluzione

Bisogna notare come i dati in uscita siano "legati" ai dati in ingresso: in particolare avendo in ingresso una sequenza di zeri (silenzio), seguita da numeri non nulli a loro volta seguiti da zeri, in uscita si avrà una sequenza simile alla prima salvo che per il numero di zeri all'inizio ed alla fine e per i valori dei campioni.

Questa disparità nel numero di zeri precedenti e seguenti i due segnali deriva dal fatto che la risposta del sistema non è istantanea, né quando il sistema viene eccitato (attacco del suono), né quando il sistema ritorna allo stato iniziale (coda del suono).

In termini matematici questo si esprime dicendo che y_n non è funzione solo di x_n ma di un certo numero di campioni in entrata precedenti a x_n . Nel dominio digitale questo viene espresso dall'equazione:

$$y_n = x_n h_1 + x_{n-1} h_2 + x_{n-2} h_3 + x_{n-3} h_4 + \dots + x_{n-m} h_m$$

dove m è il numero dell'ultimo elemento di cui si ha memoria.

Questa operazione si definisce convoluzione e si usa la notazione:

$$y = x \otimes h$$

I coefficienti rappresentano quindi la "caratteristica" del sistema. Guardando a questi coefficienti come ad una forma d'onda, questa rappresenta la risposta all'impulso del sistema ed ha la stessa frequenza di campionamento dei dati in ingresso. Per riprodurre questo comportamento del sistema i sistemi digitali di riproduzione e studio degli ambienti hanno a disposizione della memoria in cui i campioni vengono immagazzinati. La natura del problema studiato implica la necessità di avere una memoria "equivalente" a tempi che possono arrivare ad alcuni secondi. Con un semplice calcolo, ipotizzando una frequenza di campionamento di 44.100 Hz (la frequenza dei CD audio commerciali), si nota come una memoria di 5 secondi richieda l'utilizzo di $5 \cdot 44100 = 220500$ bytes.

Tale calcolo è realistico, considerando il fatto che valori tipici di memorie dei sistemi digitali di studio e simulazione degli ambienti, variano nell'ordine di $10^5 - 10^6$ bytes.

Si possono fare due considerazioni: intanto un notevole aumento della memoria richiesta rispetto a sistemi digitali dedicati allo studio ad esempio di fenomeni elettrici, inoltre la notevole mole di calcoli che questo tipo di trattazione richiede, così notevole da non poter essere svolta tuttora in tempo reale e comunque tecnologicamente irrealizzabile fino a pochi anni fa.

La tecnica di misura dei coefficienti h è teoricamente molto semplice: basta infatti immettere nel sistema la δ di Dirac ovvero un segnale costituito da un uno seguito solo da numeri nulli.

$$1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0$$

Convolvendo la δ di Dirac si ottengono quindi:

$$y_1 = 1 \cdot h_1 + 0 \cdot h_2 + 0 \cdot h_3 = h_1$$

$$y_2 = 0 \cdot h_1 + 1 \cdot h_2 + 0 \cdot h_3 = h_2$$

$$y_3 = 0 \cdot h_1 + 0 \cdot h_2 + 1 \cdot h_3 = h_3$$

Ad ogni passo di campionamento si ottiene quindi un coefficiente h .

Questo tipo di misurazione, benché teoricamente semplicissima, è praticamente irrealizzabile: bisogna essere capaci di produrre un impulso brevissimo e di elevata potenza (per una buona misurazione ho bisogno di almeno 60 db oltre il rumore di fondo).

Una possibile soluzione è l'uso di esplosivi (una pistola caricata a salve) che, però, non dà un segnale che dura solo un periodo di campionamento ma arriva a qualche decina di periodi.

Per ovviare a questo inconveniente si può convolvere il segnale di risposta dell'ambiente con sé stesso rovesciato sull'asse dei tempi in modo da far diventare primo l'ultimo campione etc.

Questa tecnica prende il nome di Time Reversal Mirror. Questo porta ad avvicinarsi alla δ di Dirac ma non al suo raggiungimento. Se con il TRM si ottenesse effettivamente la δ di Dirac viste le proprietà di commutatività della convoluzione, basterebbe convolvere il segnale in uscita con il segnale in ingresso rovesciato sui tempi per ottenere la risposta all'impulso del sistema.

Da queste considerazioni nasce l'idea di deconvolvere il segnale con un segnale casuale come ad esempio il rumore bianco. Questo si fa sperando di trovare un \mathbf{x}^{-1} tale che $\mathbf{x} \otimes \mathbf{x}^{-1}$ sia uguale alla δ di Dirac. Ammesso di riuscire a trovare \mathbf{x}^{-1} si otterrebbe:

$$\mathbf{x} \otimes \mathbf{h} \otimes \mathbf{x}^{-1} = \mathbf{y} \otimes \mathbf{x}^{-1} = \mathbf{h} \otimes \delta = \mathbf{h}$$

purtroppo è molto difficile riuscire a trovare \mathbf{x}^{-1} .

$$\mathbf{y} = \mathbf{x} \otimes \mathbf{h} \xrightarrow{\text{FFT}} \mathbf{Y} = \mathbf{X} * \mathbf{H}$$

Analisi in frequenza

Si può però provare a passare nel dominio delle frequenze (in maiuscolo i vettori rappresentati in frequenza):

$$\mathbf{y} = \mathbf{x} \otimes \mathbf{h}$$

Si nota che mentre nel dominio del tempo sono necessarie m convoluzioni (a loro volta costituite da m somme ed m prodotti) quindi un numero di operazioni dell'ordine di m^2 , in frequenza ogni armonica viene semplicemente moltiplicata per un coefficiente per un totale di m operazioni (si noti come la trasformazione e antitrasformazione siano operazioni con un costo di calcolo limitato). Inoltre in frequenza l'operazione di ricerca dei coefficienti H è estremamente semplice, infatti essi sono semplicemente il quoziente fra Y e X .

$$H = \frac{Y}{X} \xrightarrow{\text{IFFT}} \mathbf{h}$$

In particolare si definisce H funzione di trasferimento del sistema mentre h è la risposta all'impulso.

Esiste però un problema fondamentale: il risultato è instabile; infatti nel caso di una frequenza con coefficiente nullo il coefficiente H relativo diverge. Si può ovviare a questo problema facendo una media con varie misurazioni e scegliendo un segnale in ingresso che dia energia su tutte le frequenze, tipicamente il rumore. Un altro problema: si deve usare una FFT di lunghezza adeguata al tempo di risposta del sistema. Infatti, come detto prima, si deve ripetere il segnale nel tempo per ottenere dati per la media e usare una frequenza di ripristino che non porti sovrapposizioni dei segnali in uscita ed inoltre una

FFT che comprenda ogni volta la risposta all'impulso di una singola ripetizione; altrimenti si verifica il fenomeno del Time Aliasing.

Segnali particolari: MLS

Al posto del rumore si può però scegliere il segnale in ingresso in modo opportuno, mantenendo le caratteristiche utili (la presenza di tutte le frequenze con coefficienti non nulli) ma cercando di ottenere ulteriori vantaggi.

Un segnale particolarmente "intelligente" deriva dalla teoria dei numeri e prende il nome di MLS (Maximum Length Sequence). Questo segnale è una sequenza binaria costruita tramite uno Shift register (registro a scorrimento):

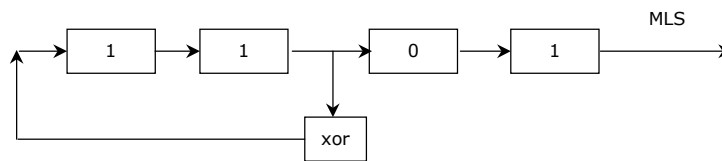


Figura 4.3 - Esempio di Shift register a 4 bit, i numeri scorrono a destra ad ogni passo.

Con opportuni valori iniziali e un corretto posizionamento dell xor si ottiene un MLS. Si possono fare varie osservazioni su questa sequenza:

- essendo nota in partenza si trova matematicamente MLS^{-1} con l'inversione di Hadamard; essendo inoltre MLS^{-1} un segnale binario è particolarmente facile la convoluzione (poiché spariscono le moltiplicazioni);
- grazie alla semplicità di questo segnale è possibile fare la convoluzione nel dominio dei tempi in tempo reale;
- siccome il segnale in entrata è generato direttamente non deve più essere campionato e quindi si può usare una scheda di campionamento ad una sola entrata
- inoltre essendo lo spettro sonoro dell'MLS piatto come quello del rumore, è possibile eseguire l'analisi in frequenza.

Si dimostra che la lunghezza di una MLS è di $2^n - 1$ dove n è il numero di celle dello shift register. In acustica in particolare, si usano normalmente shift registers a 16 o più celle ricavando MLS di 65535 campioni (caso 16 celle) o superiori (ovviamente però si può ripetere il segnale).

Questa tecnica, inventata nel 1975 dal tedesco Alrutz, è stata la più usata fino a circa un anno fa. Questo anche grazie alla diffusione di una scheda, costruita nel 1989 dall'americano Douglas Rife, che implementava per via hardware lo shift register e corredata da un software chiamato MLSSA (MELISSA) particolarmente potente, tanto da essere tuttora usato. Gli svantaggi di questa tecnica sono la fortissima dipendenza dalle caratteristiche di linearità del sistema. Infatti nel caso di non-linearità anche lievi possono apparire echi inesistenti e si può avere la cancellazione delle alte frequenze a causa del loro sfasamento reciproco che le può portare in controfase. Purtroppo quindi questa tecnica può essere usata solo con riproduttori audio molto fedeli e per un tempo non troppo lungo, infatti il riscaldamento degli al-

toparlanti comporta una perdita di fedeltà e conseguentemente di linearità.

Segnali particolari: sine sweep

Attualmente la tecnica migliore è quella dello sweep. Il segnale così definito è composto da un seno che parte dalle frequenze basse e sale verso l'alto. Questa tecnica presenta il vantaggio che

S = segnale di sweep

S^{-1} = S rovesciato sull'asse dei tempi.

Vengono però persi i vantaggi sulla convoluzione. Per riottenere la velocità di calcolo della convoluzione si deve passare in frequenza.

Prime misure e misure con MLS

Un primo tentativo può essere quello di generare una δ di Dirac mediante l'editor audio. Questo tentativo ha portato al risultato qui sotto.

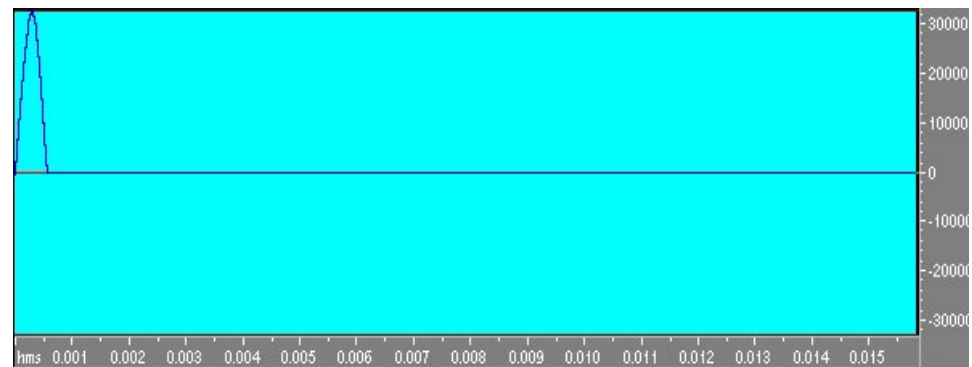


Figura 4.4 - Il primo tentativo di generare una δ vista con uno zoom sull'attacco.

Si rivela infruttuoso ai fini pratici; l'impulso così generato non ha l'energia richiesta per fare una misura e quindi non si procede su questa strada.

Una seconda possibilità è quella di generare un segnale di eccitazione con la tecnica MLS grazie ai plug-in AURORA utilizzabili nell'ambiente di Cool Edit.

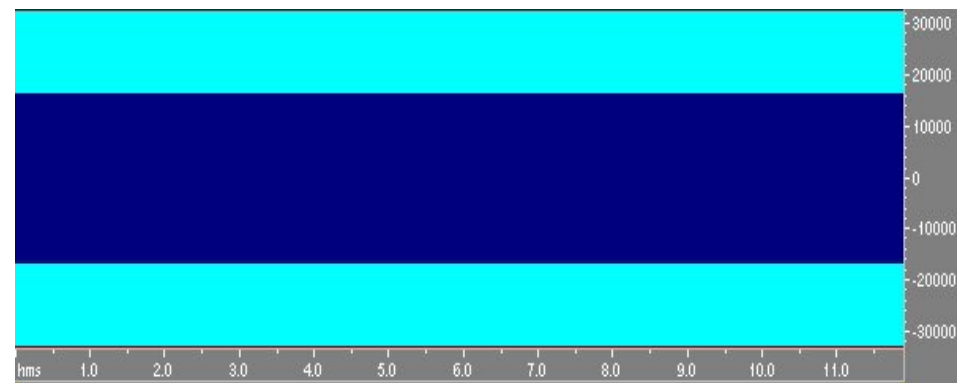


Figura 4.5 - Segnale MLS multiplo di ordine 16 ripetuto 8 volte.

Come si può vedere dall'analisi delle frequenze, questo segnale è molto simile al rumore bianco.

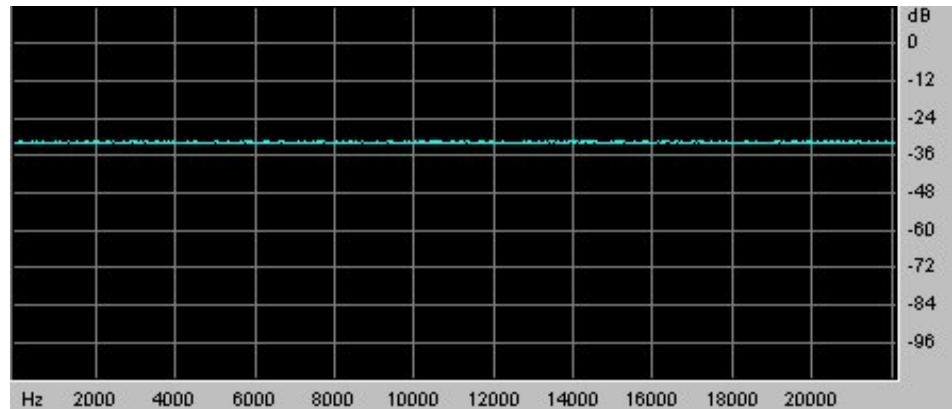


Figura 4.6 - Spettro della MLS visto con Cool Edit

Ora si procede alla misura: in Cakewalk la prima traccia contiene il segnale MLS mentre la seconda registra simultaneamente la risposta dal microfono.

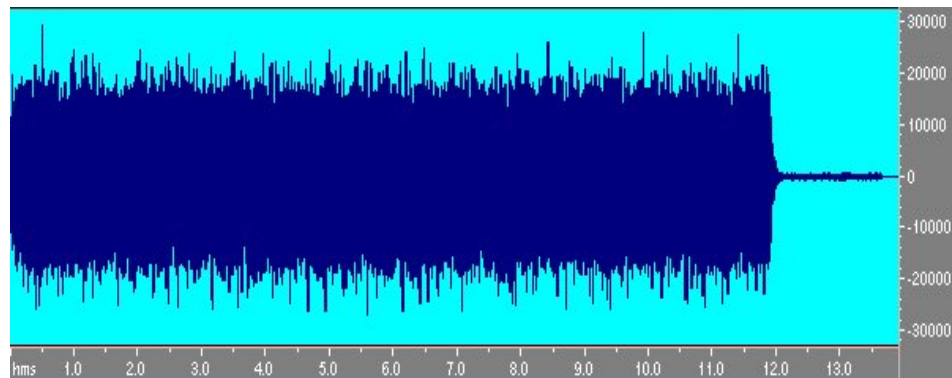


Figura 4.7 - Segnale di risposta registrato con il microfono.

Ora per ottenere la risposta della stanza alla δ si deve deconvolvere il file di risposta con l'inverso della MLS. Per fare questo si usa la funzione deconvolve multiple MLS di AURORA. Il risultato é un'onda che rappresenta la risposta dell'ambiente alla δ di Dirac. Il risultato della deconvoluzione, mostrato in figura, si trova nella clipboard di Windows.

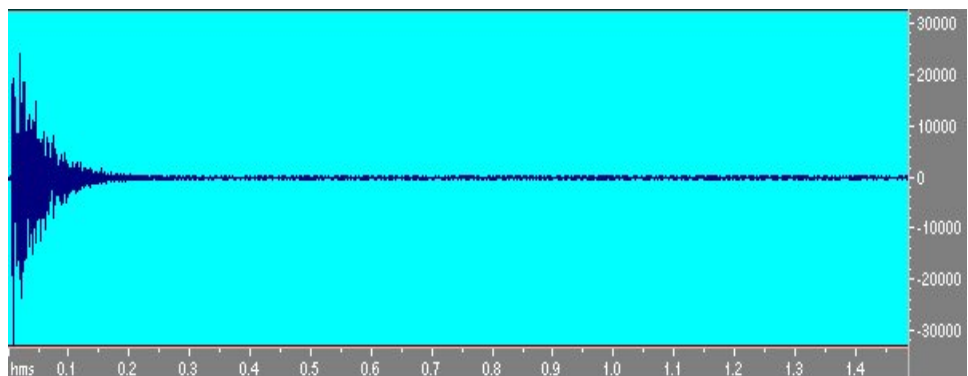


Figura 4.8 - Risposta della stanza ottenuta con MLS.

Notiamo che nella deconvoluzione bisogna usare un filtro che rimuova la componente continua, già integrato nel plug-in che deconvolve.

Misure con la tecnica Sweep

Una terza possibilità é quella dello sweep: ovvero un segnale composto da un tono puro la cui frequenza si alza nel tempo.

In questo caso si genera sempre mediante AURORA un segnale composto da tre sweep logaritmici con frequenze che partono da 100 hz ed arrivano a 20000 hz separati da 1 secondo di silenzio ognuno.

Lo sweep usato é logaritmico (la frequenza sale nel tempo in maniera logaritmica) perché in questo modo si fornisce piú energia alle basse frequenze (zona critica) e si procede piú svelti sulle alte in modo da non bruciare i tweeter. Nella figura sotto si vede il sonogramma (rappresentazione tempo-frequenza) del segnale sweep

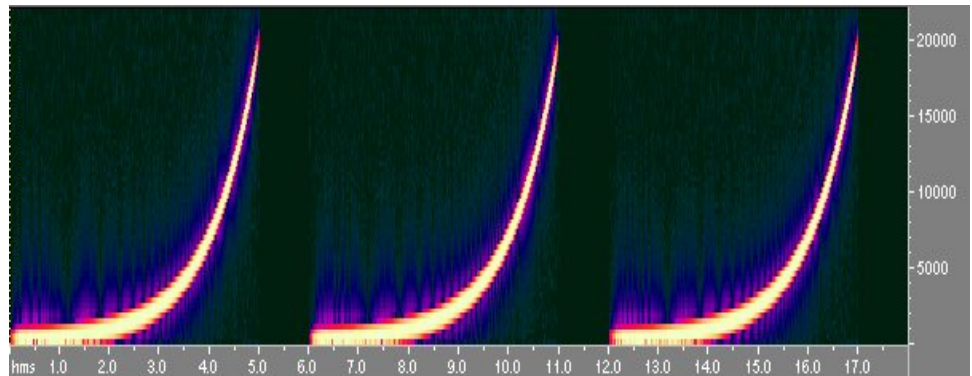


Figura 4.9 - Sequenza di tre sweep vista con Cool Edit spectral view

Le zone piú chiare indicano maggiore energia del segnale: come si vede la variazione di frequenza é logaritmica. Si fa la misura come in precedenza e si ottiene la risposta della stanza allo sweep:

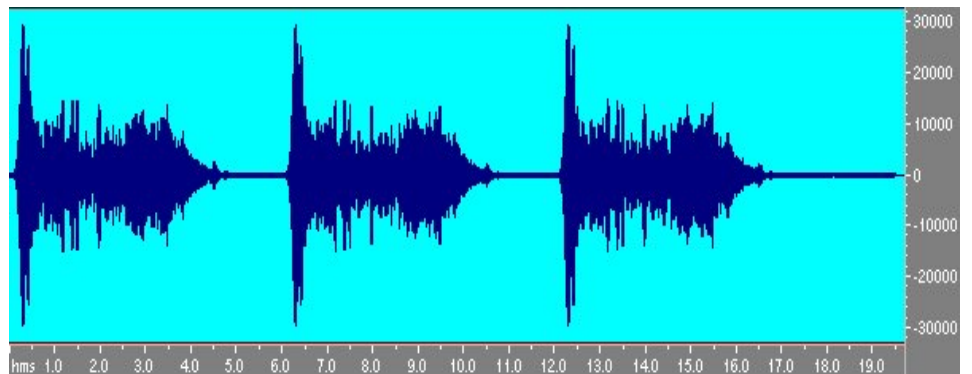


Figura 4.10 - Risposta alla sequenza di tre sweep.

Ora si convolve con lo sweep inverso, che il plug-in di generazione dello sweep ha già posto nella clipboard.

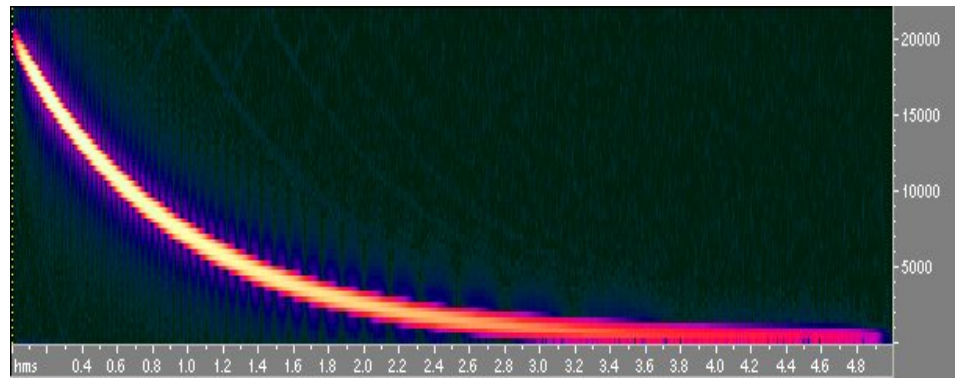


Figura 4.11 - Sweep inverso visto con Cool Edit spectral view

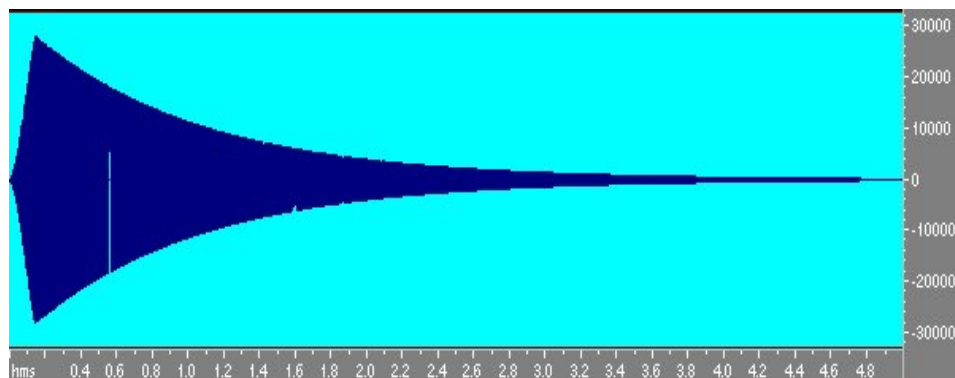
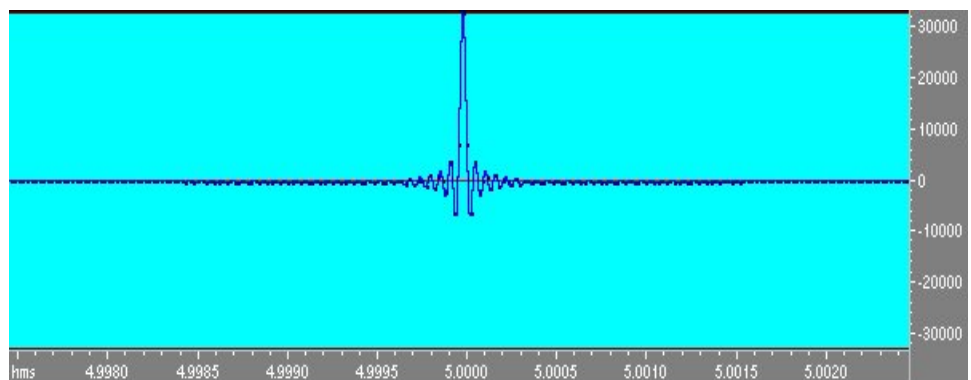


Figura 4.12 - Sweep inverso visto come ampiezza-tempo.

Si nota che, mentre lo sweep diretto ha ampiezza costante, lo sweep inverso è calante in ampiezza in modo da compensare la logaritmicità dello sweep che dava più energia alle basse frequenze. Infatti lo sweep logaritmico ha energia crescente di 3 db per ottava e quindi è un rumore rosa, se si convolve questo segnale con sé stesso non si ottiene la δ di Dirac ma un segnale che scende di 6 db per ottava. Questo viene compensato dal guadagno variabile dell'inverso. Convolverlo lo sweep col suo inverso si ottiene invece la δ di Dirac; osservando attentamente si nota però che non è esattamente la δ ma un filtro passabanda da 100 hz a 20000 hz. Questo perché la banda passante non arriva alla frequenza di Nyquist ma si ferma a 20000.

Figura 4.13 - δ ottenuta dallo sweep.

Convolverlo invece il segnale di risposta della stanza con lo sweep inverso si ottiene la risposta della stanza. Si nota intanto che avendo tre sweep si hanno anche tre risposte che è possibile mediare fra lo-

ro, ed inoltre come queste tre risposte sono assai simili a quella ottenuta con la tecnica MLS, ma molto piú pulite.

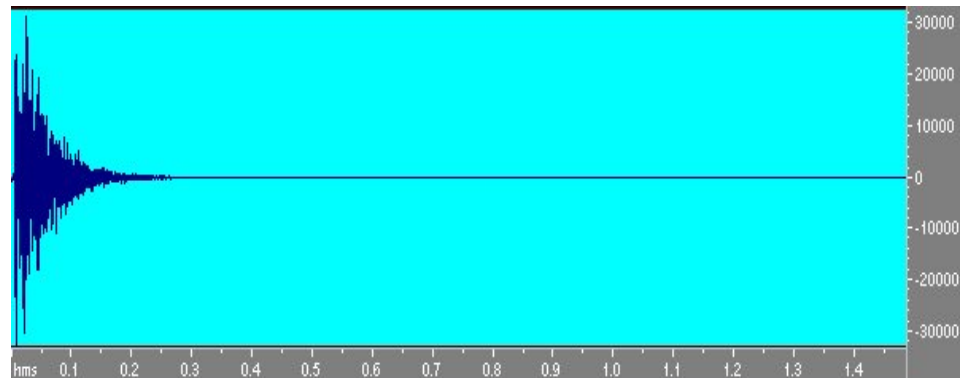


Figura 4.14 - Risposta della stanza ottenuta da sweep e mediata.

Analisi dei risultati

Terminata la parte di misura si può fare un'analisi cominciando dalla risposta ottenuta con MLS: intanto si osserva l'effetto filtrante della sala nel dominio della frequenza. Sotto si può vedere l'analisi in frequenza del sistema studiato, che comprende però anche il sistema di riproduzione essendo esso non perfettamente lineare.

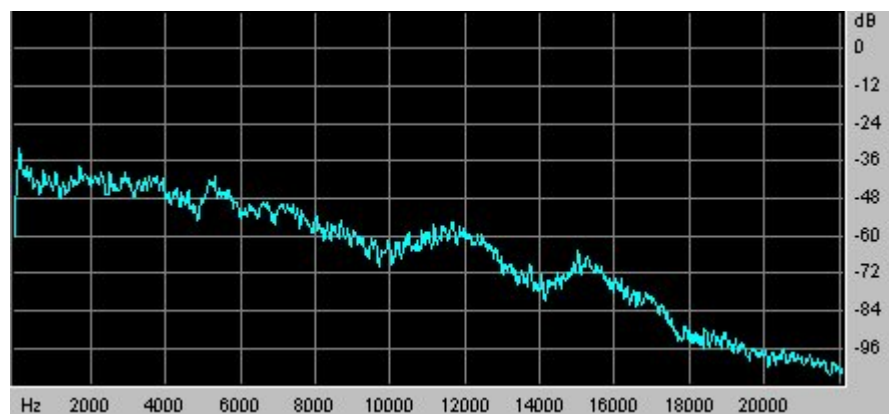


Figura 4.15 - Risposta della stanza ottenuta da MLS e vista in frequenza.

Si nota come il sistema, essendo formato da woofer molto grossi, riproduca meglio le basse e medie frequenze e cali sulle alte. Con Spectra Lab é possibile vedere il segnale con in ordinata i dB: si noti come fra picco del segnale e rumore ci sia una differenza di circa 45-50 db, non particolarmente buona per un sistema digitale.

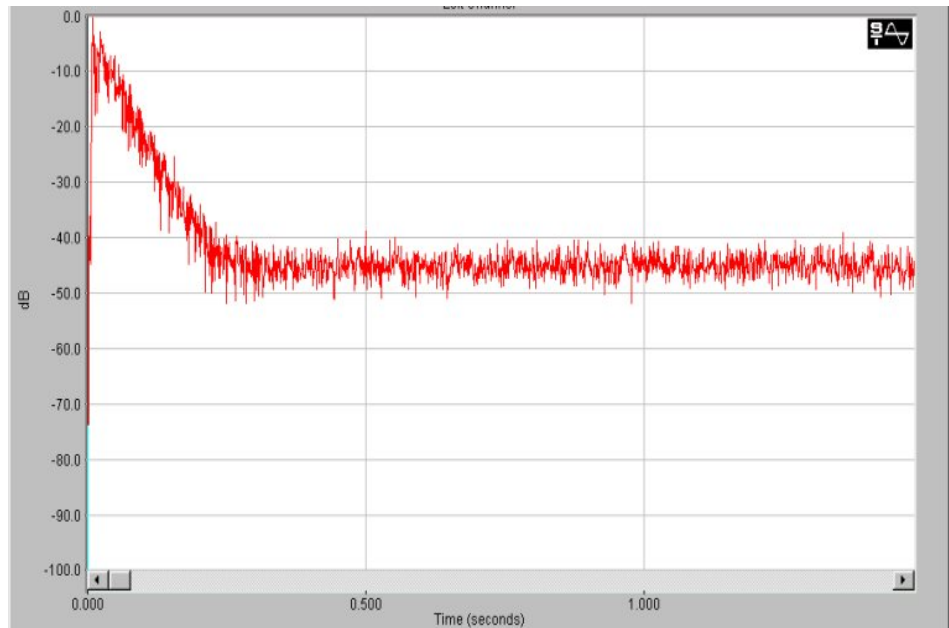


Figura 4.16 - Risposta della stanza ottenuta da MLS vista con Spectra Lab

Bisogna considerare che le uscite del computer sono state preamplificate da un mixer per ottenere un volume di uscita congruo con quello richiesto dall'amplificatore, questo procedimento aggiunge del rumore. Con l'integrazione di Schroeder (che Spectra Lab è in grado di fare) si può calcolare il tempo di riverbero.

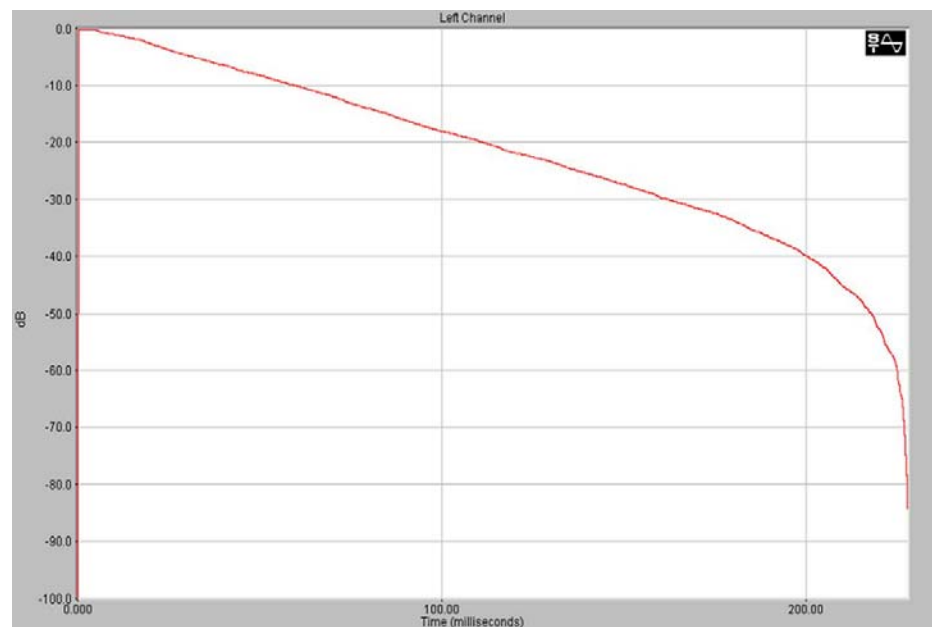


Figura 4.17 - Integrale di Schroeder

Siccome il picco è già riportato a 0 db si può vedere il tempo a cui il segnale vale -5 db e quello a cui vale -35 db sottrarre il primo al secondo e ottenere il tempo in cui il segnale ha un decadimento di 30 db. Siccome si vuole il tempo di riverbero riportato a 60 db si può avere un risultato approssimato moltiplicando tale tempo per due. Si ottiene quindi:

Tempo1 (-5 db)	33.01 ms
Tempo2 (-35 db)	243.75 ms
Tempo decadimento di 30 db	210.74 ms
Tempo decadimento di 60 db	421.48 ms approssimato!

Il tempo ottenuto é congruente con i tempi tipici di stanze di piccole dimensioni.

Per fare lo stesso studio in relazione alle singole frequenze si deve filtrare il segnale di risposta con un filtro passa banda e usare lo stesso metodo di prima. Dato che avere un'analisi su tutte le frequenze sarebbe laborioso sono stati scritti dei programmi che fanno tali calcoli automaticamente; fra questi uno dei moduli di AURORA. Tale modulo si chiama Calculate Acoustical Parameters e calcola fra gli altri i parametri RT20 e RT30 sulle frequenze di centro banda ad ottave.

OCTAVE BAND ACOUSTICAL PARAMETERS

Band	31.5	63	125	250	500	1K	2K	4K	8K	16K
C50 [dB]	4.121	10.25	7.583	4.832	8.746	7.321	5.862	8.12	9.469	12.
C80 [dB]	12.39	18.65	12.71	10.39	14.	14.86	11.28	14.76	15.48	19.69
D50 [%]	72.09	91.38	85.15	75.26	88.22	84.37	79.41	86.64	89.85	94.06
TS [ms]	41.96	21.68	33.48	36.33	32.26	29.27	30.74	28.5	23.49	22.87
EDT [s]	0.3191	0.226	0.3534	0.4178	0.2593	0.3305	0.4187	0.3195	0.2912	0.2255
RT20 [s]	0.6298	0.2825	0.343	0.4307	0.3156	0.3125	0.3052	0.3062	0.2888	0.2278
r RT20	0.9625	0.9751	0.9682	0.9863	0.9922	0.9913	0.9914	0.9965	0.9985	0.9959
RT30 [s]	0.6633	0.3402	0.3957	0.4487	0.3182	0.334	0.3147	0.3134	0.2911	--
r RT30	0.9878	0.9827	0.957	0.9947	0.9956	0.9949	0.9964	0.9985	0.9986	--
RTU [s]	0.4439	0.3748	--	0.3819	0.3072	0.2947	0.3671	0.284	0.2958	0.2476
r RTU	0.8523	0.9575	--	0.9551	0.9626	0.9773	0.9939	0.9937	0.995	0.9877
Noise Correction:	no	no	yes	yes	yes	yes	yes	yes	yes	yes

I parametri calcolati sono relativi alla risposta allo sweep perché quella relativa alla MLS non permette di calcolare il parametro RT30 a varie frequenze.

Tale calcolo é compreso anche nel software MLSSA, che in aggiunta ha molte altre funzionalità. Queste lo rendono ancora oggi un potente strumento nell'analisi delle sale. Si possono vedere gli spettri dei principali parametri acustici. Inoltre MLSSA consente di fare l'analisi di Fourier, di visualizzare il segnale per terzi di ottava e esportare i dati in formato testo, inoltre può fare le rappresentazioni di tipo waterfall. Questo é utile perché l'orecchio percepisce decine di spettri diversi durante l'evolversi della risposta. La possibilità di fare tali grafici é comunque compresa anche su Spectra Lab. Un altro modo di rappresentare il decadimento é la curva ETC (energy time curve).

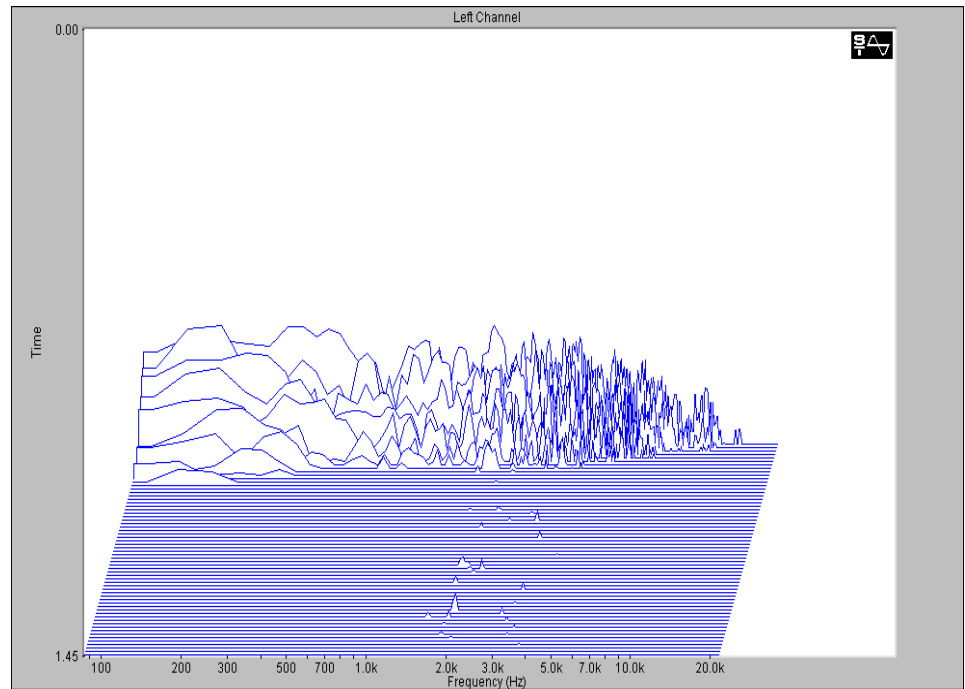


Figura 4.18 - Risposta della stanza in 3d ottenuta da Spectra Lab

4.4 Auralizzazione

L'auralizzazione (analogamente alla visualizzazione nel caso delle immagini) è il processo che prevede, tramite il filtraggio di un segnale anecoico, la riproduzione alle orecchie dell'ascoltatore della sensazione psicoacustica propria di un determinato ambiente, inclusa la riverberazione, gli echi, la risposta in frequenza e le impressioni spaziali. L'idea su cui si basa la tecnica binaurale è la seguente:

Lo stimolo all'udito avviene tramite due segnali: le pressioni sonore all'ingresso delle orecchie. Se si registrano tali pressioni tramite microfoni miniaturizzati posti nelle orecchie dell'ascoltatore e si riproducono tramite cuffie interaurali, si può considerare di aver riprodotto esattamente l'esperienza uditiva, compresi gli aspetti timbrici e spaziali. Il termine registrazione "binaurale" si riferisce al fatto che i due segnali, inviati alle rispettive orecchie, sono riprodotti correttamente.

La registrazione può avvenire ponendo i microfoni all'interno dell'orecchio umano, anche se, normalmente, si usano teste artificiali che riproducono sia dal punto di vista geometrico, sia come consistenza del materiale, la parte superiore del torso umano, il collo, il viso incluso naso e bocca, e le orecchie che, in alcuni modelli, conservano anche l'impedenza acustica reale. Usando una esatta riproduzione di una testa umana si è certi che le onde sonore raggiungono il canale uditivo seguendo gli stessi percorsi del caso di un ascoltatore reale.

La riproduzione di solito è ottenuta tramite cuffie stereofoniche; recentemente sono state realizzate anche microcuffie interaurali di elevata qualità che, potendosi inserire totalmente nell'orecchio, proprio nella stessa posizione in cui erano stati posti i microfoni, riproducono esattamente la situazione reale.

La riproduzione, in camera anecoica, tramite altoparlanti, introduce un effetto di crosstalk indesiderato, poiché il suono riprodotto da ciascun altoparlante viene percepito da entrambe le orecchie.

L'idea base della tecnica della registrazione binaurale non è nuova. La descrizione dell'idea, delle sue applicazioni e dei dettagli sulla trasmissione del suono dalla registrazione con testa artificiale fino alla riproduzione nell'orecchio dell'ascoltatore si possono trovare nella letteratura scientifica da circa 60 anni.

Sebbene si sia già dimostrato come, tramite un'opportuna equalizzazione della registrazione microfonica, si possano preservare le caratteristiche timbriche dell'evento sonoro originale anche se la riproduzione avviene con un comune sistema di casse acustiche, c'è ancora disaccordo su come farla, in particolare se sia corretto equalizzare il campo diretto o quello diffuso. In ogni caso, la riproduzione spaziale della tecnica binaurale non è ottenibile, ma, nonostante ciò, la qualità d'ascolto ottenibile anche con normali impianti hi-fi è comparabile, sia dal punto di vista timbrico, sia da quello della ricostruzione sonora, a quella ottenibile con le normali tecniche di registrazione dei CD.

Il più grande impedimento al successo della tecnica binaurale è probabilmente legato alla localizzazione frontale del suono. Le sorgenti che in origine sono collocate nella semisfera frontale sono spesso percepite come se fossero dietro all'ascoltatore o appaiono più vicine all'ascoltatore di quanto non lo fossero in fase di registrazione. A volte capita che siano percepite come provenienti dall'interno della testa stessa. Tale problema è noto come "eccesso di stereofonia" e si presenta, in misura minore, anche nella riproduzione in cuffia di CD registrati con la tecnica tradizionale.

Una spiegazione di questo fenomeno è che le caratteristiche, geometriche e di impedenza acustica, sono diverse da soggetto a soggetto, e solo una registrazione effettuata usando la testa dell'ascoltatore stesso può garantire una precisa localizzazione frontale.

Un'altra spiegazione è che l'uomo utilizza piccoli movimenti della testa per distinguere il davanti e il dietro. Una rotazione del capo verso destra permette che il suono che arriva da una sorgente frontale arrivi prima all'orecchio sinistro e dopo a quello destro. Per una sorgente posta dietro l'ascoltatore avviene l'opposto. Una registrazione binaurale non reagisce ai movimenti della testa e ciò può spiegare i problemi legati alla riconoscibilità fronte/retro. Comunque è stato dimostrato che l'uomo riesce a distinguere se una sorgente a larga banda sia davanti o dietro anche tenendo la testa ferma. Di conseguenza, la mancanza di reattività ai movimenti della testa non può essere l'unico motivo del problema.

Sebbene gravata dai problemi sopra esposti, la tecnica binaurale è superiore alle altre tecniche di registrazione. Se usata propriamente produce in modo molto realistico l'impressione di essere stati presenti durante la registrazione, e gli ascoltatori rimangono molto spesso sorpresi della verosimiglianza della sensazione.

Possibili utilizzi della tecnica binaurale

Si assume che tutto il processo di registrazione e di riproduzione binaurale avvenga in modo corretto e che la testa artificiale usata riproduca esattamente le caratteristiche della testa dell'ascoltatore.

In questo caso, per una caratteristica intrinseca alla tecnica binaurale stessa, ciò che sente l'ascoltatore è esattamente quello che avrebbe sentito se fosse stato presente nel momento della registrazione. Ciò

rende la tecnica di registrazione binaurale superiore alle altre tecniche di registrazione. Essa dà la migliore rappresentazione del suono originale, non solo dal punto di vista timbrico, ma anche dal punto di vista della ricostruzione spaziale.

Le applicazioni della tecnica binaurale si possono dividere in tre categorie:

- applicazioni che implicano la registrazione con testa artificiale;
- applicazioni che utilizzano segnali binaurali elaborati da un computer;
- applicazioni varie.

Registrazioni con testa artificiale

L'applicazione più ovvia della tecnica binaurale è la semplice registrazione con una testa artificiale di un evento musicale, per esempio un concerto, che poi verrà riprodotta mediante l'uso di una cuffia stereofonica. Questo è il normale utilizzo che va dalla registrazione alla riproduzione. E ci sono molte possibilità per il suo uso nella produzione di film e di video, connesse anche con la resa degli "effetti speciali". Ma oltre a ciò ci sono molte altre possibilità di utilizzo della tecnica binaurale in sé e del know-how ad essa associato.

- **Valutazione del rumore**
Per la valutazione del rumore è ben noto che il livello del suono ponderato A o altre misure oggettive sono spesso insufficienti per la caratterizzazione di un rumore. La valutazione soggettiva dovrà diventare più diffusa, sebbene difficile da realizzare. Normalmente il gruppo di prova deve ascoltare molti tipi di rumore, e si devono utilizzare diverse tecniche di scala psicometrica. E' molto scomodo dover spostare tutto il gruppo di prova nei luoghi dove sono le reali sorgenti di rumore, è invece molto comodo poter disporre della registrazione di tutte le fonti di rumore in un laboratorio e procedere alla prova senza doversi mai spostare. Ma una semplice riproduzione mediante casse acustiche non dà al gruppo d'ascolto la sensazione di essere "immersi nel rumore".
Con la tecnica di registrazione binaurale può essere riprodotta in laboratorio la corretta impressione sonora, incluse anche le caratteristiche di direttività. Quindi si possono riprodurre facilmente anche molti rumori diversi e sottoporli al gruppo d'ascolto e le scale di valutazione psicologiche possono essere opportunamente calibrate. E' anche possibile riprodurre il rumore dopo aver modificato o filtrato la registrazione, ottenendo così la riproduzione dello stesso rumore dopo qualche tipo di intervento di bonifica.
- **Confronto di sale da concerto**
E' normalmente abbastanza difficile confrontare l'acustica di diverse sale da concerto, perché la memoria umana per le impressioni sonore è molto corta e, in ogni caso, non abbastanza lunga per spostarsi da una sala all'altra. Registrazioni binaurali effettuate in sale diverse, se possibile con la stessa orchestra, rendono possibili confronti diretti fra due o più sale da concerto.
- **Teleconferenze e registrazioni di convegni**
Normalmente i partecipanti ad un congresso non hanno problemi nel distinguere le parole di un particolare relatore, anche quando più persone parlano simultaneamente. Ciò è dovuto al "cocktail-party effect", ovvero a quel fenomeno per il quale il cervello riesce ad estrarre informazioni riguardo i segnali di una specifica sorgente sonora anche quando sono presenti più sorgenti, anche di livello superiore. Se un convegno è registrato o trasmesso con le tec-

niche tradizionali, tale effetto va perso, e può diventare difficile comprendere ciò che sta dicendo una particolare persona. La trasmissione binaurale conserverebbe il "cocktail-party effect" e ristabilirebbe le capacità di ascolto selettivo all'ascoltatore. La tecnica binaurale può diventare, quindi, molto utile per la trasmissione di convegni e dibattiti. Può risultare molto utile anche per una segretaria che debba trascrivere la registrazione di un convegno.

- Valutazione del parlato in un ambiente
Sono state sviluppate molte nuove tecniche per la caratterizzazione oggettiva dell'intelligibilità e della qualità del parlato negli ambienti chiusi. La registrazione binaurale può costituire un mezzo per la valutazione soggettiva.

- Valutazione dell'effetto di un impianto di amplificazione e diffusione sonora in un ambiente

La qualità del suono emesso da un sistema di altoparlanti è molto influenzata dall'ambiente in cui sono installati. Per giudicare un sistema di altoparlanti sarebbe quindi necessario ascoltarlo in ambienti ed in posizioni diverse, in più tale ascolto si dovrebbe ripetere in più posizioni dell'ascoltatore per ogni ambiente. Gli spostamenti necessari, uniti al poco tempo per cui una persona può ricordare un'impressione sonora, rendono un tale tipo di prova praticamente impossibile.

Le registrazioni binaurali delle emissioni del sistema di diffusori nelle varie posizioni dei vari ambienti e nelle varie posizioni dell'ascoltatore, possono risolvere il problema. Semplificano la presentazione della resa acustica a più persone e rendono possibile cambiare ambiente o posizione in tempo reale.

- Controllo della diffusione sonora
Nei teatri e durante concerti amplificati elettronicamente, tramite la consolle di mixaggio, l'impianto deve essere regolato per ottenere le migliori prestazioni. La consolle di mixaggio può essere collocata in una cabina di regia o nella sala. Nel primo caso, l'operatore non può controllare direttamente la qualità del suono che deve regolare. Ma installare la consolle di regia audio nella sala da concerto ha la conseguenza economica che l'operatore e le apparecchiature occupano alcuni fra i migliori posti. Ciò comporta una diminuzione della capienza della sala e l'inconveniente di far lavorare "in pubblico" l'ingegnere del suono. Una testa artificiale collocata fra il pubblico e la trasmissione del suo segnale binaurale alla cabina di regia può risolvere il problema.

Utilizzo di segnali binaurali artificiali

La funzione della testa, artificiale od umana che sia, è di trasferire il suono da una determinata direzione e distanza in due segnali. Se sono note la direzione e la distanza dalla sorgente sonora, la funzione di trasferimento per quella direzione e quella distanza per entrambi i segnali, allora la funzione di trasferimento binaurale può essere simulata con filtri elettronici. Tali filtri possono essere implementati in circuiti sia analogici, sia digitali. Se sono disponibili un numero sufficiente di funzioni di trasferimento per diverse direzioni e distanze, una sorgente sonora può essere posta elettronicamente in un punto qualsiasi dello spazio. Un'applicazione molto importante della tecnologia binaurale è quindi costituita dalla realizzazione di "ambienti virtuali".

- Consolle di mixaggio binaurale
Una consolle di mixaggio binaurale è un apparecchio elettronico che contiene filtri che sono in grado di trasformare un segnale a-

necoico in un segnale binaurale corrispondente ad una data direzione e distanza. E' possibile variare distanza e direzione proprio come con un potenziometro del panpot di una normale consolle di mixaggio. Naturalmente è possibile mixare contemporaneamente più sorgenti sonore in posizioni diverse.

La consolle di mixaggio binaurale può essere ancora migliorata ed includere anche qualcuna delle prime e più importanti riflessioni, oltre all'onda diretta. Le riflessioni possono essere ritardate a seconda delle caratteristiche di una sala da concerto e filtrate a seconda del loro corretto angolo di incidenza. Il mixer può anche aggiungere un effetto di riverberazione diffusa.

- Simulatore di ambienti binaurale

Il concetto di mixer binaurale si può sviluppare sempre più fino a diventare un simulatore di ambienti binaurale. In questo apparecchio i filtri, che sono introdotti a valle del segnale della sorgente, monofonico e anecoico, contengono le esatte funzioni di trasferimento fra la posizione della sorgente e ciascuna delle due orecchie dell'ascoltatore in una specifica posizione e in una specifica sala. Queste funzioni di trasferimento sono chiamate "risposte all'impulso binaurali della sala" (da non confondere con le "risposte all'impulso della sala" che si riferiscono alla trasmissione da una sorgente ad una pressione sonora in un dato punto senza la testa). Con un simulatore di ambienti binaurale è possibile ascoltare lo stesso suono in diverse sale da concerto, e confrontare diverse posizioni della sorgente e/o dell'ascoltatore nella stessa sala.

La distinzione fra una consolle di mixaggio binaurale ed un simulatore di ambiente binaurale non è precisa. Se il mixer include le riflessioni ed il riverbero, si avvicina ad un simulatore di ambienti, e se un simulatore di ambienti permette di gestire più di una sorgente per volta, ridiventa un mixer. Le caratteristiche di un mixer sono la possibilità di avere molti canali e la possibilità di cambiare in tempo reale la posizione delle sorgenti. Un simulatore di ambienti è caratterizzato da un'accurata ricostruzione del campo sonoro di una specifica sala e da un accesso meno immediato ai cambiamenti nella stessa (o al cambio di sala per confronti).

- Strumento di progettazione in acustica architettonica

Le risposte all'impulso binaurali di una sala, usate da un simulatore di ambienti, possono essere ottenute da misure sperimentali, ma si possono ottenere anche tramite un programma di simulazione acustica di tipo Ray Tracing o Sorgenti Immagine. Se si usa uno di questi programmi, è possibile verificare come si sente in una sala da concerto prima ancora che sia costruita. E' possibile valutare soggettivamente il suono con diverse forme dell'architettura della sala, ed è possibile confrontarla con sale già esistenti.

E' anche possibile valutare l'affidabilità di un programma di simulazione acustica. Infatti, se si fanno calcolare al programma le risposte all'impulso binaurali, è possibile convolverle con un segnale anecoico per poi confrontare il risultato con la convoluzione dello stesso brano anecoico con le risposte all'impulso binaurali sperimentali, cioè misurate realmente in quella sala (già esistente). In questo modo si possono fare confronti soggettivi.

Per l'uso insieme ad un simulatore di ambienti, bisogna conoscere le esatte risposte all'impulso, e tutto deve essere basato sui segnali nel tempo, incluse anche le informazioni di fase per ogni singolo cammino del segnale. Ma, normalmente, i programmi di simulazione acustica calcolano i parametri oggettivi di acustica ar-

chitettonica come il tempo di riverbero, l'indice di chiarezza, e le stesse risposte all'impulso su base energetica. La strategia deve quindi essere cambiata se si vogliono poter creare delle risposte all'impulso adatte per la convoluzione.

La difficoltà di creare questo anello di congiunzione è data anche dal fatto che le proprietà di riflessività dei materiali sono solitamente note come coefficienti di assorbimento energetico. Potrebbe essere necessario rimisurare i dati dei materiali secondo metodi che diano le corrette informazioni direttamente sia nel dominio del tempo, come una "risposta all'impulso delle riflessioni" o, nel dominio della frequenza, come una riflessione in ampiezza e fase come funzioni della frequenza.

I metodi di misura che possono fornire tali informazioni sono tecniche di various gating o di impulse averaging, di sequenza di massima lunghezza (*MLS*) e la time-delay spectrometry (*TDS*).

- Riproduzione simulata di sistemi di altoparlanti
Quando dei segnali stereofonici tradizionali, quali quelli di tipo X/Y e A/B, sono riprodotti tramite cuffia, si ha come risultato una eccessiva localizzazione dei suoni "dentro la testa". Questo è dovuto alla mancanza di crosstalk e di riflessioni da parte dell'ambiente. Basandosi sulla conoscenza delle funzioni di trasferimento fra ogni sorgente e le due orecchie in campo libero, si può aggiungere il crosstalk al segnale stereo originale. In più possono essere aggiunte le riflessioni corrispondenti ad una normale sala d'ascolto. Fino ad oggi non si conosce precisamente quanto si debbano correggere i normali segnali stereofonici per ottenere una corretta localizzazione "fuori dalla testa". Questa applicazione può essere di grande interesse per il mercato consumer, ma lo è soprattutto per la simulazione delle sale di regia audio. In tutti i campi della produzione della registrazione audio (industria discografica, cinematografica etc...) è attualmente necessario investire grossi capitali per realizzare apposite sale di ascolto dove tramite specifiche casse acustiche (dette monitor da studio) viene effettuato il controllo di qualità sui master delle registrazioni. E quasi sempre l'ingegnere del suono sente la necessità di riascoltare il master su vari tipi di impianto e in vari tipi di sala poiché non conosce quale tipo di impianto sarà usato dall'acquirente del disco. La tecnica binaurale può competere con le sale di ascolto o esserne un valido complemento.
- Simulazione di un sistema di diffusione sonora in un ambiente
Una combinazione di un simulatore di ambienti e di un simulatore della riproduzione di un sistema di altoparlanti può essere usata per valutare i risultati ottenibili con l'installazione di un sistema di amplificazione e riproduzione sonora in auditorium, teatri, aeroporti, stazioni ferroviarie, stadi, etc. L'apporto a tale tipo di progettazione è ovvio.

Altre applicazioni

- Protesi acustica binaurale
E' ben noto come soggetti audio-lesi abbiano grande difficoltà a seguire conversazioni che coinvolgano un certo numero di persone. Se sono sofferenti di una perdita uditiva, è teoricamente possibile correggere completamente il loro deficit uditivo con una protesi acustica. Se la protesi acustica raccoglie il suono in un punto che include tutte le informazioni sulla direzionalità e se viene fatta una corretta equalizzazione, allora viene ristabilita la capacità di

localizzare i suoni, e anche l'effetto "cocktail-party" dovrebbe venir ripristinato.

Un modo pratico per realizzare una tale protesi acustica consiste nel chiudere il canale uditivo e metterci davanti un microfono. La posizione della capsula microfonica deve essere così vicina all'ingresso del canale uditivo che la propagazione del suono da questo punto sia unidimensionale. L'elettronica e il microaltoparlante devono essere inclusi nel tappo. Questo apparecchio può essere definito protesi acustica binaurale (anche se di solito solo un orecchio ne ha bisogno e quindi se ne usa uno solo).

Non è chiaro se una protesi acustica binaurale possa migliorare la possibilità di localizzare le sorgenti sonore nei casi di una perdita neurale non lineare ottenendo un recupero del danno uditivo.

- Soppressione binaurale del rumore

Il "cocktail-party effect" fa uso di processi neuronali nel sistema uditivo. Se si riuscisse a descrivere in modo appropriato un modello di questo processo, si potrebbe implementare su di un computer che potrebbe realizzare la soppressione del rumore sulle basi di segnali registrati con testa artificiale.

'Il processo di auralizzazione: metodologia ed esemplificazione'; A. Farina, P. Galaverna, M. Giabbani

5. CONVOLVER PER RAMSETE

L'idea di fondo è quella di riuscire ad ottenere la convoluzione, in tempo reale, di due segnali audio con la qualità del CD utilizzando computer general purpose.

Il giudizio finale di ogni sistema di auralizzazione può essere dato solo con le orecchie, ascoltando realmente i segnali ottenuti. In più la "naturalità" del suono può essere valutata solo per confronto con il vero campo sonoro di una sala da concerto.

In una prima fase, si è preferito utilizzare come filtri *FIR* le risposte all'impulso sperimentali e non quelle di Ramsete

Le risposte all'impulso sperimentali contengono anche la risposta dell'altoparlante e dei microfoni: esse, in linea di principio, possono essere rimosse tramite una deconvoluzione, ma sono state lasciate sia nelle risposte all'impulso sia nelle registrazioni "live" della musica (riprodotta con lo stesso altoparlante e registrata con gli stessi microfoni). In questo modo, il risultato della convoluzione può essere confrontato direttamente con la registrazione digitale dello stesso brano musicale, emesso attraverso lo stesso altoparlante, posto in una sala da concerto e registrato con un DAT attraverso i microfoni della testa artificiale. Entrambe le registrazioni (quella derivante dalla convoluzione e la registrazione diretta), in via di principio, dovrebbero poter ricreare nell'ascoltatore le stesse sensazioni psicoacustiche che percepirebbe se si trovasse in quella sala da concerto.

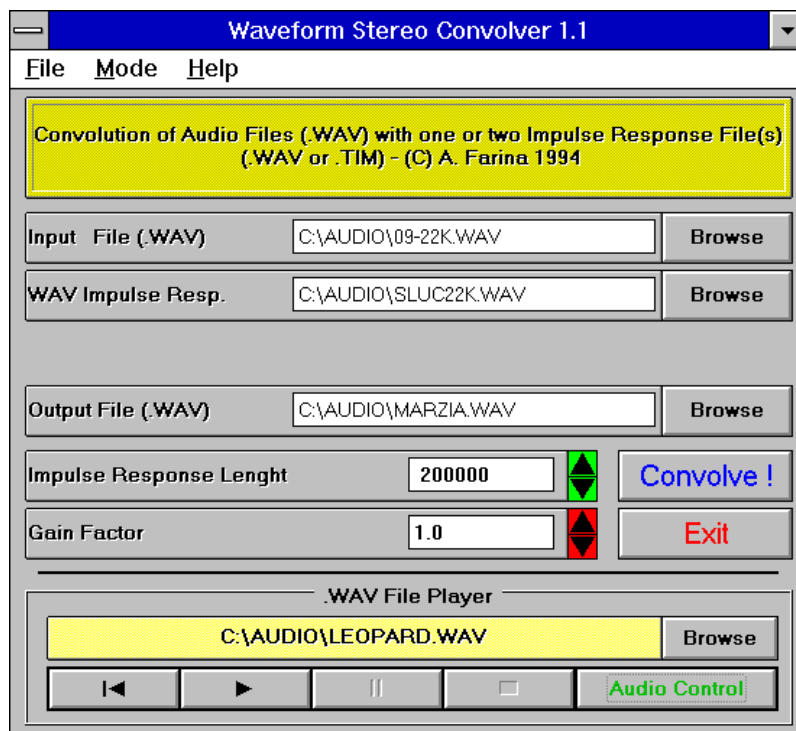


Figura 5.1 - L'interfaccia di Aurora.

Come visto al cap. 4.3 la convoluzione di un segnale continuo $x(\tau)$ con un filtro lineare caratterizzato dalla risposta all'impulso $h(\tau)$ produce un segnale d'uscita $y(\tau)$ dato dalla ben nota relazione:

$$y(\tau) = x(\tau) * h(\tau) = \int_0^{\infty} x(\tau - t) \cdot h(t) dt$$

Quando il segnale in ingresso e la risposta all'impulso sono segnali digitali, e quindi campionati ($\tau = i \cdot \Delta\tau$), e la risposta all'impulso ha una lunghezza finita N , tale integrale si riduce ad una sommatoria:

$$y(i) = \sum_{j=0}^{N-1} x(i-j) \cdot h(j)$$

Bisogna effettuare la somma di N prodotti per ciascun dato campionato, e ciò produce un numero molto alto di somme e moltiplicazioni. Questo calcolo avviene in virgola mobile per evitare overflow ed un rumore numerico eccessivo.

Il calcolo della convoluzione può essere significativamente semplificato sfruttando la *FFT* (Fast Fourier Transform) e la *IFFT* (Inverse Fast Fourier Transform), poiché una convoluzione nel dominio del tempo si riduce ad una semplice moltiplicazione nel dominio della frequenza tramite lo spettro complesso di Fourier del segnale di ingresso e della risposta all'impulso. Poiché l'algoritmo di *FFT* suppone che il segmento di segnale analizzato sia periodico, la semplice implementazione del calcolo nel dominio della frequenza produce risultati insoddisfacenti: bisogna rimuovere dal segnale d'uscita la periodicità causata dall'*FFT*.

Questo può essere fatto mediante due algoritmi, chiamati "overlap-add" e "select-save". Sotto si riporta il diagramma di flusso del secondo metodo:

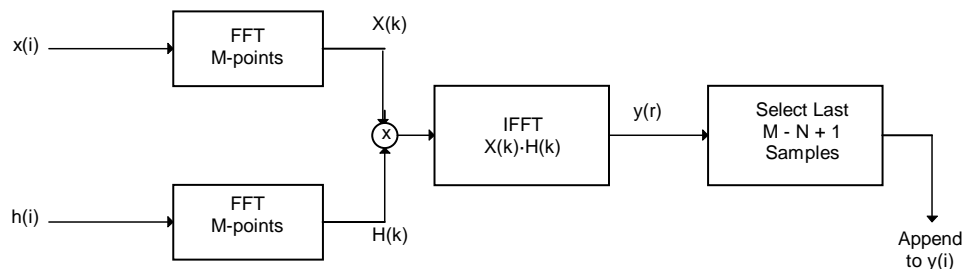


Figura 5.2 - Flow Chart dell'algoritmo di Select-Save.

Poiché l'algoritmo produce in uscita come risultato solo $M-N+1$ dati convoluti, la finestra iniziale di M punti deve essere shiftata verso destra sulla sequenza di ingresso esattamente di $M+N-1$ punti, prima di effettuare la convoluzione del segmento successivo.

Il tradeoff significa che è richiesta una *FFT* di lunghezza $M > N$. Tipicamente, il fattore 4 ($M = 4N$) è quello che consente la maggior efficienza all'algoritmo "select save": se N è 65536 (2^{16}), è necessario calcolare la *FFT* su sequenze di dati di lunghezza $M = 65536 \cdot 4 = 262144$ punti. Viene quindi richiesta un'allocazione di memoria molto grande, tipicamente 1 Mbyte, per memorizzare la sequenza di ingresso o il segnale d'uscita. Il fabbisogno totale di memoria per l'intero algoritmo di "select save" è quindi di qualche Mbytes.

Principalmente, l'algoritmo di "select save" riduce di molto il numero di moltiplicazioni in virgola mobile necessarie per la convoluzione. Ciascuna *FFT* o *IFFT* richiede $M \cdot \log_2(M)$ moltiplicazioni: una coppia di *FFT* e *IFFT* produce comunque in uscita $\frac{3}{4} M$ nuovi dati, e così il numero di moltiplicazioni per ciascun dato in uscita è intorno a 50, invece di 65536.

D'altra parte, la *FFT* è un'operazione molto pericolosa: il flusso dei dati è segmentato, e ciascun segmento è processato separatamente.

Particolarità numeriche possono influenzare in modi diversi i vari segmenti, ed il "rumore" di calcolo (che è udibile quando l'ampiezza del segnale è piccola) cambia da segmento a segmento. Per queste ragioni, la qualità dei segnali filtrati con il metodo dell'"true convolution" può risultare migliore di quella ottenuta con l'algoritmo di "select save", a patto che l'elevato numero di moltiplicazioni non introduca un rumore numerico aggiuntivo.

In più, l'algoritmo di "true convolution" è molto adatto per la codifica ed il calcolo parallelo: allocando una schiera di N processori, tutte le N moltiplicazioni possono essere eseguite simultaneamente; quindi la somma degli N risultati può essere realizzata in modo molto efficiente attraverso l'assegnamento dei dati in parallelo in una variabile scalare del C* "+=".

Post-processing per evitare il crosstalk nella riproduzione tramite altoparlanti

La figura mostra ciò che accade quando una coppia di altoparlanti è posta di fronte ad un ascoltatore: il segnale proveniente da ciascun altoparlante raggiunge entrambe le orecchie, in modo tale che all'ingresso del canale uditivo i canali destro e sinistro arrivano mixati.

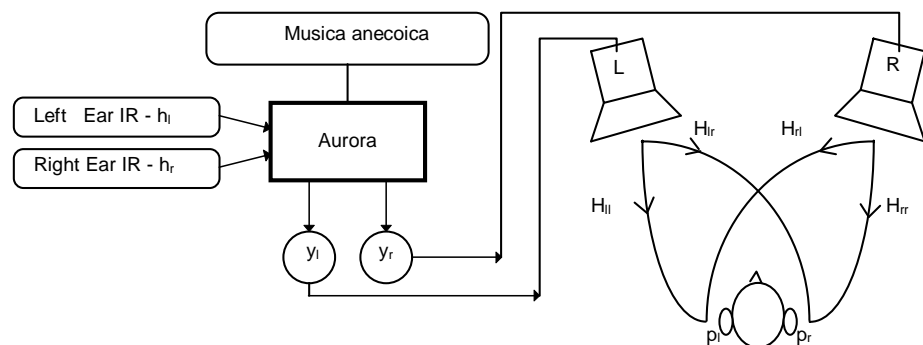


Figura 5.3 - Diagramma del percorso dei segnali che giungono alle orecchie attraverso gli altoparlanti.

In più la funzione di trasferimento relativa alla testa *HRTF* (Head Related Transfer Functions) è già compresa nelle risposte all'impulso usate per la convoluzione, ma il segnale che arriva dagli altoparlanti interferisce nuovamente con la testa dell'ascoltatore, e in questo modo il filtraggio dovuto alla testa viene operato due volte.

Per queste ragioni, è necessario usare un algoritmo che sottragga il filtraggio della testa contenuto nelle risposte all'impulso. In notazioni matematiche i segnali che arrivano ai canali uditivi sinistro e destro dell'ascoltatore possono essere così descritti:

$$\begin{aligned} p_l &= y_l * h_{ll} + y_r * h_{rl} \\ p_r &= y_r * h_{rr} + y_l * h_{lr} \end{aligned}$$

Notando che i segnali stereo y_l e y_r sono derivati da un singolo segnale d'ingresso anecoico mono (x) attraverso la convoluzione con le due risposte all'impulso h_l e h_r , e passando al dominio delle frequenze tramite *FFT*, le precedenti relazioni si possono così riscrivere:

$$\begin{aligned} P_L &= X \cdot (H_L \cdot H_{LL} + H_R \cdot H_{RL}) \\ P_R &= X \cdot (H_R \cdot H_{RR} + H_L \cdot H_{LR}) \end{aligned}$$

Ora, sostituiamo due risposte all'impulso adeguate (h_l e h_r) al posto delle originali: esse devono essere tali da rendere i termini fra parentesi uguali alle risposte all'impulso desiderate H_l e H_r . Dopo pochi passaggi si ha:

$$H'_L = \frac{H_{RR} \cdot H_L - H_{RL} \cdot H_R}{H_{LL} \cdot H_{RR} - H_{LR} \cdot H_{RL}}$$

$$H'_R = \frac{H_{LL} \cdot H_R - H_{LR} \cdot H_L}{H_{LL} \cdot H_{RR} - H_{LR} \cdot H_{RL}}$$

Questa semplice relazione permette di valutare le risposte all'impulso desiderate, eseguendo il calcolo nel dominio della frequenza per poi ritornare, se richiesto, nel dominio del tempo con una *FFT*.

Bisogna quindi misurare le quattro risposte all'impulso relative agli altoparlanti e alla testa dell'ascoltatore, il cui effetto è così completamente eliminato dal segnale audio da riprodurre. Ciò significa che, così facendo, viene eliminato non solo il crosstalk, ma anche qualsiasi altro filtraggio indesiderato dovuto alla risposta in frequenza degli altoparlanti. In più, viene eliminata anche la funzione di trasferimento relativa alla testa, evitando il doppio filtraggio esaminato precedentemente.

In teoria questo post-processing dovrebbe anche eliminare le riflessioni dell'ambiente d'ascolto, ed in questo modo non sarebbe più necessario trattare acusticamente la sala. Comunque, l'esperienza mostra che avviene un fenomeno sfortunato: la presenza delle riflessioni dovute all'ambiente d'ascolto, causano un effetto di comb filter nel dominio della frequenza, producendo difficoltà numeriche nel calcolo e risposte all'impulso modificate con ampi picchi che distruggono l'originale equilibrio della risposta in frequenza.

Esistono due soluzioni per risolvere tale limitazione: se l'ascolto non può svolgersi in una sala appositamente trattata, tramite i calcoli illustrati precedentemente, si fa in modo di estrarre solo la parte anecoica delle quattro risposte all'impulso sperimentali. In questo modo la compensazione digitale del crosstalk non tiene conto della presenza delle riflessioni dell'ambiente, ma esse sono ancora presenti, e causano ancora alterazioni spettrali e una certa quantità di crosstalk.

In alternativa, si può usare una piccola camera anecoica: anche se non è perfettamente anecoica (in particolare alle basse frequenze), è sufficiente per minimizzare le instabilità numeriche nel filtraggio. In questo caso le quattro risposte all'impulso non sono troncate, ma la loro lunghezza è ridotta applicando una finestra semi-gaussiana, per tagliare fuori il rumore dalla coda riverberante.

Tutte le risposte all'impulso richieste per la compensazione del crosstalk possono essere facilmente misurate con lo stesso sistema già introdotto per la misura in sale da concerto.

Ramsete (come molti altri programmi di Ray Tracing o Beam Tracing) produce delle risposte all'impulso energetica di lunghezza illimitata, campionate ad intervalli tipicamente di 10, 5 o 1 ms, per ciascuna banda di ottava (da 31.5 a 16000 Hz). Per ottenere una simulazione binaurale di qualità soddisfacente, si collocano due ricevitori ai poli di una sfera diffrangente, che approssimano le orecchie, come mostrato in figura. In questo modo, la risposta in un punto ricevitore è formata da una coppia di matrici, ciascuna con dieci colonne (le bande di frequenza) e da qualche centinaio di migliaia di righe (gli intervalli tem-

porali). ('Tecniche di convoluzione applicata al trattamento di segnali sonori per prove di ascolto', A. Farina).

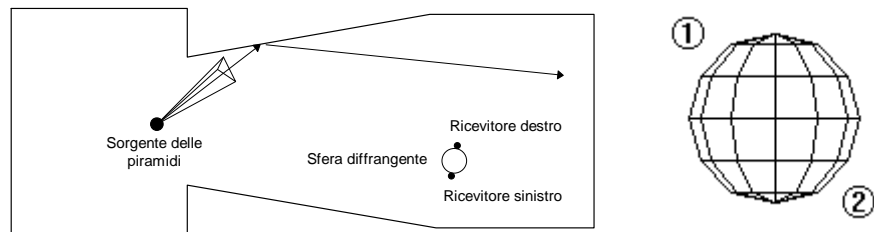


Figura 5.4 - Schema di Pyramid Tracing con ricevitore binaurale.

E' necessario convertire ciascuna di queste due matrici in una risposta all'impulso in pressione, con una banda estesa dai 20 ai 20000 Hz e campionata a 44.1 o (meglio) a 48 kHz. Solitamente queste risposte vengono memorizzate nel formato .WAV standard, o in un formato proprietario come il .TIM del sistema *MLSSA* e convertite con l'audioconverter.

Audioconverter

Il processo di conversione comincia con la generazione di un rumore bianco di lunghezza e frequenza di campionamento appropriate, in modo da coprire l'intera durata della risposta all'impulso. Tale segnale viene copiato dieci volte, e ciascuna copia viene modulata in ampiezza con l'involuppo di un segnale ottenuto dalla radice quadrata della risposta all'impulso energetica della corrispondente banda di frequenza. Successivamente ciascun rumore bianco modulato viene filtrato con un filtro *IIR* digitale a sei poli, con larghezza di banda pari ad un'ottava e centrato alla frequenza corrispondente. I dieci segnali così ottenuti sono poi sommati insieme, ottenendo così una risposta all'impulso in pressione a larga banda. Tale risposta viene poi normalizzata e convertita in interi a 16 bit.

Sebbene questo procedimento possa essere criticato sotto molti aspetti, esso risulta molto veloce e produce risposte all'impulso con le stesse curve energia/tempo, gli stessi tempi di riverbero, e praticamente gli stessi valori dei parametri acustici oggettivi delle risposte all'impulso energetiche originali. Per dimostrare che questo procedimento da risultati affidabili è necessario un test di ascolto soggettivo. Esiste un problema sulla risposta in frequenza complessiva. Spesso le simulazioni con Ramsete sono fatte con una sorgente omnidirezionale teorica (*omni.spk*), con un livello di potenza di 110 dB per ciascuna ottava. E' anche possibile usare delle sorgenti che emulano altoparlanti delle marche più note, con il proprio caratteristico balloon di direttività: anche in questo caso il calcolo viene effettuato come se la sorgente fosse di rumore rosa, poiché ciò rende più semplice controllare l'andamento dello spettro finale in ottave. D'altra parte, una risposta all'impulso non filtrata a nessuna frequenza risulta bianca e non rosa. Si rende quindi necessaria una conversione rosa-bianco che è direttamente implementata nel procedimento descritto prima.

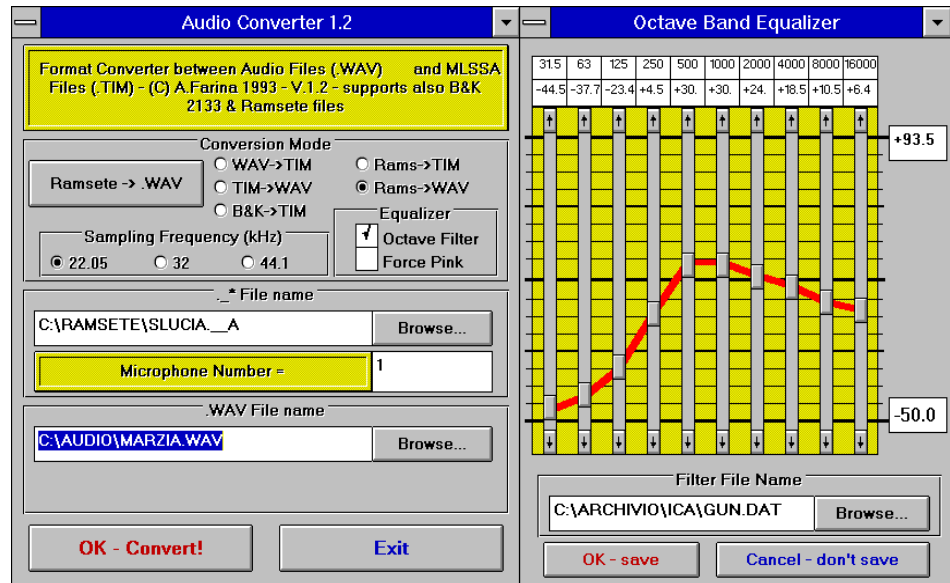


Figura 5.5 - Il convertitore di formato di Aurora.

Ma ci sono dei casi in cui nelle simulazioni effettuate con Ramsete si utilizzano sorgenti derivate da dati già bianchi: per esempio se il file dati della sorgente proviene da misure realizzate con il sistema *MLSSA*. In questo caso la procedura descritta introduce un artefatto, aggiungendo un incremento di 3 dB/ottava della risposta all'impulso: tale errore può essere rimosso selezionando l'apposito flag durante il processo di conversione ("force pink").

Un altro caso sono le simulazioni effettuate con la sorgente *omni.spk*, che vengono poi confrontate con misure sperimentali ottenute con una sorgente sonora non perfettamente piatta in frequenza. Il software di conversione è provvisto di un equalizzatore grafico a dieci bande che permette di riequilibrare la risposta in frequenza banda per banda, prima della loro somma. Se l'equalizzatore viene regolato conformemente all'analisi della risposta in frequenza della risposta all'impulso sperimentale (limitata all'onda diretta), si tiene già conto che lo spettro sia di tipo bianco, e così bisogna anche in questo caso settare il flag "force pink" per avere un incremento di 3 db/ottava nella risposta in frequenza.

Ascolto in cuffia

I risultati dei calcoli precedentemente illustrati, sono sequenze di dati che, dopo essere trasformati in formato intero, sono ascoltati. Ascoltando la registrazione con delle cuffie interaurali, si può cogliere tutta la naturalezza che il procedimento prima descritto consente alla riproduzione sonora. Il segnale convoluto è praticamente identico a quello registrato dal vivo, eccetto per il rumore di fondo, presente solo nel secondo. L'effetto della segmentazione provocato dall'algoritmo "select save", si può avvertire solo in istanti di silenzio seguiti da musica, e solo da audiofili molto allenati.

Ascolto tramite altoparlanti

Per la riproduzione tramite altoparlanti, si utilizza una camera anecoica miniaturizzata, per poter facilmente rimuovere il crosstalk, come già spiegato in precedenza, ma rimane comunque difficile togliere il riverbero.

Il processo di convoluzione deve essere eseguito in modo diverso per l'ascolto tramite altoparlanti, dato che si devono utilizzare le risposte all'impulso modificate al posto di quelle originali. Ma anche la musica registrata "dal vero" necessita dell'eliminazione del crosstalk, ma, in questo caso, essendo il segnale stereo già di partenza, il tempo di calcolo è il doppio di quello necessario per le registrazioni anecoiche monofoniche.

Le risposte all'impulso H_{LL} , H_{LR} , H_{RL} e H_{RR} , necessarie per il calcolo, variano con la posizione e le dimensioni dell'ascoltatore dentro la cabina. Per questa ragione, prima di ogni test d'ascolto, si effettuano le misure delle quattro risposte all'impulso con il sistema MLSSA, mentre l'ascoltatore è già seduto ed indossa una cuffia per registrazioni e riproduzioni binaurali. Tali risposte all'impulso sono utilizzate per produrre le due risposte all'impulso per la convoluzione, e la musica verrà così convoluta in modo diverso per ciascun soggetto, anche se la musica preparata per un soggetto può, in prima approssimazione, andare bene anche per altri ascoltatori. Per questa ragione, non è ancora stato prodotta nessun nastro DAT per l'ascolto tramite altoparlanti.

L'esperienza di ascolto nella cabina porta a risultati entusiastici, poiché si è eliminato il problema della confusione fra la localizzazione delle sorgenti sonore anteriori e posteriori, ottenendo una fedeltà d'ascolto eccezionale.

Test soggettivi – validazione del processo di convoluzione

Per validare l'algoritmo di conversione da risposta all'impulso energetica a risposta all'impulso in pressione del Convolver sono quindi necessari test soggettivi, poiché è il punto più delicato della simulazione numerica del campo acustico al fine della sua auralizzazione.

Per il test sono stati impiegati due ambienti chiusi: il primo è la Basilica di S. Domenico a Foligno ed il secondo è il palazzetto dello sport di Modena. Gli stessi ambienti sono stati utilizzati per la verifica dei parametri oggettivi. La figura mostra una vista prospettica dei due edifici. Nella Basilica sono stati scelti due ricevitori: uno molto vicino alla sorgente (con un alto rapporto fra campo diretto e campo riverberato), e uno molto distante. Nel Palazzetto, invece, è stato utilizzato un solo ricevitore poiché il campo sonoro è abbastanza simile ovunque.

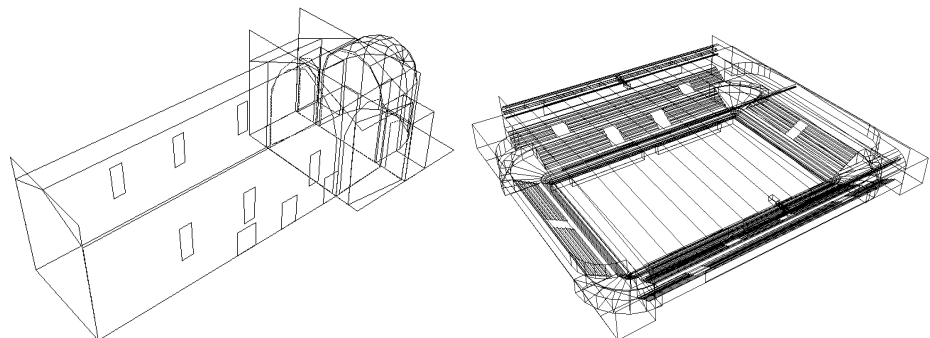


Figura 5.6 - Prospettiva delle due sale.

Sono stati utilizzati due spezzoni di brani anecoici: uno di musica (Ouverture "Le Nozze di Figaro" di Mozart, bars 1-18) e l'altro di parlato (una nota poesia del Leopardi "Il Sabato del Villaggio"). I due brani sono stati trasferiti digitalmente sull'hard disk tramite CD-ROM.

Sono stati convoluti entrambi sia con le risposte all'impulso simulate numericamente, sia con quelle misurate sperimentalmente. Si sono quindi ottenute sei coppie di brani, il cui ordine di presentazione è stato scelto casualmente. Per ottenere dei dati comparabili, sono stati mescolati due "gruppi di controllo" ciascuno composto da sei coppie di brani con i "confronti veri", sei coppie "veramente uguali" (ottenute riproducendo due volte lo stesso brano) e sei "realmente diverse" (ottenute con associazioni casuali di brani provenienti da sale diverse).

La cassetta è stata fatta ascoltare ad un campione di persone fra i 24 ed i 40 anni, per metà ascoltatori allenati e per metà ascoltatori comuni. A quest'ultima metà sono stati illustrati i concetti base per poter rispondere alle domande ed è stato necessario, in qualche caso, spiegare loro che i tipi di coppie erano tre e questo ha portato ad un aumento della sensazione di differenza fra i brani, prova ne è il fatto che la percentuale che ha riconosciuto come uguali brani veramente diversi è molto inferiore a quella che ha riconosciuto come diversi brani veramente uguali e ciò si è ripercosso negativamente anche nella discriminazione delle coppie vere a danno di Ramsete.

Un elemento di disturbo che ha portato ad un aumento del riconoscimento di brani come diversi è stato il montaggio della cassetta. Fra i due brani da confrontare passava troppo tempo e nella pausa (in cui veniva annunciato il secondo brano: "brano B") era presente un forte rumore di fondo, se raffrontato a quello virtualmente nullo dei brani da confrontare, questo è stato comunemente giudicato come negativo ai fini di una corretta valutazione.

Al gruppo di ascolto è stato chiesto di compilare per ciascuna coppia il seguente questionario:

Pair no. Are the two samples A and B equal? yes no

If Your response is no, explain why:

	a lot (-2)	slightly (-1)	no difference (0)	slightly (+1)	a lot (+2)	
A is more reverberant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	B is more reverberant
A has pronounced bass	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	B has pronounced bass
A has pronounced treble	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	B has pronounced treble
A has wide spatial impression	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	B has wide spatial impression

La seguente tabella riassume i risultati delle risposte alla prima domanda (percentuale di "uguaglianza"):

Ramsete/Experimental comparison	Truly equal samples	Truly different samples
27.1 %	70.8 %	4.2 %

Risultati delle risposte alla prima domanda.

Queste percentuali dimostrano che le simulazioni effettuate con Ramsete ingannano un soggetto su quattro. Di seguito sono riportate le

differenze medie e le deviazioni standard ottenute dal test, entrambe presentate in forma grafica e tabellare:

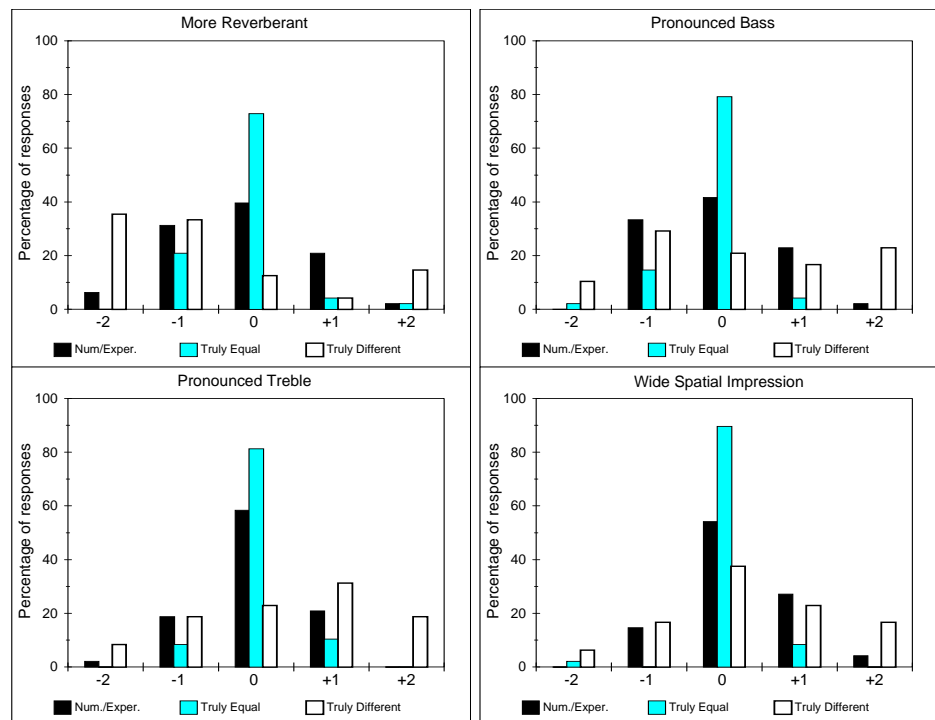


Figura 5.7 - Istogrammi delle statistiche delle quattro risposte.

	Ramsete/Exp. comparison	Truly equal samples	Truly different samples
more reverberant	-0.19 ± 0.90	-0.13 ± 0.56	-0.71 ± 1.37
pronounced bass	-0.06 ± 0.80	-0.15 ± 0.50	$+0.13 \pm 1.33$
pronounced treble	-0.02 ± 0.69	$+0.02 \pm 0.43$	$+0.33 \pm 1.21$
wide spatial impression	$+0.21 \pm 0.73$	$+0.04 \pm 0.41$	$+0.27 \pm 1.11$

Sorprendentemente le simulazioni ottenute da Ramsete danno quasi la corretta impressione dei quattro parametri soggettivi, poiché in media sono stati giudicati quasi uguali alle risposte sperimentali. Come ovvio si è manifestata una differenza più grande nel gruppo "realmente diverse", mentre nel gruppo "veramente uguali" i valori medi sono simili a quelli del confronto fra risposte numeriche e sperimentali. Questo risultato, che potrebbe sembrare strano, è probabilmente dovuto al fatto che le differenze nelle risposte "realmente uguali" sono state rilevate da soggetti poco affidabili, mentre le differenze fra i campioni numerici e quelli sperimentali sono state trovate dai soggetti più allenati all'ascolto. La deviazione standard del confronto numerico/sperimentale è maggiore di quella delle coppie "realmente uguali", mentre la deviazione standard del gruppo "veramente diverse" è ancora più grande.

Il risultato del test soggettivo dimostra che le risposte simulate non sono perfettamente indistinguibili da quelle sperimentali: un ascoltatore allenato può facilmente notare le differenze in un test comparativo. D'altra parte, tali differenze sono veramente molto piccole, poiché i principali parametri soggettivi sono giudicati uguali. Chiedendo ai soggetti del test di specificare la vera natura delle differenze, la mag-

gior parte di essi ha risposto "Ho avvertito una lieve differenza ma non sono in grado di identificare quale".

Probabilmente è necessario un test più approfondito, con un numero di soggetti più grande e che includa altri ambienti con differenti caratteristiche acustiche. Anche il questionario può essere migliorato basandosi sui suggerimenti dei componenti il gruppo d'ascolto.

Comunque il sistema Aurora ha dimostrato la capacità di produrre simulazioni binaurali realistiche che rendono possibile apprezzare anche piccole modifiche nel modello numerico della sala.

6. POTERE FONOISOLANTE, FONOASSORBENTE E FONODIFFONDENTE

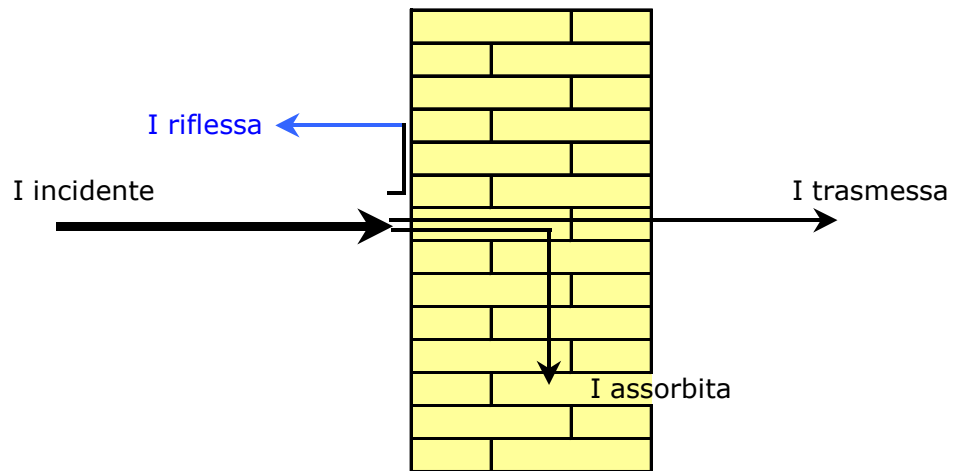


Figura 6.1 – Potere fonoisolante e fonoassorbente

Quando un'onda sonora prodotta all'interno di un ambiente incontra una parete la sua intensità sonora viene in parte riflessa, in parte assorbita dal muro stesso e in parte è trasmessa nell'ambiente adiacente; da cui derivano i coefficienti di riflessione, assorbimento e di trasmissione.

$$I_R + I_A + I_T = I_{INC}$$

da cui

$$r + \alpha + t = 1$$

Il coefficiente d'assorbimento acustico α ('Standing wave tube techniques for measuring the normal incidence absorption coefficient: comparison of different experimental setups', A. Farina, P. Fausti) contiene al suo interno sia l'assorbimento vero e proprio, sia il coefficiente di trasmissione. Una finestra aperta è un buon assorbente per un osservatore posto nell'ambiente che produce il suono (α grande) ma in realtà trasmette tutta l'intensità che riceve verso l'esterno: si può quindi pensare ad α come ad un assorbimento acustico apparente.

Il potere fonoisolante R ('Transmission loss measurements: validity of the sound intensity technique in laboratory and in the field', A. Farina, P. Fausti, R. Pompoli) si può definire come:

$$R = -10 \cdot \lg(t) \text{ [dB]}$$

Il segno meno fa sì che R sia positivo (infatti, $t < 1$ quindi il suo logaritmo è negativo). Il potere fonoisolante indica l'abbattimento in dB che il suono subisce passando attraverso una parete.

Da questa definizione si può ben capire che α e R sono due parametri distinti; i concetti che esprimono sono:

assorbire = non riflettere
isolare = non far uscire

Fonoassorbenti sono quindi materiali solitamente leggeri, morbidi, con impedenza simile a quella dell'aria (il massimo assorbimento si ha quando l'impedenza dei due materiali è uguale); perciò una finestra aperta ha assorbimento 1, che è l'assorbimento ideale.

Fonoisolanti sono invece materiali duri e pesanti (le onde sonore non riescono a deformarli) come per esempio una lastra di piombo.

Se si mette uno strato di materiale assorbente contro le pareti si otterrà un ambiente che assorbe meglio ma probabilmente fuori si sentirà più rumore. L'isolamento acustico presenta necessità differenti rispetto all'isolamento termico. Un ottimo isolante termico come il polistirolo (anche con l'aggiunta di materiali rigidi come il poliuretano) si comporta con il suono in modo da amplificarlo; spesso, infatti, gli isolanti acustici, che sono materiali a cellule aperte, sono pessimi isolanti termici (questi ultimi sono invece materiali a cellule chiuse) e viceversa.

Potere fonoisolante, in funzione della frequenza, per materiali e strutture diverse								
Materiale	Kg/m ²	Frequenze di centro banda[Hz]						Media
		125	250	500	1000	2000	4000	
Vetro (4 mm) in infisso di alluminio apribile		10	10	11	12	12	13	11
Lastra di vetro (4 mm)	10	20	22	28	34	34	29	28
Lastra di vetro (6 mm)	15	18	25	31	36	30	38	29
Lastra di vetro (12 mm)	30	26	30	35	34	39	47	35
Lastra di vetro-aria-vetro sigillata (4/12/4)		22	17	24	37	41	38	30
Lastra di vetro-aria-vetro sigillata (4/12/12)		25	22	33	41	44	44	35
Vetri doppi montati su infissi (6/150/4)		29	35	45	56	52	51	44
Vetri doppi montati su infissi (4/200/4)		27	33	39	42	46	44	39
Vetri su ante apribili da 25 mm (4/200/4)		15	23	34	32	28	32	27
Singolo laterizio intonacato da entrambi i lati	240	34	37	41	51	58	60	47
Muratura, doppia parete intonacata su due lati	480	41	45	48	56	58	60	61
Muratura leggera intonacata sui due lati (100 mm)		32	34	37	45	52	57	43
Muro leggero in laterizi liscio (200 mm)	250	35	38	43	49	54	58	46
Muro leggero intonacato su due lati (200 mm)		37	39	56	53	57	61	49
Muro in laterizi, tre strati, intonacato su due lati	720	44	43	49	57	66	70	55
Due strati di calcestruzzo, 100 mm, cavità 50 mm		35	41	49	58	67	75	52
Condotto rivestito gessolana minerale	30	11	13	12	12	12	21	12
Truciolato 50 mm, laminato dal lato sorgente	28	26	28	30	32	33	36	30
Truciolato 100 mm, laminato dal lato sorgente	50	28	28	32	34	33	38	31
Porta di legno pieno, 43 mm	28	17	21	26	29	31	34	26
Porta di acciaio, 50 mm, ben sigillata		21	27	32	34	36	39	32
Porta acustica di metallo, doppi sigilli		36	39	44	49	54	57	47
Pavimento in calcestruzzo rinforzato (200 mm)	460	42	41	50	57	60	65	53
Pavimento in calcestruzzo rinforzato (300 mm)	690	40	45	52	59	63	67	54

Il valore trovato con il metodo di normativa, R , deve essere poi corretto per essere inserito nel modello; bisogna quindi adattarlo all'ambiente in cui si deve eseguire il calcolo previsionale. Adattare vuol dire valutare il cosiddetto R' - potere fonoisolante apparente (apparente perché si considera della parete mentre il suono può essere arrivato nell'ambiente B percorrendo anche strade diverse). R' è il valore che si ottiene eseguendo le misurazioni in loco.

La legge stabilisce i valori minimi del potere fonoisolante apparente R' , non di quello misurato in laboratorio.

Solitamente per aumentare l'isolamento acustico si mette una parete doppia con in mezzo un materiale fonoisolante.

Quanto illustrato fino a questo punto rientra nella descrizione che generalmente si dà dei coefficienti che caratterizzano un materiale; in realtà, negli ultimi tempi, sta acquisendo sempre più importanza il concetto di potere diffondente di una parete e quindi anche la sua implementazione negli algoritmi di calcolo.

Il concetto di diffusione nasce dall'osservazione che i materiali non sono esattamente degli specchi tali per cui il raggio incidente viene semplicemente riflesso secondo la legge di Snell; in realtà il raggio viene sparato in tante direzioni, di cui quella principale è quella che segue le leggi della geometria ottica.

La figura riportata all'inizio di questo capitolo può essere così modificata:

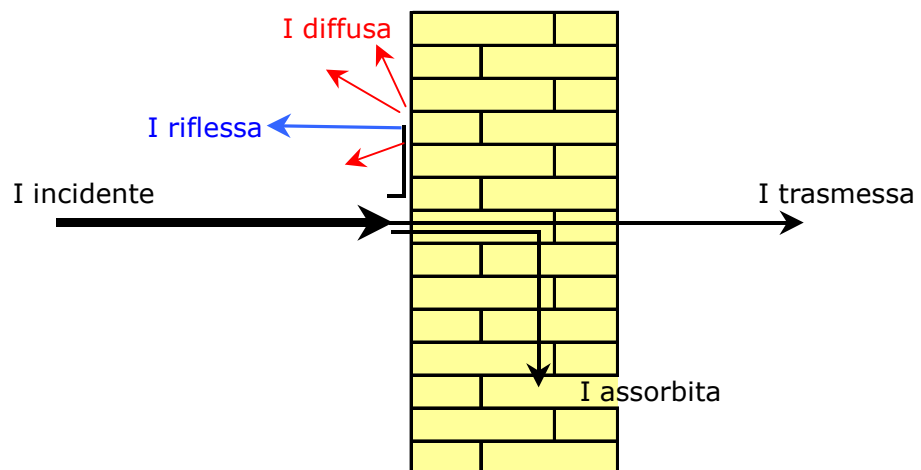


Figura 6.2 – Potere fonoisolante, fonoassorbente e fonodiffondente

La definizione di coefficiente di diffusione (o di scattering) non è semplice. Ad oggi sono diversi i gruppi di lavoro che hanno tentato di caratterizzare il comportamento diffusivo dei pannelli con un solo numero. Per l'AES tutto ciò 'che torna indietro' da una superficie finita va a costituire la diffusione; la stima del valore è effettuata ponendo N ricevitori omogeneamente spazati su una semisfera centrata nel punto di mezzo del pannello diffusivo e si definisce (I =intensità)

$$d = \frac{(\sum_j I_j)^2 - \sum_j (I_j)^2}{(N-1)(\sum_j I_j^2)}$$

pertanto d è nullo quando il suono riflesso si concentra su pochi punti di misura e si avvicina ad 1 quando si distribuisce uniformemente sulla semisfera.

ISO in realtà definisce il coefficiente di scattering s come il rapporto fra l'energia riflessa per diffusione e l'energia totale riflessa, ovvero:

$$s = \frac{E_{diff}}{E_{diff} + E_{spec}}$$

Il coefficiente così proposto si avvicina di più a quello che in genere i programmi di previsione acustica richiedono, ma il metodo per la determinazione del coefficiente è costoso e difficile.

Già dalle definizioni è possibile distinguere i diversi approcci adottati dai due gruppi. Per la determinazione del parametro è anche possibile fare riferimento anche alla tecnica del Wave Field Synthesis Analysis inizialmente sviluppato dalla Tech. University of Delft. Gli autori di Ramsete hanno preso spunto da questo metodo per separare l'onda diretta da quella riflessa e per definire la parte diffusa. Il problema che gli autori si sono posti è determinare quel valore tale che, inserito nel modello matematico, dia come risultati valori molto prossimi a quelli misurati. L'idea è di individuare il d tale da inglobare in se i fenomeni di diffusione e di bordo, questi ultimi dovuti al fatto che tutte le superfici, viste da vicino, hanno una struttura tale che i raggi sonori sono riderizzati. I valori così trovati saranno inseriti nella banca dati del Material Manager.

Nel caso di un modello basato sul tracciamento delle piramidi quale è Ramsete si verifica che:

tutti i ricevitori che si trovano all'interno del fascio riflesso della piramide riceveranno l'energia riflessa e sarà valutata come

$$I_{spec} = \frac{W \cdot Q_0}{4\pi(r_1 + r_2)^2} (1 - \alpha)(1 - s)$$

essendo

W	la potenza della sorgente sonora
Q_0	fattore di direttività della sorgente
r_1	distanza tra la sorgente e la superficie riflettente
r_2	distanza tra il ricevitore e la superficie riflettente
α	coefficiente di assorbimento della parete
s	coefficiente di scattering della superficie

tutti i ricevitori saranno interessati dall'arrivo di energia diffusa che si è prodotta con ogni riflessione

$$Idiff = \int \frac{W \cdot Q_0}{4\pi r_1^2} \cos(\varphi_1) \frac{(1 - \alpha) \cdot s}{2\pi r_2^2} Q_{\varphi_2} dA$$

essendo

φ_1	angolo tra la normale alla superficie ed il raggio incidente la superficie
φ_2	angolo tra la normale alla superficie ed il raggio che segue la legge di Snell
Q_{φ_2}	valore della direttività in direzione φ_2

A area della superficie.

Per approfondire l'argomento e per verificare come è stato implementato questo fenomeno fisico in un modello di previsione acustica basato sul tracciamento di piramidi (Ramsete) si rimanda alla lettura dei seguenti articoli: 'Validation of the numerical simulation of the scattered sound field with a geometrical pyramid tracing approach', 'Introducing the surface diffusion and edge scattering in a pyramid-tracing numerical model for room acoustics', 'Measurement of the surface scattering coefficient: comparison of the Mommertz/Vorlander approach with the new Wave Field Synthesis method', A. Farina; 'Measurement of sound scattering properties of diffusing panel through the Wave Field Synthesis approach', A. Farina, M. Zanolin, E. Crema

7. EFFETTI DIFFRATTIVI

Le onde sonore, come le onde elettromagnetiche, interagiscono con la materia, ma a differenza di queste ultime si ha un comportamento diverso se analizzate a bassa o alta frequenza. Questo comportamento è conseguenza del fatto che nel caso delle onde elettromagnetiche le fenditure sono un migliaio di volte maggiori della lunghezza d'onda incidente, mentre in questo caso, dove le grandezze in gioco hanno pressappoco la stessa dimensione, ci possono essere risposte diverse a seconda delle diverse lunghezze d'onda.

Bassa frequenza	Alta Frequenza
Ostacolo (un muro che si propaga perpendicolare al foglio)	
<p>Original wave front and secondary wave fronts from P interfere</p> <p>Original wave fronts from distant source</p> <p>P</p> <p>New wave fronts from P spread out cylindrically</p>	<p>Original wave front from distant source</p> <p>P</p> <p>Shadow region</p>
Il bordo esterno P diventa esso stesso sorgente di un'onda cilindrica	L'onda cilindrica non è completa, con la formazione di una zona d'ombra
Fenditura	
<p>New wave fronts spread out spherically from point P as source</p> <p>Original wave front from distant source</p> <p>P</p>	<p>Original wave front</p> <p>Shadow region</p> <p>Sound Beam</p> <p>Shadow region</p>
La fenditura è sorgente di un'onda sferica	Si forma un raggio, il quale è più colimato più è alta la frequenza.

Curve di Maekawa

Avevano ricavato che in campo libero il livello sonoro decresce con la formula

$$L_{dir} = L_w - 11 - 20 \log d .$$

Se tra la sorgente e il ricevitore poniamo un ostacolo, come cambia questo valore? L'ostacolo è uno schermo di altezza definita, sufficientemente spesso per evitare la propagazione attraverso la parte solida, e invece molto lungo. Un muro è un esempio. In questa condizione, a parità di dimensioni, si ottiene (per via sperimentale) che il materiale di cui è costituito il muro non influenza il decadimento.

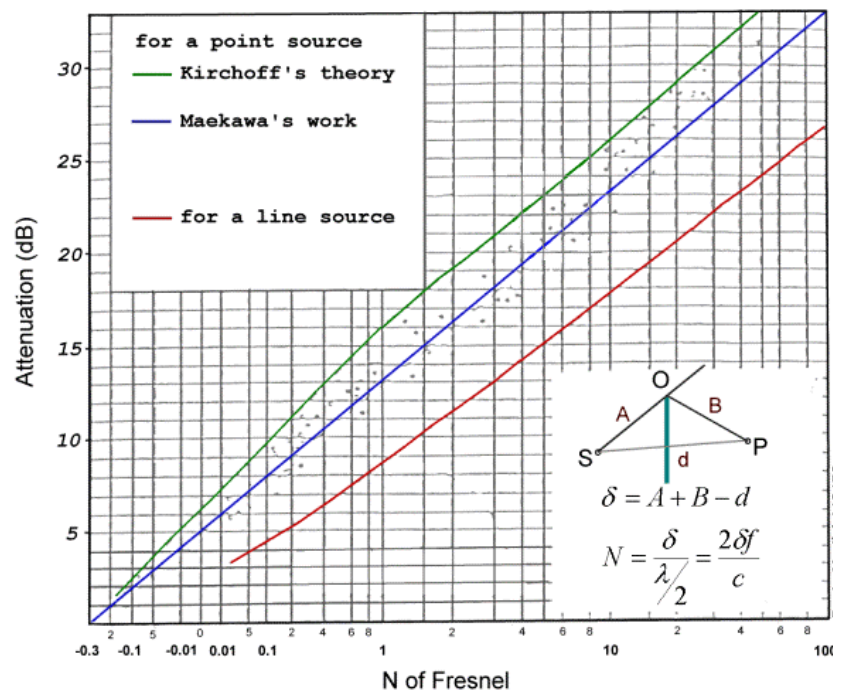
Maekawa per via teorica propose questa relazione:

$$\Delta L = L_{dir} - L_{schermo}$$

e

$$\Delta L = 10 \log(3 + 20N)$$

dove N è un numero puro detto di Fresnel:



Se la sorgente è lineare invece che puntiforme l'assorbimento è inferiore di circa 5 dB. Infatti, una sorgente lineare (l'autostrada) è più difficile da schermare con una barriera rispetto a una sorgente concentrata (la fabbrica).

8. SORGENTE DIRETTIVA

Si parla di sorgente direttiva quando l'intensità sonora emessa dalla sorgente non è distribuita uniformemente su tutte le direzioni e nello spazio.

Si definisce allora la grandezza direttività Q , nella particolare direzione θ come l'intensità emessa in questa direzione diviso l'intensità media.

$$Q = \frac{I_{\theta}}{I_m}$$

Quindi esiste un coefficiente numerico $Q(\theta)$ che dice quanto l'altoparlante guadagna rispetto alla sua emissione media in una particolare direzione e normalmente un altoparlante è dotato di una direttività che privilegia l'emanazione in direzione frontale

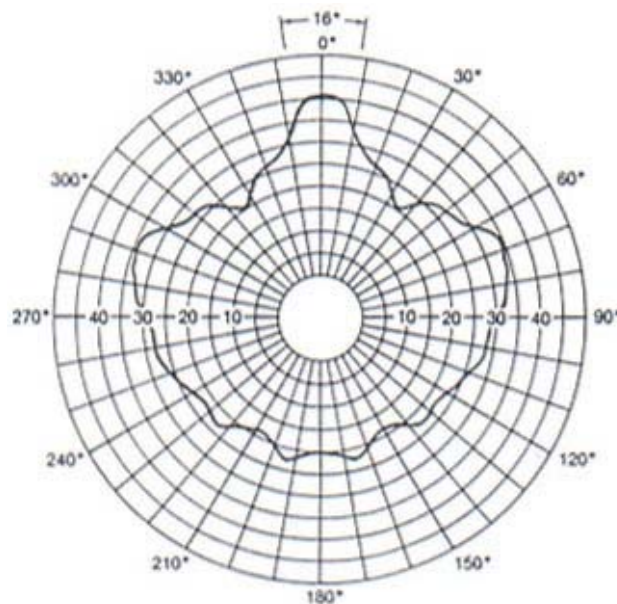


Figura 8.1 – Sorgente direttiva

La figura mostra quello che si dice un diagramma di radiazione, o diagramma polare, in cui si legge il fattore di direttività e l'angolo utile di radiazione, ovvero l'angolo entro il quale è massima l'emissione di radiazione sonora. Questo tipo di diagrammi sono ottenuti alimentando l'altoparlante ad una data frequenza e facendolo ruotare rispetto ad un microfono, tenuto fisso ad una assegnata distanza, il quale misura un parametro, in genere la pressione acustica. La lunghezza del segmento che unisce il centro del diagramma polare con un punto qualunque del diagramma stesso misura la differenza ΔL fra il livello di pressione in quella direzione e il corrispondente livello in direzione dell'asse del disco.

Considerando la direttività si ha che

$$L_D = L_W + 10 \log \left[\frac{Q}{4\pi d^2} + \frac{4}{\alpha S} \right]$$

Ovviamente Q può essere maggiore o minore di uno a seconda che ci si trovi entro o fuori dall'angolo di massima emissione. E' chiaro che essendo l'energia totale fissa se questa è maggiore in una direzione deve essere minore nell'altra. Quindi nella maggior parte delle zone il coefficiente Q è minore di 1.

ISO 3744

Come è stato illustrato nella Sezione 1 relativamente alla possibilità di produrre sorgenti di rumore, qui di seguito si illustrano i concetti della norma ISO 3744. Tale norma prevede la possibilità di effettuare i rilievi anche in spazi chiusi, apportando ai dati rilevati una opportuna correzione onde tenere conto della riverberazione dell'ambiente. Per sorgenti sonore di dimensioni medio-piccole, la norma prevede la possibilità di impiegare diverse superfici per involupparla: parallelepipedo con 5 o 9 microfoni, semisfera con 10 microfoni e superficie conforme ad 8 microfoni, come mostrato dalle figure seguenti:

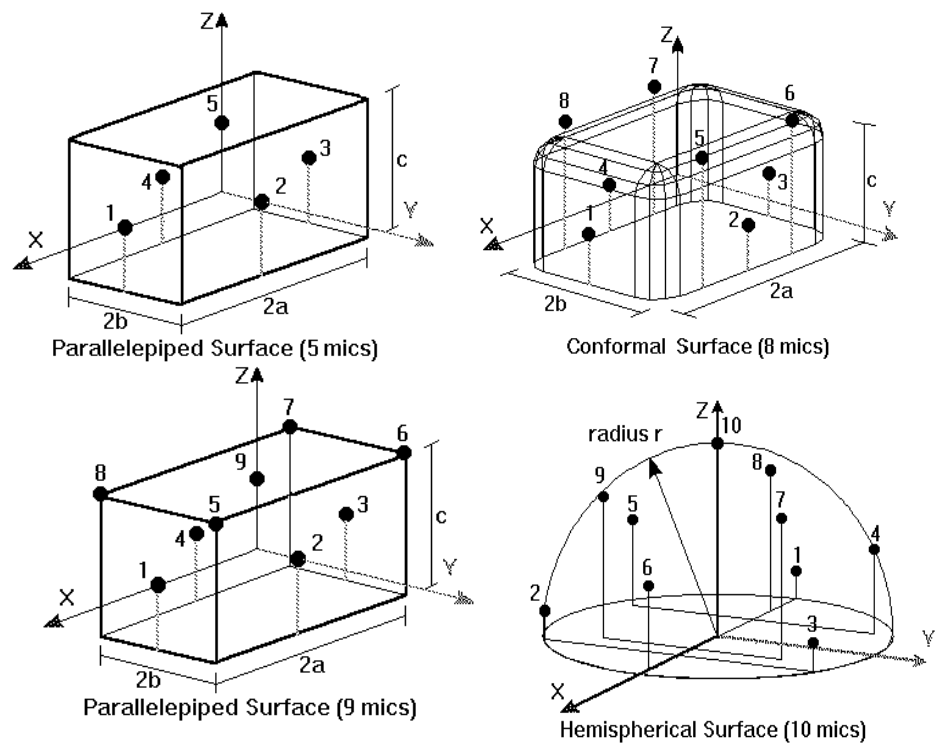


Figura 8.2 - Le quattro superfici di involuppo standard.

Per avere un dettaglio sulle modalità di calcolo dei 'baloon' a partire dai livelli di pressione sonora misurati in banda d'ottava nei punti indicati nelle direzioni definite dalla congiungente centro della sorgente-microfono si suppone che il livello sonoro rilevato sia costituito dalla somma energetica di due contributi. Il primo è l'onda diretta, pari a:

$$L_{dir} = L_W + 10 \log \left(\frac{Q_s}{4\pi r^2} \right)$$

il secondo è invece il contributo riflesso dal pavimento, e si suppone che in tale direzione si sviluppi un fattore di direttività pari alla metà di quello in direzione diretta. Pertanto il livello complessivo è dato dalla seguente espressione:

$$L_{tot} = L_w + 10 \log \left(\frac{Q_g}{4\pi r^2} + \frac{Q_g/2}{4\pi r'^2} + \frac{4T_{60}}{0.161V} \right)$$

ed, essendo note tutte le altre grandezze, si può ricavare il valore di Q_θ . La seguente figura illustra uno schema di questa costruzione:

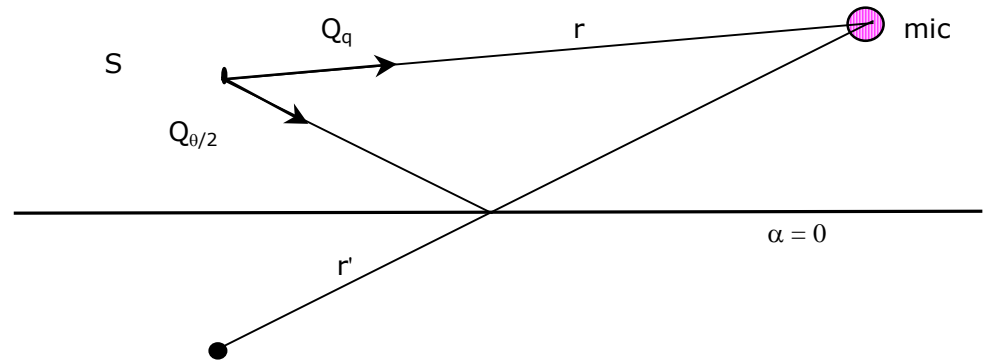


Figura 8.3 - Contributo dell'onda diretta e di quella riflessa dal pavimento.

Nelle altre direzioni si procede per interpolazione, assegnando la direttività come media pesata fra la direttività nella direzione del microfono più vicino e la direttività media (1 per definizione), impiegando come fattore di peso il coseno dell'angolo fra la direzione considerata e quella del microfono.

Al termine della costruzione del balloon, si è sicuri che esso è tale da riprodurre i livelli sonori nei punti di rilievo. Sorge tuttavia un problema: ricalcolando la potenza sonora come integrale sulla superficie del balloon dell'intensità emessa, esce in generale un risultato diverso dai valori di L_w ottenuti dalla applicazione diretta della ISO 3744.

Tale scostamento è generalmente piccolo, e tende a scomparire al crescere del numero di punti di rilievo; tuttavia, nel caso di rilievi con 5 microfoni e sorgente molto direttiva (utensili a disco), si sono trovati anche 4 dB di differenza! In queste condizioni, tuttavia, non è affatto chiaro se sia più corretto il valore indicato dalla ISO 3744 (che assegna implicitamente una uguale apertura angolare a ciascuna postazione microfonica) oppure quello ottenuto con la metodologia proposta (che viceversa assegna ad ogni microfono un'apertura angolare dipendente dalla sua posizione, in relazione agli altri).

In ogni modo, impiegando poi i dati in un programma di tracciamento di fasci divergenti (piramidi), si ottengono risultati corretti accoppiando i livelli di potenza definiti dalla ISO 3744 ai dati di direttività ottenuti con il metodo proposto. Infatti in tal modo, collocando dei ricevitori nei punti di rilievo, si ritrovano con buona approssimazione i livelli sperimentali.

Si verifica inoltre che, anche in punti situati più lontano dalla sorgente, la stima ottenuta con la tecnica proposta è sempre più accurata di quella ottenibile con l'ipotesi di sorgente omnidirezionale.

9. PARAMETRI ACUSTICI PER GLI AMBIENTI CHIUSI

C_{50} Klarheitsmass,
Gruppo di Dresda

Detta p la pressione si definisce:

$$C_{50} = 10 \log \frac{\int_0^{50ms} p^2(t) dt}{\int_{50ms}^{\infty} p^2(t) dt}$$

Si usa il logaritmo perché questo parametro è in db. È da notare come l'energia sia espressa come quadrato della pressione anche se come si è visto è una maniera impropria. Valori ottimali di questo indice variano fra -1 e 1 db, valori superiori a 1 db indicano un suono troppo secco, simile al suono all'aperto mentre valori inferiori a -1 indicano una eccessiva riverberazione del suono.

C_{80} Klarheitsmass,
Gruppo di Dresda

Il secondo indice è esattamente uguale al primo con l'unica differenza che l'intervallo di integrazione è esteso a 80 ms.

$$C_{80} = 10 \log \frac{\int_0^{80ms} p^2(t) dt}{\int_{80ms}^{\infty} p^2(t) dt}$$

Mentre il primo parametro è destinato alla valutazione di sale per il parlato il secondo è utilizzato nelle sale adibite all'ascolto di musica. Questo perché per una percezione gradevole del parlato si ha necessità di un legame meno presente che nel caso della musica.

D_{50}

Si definisce inoltre un indice di definizione

$$D_{50} = \frac{\int_0^{50ms} p^2(t) dt}{\int_0^{\infty} p^2(t) dt}$$

Questo è un parametro simile a C_{50} e che non aggiunge particolari informazioni allo stesso. La scelta di uno dei due dipende solo dall'abitudine del tecnico

T_{bar} Tempo Baricentrico,
Cremer, Kürer

Si usa il tempo baricentrico il cui pregio è la mancanza di un limite netto nell'integrazione che permette di ottenere valori più costanti.

$$t_s = \frac{\int_0^{\infty} t \cdot p^2(t) dt}{\int_0^{\infty} p^2(t) dt}$$

Questo parametro si misura in ms ed i suoi valori ottimali sono 30-80 ms per la parola e 50-120 ms per la musica.

Inoltre si definiscono dei parametri spaziali. L'aspetto spaziale del suono è infatti particolarmente importante nell'ascolto della musica in quanto le buone sale da concerto involuppano l'ascoltatore nel campo sonoro. Per fare misure relative al grado di spazialità del suono bisogna però fare misure con un altro tipo di microfono. In particolare si usano microfoni dotati di due capsule coincidenti, una omnidirezionale ed una a figura di otto orientata con il piano di sensibilità parallelo alle orecchie dell'ascoltatore. Questo microfono ha

due uscite indipendenti chiamate omni e figura di 8, la prima ci dá P^0 la seconda invece ci dá una velocità che viene erroneamente chiamata P^8 . Esiste una categoria di parametri basati su queste misure, questi parametri sono definiti dalla norma ISO 3382/97.

L_F , lateral fraction

$$L_f = \frac{\int_{5ms}^{80ms} p_8^2 d\tau}{\int_0^{\infty} p_0^2 d\tau}$$

L_E , Lateral Efficiency
Jordan

$$L_e = \frac{\int_{25ms}^{\infty} p_8^2 d\tau}{\int_0^{\infty} p_0^2 d\tau}$$

ITDG_{eq}, Initial Time Delay Gap Beranek
Tempo di ritardo fra suono diretto e la prima forte riflessione
D, indice di definizione

$$D = \frac{\int_0^{50ms} p^2(t) dt}{\int_0^{\infty} p^2(t) dt}$$

dove:

$p(t)$ pressione sonora, risposta all'impulso emesso da una sorgente non direttiva
 $t = 0$ istante in cui il suono diretto giunge nel punto di osservazione.

La proposta del parametro D scaturiva principalmente dai lavori di Haas sui legami fra percepibilità, disturbo, livello e ritardo di una riflessione che segue il suono diretto. I risultati di Haas, e di altri più tardi, confermarono che l'energia delle prime riflessioni, integrandosi con quella del suono diretto, contribuisce positivamente all'intensità soggettiva e quindi alla nitidezza del messaggio vocale. D è sempre compreso fra 0 ed 1, ma il suo valore ottimale per una sala da concerto è di circa 0.34.

R , Beranek e Schultz

Rapporto R fra l'energia riverberante e l'energia iniziale espresso in dB:

$$R = 10 \log \frac{\int_{50ms}^{\infty} p^2(t) dt}{\int_0^{50ms} p^2(t) dt} = 10 \log \frac{1 - D}{D}$$

Secondo Schultz, il criterio R rappresenta una misura fisica dell'equilibrio fra il mescolamento desiderato dei suoni dei vari strumenti dell'orchestra (Blend) e la chiarezza (Definition, Clarity).

EDT, Early Decay Time

Tempo di decadimento della curva di Schroeder da 0 a -10 dB

G , strenght

Differenza tra il livello sonoro misurato e quello prodotto da una sorgente sonora in campo libero omnidirezionale misurato a 10 m dalla stessa.

STI
Speed Transmission Index

Per definire questi parametri bisogna passare attraverso il concetto di intelligibilità.

e
RASTI
Rapid Speech Transmission Index

L'intelligibilità è solitamente definita come una proprietà del campo sonoro nel punto di ascolto, intendendo con ciò il campo sonoro all'esterno dell'orecchio dell'ascoltatore (che si presume normoudente). Recentemente, tuttavia, è stato fatto notare come la presenza dell'ascoltatore, ed, in particolare, la direttività del ricevitore "orec-

chio", siano in grado di influire sulla intelligibilità, che andrebbe, pertanto, valutata "al timpano" dell'ascoltatore (utilizzando, ad esempio, una testa artificiale). Si può pertanto ritenere che l'intelligibilità sia una variabile scalare, il cui valore varia nello spazio (e, normalmente, cresce in prossimità della sorgente sonora), e che dipende, oltre che dalla posizione, anche dall'orientamento del ricevitore. Tuttavia non esiste una "intelligibilità della sorgente", ma solo sorgenti che, in un determinato ambiente, danno luogo ad una intelligibilità più o meno elevata (a seconda del punto di ascolto considerato).

Si può dunque concludere che l'intelligibilità è sempre una proprietà del solo "canale di trasmissione", comprendendo in esso tutto l'insieme di apparati fisici, elettronici ed acustici che trasferiscono il suono originariamente emesso alle orecchie dell'ascoltatore. In tale "canale" deve essere compresa la particolare collocazione e l'orientamento dell'ascoltatore, ma non eventuali difetti uditivi dello stesso. Poiché normalmente si verifica una risposta lineare del canale di trasmissione, si può scomporre lo stesso in sezioni disposte in cascata, ciascuna delle quali produrrà un suo effetto sull'intelligibilità, che può essere quantificato misurando la stessa prima e dopo ciascuna sezione.

Si deve però precisare che l'intelligibilità dipende anche, in certa misura, dal tipo di segnale che deve essere trasferito: una sala da concerto può presentare un'ottima intelligibilità alla musica, ma quella per la parola può essere non così buona.

Come per la musica esistono diversi generi musicali, per ciascuno dei quali può verificarsi una diversa intelligibilità, così anche per il parlato esistono diverse lingue, inflessioni, intonazioni, esiste il canto in varie versioni, esiste la recita e la conversazione. Uno stesso canale di trasmissione può presentare diverse intelligibilità per ciascun diverso tipo di segnale acustico che vi viene introdotto. Ciò, però, non vuol dire che si possa definire una "intelligibilità all'origine", anche se corrisponde a verità che alcune lingue sono più facilmente comprensibili di altre. In realtà si osserva, allo stato attuale, che ben pochi metodi di misura oggettiva, tengono conto delle peculiarità del segnale originario e dell'orientamento della testa dell'ascoltatore: pertanto la misura dell'intelligibilità, con tecnica oggettiva, tende ad essere effettivamente funzione del solo canale di trasmissione.

Per le misure di intelligibilità della parola è di comune utilizzo l'indice *RASTI*, introdotto da T. Houtgast e H. J. M. Steeneken quale derivazione dello *STI*. Essi descrivono la trasmissione del suono del sistema tramite la *MTF*, (Modulation Transfer Function) ossia una funzione di trasferimento della modulazione: essa quantifica fino a che punto le modulazioni in ampiezza del segnale originale sono ridotte. Il calcolo della *MTF* viene effettuato paragonando la modulazione di intensità fra il segnale di uscita ed il segnale di entrata del sistema, e calcolando il fattore di riduzione $m(F)$, la cui relazione con la risposta all'impulso è stata ricavata da M. Schroeder, oltre che da T. Houtgast e H. J. M. Steeneken, e nel caso del campo riverberante (sotto le ipotesi di sabinianità) è così definito:

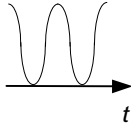
$$m(F) = \frac{1}{\sqrt{1 + \left(2\pi F \cdot \frac{T}{13.8}\right)^2}} \cdot \frac{1}{1 + 10^{(-S/N)/10}}$$

dove è:

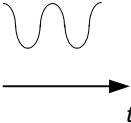
F frequenza di modulazione (per $1/3$ d'ottava)
 T tempo di riverberazione (s)
 S/N rapporto segnale/rumore (dB)

Il fattore di riduzione risulta essere il prodotto di due termini, ciascuno dei quali mette in relazione l'effetto negativo prodotto da eccessiva riverberazione e/o da eccessivo rumore di fondo.

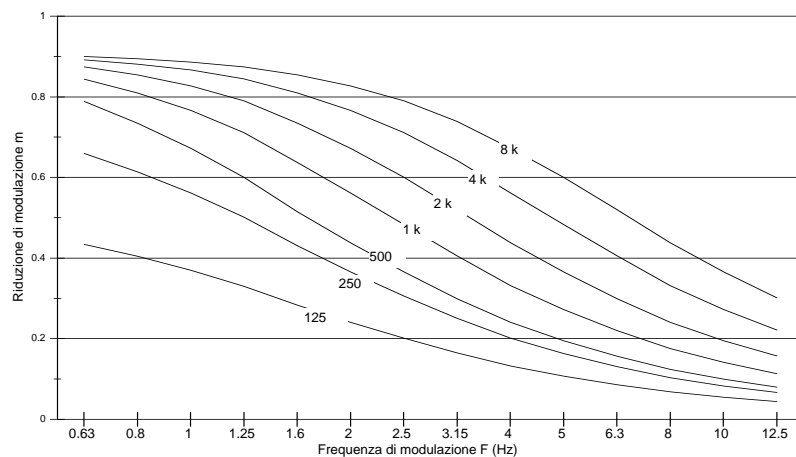
Segnale originale
 $x(t) = A[1 + \cos 2\pi Ft]$



Segnale trasmesso
 $x_t(t) = A[1 + \cos 2\pi Ft + \varphi]$



Banda d'ottava (Hz)	T (s)	S/N (dB)
125	2.0	0
250	2.0	+5
500	2.0	+10
1 k	1.4	+10
2 k	1.0	+10
4 k	0.7	+10
8 k	0.5	+10



Esempio di come si può realizzare un'analisi *MTF* usando un rumore bianco filtrato per ottave come portante, modulato al 100% con ciascuna frequenza modulante. Ciò porta ad una famiglia di curve *MTF*. Nell'esempio, ciascuna curva è derivata da differenti valori di T e di S/N .

L'analisi attraverso $m(F)$ va compiuta per bande d'ottava, da 125 Hz fino a 8 kHz, a frequenze di modulazione precedenti per terzi d'ottava, a partire da 0.63 fino a 16 Hz, per un totale quindi di 98 dati.

Ogni valore di $m(F)$ rappresenta un punto della matrice e, al termine del calcolo, si possono rappresentare delle curve per ogni banda d'ottava.

Al fine di ridurre la matrice così ottenuta (7x14) ad un solo valore, Houtgast e Steeneken propongono il calcolo di un "rapporto S/N apparente", così definito:

$$(S/N)_{app} = 10 \log \frac{m(F)}{1 - m(F)}$$

Esistono 98 (7x14) valori di $(S/N)_{app}$, essi vengono normalizzati entro l'intervallo ± 15 dB (se eccedono si assegna il valore limite) per ciascuna banda di frequenza portante; per cui si può definire lo *STI* (Speed Transmission Index) nella maniera seguente:

$$STI = \frac{\left(\frac{1}{14} \sum_{i=1}^{14} (S/N)_i \right) + 15}{30}$$

Si ottengono dunque sette valori di $STI(f)$, la cui media ponderata dà lo *STI* totale:

$$\sum_{f=1}^7 STI(f) \cdot w(f)$$

IACC,
inter aural cross correlation

Definita la funzione di correlazione mutua interaurale *K*

$$K(\tau) = \frac{\int_{-\infty}^{+\infty} p_L(t) p_R(t + \tau) dt}{\sqrt{\int_{-\infty}^{+\infty} p_L^2(t) dt} \sqrt{\int_0^{t_0=\infty} p_R^2(t) dt}}$$

dove:

$p_L(t)$ pressione sonora all'orecchio sinistro
 $p_R(t)$ pressione sonora all'orecchio destro
 τ ritardo corrente.

Il parametro IACC è il il valore massimo di $K(\tau)$ assunto nell'intervallo $[-1,+1]$ ms; teoricamente i limiti di integrazione sono infiniti, ma nella realtà si calcolano le risposte all'impulso per i primi 80 ms.

La *IACC* esprime il grado di somiglianza fra le risposte all'impulso dell'orecchio destro e di quello sinistro. Qualora la testa sia orientata correttamente verso la sorgente e se l'energia sonora raggiunge le orecchie prevalentemente nello stesso istante, $K'(\tau)$ finisce per assumere il valore massimo per $\tau = 0$.

Qualora la maggior energia provenisse lateralmente (anche per un non corretto orientamento della "testa"), il valore massimo si ha per un $\tau \neq 0$, comunque mai superiore in valore assoluto a 1 ms, dal momento che questo è il massimo tempo impiegato da un suono perfettamente laterale, per aggirare la testa.

I valori della IACC che assicurano una buona sensazione di immersione nel suono, si mantengono attorno allo 0.4-0.5; valori superiori (comunque inferiori o uguali all'unità) comportano un ascolto privo di spazialità, mentre i valori inferiori (comunque maggiori o uguali allo 0), testimoniano un'eccessiva direttività laterale del messaggio sonoro, ugualmente poco gradevole.