



Audio Engineering Society Italian Section

Seminario Top Audio 2001 “L’Acustica dei piccoli ambienti di ascolto”

Questo documento è stato stampato in base a quanto fatto pervenire dall’autore; l’AES Italiana non ha apportato alcuna modifica al testo pertanto non si assume responsabilità in merito ai contenuti.

E’ vietata la riproduzione di questo documento, intero o anche in parte, senza l’esplicito consenso dell’autore e della sezione italiana.

Metodologie di misura e valutazione

Prof. Angelo Farina

Università degli Studi di Parma

Parma (PR)

e-mail: angelo.farina@unipr.it

ABSTRACT

0. INTRODUZIONE

1. PROPAGAZIONE DEL SUONO IN AMBIENTI CHIUSI.

Supponiamo di avere una sorgente omnidirezionale che emette fronti d’onda sferici con uguale intensità in ogni direzione e supponiamo inoltre di collocarla in una stanza di forma perfettamente parallelepipedica (in modo da semplificarci i calcoli). All’interno di questa stanza disponiamo la sorgente S e il ricevitore R come in figura 1.

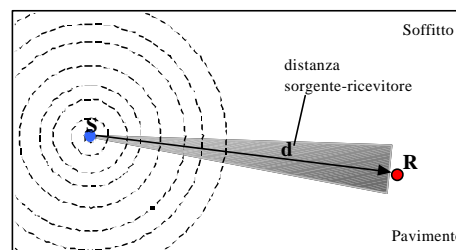


Fig 1: Propagazione sonora in una stanza

All’istante $t=0$ la sorgente emette un impulso che si propaga con dei fronti d’onda sferici sempre più grandi. Trascorso un tempo:

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

($c=343$ m/s è la velocità del suono nell'aria) necessario per la propagazione, il suono arriva al ricevitore R che per un attimo rivela un livello sonoro piuttosto elevato e poi ricade rapidamente a 0 a causa della breve durata dell'impulso. Questo livello chiamato L_1 è il livello del suono diretto e rappresenta quel suono che ha viaggiato senza subire l'influenza dell'ambiente esterno, non ha, di fatto, subito fenomeni di riflessione e/o di assorbimento, in quanto ha percorso la minima distanza possibile tra sorgente e ricevitore.

Nella sua corsa il raggio sonoro interagisce anche con le pareti e ad esempio si riflette contro il pavimento. Questo suono ha percorso una distanza maggiore del suono diretto in quanto è come se provenisse da una sorgente immagine S^1 , lo stesso vale per il suono che si riflette sul soffitto che compie una distanza ancora maggiore ed è come se provenisse da una sorgente immagine S^2 (vedi figura 2).

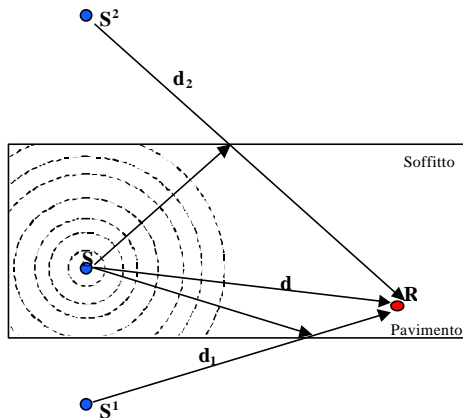


Fig 2: Sorgenti immagine e riflessioni.

Questi due raggi sonori riflessi dal pavimento e dal soffitto arriveranno con dei ritardi maggiori, se il suono diretto impiegava un tempo t_1 per arrivare alla sorgente, il suono riflesso sul pavimento arriverà con un ritardo:

$$t_2 = \frac{d_2}{c}$$

maggiore di t_1 , perché è maggiore la distanza d_2 della sorgente immagine. Analogamente arriverà la riflessione del soffitto con un ulteriore ritardo poiché il percorso è stato maggiore.

Risulta quindi evidente che i livelli sonori sono via via più bassi a causa della legge della divergenza sferica per cui il livello sonoro cala di 6dB per ogni raddoppio di distanza, quindi maggiore è la distanza e più basso sarà il livello sonoro.

Inoltre se ad esempio il pavimento ha un coefficiente di assorbimento α_1 e il soffitto ha un

coefficiente α_2 , è chiaro che il suono riflesso risultante sarà dato dal suono che sarebbe stato riflesso da una superficie perfettamente speculare moltiplicato per il coefficiente $(1-\alpha_1)$ e lo stesso per il suono riflesso dal soffitto.

Se ad esempio il soffitto è costituito da un materiale molto fonoassorbente con un coefficiente di assorbimento $\alpha_2=0,8$ allora $(1-\alpha_2)=0,2$ il che significa che solo il 20% dell'energia sonora incidente sul soffitto viene riflessa ed è anche per questo motivo che i raggi riflessi arrivano al ricevitore con una forte attenuazione.

In questo semplice caso si sono presi in esame solamente i raggi riflessi dal pavimento e dal soffitto, ma nella realtà bisognerebbe prendere in esame anche tutte le altre riflessioni prodotte dalla parete posteriore, dalle pareti laterali e via dicendo. I tempi di arrivo di questi raggi tendono ad uniformarsi e l'energia da essi trasportata a calare. Nulla vieta però ai raggi di continuare a riflettersi sulle altre pareti creando raggi riflessi di ordine multiplo che si ottengono geometricamente continuando a specularizzare le sorgenti rispetto al loro piano fino a creare una vera e propria rappresentazione geometrica dei percorsi dei vari raggi sonori.

Di sorgenti immagine di ordine successivo ce ne sono una moltitudine, esse rappresentano cammini di raggi sonori dal percorso molto complicato che hanno rimbalzato un po' su tutte le pareti accumulando ritardi e perdite di energia via via crescenti. I livelli sonori continuano a decrescere finché ad un certo punto si ha una sorta di **coda sonora continua** dove non è più possibile separare un'onda riflessa da un'altra in quanto ne arrivano una moltitudine negli stessi istanti di tempo provenienti da direzioni diverse ma con lo stesso centro d'arrivo (vedi figura 3).

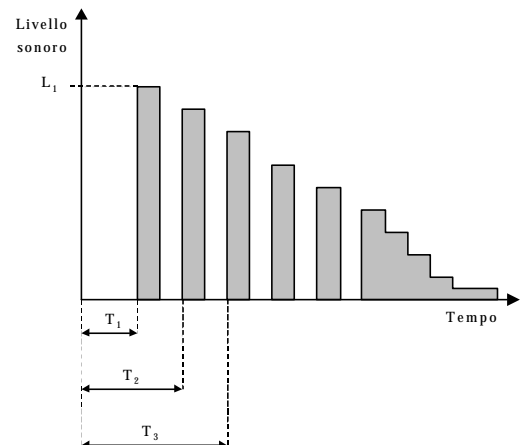


Figura 3 - Livelli sonori di una sorgente impulsiva, da notare con il passare del tempo il comparire di una **coda sonora continua**.

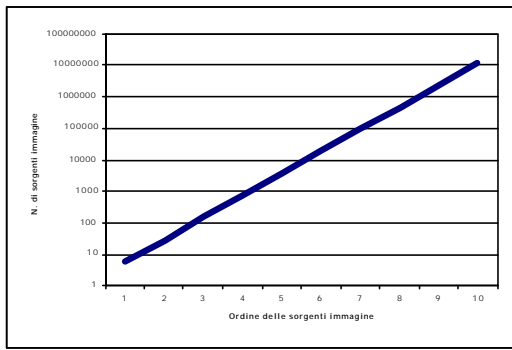


Fig. 4 - Crescita esponenziale del numero delle sorgenti immagine.

Dato un ambiente parallelepipedo con 6 facce il numero di sorgenti del primo ordine è 6, il numero di sorgenti del secondo ordine si ottiene specularizzando ciascuna sorgente del primo ordine rispetto alle altre 5 possibili facce su cui si può riflettere e cioè $5 \times 6 = 30$, il processo continua in questo modo ed il numero di sorgenti immagine di ordine elevato tende ad essere enorme. Il numero di sorgenti immagine esplose in maniera geometrica con l'ordine della riflessione. Quindi se di suoni riflessi una volta ne arrivano 6, di suoni riflessi due volte ne arrivano 30, di suoni riflessi 10 volte ne arrivano milioni e questi milioni arrivano con intervalli di tempo casuali creando una coda sonora statistica il cui andamento energetico è decrescente col tempo ma la cui energia decade in maniera proporzionale alla corsa del raggio ed al numero di rimbalzi (ad ogni rimbalzo perde una quota di energia pari a $1-\alpha$).

Il coefficiente di assorbimento α può variare da rimbalzo a rimbalzo, ma su di un gran numero di impatti è plausibile usare il suo valore medio, $\bar{\alpha}$. Non tutti i raggi rimbalzeranno nello stesso ordine, ne risulta che la coda sonora non sarà perfettamente liscia ma presenterà delle imperfezioni locali, mediamente però l'andamento della curva sarà liscio e decrescente.

Questo è il fenomeno noto come **coda sonora o riverberazione** che con una sorgente impulsiva è molto chiaro e ben individuabile. Le sorgenti impulsive però non sono in uso da molto, tradizionalmente il campo sonoro veniva studiato mediante sorgenti stazionarie che venivano bruscamente interrotte.

Per esempio se si smette bruscamente di suonare l'organo di una chiesa, si continua a sentirsi una coda che diminuisce sempre più di intensità. Questa situazione non corrisponde esattamente a quella precedentemente analizzata, in quanto la situazione dell'organo rappresenta una risposta in regime stazionario interrotto, che è diversa dal regime di transiente analizzato all'inizio.

Ci si può però ricondurre matematicamente alla situazione iniziale prendendo in esame la situazione di accensione di una sorgente stazionaria e non il suo spegnimento improvviso. Costruiremo quindi un diagramma con i valori dei livelli sonori che si sviluppano per l'effetto di una sorgente non più impulsiva ma stazionaria che viene accesa al tempo 0.

Trascorso il tempo t_1 arriva un fronte d'onda e questo assume il valore L_1 che assumeva anche con la sorgente impulsiva ammesso che la potenza delle due sorgenti sia la stessa. Siccome la sorgente non cessa dopo 1 ms come succedeva con la sorgente impulsiva ma continua ad erogare energia il livello rimane costante una volta che è arrivato il suono diretto e rimane costante fino al tempo t_2 al quale arriva anche il suono prodotto dalla prima riflessione che va a sommarsi al suono dell'onda diretta.

Fatta l'ipotesi che la sorgente sia rumore incoerente, i due segnali si sommano energeticamente senza fenomeni di interferenza e quindi il livello sonoro che si raggiunge è la somma di L_1 e di L_2 , dove L_1 è il livello sonoro dell'onda diretta e L_2 è il suono della prima onda riflessa.

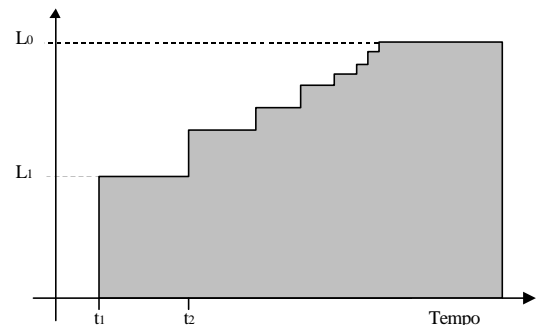


Figura 5 – Livelli sonori di una sorgente stazionaria

A questo punto il livello sonoro rimane costante fino a quando non arriva la terza onda riflessa che fa ulteriormente salire di un gradino il livello sonoro, anche se col passare del tempo e con l'arrivo di nuovi suoni i gradini diventano sempre meno alti in quanto è esperienza nota che se sommo ad un suono elevato dei suoni sempre più deboli il contributo sonoro finale tende a diventare trascurabile. Quindi con una serie di piccoli gradini corrispondenti ai singoli arrivi di energia delle singole riflessioni si giunge asintoticamente ad un livello costante L_0 che rappresenta la condizione di equilibrio energetica dell'ambiente.

Volendo paragonare questa situazione con l'idraulica si potrebbe pensare ad un rubinetto che scarica acqua in un recipiente. A regime il livello che si stabilisce dentro al recipiente è tale per cui la portata di acqua che entra nel secchio è uguale alla portata di acqua che esce dal foro.

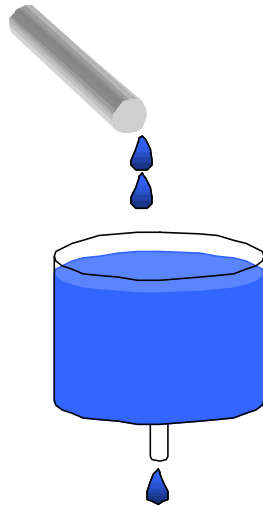


Figura 6 – Situazione di regime in idraulica

L'energia sonora in un ambiente si comporta allo stesso modo, si stabilizza ad un livello costante che corrisponde al bilancio tra la potenza che continuamente la sorgente emette e quella che continuamente le riflessioni sulle pareti stanno togliendo all'ambiente.

E' possibile ricavare matematicamente la situazione di regime ragionando su base energetica. La grandezza che mi permette di valutare un bilancio tra l'energia immessa e quella assorbita è la **densità di energia sonora** e la soluzione a cui si giunge è data dalla seguente equazione:

$$D_{regime} = \frac{4W}{\bar{a} \cdot S \cdot C}$$

Attraverso questa relazione diventa facile il calcolo del livello sonoro dentro l'ambiente che teoricamente è uguale dappertutto e passando da questa relazione alla corrispondente che ci fornisce il livello di intensità, troviamo la **formula del campo riverberante**, che ci fornisce il livello sonoro prodotto da una sorgente di rumore stazionario in un campo perfettamente diffuso e riverberante.

$$L_D = L_W + 10 \cdot \lg \frac{4}{\bar{a} \cdot S}$$

Questo particolare caso teorico si discosta molto dai casi reali in quanto sono molto rare le sorgenti di rumore stazionario ed anche perché è esperienza comune che il livello sonoro non sia sempre lo stesso in tutta la stanza ma vari a seconda delle distanza dalla sorgente di emissione, il concetto è che comunque esiste un suono diretto

che cala di 6dB ogni raddoppio della distanza e che sono le componenti riflesse ad essere mediamente sempre le stesse in ogni punto.

Da quanto appena puntualizzato si può comprendere come in generale la formula del campo riverberante sia falsa.

Ne esiste però una versione modificata che considera la sovrapposizione energetica del livello dell'onda sferica del campo libero con il livello del campo diffuso riverberante, svolgendo i calcoli si ottiene un livello d'intensità totale dato dalla **formula del campo semiriverberante**:

$$L_{D.TOT} = L_W + 10 \cdot \lg \left[\frac{4}{\bar{a} \cdot S} + \frac{1}{4 \cdot p \cdot d^2} \right]$$

Questa relazione è da tutti comunemente impiegata per valutare il campo sonoro dentro agli ambienti chiusi.

La figura 7 mostra l'andamento del livello sonoro dentro ad una stanza al variare della distanza dalla sorgente nel caso del livello sonoro della formula semiriverberante. Si tratta di un grafico in scala doppio logaritmica poiché sia la distanza dalla sorgente che il livello sonoro sono grandezze logaritmiche.

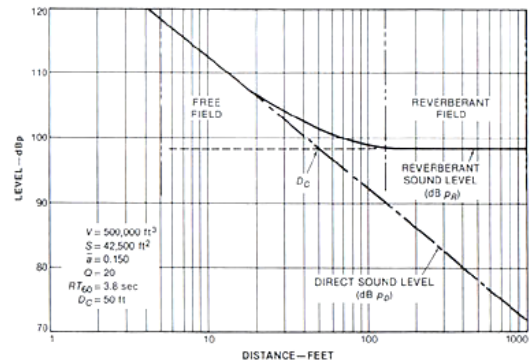


Figura 7 – Andamento del livello del suono diretto e del campo riverberante

Notiamo che la relazione teorica del campo libero corrispondente a 6dB ogni raddoppio della distanza proseguirebbe a decadere inalterata, viceversa in campo riverberante si ha questa sorta di livellamento ad un valore costante che dipende dalla superficie complessiva dell'ambiente e dal suo coefficiente di assorbimento medio, la relazione che li viene riportata è:

$$L_D = l_W + 10 \cdot \lg \left[\frac{4}{\bar{a} \cdot S} + \frac{1}{4 \cdot p \cdot d^2} \right]$$

Questa relazione viene scritta in maniera leggermente diversa da chi si occupa di altoparlanti in quanto l'altoparlante non è in

generale una sorgente omnidirezionale cioè che emette un'intensità uguale in tutte le direzioni. Gli altoparlanti sono di solito sorgenti direttive e si definisce la direttività Q nella direzione θ che sto considerando come il rapporto tra l'intensità emessa in una certa direzione e l'intensità media emessa in tutte le direzioni. Esiste quindi un coefficiente numerico Q_θ che indica 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.

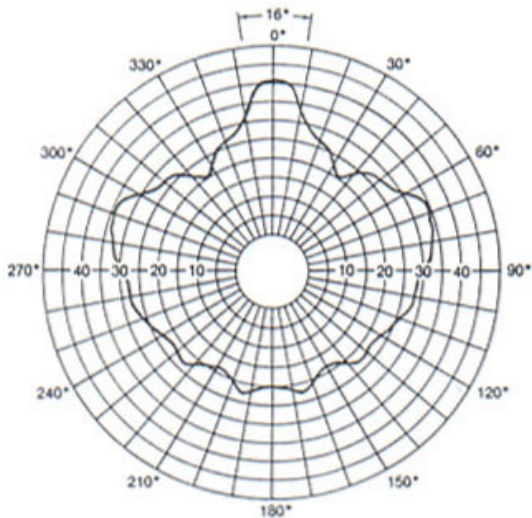


Figura 8 – Curve di direttività di un altoparlante reale

Il fattore di direttività Q può variare da poche unità a valori estremamente elevati. Quindi in presenza di una sorgente non omnidirezionale in campo diffuso mantiene la stessa rappresentazione

$$\frac{4}{\bar{a} \cdot S}$$

La modifica da apportare alla formula del campo riverberante per tenere conto della direttività è data semplicemente dalla sostituzione del fattore 1 che rappresentava l'emissione del campo diretto con il fattore Q che rappresenta ora la direttività. Ovviamente Q può essere maggiore o minore di 1. E' maggiore di 1 se ci troviamo nel lobo di emissione dell'altoparlante, ma se ne siamo dietro o fuori sarà minore 1.

Da notare che la potenza dell'altoparlante è costante quindi se ci sono zone in cui la potenza è maggiore di 1 ci devono essere zone, la maggior parte, in cui il coefficiente sarà minore di 1. A questo punto occorre definire una grandezza fisica importante che è la **distanza critica** che è quella distanza dalla sorgente alla quale il campo sonoro

diretto e il campo sonoro riverberante assumono lo stesso valore.

Analiticamente vale:

$$D_c = \sqrt{\frac{Q_j \cdot S \cdot \bar{a}}{16p}}$$

si scopre che la distanza critica dipende dal grado di assorbimento dell'ambiente, dalla sua superficie complessiva e dalla direttività e direzione dell'altoparlante.

Il concetto di distanza critica è estremamente importante in termini di qualità e comprensione del messaggio che viene riprodotto da un altoparlante. Infatti se ci si trova entro la distanza critica dalla sorgente allora si è in una situazione dove il suono diretto è predominante sul residuo campo diffuso riflesso-riverberante (il suono diretto è chiaro e nitido viceversa il suono riverberante è torbido e confuso), quindi per quanto riguarda soprattutto la comprensione della parola è bene che l'ascoltatore si trovi entro la distanza critica.

In condizioni teoriche si potrebbe pensare ad un diffusore così direttivo che indirizzi il suono solo sul pubblico e non ne sprechi niente altrove.

Tale situazione può essere rappresentata come in figura 9:

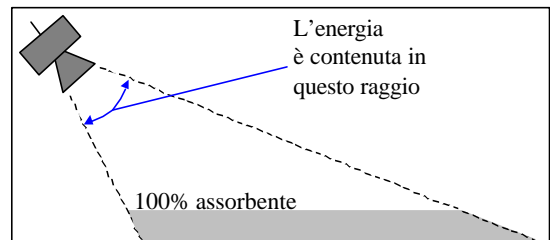


Figura 9 – Altoparlante molto direttivo

L'altoparlante è molto direttivo, il suono arriva sulle persone che hanno un forte potere fonosorbente e non va ad eccitare il campo riverberante della stanza in quanto la stessa non riceve energia né diretta né riflessa. In tale situazione il campo sonoro è tutto diretto anche se il tempo di riverbero dell'ambiente è alto.

Viceversa la situazione rappresentata in figura 10 è un caso molto felice perché è il caso di un altoparlante poco direttivo posizionato in un ambiente senza riverbero, quindi riesce sempre ad emettere un suono diretto pulito poiché il riverbero dell'ambiente è nullo. L'altoparlante direttivo ha efficacia solo se è ben posizionato.

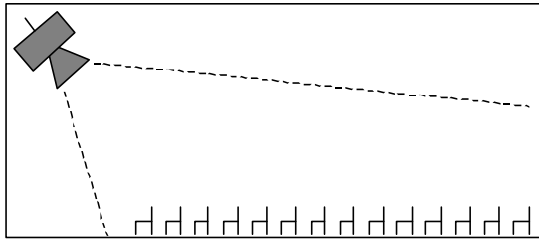


Figura 10 – Altoparlante poco direttivo

La problematica analizzata fino a questo punto è stata esaminata con l'aiuto di un segnale stazionario, ma la situazione in cui l'analisi si complica enormemente è quando il segnale varia nel tempo, quando cioè c'è un forte altalenarsi di suoni deboli e forti a diverse frequenze come ad esempio la voce umana.

In queste situazioni i concetti di livello sonoro (diviso nella sua parte diretta e riverberante) e distanza critica non sono più sufficienti a descrivere il fenomeno nella sua completezza.

Accingiamoci dunque a rifare il processo che ci ha portato alla costruzione del campo sonoro riverberante complessivo. Giunti a regime e quindi a livello complessivo L_0 ad un certo istante t_1 spegniamo la sorgente sonora che emanava il segnale stazionario. L'ascoltatore posto alla distanza d istantaneamente non si accorge che la sorgente è stata spenta e fino al tempo uguale a t_1 il suono continua ad arrivare (non viene più emesso ma quello che era nell'aria continua ad arrivare).

Trascorso il tempo t_1 viene a mancare il contributo del suono diretto e c'è da notare che quando dal silenzio è arrivato il suono diretto la differenza si è sentita molto, mentre adesso tra tutti gli altri suoni il fatto che non ci sia più il suono diretto non è così evidente e questo spiega il fatto per cui il grafico ha in salita dei gradini molto evidenti ed in discesa sia così attenuato.

Trascorre poi il tempo tra t_1 e t_2 con il livello che rimane costante, dopo t_2 se ne va la prima onda riflessa che fa un gradino ancora più minuscolo, dopo di che se ne va la seconda onda riflessa e via così se ne vanno tutte le altre dando origine ad una curva di decadimento che è molto liscia.

Mentre la risposta all'impulso era molto irregolare, la coda di decadimento del rumore stazionario interrotto è molto liscia e si dimostra matematicamente che questa curva su di un diagramma lineare nel tempo è una retta.

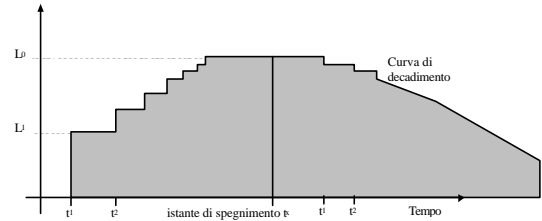


Figura 11 – Curva di decadimento di una sorgente sonora stazionaria

Se la scala verticale è in decibel c'è un decadimento sonoro che se ne va linearmente col tempo (se ne vanno tot decibel ogni secondo). Questo è il fenomeno che fu studiato da Sabine all'inizio del secolo e che porta alla formulazione della durata convenzionale della coda sonora chiamata **anche tempo di riverberazione o di riverbero**.

Il tempo di riverbero è definito come il tempo necessario affinché il livello sonoro decada di 60 dB ed è chiamato T_{60} .

Ovviamente nel fare questa valutazione è bene uscire dalla parte irregolare del grafico e farla nel tratto liscio che segue. Secondo la normativa ISO3382 del 1997 il tratto su cui valutare la pendenza media della curva deve iniziare a -5 dB rispetto al livello L_0 che si aveva con la sorgente in funzione. In teoria dunque bisogna spegnere la sorgente, attendere che il livello decresca di 5dB, far partire un cronometro, aspettare che il livello arrivi a -65 dB ed infine fermare il cronometro.

Il tempo che è passato viene detto tempo di riverberazione T_{60} ed è il parametro primario che caratterizza quanto un ambiente è riverberante o meno.

Nella pratica, invece, è estremamente difficile che si possa realizzare un decadimento "pulito" di oltre 60 dB: si misura dunque il tempo di riverbero estrapolando la pendenza del tratto iniziale della curva di decadimento, considerando il tratto compreso fra -5 e -20 dB rispetto al livello stazionario (T_{20}) oppure fra -5 e -30 dB (T_{30}). Va qui chiarito che, anche se tali tempi di riverbero vengono ottenuti analizzando solo il tratto di curva compresa fra i limiti suddetti, i valori numerici di T_{20} e T_{30} vengono sempre riportati ad un teorico decadimento di 60 dB. Se il decadimento è perfettamente lineare, si ottiene così che $T_{20} = T_{30} = T_{60}$.

Va infine detto che normalmente, onde rendere più robusta la misura, la pendenza del tratto di decadimento analizzato viene misurata non semplicemente dal tempo di ritardo fra i punti all'inizio ed alla fine del campo di livelli considerati, ma effettuando una vera e propria regressione lineare, che così fornisce un risultato dipendente da tutti i campioni compresi nell'intervallo, e non solo dai due estremi dello stesso.

Il coefficiente di assorbimento medio è in qualche modo legato al tempo di riverberazione (mediante la formula di Sabine). Però il coefficiente di assorbimento medio non fornisce direttamente il valore del tempo di riverbero, perché dipende anche dalla superficie dell'ambiente. Il tempo di riverbero indica invece direttamente l'effetto percepibile dall'uomo della durata della coda sonora in secondi.

L'ambiente è molto riverberante quando il tempo di riverbero supera i 2 secondi, viceversa è molto asciutto se il tempo di riverbero è sotto al secondo. Quindi la regolazione del tempo di riverbero in un ambiente è uno dei principali parametri di progettazione acustica ad esempio di una sala concerti. Vari tipi di utilizzo della sala richiedono valori diversi del tempo di riverberazione ottimale.

Utilizzo	T60 ottimo
Aula piccola	0,5s
Aula grande	1,0s
Cinema	0,7s ÷ 0,8 s
Teatro d'epoca	1,3s ÷ 1,5 s
Concert Hall	1,7s ÷ 2,3 s
Chiesa	8,0s ÷ 10 s

Figura 12 – Valori tipici di T60 per vari ambienti.

Quindi esiste un tempo di riverbero ottimo a seconda del tipo di utilizzo della sala.

Dal punto di vista ingegneristico l'unica cosa che occorre è poter quantificare il valore di tempo di riverbero partendo dalle caratteristiche acustiche dell'ambiente e questo lo si fa con la formula di Sabine:

$$T_{60} = 0,16 \cdot \frac{V}{\bar{a} \cdot S}$$

l'unica difficoltà è conoscere il coefficiente di assorbimento. Il coefficiente medio lo si troverà poi come media dei coefficienti di assorbimento delle varie parti della stanza:

$$\bar{a} = \frac{\sum_i a_i \cdot S_i}{S}$$

e quindi posso scrivere

$$T_{60} = 0,16 \cdot \frac{V}{\sum_i a_i \cdot S_i}$$

2. MISURE ACUSTICHE NELLE SALE

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

2.2. Rappresentazione numerica

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) Δt .

Come si è visto il segnale entra nel sistema come un vettore di numeri (il nome dei vettori verrà 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:

X ₁	Y ₁
X ₂	y ₂
X ₃	y ₃
X ₄	y ₄
X ₅	y ₅
X ₆	y ₆
X ₇	y ₇
X ₈	y ₈

2.3. 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:

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

dove m è il numero dell’ultimo elemento di cui si ha memoria (lunghezza della risposta all’impulso, anche chiamata “numero di taps”).

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 cam-

pionamento di 44.100 Hz (la frequenza dei CD audio commerciali), si nota come una memoria di 5 secondi richieda l’utilizzo di 5 * 44100=220500 locazioni di memoria (normalmente a 16 o a 32 bits).

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 10⁵-10⁶ 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.

Teoria di misura

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	0	0	0	0	0	0	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: dovrei essere capace 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 col TRM ottenessimo 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 \mathbf{d} di Dirac. Ammesso di riuscire a trovare \mathbf{x}^{-1} si otterrebbe:

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

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

Analisi in frequenza

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

$$y = x \otimes h \xrightarrow{FFT} Y = X \cdot H$$

Notiamo 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 anti-trasformazione 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 .

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

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 è che 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 ripetizione 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 (la coda della risposta all'impulso non "sta dentro" la finestra di analisi, e la parte che deborda viene risvoltata all'inizio).

Un programma commerciale che è basato sull'uso del rumore è Smaart Pro della JBL.

Segnali particolari: MLS

Al posto del rumore si può però scegliere il segnale in ingresso in maniera più furba, 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):

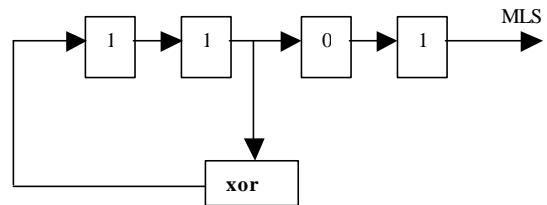


Figura 13 – Esempio di Shift register a 4 bit, i numeri scorrono a destra ad ogni passo.

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 della MLS piatto come quello del rumore bianco posso fare anche 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 poi si può ripetere il segnale e mediare, onde migliorare ulteriormente il rapporto segnale-rumore).

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, sebbene si tratti di un vecchio programma DOS. 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 altoparlanti comporta una perdita di fedeltà e conseguentemente di linearità.

Segnali particolari:sine sweep

Attualmente, comunque, 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

$$S = \text{segnale di sweep}$$

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

Come contropartita di questa comodità però i vantaggi sulla convoluzione. Per riottenere la velocità di calcolo della convoluzione posso passare in frequenza.

2.4. LA PRATICA

Descrizione dell'impianto di misura

Per effettuare le misure è necessario avere: un computer dotato di scheda audio full duplex (ovvero abilitata a registrare da un canale mentre riproduce su di un altro), un software di hard disk recording funzionante in modalità multitraccia (Cool Edit pro oppure Cakewalk Pro Audio o altri), un impianto di riproduzione ed un microfono.Inoltre è utile un editor audio abbastanza completo per la visualizzazione dello spettrogramma e della ETC(energy time curve).

Cablaggio e funzionamento dell'impianto usato per le misure: il software funge sia da riproduttore del segnale di eccitazione sia da registratore del segnale di risposta dell'ambiente; questo fa nascere la necessità di un computer che supporti il full duplex.

L'impianto è molto semplice, infatti alla uscita della scheda audio si collega l'impianto di riproduzione (nel mio caso due speaker da concerto amplificati con un amplificatore esterno) e all'entrata di registrazione della scheda audio si collega il microfono (che se passivo potrebbe necessitare di una preamplificazione).

Collegato l'impianto si procede a fare le misure.

Prime misure e misure con MLS

Un primo tentativo può essere quello di generare una δ di Dirac mediante l'editor audio.Questo tentativo il cui risultato si può vedere in fig. 1

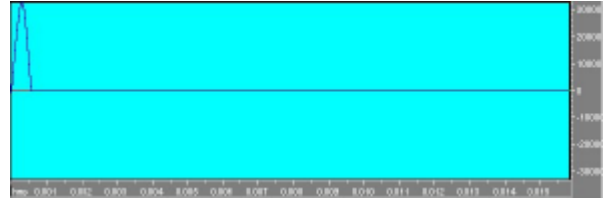


Fig 14 – Il primo tentativo di generare una δ visto con uno zoom sull'attacco.

si rivela però 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.

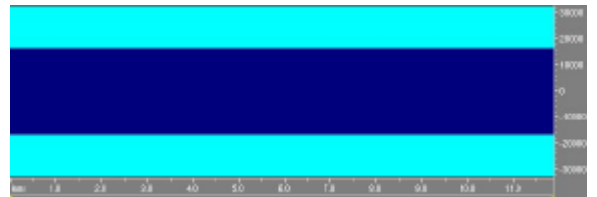


Fig. 15 - Segnale MLS multiplo di ordine 16 ripetuto 8 volte.

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

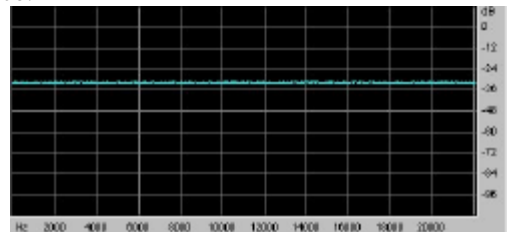


Fig 16 - Spettro della MLS visto con Cool Edit

Ora si può procedere ad una misura: in Cakewalk la prima traccia contiene il segnale MLS mentre la seconda registra simultaneamente la risposta dal microfono.

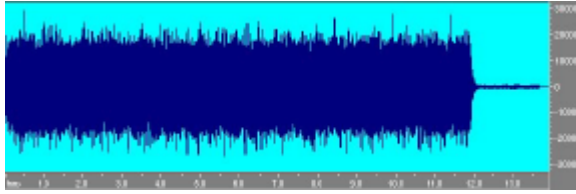


Fig. 17 - Segnale di risposta registrato con il microfono.

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 18, si trova nella clipboard di Windows.

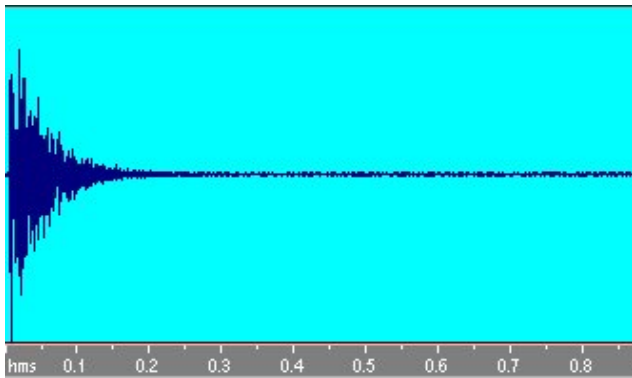


Fig 18 - Risposta della stanza ottenuta con MLS.

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 dà piú energia alle basse frequenze (zona critica) e si procede piú svelti sulle alte in modo da non bruciare i tweeter. In figura 19 si vede il sonogramma (rappresentazione tempo-frequenza) del segnale sweep

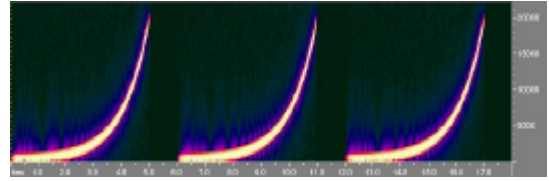


Fig 19 - Sequenza di tre sweep vista con Cool Edit spectral view

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 di fig. 20:

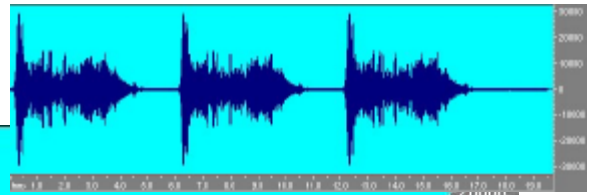


Fig 20 - Risposta alla sequenza di tre sweep.

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

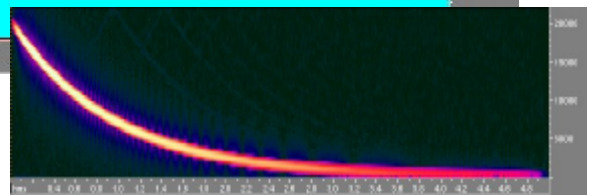


Fig 21 - Sweep inverso visto con Cool Edit spectral view

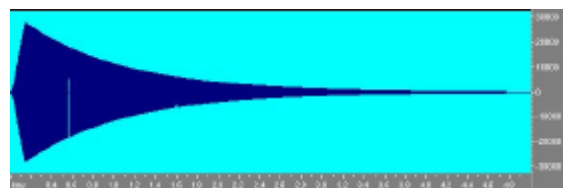


Fig 22 - Sweep inverso visto come ampiezza-tempo.

Notiamo che mentre lo sweep diretto ha ampiezza costante lo sweep inverso é calante in ampiezza in modo da compensare la logaritmicità dello sweep che dava piú energia alle basse frequenze. Infatti lo sweep logaritmico ha energia crescente di 3 db per ottava e quindi é un rumore rosa, se si convolve questo segnale con sé stesso non si ottiene la δ di Dirac ma un segnale che va giú di 6 db per ottava. Questo viene compensato dal guadagno variabile dell'inverso. Convolvere 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.

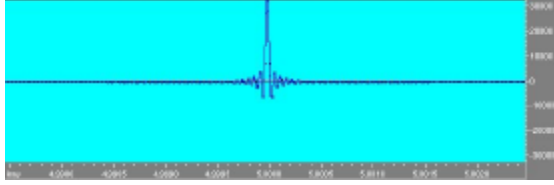


Fig 23 - δ ottenuta dallo sweep.

Convolgendo invece il segnale di risposta della stanza con lo sweep inverso otteniamo la risposta della stanza. Notiamo intanto che avendo tre sweep si hanno anche tre risposte che è possibile mediare fra loro ed inoltre come queste tre risposte siano assai simili a quella ottenuta con la tecnica MLS, ma molto più pulite. In fig. 24 la risposta dell'ambiente ottenuta con lo sweep.

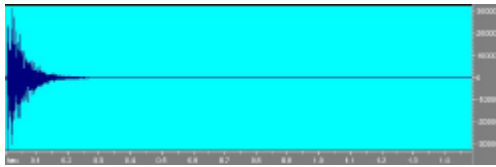


Fig 24 Risposta della stanza ottenuta da sweep e mediata.

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. In fig. 12 l'analisi in frequenza del sistema studiato, che comprende però anche il sistema di riproduzione essendo esso non perfettamente lineare.

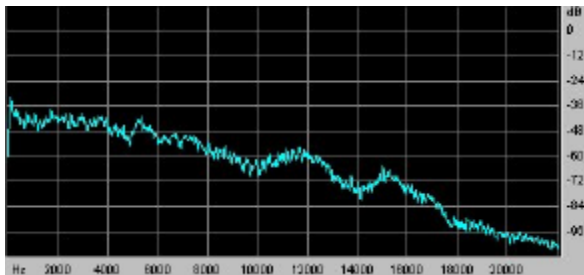


Fig 25 - Risposta della stanza ottenuta da MLS e vista in frequenza.

Notiamo come il sistema, essendo formato da woofer molto grossi, riproduca meglio le basse e medie frequenze e cali sulle alte.

Con Spectra Lab è 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.

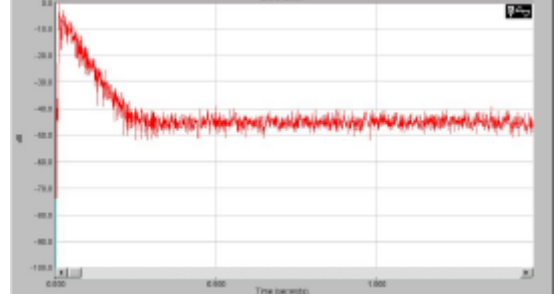


Fig 26 - Risposta della stanza ottenuta da MLS vista con Spectra Lab

Bisogna però contare 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.

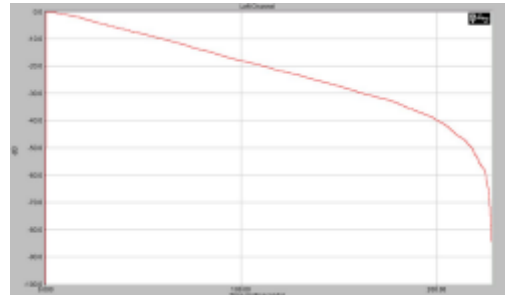


Fig. 27 - Integrale di Schroeder

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 è congruente con i tempi tipici di stanze di piccole dimensioni.

Per fare lo stesso studio in relazione alle singole frequenze dovrei (ad esempio vedere il tempo di decadimento ad 1 KHz dovrei filtrare il segnale di risposta con un filtro passa banda) e usare lo stesso metodo di prima. Siccome 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.

[Vedi Tabella in fondo al documento]

Tale calcolo è 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 posso esportare i dati in formato testo, inoltre può fare le rappresentazioni di tipo waterfall. Questo è utile perché il nostro orecchio percepisce decine di spettri diversi durante l'evolversi della nostra risposta. La possibilità di fare tali grafici è comunque compresa anche su Spectra Lab (Fig. 15). Un altro modo di rappresentare il decadimento è la curva ETC (energy time curve).

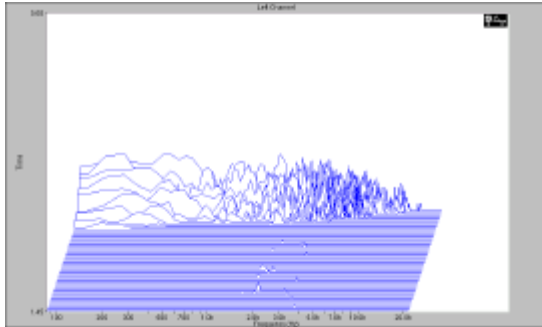


Fig.28 - Risposta della stanza in 3d ottenuta da Spectra Lab

Definizione di alcuni parametri acustici

Infine esistono alcuni altri parametri utilizzati come descrittori di qualità del campo sonoro, principalmente in riferimento all'attività musicale. La base concettuale di questi parametri è la misurazione della parte di suono riverberato che arriva in concomitanza con una parte dell'onda diretta integrandovisi (considerato "utile") a confronto con la parte che arriva dopo (una parte più propriamente di riverberazione che però tende a

sporcare il suono). Questo ha portato alla definizione dell'indice di chiarezza, esistente in due versioni ovvero C50 e C80 (in relazione all'ampiezza della finestra temporale che si assegna al suono utile). In particolare 50 e 80 sono i tempi espressi in millisecondi del massimo ritardo del suono utile.

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.

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. 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} che non aggiunge particolari informazioni allo stesso. La scelta di uno dei due dipende solo dall'abitudine del tecnico. Infine 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.

Esistono altri parametri molto più avanzati che però non vengono trattati in questo corso.

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 dà P_0 la seconda invece ci dà un segnale proporzionale alla velocità particellare che viene erroneamente chiamata P_8 . Esiste una categoria di parametri basati su queste misure:

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

Questo parametro si chiama lateral fraction ed è una percentuale. Un altro parametro simile è l'efficienza laterale

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

Questi parametri sono definiti dalla norma ISO 3382/97.

Un ultimo parametro derivante dal Giappone è lo IACC (inter aural cross correlation), definito come il valore massimo della funzione di cross-correlazione misurata fra i due segnali captati dalle orecchie di una testa artificiale (microfono stereo binaurale, ad esempio Neumann KU-100 o Bruel & Kjaer tipo 4100).

Numerosi dei parametri acustici suddetti possono essere facilmente calcolati in modo automatico grazie ad un apposito modulo di Aurora. La fig. 29 mostra un esempio di calcolo dei parametri acustici a partire da una risposta all'impulso misurata con la tecnica del segnale sinusoidale sweepato.

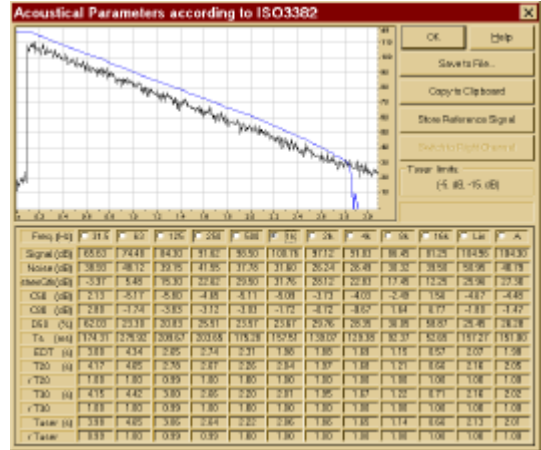


Fig. 29 – Calcolo dei parametri acustici.

Auralizzazione

Grazie a questa tecnica si può ricostruire il suono di un ambiente coinvolgendo la sua risposta con una registrazione anecoica (priva di qualsiasi effetto dovuto alla riflessione del suono). Proviamo quindi ad auralizzare un segnale anecoico con la risposta della stanza studiata.

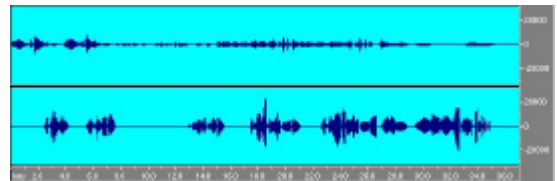


Fig. 30 - Segnale anecoico

E con il comando convolve with clipboard lo si auralizza.

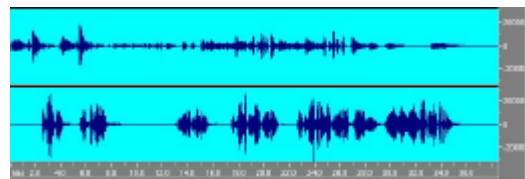


Fig. 31 - Il segnale di fig. 30 auralizzato con le caratteristiche della stanza misurate precedentemente.