

UNIVERSITÀ DEGLI STUDI DI TRIESTE
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

TESI DI LAUREA IN FISICA
INDIRIZZO FISICA DELLA MATERIA

**Applicazione dell'intensimetria
per la registrazione e
riproduzione del suono: la
“Quadrifonia Acustica”**

LAUREANDO:
Giulio CENGARLE

RELATORE:
Prof. Domenico STANZIAL

CORRELATORE:
Dott. Davide BONSI

ANNO ACCADEMICO 2005/2006

Indice

Prefazione	5
Introduzione	6
Cenni storici sulla registrazione e riproduzione del suono	7
Cenni storici sull'intensimetria	18
1 Fondamenti teorici della Quadrifonia Acustica	21
1.1 Derivazione dell'equazione d'onda	21
1.1.1 Equazioni del moto di un fluido	21
1.1.2 Equazione d'onda dell'acustica lineare	24
1.2 Il campo acustico nello spazio quadri-dimensionale	27
1.2.1 Lo spazio-tempo acustico	28
1.2.2 L'equazione d'onda in forma quadridimensionale	28
1.3 Proprietà energetiche del campo acustico	30
1.3.1 Caso monocromatico	30
1.3.2 Caso generale	32
1.4 La Quadrifonia Acustica	33
1.4.1 Dati Quadrifonici	33
1.4.2 Risposta quadrifonica all'impulso	34
1.4.3 Indicatori del campo e metadati	35
1.4.4 Il processo di misura quadrifonica	36
2 Tecniche di misura quadrifoniche acustiche e risultati sperimentali	40
2.1 La sonda Microflow n	40
2.1.1 La procedura di calibrazione della sonda	42
2.1.2 Il modello USP	45
2.2 Il metodo di misura della risposta all'impulso	46
2.2.1 Misura diretta della risposta all'impulso	46
2.2.2 Tecniche di cross-correlazione	47
2.2.3 Tecnica "Sine Sweep"	48
2.3 La catena di acquisizione di dati e metadati	54
2.4 L'analisi energetica di campi acustici a partire dalla risposta all'impulso	59
2.4.1 Il tempo di riverberazione	60
2.4.2 Il metodo dell'integrale di Schroeder	62
2.4.3 Parametri acustici oggettivi	64
2.4.4 Considerazioni sulle incertezze	65

2.5	Misura della risposta quadrifonica in diversi ambienti	66
2.5.1	La Sala Carnelutti	66
2.5.2	Il Dormitorio del Buora	73
2.5.3	L'auletta didattica FSSG	79
2.5.4	La Basilica di San Marco a Venezia	85
3	Le tecniche di spazializzazione attualmente usate	91
3.1	La Wavefield Synthesis	92
3.2	Ambisonics	95
3.2.1	Il B-Format	95
3.2.2	Il microfono SoundField	96
3.2.3	La decodifica dei segnali per la riproduzione	99
3.3	La compatibilità della Quadrifonia Acustica con Wavefield Synthesis e Ambisonics	102
3.3.1	Il confronto tra la sonda Microflown e il microfono SoundField	107
4	La restituzione di registrazioni quadrifoniche con le attuali tecnologie	111
4.1	Il processo di convoluzione quadrifonica	111
4.2	Il problema della transcodifica e della riproduzione	115
4.2.1	Le Matrici di Gerzon	119
4.3	DVD-Audio e DVD-Video Authoring	120
4.4	Il confronto oggettivo tra evento reale e campo risintetizzato	122
4.4.1	Proposta oggettiva per il confronto	122
4.4.2	Audio quality test	123
	Conclusioni e sviluppi futuri	131
	Appendice A - Segnali e sistemi lineari	134
	Appendice B - Le grandezze acustiche	138
	Bibliografia	140
	Ringraziamenti	142

Prefazione

L'occasione per svolgere questo lavoro è nata da una serie di circostanze favorevoli nelle quali l'autore si è imbattuto dopo essersi messo alla ricerca di opportunità in un settore, quello dell'audio e dell'acustica, che è sempre stato in qualche modo al centro dei suoi interessi. La prima di queste circostanze fu il venire a conoscenza di un'importante realtà all'avanguardia nel settore dell'acustica, quale il Laboratorio di Acustica Musicale e Architettonica della Fondazione Scuola di San Giorgio - Consiglio Nazionale delle Ricerche (FSSG-CNR) di Venezia. Tutto questo è successo proprio mentre era in corso d'opera il perfezionamento tecnologico e l'implementazione pratica di alcune applicazioni della Quadrifonia Acustica, sviluppata dallo stesso laboratorio.

La partecipazione del laboratorio al progetto europeo di ricerca *IP-Racine*, per lo sviluppo di nuove tecnologie per il cinema digitale, e al progetto multidisciplinare *C.A.M.E.R.A.*, promosso dall'Università di Cambridge per lo studio dell'acustica delle chiese di Venezia in relazione alla composizione e all'esecuzione di musica sacra nel Rinascimento, ha fornito l'occasione per lo svolgimento di alcune delle misure e dei test che verranno qui presentati.

Play... ▶

Introduzione

In questo lavoro viene presentata una nuova applicazione dell'intensimetria che prende il nome di Quadrifonia Acustica. Questa tecnica, come vedremo, risulta di notevole interesse sotto diversi punti di vista, in quanto abbraccia vari aspetti e costituisce una sorta di denominatore comune tra discipline apparentemente "distanti" tra di loro; essa infatti fonde l'aspetto squisitamente teorico della fisica acustica con quello applicativo dell'acustica architettonica, fino addirittura a trovare un importante utilizzo, quasi un naturale sbocco, nel settore artistico della produzione musicale e cinematografica. Questa tesi vuole guidare alla scoperta di questi legami esplorando passo per passo i settori sopra citati, pur rimanendo costantemente ancorata agli aspetti prettamente sperimentali e alle misurazioni, che ne costituiscono il nucleo centrale.

Per prima cosa verrà fatta una breve panoramica sulla storia della registrazione e riproduzione del suono, nel corso della quale verranno descritte alcune delle tecnologie e procedure standard dell'ingegneria audio; ciò non viene fatto con la pretesa di trattare esaurientemente l'argomento, cosa al di fuori degli scopi di questo lavoro, ma semplicemente per fornire quegli elementi cui si farà riferimento nel seguito e per trattare la Quadrifonia Acustica come una sorta di sviluppo coerente delle attuali tecnologie.

Seguirà un'analoga introduzione allo sviluppo dell'intensimetria, che ha seguito un percorso in un certo senso parallelo e affine a quello dell'ingegneria audio, trovandosi le due discipline a condividere molti aspetti tecnologici.

A questo punto prenderà il via il corpo principale della tesi, che è così strutturata:

Nel capitolo 1 verranno presentati i fondamenti teorici della Quadrifonia Acustica, a partire dalla formulazione dell'acustica lineare e dalla derivazione fisica della ben nota equazione di D'Alembert, per giungere ad un formalismo quadri-dimensionale del campo acustico lineare nel quale si collocano gli aspetti energetici, fino a definire una nuova procedura di misura denominata appunto Quadrifonia Acustica.

Nel secondo capitolo verranno trattati gli aspetti sperimentali della procedura di misurazione, partendo da una descrizione teorica delle tecniche e degli strumenti usati per poi passare ai dettagli tecnici del setup sperimentale; tutto ciò servirà ad introdurre la parte più importante del capitolo, nella quale esporremo i risultati ottenuti applicando la procedura di misura della risposta

quadrifonica all'impulso in quattro ambienti diversi, alcuni dei quali di notevole interesse storico, artistico ed architettonico. Come vedremo, questa tecnica permette di "campionare" le caratteristiche acustiche degli ambienti, rendendole disponibili per svariati utilizzi.

Conclusa la trattazione della misura e registrazione di un evento sonoro, nel capitolo 3 verranno presentate le due tecniche più avanzate di registrazione e spazializzazione del suono sviluppate nel corso degli ultimi decenni da un approccio di tipo scientifico: la Wavefield Synthesis e l'Ambisonics; come si vedrà, la Quadrifonia Acustica ha il pregio di essere compatibile con queste due tecnologie antecedenti.

Nel capitolo 4 affronteremo il problema della ricostruzione di un campo acustico reale e della sintesi di un evento sonoro virtuale, applicando le tecniche quadrifoniche ai sistemi standard di riproduzione attualmente presenti sul mercato. Verrà inoltre qui proposto un metodo per la valutazione oggettiva della fedeltà della ricostruzione di un evento sonoro.

A conclusione del lavoro, citeremo i futuri sviluppi tecnologici e le ulteriori possibilità offerte nei settori nei quali la Quadrifonia Acustica sta cominciando ad affermarsi.

Cenni storici sulla registrazione e riproduzione del suono

L'inizio della storia della registrazione del suono viene fatto risalire all'anno 1877, quando Edison registrò "*Mary had a little lamb*" sul proprio fonografo a cilindro. I primi dispositivi per la registrazione sfruttavano principi di trasduzione acustico-meccanici, nei quali la vibrazione dell'aria veniva usata per mettere in movimento uno stilo in grado di incidere un solco su un cilindro o su un disco in rotazione. L'applicazione del principio inverso consentiva di riprodurre la traccia incisa convertendo il moto dello stesso stilo nella vibrazione dell'aria attraverso un diffusore acustico; su questo principio si basa il funzionamento del grammofono, che ottenne larga diffusione agli inizi del secolo scorso grazie all'entrata in produzione dei dischi in vinile.

La vera rivoluzione nel settore giunse grazie all'applicazione dell'elettromagnetismo, che rese possibile l'invenzione del telegrafo, della radio e degli altoparlanti; da quel momento in poi cominciò lo sviluppo tecnologico, spinto principalmente dal settore delle telecomunicazioni, che portò all'introduzione dell'elettronica nella registrazione audio grazie all'invenzione del più importante strumento di trasduzione acustico-elettrica: il microfono.

Il microfono

I primi a fare la comparsa furono i microfoni a carbone, nei quali della polvere di carbone veniva collocata tra due placche di metallo, una delle quali era esposta alla pressione sonora. Quando il suono raggiungeva una placca, mettendola in movimento, questa trasmetteva una variazione di pressione alla polvere di

carbone, a seguito della quale si verificava una variazione di conducibilità elettrica. Facendo passare una corrente continua tra le due placche, questa veniva quindi modulata in presenza di onde acustiche, dando luogo così ad un segnale che rifletteva l'andamento della pressione sonora. Questi microfoni trovarono largo utilizzo nella telefonia e furono impiegati anche per le trasmissioni radiofoniche fino agli anni venti, quando furono soppiantati da una nuova tecnologia in grado di portare notevoli migliorie nella risposta in frequenza e nel rapporto segnale/rumore.

L'innovazione fu rappresentata dal microfono a condensatore, entrato in produzione verso la fine degli anni venti. In questo microfono, le due armature di un condensatore sono alimentate da una tensione costante. La presenza di un campo acustico fa variare la distanza tra le due armature, producendo una variazione di capacità elettrica che dà luogo ad un flusso di carica in un circuito collegato alle due armature. Questo sistema permise di estendere la risposta in frequenza su tutto lo spettro udibile¹ e di migliorare notevolmente il rapporto segnale/rumore, tanto da portarlo a valori superiori agli $80dB$. I microfoni a condensatore costituiscono tuttora lo stato dell'arte per la registrazione e amplificazione del suono.

Segnaliamo, per completezza, la presenza di altri due tipi di trasduttore microfonico: il microfono dinamico, che sfrutta l'induzione elettromagnetica di una bobina mobile inserita in un campo magnetico, e il microfono a nastro, nel quale il segnale elettrico viene generato in seguito al movimento di un nastro metallico posto tra le espansioni polari di un magnete.

La figura 1 riassume i tipi di trasduttore citati e ne schematizza i principi di funzionamento.

Oltre che per tipologia di trasduttore, i microfoni vengono classificati in base alla loro direzionalità, cioè al modo in cui rispondono ai suoni provenienti dalle diverse direzioni².

Analizziamo il funzionamento di una capsula nella quale entrambe le facce del diaframma siano esposte alla pressione sonora; lo spostamento della membrana (quindi il segnale di uscita, ad esso proporzionale) è proporzionale alla differenza di pressione dell'aria sulle due facce, cioè a $\hat{n} \cdot \vec{\nabla}p$, dove \hat{n} è il versore ortogonale alla superficie della membrana. Detto θ l'angolo tra la direzione di propagazione dell'onda e il versore \hat{n} , il segnale in uscita vale $V_{out} = |\vec{\nabla}p| \cos(\theta)$. Una capsula di questo tipo ha una sensibilità direzionale simmetrica rispetto al piano della membrana. Il valore massimo si ottiene nella direzione perpendicolare alla membrana. Un microfono basato su questo tipo di capsula viene detto *bi-direzionale*, ovvero a *figura di otto* per la forma della funzione di risposta polare. Come vedremo nel capitolo 1, il gradiente di pressione è proporzionale alla velocità del flusso d'aria in condizioni di onda piana. Per questo motivo il microfono a gradiente di pressione viene anche detto, impropriamente, *velocity microphone*.

Se consideriamo invece una capsula nella quale la membrana sia esposta alla

¹La gamma di suoni udibili dal nostro orecchio copre un intervallo di frequenze che si estende da $20Hz$ a $20kHz$.

²Si noti che storicamente è stata privilegiata non tanto la ricerca del microfono ideale per la misura e riproduzione fedele del campo sonoro, quanto la selezione di particolari caratteristiche, ad esempio la soppressione di suoni indesiderati provenienti da certe direzioni.

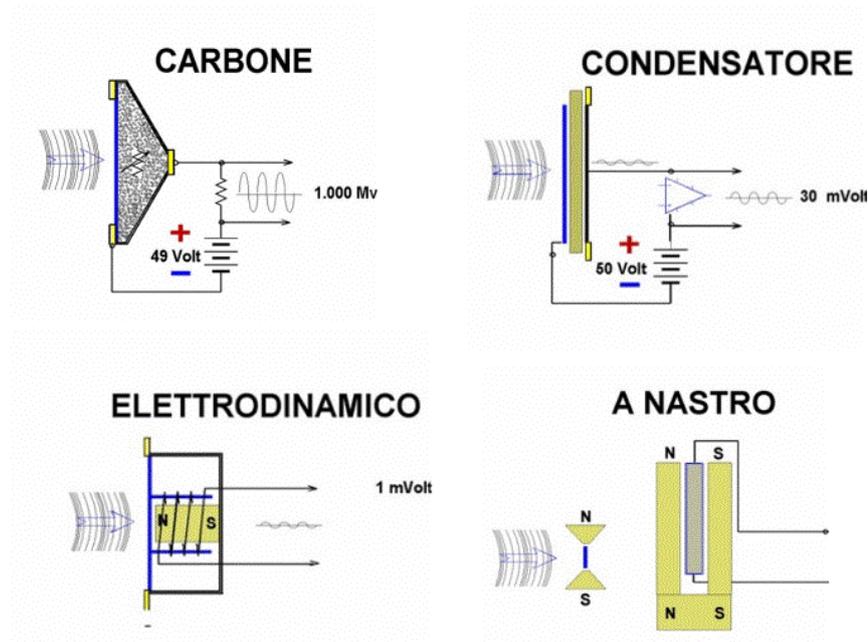


Figura 1: Le principali tipologie e principi di funzionamento delle capsule microfoniche. Tratto da: www.smpte.it

pressione sonora su un lato solo e sia chiusa ermeticamente dall'altro, nel limite in cui le dimensioni della capsula siano inferiori alla lunghezza d'onda del suono incidente, il microfono risponderà al campo d'onda in maniera indipendente dalla sua orientazione rispetto ai fronti d'onda, con un segnale proporzionale alla pressione stessa. Un microfono di questo tipo viene detto *omnidirezionale*.

Osserviamo che una capsula microfonica ha, nella maggiorparte dei casi, una simmetria cilindrica, pertanto tratteremo la direzionalità limitandoci allo studio della sensibilità sul piano dell'orizzonte apparente. Si immagini di porre il microfono in un campo di onde piane di intensità costante e di registrarne la tensione di uscita in funzione della direzione d'arrivo dell'onda. Effettuando il plot della sensibilità su un grafico polare si ottiene il cosiddetto *pattern polare* del microfono. La risposta polare di un microfono viene parametrizzata con l'espressione

$$f(\theta) = A + B \cos(\theta) \quad (1)$$

dove θ è l'angolo di provenienza del suono sul piano orizzontale, con la convenzione che a $\theta = 0$ corrisponda la direzione di massima sensibilità. A e B sono due numeri reali compresi tra 0 e 1, tali che $A + B = 1$. Per $A = 1$, $B = 0$ la sensibilità non dipende dalla direzione di arrivo dell'onda, per cui il microfono è omnidirezionale; per $A = 0$, $B = 1$ si ottiene invece la figura di otto. Si osserva pertanto che il segnale proveniente da un microfono può, in generale, essere espresso come combinazione lineare di un microfono omnidirezionale e di uno bi-direzionale; in altri termini, ogni microfono può essere scomposto in una componente proporzionale alla pressione sonora e in una proporzionale

alla proiezione del gradiente di pressione sul piano polare. La figura 2 mostra i quattro più comuni pattern che possiamo trovare nei microfoni presenti in commercio. Il tipo di microfono più usato è sicuramente quello direzionale, detto anche cardioide per il pattern che richiama la forma del cuore. Tale microfono offre la massima attenuazione dei suoni provenienti da dietro e il suo segnale in uscita può essere considerato la combinazione, in parti uguali, di pressione e gradiente di pressione.

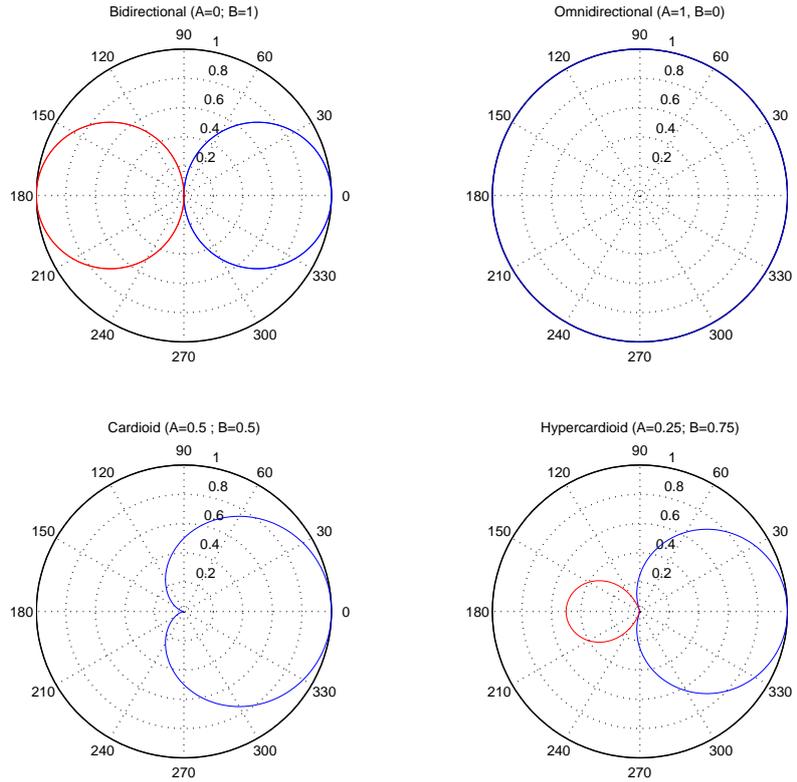


Figura 2: Curve di sensibilità (polar pattern) delle capsule microfoniche più diffuse. In senso antiorario, partendo dalla figura in alto a sinistra: pattern bi-direzionale (figura di otto), pattern direzionale (cardioide), pattern ipercardioide, pattern omnidirezionale. Le curve in rosso indicano che il segnale di uscita ha polarità invertita (è in opposizione di fase) rispetto al segnale acustico.

Una trattazione più esauriente, in cui la capsula microfonica viene trattata come una cavità che presenta la membrana esposta all'aria da un lato e una apertura con una resistenza acustica dall'altro lato viene proposta da Beranek in [1] sulla base di un'analogia elettrodinamica. Beranek trova che, nel caso in cui il campo acustico sia l'onda piana progressiva $p(\vec{r}, t) = \hat{p} \exp[i(\vec{k} \cdot \vec{r} - \omega t)]$, il segnale in uscita dal microfono può essere espresso come

$$V_{out} = A[p + iB(\hat{n} \cdot \vec{\nabla} p)/k] \quad (2)$$

dove p è la pressione sulla membrana, \hat{n} è il versore ortogonale alla membrana, A determina la sensibilità del microfono e B è il coefficiente che ne determina il pattern polare. I coefficienti A e B dipendono dal valore della resistenza acustica posta dietro la membrana, dalla cedevolezza della membrana e dalla sua impedenza acustica.

Il pattern polare di un microfono, in generale, può dipendere dalla frequenza. La figura 3 mostra la sensibilità di un microfono omnidirezionale in funzione della frequenza: man mano che la frequenza sale, la sensibilità laterale e posteriore diminuisce a causa delle dimensioni fisiche della capsula e del corpo del microfono, che costituiscono un ostacolo alla propagazione di onde aventi lunghezza d'onda comparabile con le loro dimensioni geometriche. Il microfono tende a diventare più direzionale alle alte frequenze. Ciò si traduce in una deviazione dalla risposta in frequenza ottimale tanto maggiore quanto più ci si sposta fuori asse; questo ha l'effetto di introdurre "colorazione" nei suoni ripresi fuori asse.

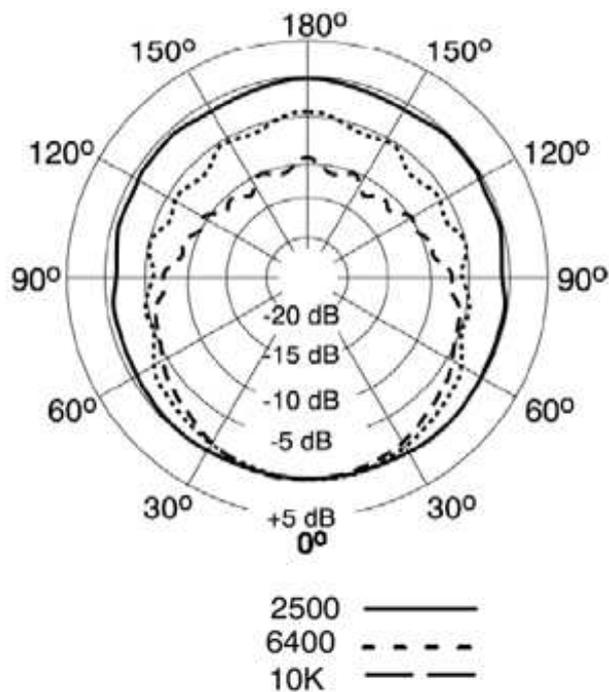


Figura 3: Variazione della risposta polare in funzione della frequenza.

L'effetto prossimità

La misura del gradiente di pressione anziché della pressione dà luogo ad una anomalia nella risposta in frequenza del microfono quando esso viene posto nelle immediate vicinanze di una sorgente sonora. Questa anomalia consiste nella crescente amplificazione delle basse frequenze al diminuire della distanza dalla sorgente ed è nota come *effetto prossimità*. Vogliamo qui brevemente giustificare questo effetto dal punto di vista fisico. Nel caso di un'onda piana progressiva in una dimensione $p(x, t) = \hat{p} \exp[i(kx - \omega t)]$, il gradiente è $\partial p / \partial x = (ik)\hat{p} \exp[i(kx - \omega t)]$ e il suo modulo non dipende dalla posizione

x . La dipendenza dal numero d'onda $k = \omega/c = 2\pi/\lambda$ può essere compensata mediante un'operazione di filtraggio elettronico corrispondente a integrazione temporale e moltiplicazione per $-c$; si ha infatti:

$$-c \int \frac{\partial p}{\partial x} dt = \frac{-c}{-i\omega} (ik) \widehat{p} e^{i(kx-\omega t)} = p e^{i(kx-\omega t)} \quad (3)$$

Questa operazione viene normalmente implementata nei microfoni a gradiente di pressione per ottenere un segnale che abbia lo stesso spettro di quello che si otterrebbe da una misura di pressione, in modo da rispettare le caratteristiche timbriche dell'evento sonoro da riprendere. Se nel caso di un'onda piana progressiva il modulo della pressione e del gradiente di pressione non dipendono dalla posizione, altrettanto non si può dire per un'onda divergente, nella quale l'ampiezza del segnale deve diminuire man mano che ci si allontana dalla sorgente per il principio di conservazione dell'energia. Nel caso unidimensionale, la pressione per un'onda sferica ha la forma

$$p(x, t) = \frac{\widehat{p}}{x} e^{i(kx-\omega t)} \quad (4)$$

cui corrisponde una riduzione di livello di $6dB$ ad ogni raddoppio di distanza a tutte le frequenze. Se calcoliamo il gradiente, si trova

$$\frac{dp(x, t)}{dx} = \frac{ikx - 1}{x^2} \widehat{p} e^{i(kx-\omega t)} \quad (5)$$

Applicando il filtro 3 al segnale 5, otteniamo un segnale

$$V_{out} = \frac{ikx - 1}{ikx^2} \widehat{p} e^{i(kx-\omega t)} \quad (6)$$

Se calcoliamo il rapporto B tra le ampiezze dei segnali prodotti dal microfono a gradiente di pressione filtrato e da un microfono a pressione, collocati nello stesso punto dello spazio, si ha:

$$B = \frac{x \widehat{p} \sqrt{1 + k^2 x^2}}{\widehat{p} kx^2} = \frac{\sqrt{1 + k^2 x^2}}{kx} \quad (7)$$

Il coefficiente B prende il nome di *Bass Boost*. Nel caso in cui la distanza x dalla sorgente sia considerevolmente maggiore della lunghezza d'onda del suono si ha $kx \gg 1$, da cui segue

$$B \simeq \frac{\sqrt{k^2 x^2}}{kx} = 1 \quad (8)$$

pertanto non si ha enfasi delle basse frequenze³. In condizioni di campo vicino, dove la distanza dalla sorgente è piccola rispetto alla lunghezza d'onda, si ha invece

$$B \simeq \frac{\sqrt{1}}{kx} = \frac{1}{kx} \quad (9)$$

Questa relazione prevede, fissata una data distanza x dalla sorgente, un aumento di $6dB$ del segnale prodotto dal microfono al dimezzare della frequenza, a

³In un certo senso, questa approssimazione equivale a considerare l'onda come un'onda piana progressiva.

partire dalla frequenza cui corrisponde una lunghezza d'onda comparabile alla distanza x dalla sorgente. Nelle vicinanze delle sorgenti sonore più comuni, il campo d'onda non è certo assimilabile ad un'onda piana. In realtà tale campo può essere considerato in buona approssimazione sferico, almeno per lunghezze d'onda superiori alle dimensioni della sorgente, condizione nella quale la sorgente viene trattata come "puntiforme". La relazione 9 spiega e quantifica dunque l'effetto prossimità ben noto a tutti gli operatori alle prese con la registrazione ravvicinata di sorgenti sonore con microfoni direzionali. Tale effetto può essere compensato con l'inserimento di un filtro passa alto con pendenza di $6dB/oct$.

I supporti per la registrazione e la riproduzione

Le prime registrazioni, come abbiamo visto, venivano effettuate incidendo direttamente il supporto destinato alla riproduzione, fosse esso un cilindro di gomma o un disco di vinile, con un solco analogo al segnale da riprodurre. Agli inizi degli anni trenta fecero la loro comparsa i primi registratori magnetici, nei quali il segnale veniva registrato magnetizzando un nastro con una testina. Questa tecnica permetteva di migliorare notevolmente il rapporto segnale/rumore rispetto alle incisioni su lacca, tant'è che il sistema di registrazione analogica su nastro magnetico fu la tecnologia dominante fino agli anni novanta. Mentre i primi registratori magnetici, a filo o a nastro, permettevano la registrazione di un solo canale, ben presto si pensò di sfruttare la larghezza del nastro per registrarci più tracce: dapprima due, poi tre, poi quattro, fino a rendere disponibili, nel corso degli anni settanta, registratori in grado di registrare sedici o ventiquattro tracce su un nastro da 2", per venire incontro alle crescenti esigenze di produzione della musica pop.

Gli anni ottanta videro l'affermarsi dei supporti digitali di registrazione su nastro magnetico, tra i quali ricordiamo il multitraccia DASH, il DAT a due canali e l'ADAT a otto tracce. Negli anni novanta comparvero i sistemi di Hard Disk Recording, nei quali il flusso di dati digitali (*streaming*) provenienti dal convertitore A/D viene immagazzinato sul disco rigido di un computer o di un sistema dedicato, permettendo l'accesso non sequenziale (*non linear access*) in fase di playback, che elimina i tempi morti dedicati al riavvolgimento del nastro, e una durata di registrazione continua di ordini di grandezza superiore a quella ottenibile su nastro. L'Hard Disk Recording rappresenta oggi lo stato dell'arte per la registrazione digitale multitraccia, grazie anche alla sempre crescente capacità dei dischi rigidi e al loro costo relativamente basso. Negli ultimi anni, all'Hard Disk Recording si è affiancata la registrazione su memorie Flash, che permette un'ulteriore miniaturizzazione dei dispositivi di registrazione e ne consente l'utilizzo anche in condizioni estreme di shock meccanici, essendo questi apparecchi decisamente meno sensibili alle vibrazioni meccaniche rispetto ai sistemi con disco rigido.

Nel settore consumer, il disco in vinile rimase il supporto più diffuso e fedele per la riproduzione fino alla fine degli anni ottanta. Ai dischi si affiancarono i nastri analogici, ovvero le cassette, di ridotte dimensioni, portatili e quindi adatte anche all'uso in automobile o in viaggio.

Nel 1982 fece la sua comparsa il Compact Disc, il supporto digitale tuttora più venduto. Lo standard *Red Book* definisce le caratteristiche del cd-audio, che contiene audio digitale in formato PCM lineare con una quantizzazione a 16 bit e una frequenza di campionamento di $44.1kHz$. In seguito fu introdotto il

supporto DVD (*digital versatile disc*), con una capacità estesa fino a 9.6 GByte e quindi adatto a contenere video e film digitali, nonché audio multicanale ad una risoluzione superiore a quella del cd. Dei formati DVD si parlerà un po' più dettagliatamente nel capitolo 4.

Citiamo infine l'esistenza del Super Audio CD (*SACD*) destinato a contenere audio ad alta risoluzione con codifica DSD (*direct stream digital*), che rispetto al DVD è compatibile con gli attuali lettori CD grazie alla possibilità di inserire un layer contenente l'audio nel formato previsto dallo standard *Red Book*.

La stereofonia e il surround

Una volta disponibile l'apparecchiatura per poter registrare i suoni con la necessaria risposta in frequenza e un sufficiente rapporto segnale/rumore, si pose all'attenzione il problema di come riuscire a registrare e riprodurre in maniera realistica un evento sonoro (*alta fedeltà*).

Nell'ascolto di un concerto dal vivo, le nostre due orecchie ci permettono di identificare la posizione delle sorgenti nelle tre dimensioni e ci forniscono inoltre la percezione delle caratteristiche acustiche del luogo in cui ci troviamo. Idealmente, una corretta registrazione e riproduzione del suono dovrebbe essere in grado di conferire all'ascoltatore l'illusione di essere presente ad una esecuzione dal vivo. Per poter raggiungere questo obiettivo è necessaria la comprensione dei meccanismi psicoacustici che determinano la localizzazione delle sorgenti.

Il nostro sistema uditivo si basa essenzialmente su differenze di ampiezza e di fase tra i segnali che giungono alle due orecchie per decodificare la direzione di provenienza di un suono. Esperimenti effettuati dimostrano che i due meccanismi dominano in regioni diverse dello spettro: in particolare, Lord Rayleigh dimostrò che la localizzazione alle basse frequenze (indicativamente al di sotto dei 700Hz) è data dalla differenza di fase tra i suoni che giungono alle orecchie, mentre al di sopra dei 700Hz è la differenza di intensità a determinare la direzione di provenienza. Al di sopra dei 5000Hz entra in gioco anche la colorazione del suono, ovvero il risultato del filtraggio in frequenza operato dal padiglione auricolare e dalla testa, che presentano dimensioni comparabili con o maggiori della lunghezza d'onda del suono. Per quanto riguarda la percezione della distanza, essa è legata al livello di pressione sonora e al rapporto tra il suono diretto e le riflessioni generate dall'ambiente.

La stereofonia nacque negli anni trenta come tecnica per la registrazione e riproduzione del suono basata su due canali. In figura 4 è schematizzata la disposizione degli altoparlanti e dell'ascoltatore in un sistema stereofonico. La stereofonia si basa sulla ricostruzione virtuale di un'immagine sonora nell'arco delimitato dai due altoparlanti. Per una ottimale ricostruzione dell'immagine, gli altoparlanti vengono collocati a $+30^\circ$ e -30° rispetto all'ascoltatore (figura 4). È dimostrato che una maggiore separazione degli altoparlanti produce l'effetto "hole in the middle", per il quale l'ascoltatore percepisce due sorgenti sonore in corrispondenza degli altoparlanti e un buco nel mezzo; una minore separazione si traduce invece in una immagine sonora schiacciata.

Alla base della stereofonia stanno i seguenti fenomeni psicoacustici:

- se i due altoparlanti emettono lo stesso segnale, l'ascoltatore percepisce una sorgente puntiforme posta al centro

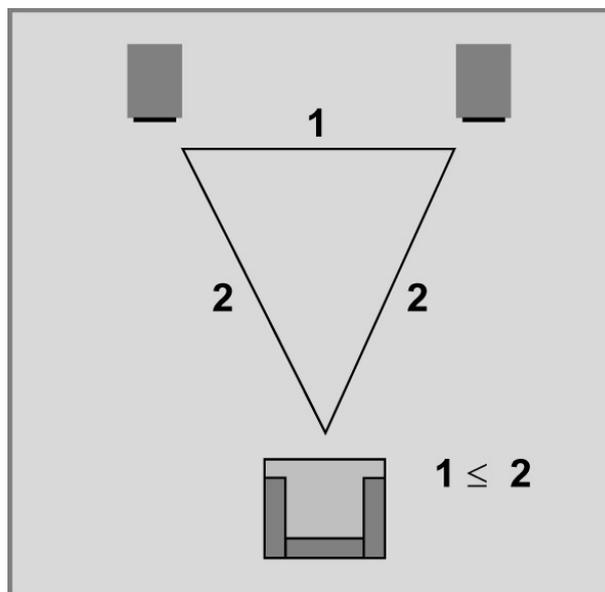


Figura 4: Disposizione degli altoparlanti e dell'ascoltatore in un sistema stereofonico.

- se si aumenta l'intensità di un canale, l'immagine si sposta verso l'altoparlante che emette il suono con maggiore intensità
- se si inserisce una linea di ritardo su un canale, l'immagine si sposta nella direzione dell'altoparlante che emette il suono per primo

La posizione di una sorgente nell'immagine sonora dipende dalla differenza di intensità e dalla differenza di tempo di emissione tra i due canali; la figura 5 mostra la posizione dalla quale sembra provenire una sorgente sonora quando il suo suono viene riprodotto da entrambi gli altoparlanti con una differenza di livello o un ritardo temporale. Notiamo come con due canali non sia possibile estendere la posizione di una sorgente al di fuori dell'arco sotteso dai due altoparlanti⁴.

Le tecniche stereofoniche di ripresa sfruttano la codifica di differenze di livello o di tempo di arrivo tra i canali per creare l'immagine sonora. Nel 1933 Alan Blumlein brevettò un sistema di ripresa basato sull'utilizzo di due microfoni a figura di otto in configurazione coincidente⁵ angolati di 90° l'uno rispetto all'altro. Le direzioni di massima sensibilità vengono quindi a trovarsi a $\pm 45^\circ$ rispetto alla direzione frontale. Ricordando che la curva di sensibilità polare di un microfono a figura di 8 è parametrizzata dall'equazione $f(\theta) = \cos(\theta)$, in presenza di una sorgente sonora collocata entro l'arco di 90° delimitato dalle

⁴Quando i segnali mandati agli altoparlanti hanno polarità invertita, non si riesce più a determinare la direzione di provenienza del suono, che talvolta sembra provenire dai lati o addirittura circondare l'ascoltatore; si potrebbe pensare che sfruttando questo effetto e manipolando la relazione di fase tra i canali sia possibile "estendere" l'immagine stereofonica, tuttavia questo metodo porta più svantaggi che miglioramenti: su tutti, il fatto che l'immagine oltre gli altoparlanti non sia né localizzabile né prevedibile e stabile!

⁵Cioè posizionati il più vicino possibile, in modo che non ci siano differenze di fase tra i due canali.

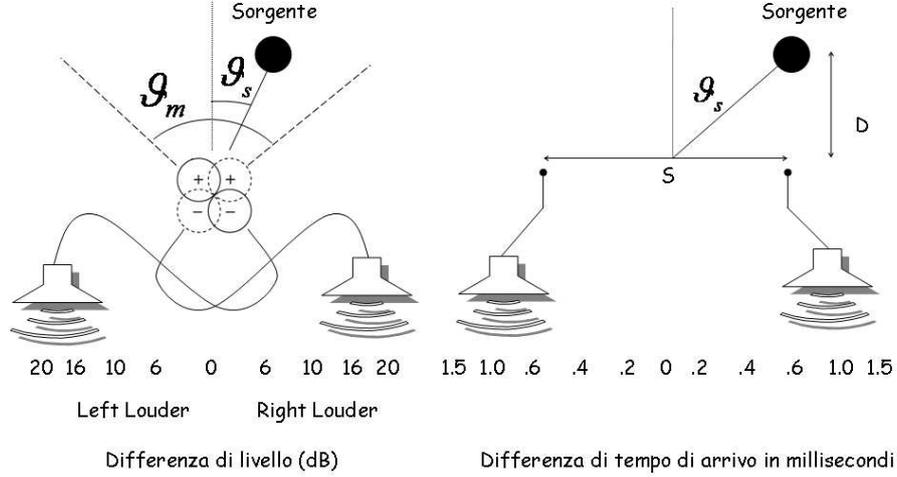


Figura 5: La posizione delle sorgenti sonore nell'immagine stereofonica in funzione della differenza di livello tra i canali e della differenza di tempo di emissione tra i due altoparlanti.

due direzioni di massima sensibilità dei microfoni, ad un angolo θ_s rispetto alla direzione frontale (si veda la figura 5), la differenza di livello tra i due canali è data da

$$\Delta dB = 20 \text{Log} \left[\frac{\cos(45^\circ - \theta_s)}{\cos(45^\circ + \theta_s)} \right] \quad (10)$$

I segnali provenienti dai microfoni orientati a $+45^\circ$ e -45° vengono mandati rispettivamente ai canali destro e sinistro del sistema d'ascolto. Notiamo che una sorgente posta a $\pm 45^\circ$ dà luogo ad una differenza di 20dB e viene riprodotta agli estremi del triangolo d'ascolto, mentre una sorgente posta al centro genera lo stesso segnale su entrambi i canali e verrà quindi riprodotta al centro. Questa tecnica di ripresa, tuttora utilizzata con successo in stereofonia, è universalmente conosciuta come "Blumlein", dal nome del suo inventore.

A partire dalla tecnica Blumlein, vennero proposte e utilizzate altre tecniche basate sull'utilizzo di microfoni in configurazione coincidente, con diversi pattern polari e angoli. Se θ_M è l'angolo tra i due microfoni e A e B i parametri che determinano il pattern, la differenza di livello tra i due canali prodotta da una sorgente in direzione θ_S è data da

$$\Delta dB = 20 \text{Log} \left[\frac{A + B (\cos((\theta_M/2) - \theta_S))}{A + B (\cos((\theta_M/2) + \theta_S))} \right] \quad (11)$$

Le tecniche spaziate utilizzano due microfoni (tipicamente sensori di pressione, quindi omnidirezionali) separati da una certa distanza, in modo che suoni

provenienti da diverse direzioni raggiungano i due microfoni in tempi diversi, permettendo di ricostruire l'immagine sonora in base alle conseguenti differenze di tempo di emissione degli altoparlanti. Questo si basa sull'assunzione, peraltro raramente verificata nell'utilizzo pratico, che il suono giunga ad entrambi i microfoni con la stessa intensità. In queste condizioni, la differenza di tempo d'arrivo in funzione della direzione della sorgente, della sua distanza D dai microfoni e della spaziatura S tra i microfoni (si veda la figura 5) è data da

$$\Delta t = \frac{\sqrt{D^2 + (S/2 + D \tan \theta_S)^2} - \sqrt{D^2 + (S/2 - D \tan \theta_S)^2}}{c} \quad (12)$$

Citiamo inoltre l'esistenza di tecniche di ripresa miste, facenti uso di microfoni direzionali, angolati e separati da distanze generalmente inferiori ai 30cm, in grado di ricostruire l'immagine sonora sfruttando sia differenze di livello che di tempo di arrivo tra i due canali.

La stereofonia richiede supporti e sistemi di diffusione in grado di gestire due canali indipendenti simultaneamente. Dischi e nastri stereofonici vennero commercializzati a partire dagli anni sessanta e la stereofonia si affermò con successo come standard per l'ascolto della musica fino ai giorni nostri.

Il Surround nasce dall'ambizione di circondare l'ascoltatore con il suono. Già negli anni cinquanta, importanti compositori come Pierre Boulez avevano esplorato le possibilità offerte dalla spazializzazione sonora, componendo opere destinate ad essere riprodotte sia da musicisti che da altoparlanti posizionati nella sala da concerto attorno agli ascoltatori, utilizzando la direzionalità e lo spazio sonoro come strumento compositivo e artistico. Le potenzialità dei sistemi surround vennero prevalentemente utilizzate e sviluppate dall'industria cinematografica, come ulteriore strumento per stupire lo spettatore e catturarne l'attenzione. Nel 1940 il film *Fantasia* della Walt Disney porta alla ribalta il suono multicanale [2] dando il via allo sviluppo del surround multicanale nel cinema. Verso la fine degli anni sessanta fecero la loro comparsa i sistemi dedicati alla ricostruzione tridimensionale di un evento sonoro, mossi dal desiderio di riuscire a proporre (e commercializzare) sistemi per un ascolto più realistico, più fedele e più coinvolgente. I due sistemi che ottennero i migliori risultati sono l'Ambisonics e la Wavefield Synthesis, che verranno discussi più approfonditamente nel capitolo 3.

Gli anni settanta videro la comparsa - e, subito dopo, la scomparsa - di una tecnologia nota come quadrifonia, che nulla ha a che vedere con l'argomento di questa tesi. Questa prevedeva la restituzione di un evento sonoro mediante quattro altoparlanti collocati ai vertici di un ideale quadrato, di cui l'ascoltatore doveva occupare il centro. Due furono i sistemi proposti: uno, il sistema discreto, prevedeva la registrazione di quattro canali separati destinati ad essere riprodotti dai quattro altoparlanti; l'altro, il sistema a matrice, pretendeva di derivare i segnali dei quattro canali mediante delle somme e differenze dei due canali di una registrazione stereofonica. Nessuno dei due sistemi ottenne successo in quanto, al di là dello stupore iniziale provocato nell'ascoltatore dalla presenza di suoni provenienti dalle svariate direzioni, si dimostrò incapace di fornire una adeguata ricostruzione dell'immagine sonora non solo a 360° , ma addirittura su ogni singolo lato del quadrato, a causa dell'affetto "hole in the middle" già noto in stereofonia. Inoltre, l'ascolto richiedeva l'installazione di un sistema

di quattro altoparlanti anzichè due, con conseguenti svantaggi economici per l'utenza, ragione principale che ne determinò il fallimento commerciale.

Nel 1977 uscì *Star Wars*, il primo film con colonna sonora Dolby surround che sfruttava i canali posteriori per gli effetti sonori. Negli anni seguenti, il surround si affermò come standard dell'audio per il cinema. In seguito, grazie all'introduzione del DVD, l'audio surround venne esteso anche al settore consumer, con la comparsa dei sistemi *Home Theater*. Il formato attualmente più in uso nell'*Home Theater* è il 5.1 surround, che prevede l'utilizzo di cinque altoparlanti collocati su una circonferenza sul piano orizzontale e un subwoofer per gli effetti a bassa frequenza. Il sistema, illustrato nella figura 6, prevede un altoparlante centrale, due altoparlanti collocati in configurazione stereofonica e due canali posteriori per gli effetti surround. La disposizione degli altoparlanti è dettata dalle esigenze dell'audio per il cinema: il canale centrale fornisce un solido riferimento per il parlato ed i dialoghi, mentre i canali posteriori, collocati nella direzione in cui il nostro sistema uditivo offre la minore capacità di localizzare le sorgenti, nascono per poter restituire effetti sonori vagamente localizzabili. La notevole spaziatura fra i canali posteriori e fra frontali e posteriori limita la capacità del sistema di ricostruire una accurata immagine laterale e posteriore, come vedremo nel capitolo 4. Per questo motivo e per migliorare la restituzione di materiale musicale surround con lo scopo di creare un ascolto più realistico, sono state proposte evoluzioni del 5.1 Surround come il 6.1 o il 7.1 che aggiungono altoparlanti posteriori e laterali per rendere l'immagine sonora più accurata in tutte le direzioni.

Cenni storici sull'intensimetria

Come vedremo nel capitolo 1, le onde sonore sono perturbazioni di pressione e densità del fluido nel quale si propagano (l'aria), accompagnate dal moto degli elementi del fluido stesso. Il moto del fluido comporta una velocità dei suoi elementi, alla quale è associata una certa quantità di energia cinetica. Inoltre, nelle regioni in cui la pressione è superiore al valore di equilibrio c'è immagazzinamento e seguente scambio di energia potenziale. Questa energia viene trasportata dall'onda sonora e si propaga nel campo sonoro.

L'intensimetria nasce per la misura del trasporto di energia nell'aria ad opera dei campi sonori. Come vedremo, l'intensità sonora è un vettore dato dal prodotto della pressione scalare e della velocità vettoriale associate alle particelle acustiche. Misurare l'intensità richiede quindi la capacità di misurare la velocità acustica dell'aria. Se una misura di pressione sonora è un compito relativamente semplice, a partire dall'invenzione del microfono a condensatore, altrettanto non si può dire per la misura di velocità e di intensità, come dimostrato dal fatto che sono trascorsi circa cinquant'anni dal primo brevetto di un dispositivo per la misura del flusso di energia sonora (1932) alla comparsa sul mercato di apparecchiature per le misure intensimetriche. Questo ritardo è dovuto principalmente alle difficoltà tecniche di realizzare un trasduttore accurato e affidabile in grado di convertire la velocità dei flussi d'aria in un segnale elettrico analogico. Lo scoglio fu superato agli inizi degli anni ottanta, grazie anche all'avvento dei calcolatori in grado di effettuare operazioni di filtraggio sui segnali, e da quel momento fiorirono le applicazioni dell'intensimetria, che trovò uno sbocco

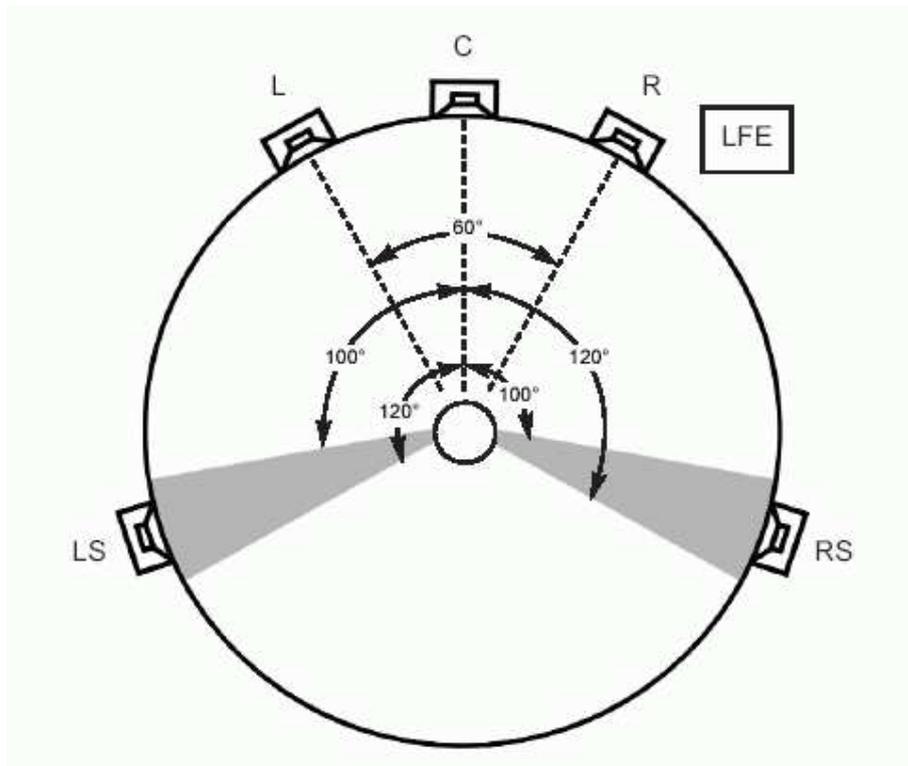


Figura 6: Schema della collocazione degli altoparlanti in un sistema 5.1 Surround.

diretto nella caratterizzazione della potenza sonora delle sorgenti e negli studi sul controllo e la riduzione del rumore.

L'inizio della storia dell'intensimetria si può probabilmente far risalire al 1931, quando Harry Olson presentò la richiesta di brevetto per un dispositivo sensibile al flusso di energia di un'onda sonora. Il sistema prevedeva l'utilizzo di un microfono a pressione e uno a velocità (un microfono a nastro) per ricavare il prodotto pv lungo una sola direzione a partire dall'elevamento al quadrato della somma e differenza dei segnali di pressione e velocità, sfruttando il fatto che $(p + v)^2 - (p - v)^2 = 4pv$. Non risulta però che il dispositivo abbia trovato alcuna applicazione pratica. Furono Enns, Firestone e Clapps, agli inizi degli anni quaranta, ad usare una combinazione di microfono a pressione e microfono a gradiente di pressione per la misura di intensità sonora. Essi compirono studi in un tubo a onde stazionarie e in una camera riverberante, ottenendo buoni risultati solo all'interno del tubo, dove le condizioni di onda piana consentono di identificare la velocità col gradiente di pressione. Al 1955 risale il tentativo di utilizzare il principio dell'anemometria a filamento caldo per la misura diretta della velocità del flusso d'aria, tuttavia il sistema risultò inutilizzabile perché troppo affetto da rumori esterni per poter fornire dati significativi. Tale sistema verrà riscoperto e perfezionato nel 1994 portando alla realizzazione dei sensori Microflown, utilizzati nel corso del presente lavoro.

Nel 1956 fu Schultz a dare un importantissimo contributo all'intensimetria

implementando il principio secondo cui la velocità delle particelle d'aria può essere calcolata integrando il gradiente di pressione ottenuto dalla differenza dei segnali provenienti da due microfoni di pressione in posizione ravvicinata. Questa tecnica fu in seguito perfezionata, e contemporaneamente si accrebbero le conoscenze teoriche sui campi acustici e sull'intensità sonora. Negli anni settanta comparvero sul mercato i primi strumenti analogici per la misura dell'intensità sonora basati su una coppia di sensori di pressione. Negli anni ottanta, grazie allo sviluppo dei DSP e dei calcolatori, si diffusero i sistemi digitali di misura, sempre basati sul doppio sensore di pressione. Grazie a questi strumenti fu possibile effettuare analisi in bande di frequenza, mediante l'applicazione di filtri digitali.

Citiamo la comparsa, alla fine degli anni ottanta, di una sonda per la misura diretta di velocità lungo un asse, basata sulla convezione di un fascio di ultrasuoni da parte del campo acustico. Due fasci paralleli di ultrasuoni vengono lanciati in direzioni opposte; la presenza di moto oscillatorio dell'aria a frequenze udibili provoca una differenza di fase tra i due fasci ultrasonici, che può essere rilevata da due ricevitori. Questa differenza di fase è proporzionale alla velocità dell'onda sonora nella direzione del fascio. Abbinando un microfono di pressione in prossimità del fascio di ultrasuoni è possibile effettuare la misura simultanea e coincidente di p e v . Questa tecnologia non ha riscosso lo stesso successo della tecnica basata su coppie di sensori di pressione.

Come vedremo, solo nell'ultimo decennio ha avuto luogo l'innovazione che ha permesso una misura diretta e accurata di velocità.

Capitolo 1

Fondamenti teorici della Quadrifonia Acustica

1.1 Derivazione dell'equazione d'onda

In questa sezione introduttiva, che segue lo svolgimento trattato in [3], si ripercorre lo schema logico-concettuale che porta all'equazione d'onda di D'Alembert a partire dai principi della fluidodinamica. Altri testi di riferimento sono riportati in bibliografia [4], [5], [6], [7].

1.1.1 Equazioni del moto di un fluido

Per determinare lo stato di moto di un fluido è necessario conoscere in ogni punto dello spazio $\vec{x} = (x_1, x_2, x_3)$ occupato dal fluido e in ogni istante di tempo t non solo la velocità del fluido $\vec{v}(\vec{x}, t)$, ma anche il valore di altre due grandezze fisiche che descrivono le proprietà termodinamiche del fluido stesso: la densità $\rho(\vec{x}, t)$ e la pressione $p(\vec{x}, t)$. Definire queste quantità è possibile perchè i fenomeni trattati dalla fluidodinamica - in particolare dall'acustica, che ne costituisce un importante capitolo - sono di natura macroscopica, e il fluido è perciò trattato come una *sostanza continua*. Di conseguenza, per descrivere matematicamente lo spostamento di una *particella di fluido* che occupa la posizione \vec{x} all'istante t rispetto ad un sistema di riferimento inerziale, si deve ricorrere, come d'uso nella teoria dell'Elasticità, ad un vettore $\vec{s}(\vec{x}, t) = \hat{i}\xi + \hat{j}\eta + \hat{k}\zeta$, dove ξ, η, ζ sono le tre componenti dello spostamento nelle direzioni x, y e z , mentre $\hat{i}, \hat{j}, \hat{k}$ sono i tre versori che definiscono il sistema di riferimento cartesiano.

La massa m di una generica particella a riposo nel punto \vec{x} è data da $m = \rho_0 V$, dove ρ_0 è la densità del fluido in condizioni statiche e $V = \Delta x \Delta y \Delta z$ è l'elemento di volume che delimita la particella di fluido nel punto \vec{x} . Supponendo che tale massa debba rimanere invariata anche durante il moto della particella, è chiaro che uno scostamento della densità del fluido $\rho(\vec{x}, t)$ dal suo valore di equilibrio ρ_0 è necessariamente dovuto ad una variazione ΔV del volume V della particella che si trova nel punto \vec{x} all'istante t . Si può quindi scrivere la relazione tra la dilatazione di volume Θ cui la particella è soggetta durante il

proprio moto e il vettore spostamento \vec{s} tramite l'operatore divergenza:

$$\Theta \equiv \frac{\Delta V}{V} = \vec{\nabla} \cdot \vec{s} \quad (1.1)$$

Questa dilatazione di volume, indotta dalla variazione della densità di massa della particella, è a sua volta legata alla pressione incrementale associata alla stessa particella $p'(\vec{x}, t) = p - p_0$, che nel caso di deformazioni elastiche si esprime tramite la legge di Hooke, che in questo caso assume la forma

$$p' = -B\Theta \quad (1.2)$$

dove B è la costante di compressibilità di volume del fluido, ovvero il *bulk modulus*¹.

Le variabili coinvolte nella descrizione dei fenomeni acustici sono dunque cinque: la pressione, la densità e le tre componenti della velocità vettoriale associate alle particelle di fluido. Queste grandezze obbediscono ad un set di cinque equazioni che esprimono tre principi fondamentali della fisica: la conservazione della massa, la conservazione della quantità di moto e la conservazione dell'energia.

1.1.1.1 Equazione di continuità della massa

La conservazione della massa richiede che la rapidità con cui la massa del fluido aumenta in un volume fissato V_0 sia uguale al flusso netto entrante dovuto alla convezione delle particelle di fluido attraverso le "pareti" del volume V_0 . Perché ciò sia verificato a livello globale (cioè su tutto il volume V_0), la velocità di ogni singola particella di fluido \vec{v} e la sua densità ρ devono soddisfare a livello locale (cioè in ogni punto del volume V_0) l'*equazione di continuità*, che si può scrivere nelle tre forme equivalenti

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (1.3)$$

$$\frac{1}{\rho} \frac{D\rho}{Dt} + \vec{\nabla} \cdot \vec{v} = 0 \quad (1.4)$$

$$\vec{\nabla} \cdot \vec{v} = \rho \frac{D}{Dt} \frac{1}{\rho} \quad (1.5)$$

dove $\vec{j} = \rho\vec{v}$ è la *densità di flusso di massa* e

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla} \quad (1.6)$$

è la cosiddetta *derivata sostanziale*.

Il significato fisico dell'equazione di continuità si può interpretare a livello locale analizzando l'equazione 1.3, considerando che il vettore \vec{j} che ivi compare rappresenta la quantità di massa di fluido che scorre nell'unità di tempo attraverso l'unità di superficie istantaneamente ortogonale a \vec{v} , e che per esso vale il teorema della divergenza. La forma 1.5 dell'equazione di continuità per la massa è invece particolarmente importante per interpretare in termini di volume il

¹Nel caso dei gas, si ha $B = \gamma p_0$, dove γ è l'indice adiabatico, definito come rapporto tra i calori specifici a pressione costante e a volume costante.

significato della conservazione della massa. Infatti, poichè come si deduce dalla 1.1, $\vec{\nabla} \cdot \vec{v}$ esprime la rapidità con cui il volume della particella di fluido che scorre con velocità \vec{v} si dilata per unità di volume seguendo la traiettoria della particella stessa²³, un incremento/decremento della massa del fluido all'interno del volume fissato V_0 deve necessariamente avvenire con la stessa rapidità con cui il volume complessivo di tutte le particelle che si muovono all'interno del volume V_0 , cioè $\Delta V_0/V_0$ aumenta/diminuisce. Ne segue che un fluido è *incompressibile* quando $\vec{\nabla} \cdot \vec{v} = 0$ in ogni punto \vec{x} del fluido ad ogni istante di tempo t .

1.1.1.2 Equazione della quantità di moto

Questa equazione, nota anche col nome di *equazione di Navier-Stokes*, esprime la rapidità sostanziale con cui cambia la quantità di moto di una particella di fluido in termini della sua pressione p , della forza *viscosa* (attrito interno) e della risultante \vec{F} di tutte le altre forze esterne che agiscono sulla particella per unità di volume (*body forces*). I fluidi di interesse acustico, ad esempio l'aria e l'acqua, sono detti *Stokesiani* in quanto le loro forze di attrito interno possono essere espresse in termini del coefficiente di viscosità allo scorrimento η , che nella descrizione del moto viene sempre considerato costante. L'equazione della quantità di moto di questi fluidi si scrive come

$$\rho \frac{D\vec{v}}{Dt} = -\vec{\nabla}p + \eta \left(\nabla^2 \vec{v} + \frac{1}{3} \vec{\nabla}(\vec{\nabla} \cdot \vec{v}) \right) + \vec{F} \quad (1.7)$$

Ponendo $\eta = 0$ e $\vec{F} = 0$ nell'equazione 1.7, cioè assumendo l'assenza di viscosità e di forze esterne, si ottiene un'equazione nota come *equazione di Eulero*.

La tabella seguente riporta i valori sperimentali di riferimento per ρ_0 , η e del rapporto $\nu = \eta/\rho_0$ detto *coefficiente di viscosità cinematica*, per l'aria e per l'acqua a 10°C e 1 atm

	ρ_0 [Kg/m ³]	η [Kg/m s]	ν [m ² /s]
Aria	1.23	1.764×10^{-5}	1.433×10^{-5}
Acqua	1000	1.284×10^{-2}	1.284×10^{-5}

1.1.1.3 Equazione di continuità dell'entropia

Quando nella descrizione del moto di un fluido è possibile trascurare gli effetti di dissipazione energetica dovuti alla viscosità del fluido e di scambio di calore per conduzione termica tra le sue diverse parti, il fluido si dice *ideale*. Il moto di un fluido ideale, descritto dall'equazione di Eulero, è quindi *adiabatico* per definizione, pertanto l'*entropia* di ogni singola particella di fluido rimane costante durante il suo movimento. In particolare, i piccoli moti vibrazionali dell'aria che avvengono quando l'aria stessa si comporta come mezzo di propagazione ideale delle onde sonore, possono essere considerati adiabatici. Ai fini dell'adiabaticità del moto della "particella acustica", l'aria è dunque considerata come un

²La stessa cosa si può esprimere anche dicendo che essa rappresenta il "tasso sostanziale di variazione relativa del volume della particella"

³N.B. Si noti che le traiettorie delle particelle non coincidono sempre con le linee di flusso: ciò avviene soltanto quando il flusso è stazionario.

fluido non viscoso, e le velocità del moto sono piccole rispetto alla velocità di propagazione dell'onda ($|\vec{v}| \ll c$).

Se una particella d'aria si muove con velocità acustica \vec{v} , allora la sua *entropia specifica*⁴ s deve soddisfare l'equazione

$$\frac{Ds}{Dt} = 0 \quad (1.8)$$

L'equazione 1.8 esprime il fatto che la derivata di s calcolata lungo la traiettoria della particella, ovvero la derivata sostanziale 1.6 dell'entropia specifica, deve essere sempre nulla. La condizione di moto adiabatico 1.8 si può anche scrivere come un'equazione di continuità per l'entropia della particella di fluido, nella forma

$$\frac{\partial \rho s}{\partial t} + \vec{\nabla} \cdot (\rho s \vec{v}) = 0 \quad (1.9)$$

dove la quantità $\rho s \vec{v}$ è la densità di flusso di entropia. Si noti che quando $s(\vec{x}, t) = \text{costante}$, l'equazione 1.9 si riduce all'equazione 1.3. In tal caso il moto del fluido è detto *isentropico* e si può assumere che in ogni punto all'interno del volume macroscopico V_0 e in ogni istante di tempo, la relazione che intercorre tra la pressione la densità di ogni singola particella di fluido (in moto oppure ferma) sia data da un'equazione della forma

$$p = p(\rho, s); \quad s = \text{costante} \quad (1.10)$$

Come noto, una trasformazione adiabatica in un gas perfetto è caratterizzata da $pV^\gamma = \text{costante}$; poichè la densità di una particella è inversamente proporzionale al suo volume, questa relazione si scrive anche come

$$p = \text{costante} \times \rho^\gamma \quad (1.11)$$

dove $\gamma = C_p/C_V > 1$ è il rapporto dei calori specifici, o *indice adiabatico*. L'indice adiabatico dell'aria misurato sperimentalmente risulta uguale a 1.403.

1.1.2 Equazione d'onda dell'acustica lineare

In riferimento all'Appendice B, osserviamo che il rapporto tra la pressione sonora *rms* corrispondente ad un livello di 120dB SPL e la pressione di equilibrio vale

$$\frac{(p_{rms})_{120dB\ SPL}}{p_0} \approx \frac{2 \times 10^{-5}}{10^5} \times 10^{(120/20)} = 2 \times 10^{-4} \ll 1 \quad (1.12)$$

Si deduce quindi che anche i suoni più forti che percepiamo sono comunque causati da piccole perturbazioni della pressione atmosferica: precisamente da perturbazioni che sono circa diecimila volte più piccole della pressione di equilibrio p_0 . Con lo stesso procedimento, per un suono assordante di 160dB SPL di intensità si troverebbe $(p_{rms})_{160dB\ SPL}/p_0 \sim 0.02$. In questo caso, tuttavia, le perturbazioni di pressione che generano il suono, essendo soltanto un centesimo di p_0 , non possono essere propriamente considerate "piccole perturbazioni". Di conseguenza, quando si cerca di applicare l'equazione generale 1.7 nella sua forma di Eulero (cioè considerando l'aria come un fluido ideale) per ricavare il

⁴Cioè la quantità di calore per unità di temperatura e di massa dovuta al cambiamento dello stato termodinamico della particella conseguente al suo moto.

moto di queste perturbazioni ondose, il termine non lineare $(\vec{v} \cdot \vec{\nabla}) \vec{v}$ che compare nella derivata sostanziale della velocità della particella comincia a dare un contributo significativo alla descrizione del moto. In pratica, il livello di pressione sonora di 160dB SPL è da considerarsi come livello limite per il regime di linearità dell'equazione di Eulero. L'*Acustica lineare* propriamente detta descrive dunque i moti ondosi che si instaurano nell'aria a causa di perturbazioni della sua pressione di equilibrio comprese tra $(p_{rms})_{0dB\ SPL} = 2 \times 10^{-5} Pa$ e $(p_{rms})_{140dB\ SPL} = 2 \times 10^2 Pa$. A questi valori di pressione corrisponde rispettivamente una *velocità rms della particella acustica* pari a circa $5 \times 10^{-8} m/s$ e $5 \times 10^{-1} m/s$, mentre la soglia di dolore corrisponde ad una velocità *rms* della particella d'aria di circa $5 \times 10^{-2} m/s$, ovvero $5 cm/s$.

Vediamo ora come si ricava l'equazione delle onde (equazione di D'Alembert) dal sistema di cinque equazioni scalari 1.3, 1.7 e 1.8, che definiscono il moto della particella d'aria o *particella acustica* in un generico punto \vec{x} e in tutti gli istanti di tempo t misurati con un orologio posto in quel punto. In ciascuno di questi punti, un'onda sonora causa sempre dei cicli alternati di compressione e rarefazione dell'aria. Quando ciò avviene, si dice che lo spazio misurato con \vec{x} è sede di un *campo acustico*.

Poichè stiamo supponendo che lo spostamento delle particelle acustiche nel loro moto oscillatorio intorno alla posizione di equilibrio sia piccolo, sarà piccola anche la loro velocità \vec{v} . Allora l'ipotesi delle "piccole oscillazioni" in fluidodinamica equivale a trascurare gli effetti sul moto del termine $(\vec{v} \cdot \vec{\nabla}) \vec{v}$ - detto anche *accelerazione convettiva* - che compare nell'equazione di Eulero. Per la stessa ragione, possiamo supporre che i cambiamenti differenziali della densità e della pressione della particella di fluido siano piccoli. Possiamo perciò scrivere le variabili p e ρ nella forma

$$p = p_0 + p' \quad \rho = \rho_0 + \rho' \quad (1.13)$$

dove p_0 e ρ_0 sono i valori costanti di pressione e densità all'equilibrio, mentre p' e ρ' sono le loro variazioni nel campo d'onda sonora ($p' \ll p_0$, $\rho' \ll \rho_0$). Ponendo quindi $\eta = 0$ e $\vec{F} = 0$ nell'equazione 1.7, ed eseguendo il cambiamento di variabile definito in 1.13, si ottiene la forma linearizzata dell'equazione di Eulero:

$$\rho_0 \frac{\partial \vec{v}}{\partial t} + \vec{\nabla} p' = 0 \quad (1.14)$$

Questa equazione rappresenta tre delle equazioni scalari necessarie a risolvere il problema del moto delle particelle acustiche. Poichè la densità di forza $\vec{\nabla} p'$ dovuta alla variazione di pressione della particella soddisfa l'identità vettoriale $\vec{\nabla} \times (\vec{\nabla} p') \equiv 0$, se per la 1.14 vale anche la condizione di irrotazionalità del moto delle particelle acustiche, cioè

$$\vec{\nabla} \times \vec{v} = 0 \quad (1.15)$$

allora il campo di velocità \vec{v} può essere ricavato come gradiente di un potenziale scalare ϕ . Questo è il potenziale di velocità, o *potenziale cinetico*, la cui esistenza permette di scrivere i campi d'onda (p', \vec{v}) nella forma

$$p' = -\rho_0 \frac{\partial \phi}{\partial t} \quad (1.16)$$

$$\vec{v} = \vec{\nabla} \phi \quad (1.17)$$

annullando così identicamente l'equazione 1.14. La risoluzione del problema del moto delle particelle acustiche è così ricondotto alla sola determinazione del loro potenziale cinetico ϕ .

Poichè non è tanto la forma funzionale di ϕ ma semplicemente la sua esistenza che permette - tramite le relazioni 1.16 e 1.17 - di soddisfare l'equazione 1.14, per determinare ϕ sarà necessario sfruttare le altre due condizioni del moto: l'equazione di continuità per la massa 1.3 e la condizione di adiabaticità 1.11.

Eseguendo il cambiamento di variabili 1.13 nell'equazione di continuità della massa, si ottiene la sua forma linearizzata come

$$\frac{1}{\rho_0} \frac{\partial \rho'}{\partial t} + \vec{\nabla} \cdot \vec{v} = 0 \quad (1.18)$$

e sostituendo in questa equazione il valore di $\rho' = (\rho_0/\gamma p_0)p'$ che si ottiene dall'equazione 1.11 con lo stesso procedimento di linearizzazione, si ricava⁵:

$$\frac{1}{\gamma p_0} \frac{\partial p'}{\partial t} + \vec{\nabla} \cdot \vec{v} = 0 \quad (1.19)$$

Soltanto quando p' e \vec{v} sono espresse tramite ϕ l'equazione 1.14 è identicamente soddisfatta, per cui effettuando quest'ultima sostituzione nella 1.19 si ottiene infine

$$\Delta \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0 \quad (1.20)$$

che è l'equazione delle onde sonore, ovvero l'unica equazione scalare che soltanto quando risolta nell'incognita ϕ determina completamente il problema del moto delle particelle acustiche.

Infine, si noti che la costante

$$c = \sqrt{\frac{\gamma p_0}{\rho_0}} \quad (1.21)$$

che compare nella 1.20 ha le dimensioni di una velocità e dipende soltanto dalla temperatura del fluido, non dalla sua pressione o densità. Ciò è evidente analizzando le proprietà termodinamiche del fluido stesso, che nel caso dei gas si esprimono con la nota legge $PV = NkT$. Si ha infatti $c^2 = \gamma p_0 V / \rho_0 V = \gamma kT / m$, dove m è la massa di una molecola di gas.

Nel caso dell'aria si trova $c(20^\circ C) = 343.4 \text{ m/s}$.

Per comprendere il significato fisico di c occorre considerare che se il profilo geometrico di un'onda deve mantenersi uguale a se stesso mentre l'onda si propaga nello spazio con velocità costante c' , essa deve essere necessariamente rappresentata da una funzione continua di forma $W(x, t) = f(x - c't)$. Sostituendo questa funzione di prova nell'equazione d'onda si scopre che $f(x - c't)$ soddisfa la 1.20 se e solo se $c' = c$. La costante c rappresenta dunque la velocità con cui il potenziale cinetico ϕ che governa il moto delle particelle acustiche si sposta nello spazio sotto forma di onda. Tale potenziale non è però direttamente osservabile, ma è possibile soltanto misurarne gli effetti nello spazio-tempo

⁵Infatti sviluppando $p(\rho)$ in serie di Taylor intorno al valore di equilibrio $p_0 = p(\rho_0)$ si ha $p(\rho) = p(\rho_0) + (\partial p / \partial \rho)_0 (\rho - \rho_0) + \dots = p_0 + p' + \dots$

ma, per la 1.11, si trova: $(\partial p / \partial \rho)_0 = \text{costante} \times (\gamma \rho^{\gamma-1})_0 = \text{costante} \times \left(\frac{\gamma}{\rho} \rho^\gamma \right)_0 = \frac{\gamma p_0}{\rho_0}$, ovvero $p' = (\gamma p_0 / \rho_0) \rho'$.

a quattro dimensioni mediante il quadrivettore (p', \vec{v}) che, come verrà discusso nella sezione successiva, costituisce il vero evento fisico nello spazio acustico di Minkovsky.

1.2 Il campo acustico nello spazio quadri-dimensionale

D'ora in avanti, nel riferirci alle grandezze acustiche di pressione e densità abbandoneremo la scrittura con l'apice e useremo i simboli p e ρ per riferirci alla differenza rispetto ai valori all'equilibrio. La densità di equilibrio verrà indicata col simbolo ρ_0 . Come visto nella sezione 1.1, le grandezze fisiche dalle quali si ricavano le proprietà del campo acustico, ossia $p(\vec{x}, t)$ e $\vec{v}(\vec{x}, t)$, possono essere espresse introducendo un campo scalare reale, il potenziale cinetico $\phi(\vec{x}, t)$, cosicchè risulta possibile derivarle direttamente da questo nel seguente modo:

$$p(\vec{x}, t) = -\rho_0 \frac{\partial \phi(\vec{x}, t)}{\partial t} \quad (1.22)$$

$$\vec{v}(\vec{x}, t) = \vec{\nabla} \phi(\vec{x}, t) \quad (1.23)$$

$$\rho(\vec{r}, t) = -\frac{\rho_0}{c^2} \frac{\partial \phi(\vec{x}, t)}{\partial t} \quad (1.24)$$

A partire dalle grandezze fisiche introdotte, è possibile derivare le espressioni per l'energia contenuta nel campo acustico. La densità di energia cinetica dovuta al moto acustico del fluido, ossia l'energia cinetica per unità di volume, è pari a

$$K = \frac{1}{2} \rho_0 v^2 \quad (1.25)$$

mentre l'energia potenziale associata ad un elemento di volume è $dU = -p(dV/V)$. Siccome $dV/V = -d\rho/\rho$ e $p/(\rho - \rho_0) = c^2$, si ha $d\rho = dp/c^2$, per cui si ottiene $dU = p dp/\rho_0 c^2$. L'energia potenziale per unità di volume si può pertanto esprimere come

$$U = \frac{p^2}{2\rho_0 c^2} \quad (1.26)$$

La densità di energia totale del campo d'onda, cioè la sua energia meccanica totale per unità di volume, è pari alla somma dell'energia cinetica e di quella potenziale e può essere espressa come

$$w = K + U = \frac{1}{2} \rho_0 \left(\frac{p^2}{z^2} + v^2 \right) \quad (1.27)$$

ove $z = \rho_0 c$ è l'impedenza caratteristica dell'aria. La densità di energia è dunque costituita da un termine cinetico, proporzionale al quadrato della velocità della particella acustica, e da un termine potenziale proporzionale al quadrato della sua pressione.

La *densità di quantità di moto (impulso)* del campo acustico, definita come la densità di quantità di moto del fluido, è data da

$$\vec{q} = \rho \vec{v} \quad (1.28)$$

1.2.1 Lo spazio-tempo acustico

Consideriamo lo spazio euclideo tridimensionale $E_3 = (\mathbb{R}^3, d)$, costituito dallo spazio vettoriale tridimensionale \mathbb{R}^3 e dalla metrica d . Un vettore in \mathbb{R}^3 viene rappresentato per mezzo di tre vettori linearmente indipendenti $\{e_i\}_{i=1,2,3}$ e per mezzo delle sue componenti x^i come

$$\mathbf{x} = \sum_{i=1}^3 x^i \mathbf{e}_i := x^i \mathbf{e}_i \quad (1.29)$$

La metrica d è definita come

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})} = |\mathbf{x} - \mathbf{y}| \quad (1.30)$$

con

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{e}_i x^i \cdot \mathbf{e}_j y^j = e_{ij} x^i y^j \quad (1.31)$$

e

$$e_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (1.32)$$

Siccome un campo sonoro è caratterizzato dalla dipendenza sia spaziale che temporale, si estende lo spazio alla quarta dimensione aggiungendo la variabile $x^0 = ct$, dove c indica la velocità del suono. La coordinata x^0 rappresenta la distanza percorsa dal suono in un tempo t . Lo *spazio-tempo acustico* è definito come lo spazio vettoriale \mathbb{R}^4 i cui vettori \mathbf{x} hanno componenti x^0, x^1, x^2, x^3 . I vettori \mathbf{x} vengono rappresentati in funzione dei vettori di base g_λ come

$$\mathbf{x} = \sum_{i=0}^3 x^i \mathbf{g}_i := x^i \mathbf{g}_i \quad (1.33)$$

Si introduce una metrica pseudo-euclidea così definita:

$$[m(x, y)]^2 = (x - y) \odot (x - y) = g_{\lambda\mu} (x^\lambda - y^\lambda)(x^\mu - y^\mu) \quad (1.34)$$

con

$$x \odot y = x^\lambda g_\lambda \odot x^\mu g_\mu = g_{\lambda\mu} x^\lambda y^\mu \quad (1.35)$$

e

$$g_{\lambda\mu} = \begin{cases} 1 & \lambda = \mu = 0 \\ 0 & \lambda \neq \mu \\ -1 & \lambda = \mu = 1, 2, 3 \end{cases} \quad (1.36)$$

Lo spazio \mathbb{R}^4 e la metrica m appena introdotti definiscono lo spazio quadridimensionale M_4 . Gli operatori di derivata rispetto alle coordinate degli spazi E_3 ed M_4 sono indicati da $\vec{\nabla} = \mathbf{e}^k \partial_k$ ($k = 1, 2, 3$), $\partial_\mu = \partial/\partial x^\mu$, $\partial_0 = (1/c)\partial/\partial t$ ($\mu = 0, 1, 2, 3$).

1.2.2 L'equazione d'onda in forma quadridimensionale

Nello spazio quadridimensionale appena introdotto, le variabili acustiche possono essere espresse come [8]

$$p = -z\partial_0\phi, \quad \mathbf{v} = v_k \mathbf{e}^k = \mathbf{e}^k \partial_k \phi = \vec{\nabla} \phi \quad (1.37)$$

Le densità di energia cinetica, potenziale e totale si scrivono come⁶

$$K = \frac{1}{2}\rho_e e_{ij}(\partial_i\phi)(\partial_j\phi) = \frac{1}{2}\rho_e |\nabla\phi|^2, \quad (1.38)$$

$$U = \frac{1}{2}\rho_e(\partial_0\phi)^2 = \frac{1}{2}\rho_e \frac{1}{c^2} \left(\frac{\partial\phi}{\partial t}\right)^2, \quad (1.39)$$

$$w = K + U = \frac{1}{2}\rho_e [e_{ij}(\partial_i\phi)(\partial_j\phi) + (\partial_0\phi)^2] \quad (1.40)$$

A questo punto possiamo scrivere la *densità di Lagrangiana* $L := K - U$, che può essere espressa come

$$L = -\frac{1}{2}\rho_e g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) \quad (1.41)$$

L'equazione d'onda può essere ottenuta applicando il principio variazionale alla densità di lagrangiana L , ottenendo le seguenti relazioni

$$\frac{\partial L}{\partial\phi} - \partial_\mu \frac{\partial L}{\partial\phi_\mu} = 0, \quad \phi_\mu = \partial_\mu\phi \quad (1.42)$$

da cui si giunge all'equazione di D'Alembert

$$\square\phi = 0 \quad (1.43)$$

con

$$\square := g^{\mu\nu}\partial_\mu\partial_\nu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \quad (1.44)$$

Anche il vettore quantità di moto può essere esteso al caso quadridimensionale definendo il vettore generalizzato *quadrimpulso* $P \in M_4$ di componenti

$$P^\mu := \frac{\partial L}{\partial\phi_\mu} = -\rho_e g^{\mu\nu}\phi_\nu = -\rho_e\phi^\mu \quad (1.45)$$

Si ha

$$P^0 = -\rho_e g^{00}\phi_0 = -\frac{\rho_e}{c}\frac{\partial\phi}{\partial t} = \frac{p}{c} = \rho c \quad (1.46)$$

e

$$P^i = -\rho_e g^{i\nu}\phi_\nu = -\rho_e\phi_i = \rho_e \frac{\partial\phi}{\partial x^i} = \rho_e v_i \quad (1.47)$$

cioè la prima componente del quadrimpulso corrisponde alla pressione acustica divisa per la velocità del suono, mentre le altre componenti sono uguali alla densità di flusso di massa del gas considerato come incomprimibile.

Per quanto visto nella sezione 1.1, il quadrivettore (p, v_x, v_y, v_z) costituisce un set completo di soluzioni dell'equazione d'onda ed è necessario e sufficiente ad identificare fisicamente lo stato meccanico di un suono nel tempo, in un arbitrario punto dello spazio all'interno di un campo acustico.

⁶Per evitare confusione coi valori degli indici μ, ν , in questo paragrafo poniamo sistematicamente $\rho_0 \equiv \rho_e$. Nel resto della tesi invece la densità di equilibrio sarà sempre indicata con ρ_0 .

1.3 Proprietà energetiche del campo acustico

Nella sezione 1.2 sono state definite le densità di energia cinetica, potenziale e totale K , U e w . Vogliamo ora considerare la derivata temporale della densità di energia totale:

$$\begin{aligned}\frac{\partial w}{\partial t} &= \rho_e \left(\frac{1}{z^2} p \frac{\partial p}{\partial t} + \vec{v} \cdot \frac{\partial \vec{v}}{\partial t} \right) \\ &= \rho_e \left(\frac{c^2}{z^2} p \frac{\partial p}{\partial t} + \vec{v} \cdot \frac{\partial \vec{v}}{\partial t} \right)\end{aligned}\quad (1.48)$$

Applicando l'equazione di conservazione della massa 1.18 e l'equazione di Eulero 1.14, si ottiene

$$\frac{\partial w}{\partial t} = -p \vec{\nabla} \cdot \vec{v} - \vec{v} \cdot \vec{\nabla} p = -\vec{\nabla} \cdot (p \vec{v}) \quad (1.49)$$

Nell'intensimetria, l'intensità sonora istantanea $\vec{I}(t)$ viene definita come prodotto della pressione e della velocità [9]

$$\vec{I}(t) = p(t) \vec{v}(t) \quad (1.50)$$

Pertanto, in base all'equazione 1.49, l'intensità sonora è legata alla densità di energia da una equazione di bilancio energetico:

$$\vec{\nabla} \cdot \vec{I}(t) = -\frac{\partial w(t)}{\partial t}, \quad (1.51)$$

che diventa

$$\vec{\nabla} \cdot \vec{I}(t) = -\frac{\partial w(t)}{\partial t} + w' \quad (1.52)$$

nel caso in cui un agente esterno, per esempio una sorgente sonora, compia lavoro sul sistema. L'equazione 1.51 associa all'intensità acustica \vec{I} il flusso di densità di energia del campo acustico.

1.3.1 Caso monocromatico

Consideriamo ora un campo monocromatico unidimensionale: la pressione può essere espressa come [10]

$$p(x, t) = P(x) \exp[i(\omega t + \phi(x))] \quad (1.53)$$

Calcolandone il gradiente otteniamo

$$\frac{\partial p}{\partial x} = \left[\frac{dP}{dx} + iP \frac{d\phi}{dx} \right] \exp[i(\omega t + \phi(x))] \quad (1.54)$$

Usando l'equazione di Eulero

$$\vec{\nabla} p = -\rho_0 \frac{\partial v}{\partial t} \quad (1.55)$$

possiamo ricavare la velocità delle particelle d'aria come

$$v = \frac{i}{\omega \rho_0} \frac{\partial p}{\partial x} = \frac{1}{\omega \rho_0} \left[-P \frac{d\phi}{dx} + i \frac{dP}{dx} \right] \exp[i(\omega t + \phi(x))] \quad (1.56)$$

Si osserva che v contiene un termine in fase con la pressione e un termine in quadratura. A questo punto l'intensità sonora, definita come prodotto delle due grandezze p e v in rappresentazione complessa, può essere suddivisa in una parte reale:

$$I_a(x, t) = Re \{I\} = -\frac{1}{\omega\rho_0} \left[P^2 \frac{d\phi}{dx} \right] \cos^2(\omega t + \phi) \quad (1.57)$$

e una parte immaginaria:

$$I_r(x, t) = Im \{I\} = -\frac{1}{4\omega\rho_0} \left(\frac{dP}{dx} \right)^2 \sin 2(\omega t + \phi) \quad (1.58)$$

Vogliamo ora considerare la media delle due componenti dell'intensità su un intervallo di tempo sufficientemente grande. La media stazionaria di una grandezza acustica $A(t)$ viene definita come

$$\langle A \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T A(t) dt \quad (1.59)$$

Essendo gli stimoli sonori di durata finita, la loro media viene di solito calcolata a partire da un istante precedente all'instaurarsi della perturbazione fino al suo cessare, oppure, per segnali tempo-stazionari, per un intervallo di tempo entro il quale le grandezze medie caratteristiche equivalgono alle grandezze medie calcolate su tutta la durata del segnale. Per l'intensità sonora di un campo monocromatico si ottiene:

$$\langle I_a(x) \rangle = -\frac{1}{2\omega\rho_0} P^2 \frac{d\phi}{dx} \quad (1.60)$$

e

$$\langle I_r(x) \rangle = 0 \quad (1.61)$$

$I_a(x)$ prende il nome di *intensità attiva*, $I_r(x)$ viene definita *intensità reattiva*⁷. Estendendo la trattazione al caso tridimensionale, si trova

$$\vec{I}_a(\vec{x}, t) = Re \{ \vec{I} \} = -\frac{1}{\omega\rho_0} \left[P^2 \vec{\nabla} \phi(\vec{x}, t) \right] \cos^2(\omega t + \phi) \quad (1.62)$$

$$\vec{I}_r(\vec{x}, t) = Im \{ \vec{I} \} = -\frac{1}{2\omega\rho_0} P(\vec{x}, t) \vec{\nabla} P(\vec{x}, t) \sin 2(\omega t + \phi) \quad (1.63)$$

Essendo $\vec{I}_a(\vec{x}, t) \propto \vec{\nabla} \phi(\vec{x}, t)$, l'intensità attiva risulta ortogonale alle superfici a fase costante, cioè ai fronti d'onda.

L'intensità attiva è legata al flusso di energia del campo acustico. In particolare, il vettore $\langle \vec{I}_a(\vec{x}) \rangle$ identifica la direzione nella quale avviene il trasferimento di energia nel punto \vec{x} .

A partire dalla rappresentazione complessa delle variabili $p(\omega)$ e $\vec{v}(\omega)$, viene definita per ogni componente d'onda di Fourier la grandezza complessa *impedenza acustica specifica*

$$Z_S(\omega) = \frac{p(\omega)}{v(\omega)} = R_S + iX_S \quad (1.64)$$

⁷N.B. La terminologia *attiva* e *reattiva* è stata introdotta in letteratura sulla base della trattazione dei campi monocromatici e viene abbandonata nel caso generale, trattato nella sezione 1.3.2.

dove $v(\omega)$ è la velocità dell'aria nella direzione definita dal vettore \vec{I}_a , cioè nella direzione di propagazione dell'energia. L'impedenza specifica ha le stesse dimensioni di un'impedenza meccanica per unità di superficie; l'unità di superficie viene scelta ortogonale al vettore \vec{I}_a . Notiamo l'analogia con l'elettronica, dove l'impedenza di un componente è data dal rapporto tra tensione e corrente, mentre la potenza dissipata è pari al loro prodotto. La parte reale dell'impedenza R_S viene definita *resistenza acustica*. La direzione determinata dall'intensità attiva risulta per definizione la direzione che definisce la componente della velocità in fase con la pressione, ovvero quella che determina un trasporto netto di energia.

1.3.2 Caso generale

Recentemente un nuovo approccio all'analisi energetica dei campi acustici, basato sul formalismo quadrifonico e sull'analisi nel dominio del tempo [11], ha portato alla suddivisione dell'intensità istantanea $\vec{I}(t) = p(t)\vec{v}(t)$ in *intensità radiante* $\vec{a}(t)$ e *intensità oscillante* $\vec{r}(t)$, così definite:

$$\vec{a}(t) = \frac{p^2 \langle p\vec{v} \rangle}{\langle p^2 \rangle} \quad (1.65)$$

$$\vec{r}(t) = \vec{I}(t) - \vec{a}(t) = p\vec{v} - \frac{p^2 \langle p\vec{v} \rangle}{\langle p^2 \rangle} \quad (1.66)$$

Si dimostra che l'intensità attiva \vec{I}_a coincide con la media temporale dell'intensità radiante, infatti:

$$\vec{I}_a = \langle \vec{I}(t) \rangle = \langle p\vec{v} \rangle = \left\langle \frac{p^2 \langle p\vec{v} \rangle}{\langle p^2 \rangle} \right\rangle = \langle \vec{a} \rangle \equiv \vec{A} \quad (1.67)$$

L'intensità radiante è dunque la parte di intensità sonora responsabile della propagazione dell'energia lungo le linee di flusso del campo $p\vec{v}$. Per definizione, invece, risulta $\langle \vec{r} \rangle = 0$. $\vec{r}(t)$ rappresenta la frazione di energia che oscilla attorno al punto di misura in modo da dare un valor medio temporale nullo. Per poter avere una misura della quantità di energia oscillante è necessario calcolare il primo momento statistico non nullo del vettore intensità oscillante. L'espressione adatta a dare una misura dell'intensità oscillante è il tensore del secondo ordine $\mathfrak{R} = \sqrt{2} \langle \vec{r} \times \vec{r} \rangle$. Gli autovalori del tensore \mathfrak{R} possono essere interpretati come gli assi di un ellissoide che rappresenta la polarizzazione dell'energia oscillante [11]. Le grandezze di intensità radiante ed oscillante possono essere anche espresse mediante il loro modulo

$$|\vec{A}| = \sqrt{\sum_i A_i^2}, \quad R := \|\mathfrak{R}\| = \sqrt{\sum_{i,j} \mathfrak{R}_{ij}^2}, \quad i, j = 1, 2, 3 \quad (1.68)$$

e convertite in *dB re* 10^{-12}W/m^2 . I loro livelli di intensità così espressi vengono indicati con L_I ed L_R .

Sulla base della suddivisione 1.65, 1.66, la resistenza acustica specifica R_S può essere espressa in funzione dell'intensità attiva e dell'energia potenziale come:

$$R_S = \frac{\langle p^2 \rangle}{A} \quad (1.69)$$

Notiamo che l'equazione 1.69 prevede un valore infinito della resistenza acustica quando l'intensità attiva vale zero: questo si interpreta ammettendo che nei punti in cui il campo acustico ha resistenza infinita non ci sia trasporto di energia, analogamente a quanto succede in elettronica, dove in presenza di una resistenza infinita (circuito aperto) non c'è trasferimento o dissipazione di potenza elettrica.

Introduciamo qui alcuni indicatori energetici che verranno utilizzati nel seguito. L'indicatore σ , definito nel modo seguente:

$$\sigma = \frac{2\sqrt{W_{Pot}W_{Kin}}}{W_{Pot} + W_{Kin}} \quad (1.70)$$

ove $W_{Pot} = \frac{1}{2}\rho_0 \langle p^2 \rangle / z^2$ è la media temporale della densità di energia potenziale U e $W_{Kin} = \frac{1}{2}\rho_0 \langle v^2 \rangle$ è la media temporale della densità di energia cinetica K , rappresenta il bilanciamento tra le densità di energia cinetica e potenziale in un campo acustico. Si vede che σ vale 1 quando $W_{Kin} = W_{Pot}$, vale zero quando una delle due componenti dell'energia è nulla e, in generale, risulta $0 \leq \sigma \leq 1$. Si definisce poi l'indicatore

$$\eta = \frac{|\vec{A}|}{cW} \quad (1.71)$$

che rappresenta la frazione di energia che viene irradiata alla velocità c . Si ha $0 \leq \eta \leq 1$.

Esempi di misura delle principali grandezze energetiche introdotte in questo capitolo verranno presentati nel capitolo 2.

1.4 La Quadrifonia Acustica

Nella sezione 1.2 sono state presentate le basi teoriche che portano all'identificazione di un fenomeno acustico mediante le quattro grandezze fisiche pressione e velocità vettoriale; inoltre, come spiegato nella sezione 1.3, anche le proprietà energetiche del campo dipendono dalle stesse quattro quantità. Partendo da questi presupposti è sorta la necessità di fondare una teoria e definire un approccio sperimentale atto a fornire un metodo generale e completo per l'identificazione, la misura e la restituzione di un evento sonoro [12]. L'invenzione della Quadrifonia Acustica risponde a queste esigenze prevedendo l'acquisizione di una serie di dati e metadati che consentono, oltre alla cattura completa dell'evento sonoro, anche il controllo della fedeltà in fase di riproduzione. La procedura, come vedremo, permette anche di misurare le caratteristiche acustiche di un ambiente, dimostrandosi non solo completamente compatibile con le tecniche classiche usate in acustica architettonica, ma addirittura superiore dal punto di vista dell'accuratezza grazie all'uso dell'intensimetria.

1.4.1 Dati Quadrifonici

Il processo di registrazione di un evento sonoro viene definito dalla Quadrifonia Acustica come la trasduzione elettro-acustica e la memorizzazione dei quattro segnali corrispondenti alle grandezze fisiche di pressione e velocità, per tutta la

durata dell'evento sonoro considerato. Dalla registrazione si ottiene così un set di quattro segnali

$$\Delta_{Quad} := \{s_\alpha(t)\}_{\alpha=0,1,2,3} \quad (1.72)$$

che possono essere rappresentati matematicamente nel seguente modo:

$$s_0(t) = p(\vec{r}_M, t) = -\rho_0 \frac{\partial \phi(\vec{r}, t)}{\partial t}; \quad \{s_i(t)\}_{i=1,2,3} = \vec{v}(\vec{r}_M, t) = \nabla \phi(\vec{r}_M, t) \quad (1.73)$$

L'insieme $\{s_\alpha(t)\}_{\alpha=0,1,2,3}$ costituisce i dati quadrifonici.

1.4.2 Risposta quadrifonica all'impulso

In un ambiente, tutto ciò con cui le onde sonore possono interagire, venendo riflesse, diffratte o assorbite, costituisce la condizione al contorno che definisce le soluzioni dell'equazione d'onda quadridimensionale. Il contributo delle condizioni al contorno, per un suono che viene generato e recepito in una coppia di punti all'interno dell'ambiente, può essere riassunto nella risposta all'impulso corrispondente, che, in un sistema lineare, rappresenta la funzione di trasferimento (si veda in proposito l'appendice A).

Consideriamo, all'interno di un ambiente, un punto \vec{r}_S in cui sia collocata una sorgente sonora in grado di generare un arbitrario segnale di pressione $S_p(\tau)$. Sia \vec{r}_M l'origine di un sistema di riferimento cartesiano, in cui vogliamo misurare le soluzioni dell'equazione d'onda relative al campo prodotto dalla sorgente. Il campo $\{s_\alpha(t)\}_{\alpha=0,1,2,3}$ può essere espresso come convoluzione temporale del segnale di pressione della sorgente $S_p(\tau)$ con le quattro componenti della risposta quadrifonica all'impulso $\{g_p, g_{vx}, g_{vy}, g_{vz}\}$, secondo le seguenti relazioni:

$$\begin{aligned} p(\vec{r}_M, \vec{r}_S; t) &= \int_{-\infty}^t g_p(\vec{r}_M, \vec{r}_S, t - \tau) S_p(\tau) d\tau \\ v_x(\vec{r}_M, \vec{r}_S; t) &= \int_{-\infty}^t g_{vx}(\vec{r}_M, \vec{r}_S, t - \tau) S_p(\tau) d\tau \\ v_y(\vec{r}_M, \vec{r}_S; t) &= \int_{-\infty}^t g_{vy}(\vec{r}_M, \vec{r}_S, t - \tau) S_p(\tau) d\tau \\ v_z(\vec{r}_M, \vec{r}_S; t) &= \int_{-\infty}^t g_{vz}(\vec{r}_M, \vec{r}_S, t - \tau) S_p(\tau) d\tau \end{aligned} \quad (1.74)$$

Riassumiamo brevemente la giustificazione di questa affermazione, seguendo la procedura svolta in [13]. Abbiamo visto come la pressione si possa esprimere come derivata temporale del potenziale cinetico, come descritto nell'equazione 1.23. Nel caso in cui la sorgente sonora sia puntiforme e presenti un andamento temporale descritto dalla funzione $s(t)$ (termine di disomogeneità dell'equazione d'onda), il potenziale cinetico può essere scritto come convoluzione temporale del segnale della sorgente con la risposta all'impulso del potenziale

$$\phi(\vec{r}, t) = \int_{-\infty}^t g(\vec{r}, t - \tau) s(\tau) d\tau \quad (1.75)$$

dove la funzione $g(\vec{r}, t)$ rappresenta la risposta all'impulso del potenziale cinetico e soddisfa l'equazione d'onda

$$\nabla^2 g(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 g(\vec{r}, t)}{\partial t^2} = \delta(\vec{r}) \delta(t) \quad (1.76)$$

Dalle equazioni 1.75 e 1.22, effettuando il cambiamento di variabili $t' = t - \tau$ si ricava

$$p(\vec{r}, t) = \int_{-\infty}^t g(\vec{r}, t - \tau) s_p(\tau) d\tau \quad (1.77)$$

dove $s_p(\tau) = -\rho_0 \partial s(t-t')/\partial t$ è il segnale di pressione emesso dalla sorgente. La velocità delle particelle d'aria si ricava integrando l'equazione di Eulero 1.55

$$\vec{v}(\vec{r}, t) = -\frac{1}{\rho_0} \int_{-\infty}^t \vec{\nabla} p(\vec{r}, \tau) d\tau \quad (1.78)$$

da cui si ha

$$\vec{v}(\vec{r}, t) = \int_{-\infty}^t \left(-\frac{1}{\rho_0} \int_0^{t-\tau} \vec{\nabla} g(\vec{r}, t') dt' \right) s_p(\tau) d\tau \quad (1.79)$$

Definendo la funzione

$$g_{\vec{v}}(\vec{r}, t) = -\frac{1}{\rho_0} \int_0^t \vec{\nabla} g(\vec{r}, \tau) d\tau \quad (1.80)$$

e sostituendo nell'equazione 1.77, si ottiene per la velocità la seguente espressione:

$$\vec{v}(\vec{r}, t) = \int_{-\infty}^t g_{\vec{v}}(\vec{r}, t-\tau) s_p(\tau) d\tau \quad (1.81)$$

in cui la velocità può essere interpretata come convoluzione della funzione $g_{\vec{v}}$ e dello stesso segnale di pressione $s_p(t)$.

L'insieme $\Gamma_{Quad} = \{g_p, g_{vx}, g_{vy}, g_{vz}\}$ è costituito dai quattro segnali corrispondenti alle risposte di pressione e velocità all'impulso, cioè ad un segnale rappresentato dalla funzione $\delta(\vec{r}_S, t_0)$. Γ_{Quad} rappresenta il propagatore quadrimensionale del campo acustico, e costituisce uno dei metadati del processo di registrazione quadrifonica. Tecnicamente esso viene acquisito e memorizzato alla stregua dei dati quadrifonici, secondo una procedura che verrà descritta nella sezione 2.2.

1.4.3 Indicatori del campo e metadati

Dopo aver definito i dati, vengono presentati i metadati associati all'evento sonoro. Si introduce la media statistica $\overline{\Gamma_{Quad}}$ associata al propagatore:

$$\overline{\Gamma_{Quad}} := \{\Gamma_p, \Gamma_v\} \equiv \left\{ \sqrt{\int_{t_0}^{\infty} g_p^2(t) dt}, \sqrt{\int_{t_0}^{\infty} [g_{vx}^2(t) + g_{vy}^2(t) + g_{vz}^2(t)] dt} \right\} \quad (1.82)$$

ove t_0 rappresenta l'istante in cui viene generato l'impulso nel punto \vec{r}_S , mentre Γ_{Quad} viene misurato in \vec{r}_M . Vengono definiti due fattori di normalizzazione

$$\Gamma_A := \frac{1}{2} [\rho_0 c \Gamma_v^2 + (\rho_0 c)^{-1} \Gamma_p^2] \quad (1.83)$$

$$\Gamma_F := \Gamma_v \Gamma_p \quad (1.84)$$

Anche la risposta quadrifonica all'impulso rappresenta un evento sonoro, pertanto è possibile calcolarne le quantità energetiche descritte nella sezione 1.3. Si definisce *intensità impulsiva* l'intensità acustica associata alla risposta all'impulso; essa è un vettore che ha per componenti e modulo

$$J_i = \int_{t_0}^{\infty} g_p(t) g_{v_i}(t) dt; \quad J := \sqrt{J_1^2 + J_2^2 + J_3^2} \quad i = 1, 2, 3 \quad (1.85)$$

Da J , Γ_A e Γ_F si definiscono i due principali indicatori energetici η e ξ e da questi altri cinque indicatori vengono ricavati, nel modo seguente:

$$\eta := J/\Gamma_A; \quad \xi := J/\Gamma_F; \quad \sigma := \Gamma_F/\Gamma_A; \quad (1.86)$$

$$\mu := \sqrt{1 - \eta^2}; \quad \kappa := \sqrt{1 - \xi^2}; \quad \zeta := \sigma\sqrt{1 - \xi^2}; \quad \tau := \sigma\sqrt{(1 - \xi^2)/(1 - \eta^2)}; \quad (1.87)$$

Notiamo che la definizione degli indicatori η e σ è equivalente a quella già data dalle relazioni 1.70 e 1.71. η rappresenta la frazione di energia che si propaga nel campo acustico nel punto di misura, σ rappresenta il bilanciamento tra le componenti cinetica e potenziale della densità di energia. L'indicatore ξ può essere ricavato a partire da η e σ : dalla definizione segue infatti $\xi = \eta/\sigma$. Gli altri indicatori quadratici sono tutti funzione di η , σ e ξ e vengono presentati soltanto per completezza.

A questo punto si può definire l'insieme completo dei metadati quadrifonici come

$$M_{Quad} := \{\Gamma_{Quad}, \sigma, \eta, \xi, \mu, \kappa, \zeta, \tau\} \quad (1.88)$$

Come si vede, una volta misurato Γ_{Quad} gli altri metadati possono essere calcolati a partire da g_p e $g_{\vec{v}}$ con tecniche di calcolo numerico. Una registrazione quadrifonica dunque prevede l'acquisizione di tutto l'insieme di dati e metadati

$$\Sigma_{Quad} := \{\Delta_{Quad}, M_{Quad}\} \quad (1.89)$$

1.4.4 Il processo di misura quadrifonica

Vogliamo ora descrivere schematicamente il metodo adottato per la misura quadrifonica di un evento sonoro, seguendo le disposizioni definite in [12]. In questa sezione verrà presentata la procedura generale, mentre i dettagli tecnici e i particolari sull'apparato utilizzato per le misure presentate in questo lavoro verranno trattati nel capitolo 2. La figura 1.1 riporta lo schema dei processi che costituiscono il metodo di registrazione quadrifonica. In primo luogo è necessario misurare le grandezze fisiche di pressione e velocità per tutta la durata dell'evento sonoro. Un'apposita sonda permette di rilevare la pressione e tre componenti ortogonali della velocità, che vengono convertite in tensioni elettriche aventi lo stesso andamento temporale (frequenza e fase) dei segnali fisici. I segnali analogici vengono a questo punto digitalizzati mediante convertitori A/D a quattro canali sincroni, acquisiti e memorizzati, permettendone contemporaneamente la visualizzazione e il monitoraggio audio.

La procedura di misurazione di Γ_{Quad} , ovvero della risposta quadrifonica all'impulso, è schematizzata in figura 1.2. L'apparato è composto da un dispositivo per la generazione di un segnale di eccitazione monofonico (il segnale di pressione $s_p(t)$) con contenuto spettrale uniforme; un trasduttore elettro-acustico che costituisce la sorgente sonora viene collocato nel punto \vec{r}_S , dal quale emette il segnale di pressione $s_p(t)$. L'ambiente nel quale sono collocati la sorgente ed il ricevitore costituisce la condizione al contorno che, unitamente all'equazione d'onda, determina il campo acustico nell'ambiente. Una sonda quadrifonica viene posizionata nel punto di ricezione \vec{r}_M , dove la risposta al segnale di eccitazione viene acquisita e memorizzata con il sistema descritto precedentemente

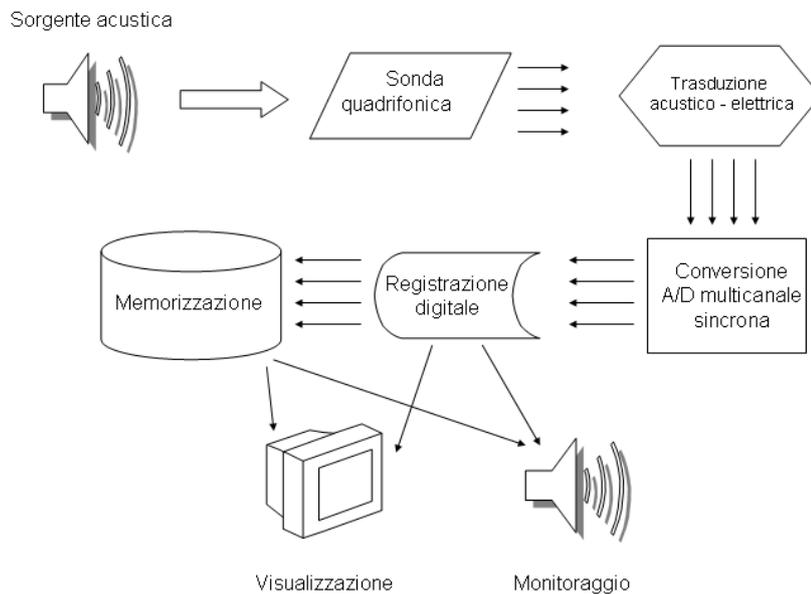


Figura 1.1: Schema del processo di registrazione quadrifonica.

e mostrato nella figura 1.1. La risposta quadrifonica all'impulso viene calcolata applicando un algoritmo di cross-correlazione tra il segnale di eccitazione e i quattro segnali di risposta. A questo punto la risposta all'impulso viene memorizzata e resa disponibile per le fasi successive del processo quadrifonico.

Il passo successivo consiste nel calcolo degli indicatori e delle proprietà energetiche del campo. La figura 1.3 riporta la procedura per il calcolo (misura indiretta) dei metadati a partire dalla risposta all'impulso. La risposta all'impulso viene filtrata e suddivisa in bande di terzi d'ottava⁸; in ogni banda vengono applicati degli algoritmi per il calcolo numerico degli indicatori e dei parametri energetici, secondo le relazioni descritte nella sezione 1.4.3. A questo punto gli indicatori, che assieme alla risposta quadrifonica all'impulso completano l'insieme dei metadati, vengono immagazzinati in memoria.

⁸Per alcune applicazioni si possono scegliere bande diverse, ad esempio bande di ottava, oppure effettuare una analisi in banda larga.

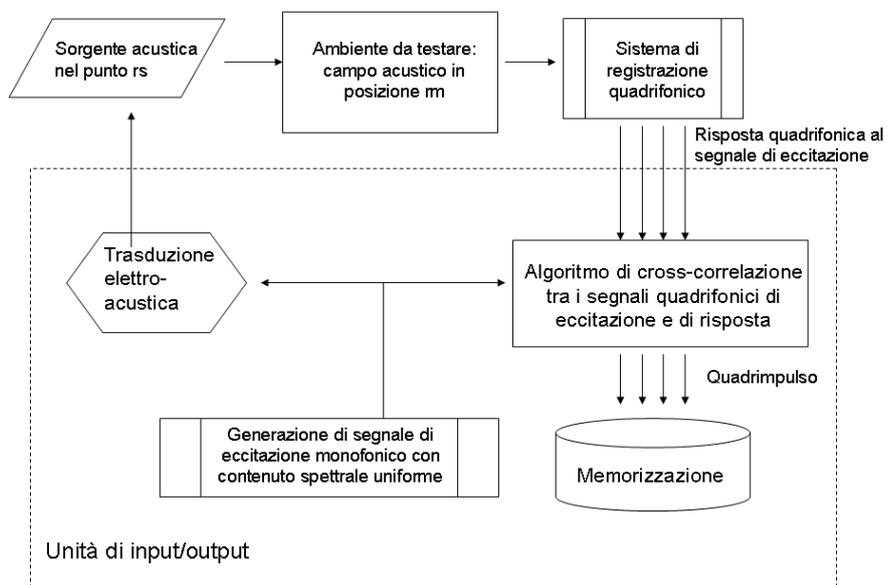


Figura 1.2: Schema del processo di misurazione della risposta quadrifonica all'impulso Γ_{Quad} .

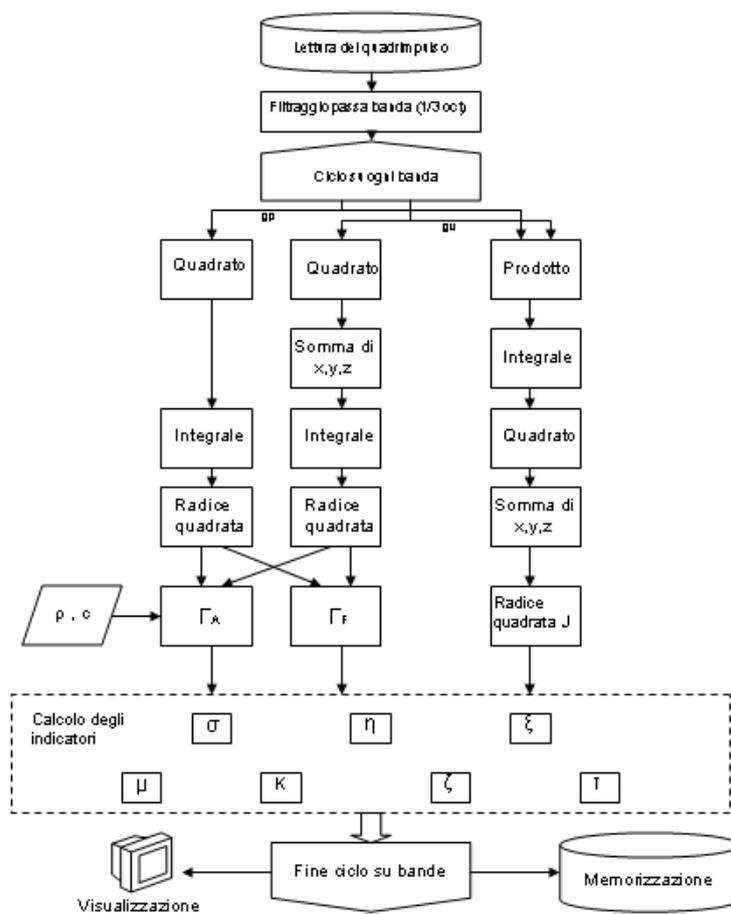


Figura 1.3: Schema del processo per il calcolo degli indicatori a partire dalla risposta quadrifonica all'impulso.

Capitolo 2

Tecniche di misura quadrifoniche acustiche e risultati sperimentali

2.1 La sonda Microflown

Nel 1994 fu inventato un sensore in grado di misurare direttamente la velocità vettoriale delle particelle d'aria [14], basato sul principio dell'anemometria a doppio filamento caldo. L'anemometria sfrutta la variazione di resistenza di un conduttore elettrico in funzione della temperatura: un flusso d'aria comporta trasferimento di calore per convezione in prossimità di un filamento riscaldato da corrente. Disponendo una coppia di filamenti paralleli riscaldati alla stessa temperatura in presenza di un flusso d'aria, si verifica una differenza di temperatura tra i filamenti dovuta al fatto che la corrente d'aria trasporta calore nella direzione del flusso, portando ad un riscaldamento del filamento "coperto" rispetto a quello direttamente esposto al flusso (si veda l'illustrazione di figura 2.1). A ciò consegue una variazione della resistenza elettrica dei due filamenti, che è possibile misurare inserendo i due conduttori in un circuito percorso da corrente. Dalla variazione di resistenza è possibile ottenere la variazione di temperatura e da essa la velocità del flusso d'aria.

Il sensore Microflown è costituito da due filamenti paralleli di platino, di $1mm$ di lunghezza e $200nm$ di spessore, ricoperti da uno strato di nitrato di silicio. I filamenti vengono riscaldati da una corrente elettrica e raggiungono una temperatura stazionaria compresa tra $200^{\circ}C$ e $400^{\circ}C$. Un aumento della temperatura dei sensori provoca un aumento della resistenza elettrica. Un singolo filamento riscaldato, in presenza di un flusso d'aria, subisce un raffreddamento a seguito dello scambio di calore per convezione con l'aria più fredda. Naturalmente un singolo filamento misura il modulo della velocità del flusso d'aria ortogonale ad esso, non essendo in grado di discriminare le varie direzioni d'arrivo sul piano ortogonale al filamento. Il sensore Microflown invece sfrutta la differenza di temperatura che si viene a creare tra i due sensori paralleli quando sono esposti al flusso d'aria. Il sistema ha una risposta polare a figura di otto,

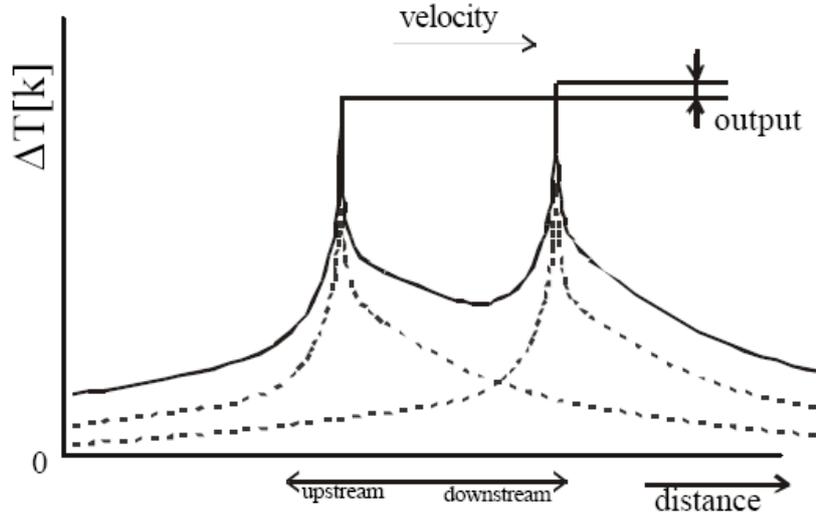


Figura 2.1: Profilo di temperatura in prossimità della coppia di filamenti in presenza di moto dell'aria. L'asimmetria nel profilo di temperatura provoca la differenza di temperatura tra i due sensori, che dà origine al segnale misurabile.

descritta dalla legge

$$S(\theta) = A \cos(\theta) \quad (2.1)$$

Il sensore permette di misurare la velocità delle particelle d'aria in un range approssimativamente compreso tra i 100 nm/s e $0.1 \div 1 \text{ m/s}$ per frequenze variabili da 0 Hz (flusso costante) fino a circa 20 kHz . Il sistema diventa meno efficace alle alte frequenze per due motivi: in primo luogo, l'effetto di convezione del calore tra i due filamenti richiede un certo tempo necessario alla propagazione del calore da un filo all'altro; in secondo luogo, la capacità termica dei filamenti ne ostacola l'immediato riscaldamento e raffreddamento. Il primo effetto può essere approssimato da un filtro passa basso del primo ordine, con una frequenza di taglio f_d compresa tra 500 Hz e 2 kHz ; il secondo effetto corrisponde esattamente ad un filtro passa basso del primo ordine, con frequenza di taglio f_h compresa tra 2 kHz e 15 kHz .¹ La tipica curva di risposta del trasduttore in funzione della frequenza, mostrata in figura 2.2, è descritta dalla seguente funzione:

$$S_{\text{Microflow}}(f) = \frac{LFS}{\sqrt{1 + \frac{f_e^2}{f^2}} \sqrt{1 + \frac{f^2}{f_d^2}} \sqrt{1 + \frac{f^2}{f_h^2}}} \quad (2.2)$$

dove LFS [$\frac{\text{mV}}{\text{Pa}} \cdot \rho c$] è la sensibilità misurata a 250 Hz . Questa legge descrive anche il *low frequency rise* di 6 dB/oct caratteristico dei sensori di velocità, associato ad una frequenza f_e compresa tra 30 Hz e 100 Hz .

La corrispondente risposta in fase è descritta dalla funzione

$$P_{\text{Microflow}}(f) = \arctan \frac{C_1}{f} - \arctan \frac{f}{C_2} - \arctan \frac{f}{C_3} \quad (2.3)$$

¹Le frequenze di taglio variano a seconda del tipo di modello considerato, in quanto dipendono dalla geometria del trasduttore e dalla temperatura di utilizzo. I valori reali vengono ottenuti mediante una procedura di calibrazione.

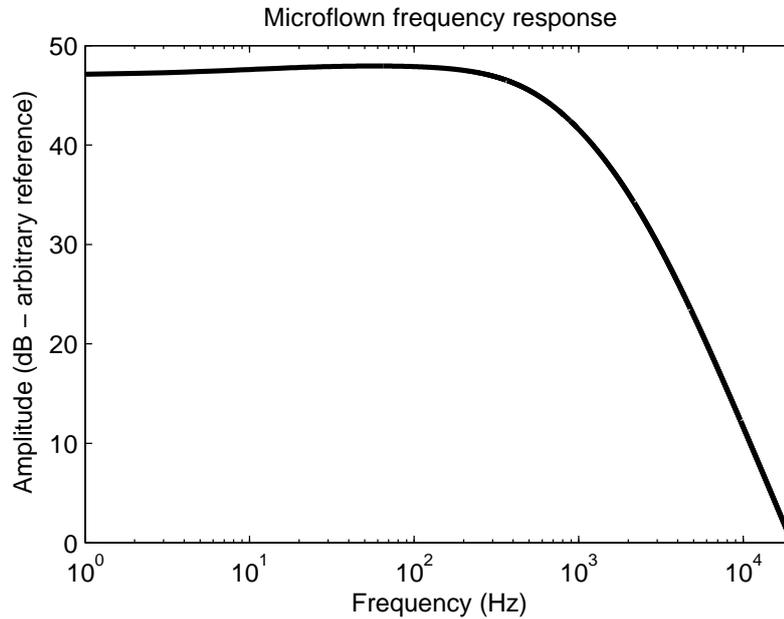


Figura 2.2: Risposta in frequenza del trasduttore Microflow.

con $C_1 \simeq f_e$, $C_2 \simeq f_d$, $C_3 \simeq f_h$.

2.1.1 La procedura di calibrazione della sonda

Per utilizzare la sonda per misurazioni intensimetriche è necessario effettuare la calibrazione, cioè misurarne la curva di risposta e ricavare mediante *fitting* i valori dei parametri f_e , f_d ed f_h . Vogliamo precisare che la calibrazione completa di una sonda intensimetrica è dettata da due necessità: una è quella di conoscere le curve di risposta e i livelli relativi dei segnali di pressione e di velocità, per poter poi deconvolvere la curva di risposta dai dati misurati; l'altra è quella di poter tradurre questi dati in unità fisiche e livelli assoluti, esprimendo la pressione in Pa e la velocità in m/s . La calibrazione dei livelli relativi e la misura della curva di risposta viene effettuata ponendo la sonda in un campo sonoro di cui si conosca la relazione di ampiezza e fase esistente tra pressione e velocità. Tra i vari metodi possibili, presentiamo i due più frequentemente utilizzati: il metodo del tubo di Kundt e il metodo della camera anecoica.

2.1.1.1 Il metodo del tubo di Kundt

Il tubo di Kundt è un cilindro metallico la cui superficie offre un coefficiente di riflessione sonora prossimo all'unità. Ad un'estremità è collocato un altoparlante, mentre l'altra estremità viene chiusa con una superficie metallica riflettente.

In riferimento alla figura 2.3, la pressione nel punto x è data dalla somma della pressione associata all'onda diretta emessa dall'altoparlante e da quella dell'onda riflessa dall'estremità metallica. La velocità è ricavabile dall'equazione

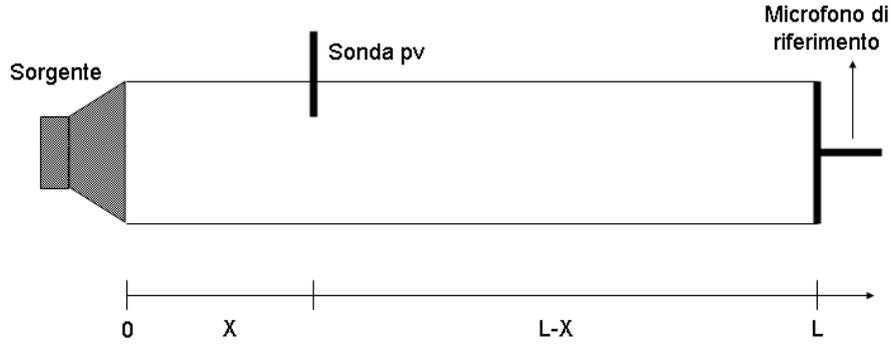


Figura 2.3: Schema dell'apparato per la calibrazione della sonda nel tubo di Kundt.

di Eulero 1.55. Usando la notazione esponenziale, si ha:

$$p(x, t, \omega) = (Ae^{-ikx} + Be^{ikx}) e^{i\omega t} \quad (2.4)$$

$$v(x, t, \omega) = -\frac{1}{i\omega\rho_0} \frac{\partial p(x, t, \omega)}{\partial x} = \frac{1}{\rho_0 c} (Ae^{-ikx} + Be^{ikx}) e^{i\omega t} \quad (2.5)$$

dove ω è la frequenza del suono e $k = 2\pi/\lambda = \omega/c$ è il numero d'onda. Un microfono di riferimento viene posto all'estremità riflettente del tubo, dove si ha $v(L) = 0$ a causa della riflessione totale del suono, mentre all'altra estremità la condizione al contorno è $p(0) = p_0$, dove p_0 è il segnale generato dall'altoparlante. La pressione e la velocità nella posizione x risultano essere

$$p(x, \omega) = p_0 \frac{\cos(k(L-x))}{\cos(kL)} \quad (2.6)$$

$$v(x, \omega) = \frac{p_0}{\rho_0 c} \frac{i \sin(k(L-x))}{\cos(kL)} \quad (2.7)$$

Se la sonda pv da calibrare è posta nella posizione x , il rapporto tra il segnale di pressione p_{sonda} misurato in x ed il segnale di riferimento p_{ref} misurato in L è dato da

$$\frac{p_{sonda}}{p_{ref}} = \cos(k(L-x)) \quad (2.8)$$

mentre il rapporto tra il segnale di velocità della sonda v_{sonda} ed il segnale di riferimento deve valere

$$\frac{v_{sonda}}{p_{ref}} = \frac{i}{\rho_0 c} \sin(k(L-x)) \quad (2.9)$$

Misurando questi rapporti in risposta ad uno sweep sinusoidale² è possibile ricavare la risposta in frequenza della sonda pv [15]. Per quanto riguarda la risposta in fase, facendo il rapporto tra le grandezze di velocità e pressione nel punto x si ha

$$\frac{v_{sonda}}{p_{sonda}} = \frac{i}{\rho_0 c} \tan(k(L - x)) \quad (2.10)$$

il che indica che il campo di pressione e quello di velocità sono sfasati di 90° . Dal confronto tra lo sfasamento previsto e lo sfasamento misurato si ricostruisce la curva di risposta in fase.

Il metodo del tubo di Kundt è applicabile fintantochè all'interno del tubo ci sono le condizioni per l'instaurarsi di un campo unidimensionale di onde piane stazionarie. Questo vale se la lunghezza λ dell'onda sonora è grande rispetto alla dimensione d della sezione del tubo, cioè se $c/f \gtrsim d$ ³. Considerando che il posizionamento dell'altoparlante pone un limite inferiore di una decina di cm al diametro del tubo, il metodo proposto per la calibrazione risulterà inapplicabile alle alte frequenze. Il metodo infatti viene solitamente usato nel range $20Hz < f < 4kHz$.

2.1.1.2 Il metodo della camera anecoica

Una camera anecoica, ossia un ambiente all'interno del quale sia impedita ogni riflessione del suono, si presta all'effettuazione di misurazioni in campo libero, cioè in condizioni in cui il campo sonoro coincide col campo diretto generato dalla sorgente. Una sorgente puntiforme che emette suono in ambiente anecoico genera un campo di onde sferiche, per il quale il rapporto tra la pressione e la velocità nella direzione di propagazione alla distanza r dalla sorgente è espresso dalla seguente relazione:

$$\frac{p(r, \omega, t)}{v(r, \omega, t)} = \rho_0 c \frac{ikr}{1 + ikr} = Z(r, \omega) \quad (2.11)$$

dove Z è l'impedenza acustica, definita nel dominio della frequenza come rapporto tra pressione e velocità. Ricordando che $k = 2\pi/\lambda$, per $r \gg \lambda$ si ha $kr \gg 1$, pertanto è possibile approssimare l'impedenza al valore $Z = \rho_0 c$. Questa approssimazione equivale a considerare il suono prodotto dalla sorgente come un campo di onde piane progressive. Effettuando la misura del rapporto tra pressione e velocità si può dunque ricavare la funzione di trasferimento della sonda. Il metodo si basa sull'approssimazione di onda piana progressiva, che è lecita purchè la distanza dalla sorgente sia di un ordine di grandezza superiore alla lunghezza d'onda del suono. Considerando dimensioni tipiche di camere anecoiche inferiori ai $10m$, è evidente che il metodo può essere applicato alle frequenze medio alte ma non nella regione bassa dello spettro.

Questa caratteristica rende il metodo della camera anecoica complementare, anzichè alternativo, al metodo del tubo di Kundt; una calibrazione accurata richiederebbe quindi l'effettuazione di una misura con entrambe le tecniche descritte.

²Cioè ad un segnale sinusoidale di frequenza variabile nel tempo entro un certo intervallo prestabilito.

³Per un tubo di sezione circolare la frequenza di *cutoff* in funzione del diametro si esprime come $\omega = 1.841c/r$, dove r è il raggio del tubo.

2.1.2 Il modello USP

Esistono in commercio diversi modelli di sonde che utilizzano il trasduttore di velocità Microflown. Le sonde assiali (o unidimensionali) utilizzano un trasduttore per la misura della componente di velocità in una dimensione; le sonde pv abbinano al trasduttore di velocità un trasduttore di pressione, nella forma di un microfono ad elettrete di piccole dimensioni ($1/10''$) posizionato in prossimità del trasduttore di velocità, in modo da consentire la misura simultanea e coincidente di p e v ; il modello più avanzato è la sonda *USP* (ultimate sound probe), mostrata in figura 2.4 e utilizzata per le misure effettuate in questo lavoro, che dispone di tre sensori di velocità montati in prossimità l'uno dell'altro e orientati lungo tre direzioni ortogonali tra di loro, abbinati ad un trasduttore di pressione ad elettrete.



Figura 2.4: La sonda *Microflown USP* è composta da tre trasduttori di velocità basati sul principio dell'anemometria a doppio filamento caldo, disposti lungo gli assi ortogonali di un sistema di riferimento cartesiano. A sinistra osserviamo il dettaglio dei trasduttori: il sensore rosso misura la velocità lungo l'asse x , il sensore verde lungo l'asse y e il sensore blu lungo l'asse z . Al centro si nota il trasduttore di pressione dalla membrana circolare. Nell'immagine di destra vengono evidenziate le ridotte dimensioni della sonda, che permettono di ridurre l'interferenza col campo acustico da misurare.

Questa sonda innovativa permette la misura diretta e simultanea delle tre componenti del vettore \vec{v} e della pressione p in un punto dello spazio. La sonda, di forma cilindrica, ha una lunghezza di 13cm ed un diametro di $1,25\text{cm}$, con un peso di 43gr . Il trasduttore di pressione è costituito da un microfono ad elettrete con membrana da $1/10''$, risposta in frequenza $20\text{Hz} - 20\text{kHz}$, caratteristiche omnidirezionali e sensibilità di $20\text{mV}/\text{Pa}$. I trasduttori di velocità funzionano in un range di frequenza $1\text{Hz} - 20\text{kHz}$, sono bi-direzionali (pattern a figura di 8) e hanno una sensibilità di $20\text{V}/(\text{m}/\text{s})$. Il costruttore opera la procedura di calibrazione e fornisce i valori dei parametri LFS , f_e , f_d ed f_h . Nel caso della sonda utilizzata in questo lavoro, costituita da tre trasduttori di velocità, il costruttore fornisce i dati di calibrazione per ogni sensore. Assieme alla sonda viene fornito un condizionatore che ha la duplice funzione di alimentare la sonda, fornendo le correnti necessarie al riscaldamento dei filamenti, e di operare la deconvoluzione

della risposta mediante filtraggio elettronico analogico, implementato tenendo conto dei parametri ricavati dal *fitting* dei dati in fase di calibrazione. I filtri presenti nel condizionatore si limitano a regolarizzare la curva di risposta, senza variare il guadagno dei segnali per compensare le differenze di sensibilità. Tale correzione verrà pertanto implementata nel dominio digitale attraverso un apposito software, tenendo conto dei dati forniti dal costruttore per la regolazione della sensibilità relativa tra i quattro canali. Un apposito interruttore permette di disinserire il circuito di filtraggio lasciando all'utente la possibilità di implementarlo nel modo preferito, per esempio nel dominio digitale. Il condizionatore fornisce le uscite dei quattro segnali su connettori BNC. La sonda viene collegata all'alimentatore mediante un cavo multipolare con connettore LEMO a 12 poli. L'ampiezza del segnale di velocità è normalizzato a quello di pressione, cioè viene moltiplicato per il modulo dell'impedenza caratteristica dell'aria $\rho_0 c$ in modo che i sensori di pressione e velocità forniscano lo stesso valore in presenza di un'onda piana progressiva.

Questa sonda, rispetto alle tecniche tradizionali di misura indiretta della velocità a partire dal gradiente di pressione per mezzo dell'equazione di Eulero, ha il pregio di evitare le approssimazioni e le inaccurately intrinseche delle sonde p-p. Inoltre le sue ridotte dimensioni limitano e rendono pressochè trascurabili gli effetti di diffrazione delle onde sonore dovuti alla presenza del trasduttore stesso nel campo acustico.

2.2 Il metodo di misura della risposta all'impulso

Come visto nella sezione 1.4, un evento sonoro può essere espresso mediante la convoluzione temporale del segnale di pressione generato dalla sorgente e della risposta quadrifonica all'impulso del sistema in cui si trova la sorgente. La risposta all'impulso pertanto contiene tutte le informazioni necessarie per studiare il comportamento del sistema fisico che costituisce la condizione al contorno dell'equazione d'onda. Come descritto nell'appendice A, la risposta $s'(t)$ di un sistema lineare ad un segnale $s(t)$ è data da

$$s'(t) = s(t) \otimes g(t) = \int_{-\infty}^{\infty} s(\tau)g(t - \tau)d\tau \quad (2.12)$$

Passando al dominio della frequenza mediante la trasformata di Fourier, l'equazione 2.12 può essere tradotta nella forma algebricamente più semplice

$$S'(\omega) = S(\omega)G(\omega) \quad (2.13)$$

2.2.1 Misura diretta della risposta all'impulso

Osserviamo che se il segnale fornito in ingresso al sistema è di tipo impulsivo, cioè $s(t) = \delta(t)$, la risposta sarà data da $s'(t) = \delta(t) \otimes g(t) = g(t)$. Questo significa che la risposta all'impulso di un sistema può essere misurata direttamente eccitando il sistema con un segnale impulsivo. Produrre un segnale acustico impulsivo significa generare, ad un certo istante di tempo e in un dato punto dello spazio, un suono di durata "infinitamente" breve. Ciò ovviamente non risulta

in pratica fattibile, per cui si ricorre a metodi che forniscono suoni della durata più breve possibile, seppur inevitabilmente finita. Svariate tecniche furono utilizzate nel corso degli anni per la generazione di suoni impulsivi. L'esempio probabilmente più antico, anche se non assimilabile a metodologia scientifica, è il battito di mani, ancora oggi consuetudine diffusa tra musicisti e direttori d'orchestra per saggiare empiricamente le caratteristiche delle sale da concerto. Tale metodo, pur non essendo usato per ottenere informazioni quantitative, risulta se non altro efficace per fornire una stima percettiva del tempo di riverberazione. Altre tecniche di misura diretta fanno ricorso allo scoppio di palloncini o, caso più frequente, allo sparo di pistole ad aria compressa, del tipo usato dagli starter nelle gare di atletica, o di armi caricate a salve. Un sistema meno diffuso sfrutta l'impulso sonoro associato ad una intensa scarica elettrostatica prodotta da dispositivi in grado di generare scintille. Citiamo infine anche la possibilità di generare elettronicamente un segnale impulsivo (ad esempio, si può definire un impulso nel dominio digitale assegnando valore unitario ad un campione di una sequenza e valore nullo a tutti gli altri campioni) e di diffonderlo in ambiente mediante un altoparlante.

Il segnale generato nell'ambiente viene poi registrato da un dispositivo analogico o digitale ed il suo andamento temporale viene stampato o reso disponibile per l'analisi tramite calcolatore. Nel campo dell'acustica architettonica, il metodo di misura diretta di risposta all'impulso fu applicato sistematicamente in tempi in cui non erano state ancora sviluppate tecniche più accurate, o in situazioni in cui non fosse possibile disporre dell'apparecchiatura necessaria per misurazioni più sofisticate. Senz'altro la misura diretta di risposta all'impulso ha il vantaggio di non richiedere strumentazione sofisticata, di essere di facile implementazione e relativamente economica. Può essere adeguata agli scopi nel caso in cui si voglia ricavare l'ecogramma del suono, cioè l'andamento temporale dell'intensità delle riflessioni generate in un ambiente. Tuttavia, questo metodo è affetto da limiti invalicabili che determinano una sostanziale inadeguatezza in situazioni che richiedono risultati accurati. In primo luogo, un impulso così generato ha una durata finita e uno spettro che non contiene tutte le frequenze del range udibile. In secondo luogo, il segnale di eccitazione non è ripetibile, poiché due spari o scoppi di palloncino avranno, in genere, durate, intensità e contenuto spettrale diversi, pertanto la misura stessa non darà risultati ripetibili. Il limite più severo è dato dall'impossibilità di conoscere a priori l'andamento temporale e lo spettro del segnale generato, negando così la possibilità di deconvolvere la risposta della sorgente dalla misura o di stimare le inaccurattezze da cui la risposta misurata è affetta. Inoltre, in ambienti di grandi dimensioni, può essere difficile riuscire a generare un segnale di potenza sufficientemente elevata da garantire il necessario rapporto segnale/rumore nel punto di ricezione durante tutto il decadimento sonoro.

2.2.2 Tecniche di cross-correlazione

Si definisce *cross-correlazione* di due segnali $a(t)$ e $b(t)$ l'espressione

$$\phi_{ab}(\tau) = \lim_{T_0 \rightarrow \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} a(t)b(t + \tau)dt \quad (2.14)$$

Calcolando la cross-correlazione tra il segnale di eccitazione $s(t)$ e la risposta del sistema $s'(t)$, nell'ipotesi che la risposta del sistema sia data dall'equazione

2.12, si ha

$$\begin{aligned}
\phi_{ss'}(\tau) &= \lim_{T_0 \rightarrow \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} s(t)s'(t+\tau)dt \\
&= \lim_{T_0 \rightarrow \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} s(t) \int_{-\infty}^{\infty} g(t')s(t+\tau-t')dt' dt \\
&= \int_{-\infty}^{\infty} g(t') \left[\lim_{T_0 \rightarrow \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} s(t)s(t+\tau-t')dt \right] dt' \\
&= \int_{-\infty}^{\infty} g(t')\phi_{ss}(\tau-t')dt' \tag{2.15}
\end{aligned}$$

La funzione

$$\phi_{ss}(\tau) = \lim_{T_0 \rightarrow \infty} \frac{1}{2T_0} \int_{-T_0}^{T_0} s(t)s(t+\tau)dt \tag{2.16}$$

viene definita *autocorrelazione*. Nel caso in cui la funzione di eccitazione abbia autocorrelazione $\phi_{ss}(\tau) = \delta(\tau)$, si ottiene

$$\phi_{ss'}(\tau) = g(\tau) \tag{2.17}$$

cioè il calcolo della funzione di cross-correlazione tra l'ingresso e l'uscita del sistema fornisce direttamente la risposta all'impulso.

Come esempio, consideriamo il caso in cui il segnale di eccitazione sia rumore bianco, che per definizione ha autocorrelazione $\phi(\tau) = \delta(\tau)$. Per misurare la risposta all'impulso di un sistema basterebbe eseguire l'algoritmo di cross-correlazione con la risposta al rumore bianco, tuttavia si nota che il calcolo risulterebbe dispendioso in termini di tempo, in quanto richiede un'operazione di media temporale per ogni istante di tempo τ della risposta che si vuole misurare. Nel caso di implementazione nel dominio digitale, con una risoluzione temporale tipica di 1/44100s (valore necessario per ottenere una risposta in frequenza estesa su tutta la banda udibile) ed una risposta impulsiva dell'ordine di alcuni secondi, l'operazione richiederebbe il calcolo di oltre centomila medie temporali!

2.2.3 Tecnica "Sine Sweep"

La risposta all'impulso può essere ricavata nel dominio della frequenza, grazie a tecniche basate sulla trasformata di Fourier FT e sulla trasformata di Fourier inversa IFT . Digitalizzando i segnali di eccitazione e risposta dell'ambiente e sfruttando l'algoritmo della FFT si ha, per i segnali $s(t)$ ed $s'(t)$, $S(\omega) = FT[s(t)]$ e $S'(\omega) = FT[s'(t)]$. Sfruttando la relazione 2.13, si può scrivere la risposta all'impulso $g(t)$ come

$$g(t) = IFT \left[\frac{FT[s'(t)]}{FT[s(t)]} \right] \tag{2.18}$$

L'equazione 2.18 sta alla base delle tecniche che sfruttano l'algoritmo della trasformata di Fourier per il calcolo della risposta all'impulso. Il segnale di eccitazione $s(t)$ immesso nell'ambiente deve avere un contenuto spettrale tale da coprire tutta la gamma dei suoni udibili. Il segnale di risposta $s'(t)$ deve essere acquisito e memorizzato in sincronia col segnale di eccitazione, cioè i due segnali

devono essere allineati nel tempo e la misura di s' deve partire esattamente all'istante in cui s viene emesso. Ai due segnali viene poi applicato l'algoritmo FFT, che calcola ampiezza e fase delle relative componenti spettrali. Successivamente viene calcolato il rapporto tra i numeri complessi $S'(\omega)$ ed $S(\omega)$, che fornisce la risposta in frequenza e in fase della funzione g , e poi viene applicata la trasformata di Fourier inversa per risalire alla risposta all'impulso nel dominio del tempo $g(t)$. Per ottenere un maggiore rapporto segnale/rumore, la misura viene ripetuta diverse volte e viene calcolata la media di tutte le risposte misurate, dopo averle sincronizzate. Le diverse risposte, teoricamente uguali per un sistema tempo-invariante, si sommano in maniera coerente, a differenza del rumore casuale, che quindi risulterà ridotto rispetto al segnale al termine del processo di media. È necessario osservare che, in generale, il segnale di risposta ha una durata temporale maggiore rispetto al segnale di eccitazione. Intuitivamente, questo è dovuto al fatto che il suono in un ambiente è soggetto al fenomeno della riverberazione, che ha come effetto la presenza di una coda di riflessioni a seguito di uno stimolo sonoro. Nell'ambito dell'analisi del sistema fisico, il fenomeno è dovuto alla presenza di effetti di ritardo e filtraggio nella risposta del sistema, dovuti alle riflessioni e alla velocità finita di propagazione delle onde sonore. Questo fatto implica la necessità di allungare il segnale di eccitazione inserendo un intervallo di silenzio alla fine del segnale, fino a raggiungere la durata del segnale di risposta. Questa operazione prende il nome di *zero padding*, in quanto, nel dominio digitale, aggiungere silenzio significa aggiungere campioni che assumono un valore nullo, cioè composti da soli zeri.

Un caso particolare di tecnica FT di misura di risposta all'impulso va sotto il nome di *Sine Sweep*, o sweep sinusoidale. Lo sweep sinusoidale è un segnale sinusoidale di ampiezza costante e frequenza variabile nel tempo, descritto da una funzione del tipo $s(t) = \sin[f(t)]$. La frequenza istantanea è pari alla derivata temporale della funzione $f(t)$. Gli sweep sinusoidali più frequentemente usati sono lo sweep lineare e lo sweep logaritmico. In entrambi i casi la frequenza è una funzione crescente del tempo. Per misurazioni di acustica, generalmente si sceglie un intervallo di frequenze che copra tutta la gamma dei suoni udibili, da $\omega_1 = 20Hz$ a $\omega_2 = 20kHz$, entro un tempo T di poche decine di secondi.

Per uno sweep lineare si ha

$$\frac{df(t)}{dt} = \omega_1 + \frac{\omega_2 - \omega_1}{T}t \quad (2.19)$$

da cui otteniamo

$$f(t) = \omega_1 t + \frac{\omega_2 - \omega_1}{2T}t^2 \quad (2.20)$$

e

$$s(t) = \sin \left[\omega_1 t + \frac{\omega_2 - \omega_1}{2T}t^2 \right] \quad (2.21)$$

Lo sweep lineare garantisce l'emissione di una potenza costante in ogni banda di frequenza di larghezza costante: la potenza contenuta nell'intervallo $100Hz - 200Hz$ è la stessa contenuta nell'intervallo $1000Hz - 1100Hz$, e così via. Ciò significa che

$$\int_{t_1}^{t_1+\Delta t} s^2(\tau) d\tau = \int_{t_2}^{t_2+\Delta t} s^2(\tau) d\tau \quad (2.22)$$

o, analogamente, che

$$\int_{\omega(t1)}^{\omega(t1+\Delta t)} S^2(\omega) d\omega = \int_{\omega(t2)}^{\omega(t2+\Delta t)} S^2(\omega) d\omega \quad (2.23)$$

dove $S(\omega)$ è la trasformata di Fourier della funzione $s(t)$. Osserviamo inoltre che, essendo lo sweep lineare, si ha $\omega(t1 + \Delta t) - \omega(t1) = \omega(t2 + \Delta t) - \omega(t2)$.

Nel caso di uno sweep logaritmico abbiamo

$$s(t) = \sin \left[K \left(e^{t/L} - 1 \right) \right] \quad (2.24)$$

Le condizioni al contorno per ottenere i valori di K ed L sono che la frequenza iniziale e finale valgano rispettivamente ω_1 e ω_2 . Si ha

$$\left. \frac{d}{dt} K \left(e^{t/L} - 1 \right) \right|_{t=0} = \omega_1, \quad \left. \frac{d}{dt} K \left(e^{t/L} - 1 \right) \right|_{t=T} = \omega_2 \quad (2.25)$$

Risolvendo, si trova

$$K = \frac{T\omega_1}{\log(\omega_2/\omega_1)}, \quad L = \frac{T}{\log(\omega_2/\omega_1)} \quad (2.26)$$

per cui lo sweep logaritmico avrà la seguente espressione:

$$s(t) = \sin \left[\frac{T\omega_1}{\log \frac{\omega_2}{\omega_1}} \left(e^{(t/T) \log(\omega_2/\omega_1)} - 1 \right) \right] \quad (2.27)$$

Lo sweep logaritmico è caratterizzato da una potenza costante in ogni banda di larghezza logaritmica costante, cioè si ha

$$\int_{t1}^{k \cdot t1} s^2(\tau) d\tau = \int_{t2}^{k \cdot t2} s^2(\tau) d\tau \quad (2.28)$$

e, passando nel dominio della frequenza,

$$\int_{\omega(t1)}^{\omega(k \cdot t1)} S^2(\omega) d\omega = \int_{\omega(t2)}^{\omega(k \cdot t2)} S^2(\omega) d\omega \quad (2.29)$$

Per le proprietà dello sweep logaritmico, si ha $\omega(k \cdot t1)/\omega(t1) = \omega(k \cdot t2)/\omega(t2)$.

Osserviamo che, nel caso si riesca a trovare una funzione $i(t)$ tale che $s(t) \otimes i(t) = \delta(t)$, si ha

$$s'(t) = s(t) \otimes g(t) \Rightarrow i(t) \otimes s'(t) = i(t) \otimes s(t) \otimes g(t) \Rightarrow g(t) = i(t) \otimes s'(t) \quad (2.30)$$

La funzione $i(t)$ viene chiamata *funzione inversa* del segnale di eccitazione $s(t)$. In questo caso, la risposta all'impulso del sistema è ricavabile direttamente come convoluzione temporale della risposta del sistema con la funzione inversa del segnale di eccitazione. In alternativa, si può passare attraverso il dominio della frequenza e scrivere

$$g(t) = IFT [FT [i(t)] \cdot FT [s'(t)]] \quad (2.31)$$

Vogliamo ora trovare la funzione $i(t)$ nel caso in cui la funzione $s(t)$ sia lo sweep lineare 2.21. Essendo $s(t)$ una funzione reale, la sua trasformata di Fourier $S(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t)e^{-i\omega t} dt$ gode della proprietà $S(-\omega) = S^*(\omega)$. Consideriamo la funzione $s(-t)$ ottenuta operando l'inversione temporale della funzione $s(t)$. La sua trasformata di Fourier è la funzione

$$S'(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(-t)e^{-i\omega t} dt \quad (2.32)$$

$$= -\frac{1}{\sqrt{2\pi}} \int_{+\infty}^{-\infty} s(\tau)e^{i\omega\tau} d\tau \quad (2.33)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(\tau)e^{i\omega\tau} d\tau = S(-\omega) = S^*(\omega) \quad (2.34)$$

Facendo il prodotto di convoluzione di $s(t)$ ed $s(-t)$, otteniamo

$$s(t) \otimes s(-t) = IFT[S(\omega) \cdot S^*(\omega)] = IFT[|S(\omega)|^2] = IFT[const] = k \cdot \delta(t) \quad (2.35)$$

grazie al fatto che lo sweep lineare ha uno spettro di potenza indipendente dalla frequenza. Pertanto, nel caso dello sweep lineare, la funzione inversa $i(t)$ si ottiene dall'inversione temporale dello sweep.

Si può dimostrare che, nel caso in cui la funzione di eccitazione sia lo sweep logaritmico 2.27, la funzione inversa si ottiene operando la sostituzione $t \rightarrow -t$, cioè invertendone l'andamento temporale e modulandone l'ampiezza in modo da ottenere una diminuzione logaritmica di $3dB$ per ottava decrescente con la frequenza, cioè un dimezzamento dell'involuppo al dimezzare della frequenza. Questa necessità è dettata dal fatto che lo sweep logaritmico, per quanto visto sopra, ha una densità di potenza spettrale proporzionale al logaritmo della larghezza di banda, ovvero ha lo stesso spettro di potenza del rumore rosa, cui corrisponde una diminuzione di ampiezza di $3dB/oct$ al crescere della frequenza. Questa modulazione di ampiezza introdotta nella funzione inversa fa sì che, nel fare la convoluzione tra lo sweep logaritmico $s(t)$ e lo sweep invertito e modulato in ampiezza si ottenga, nel dominio della frequenza, una funzione indipendente dalla frequenza, la cui trasformata di Fourier inversa produce la funzione $\delta(t)$.

La funzione inversa dello sweep logaritmico ha la seguente espressione:

$$i(t) = 10^{-(3/10) \log_2(e^{-(t/T) \log(\omega_2/\omega_1)})} \sin \left[\frac{T\omega_1}{\log(\omega_2/\omega_1)} \left(e^{-(t/T) \log(\omega_2/\omega_1)} - 1 \right) \right] \quad (2.36)$$

L'invertibilità dello sweep logaritmico e la conseguente facilità nel calcolo della risposta all'impulso non è l'unico vantaggio di questa tecnica: è stato recentemente dimostrato che questa procedura permette di separare eventuali componenti non lineari presenti nel sistema dalla risposta lineare all'impulso [16]. Effettuando una misura in presenza di componenti non lineari, si osserva che queste danno luogo alla comparsa di segnali, simili a risposte all'impulso, temporalmente distinti dalla risposta del sistema e ben separati da essa (figura 2.5).

Questo effetto trova una duplice applicazione: da un lato, ripulire la misura dalle non linearità introdotte nel sistema dall'apparato di misura, quali ad esempio la distorsione armonica generata da amplificatori ed altoparlanti; dall'altro, permette contemporaneamente di misurare tali componenti non lineari

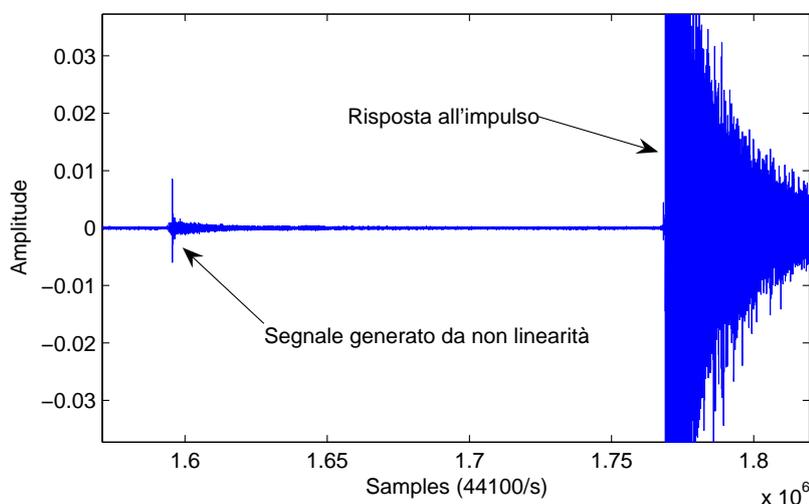


Figura 2.5: Segnale corrispondente a componente non lineare del sistema, dovuto a distorsione armonica dell'altoparlante. Il processo di deconvoluzione con tecnica sine sweep separa le componenti di distorsione armonica e le colloca prima della risposta all'impulso del sistema lineare.

per caratterizzarle.

Nei sistemi di misura è sempre presente del rumore, sia che si tratti del rumore di fondo acustico dell'ambiente in cui si svolgono le misure, sia che si tratti del rumore termico dell'elettronica nelle apparecchiature. Anche la risposta misurata pertanto sarà affetta dalla presenza di rumore casuale, che in genere copre uniformemente tutto lo spettro delle frequenze udibili. Pur essendo impossibile rimuovere completamente il rumore di fondo da una misurazione, si è osservato sperimentalmente che la tecnica *sine sweep* è in grado di fornire un rapporto segnale/rumore più elevato rispetto alle altre tecniche di misura descritte. Tale rapporto è sufficientemente alto da rendere sufficiente una sola misurazione, evitando la necessità di calcolare medie sincrone su diverse misurazioni. Inoltre, i dati evidenziano che il rapporto segnale/rumore migliora all'aumentare della durata temporale dello sweep, pertanto risulta possibile effettuare misurazioni con un dato rapporto S/N anche in presenza di rumore di fondo superiore al previsto, semplicemente aumentando la durata dello sweep. Per dimostrare la efficacia dello sweep logaritmico in presenza di rumore esterno e per ottenere un'indicazione della dipendenza del rumore di fondo presente nella risposta all'impulso dalla lunghezza dello sweep, è stata effettuata una serie di misure in presenza di una sorgente di rumore. In camera anecoica è stato utilizzato un altoparlante per generare lo sweep ad un livello di 80 dB SPL, in presenza di rumore bianco a 65 dB SPL emesso da un altro altoparlante presente nella camera. La misura è stata ripetuta con sweep di lunghezza 5s, 10s, 15s, 20s, 25s e 30s. Il rumore di fondo presente nella risposta all'impulso è stato quantificato calcolandone il valore medio *rms* su un intervallo di 5s dopo il completo decadimento della risposta all'impulso.



Figura 2.6: Spettro del rumore di fondo presente nella risposta all'impulso per tre diverse lunghezze dello sweep. L'irregolarità della curva riflette la risposta in frequenza dell'altoparlante usato per emettere il rumore bianco.

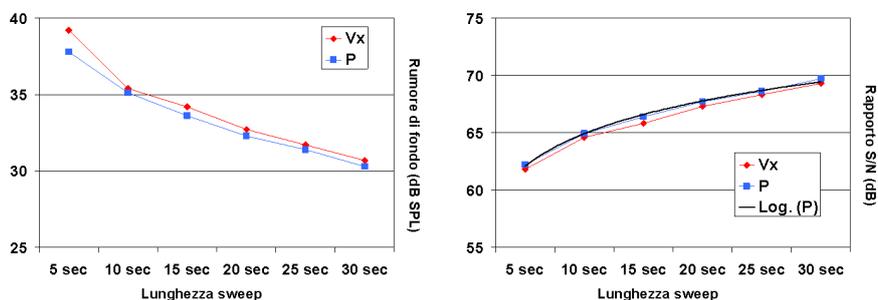


Figura 2.7: Andamento del rumore di fondo residuo contenuto nella risposta misurata con uno sweep logaritmico di durata variabile e relativo rapporto segnale/rumore in presenza di una sorgente esterna di rumore.

Come si vede dalle figure 2.6 e 2.7, il metodo dello sweep sinusoidale permette di attenuare significativamente il rumore esterno, con un andamento esponenziale decrescente in funzione della durata dello sweep. Nelle nostre misure la durata minima dello sweep è di 20s.

Bisogna sottolineare che la risposta misurata, pur essendo libera da distorsioni non lineari delle apparecchiature di misura, conserva l'impronta delle funzioni di trasferimento delle apparecchiature usate per la misura: ogni strumento, infatti, può essere interpretato come un filtro lineare la cui curva di risposta viene convoluta con la risposta del sistema. Per ricavare la risposta del sistema ripulita dalle risposte degli strumenti si può procedere in due modi: uno con-

siste nel misurare le curve di risposta in condizioni note, ricavarne la funzione inversa e deconvolverla dalle risposte misurate; l'alto metodo, utilizzato nel presente lavoro, consiste nell'utilizzare i filtri forniti direttamente dai costruttori per linearizzare le curve di risposta degli strumenti e acquisire direttamente la risposta ripulita.

2.3 La catena di acquisizione di dati e metadati

Vogliamo ora descrivere il setup sperimentale e le apparecchiature usate per le misure quadrifoniche. Come visto nella sezione 1.4, l'intero processo di misura quadrifonica può essere suddiviso in tre blocchi: la registrazione quadrifonica, la misura di risposta all'impulso ed il post-processing.

La registrazione quadrifonica prevede l'acquisizione simultanea dei segnali di pressione e velocità generati in un punto dello spazio in presenza di un evento sonoro. La trasduzione acustico-elettrica di questi segnali viene effettuata dalla sonda Microdown USP, descritta nella sezione 2.1. La pressione e le tre componenti ortogonali della velocità vengono fornite in output come quattro segnali analogici di tensione elettrica, in rapporto tra di loro come le grandezze acustiche p , $\rho_0 cv_x$, $\rho_0 cv_y$ e $\rho_0 cv_z$. Nel corso delle nostre misure ci siamo affidati al filtraggio analogico incorporato nel condizionatore per la linearizzazione della risposta dei trasduttori. Ricordando che alla pressione sonora di $1Pa$ corrisponde una intensità sonora di $94dB SPL$, la sensibilità della sonda, dell'ordine dei $20mV/Pa$, fa sì che le tensioni di uscita in presenza di campi acustici di intensità medio-alta siano dell'ordine di pochi millivolt. Prima di effettuare la digitalizzazione del segnale è pertanto necessario amplificare il segnale per poter sfruttare al meglio la gamma dinamica del convertitore ed ottenere una migliore risoluzione. L'amplificazione del segnale deve essere implementata attraverso uno stadio di guadagno lineare in grado di aumentare l'ampiezza della tensione senza introdurre significative distorsioni e rumore di fondo. Questa fase, nell'ambito dell'ingegneria audio ha preso il nome di *preamplificazione*, per distinguersi dalla *amplificazione*, che riguarda invece l'incremento di potenza elettrica di un segnale, tipicamente usato per il pilotaggio di altoparlanti. Per ottimizzare il trasferimento della tensione dalla sonda attraverso i cavi e il livello del segnale in uscita, è necessario usare un dispositivo che abbia una impedenza elettrica di ingresso pari ad almeno cinque volte l'impedenza d'uscita del condizionatore, e riesca a fornire un guadagno di decine di dB . Queste caratteristiche si ritrovano nei preamplificatori microfonic professionali, sui quali è pertanto caduta la nostra scelta per quanto riguarda la preamplificazione.

Il passo successivo consiste nella conversione analogico/digitale dei segnali. Questa viene effettuata secondo la tecnica *PCM (Pulse Code Modulation)*, che implementa il metodo *sample and hold* attraverso i processi di campionamento e quantizzazione. Il campionamento del segnale viene effettuato a intervalli regolari di tempo, regolati dalla frequenza di campionamento (*sampling frequency*) F_s . Per il teorema di Shannon-Nyquist, è necessario che la frequenza di campionamento sia almeno il doppio della frequenza più alta contenuta nel segnale da digitalizzare, per poter conservare tutta l'informazione e per non dar luogo alla comparsa di componenti indesiderate dovute al fenomeno dell'*aliasing*. Siccome la nostra analisi riguarda segnali acustici nella gamma di frequenze

udibili dall'orecchio umano, che si estende all'incirca da $20Hz$ fino a $20kHz$, una frequenza di campionamento di $44100Hz$ risulta adeguata agli scopi. Tale frequenza di campionamento risponde inoltre alle specifiche dello standard *Red Book* del formato CD-Audio, ed è compatibile con le specifiche dei formati DVD-A e DVD-V; questo assicura perciò la compatibilità del formato usato con lo standard dell'audio commerciale. Per quanto riguarda il campionamento, 16 bit sono sufficienti per fornire un adeguato rapporto segnale/rumore e una buona risoluzione. Il valore teorico del rapporto s/n a 16 bit è $96dB$. Portando la *bit depth* a 24, il valore teorico sale a $144dB$, di fatto limitato a circa $110dB$ a causa del rumore interno dell'elettronica impiegata negli stadi di conversione. Considerando che spesso il fattore limitante nelle misurazioni di acustica è il rumore di fondo dell'ambiente stesso, per le nostre misure abbiamo scelto di quantizzare il segnale a 16bit.



Figura 2.8: Interfaccia Firewire MOTU Traveler.

Per effettuare la preamplificazione e la conversione abbiamo usato l'interfaccia audio MOTU Traveler (figura 2.8). Questa interfaccia dispone di quattro preamplificatori microfonici il cui guadagno, controllato digitalmente, è variabile da 0 a 54 dB ad intervalli di 1dB. E' dotata di convertitori A/D e D/A a 16 e 24 bit e frequenze di campionamento da $44.1kHz$ a $192kHz$. Ideata per l'utilizzo con computer, è dotata di un bus *Firewire* per il collegamento e il trasferimento di dati in tempo reale con i pc.

La registrazione digitale di un evento sonoro viene fatta con un pc portatile dotato del software audio multitraccia *Adobe Audition*, versione 2.0; questo software permette l'acquisizione simultanea e il salvataggio in formato *wav* PCM lineare dei segnali digitali provenienti dai quattro canali dell'interfaccia. Durante l'acquisizione, la forma d'onda dei segnali viene visualizzata sul monitor assieme al valore di picco del segnale, e l'audio può essere inviato alle uscite della stessa scheda audio per permettere all'operatore di monitorare la qualità.

La procedura di misura della risposta quadrifonica all'impulso prevede la generazione di un segnale di eccitazione. Le nostre misure utilizzano la tecnica *sine sweep*, descritta nella sezione 2.2.3, per i vantaggi che offre in termini di rapidità di esecuzione e rapporto segnale/rumore. Per la generazione del segnale di

eccitazione ci si è serviti di una routine implementata con il software *Matlab*[®], che genera uno sweep in base ai parametri scelti dall'operatore attraverso un'interfaccia grafica (figure 2.9 e 2.10). In primo luogo è possibile optare tra uno sweep lineare o logaritmico. Le nostre misure utilizzano lo sweep logaritmico per sfruttare la separazione delle non linearità presenti nella sorgente elettroacustica utilizzata. Si devono poi scegliere gli estremi dell'intervallo di frequenza da sondare, la durata totale del segnale e la sua ampiezza digitale. L'intervallo di frequenza viene generalmente scelto tra 20Hz e 22kHz . L'ampiezza del segnale a , compresa tra 0 e 1, controlla il valore massimo della funzione $s(t) = a \sin[f(t)]$ generata nel dominio digitale. Alla fine dello sweep viene aggiunto un silenzio digitale, la cui durata dev'essere sufficiente a permettere il completo decadimento del suono, per non tagliare la coda sonora durante la simultanea registrazione. L'interfaccia grafica permette di eseguire la routine di

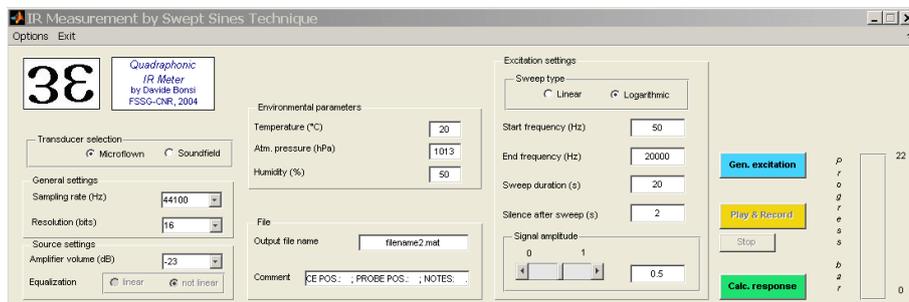


Figura 2.9: Interfaccia di misura della risposta all'impulso: i diversi blocchi e pulsanti permettono la selezione e regolazione dei parametri e l'esecuzione dei comandi per generare il segnale di eccitazione, acquisire la risposta del sistema e calcolare la risposta all'impulso.

generazione del segnale premendo il pulsante corrispondente. Successivamente viene calcolata e memorizzata la funzione inversa. Con un altro pulsante viene dato il via alla riproduzione del segnale di eccitazione e alla simultanea acquisizione dei quattro segnali di risposta. Il segnale di eccitazione precedentemente generato viene mandato dal software ad un canale del convertitore D/A della scheda, la cui uscita viene collegata alla sorgente sonora. Contemporaneamente, la risposta dell'ambiente viene misurata in un punto dello spazio dalla sonda, i cui segnali vengono acquisiti dallo stesso software di misura attraverso i quattro canali di preamplificazione e conversione A/D della scheda. A partire dalla risposta misurata (figura 2.11) e dall'eccitazione inversa calcolata, implementando il calcolo dell'espressione 2.31 per ognuno dei quattro segnali di pressione e velocità vengono calcolate e salvate in memoria le corrispondenti risposte all'impulso.

La sorgente sonora utilizzata per diffondere in ambiente il segnale di eccitazione, cioè per generare il segnale di pressione $s(t)$ è un altoparlante omnidirezionale dodecaedrico modello DL301 prodotto dalla ditta *Lookline* pilotato da un amplificatore di potenza esterno da 300W modello D301 (figura 2.12). Su ogni faccia del dodecaedro è posto un altoparlante *full range*, in grado cioè di coprire tutta la gamma di frequenze sonore al di sopra dei 40Hz . Tale disposizione

degli altoparlanti assicura che il segnale venga diffuso in tutte le direzioni. Gli altoparlanti sono pilotati da un amplificatore di potenza monofonico analogico, con controllo del livello di uscita a passi di $3dB$. E' possibile anche alimentare l'amplificatore con un pacco di batterie, per l'utilizzo in ambienti in cui non sia disponibile la rete elettrica o in presenza di forti disturbi di rete che possono essere trasmessi agli altoparlanti. L'amplificatore dispone di un filtro equalizzatore in grado di appiattire la curva di risposta degli altoparlanti, a scapito di una leggera riduzione della potenza di uscita. Il filtro è stato utilizzato durante le nostre misure. Questa sorgente è la principale fonte di non linearità presente nel nostro setup sperimentale in quanto presenta, come qualsiasi amplificatore ed altoparlante, componenti armoniche di distorsione; ciononostante, le nostre misure di risposta all'impulso non ne sono affette per quanto illustrato nella sezione 2.2.3.

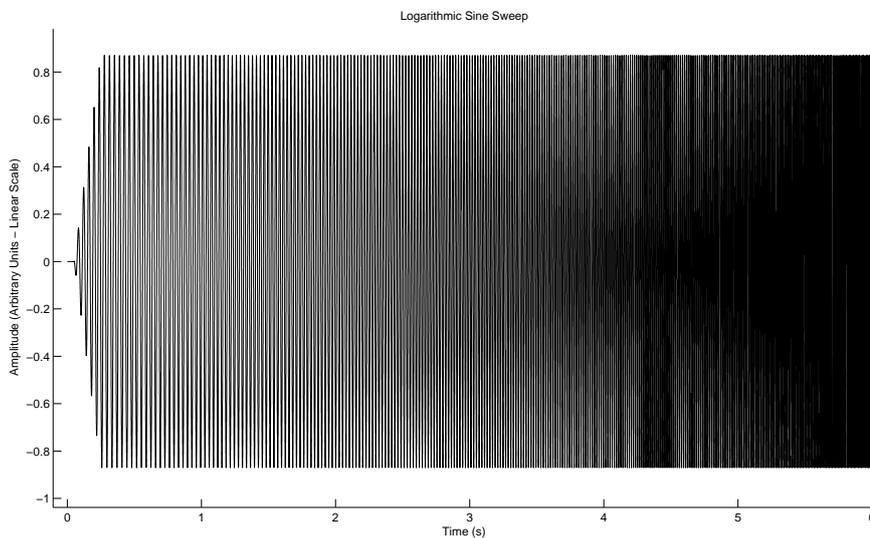


Figura 2.10: Andamento temporale dello sweep logaritmico.

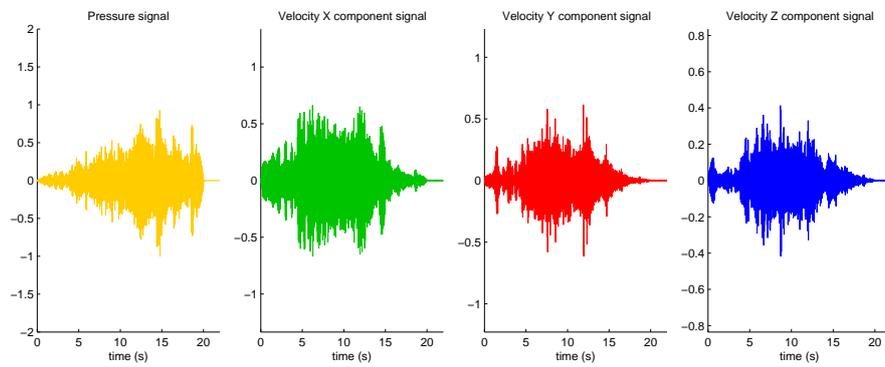


Figura 2.11: Andamento temporale delle risposte di pressione e velocità a sweep sinusoidale.



Figura 2.12: Sorgente omnidirezionale dodecaedrica con amplificatore di potenza.

2.4 L'analisi energetica di campi acustici a partire dalla risposta all'impulso

La figura 2.13 mostra l'andamento temporale del segnale di pressione di una risposta all'impulso nei primi 250ms. Il primo segnale, il più intenso, rappresenta il suono diretto che si propaga in linea retta dalla sorgente al ricevitore. Esso è immediatamente seguito dal suono riflesso dal pavimento e dalle prime riflessioni⁴ dell'ambiente, cui fa seguito una coda di decadimento sonoro nella quale le singole riflessioni sono indistinguibili.

Elevando al quadrato la pressione si ottiene l'ecogramma mostrato nella figura 2.14, una curva che fornisce una misura dell'energia contenuta nelle riflessioni. L'ampiezza dell'ecogramma può essere visualizzata in scala logaritmica, ovvero in decibel, per poter essere correlata direttamente alla nostra percezione (figura 2.15).

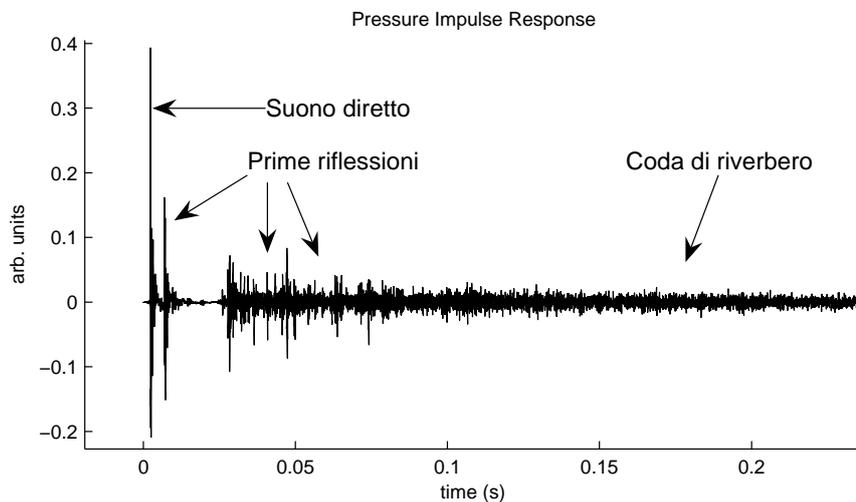


Figura 2.13: Andamento temporale di una risposta all'impulso di pressione; il segnale diretto è seguito dalle prime riflessioni e dalla coda riverberante.

⁴Per *prime riflessioni* non si intendono solamente i segnali che giungono al ricevitore dopo aver subito una sola riflessione dal contorno dell'ambiente ma, in generale, l'insieme di tutte le riflessioni, anche multiple, che giungono nelle prime decine di millisecondi, quando il loro numero per unità di tempo non è ancora sufficientemente elevato per permettere la descrizione statistica propria del campo riverberato.

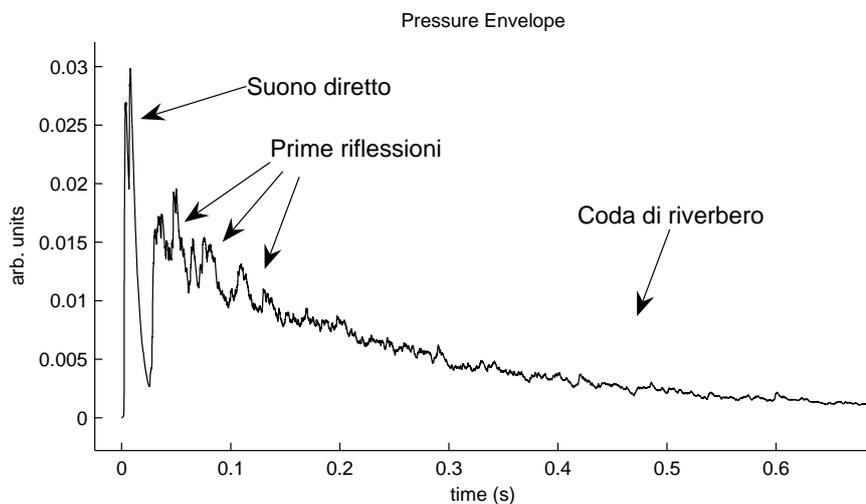


Figura 2.14: Ecogramma, ovvero andamento della pressione elevata al quadrato, calcolato dalla risposta all'impulso mostrata in figura 2.13.

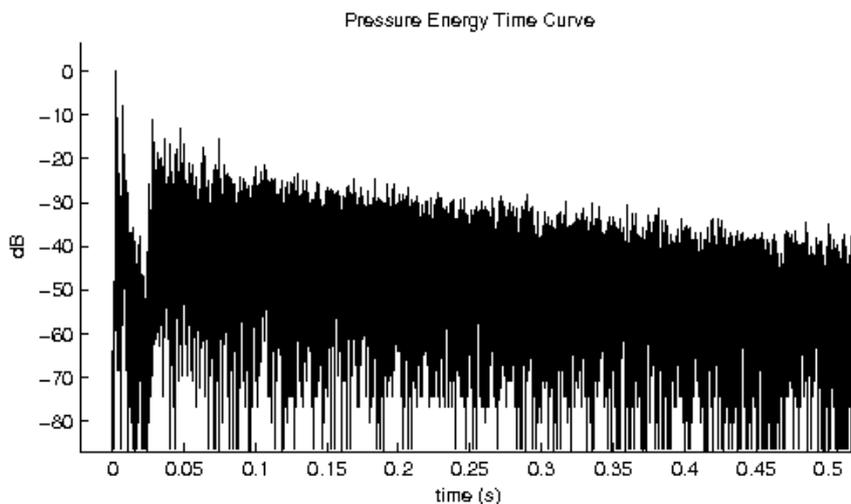


Figura 2.15: Ecogramma in scala logaritmica.

2.4.1 Il tempo di riverberazione

Tra i parametri acustici degli ambienti, il più importante è sicuramente il tempo di riverberazione. Esso è definito come l'intervallo di tempo necessario affinché, al cessare di uno stimolo sonoro, l'intensità da esso prodotta diminuisca di $60dB$. Storicamente, questa scelta fu dettata dalla necessità di poter calcolare un indicatore capace di fornire una misura della persistenza del suono in una sala. Siccome la pressione sonora prodotta nel corso di un *fortissimo* orchestrale può superare anche i $100dB SPL$, mentre il rumore di fondo di una sala da concerto silenziosa si attesta mediamente sui $40dB SPL$, la scelta di un intervallo

di $60dB$ per caratterizzare il decadimento appare la più ragionevole. Vogliamo inoltre far notare che una differenza di $60dB$ di pressione sonora corrisponde ad un fattore 1000 in ampiezza, cioè ad una variazione di tre ordini di grandezza nella pressione lineare.

Una delle tecniche più semplici per la misura del tempo di riverberazione è la misura diretta con la tecnica del rumore interrotto (*gated noise burst*): essa prevede l'emissione in ambiente di rumore bianco stazionario di intensità e durata sufficienti a saturare tutti i possibili modi propri di oscillazione dell'aria nell'ambiente. Interrompendo improvvisamente la sorgente di rumore, l'energia sonora immessa nell'ambiente comincia a decadere, essenzialmente a causa della dissipazione dell'energia che ha luogo in vicinanza del contorno, fino a scendere al di sotto della soglia di udibilità. Misurando l'andamento della pressione sonora nel corso del decadimento, è possibile ottenere il tempo di riverberazione come il tempo entro il quale l'intensità è diminuita di $60dB$. Questo valore viene indicato con $RT60$. Questa tecnica, apparentemente immediata e di facile applicazione, presenta in realtà dei limiti, sia pratici che teorici. In primo luogo, può funzionare solo in ambienti silenziosi e nei quali si possa immettere rumore ad un livello sufficientemente elevato da poterne misurare un decadimento di $60dB$. Esiste poi il problema che il punto scelto per la misura sia effettivamente un campione significativo per caratterizzare il decadimento in tutto l'ambiente. In condizioni di *campo diffuso*, cioè in presenza di un campo sonoro che possa essere trattato secondo un approccio statistico, il decadimento dell'energia può essere descritto dalla seguente legge, la cui forma è stata originariamente proposta da W.C. Sabine alla fine dell'ottocento:

$$W(t) = W_0 e^{-t/\tau} \quad (2.37)$$

W_0 è l'energia presente all'inizio del decadimento. Questa legge prevede il decadimento esponenziale con una costante di tempo τ che può assumere diverse forme a seconda del modello fisico con cui si interpreta il processo di decadimento. Definito il coefficiente di assorbimento medio $\bar{\alpha} = (\sum_i \alpha_i S_i) / S$, dove S è la superficie totale dell'ambiente, la formulazione di Eyring prevede $\tau = -(4V/cS) / \log(1 - \bar{\alpha})$, dove V è il volume dell'ambiente. Rappresentando l'andamento temporale dell'intensità sonora in scala logaritmica, ovvero in decibel, la pendenza di decadimento assume un andamento lineare. In queste condizioni il valore di $RT60$ può essere estrapolato da un intervallo di decadimento inferiore ai $60dB$ previsti dalla definizione. Qualora il campo non sia diffuso, il decadimento può differire dall'andamento teorico previsto. Vengono allora introdotti dei parametri diversi per rappresentare il tempo di riverberazione. Il più utilizzato è l' $RT30$, definito come tempo di decadimento di $60dB$ estrapolato mediante fitting della curva di decadimento tra gli istanti in cui il livello SPL assume i valori di $-5dB$ e $-35dB$ rispetto al valore presente al cessare dello stimolo stazionario. In maniera analoga vengono definiti l' $RT15$ e l'*Early Decay Time EDT*, che viene calcolato mediante fitting sui primi $10dB$ del decadimento. Nel corso del decadimento sonoro, la presenza di rumore di fondo costante può falsare o rendere impossibile la stima di $RT60$, pertanto è stato scelto l'indicatore $RT30$ per esprimere il tempo di riverberazione. Naturalmente, nelle condizioni ideali di campo diffuso tutti questi parametri coincidono. Il tempo di decadimento iniziale EDT fornisce invece una misura che si avvicina di più al tempo di riverberazione effettivamente percepito dal nostro sistema

uditivo, che è determinato principalmente dall'andamento dell'energia nel corso della prima parte del decadimento.

L'altro limite della misura diretta con tecnica *gated noise burst* è dato dal fatto che, anche nell'ipotesi di campo diffuso, la misura di un singolo evento di decadimento non è sufficiente a descrivere un fenomeno di tipo statistico: per ottenere una curva regolare sarebbe necessario effettuare una media su numerose curve di decadimento.

2.4.2 Il metodo dell'integrale di Schroeder

Si deve a Manfred Schroeder l'introduzione, negli anni sessanta, di una tecnica di misura basata su una relazione tra la risposta impulsiva di un sistema e la media su infiniti decadimenti di un segnale stazionario a banda larga, quale il rumore bianco. Grazie a questa relazione vedremo come sia possibile ricavare tutte le caratteristiche del decadimento dell'energia sonora in un ambiente da una singola misura di risposta all'impulso, e da questa ricavare i parametri oggettivi comunemente usati in acustica architettonica. Schroeder dimostrò la relazione e la applicò al quadrato della pressione per stimare il decadimento energetico, ma la tecnica può essere estesa anche alla velocità e all'intensità acustica.

Consideriamo una sorgente di rumore bianco $s(t)$ posta nel punto \vec{r}_S , un ricevitore posto in \vec{r}_M ed un ambiente in cui la funzione $h(t)$ (sia essa la pressione o la velocità) sia caratterizzata dalla risposta all'impulso $g(t)$ tra i punti \vec{r}_S ed \vec{r}_M . In base all'equazione 2.12, la funzione h può essere scritta come

$$h(t) = g(t) \otimes s(t) = \int_{-\infty}^{+\infty} g(t - \tau)s(\tau)d\tau \quad (2.38)$$

Nel caso in cui la sorgente s venga spenta all'istante $t = 0$, il valore della funzione $h^2(t)$ a tempi $t \geq 0$ è dato da

$$h^2(t) = \int_{-\infty}^0 g(t - \tau)s(\tau)d\tau \int_{-\infty}^0 g(t - \tau')s(\tau')d\tau' \quad (2.39)$$

Grazie alla proprietà di autocorrelazione del rumore bianco $\int s(t)s(t')dt dt' = \delta(t - t')$, la media su un insieme statistico di decadimenti è data da

$$\begin{aligned} \overline{h^2(t)} &= \int_{-\infty}^0 g(t - \tau)d\tau \int_{-\infty}^0 g(t - \tau')\overline{s(\tau)} \cdot \overline{s(\tau')}d\tau' \\ &= \int_{-\infty}^0 g(t - \tau)d\tau \int_{-\infty}^0 g(t - \tau')\delta(\tau - \tau')d\tau' \\ &= \int_{-\infty}^0 [g(t - \tau)]^2 d\tau = - \int_{\infty}^t [g(\tau'')]^2 d\tau'' = \int_t^{\infty} [g(\tau)]^2 d\tau \end{aligned} \quad (2.40)$$

La figura 2.16 mostra il decadimento della pressione quadratica $\overline{p^2(t)}$ calcolato a partire dalla risposta all'impulso $g_p(t)$ mostrata in figura 2.13 applicando il metodo dell'integrale di Schroeder. Confrontando il plot con quello di figura 2.15, ottenuto da un singolo decadimento di rumore bianco, si osserva un notevole miglioramento della regolarità del decadimento, che rende più accurato il

calcolo del tempo di riverberazione.

In maniera analoga alla pressione, il metodo descritto può essere impiegato per calcolare il decadimento della velocità, definito nel modo seguente:

$$\overline{v^2(t)} = \int_t^\infty [\vec{g}_v(\tau)]^2 d\tau \quad (2.41)$$

dove il vettore $\vec{g}_v = (g_{vx}, g_{vy}, g_{vz})$ ha per componenti le risposte all'impulso delle componenti della velocità.

A questo punto possiamo definire la media statistica della densità di energia impulsiva e dell'intensità sonora impulsiva nel modo seguente:

$$\overline{w(t)} = \frac{1}{2}\rho_0 \int_t^\infty \left\{ \frac{[g_p(\tau)]^2}{z^2} + [\vec{g}_v(\tau)]^2 \right\} d\tau \quad (2.42)$$

$$\overline{j_i(t)} = \int_t^\infty [g_p(\tau)g_{vi}(\tau)] d\tau, \quad i = x, y, z \quad (2.43)$$

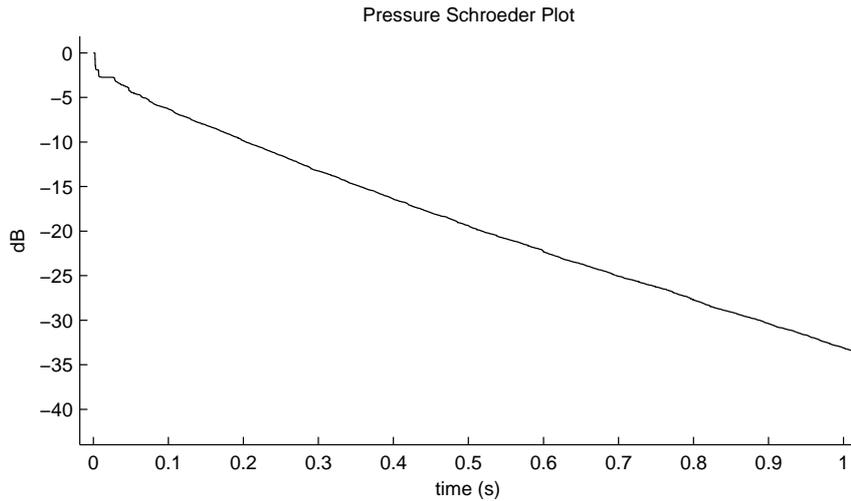


Figura 2.16: Schroeder plot ricavato dalla risposta all'impulso di pressione di figura 2.13.

Osserviamo che si ha

$$\frac{\partial \overline{w(t)}}{\partial t} = -\frac{1}{2}\rho_0 \left\{ \frac{[g_p(\tau)]^2}{z^2} + [\vec{g}_v(\tau)]^2 \right\}, \quad (2.44)$$

cioè la densità di energia ha un andamento monotono decrescente nel corso del decadimento. Questo vale anche per la pressione quadratica e per la velocità quadratica, mentre in generale non si verifica per le componenti dell'intensità impulsiva. Dalla pendenza della curva $\overline{w(t)}$, estrapolata mediante fitting tra $0dB$ e $-60dB$ rispetto al valore massimo, si può ricavare il tempo di decadimento dell'energia $WT60$, definito, analogamente al tempo di riverberazione $RT60$,

come il tempo necessario affinché l'energia sonora diminuisca di $60dB$ rispetto al valore iniziale.

Nel caso di sistemi ergodici, si dimostra che se $f(t)$ è una funzione quadratica nelle variabili acustiche si ha $\langle f \rangle = \overline{f(0)}$, quindi il metodo dell'integrale di Schroeder può essere usato in generale anche per la valutazione delle medie tempo-stazionarie delle grandezze acustiche quadratiche.

2.4.3 Parametri acustici oggettivi

A partire dalla risposta all'impulso di pressione possono essere calcolati col metodo dell'integrale di Schroeder i parametri più comunemente utilizzati per caratterizzare oggettivamente l'acustica di un ambiente in relazione alla nostra percezione uditiva.

Accanto al tempo di riverberazione $RT30$ e all' EDT , si può calcolare l'indice di chiarezza C_{80} , definito come il rapporto, espresso in decibel, tra l'energia sonora che giunge all'ascoltatore entro i primi $80ms$ e l'energia che arriva dopo $80ms$:

$$C_{80} = 10 \log_{10} \left\{ \frac{\int_0^{80ms} [g_p(t)]^2 dt}{\int_{80ms}^{\infty} [g_p(t)]^2 dt} \right\} \quad (2.45)$$

L'indice di chiarezza esprime il rapporto tra l'energia contenuta nelle prime riflessioni, in grado di supportare e rinforzare il suono diretto, e quella contenuta nella coda di riverbero. Esso è indicativo del grado con cui in un ambiente è possibile riuscire a distinguere le singole note di un passaggio musicale senza che queste si confondano nella coda riverberante. Studi effettuati indicano ottimali i valori di C_{80} compresi tra $+1dB$ e $+5dB$.

Analogo all'indice di chiarezza ma più adatto ad esprimere la resa acustica del parlato è l'indice di definizione D_{50} , pari al rapporto, espresso in percentuale, tra l'energia sonora che giunge entro i primi $50ms$ e l'energia totale:

$$D_{50} = \frac{\int_0^{50ms} [g_p(t)]^2 dt}{\int_0^{\infty} [g_p(t)]^2 dt} 100\% \quad (2.46)$$

Più grande è questo parametro, più il segnale viene percepito in maniera definita, perchè non è disturbato dalla coda di riverbero. Per il parlato, valori ottimali sono attorno al 50%.

Il tempo baricentrico t_s è il tempo tale per cui metà dell'energia sonora giunge all'ascoltatore entro t_s e metà giunge dopo. Si ha

$$t_s = \frac{\int_0^{\infty} [g_p(t)]^2 t dt}{\int_0^{\infty} [g_p(t)]^2 dt} \quad (2.47)$$

Qualitativamente, t_s fornisce una stima del tempo che separa la zona delle prime riflessioni da quella della coda riverberante. Per la musica, valori accettabili sono compresi tra $70ms$ e $200ms$, mentre per il parlato è preferibile un tempo baricentrico inferiore ai $60ms$.

La *spazialità* o *dimensione apparente della sorgente* si riferisce alla capacità di un ambiente di trasmettere l'impressione di una sorgente più grande delle sue dimensioni reali, grazie alle prime riflessioni che giungono dai lati. La spazialità

viene quantificata dai parametri LEF e $LEFC$, definiti nel seguente modo:

$$LEF = \frac{\int_{5ms}^{80ms} [g_{vy}(t)]^2 dt}{\int_0^\infty [g_p(t)]^2 dt} \quad (2.48)$$

$$LEFC = \frac{\int_{5ms}^{80ms} |g_{vy}(t) \cdot g_p(t)| dt}{\int_0^\infty [g_p(t)]^2 dt} \quad (2.49)$$

2.4.4 Considerazioni sulle incertezze

Vogliamo trattare brevemente la problematica delle incertezze presenti nei risultati che verranno presentati in questo lavoro. Innanzitutto bisogna precisare che non è ancora stata effettuata la taratura di tutta la catena in unità fisiche. Per quanto riguarda i segnali di pressione e velocità, possiamo naturalmente essere sicuri che il loro andamento temporale e le loro ampiezze relative rispecchino i valori reali, tuttavia non possiamo conoscerne i valori assoluti non avendo ancora effettuato una calibrazione in condizioni di campo note e mediante il confronto con un fonometro di riferimento calibrato. Questa operazione costituisce una delle priorità da affrontare dopo la conclusione del presente lavoro. Tuttavia, la cosa non risulta essenziale per le misurazioni e i risultati qui esposti (e per la maggiorparte delle misurazioni acustiche in generale), in quanto abbiamo preso in considerazione rapporti tra grandezze, ottenendo quindi risultati sotto forma di numeri puri.

Le incertezze possono essere dovute al sistema di misura e all'eventuale carattere stocastico del sistema da misurare. Quest'ultima possibilità in realtà è trascurabile nella maggior parte delle misurazioni acustiche, in quanto i fenomeni coinvolti nella propagazione del suono in un ambiente sono di tipo deterministico e non sono soggetti a variabilità temporale, a meno che non entrino in gioco improbabili variazioni significative in fattori quali la temperatura dell'ambiente, la densità e l'umidità dell'aria. A riprova di ciò è il fatto che ripetendo più volte una misura acustica si ottengono sempre gli stessi risultati.

La fonte maggiore di incertezza è dovuta alla presenza di apparecchiature con curve di risposta ignote all'utente o non basate su un'effettiva misura del particolare strumento da parte del costruttore (è il caso della sonda USP e della sorgente omnidirezionale). Per questo fatto ci siamo sentiti di riportare i risultati delle misure degli indicatori energetici limitandoci a due cifre significative. Per quanto riguarda i tempi di riverberazione della pressione sonora e dell'energia $RT30$ e $WT60$, essendo i valori ricavati da fitting lineare si potrebbe associare loro l'incertezza relativa al fitting, tuttavia questa risulterebbe talmente bassa da costringerci a usare un numero di cifre significative insensato per due parametri che hanno un significato percettivo e psicoacustico. Ci siamo pertanto limitati alla seconda cifra decimale, in sintonia con la maggiorparte dei lavori prodotti nell'ambito dell'acustica architettonica, ben consci del fatto che nessuno sia in grado di percepire differenze più piccole di un decimo di secondo nel tempo di riverberazione. Precisiamo inoltre che il valore di $WT60$, derivato dalla pendenza di una curva di decadimento, può essere falsato dalla presenza di rumore di fondo che causa una variazione di pendenza nella curva: la significatività del dato va pertanto valutata dal plot della curva.

Nel caso invece dei parametri D_{50} e C_{80} , espressi rispettivamente in valore per-

centuale e in dB , pur avendo a disposizione valori molto accurati⁵ ci siamo fermati rispettivamente all'intero e alla prima cifra decimale. Nei casi in cui l'indice D_{50} è sceso al di sotto del 10% abbiamo riportato anche la prima cifra decimale.

2.5 Misura della risposta quadrifonica in diversi ambienti

In questa sezione presentiamo i risultati ottenuti dalle misurazioni di risposta quadrifonica all'impulso con la tecnica dello sweep sinusoidale logaritmico. Le misure sono state effettuate in quattro ambienti dalle differenti caratteristiche acustiche: una piccola aula didattica, un ampio salone, un lungo corridoio e una cattedrale. In ogni ambiente le misure sono state eseguite collocando la sorgente in diverse posizioni. La scelta dei punti di posizionamento della sorgente e del ricevitore viene fatta prima di effettuare la misura, ed è dettata dagli scopi che la misura si prefigge. Molteplici sono i campi di applicazione della misura di risposta all'impulso e finalità diverse possono richiedere, in generale, configurazioni diverse. Per esempio, se lo scopo è quello di misurare l'intelligibilità del parlato in una sala per conferenze, sorgente e ricevitore verranno collocati in corrispondenza dei punti in cui si verranno a trovare l'oratore e gli ascoltatori durante l'effettivo utilizzo della sala; se si vogliono misurare le risonanze dovute alle onde stazionarie in una stanza parallelepipedica, è opportuno collocare sorgente e ricevitore in due angoli opposti della sala, dove è possibile intercettare contemporaneamente i ventri di pressione e i nodi di velocità di ogni modo proprio di oscillazione; se si vuole catturare la risposta all'impulso per virtualizzare la localizzazione di un attore in fase di post-produzione cinematografica, si terrà conto delle posizioni dell'attore e della camera nella scena, e così via.

Le misure nella Sala Carnelutti e nel Dormitorio del Buora, presentate nelle sezioni 2.5.1 e 2.5.2, sono state effettuate nel corso dell'*Audio Testbed* previsto nell'ambito del progetto europeo *IP Racine* [17] per lo sviluppo di nuove tecnologie per il cinema digitale. Lo scopo di queste due misurazioni è quello di testare la procedura di misurazione, fornire le risposte all'impulso per l'utilizzo in fase di post-produzione cinematografica e applicare il processo di convoluzione e transcodifica per fornire un audio realistico ad una scena virtuale.

La misura nell'auletta didattica ha lo scopo di studiarne le caratteristiche acustiche e fornire la risposta all'impulso da inserire in una libreria di risposte di ambienti diversi per l'utilizzo nella post-produzione cinematografica.

Le misure effettuate nella Cattedrale di San Marco sono state svolte nell'ambito del progetto *C.A.M.E.R.A.*, in collaborazione con l'università di Cambridge, per uno studio sulla relazione tra l'acustica delle chiese e la composizione e l'esecuzione di musica corale nel Rinascimento [18], [19].

2.5.1 La Sala Carnelutti

La Sala Carnelutti, appartenente alla Fondazione Cini di Venezia, ha una forma parallelepipedica e dimensioni di $33.9m \times 10.5m \times 7.3m$. Il pavimento è realizzato

⁵Questi valori, come quelli di *RT30* e *WT60*, non sono affetti da non linearità nella risposta in frequenza degli strumenti.

in marmo rosso, i muri sono intonacati ed è presente un controsoffitto fonoassorbente con pannelli forati di cartongesso e traverse disposte a intervalli regolari su tutta la lunghezza, come è possibile vedere nella figura 2.17.

Sono presenti quattro porte, ciascuna coperta da una tenda pesante. Altre dodici tende pesanti sono collocate sui lati lunghi della stanza, sei per parete. Al momento della misura, nella stanza era presente solo un pianoforte a coda e l'apparecchiatura di misura.

I muri ed il pavimento favoriscono la riflessione del suono; proprio per questo motivo sono stati installati i tendaggi ed il controsoffitto che, essendo fonoassorbenti, permettono di contenere il tempo di riverberazione.

Nella sala sono state effettuate otto misurazioni, tenendo fissa la sonda e spostando la sorgente. La figura 2.18 riporta le posizioni scelte. La sonda è stata orientata con gli assi x , y e z rispettivamente nelle direzioni della lunghezza, della larghezza e dell'altezza. I punti 1, 2 e 3 si trovano in direzione frontale rispettivamente a $2m$, $4m$ e $6m$ dalla sonda. Il punto 4 si trova $4m$ dietro la sonda e i punti 5, 6, 7 e 8 si trovano su una circonferenza di raggio $4m$ ad angoli di -30° , -110° , -90° e $+30^\circ$ rispettivamente rispetto alla direzione dell'asse x . La sorgente e la sonda sono state poste ad un'altezza di $1.5m$.



Figura 2.17: La Sala Carnelutti al momento della misura: sono visibili la sorgente omnidirezionale e la sonda intensimetrica.

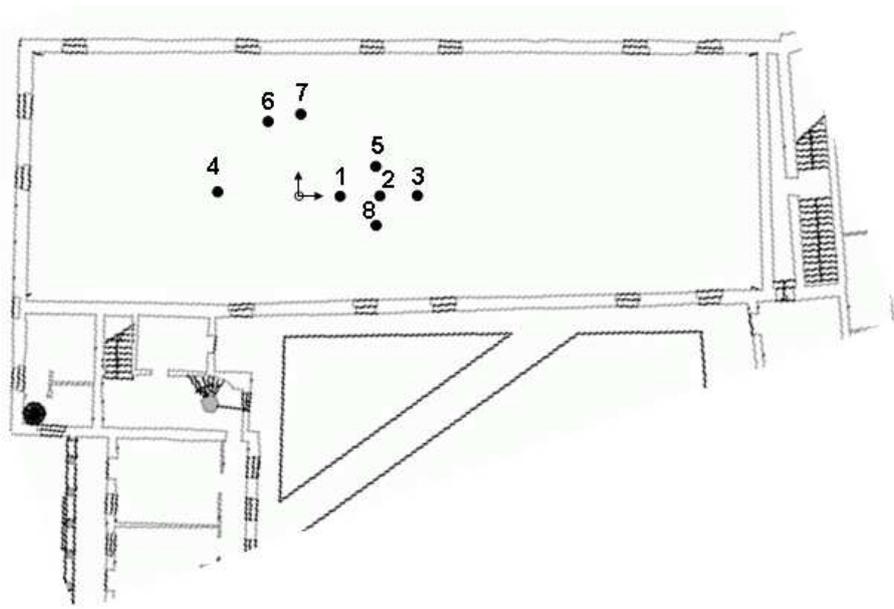


Figura 2.18: Pianta della Sala Carnelutti con posizioni sorgente-ricevitore.

Lo scopo di queste misurazioni è di testare il sistema messo a punto per la ricostruzione virtuale di un evento sonoro e la sua riproduzione. A partire dalle risposte misurate, verrà poi eseguita la convoluzione con una registrazione anecoica, sintetizzando così le componenti di pressione e velocità che avremmo misurato se il suono fosse stato realmente prodotto nell'ambiente, con la sorgente nella posizione dell'altoparlante. I punti di misura 1, 2 e 3 servono a testare la bontà della ricostruzione della distanza apparente della sorgente. Gli altri punti permettono di valutare la capacità di ricostruire l'immagine sonora attorno all'ascoltatore. Alcuni dei risultati ottenuti verranno discussi nel capitolo 4.

La figura 2.19 riporta l'andamento temporale dei segnali di risposta all'impulso di pressione e velocità misurati con sorgente nel punto 1. Notiamo innanzitutto che il segnale diretto è presente solamente nelle componenti p e v_x , in quanto l'onda diretta produce una velocità orientata lungo l'asse x , quindi non c'è movimento delle particelle d'aria nelle direzioni y e z . Nella componente v_z il primo segnale che compare è la riflessione dal pavimento, assente nella componente y . Successivamente all'arrivo del suono diretto, si notano i contributi delle prime riflessioni, seguite dalla coda riverberante. La figura 2.20 mostra il plot polare del vettore intensità istantanea sui tre piani definiti dagli assi cartesiani. I grafici riportano le proiezioni del vettore sui rispettivi piani nei primi $80ms$. Tale intervallo di tempo è scelto poichè contiene il suono diretto e le prime riflessioni; dal plot del vettore intensità possiamo quindi visualizzare la distribuzione e le direzioni nelle quali fluisce l'energia sonora impulsiva, che determinano la percezione della direzione della sorgente e della spazialità dell'ambiente [20]. Osservando il plot nel piano XY si nota che la direzione del vettore intensità è la stessa in cui è posta la sorgente, ma ha verso opposto, ad indicare che c'è un flusso di intensità dalla sorgente attraverso il punto di misura. La figura 2.21 mostra lo Schroeder plot 2.42 delle componenti dell'energia

sonora, precisamente l'energia potenziale w_p e l'energia cinetica w_K data dalle singole componenti della velocità (si vedano, in proposito, le definizioni 1.25, 1.26, 1.27). Notiamo che l'andamento delle componenti cinetiche, inizialmente esponenziale decrescente (l'asse y ha una scala logaritmica), subisce un cambiamento di pendenza dovuto alla presenza di rumore di fondo dell'apparecchiatura (in questo caso della sonda) nei canali di velocità. Filtrando la risposta all'impulso in bande di ottava secondo le norme ANSI (*American National Standards Institute*), si ricava il valore di $WT60$ in ogni banda, riportato nel grafico di figura 2.22. La tabella 2.1 riporta i valori degli indicatori energetici η , ξ e σ , assieme al tempo di decadimento dell'energia, al tempo di riverberazione e ai parametri C_{80} e D_{50} . Oltre al valore globale vengono riportati i valori in bande di ottava. Osserviamo che il valore di σ prossimo a 1 indica l'equipartizione dell'energia in energia cinetica e potenziale. L'indicatore η indica che il 40% dell'energia sonora viene irradiata, mentre 60% rimane localizzato attorno al punto di misura. Agli estremi dello spettro si nota il predominare di energia localizzata.

Si osserva nelle prime due ottave e nel valore overall una discrepanza tra il tempo di riverberazione $RT60$ ed il tempo di decadimento dell'energia $WT60$: ciò è dovuto alla presenza, al momento della misura, di una componente di rumore di fondo nei segnali di velocità che finisce per predominare sul decadimento del segnale. Il tempo di riverberazione, calcolato a partire dal solo segnale di pressione, non ne risulta influenzato.

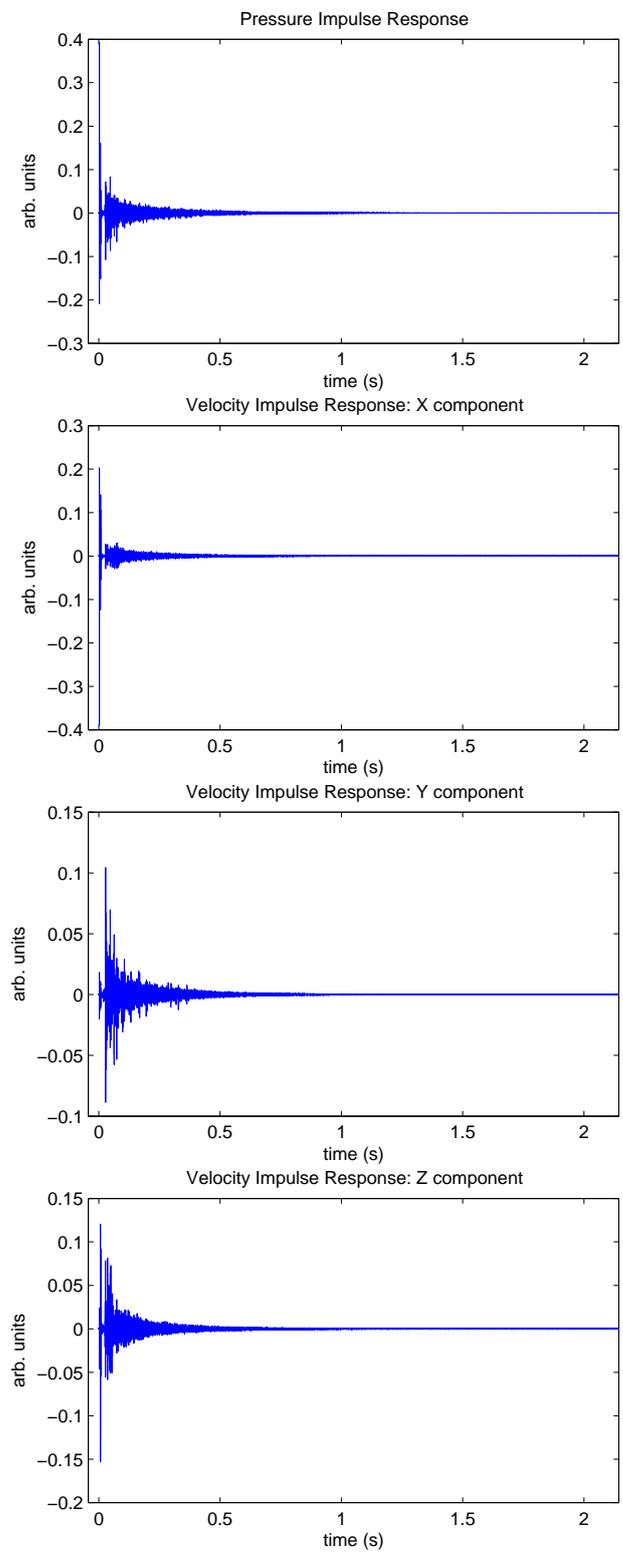


Figura 2.19: Risposta all'impulso con sorgente nel punto 1.

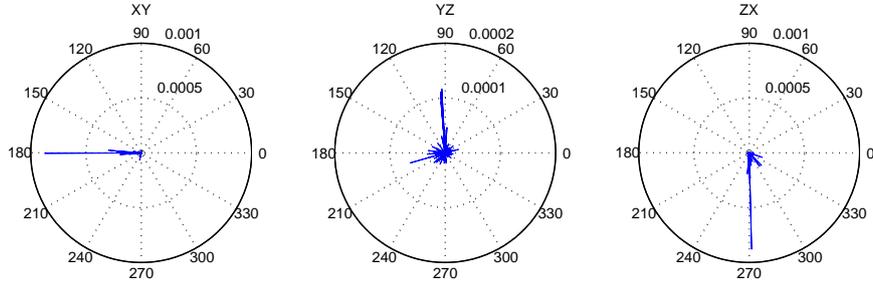


Figura 2.20: Polar Plot del vettore intensità misurato nella posizione 1 nei primi 80ms.

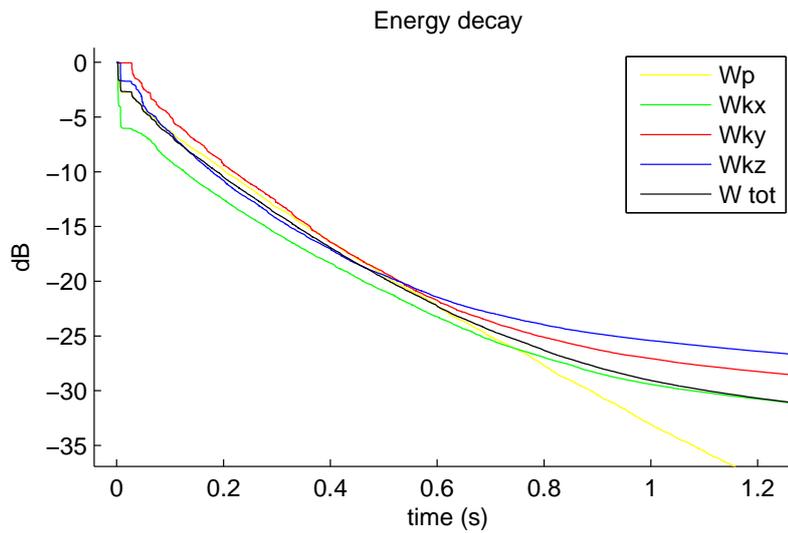


Figura 2.21: Decadimento temporale delle componenti dell'energia.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.42	0.34	0.29	0.60	0.58	0.50	0.50	0.37	0.39
ξ	0.42	0.36	0.30	0.61	0.59	0.51	0.50	0.37	0.40
σ	0.99	0.96	0.97	0.99	0.98	0.98	0.99	0.99	0.96
$WT60$ (s)	4.21	4.94	2.90	1.35	1.81	2.30	2.27	1.87	1.35
$RT30$ (s)	2.05	2.33	2.26	1.32	1.77	2.30	2.27	1.88	1.31
C_{80} (dB)	4.3	-3.1	-0.6	5.6	5.7	3.9	3.7	3.0	5.8
D_{50} (%)	64	30	38	75	72	62	65	56	66

Tabella 2.1: Indicatori energetici e parametri oggettivi calcolati dalla risposta all'impulso misurata nella Sala Carnelutti nella posizione 1.

La figura 2.23 riporta i valori in bande di ottava degli indicatori energetici η e σ .

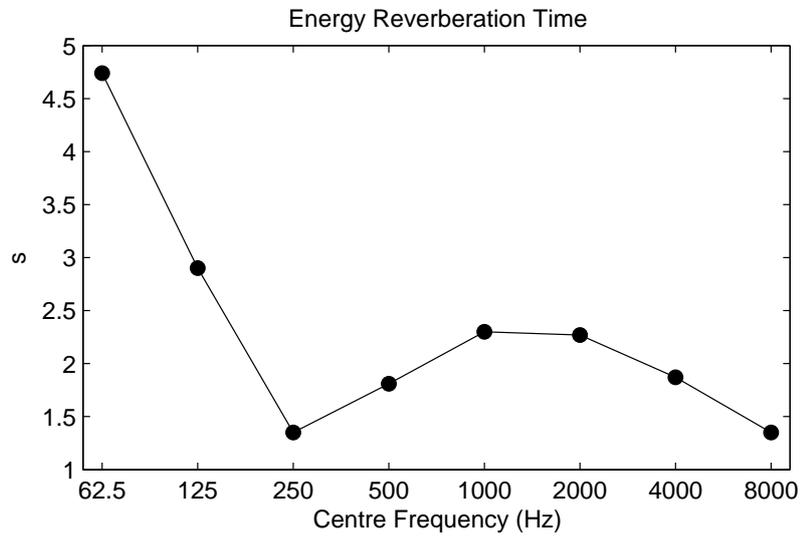


Figura 2.22: Tempo di decadimento dell'energia in bande di ottava.

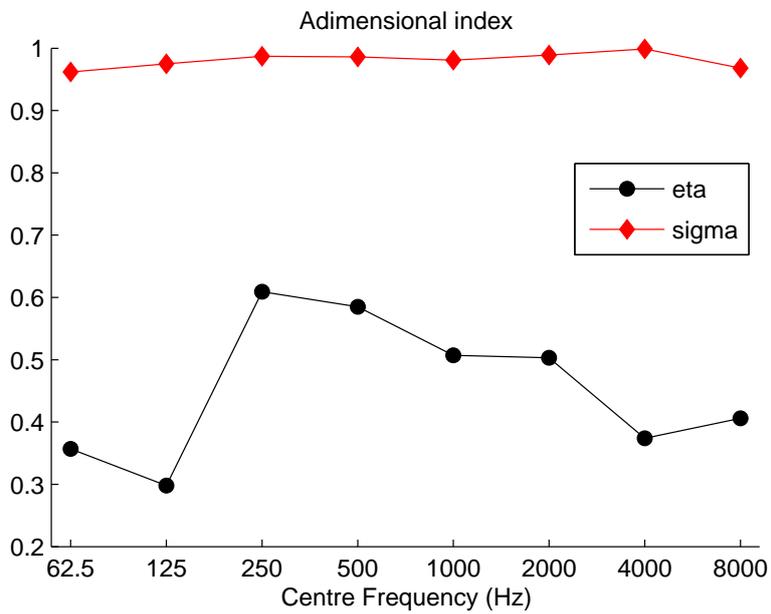


Figura 2.23: Andamento degli indicatori σ ed η in bande di ottava.

2.5.2 Il Dormitorio del Buora

Il dormitorio del Buora (figura 2.24) si trova all'interno dell'abbazia dei monaci benedettini dell'isola di San Giorgio Maggiore a Venezia. Il corridoio è lungo 128 metri e largo 6 e presenta un soffitto a volta che raggiunge l'altezza di 8.3 metri. Lungo le pareti laterali ci sono 42 porte di accesso alle celle monastiche. Il pavimento è di marmo e i muri sono in mattone intonacato. Delle ampie vetrate agli estremi del lungo corridoio, assieme a piccole finestre alla base della volta, assicurano la sufficiente illuminazione dell'ambiente.

In questo ambiente sono state effettuate quattro misure di risposta all'impulso. La sonda è stata posta al centro del corridoio, con l'asse x orientato nel senso della lunghezza. Come mostrato in figura 2.25, la sorgente è stata collocata in tre punti davanti alla sonda, alla distanza di 2, 4 e 6 metri e in un punto dietro alla sonda, a 4 metri di distanza.



Figura 2.24: Foto del Dormitorio del Buora ripresa dal punto centrale.

I valori degli indicatori misurati nella posizione 1 sono riportati nella tabella 2.2.

La figura 2.26 mostra l'andamento temporale dei segnali di pressione e di velocità della risposta all'impulso misurata nella posizione 1. Osserviamo la presenza del segnale diretto nelle componenti p e v_x , seguito dalle riflessioni e dalla coda riverberante, meno intense ma più prolungate rispetto al caso della sala Carnelutti. La componente v_z presenta come segnale più intenso la riflessione dal pavimento, che segue il segnale diretto appena identificabile. I polar plot di figura 2.27 indicano che nei primi $80ms$ l'energia è essenzialmente concentrata lungo la direzione dell'asse x . Nella figura 2.28 osserviamo che la curva di decadimento dell'energia potenziale e di quella cinetica dovuta alle componenti

x e z ha un comportamento irregolare nei primi istanti, dovuto alla notevole differenza di livello fra il suono diretto e le riflessioni successive, che causa la rapida diminuzione della funzione di Schroeder subito dopo l'arrivo del segnale diretto. Questo non si riscontra nella componente y , dove il segnale diretto è meno intenso delle riflessioni successive.

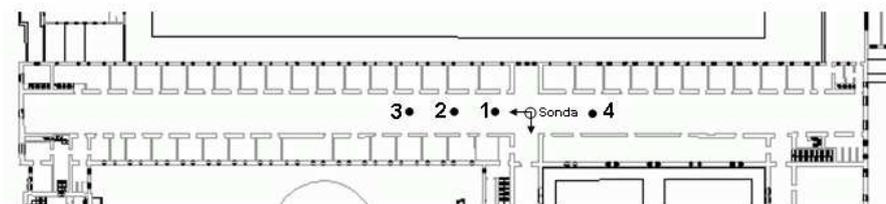


Figura 2.25: Pianta del Dormitorio con posizioni di misura.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.64	0.59	0.25	0.36	0.46	0.48	0.44	0.63	0.72
ξ	0.64	0.62	0.27	0.40	0.46	0.49	0.45	0.63	0.73
σ	0.99	0.94	0.90	0.99	0.99	0.98	0.99	1	0.99
$WT60$ (s)	5.35	3.95	5.34	5.84	6.01	5.95	5.2	3.86	2.24
$RT30$ (s)	4.82	3.47	4.66	4.79	5.26	4.32	4.71	3.89	2.21
$C80$ (dB)	5.7	4.2	-0.6	0.2	0.1	1.9	0.8	4.6	9.9
$D50$ (%)	40	40	16	41	46	53	46	68	86

Tabella 2.2: Indicatori energetici e parametri acustici misurati nella posizione 1.

Dal plot dell'indicatore η in bande di ottava di figura 2.30 osserviamo come l'indice di radiazione in questo caso assuma i valori più alti agli estremi della banda udibile, ad indicare che il trasporto di energia viene favorito alle alte e alle basse frequenze, mentre nelle frequenze medio basse l'energia sonora è prevalentemente intrappolata nei modi propri dell'ambiente. Questo è il contrario di quanto abbiamo ottenuto dalle misure nella Sala Carnelutti. Si nota inoltre che l'indicatore di radiazione η e il tempo di decadimento dell'energia $WT60$, mostrato in figura 2.29, seguono andamenti opposti: alla crescita dell'uno corrisponde la diminuzione dell'altro. Ciò si interpreta alla luce del fatto che il contributo alla riverberazione deriva dalla parte di energia che oscilla avanti e indietro nell'ambiente attraverso il punto di misura, mentre il contributo ad η deriva da quella parte di energia che viene irradiata dal campo attraverso il punto di misura. A riprova di ciò è il fatto che ambienti molto assorbenti, cui corrisponde un tempo di riverberazione basso, siano caratterizzati da un indice di radiazione prossimo a uno. Il tempo di riverberazione calcolato nella posizione 1 si attesta attorno ai cinque secondi, tuttavia la distanza di soli due metri dalla sorgente può portare ad una stima errata. La tabella 2.3 riporta i valori calcolati dalla risposta all'impulso misurata nella posizione 3. In questo caso vediamo che il tempo di riverberazione risulta pari a circa 5.5s e si avvicina ai sette secondi nell'ottava centrata sui 500Hz. Il tempo di riverberazione è alto

a causa dell'assenza di materiali in grado di fornire assorbimento. Alle alte frequenze si ha una diminuzione di $RT30$ dovuta all'assorbimento da parte dell'aria stessa che, data la lunghezza del corridoio, può dare effetti significativi.

Vogliamo infine presentare il confronto tra l'indice di chiarezza C_{80} valutato nella posizione 1 e nella posizione 3. Ricordiamo che esso esprime il rapporto tra l'energia potenziale contenuta nei primi $80ms$ e quella che arriva negli istanti successivi. Come si vede nella figura 2.31, alla distanza di due metri predominano ancora il segnale diretto e le prime riflessioni, seppur di poco, mentre a sei metri è la riverberazione a prevalere. Questo conferma il fatto che nell'ambiente sia difficile sostenere una comunicazione con un interlocutore distante più di pochi metri, come abbiamo avuto modo di sperimentare direttamente durante lo svolgimento della misura.

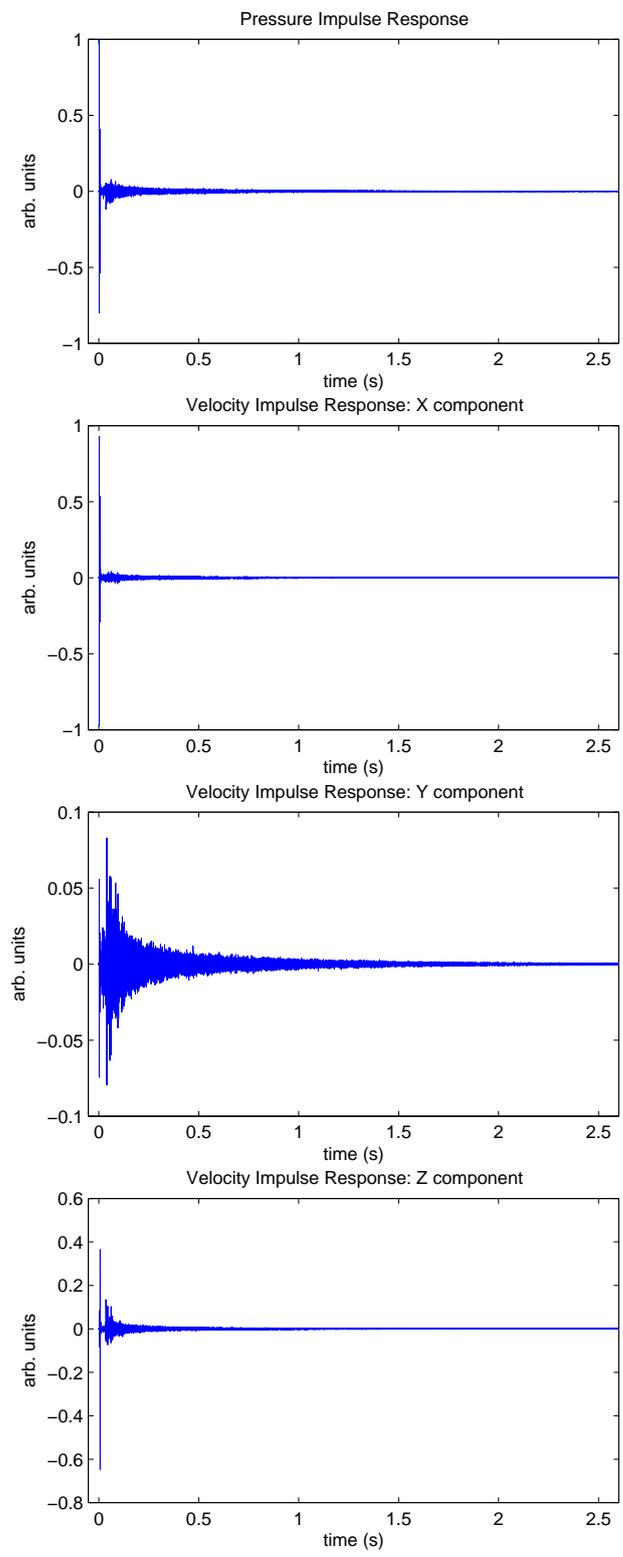


Figura 2.26: Risposta all'impulso misurata nella posizione 1.

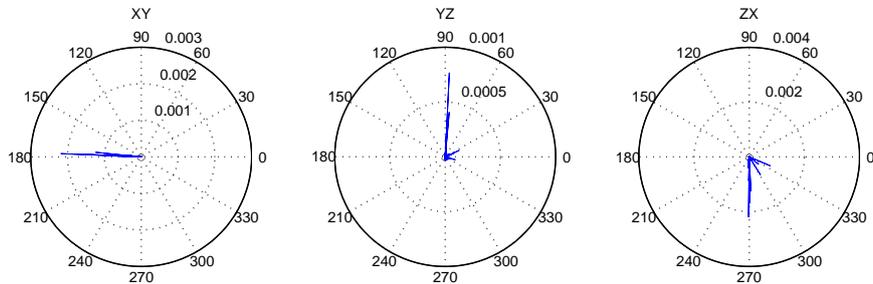


Figura 2.27: Polar plot del vettore intensità nei primi 80ms.

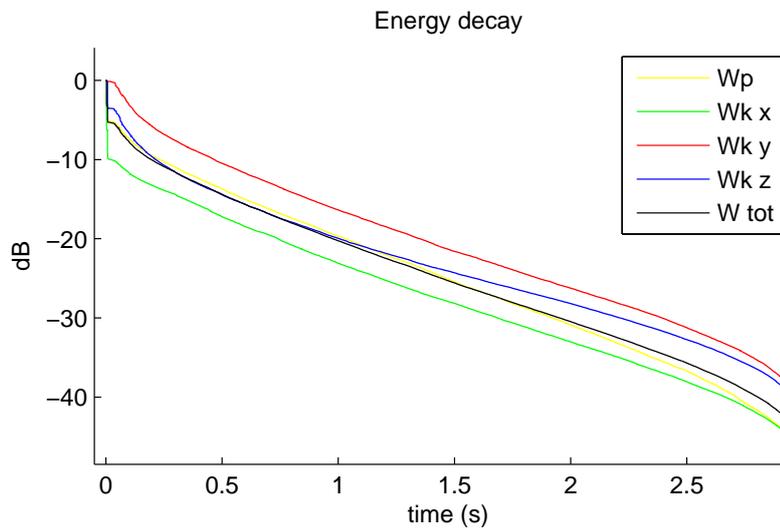


Figura 2.28: Curve di decadimento della densità di energia relative alla posizione 1.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.40	0.50	0.30	0.15	0.26	0.30	0.27	0.38	0.44
ξ	0.41	0.51	0.30	0.16	0.26	0.30	0.27	0.38	0.45
σ	0.99	0.98	0.97	0.99	0.99	0.99	0.99	1	0.97
$WT60$ (s)	6.45	16.1	7.03	6.72	6.89	6.69	5.46	4.00	4.07
$RT30$ (s)	5.48	3.97	5.67	6.57	6.89	6.67	5.46	3.98	2.44
$C80$ (dB)	0.2	-3.6	-1.6	-4.2	-3.9	-3.0	-4.3	-0.9	3.3
$D50$ (%)	43	16	26	15	23	26	20	35	58

Tabella 2.3: Indicatori energetici e parametri acustici misurati nella posizione 3.

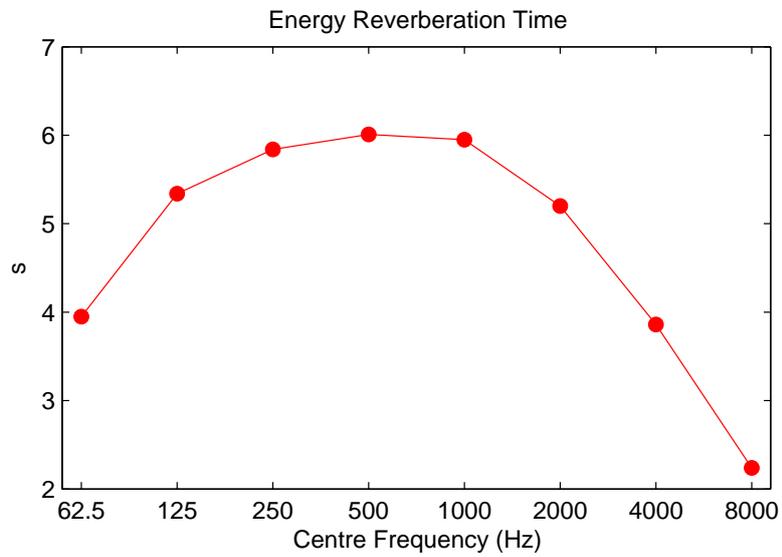


Figura 2.29: Decadimento temporale dell'energia sonora in bande di ottava.

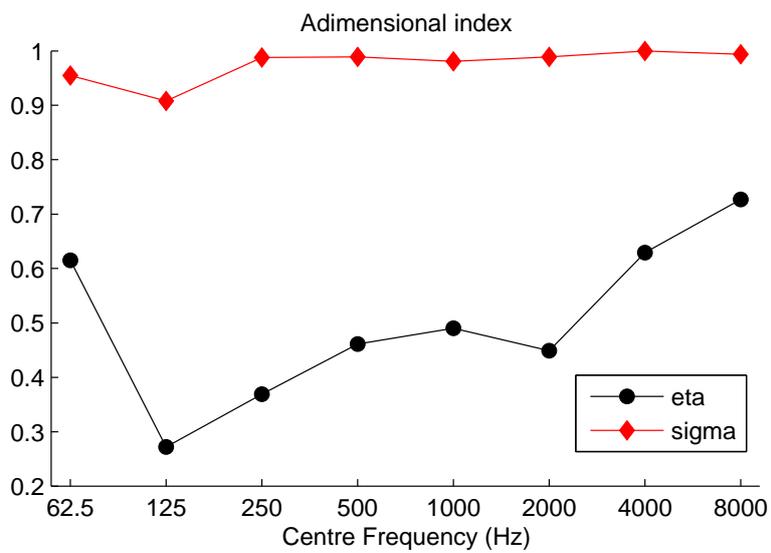


Figura 2.30: Indicatori η e σ in bande di ottava.

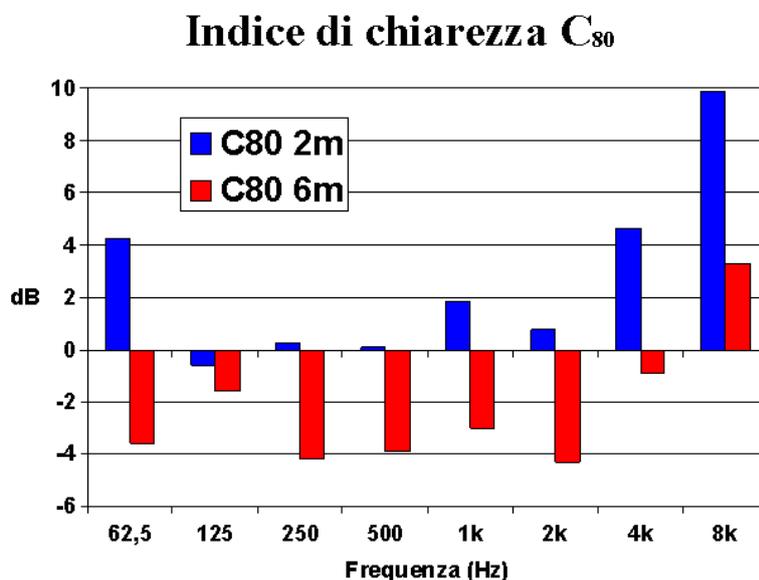


Figura 2.31: Confronto tra i valori di C_{80} nelle posizioni 1 e 3.

2.5.3 L'auletta didattica FSSG

L'auletta didattica è una piccola stanza parallelepipedica di dimensioni $5.20m \times 4.40m \times 3m$, adiacente al Laboratorio di Acustica Musicale e Architettonica FSSG-CNR a Venezia. Il pavimento è rivestito di marmo e la parete sul lato della finestra è realizzata in muratura ricoperta da intonaco, così come il soffitto. La parete frontale consiste in un armadio a muro con ante in legno, che ne copre tutta la larghezza. Le altre due pareti sono realizzate in legno nella parte inferiore, fino ad una altezza di un metro, e sono costituite da vetrate nella parte superiore. Nella stanza sono presenti una scrivania, una dozzina di sedie di plastica ed un mobiletto per attrezzature audio-video. La figura 2.32 offre uno scorcio dell'ambiente, in cui è possibile vedere anche il setup sperimentale in una delle posizioni di misura.

In questo ambiente sono state effettuate due misurazioni, secondo lo schema illustrato in figura 2.33. La sorgente è stata tenuta fissa dietro la cattedra, nella posizione occupata dall'oratore durante le lezioni didattiche, ad un'altezza di $1.5m$ dal pavimento. La sonda, anch'essa posta ad un'altezza di $1.5m$, è stata orientata con l'asse x lungo il lato corto della stanza, con il verso positivo rivolto verso la sorgente. La prima posizione vede la sonda di fronte alla sorgente, ad una distanza di $2m$, mentre nella seconda misura la sonda è stata posta nell'angolo tra le due pareti con vetrata.

La risposta ottenuta nella posizione M1 è mostrata in figura 2.34. Le figure 2.35 e 2.36 riportano il plot del vettore intensità istantanea per i primi $80ms$ rispettivamente nelle posizioni M1 ed M2. Rispetto agli ambienti precedenti, le ridotte dimensioni di questa stanza fanno sì che le prime riflessioni e il campo

riverberato arrivino al punto di misura con minor ritardo rispetto a quello che si verifica negli ambienti più grandi, pertanto esse costituiscono una parte significativa dei dati mostrati nel polar plot. Dal confronto tra le due posizioni di misura, vediamo come le riflessioni siano maggiormente evidenti nella misura effettuata nell'angolo, dove il segnale diretto è meno intenso rispetto a quello misurato in posizione M1. I valori ottenuti per gli indicatori sono riportati in tabella 2.4. Dalla figura 2.38 notiamo come il valore dell'indicatore η sia inferiore al 25% in tutte le bande, ad indicare il fatto che meno di un quarto dell'energia viene effettivamente irradiata.

Il tempo di riverberazione prossimo al secondo e i valori degli indici di chiarezza e definizione permettono una buona intelligibilità del parlato, rendendo l'ambiente adatto a lezioni didattiche.



Figura 2.32: La misura nella auletta didattica.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.23	0.20	0.27	0.12	0.17	0.12	0.13	0.21	0.26
ξ	0.23	0.23	0.30	0.12	0.17	0.12	0.13	0.21	0.27
σ	0.98	0.88	0.93	0.99	0.99	0.99	1	1	0.96
$WT60$ (s)	0.89	0.89	1.09	1.06	0.99	0.99	0.91	0.87	0.69
$RT30$ (s)	0.87	0.91	1.14	1.1	1.0	0.99	0.90	0.88	0.68
$C80$ (dB)	6.2	3.0	4.3	2.2	2.9	3.8	3.9	5.3	7.6
$D50$ (%)	64	16	47	49	44	55	48	60	71

Tabella 2.4: Indicatori energetici e parametri acustici calcolati nella posizione M1.

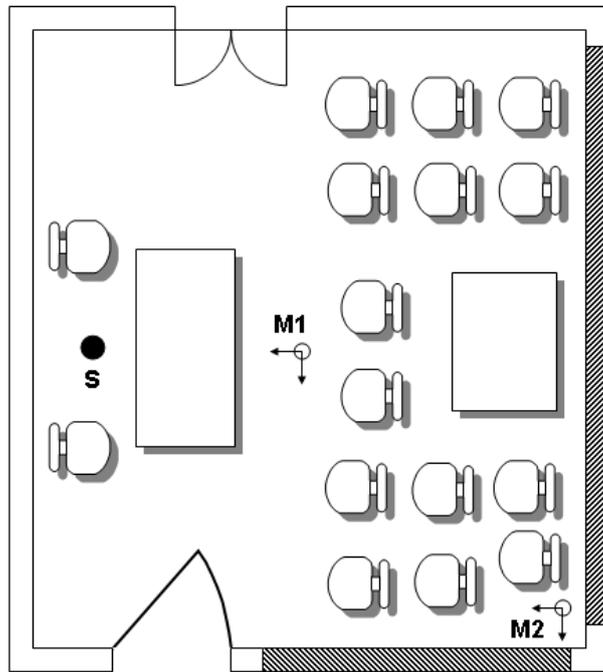


Figura 2.33: Pianta dell'auletta didattica con posizioni di sorgente e ricevitore.

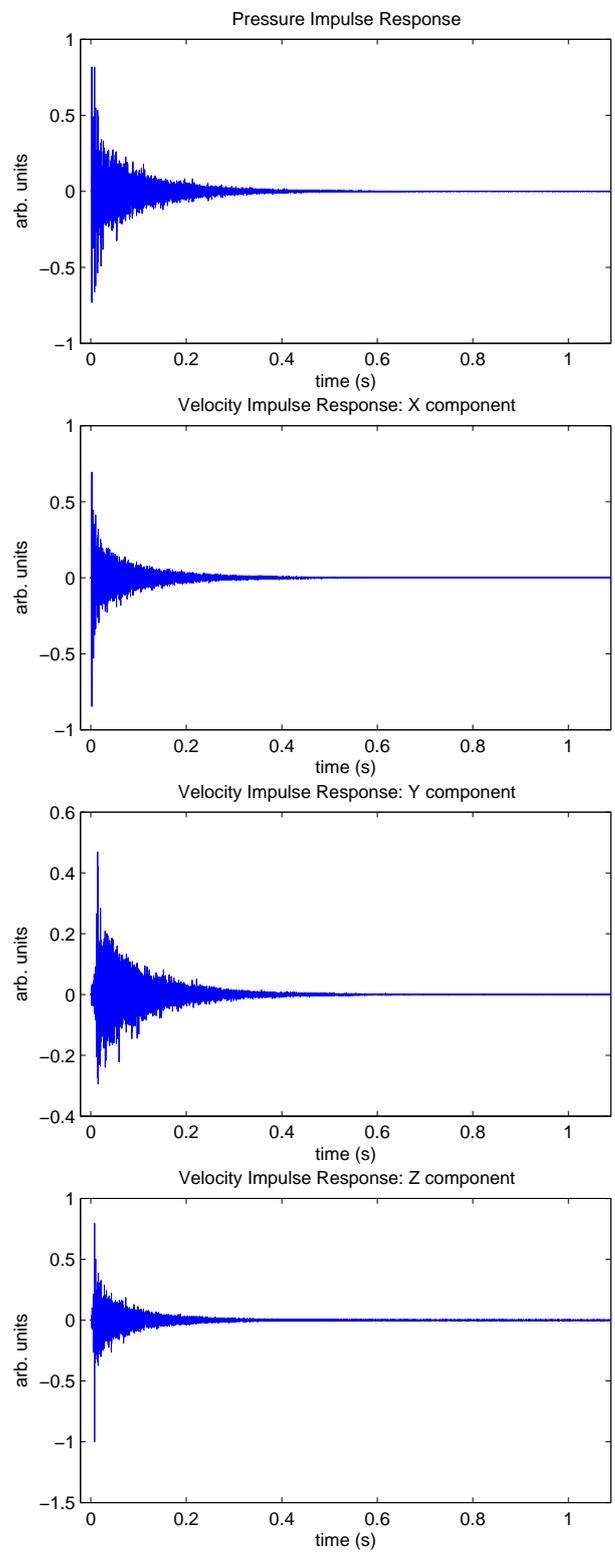


Figura 2.34: Risposta all'impulso misurata in posizione M1.

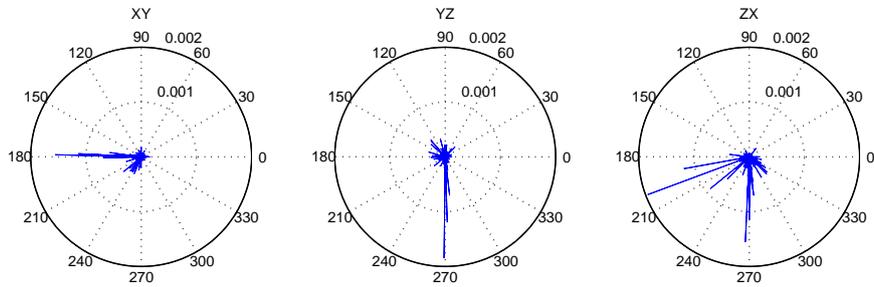


Figura 2.35: Polar plot del vettore intensità misurato in posizione M1 nei primi 80ms.

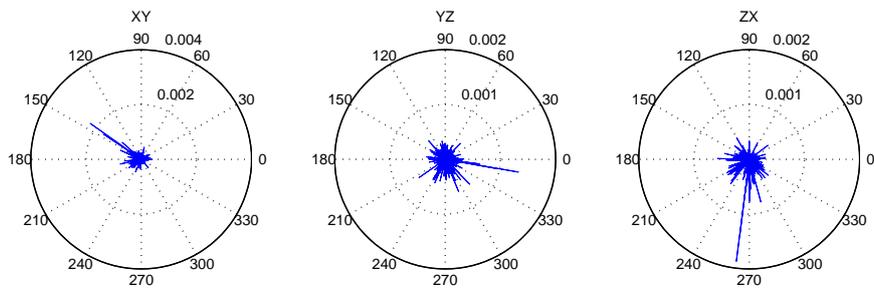


Figura 2.36: Polar plot del vettore intensità misurato in posizione M2 nei primi 80ms.

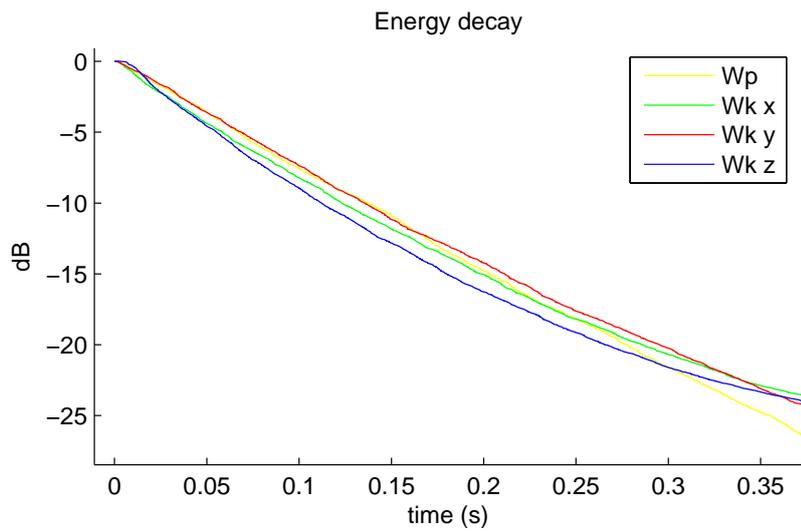


Figura 2.37: Decadimento della densità di energia nella posizione M1.

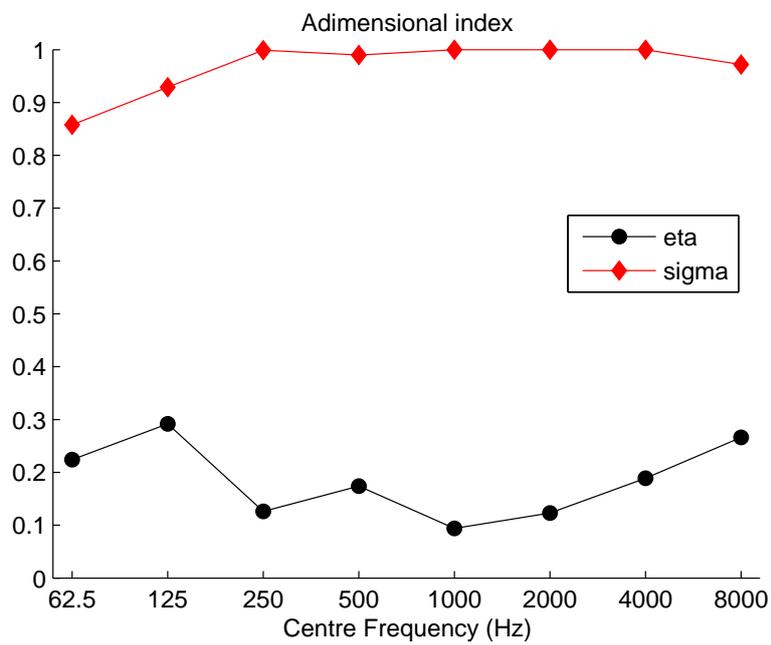


Figura 2.38: Indicatori η e σ in posizione M1 in bande di ottava.

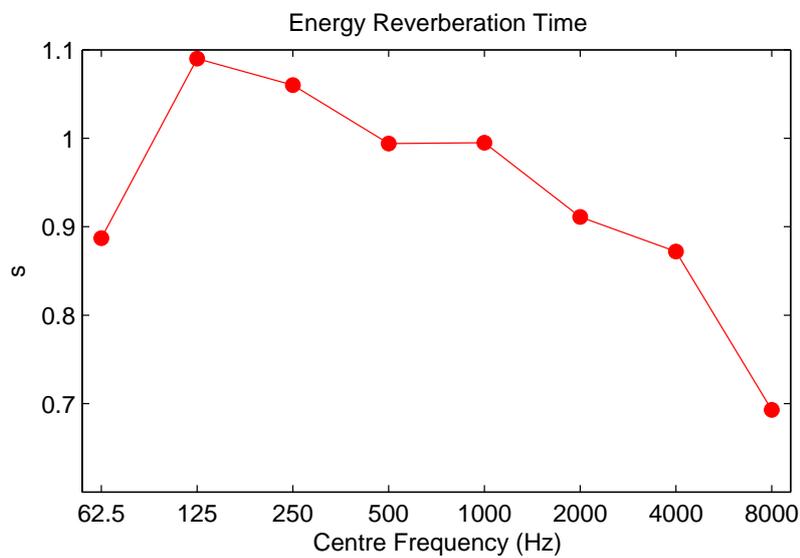


Figura 2.39: Parametro $WT60$ in posizione M1 in bande di ottava.

2.5.4 La Basilica di San Marco a Venezia

La Basilica di San Marco a Venezia è la chiesa più famosa del capoluogo veneto ed il più noto esempio di architettura bizantina in Italia. Dal 1807 è la sede del Patriarca di Venezia. La basilica si affaccia su piazza San Marco ed è adiacente e collegata al palazzo ducale.

La base dell'attuale basilica risale all'anno 1063, ma il suo attuale aspetto è frutto degli interventi compiuti tra l'undicesimo e la fine del quattordicesimo secolo, che riflettono le varie epoche artistiche cui è stata soggetta nel corso dei secoli. La pianta della basilica, illustrata in figura 2.40, è a croce greca con cinque strutture autonome, coperte da cupola, distribuite lungo gli assi della croce e raccordate da brevi volte a botte.

Le misurazioni acustiche sono state effettuate nell'ambito del progetto *C.A.M.E.R.A.* [18], con lo scopo di fornire dati sperimentali per lo studio dell'acustica delle chiese di Venezia in relazione alla composizione e all'esecuzione di musica corale nel sedicesimo secolo. Le posizioni di misura, illustrate nella pianta di figura 2.40 e nella sezione di figura 2.41, rispecchiano proprio le posizioni in cui tradizionalmente erano collocate le sorgenti sonore (il coro e gli strumenti) e i principali punti d'ascolto. In riferimento alle figure, le posizioni in cui è stata collocata la sorgente sonora sono:

- A: *Pulpitum magnum cantorum*, occupato dal coro durante messe e vesperi
- B: Presbiterio, nel quale i chierici eseguivano il *Cantus Firmus*
- C: Pergolo sansoviniano, occupato dal coro durante i vesperi
- D: Galleria dell'organo, dove assieme all'organo potevano essere collocati i musicisti dell'orchestra e il coro.

La sonda è stata posizionata nei seguenti punti:

1. *Pulpitum magnum cantorum*, dove sedeva il Doge
2. Interno dell'iconostasi destra, in corrispondenza del seggio ducale
3. Coro maggiore, in cui sedevano ambasciatori e membri del governo
4. Presbiterio, occupato dai chierici
5. Navata, occupata dai fedeli

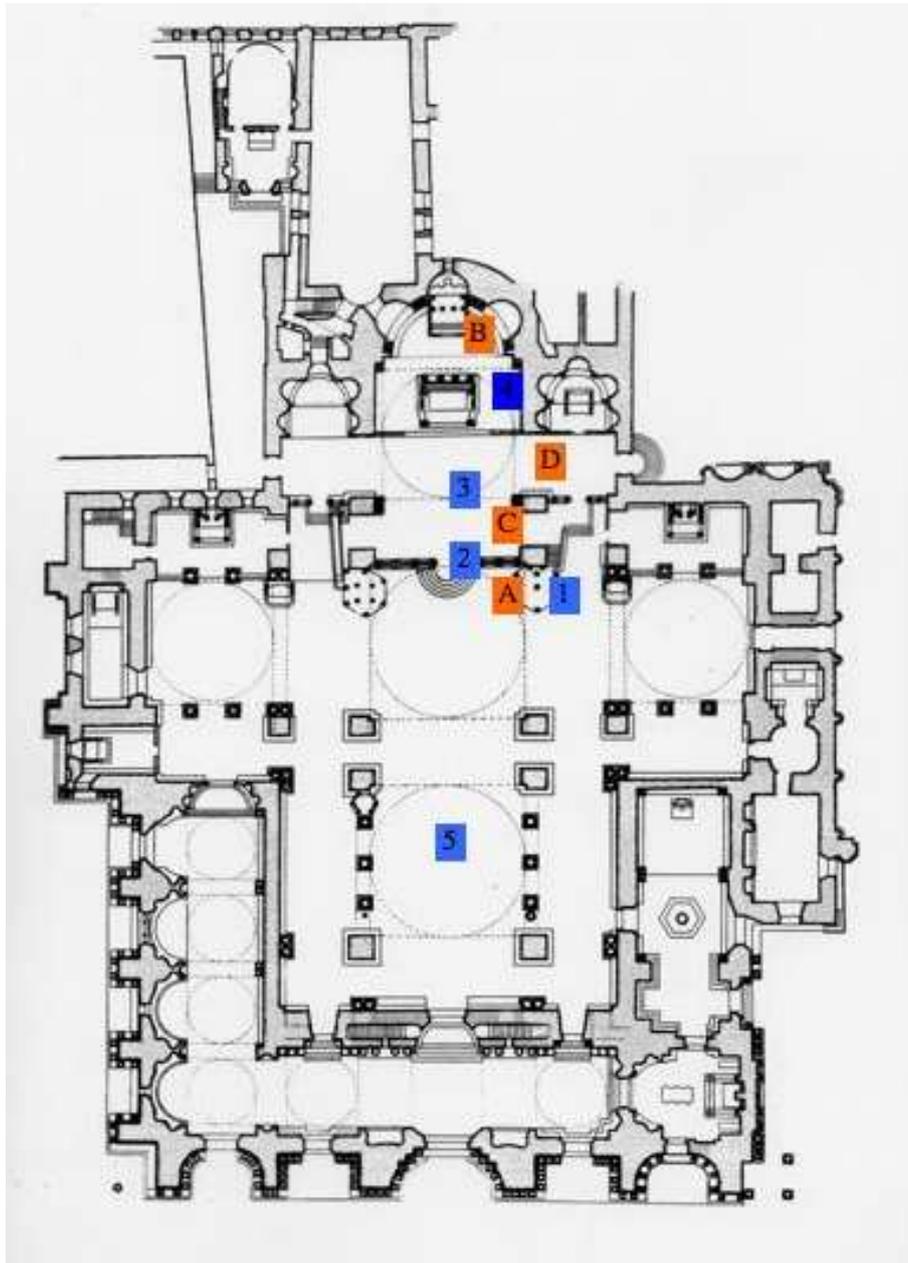


Figura 2.40: Pianta della Basilica di San Marco con posizioni di misura.

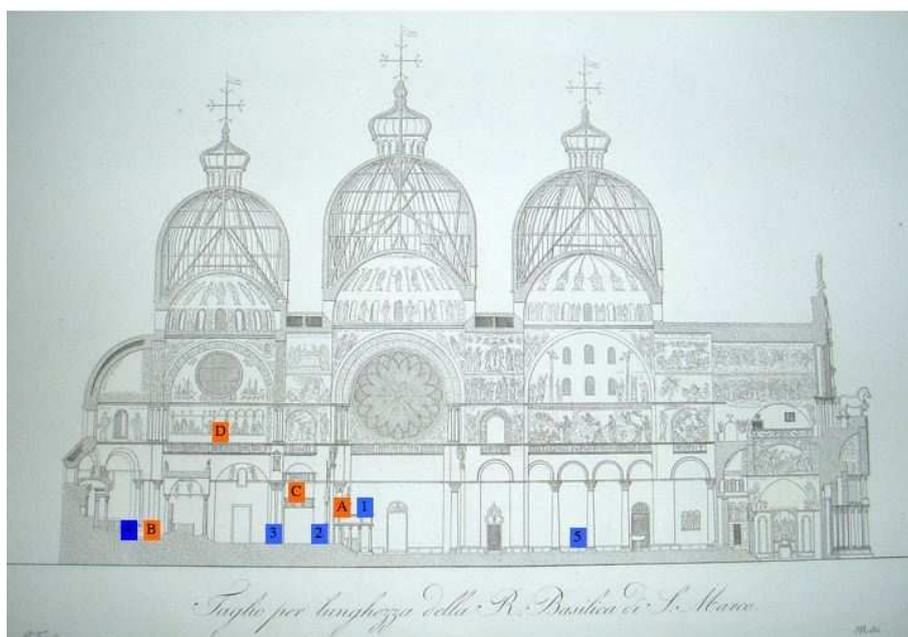


Figura 2.41: Sezione della Basilica di San Marco con posizioni di misura.

In questo caso la misura è stata effettuata nell'orario di apertura ai turisti, in presenza di diverse persone nell'ambiente. Oltre ai rumori provocati dalla presenza del pubblico, va precisato che, rispetto agli altri ambienti, le distanze medie tra sorgente e ricevitore sono di un ordine di grandezza superiori, essendo qui dell'ordine di poche decine di metri. Questi fattori, uniti all'impossibilità di usare elevati livelli di emissione del segnale di eccitazione per rispetto del luogo e per non recare eccessivo disturbo ai visitatori, si ripercuotono in un significativo peggioramento del rapporto segnale/rumore rispetto ai casi precedenti. La durata dello sweep è stata incrementata a un minuto, rispetto ai 30s utilizzati negli altri ambienti. Nonostante ciò le misure di tempo di decadimento dell'energia basate sull'estrapolazione dell'andamento della curva di Schroeder, come possiamo vedere dalla figura 2.42, risultano di difficile lettura. Per tale motivo i valori di $WT60$ sono stati omessi dalle tabelle.

La figura 2.43 riporta la risposta quadrifonica all'impulso misurata con sorgente in posizione A e ricevitore in posizione 5.

Tra i dati raccolti, presentiamo nelle tabelle 2.5 e 2.6 i risultati relativi alle posizioni A5 e B5. Nella posizione A5 la sorgente è visibile dalla posizione della sonda, mentre in posizione B5 la sorgente è nascosta dall'altare, oltre a trovarsi ad una distanza quasi doppia. Questo si ripercuote nelle diverse proporzioni tra suono diretto, prime riflessioni e campo riverberante. Vogliamo in particolare far notare come i dati relativi al tempo di riverberazione $RT30$ siano consistenti, essendo questo parametro indipendente dalla posizione, almeno finché non ci si trova nelle vicinanze della sorgente. A cambiare significativamente sono i valori degli indici di chiarezza e di definizione, che si abbassano all'aumentare della distanza dal punto di misura. In particolare, l'indice di chiarezza valutato

nella posizione B5 mostra che l'intelligibilità di una sorgente sonora in queste condizioni è seriamente compromessa.

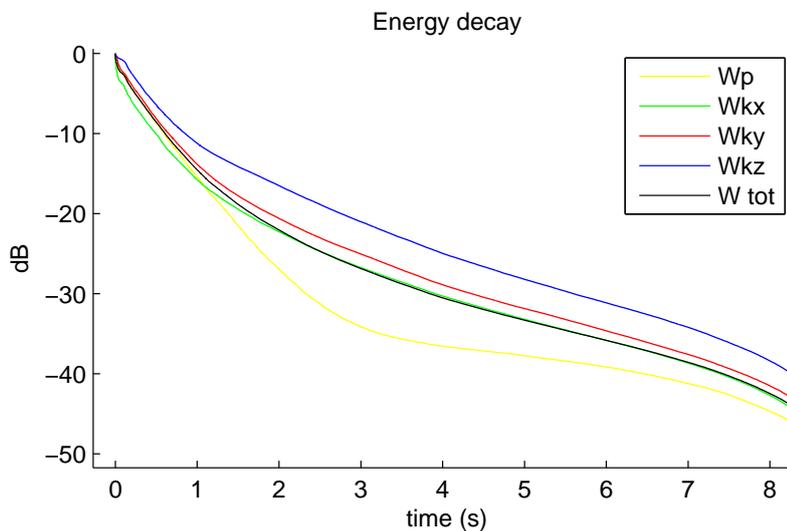


Figura 2.42: Decadimento delle componenti dell'energia potenziale e cinetica misurate in posizione A5. Si osserva la presenza di diverse pendenze nel decadimento energetico, in particolare nelle componenti dell'energia cinetica, influenzate dal rumore nei segnali di velocità.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.3	0.17	0.14	0.08	0.12	0.21	0.14	0.30	0.49
σ	1	0.99	0.99	1	1	0.99	1	1	1
$RT30 (s)$	4.89	5.8	5.6	5.6	5.9	5.8	4.5	3.0	1.6
$C80 (dB)$	-1.1	-5.6	-8.0	-10	-6.0	-3.8	-5.0	-1.1	4.5
$D50 (\%)$	39	17	11	5.7	18	26	20	38	69

Tabella 2.5: Indicatori energetici e parametri acustici calcolati nella posizione A5.

	Overall	62.5Hz	125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
η	0.17	0.21	0.27	0.06	0.08	0.13	0.12	0.23	0.36
σ	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1
$RT30 (s)$	4.9	5.5	5.4	5.6	5.8	5.4	4.5	3.1	1.9
$C80 (dB)$	-9.8	-9.8	-10	-14	-14	-13	-11	-9.7	-4.1
$D50 (\%)$	5.8	8.0	6.1	1.7	2.0	3.3	4.0	5.5	19

Tabella 2.6: Indicatori energetici e parametri acustici calcolati nella posizione B5.

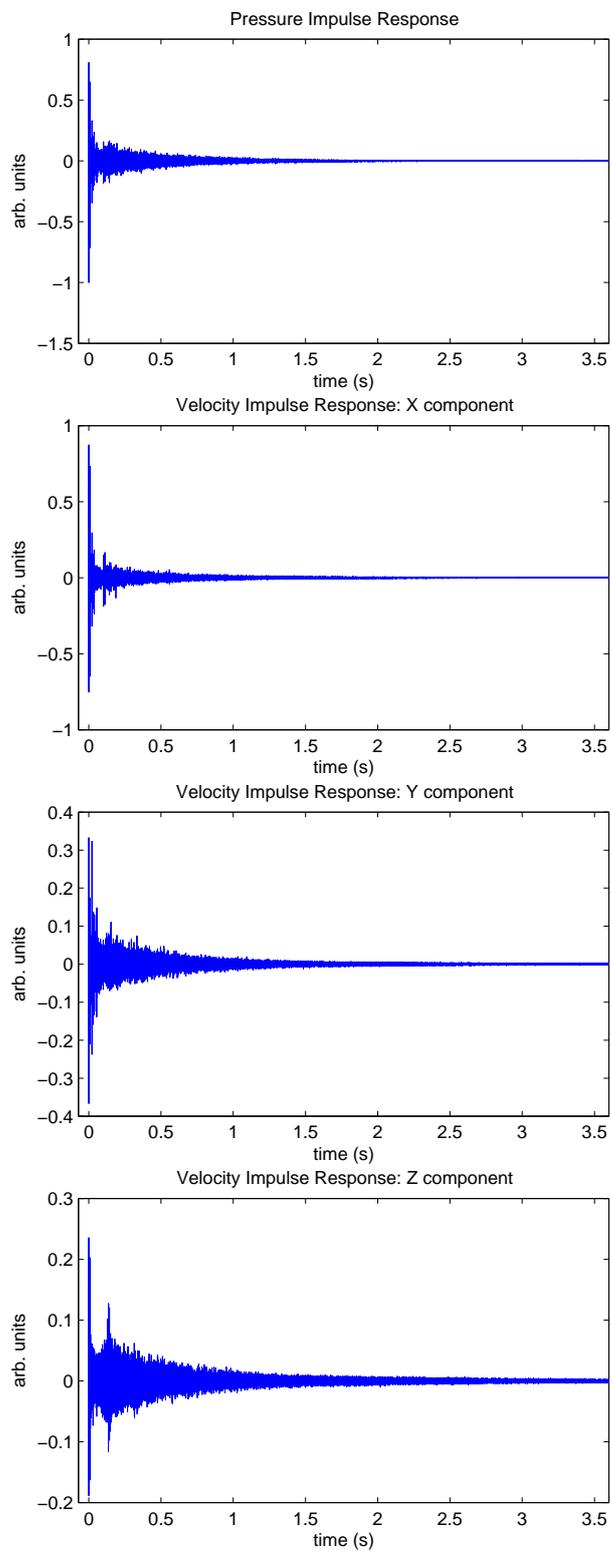


Figura 2.43: Risposta all'impulso misurata nella Basilica di S.Marco in configurazione A5.

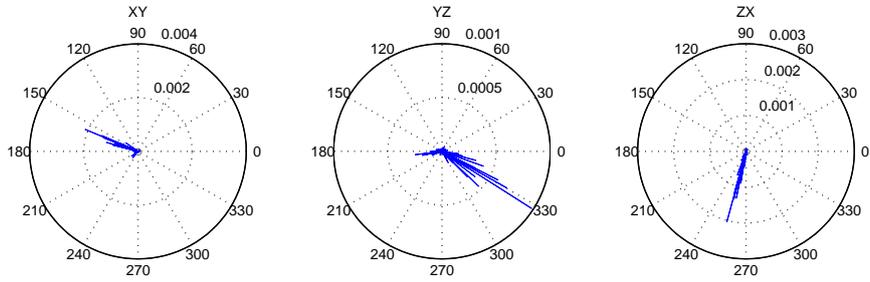


Figura 2.44: Polar plot del vettore intensità istantanea nei primi 80ms relativo alla risposta misurata in configurazione A5.

Capitolo 3

Le tecniche di spazializzazione attualmente usate

Una caratteristica fondamentale del nostro sistema uditivo è quella di permetterci di localizzare le sorgenti sonore nelle tre dimensioni. Un sistema di registrazione e riproduzione ad alta fedeltà dovrebbe riuscire a ricreare l'esperienza di un ascolto dal vivo trasmettendo all'ascoltatore l'impressione di prendere parte direttamente all'evento sonoro. Come abbiamo visto brevemente nel capitolo introduttivo, la nostra percezione della spazialità del suono sfrutta meccanismi diversi in regioni diverse dello spettro, pertanto viene naturale pensare che non sia necessariamente indispensabile codificare e restituire tutta l'informazione contenuta nel campo acustico, ma che possa essere sufficiente concentrarsi sulle grandezze più significative per la nostra percezione, sia in fase di registrazione che in fase di restituzione. Resta tuttavia il fatto che la misura completa delle proprietà fisiche di un evento sonoro sia l'obiettivo ultimo da perseguire per una sua caratterizzazione; inoltre è logico ammettere che, una volta in grado di ricostruire un campo sonoro che sia caratterizzato dagli stessi valori del set completo di osservabili misurati durante un evento reale, l'ascoltatore che si trovasse immerso in quel campo fruirebbe della stessa esperienza percettiva sperimentabile "dal vivo"¹.

In generale, una tecnica di ripresa e restituzione sonora che miri al realismo può sfruttare un approccio di tipo fisico o di tipo psicoacustico.

In questo capitolo presenteremo la Wavefield Synthesis e l'Ambisonics, due tecnologie nate e sviluppate negli scorsi decenni, che mirano ad una ricostruzione fedele delle proprietà spaziali di un campo acustico basandosi principalmente su un approccio rispettivamente di tipo fisico e di tipo psicoacustico, entrambe attraverso un certo grado di approssimazione.

¹Verrebbe ancora a mancare l'impatto visivo e il coinvolgimento emotivo dell'evento reale, che secondo l'autore sono ugualmente importanti.

3.1 La Wavefield Synthesis

La Wavefield Synthesis (WFS) è un approccio alla spazializzazione sonora introdotto negli anni ottanta e derivato dalla teoria dell'integrale di Kirchoff-Helmholtz. Essa si sviluppa a partire dalla tecnica dell'olofonia, nata con lo scopo di ricostruire un evento sonoro in un campo d'ascolto esteso, in modo da permettere a più ascoltatori simultaneamente di fruire della corretta riproduzione delle proprietà spaziali di un campo sonoro. Essenzialmente, l'olofonia si prefigge lo scopo di ricostruire un campo sonoro entro un volume, note le variabili acustiche sulla superficie che racchiude tale volume.

Nel presentare la Wavefield Synthesis seguiremo qui, nei punti essenziali, lo sviluppo trattato in [21], senza addentrarci troppo nei dettagli matematici.

Il metodo generale di ricostruzione di un campo sonoro tridimensionale (olofonia) si basa essenzialmente sul principio di Huygens, che si può esprimere dicendo che il campo sonoro prodotto da una sorgente è equivalente al campo generato da un insieme continuo di sorgenti collocate su un fronte d'onda emesso dalla sorgente primaria.

Suddividiamo l'intero spazio Ω in due sottospazi Ω_1 e Ω_2 , in modo che il primo contenga le sorgenti primarie (*dominio delle sorgenti*) e il secondo non ne contenga (*dominio di restituzione*). L'obiettivo dell'olofonia è di esprimere il campo sonoro in Ω_2 come generato non dalle sorgenti contenute in Ω_1 bensì da una distribuzione di sorgenti secondarie collocate sulla frontiera $\partial\Omega$ tra i due domini. Dal punto di vista matematico, il problema si risolve per mezzo delle funzioni di Green, e il campo di pressione al tempo t nel punto \vec{r} interno al dominio Ω_2 si può scrivere come

$$p(\vec{r}, t) = \int \int_{\partial\Omega_0} \partial S_0 \hat{n} \cdot \left\{ \int_{t_1}^{t_2} dt_0 \left[g(\vec{r} - \vec{r}_0, t - t_0) \vec{\nabla} p_0(\vec{r}_0, t_0) - p_0(\vec{r}_0, t_0) \vec{\nabla} g(\vec{r} - \vec{r}_0, t - t_0) \right] \right\} \quad (3.1)$$

dove \hat{n} è il versore ortogonale alla superficie ∂S della frontiera, \vec{r}_0 è la variabile di integrazione e t_1 e t_2 sono gli estremi dell'intervallo temporale di durata dell'evento sonoro. Pertanto, per acquisire l'informazione necessaria alla ricostruzione fisica del campo basterebbe collocare un *array* di microfoni omnidirezionali e bidirezionali su una superficie, in grado di misurare rispettivamente la pressione $p_0(\vec{r}_0, t_0)$ e il gradiente di pressione $\vec{\nabla} p_0(\vec{r}_0, t_0)$, come illustrato in figura 3.1. Ciò equivale a registrare i segnali corrispondenti alle sorgenti secondarie. Per la ricostruzione del campo, i segnali precedentemente registrati devono essere riprodotti da altoparlanti rispettivamente monopolari (i cosiddetti sistemi *baffled* ovvero a cassa chiusa) e bipolari (*non-baffled* ovvero a cassa aperta, cioè con la membrana dell'altoparlante lasciata libera di irradiare da entrambi i lati), collocati nelle stesse posizioni in cui si trovavano i microfoni in fase di registrazione (figura 3.2). L'intero processo è schematizzato nella figura 3.3.

La trattazione finora presentata prevederebbe una distribuzione continua di sorgenti (infiniti segnali) collocate su una superficie in tre dimensioni, per la ricostruzione tridimensionale del campo. Se ci si volesse limitare ad una ricostruzione bidimensionale del campo sonoro, ad esempio sul piano orizzontale, si può dimostrare che sarebbe sufficiente effettuare la misura e la restituzione

collocando microfoni ed altoparlanti su una superficie appartenente allo stesso piano. Questo risultato è dimostrato nel cosiddetto *teorema della fase stazionaria* e la semplificazione prende il nome di *approssimazione della fase stazionaria*. Per limitare ulteriormente il numero dei dispositivi in gioco, si ricorre ad un'ulteriore approssimazione che fa uso di un numero finito di ricevitori e diffusori, collocati in punti discreti di una curva su un piano, e che prevede l'utilizzo di microfoni a dipolo (bi-direzionali) per la misura e altoparlanti monopolari per la ridiffusione (formulazione di Rayleigh dell'integrale di Kirchoff-Helmholtz - si veda in proposito [21]). La tecnica della Wavefield Synthesis non è altro che l'implementazione di questa approssimazione, che si dimostra di particolare interesse per la sua applicabilità in settori quali le telecomunicazioni, le teleconferenze e il *sound reinforcement* nel cinema e nelle sale da concerto.

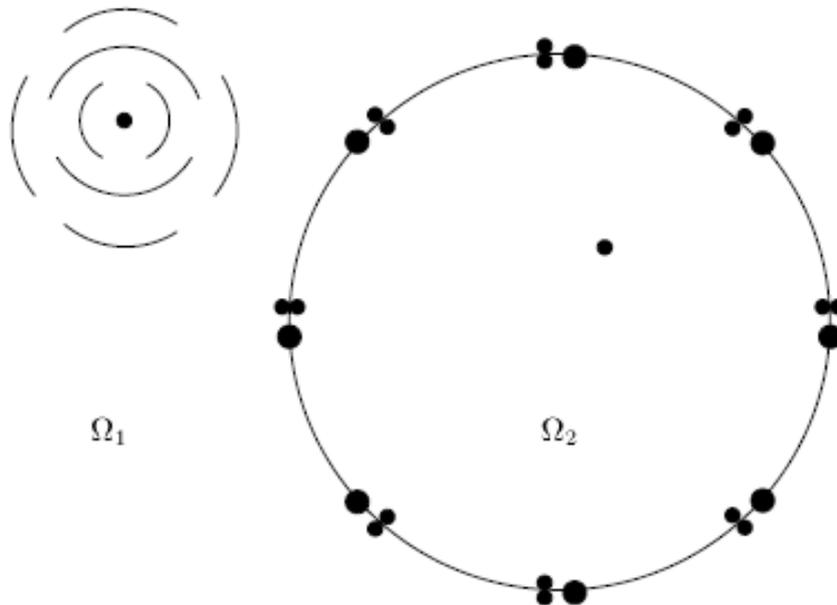


Figura 3.1: Registrazione di pressione e gradiente di pressione con microfoni omni-direzionali e bi-direzionali sulla frontiera tra il dominio delle sorgenti e il dominio di restituzione.

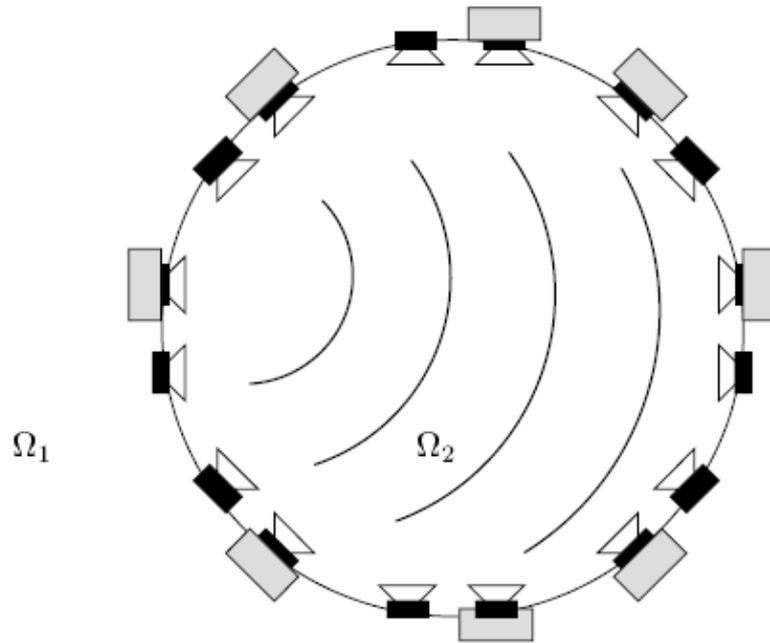


Figura 3.2: Ricostruzione olofonica del campo nel dominio di restituzione mediante altoparlanti monopolari e dipolari collocati sulla frontiera tra i domini Ω_1 e Ω_2 .

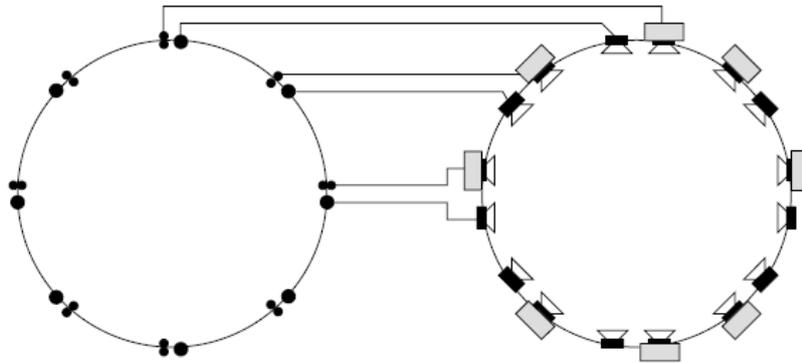


Figura 3.3: Schema concettuale del processo di registrazione e diffusione olofonica di un campo sonoro. La tecnologia può anche essere applicata nel settore delle telecomunicazioni e delle teleconferenze, trasmettendo in tempo reale i dati registrati e ricostruendo il campo sonoro in diretta per ascoltatori collocati in un altro ambiente.

3.2 Ambisonics

Studi compiuti negli anni settanta sulla spazializzazione del suono e sulla percezione uditiva portarono all'invenzione di un sistema di registrazione dei suoni e di riproduzione surround noto col nome di Ambisonics. Il sistema Ambisonics prevede la codifica di un evento sonoro in un punto mediante un set di quattro segnali $\{W, X, Y, Z\}$, che prende il nome di B-Format, ed è indipendente dal sistema usato per la riproduzione. Infatti, a partire dal B-Format è possibile ricavare i segnali corrispondenti ad una data configurazione di altoparlanti applicando delle matrici che ottimizzano la ricostruzione dell'esperienza d'ascolto reale basandosi su criteri psicoacustici.

3.2.1 Il B-Format

Il B-Format è stato proposto dai suoi sviluppatori, in primo luogo Michael Gerzon, come un insieme di quattro segnali che rappresentano tutta l'informazione necessaria a descrivere le proprietà spaziali del campo sonoro in un punto. I quattro canali del B-Format vengono indicati con W, X, Y e Z e rappresentano i segnali provenienti da quattro microfoni in configurazione coincidente². Fissato un sistema di riferimento cartesiano la cui origine coincida col punto di misura in cui vengono collocati i microfoni, W rappresenta il segnale proveniente da un microfono omnidirezionale, mentre X, Y e Z sono i segnali provenienti da tre microfoni bi-direzionali orientati lungo i corrispondenti assi cartesiani. La figura 3.4 illustra i quattro diagrammi polari dei trasduttori corrispondenti ai segnali che compongono il B-Format: il segnale W ha caratteristiche omnidirezionali, X corrisponde ad un bi-direzionale nella direzione fronte-retro, Y è il bi-direzionale orientato lungo la direzione sinistra-destra e Z è il segnale bi-direzionale nella direzione alto-basso.

Per esprimere la dipendenza funzionale della sensibilità dei microfoni bi-direzionali in uno spazio a tre dimensioni conviene effettuare una trasformazione di coordinate, passando ad un sistema di coordinate sferiche, illustrato in figura 3.5:

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arctan\left(\frac{y}{x}\right) \\ \varphi &= \arctan\left(\frac{z}{\sqrt{x^2 + y^2}}\right) \end{aligned} \tag{3.2}$$

²Per configurazione coincidente o microfoni coincidenti si intende un setup in cui due o più microfoni vengono collocati il più vicino possibile, in modo che i segnali da loro generati siano riferiti, idealmente, allo stesso singolo punto dello spazio. Naturalmente le dimensioni finite di una capsula microfonica impongono limitazioni alla distanza minima alla quale è possibile avvicinare le capsule, dando luogo a inaccurately e sfasamenti tra i segnali.

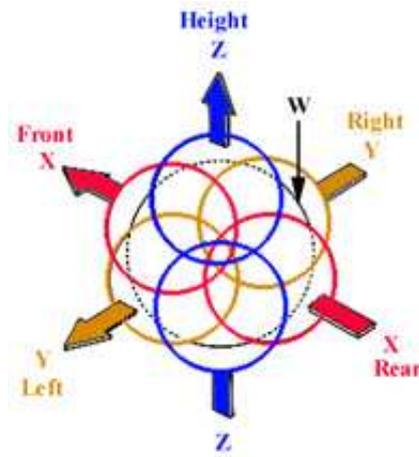


Figura 3.4: Diagrammi polari dei quattro trasduttori coincidenti che generano il segnale B-Format.

In questa rappresentazione, la dipendenza angolare della sensibilità si esprime con le seguenti relazioni:

$$\begin{aligned}
 W &= 1 \\
 X &= \cos(\theta) \cos(\varphi) \\
 Y &= \sin(\theta) \cos(\varphi) \\
 Z &= \sin(\varphi)
 \end{aligned}$$

dove i quattro segnali sono normalizzati al valore 1.

È necessario fare una nota riguardo la calibrazione dei segnali: in presenza di un'onda piana progressiva, si ha $|p| = \rho_0 c |\vec{v}|$, dove il termine $\rho_0 c = z_0$ rappresenta l'impedenza acustica caratteristica dell'aria. I segnali X, Y e Z sono normalizzati alla pressione, cioè vengono moltiplicati per z_0 per essere numericamente compatibili. Inoltre, il segnale W rappresenta la pressione sonora ridotta di $3dB$, cioè $W = \frac{1}{\sqrt{2}}p$; questa convenzione deriva dal fatto che frequentemente il segnale che presenta valori di picco più elevati nel corso della registrazione di un evento sonoro in un punto è la pressione. Per permettere di poter registrare o digitalizzare i quattro segnali con ampiezze confrontabili senza applicare guadagni diversi nelle fasi di preamplificazione e conversione digitale, si è deciso di ridurre di $3dB$ il livello di W. Naturalmente si è tenuto conto di questo nei confronti effettuati e ogni qualvolta sia stata usata una tecnologia quadrifonica acustica su un segnale B-Format o viceversa.

3.2.2 Il microfono SoundField

Lo stato dell'arte della tecnologia per la registrazione del B-Format è rappresentato dal microfono SoundField, introdotto sul mercato negli anni settanta e ora presente in diverse versioni, analogiche e digitali, abbinato a varie unità di controllo più o meno sofisticate per la decodifica e la manipolazione dei segnali. Il microfono è composto da un singolo corpo che contiene quattro capsule microfoniche a condensatore. Al contrario di quanto si potrebbe pensare,

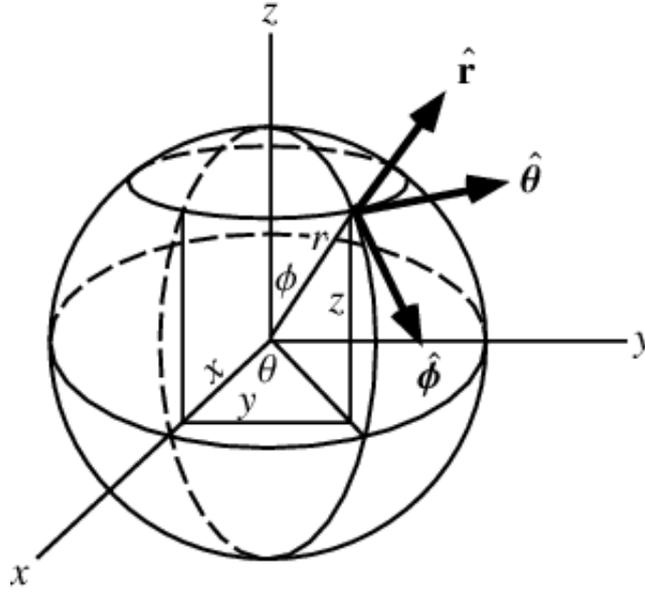


Figura 3.5: Definizione delle variabili nelle coordinate sferiche.

queste capsule non forniscono direttamente i quattro segnali di pressione e gradiente di pressione; ciascuna di esse infatti è un trasduttore misto di pressione-gradiente di pressione, con caratteristiche polari subcardioidi espresse dalla seguente equazione:

$$p_{output} = A \left[\frac{3}{4}p + \frac{1}{4}\vec{\nabla}p \cdot \hat{n} \right] \quad (3.3)$$

ove il versore \hat{n} identifica l'orientazione del massimo di sensibilità della capsula. Nel microfono SoundField quattro capsule subcardioidi vengono montate sui vertici di un tetraedro regolare. Le quattro capsule assemblate sono mostrate in figura 3.6.

Esprimendo l'orientazione delle capsule in coordinate sferiche (φ, ϑ) , i segnali corrispondenti sono: L_F (left-front): $(45^\circ; 54.7^\circ)$; L_B (left-back): $(135^\circ; 125.3^\circ)$; R_F (right-front): $(225^\circ; 54.7^\circ)$; R_B (right-back): $(315^\circ; 125.3^\circ)$. Questi quattro segnali costituiscono un insieme denominato A-Format. Essendo i quattro segnali dell'A-Format una combinazione lineare di p e $\vec{\nabla}p$, da essi può essere ricavato il B-Format mediante un'applicazione lineare, che può essere espressa in forma matriciale nel seguente modo:

$$\begin{pmatrix} W \\ X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} L_F \\ R_B \\ L_B \\ R_F \end{pmatrix} \quad (3.4)$$

Il calcolo per la sintesi del B-Format può essere implementato elettronicamente all'interno di un'unità di controllo analogica fornita con il microfono oppure attraverso un software, dopo aver acquisito l'A-Format nel dominio digitale. Il passaggio attraverso l'A-Format sembra essere una necessità dovuta ai limiti



Figura 3.6: A sinistra: schema della collocazione delle capsule microfoniche sulle quattro facce di un tetraedro regolare. A destra: immagine delle capsule di un microfono SoundField.

intrinseci della tecnologia impiegata; in primo luogo, una capsula microfonica a condensatore tradizionale per applicazioni di alta fedeltà necessita di un diametro superiore al mezzo pollice. Risulterebbe pertanto difficile riuscire a montare in configurazione coincidente quattro capsule, di cui tre orientate lungo gli assi del sistema di riferimento cartesiano, poichè inevitabilmente alcune di loro verrebbero mascherate dalla presenza delle altre, che verrebbero a creare una zona d'ombra in grado di compromettere l'accuratezza del sistema, in particolare alle frequenze più alte. La collocazione ai vertici di un tetraedro regolare fa sì che ogni capsula sia esposta direttamente al campo sonoro³. Il fatto che le capsule non siano coincidenti, bensì spaziate di poco più di un centimetro, comporta un tempo d'arrivo del suono leggermente diverso da capsula a capsula. Ad ogni frequenza, questa differenza di tempo d'arrivo si traduce in uno sfasamento dei segnali, che dà luogo ad effetti di interferenza quando vengono effettuate le somme o le differenze dei segnali per ricavare il B-Format. Per ovviare a questo problema si introduce un filtraggio sulle componenti del B-Format, filtraggio che ha quindi l'effetto di rendere le capsule (e, di conseguenza, i segnali d'uscita) virtualmente coincidenti. Per analizzare il problema, consideriamo la capsula i -esima, il cui pattern, in coordinate cartesiane, sia espresso dalla funzione $f_i(x, y, z)$. Sia (u_i, v_i, w_i) la posizione del centro della capsula, in un sistema di riferimento la cui origine corrisponde col centro del tetraedro. Il diagramma polare relativo all'origine si può scrivere come

$$F_i(x, y, z, \omega) = f_i(x, y, z) \exp[i(xu_i + yv_i + zw_i)\omega/c] \quad (3.5)$$

³Vedremo comunque che le dimensioni delle capsule impongono una limitazione alla performance del sistema ad alte frequenze.

dove ω è la frequenza del suono. Il termine esponenziale rappresenta lo sfasamento tra la capsula sulla faccia del tetraedro e la capsula ideale posta al centro; tale sfasamento è ovviamente dovuto alla distanza finita tra la faccia del solido e il centro, che determina una differenza nel tempo di arrivo del suono. Per compensare l'effetto di tale sfasamento, Gerzon ha proposto l'utilizzo di filtri caratterizzati dalle seguenti risposte:

$$F_W(\omega) = \frac{1 + i\omega\tau - \frac{1}{3}\omega^2\tau^2}{1 + \frac{1}{3}i\omega\tau} \quad (3.6)$$

$$F_{X,Y,Z}(\omega) = \sqrt{6} \frac{1 + \frac{1}{3}i\omega\tau - \frac{1}{3}\omega^2\tau^2}{1 + \frac{1}{3}i\omega\tau} \quad (3.7)$$

rispettivamente per i segnali W e X,Y,Z.

Il segnale proveniente da una capsula subcardioide, come abbiamo visto, può essere interpretato come un misto di pressione e gradiente di pressione, nelle proporzioni 3/4 e 1/4. Questo fatto impone delle limitazioni all'accuratezza con cui vengono ricostruiti i segnali del B-Format, in quanto su questi si ripercuoteranno i difetti riscontrabili nelle capsule miste, in primo luogo l'effetto prossimità e la variazione della risposta polare in funzione della frequenza, discussi nel capitolo introduttivo. Le capsule, aventi dimensioni dell'ordine del *cm*, costituiscono un ostacolo per la propagazione del suono e comportano effetti di diffrazione e ombra quando le lunghezze d'onda sono dello stesso ordine di grandezza della dimensione delle capsule.

3.2.3 La decodifica dei segnali per la riproduzione

Come si è visto nel capitolo introduttivo, nel registrare e riprodurre i suoni è possibile ricreare l'illusione di un'immagine sonora in cui le varie sorgenti vengono ricostruite rispettando quelle che erano le loro posizioni originali nell'evento reale. Per fare questo, la strada più seguita è quella che prevede la codifica di informazioni direzionali sotto forma di differenze di livello fra trasduttori e la seguente riproduzione di tali differenze attraverso sistemi di altoparlanti opportunamente configurati. Questo principio viene usato anche per la riproduzione del B-Format. In base a quanto spiegato nell'introduzione, i principali tipi di microfoni, classificati in base alle loro caratteristiche direzionali, possono essere descritti come combinazione di un segnale di pressione e uno di gradiente di pressione misurati simultaneamente nello stesso punto. Essendo il gradiente di pressione un vettore identificabile dalle tre componenti cartesiane $\vec{\nabla}p \equiv \left(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}, \frac{\partial p}{\partial z}\right)$ ed essendo una generica derivata direzionale calcolabile a partire dal gradiente con la formula $\partial_{\hat{n}}p = \hat{n} \cdot \vec{\nabla}p$, conoscendo la pressione e il gradiente di pressione in un punto è possibile ricavare il segnale corrispondente ad un microfono di pattern arbitrario orientato nella direzione \hat{n} nel modo seguente:

$$\begin{aligned} S_{out}(\hat{n}, B) &= S_{out}(\vartheta, \varphi, B) = [Bp - (1 - B)z_0\vec{v} \cdot \hat{n}] = \\ &= \left[\sqrt{2}BW - (1 - B)(X, Y, Z) \cdot (\cos \vartheta \cos \varphi, \sin \vartheta \cos \varphi, \sin \varphi) \right] \end{aligned} \quad (3.8)$$

ove (ϑ, φ) sono gli angoli che identificano la direzione in coordinate sferiche e B è il coefficiente che determina il pattern del microfono ($B = 1$ omnidirezionale,

$B = 0$ bi-direzionale, $B = 1/2$ cardioide, $B = 1/4$ ipercardioide, $B = 3/4$ sub-cardioide). Pertanto, dal set di segnali che compongono il B-Format è possibile ricostruire un numero arbitrario di *microfoni virtuali* orientati in qualsiasi direzione, con l'unico vincolo che essi corrispondono allo stesso punto dello spazio in cui si è acquisito il B-Format⁴. Creando un solo microfono virtuale si potrà ottenere un segnale monofonico adatto alle esigenze della diffusione radiofonica. Per una ricostruzione stereofonica dell'evento sonoro è sufficiente creare, attraverso un'apposita matrice di decodifica, i segnali corrispondenti ad una coppia di microfoni direzionali, scegliendo l'angolo e il pattern in maniera da ottenere l'immagine sonora desiderata. Per l'ascolto in 5.1 surround, la tecnica maggiormente usata a livello commerciale consiste nel sintetizzare cinque microfoni virtuali di pattern cardioide orientati nelle direzioni -45° , 0° , $+45^\circ$, -135° , $+135^\circ$, i cui segnali verranno poi assegnati rispettivamente ai canali L, C, R, RL, RR; il segnale per gli effetti a bassa frequenza viene ottenuto applicando un filtro passa basso al segnale di pressione W. Tale tecnica è la naturale estensione multicanale delle tecniche stereofoniche coincidenti. In mancanza di informazioni reperibili riguardo gli angoli e i pattern proposti, l'autore ritiene che la scelta sia emersa e si sia affermata principalmente per via sperimentale, attraverso prove e test d'ascolto. E' curioso notare che questa configurazione minimizza il *leakage* dei segnali tra gli altoparlanti frontali e quelli posteriori contrapposti, ovvero tra le coppie L-RR e R-RL. Infatti, un microfono cardioide presenta un minimo di sensibilità (idealmente una sensibilità nulla) a 180° rispetto al proprio asse. Di conseguenza, il segnale assegnato al canale L, ottenuto da un cardioide orientato a -45° , ha una minima sensibilità ai suoni provenienti da $+135^\circ$, direzione nella quale il segnale RR presenta il massimo di sensibilità. Questo accorgimento previene la comparsa di segnali posteriori sui canali frontali e viceversa.

Anche se nella maggior parte dei casi gli altoparlanti di un sistema d'ascolto sono collocati su un piano orizzontale, pressapoco alla stessa altezza delle orecchie dell'ascoltatore, il B-Format permette anche la riproduzione di informazioni relative all'altezza (*azimut*). Questo chiaramente richiede la presenza di altoparlanti posti al di sopra e/o al di sotto dell'ascoltatore. Studi sono stati compiuti per definire le possibili configurazioni di altoparlanti e le loro prestazioni [22], tuttavia questi sistemi non hanno mai trovato un'applicazione diffusa, soprattutto a causa della difficoltà di allestire un adeguato sistema d'ascolto destinato al mercato consumer. Per quanto riguarda la riproduzione stereofonica e surround, sono stati proposti metodi di decodifica dei segnali basati su criteri psicoacustici, a partire dai pionieristici lavori di M.Gerzon [23], [24]. Come già accennato nell'introduzione, il presupposto di partenza è che il nostro sistema uditivo utilizzi, per la localizzazione delle sorgenti sonore, dei meccanismi diversi a seconda delle bande di frequenza considerate. Fu Lord Rayleigh a scoprire una differenza tra i meccanismi di localizzazione al di sotto e al di sopra dei $700Hz$. Altre evidenze suggeriscono un'ulteriore differenziazione al di sopra dei $5kHz$.

Considerando che la distanza tra le orecchie di un essere umano corrisponde a mezza lunghezza d'onda di un suono che abbia la frequenza di $700Hz$, il fenomeno della diffrazione consente di supporre che la testa non rappresenti un ostacolo

⁴In realtà è stato previsto un controllo, denominato *dominance*, che agisce sulle componenti del B-Format permettendo di ottenere l'impressione di avvicinare o allontanare il microfono dalla sorgente, come avviene per lo zoom di una macchina fotografica. L'implementazione del controllo verrà discussa nel capitolo 4.

significativo alla propagazione delle onde sonore al di sotto dei $700Hz$. Pertanto, in condizioni di onda piana, un suono di frequenza inferiore a $700Hz$ giunge ad entrambe le orecchie senza significative differenze di ampiezza⁵. L'unica informazione sulla quale si possa basare il sistema percettivo per la localizzazione delle sorgenti rimane quindi la differenza di fase tra i segnali che giungono alle due orecchie. Considerando un ascoltatore orientato lungo la direzione positiva dell'asse x , la differenza $p_{left} - p_{right}$ tra i due segnali in corrispondenza delle due orecchie è proporzionale a $\partial p / \partial y$, che a sua volta, in base a quanto discusso nel capitolo 1, risulta proporzionale alla velocità v_y . Questo porta a concludere che la corretta ricostruzione del vettore velocità in fase di riproduzione sia il prerequisito necessario per permettere la localizzazione delle sorgenti a frequenze medio basse. Al di sopra dei $700Hz$ invece il fattore più significativo per la localizzazione dei suoni è la distribuzione del campo di intensità sonora attorno all'ascoltatore. Questo meccanismo domina solo fino a che non entrano in gioco i fattori direzionali della risposta delle orecchie. In particolare, al di sopra dei $5kHz$ il nostro cervello sfrutta le differenze di livello e di *colorazione* tra i suoni percepiti dalle due orecchie. Per colorazione si intende una alterazione della risposta in frequenza di un segnale operata da un sistema. Nel caso del nostro orecchio, quando le lunghezze d'onda dei suoni sono confrontabili con o inferiori alle dimensioni del padiglione auricolare, differenti direzioni d'arrivo si traducono in differenze nella risposta in frequenza che vengono codificate dal nostro cervello in informazioni sulla localizzazione delle sorgenti.

Vediamo ora come vengono attualmente utilizzati questi criteri psicoacustici per la ricostruzione della direzionalità del suono. Consideriamo dunque un ascoltatore e un sistema di n altoparlanti collocati sul piano dell'ascoltatore ed equidistanti dal punto d'ascolto. Ci basiamo su un sistema di riferimento cartesiano in cui l'asse x punta in avanti e l'asse y punta verso sinistra. Siano $\Phi_i (i = 1, 2, \dots, n)$ gli angoli che identificano le posizioni degli altoparlanti e siano $\alpha_i (i = 1, 2, \dots, n)$ i fattori di amplificazione complessi (cioè che comprendono ampiezza e fase) dei segnali emessi dagli altoparlanti. Nell'ipotesi in cui l'ascoltatore sia raggiunto da onde piane, l'altoparlante che si trova nella direzione Φ_i produce un vettore velocità le cui componenti sono $(\alpha_i \cos(\Phi_i), \alpha_i \sin(\Phi_i))$. Il vettore velocità generato dal sistema di altoparlanti, pari alla somma delle velocità prodotte dai singoli altoparlanti, avrà quindi componenti

$$v_x = Re \left\{ \frac{\sum \alpha_i \cos \Phi_i}{\sum \alpha_i} \right\} \quad (3.9)$$

$$v_y = Re \left\{ \frac{\sum \alpha_i \sin \Phi_i}{\sum \alpha_i} \right\} \quad (3.10)$$

E' possibile scrivere

$$v_x = r_v \cos \theta_v \quad (3.11)$$

$$v_y = r_v \sin \theta_v \quad (3.12)$$

⁵Un'onda piana non è soggetta ad attenuazione dovuta alla divergenza nella propagazione. Il suono in un punto che disti più di poche lunghezze d'onda dalla sorgente può essere considerato in buona approssimazione un'onda piana, almeno per quanto riguarda possibili differenze di intensità tra due punti distanti circa $17cm$. Nel caso in cui una sorgente fosse posta in prossimità di un orecchio, ci si troverebbe in condizioni di onda sferica e la differenza di intensità tra le due orecchie, conseguenza della attenuazione di $6dB$ al raddoppiare della distanza, sarebbe l'effetto dominante per la localizzazione del suono.

dove θ_v viene interpretato come direzione di provenienza apparente del suono. Alle basse frequenze, il criterio proposto per l'ottimizzazione dell'esperienza d'ascolto è la massimizzazione di r_v .

Per quanto riguarda le alte frequenze, l'intensità prodotta da un altoparlante è proporzionale al quadrato del guadagno e ha per componenti $(\alpha_i^2 \cos(\Phi_i), \alpha_i^2 \sin(\Phi_i))$. L'intensità complessiva, pari alla somma delle intensità, ha per componenti

$$E_x = Re \left\{ \frac{\sum \alpha_i^2 \cos \Phi_i}{\sum \alpha_i^2} \right\} \quad (3.13)$$

$$E_y = Re \left\{ \frac{\sum \alpha_i^2 \sin \Phi_i}{\sum \alpha_i^2} \right\} \quad (3.14)$$

e può essere espressa come

$$E_x = r_E \cos \theta_E \quad (3.15)$$

$$E_y = r_E \sin \theta_E \quad (3.16)$$

dove θ_E rappresenta la direzione di provenienza del suono percepita. Per le alte frequenze il criterio da seguire è la massimizzazione di r_E . Vogliamo osservare che, siccome $E_x, E_y \leq 1$, si ha $r_E \leq 1$, pertanto la condizione d'ascolto ideale si otterrebbe per $r_E = 1$; questo tuttavia è possibile solo nel caso esista i' tale che

$$\begin{cases} \alpha_i \neq 0 & i = i' \\ \alpha_i = 0 & i \neq i' \end{cases} \quad (3.17)$$

In tal caso si avrà $\theta_E = \Phi_{i'}$; questo si interpreta ammettendo che il valore massimo di r_E si può ottenere solo quando è attivo un solo altoparlante, nel qual caso la direzione percepita sarà la direzione in cui si trova l'altoparlante. Viceversa, con una tale configurazione di altoparlanti è possibile ricostruire esattamente il vettore intensità solo se si posiziona un altoparlante nella direzione in cui si vuole ricostruire la sorgente e si spengono tutti gli altri!

A partire da una registrazione B-Format e una data configurazione di altoparlanti per la riproduzione, è stato proposto un metodo di decodifica nel quale i segnali degli altoparlanti vengono ricostruiti con una matrice $\alpha_i = k_{i,1}W + k_{i,2}X + k_{i,3}Y$ calcolando i coefficienti $k_{i,j}$ in modo da massimizzare i valori di r_V ed r_E rispettivamente alle basse e alte frequenze [25], [26]. L'implementazione di tale decodifica richiede l'applicazione di un filtro crossover⁶, il processamento separato delle due regioni dello spettro ($f \gtrsim 700Hz$ e $f \lesssim 700Hz$) e la successiva ricostruzione del segnale in banda larga mediante somma.

Il processo sopra descritto può essere applicato anche alla ricostruzione di un campo acustico a partire da una registrazione quadrifonica, come verrà mostrato nel capitolo 4.

3.3 La compatibilità della Quadrifonia Acustica con Wavefield Synthesis e Ambisonics

L'esistenza di diversi metodi per la registrazione e la riproduzione degli eventi sonori pone all'utilizzatore il problema della scelta della tecnologia più vantaggiosa in termini di qualità del risultato prodotto, reperibilità dei dispositivi di

⁶Un crossover è un filtro in grado di separare un segnale in due o più bande di frequenza.

registrazione e riproduzione e compatibilità con gli standard utilizzati dall'industria e dagli altri utenti. Senz'altro la tecnologia più vantaggiosa è quella in grado di portare delle migliorie alla situazione attuale mantenendo anche la "compatibilità verso il basso" con tutti gli altri sistemi presenti. In questo senso, un grosso punto a favore della Quadrifonia Acustica è il fatto che, oltre a portare migliorie in termini della caratterizzazione fisica ed energetica di un campo sonoro, sia di fatto compatibile con i due sistemi descritti in questo capitolo, seppur entro un certo grado di approssimazione [27]. In questa sezione presenteremo le relazioni esistenti tra WFS e Ambisonics descritte in [21] e troveremo una corrispondenza tra Ambisonics e Quadrifonia Acustica. La corrispondenza tra Quadrifonia Acustica e WFS seguirà per proprietà transitiva.

Consideriamo innanzitutto un sistema di riproduzione Ambisonics costituito da N altoparlanti collocati su una circonferenza in corrispondenza dei vertici del poligono di N lati in essa inscritto. L'altoparlante n -esimo si trova ad un angolo $\theta_n = n2\pi/N$, dove $\theta = 0$ identifica la direzione frontale in cui guarda l'ascoltatore. Si dimostra che le funzioni di codifica dei segnali da inviare agli altoparlanti per la migliore restituzione del campo misurato nel punto centrale sono date da

$$S_i(\theta) = \frac{1}{N} [1 + 2 \cos(\theta - \theta_i)] \quad (3.18)$$

cioè all'altoparlante θ_n viene mandato il segnale corrispondente ad un microfono virtuale di pattern supercardioide orientato nella direzione $\theta = \theta_n$. Il campo di pressione valutato nel centro del sistema si ottiene facendo la somma dei segnali generati da tutti gli altoparlanti e può essere scritto in coordinate polari (r, ϕ) come:

$$\begin{aligned} p(r, \phi) &= \sum_{i=1}^N s_i e^{jkr \cos(\phi - \theta_i)} \\ &= \frac{1}{N} \sum_{i=1}^N [1 + 2 \cos(\theta - \theta_i)] e^{jkr \cos(\phi - \theta_i)} \end{aligned} \quad (3.19)$$

Le assunzioni fatte per derivare la corrispondenza con la WFS sono le seguenti:

- si assume che il campo da ricostruire con l'approccio olofonico sia un'onda piana progressiva
- gli altoparlanti sono disposti su un cerchio per una ricostruzione bidimensionale (su una sfera nel caso tridimensionale)
- La distanza dell'ascoltatore dagli altoparlanti deve essere maggiore della lunghezza d'onda del suono riprodotto, in modo da poter considerare l'onda proveniente da ogni altoparlante come un'onda piana.

Con riferimento alla figura 3.7, applicando l'approssimazione della fase stazionaria all'integrale di Kirchoff-Helmholtz si può esprimere la pressione in un punto interno del dominio di restituzione in funzione delle sorgenti di monopolo distribuite sulla frontiera nel modo seguente:

$$p(r, \varphi; \omega) = \sqrt{\frac{2\pi c}{j\omega}} r_s \int_0^{2\pi} \frac{\partial p_0(r_s, \varphi_s; \omega)}{\partial r_s} \text{Rect}\left(\frac{\varphi_0 - \varphi_s}{\pi}\right) \frac{e^{-j\omega R/c}}{2\pi\sqrt{R}} d\varphi_s \quad (3.20)$$

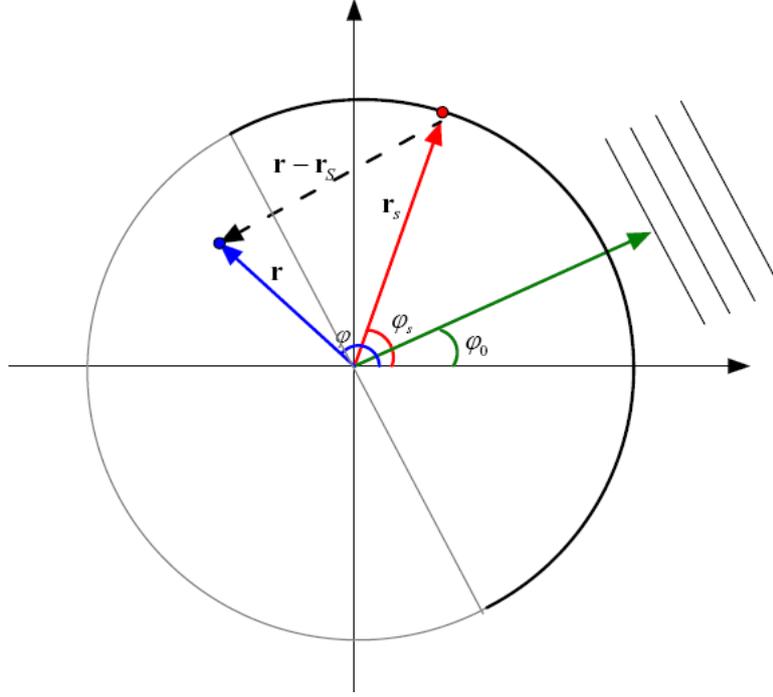


Figura 3.7: Schema di riproduzione olfonica su un piano con microfoni collocati su una circonferenza. φ_0 identifica la direzione di provenienza dell'onda piana, \vec{r}_s rappresenta il generico punto di misura sulla frontiera del dominio di restituzione e \vec{r} è il punto in cui viene valutato il campo di pressione ricostruito.

dove $R = |\vec{r} - \vec{r}_s|$ e $Rect(\varphi) = 1$ per $|\varphi| \leq 1/2$, $Rect(\varphi) = 0$ per $|\varphi| > 1/2$. Questa espressione, che esprime il campo all'interno del dominio di restituzione come sovrapposizione di un insieme continuo di sorgenti collocate sulla circonferenza di raggio r_s , può essere ridotta al caso di un numero finito di sorgenti discrete. Assumendo la presenza di N sorgenti secondarie, otteniamo:

$$p(r, \varphi; \omega) = \frac{1}{N} \sqrt{\frac{2\pi c}{j\omega}} r_s \left[\sum_{i=1}^N \frac{\partial p_0(r_s, \varphi_i; \omega)}{\partial r_s} Rect\left(\frac{\varphi_0 - \varphi_i}{\pi}\right) \frac{e^{-j\omega R_i/c}}{2\pi\sqrt{R_i}} \right] \quad (3.21)$$

dove $R_i = |\vec{r} - \vec{r}_s| = \sqrt{r^2 + r_s^2 - 2rr_s \cos(\varphi - \varphi_i)}$.

L'equazione 3.21 può essere espressa in funzione della velocità della particella acustica grazie all'equazione di Eulero linearizzata nel dominio della frequenza $\hat{n} \cdot \vec{\nabla} P(\omega) = -j\rho_0\omega V_{\hat{n}}(\omega)$, ottenendo:

$$p(r, \varphi; \omega) = \frac{1}{N} \sqrt{\frac{2\pi c}{j\omega}} r_s \left[-j\rho_0\omega \sum_{i=1}^N V_{r_s}(r_s, \varphi_i; \omega) Rect\left(\frac{\varphi_0 - \varphi_i}{\pi}\right) \frac{e^{-j\omega R_i/c}}{2\pi\sqrt{R_i}} \right] \quad (3.22)$$

Il campo di pressione può quindi essere riprodotto misurando la componente normale della velocità acustica in un insieme di N punti disposti in maniera regolare su un cerchio di raggio r_s .

Consideriamo ora la scomposizione di un'onda piana progressiva in armoniche cilindriche espressa dall'identità di Jacobi-Anger:

$$e^{iz \cos \alpha} = \sum_{n=-\infty}^{+\infty} j^n J_n(z) e^{in\alpha} = J_0(z) + 2 \sum_{n=1}^{+\infty} j^n J_n(z) \cos(n\alpha)$$

Sostituendo $z = kr_s$ e $\alpha = \varphi - \varphi_0$ otteniamo:

$$\begin{aligned} a e^{ikr_s \cos(\varphi - \varphi_0)} &= a J_0(kr_s) + 2a \sum_{n=1}^{+\infty} j^n J_n(kr_s) \cos(n(\varphi - \varphi_0)) \\ &= a J_0(kr_s) + 2a \sum_{n=1}^{+\infty} j^n J_n(kr_s) \cos(n\varphi) \cos(n\varphi_0) \\ &\quad + 2a \sum_{n=1}^{+\infty} j^n J_n(kr_s) \sin(n\varphi) \sin(n\varphi_0) \end{aligned}$$

Troncando l'espansione al primo ordine, si ottiene:

$$P_0(r_s, \varphi) \simeq a J_0(kr_s) + 2aj J_1(kr_s) \cos(\varphi) \cos(\varphi_0) + 2aj J_1(kr_s) \sin(\varphi) \sin(\varphi_0)$$

La componente normale della velocità risulta essere

$$\begin{aligned} \frac{j}{\omega \rho_0} \frac{\partial P_0(r_s, \varphi)}{\partial r_s} &\simeq \frac{a}{\rho_0 c} J_0' \left(\frac{\omega r_s}{c} \right) + \frac{2a}{\rho_0 c} j J_1' \left(\frac{\omega r_s}{c} \right) \cos(\varphi) \cos(\varphi_0) \\ &\quad + \frac{2a}{\rho_0 c} j J_1' \left(\frac{\omega r_s}{c} \right) \sin(\varphi) \sin(\varphi_0) \\ &= \frac{P_0(0)}{\rho_0 c} J_0' \left(\frac{\omega r_s}{c} \right) + \frac{2V_{x0}(0)}{\rho_0^2 c^2} j J_1' \left(\frac{\omega r_s}{c} \right) \cos(\varphi) \\ &\quad + \frac{2V_{y0}(0)}{\rho_0^2 c^2} j J_1' \left(\frac{\omega r_s}{c} \right) \sin(\varphi) \end{aligned} \quad (3.23)$$

L'equazione 3.23 esprime la componente normale della velocità sulla circonferenza in funzione dei segnali di pressione e delle due componenti ortogonali della velocità valutati al centro del cerchio. Questi segnali, per onde piane, non sono altro che le componenti W, X e Y del B-Format su cui si basano i sistemi Ambisonics, qui sommati con coefficienti proporzionali alle derivate delle funzioni di Bessel del primo tipo J_0 e J_1 . Ponendo

$$\tilde{P}_0(\omega) = P_0(\omega) J_0' \left(\frac{\omega r_s}{c} \right); \quad \tilde{V}_{x0}(\omega) = j V_{x0}(\omega) J_1' \left(\frac{\omega r_s}{c} \right); \quad \tilde{V}_{y0}(\omega) = j V_{y0}(\omega) j J_1' \left(\frac{\omega r_s}{c} \right)$$

e

$$a_p = \frac{1}{\rho_0 c}; \quad a_x = \frac{2 \cos(\varphi)}{\rho_0^2 c^2}; \quad a_y = \frac{2 \sin(\varphi)}{\rho_0^2 c^2},$$

l'equazione 3.23 può essere riformulata nel dominio del tempo applicando la trasformata di Fourier inversa per dare

$$v(r_s, \varphi; t) = a_p IFT[\tilde{P}_0(\omega)](t) + a_x IFT[\tilde{V}_{x0}(\omega)](t) + a_y IFT[\tilde{V}_{y0}(\omega)](t) \quad (3.24)$$

Questa equazione esprime infine la relazione che sussiste tra i segnali del B-Format misurati al centro del sistema e le funzioni di decodifica per la restituzione olofonica del campo secondo la WFS.

Per quanto riguarda la compatibilità dei sistemi Ambisonics con la Quadrifonia Acustica, il primo indizio viene dato dal fatto che i quattro segnali su cui si basa la Quadrifonia, cioè (p, v_x, v_y, v_z) hanno le stesse caratteristiche polari dei quattro segnali del B-Format (W, X, Y, Z) che stanno alla base dell'Ambisonics. Infatti p e W rappresentano grandezze scalari e vengono misurate da trasduttori omnidirezionali, mentre le componenti della velocità e del gradiente di pressione sono componenti di un vettore e vengono misurate da sensori bi-direzionali (pattern a figura di otto). Entrambi i sistemi si basano sulla misura simultanea e coincidente di quattro grandezze fisiche, che sono pressione e velocità nel caso della Quadrifonia Acustica e pressione e gradiente di pressione nel caso dell'Ambisonics⁷. Il collegamento tra le due tecnologie è dato dall'equazione di Eulero linearizzata

$$\vec{\nabla}p = -\rho_0 \frac{\partial \vec{v}}{\partial t} \quad (3.25)$$

che lega la velocità acustica col gradiente di pressione. Nel caso di un'onda piana progressiva $p(\vec{r}, t) = \hat{p} \exp[i(\vec{k} \cdot \vec{r} - \omega t)]$ (\vec{k} è il vettore di modulo $2\pi/\lambda = \omega/c$ e direzione e verso uguali a quelli di propagazione dell'onda), il gradiente di pressione vale $\vec{\nabla}p(\vec{r}, t) = i\vec{k}\hat{p} \exp[i(\vec{k} \cdot \vec{r} - \omega t)]$ e integrando l'equazione di Eulero 1.55 per l'onda piana otteniamo:

$$\vec{v} = -\frac{1}{\rho_0} \int_{-\infty}^t \vec{\nabla}p dt = \frac{i}{\rho_0 \omega} \vec{\nabla}p \quad (3.26)$$

Si ha quindi

$$\vec{\nabla}p(\vec{r}, t) = iz_0 k \vec{v}(\vec{r}, t) \quad (3.27)$$

dove $z_0 = \rho_0 c$ è l'impedenza acustica dell'aria. L'equazione 3.26 ci dice che i segnali di velocità e gradiente di pressione risultano proporzionali per un'onda piana. L'analogia del B-Format con la Quadrifonia Acustica consiste dunque nel fatto che i due set di segnali su cui si basano le tecnologie sono proporzionali, in approssimazione di onda piana progressiva.

Usando il risultato della 3.27, l'espressione 2 che descrive il segnale in uscita da un generico microfono può essere riscritta come somma di pressione e velocità nel seguente modo:

$$V_{out} = A[p - z_0 B(\hat{n} \cdot \vec{v})] \quad (3.28)$$

Questo significa che possiamo usare combinazioni di segnali di pressione e velocità per creare i microfoni virtuali per la decodifica Ambisonics di una registrazione Quadrifonica: questa applicazione verrà descritta nel capitolo 4. Analogamente, si potrebbe tentare di usare un microfono SoundField per misure Quadrifoniche, tuttavia il campo di applicabilità del sistema nell'ambito quadrifonico è limitato dal fatto che l'identificazione del gradiente di pressione con la velocità delle particelle d'aria è il risultato di un'approssimazione che presuppone la presenza di un campo di onde piane. Questa semplificazione non risulta sempre conforme alle caratteristiche dei campi reali; sappiamo, ad esempio, che

⁷Anche se, come visto nella sezione 3.2.2, queste non vengono misurate direttamente bensì ricavate mediante combinazione lineare di altri quattro segnali.

si può considerare il campo emesso da una sorgente come un campo di onde sferiche, perlomeno entro i limiti in cui l'estensione della sorgente è inferiore alla lunghezza d'onda del suono emesso. Le onde sferiche, a loro volta, possono essere approssimate ad onde piane solo quando la distanza dalla sorgente sia molto maggiore della lunghezza d'onda, condizione non sempre verificata durante la registrazione di un evento sonoro o l'esecuzione di una misura acustica.

Grazie a questa corrispondenza, in prima approssimazione, tra Quadrifonia Acustica e Ambisonics e alla relazione dimostrata tra WFS e Ambisonics, risulta in linea di principio possibile implementare dei sistemi di transcodifica capaci di permettere una riproduzione fedele del suono registrato con una di queste tre tecniche su un sistema basato su un'altra.

3.3.1 Il confronto tra la sonda Microflown e il microfono SoundField

In questa sezione vogliamo proporre un confronto tra la sonda intensimetrica Microflown ed il microfono SoundField per presentare i pregi ed i difetti dei due sensori, vista la relativa corrispondenza appena descritta tra i segnali prodotti dai due sistemi. In realtà le specifiche tecniche dichiarate dai costruttori e le considerazioni fatte nelle sezioni precedenti sarebbero già sufficienti a fornire indicazioni significative all'operatore che si trovasse a dover scegliere tra i due strumenti per una particolare applicazione, tuttavia approfittiamo di questa sezione per sottolineare alcuni aspetti sinora poco chiari.

Il confronto qui presentato riguarda il modello Microflown USP descritto nella sezione 2.1 e il microfono SoundField modello ST 250. Quest'ultimo è un microfono che incorpora quattro capsule subcardioidi in configurazione tetraedrica (si veda la sezione 3.2.2) ed è fornito di un'unità esterna per l'alimentazione delle capsule e la codifica del B-Format a partire dall'A-Format. Essendo quest'ultimo un apparecchio dedicato al mercato dell'audio professionale, soddisfa criteri di alta fedeltà quali una risposta in frequenza estesa da $20Hz$ a $20kHz$ entro una tolleranza di $3dB$ e ha un rumore intrinseco pari a $17dBA$. La sonda intensimetrica, al contrario, presenta una curva di risposta dei sensori di velocità che necessita di un filtraggio e produce un rumore di fondo che tende ad aumentare sia alle basse frequenze (*rumble*) che alle alte (*hiss*). Esso risulta decisamente udibile, tanto da compromettere l'utilità dello strumento per la registrazione diretta di un evento sonoro, come evidenziato da prove di registrazione e test di ascolto che abbiamo effettuato in laboratorio. Questa considerazione ci porta dunque a preferire il microfono SoundField alla sonda Microflown per la registrazione di suoni. Ricordiamo tuttavia che l'utilizzo principale della sonda è stato per noi quello di misurare le risposte all'impulso, applicazione nella quale il rumore di fondo viene sensibilmente ridotto grazie alla particolare tecnica di cross-correlazione implementata, al punto da fornire risultati di poco inferiori agli standard dell'alta fedeltà.

Per quanto già detto riguardo le dimensioni dei due sensori e il loro principio di funzionamento, ci aspettiamo invece dalla sonda Microflown risultati più accurati dal punto di vista fisico. Per testare questo aspetto, abbiamo posto entrambi i sensori in posizione coincidente in un campo acustico noto, nella fattispecie il campo generato da un altoparlante che emette uno sweep sinusoidale di frequenza variabile da $100Hz$ a $20kHz$ in camera anecoica, in posizione frontale rispetto ai sensori ad una distanza di $1.3m$ da questi. Il segnale è stato

acquisito e da esso è stato calcolato l'indicatore η e il vettore intensità istantanea $\vec{i}(t)$. In queste condizioni ci si aspetta un'intensità sonora diretta lungo la congiungente tra gli altoparlanti e i microfoni. Il campo risultante è in buona approssimazione un'onda piana progressiva al di sopra dei $1000Hz$, campo per il quale η assumerebbe il valore 1. Al di sotto dei $1000Hz$ il campo diventa sempre più assimilabile ad un'onda sferica man mano che la frequenza scende, per cui ci aspettiamo un valore di η che si abbassa al diminuire della frequenza, e una corrispondente diminuzione del modulo del vettore intensità. La tabella 3.1 riporta i valori di η e σ in banda larga e in bande di ottava.



Figura 3.8: Il confronto tra la sonda Microflow e il microfono Soundfield è stato effettuato in camera anecoica disponendo i due sensori in configurazione coincidente.

Frequenza (Hz)	<i>overall</i>	125	250	500	1k	2k	4k	8k
η Microflow	0.88	0.72	0.75	0.84	0.90	0.90	0.91	0.82
η SoundField	0.82	0.81	0.90	0.94	0.91	0.87	0.81	0.70
σ Microflow	0.92	0.87	0.85	0.89	0.93	0.92	0.94	0.93
σ SoundField	0.99	0.93	0.97	0.97	0.98	0.99	0.98	0.85

Tabella 3.1: Valore degli indicatori η e σ calcolati con sonda Microflow e microfono SoundField.

Osserviamo che nel caso della sonda Microflow l'indicatore η assume valore crescente all'aumentare della frequenza, eccezion fatta per il dato a $8kHz$.

Questi dati sono consistenti con le considerazioni di cui sopra sul tipo di campo acustico che si viene a creare in queste condizioni. Il valore ricavato con il microfono SoundField invece risulta sospettosamente alto alle frequenze medio basse e decresce a partire dai $2kHz$.

Le figure 3.9 e 3.10 mostrano i plot polar della proiezione del vettore intensità sul piano orizzontale, misurati rispettivamente con sonde Microflow e SoundField. Appare evidente la maggiore accuratezza della sonda intensimetrica in tutte le bande di frequenza considerate: il vettore intensità si mantiene mediamente diretto lungo l'asse x e il plot non presenta un significativo effetto di allargamento, ad eccezione dall'ottava centrata sulla frequenza di $8kHz$, dove la possibile spiegazione dello sparpagliamento del vettore intensità attorno alla direzione assiale è interpretabile come l'effetto della presenza di rumore di fondo nei sensori di velocità. Al contrario, dai dati ottenuti con il SoundField risulta che la misura di intensità è decisamente meno accurata e che la direzione del vettore intensità può cambiare notevolmente a seconda della banda di frequenza considerata. Questo effetto è dovuto principalmente al fatto che i quattro segnali del B-Format non vengono misurati direttamente ma sono ricavati da combinazioni lineari di segnali le cui caratteristiche polari subiscono inevitabili variazioni al variare della frequenza, come visto nella sezione in cui abbiamo descritto le caratteristiche dei microfoni.

In base a quanto mostrato, abbiamo concluso che l'utilizzo della sonda intensimetrica rappresenta al momento il metodo più accurato per l'analisi energetica dei campi acustici.

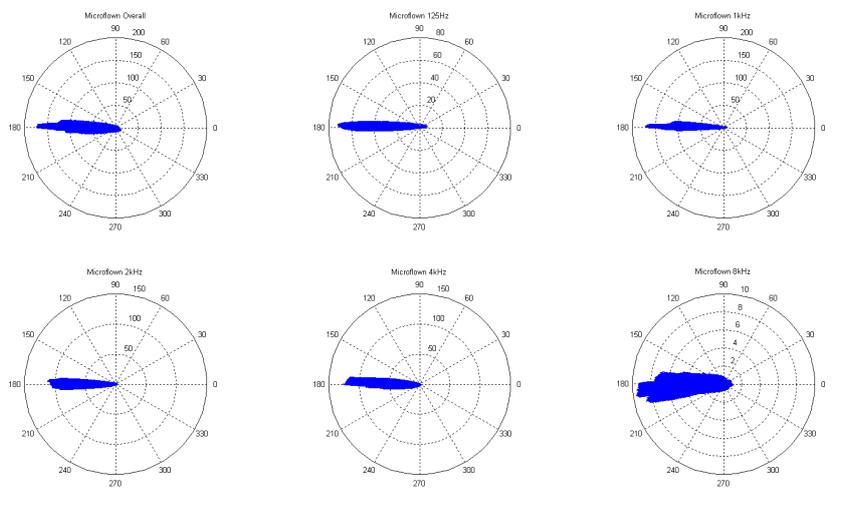


Figura 3.9: Polar plot del vettore intensità misurato con sonda Microflow. In alto, da sinistra: analisi in banda larga, 125Hz, 1kHz. In basso, da sinistra: 2kHz, 4kHz, 8kHz.

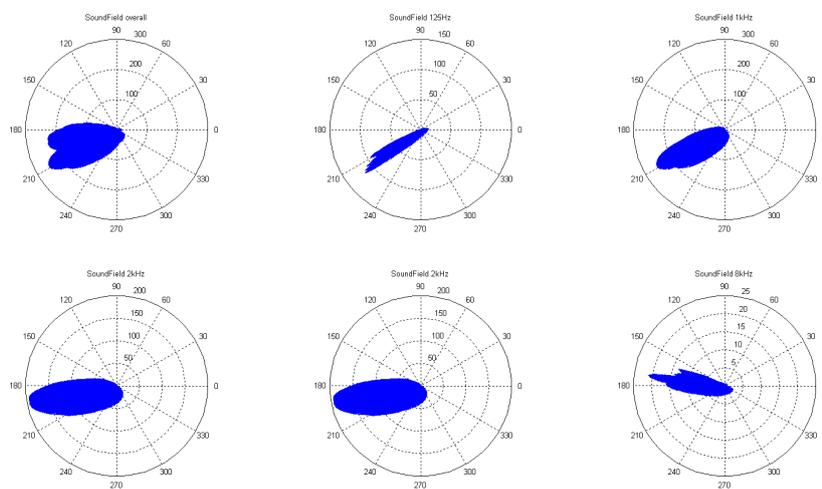


Figura 3.10: Polar plot del vettore intensità misurato con microfono SoundField. In alto, da sinistra: analisi in banda larga, 125Hz, 1kHz. In basso, da sinistra: 2kHz, 4kHz, 8kHz.

Capitolo 4

La restituzione di registrazioni quadrifoniche con le attuali tecnologie

4.1 Il processo di convoluzione quadrifonica

Il set di equazioni 1.74 identifica il campo prodotto nel punto di misura \vec{r}_M da una sorgente puntiforme collocata in \vec{r}_S con la convoluzione del segnale di pressione generato dalla sorgente e della risposta quadrifonica all'impulso del sistema acustico relativo alle posizioni \vec{r}_M ed \vec{r}_S . Essendo in grado di misurare la risposta all'impulso di un ambiente, a partire dal segnale di pressione generato da una sorgente sonora possiamo ricostruire con un'operazione di convoluzione i segnali di pressione e velocità che misureremmo se collocassimo tale sorgente nell'ambiente misurato. Tale procedura prende il nome di *convoluzione quadrifonica*. In questa sezione vogliamo illustrare il processo di convoluzione quadrifonica e presentare alcune possibili applicazioni.

Per ottenere il segnale di pressione generato da una sorgente è necessario riuscire a registrare tale sorgente in un ambiente che non apporti alcun contributo al campo sonoro, ovvero in condizioni di campo libero. In pratica, la condizione di campo libero viene approssimata effettuando misure all'interno delle *camere anecoiche*, ambienti nei quali la presenza di materiale fonoassorbente sulle superfici impedisce la riflessione sonora. Collocando una sorgente sonora in camera anecoica e registrandone il suono con un microfono di pressione (omnidirezionale) otteniamo il segnale di pressione della sorgente. Bisogna però precisare che l'espressione 1.74 è stata dimostrata nel caso di una sorgente puntiforme, mentre le sorgenti reali sono caratterizzate da un'estensione finita. Il campo prodotto da una sorgente puntiforme è un campo di onde sferiche (detto anche campo di *monopolo*), mentre nel campo prodotto dalle sorgenti estese sono presenti contributi di *multipolo* che introducono fattori di direzionalità. In generale una sorgente può essere considerata puntiforme nel limite in cui le sue dimensioni sono trascurabili rispetto alla lunghezza d'onda del suono emesso. Nel caso della voce umana, ad esempio, possiamo assumere un'emissione omnidirezionale alle basse frequenze ($f \lesssim 343Hz$, $\lambda \gtrsim 1m$), mentre al crescere della frequenza

l'emissione si fa sempre più direzionale e la presenza della testa impedisce alla voce di propagarsi all'indietro. Si pensi poi al caso di un'orchestra, dove i singoli strumenti sono collocati su una zona che supera facilmente la decina di metri di estensione: in questo caso sarebbe impensabile sperare di ottenere l'equivalenza tra il campo sonoro reale prodotto da un'orchestra e il risultato della convoluzione della risposta all'impulso misurata con sorgente puntiforme e di un singolo segnale di pressione relativo a tutta l'orchestra. In realtà, nel caso di sorgenti estese, per ottenere la ricostruzione fisica corretta del campo sarebbe necessario usare la stessa sorgente estesa per l'emissione del segnale di eccitazione durante la fase di misura. Ovviamente è impensabile sperare di ottenere uno sweep sinusoidale o un qualsiasi tipo di segnale conforme a parametri fisici ben precisi e ripetibile da una sorgente musicale, tuttavia la cosa può essere realizzata approssimando la sorgente estesa con un numero finito di sorgenti puntiformi ed effettuando le misure di risposta all'impulso con altrettante sorgenti puntiformi disposte nella stessa configurazione. Ognuna di queste risposte dovrebbe essere poi convoluta con il corrispondente segnale anecoico di pressione generato dalla sorgente reale e misurato da altrettanti microfoni posti nelle stesse posizioni relative al processo di misurazione. Nei nostri esperimenti ci siamo tuttavia limitati a sorgenti quali voce umana e piccoli strumenti musicali, che abbiamo considerato puntiformi.

Una volta registrati i segnali di pressione anecoici all'interno della camera anecoica del laboratorio FSSG-CNR di Venezia (per una descrizione più dettagliata della camera si veda la sezione 4.4.2), l'operazione di convoluzione delle quattro componenti della risposta all'impulso con il segnale di pressione è stata implementata attraverso il software *Matlab*. Anziché implementare la convoluzione nel dominio del tempo, espressa dall'equazione 1.74, abbiamo usato la relazione $s(t) \otimes g(t) = IFFT[FFT(s) \cdot FFT(g)]$, più efficiente dal punto di vista computazionale. Segnaliamo anche la possibilità di eseguire il processo all'interno delle più avanzate applicazioni professionali di post-produzione audio, grazie alla recente immissione sul mercato di *plug-in* in grado di effettuare la convoluzione di due file audio permettendo all'operatore il monitoraggio dell'operazione, che viene svolta in tempo reale con una latenza di pochi millisecondi. Su un pc attuale l'operazione di convoluzione da noi eseguita risulta più veloce rispetto all'operazione in tempo reale.

Per testare il processo di convoluzione abbiamo ideato un semplice esperimento che consiste nel confronto tra un evento sonoro registrato dal vivo in una sala e l'evento ottenuto mediante convoluzione della risposta all'impulso della sala col segnale anecoico che ha dato origine all'evento sonoro stesso. Per far questo siamo partiti dalla registrazione anecoica di un attore che ha recitato la frase: "*In language, infinitely many words can be written with a small set of letters*". Abbiamo poi diffuso questa registrazione nella sala Carnelutti attraverso la sorgente omnidirezionale usata per le misurazioni, registrando simultaneamente i quattro segnali di pressione e velocità misurati dalla sonda Microflown. Sonda e sorgente sono state collocate nella posizione di misura n°4 (si veda la sezione 2.5.1). Successivamente, per fare il confronto, abbiamo effettuato la convoluzione della risposta all'impulso misurata nella stessa posizione con lo stesso segnale anecoico precedentemente riprodotto in ambiente. A partire da questi dati, sono stati calcolati i valori medi di densità totale, potenziale e cinetica di energia e gli indicatori η e σ su tutta la durata degli eventi sonori, riportati nella tabella 4.1.

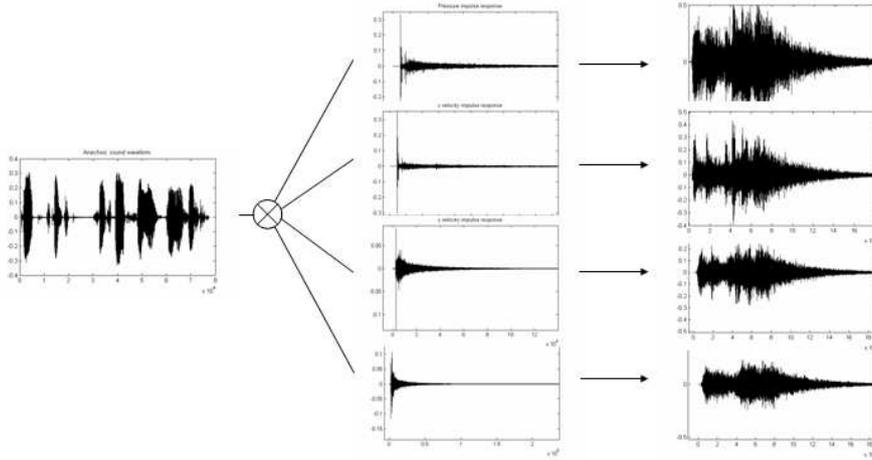


Figura 4.1: Forma d'onda dei segnali nel processo di convoluzione. A sinistra: segnale anecoico di pressione per una voce recitante maschile. Al centro: componenti di pressione e velocità della risposta all'impulso. A destra: risultato della convoluzione: si noti come le pause nel parlato anecoico siano state riempite dalla riverberazione dell'ambiente dopo il processo di convoluzione.

	W_P (A.U.)	W_K (A.U.)	W_{tot} (A.U.)	η	σ
Evento sonoro	0.0015	0.0033	0.0047	0.28	0.93
Convoluzione	0.0015	0.0012	0.0025	0.33	0.99

Tabella 4.1: Confronto di indicatori energetici relativi a evento sonoro reale e convoluzione.

Il valore di W_P risulta uguale in entrambi i casi: questa non è altro che una coincidenza numerica, in quanto il nostro sistema non è stato calibrato in unità fisiche e nell'effettuare la misura non si è tenuto conto dei guadagni di tutte le apparecchiature, in quanto eravamo interessati solo all'andamento temporale di pressione e velocità e ai relativi rapporti di ampiezza. I valori di η e σ sono invece assoluti e possono essere confrontati direttamente. Partendo dallo stesso valore di densità di energia potenziale, il confronto tra la ripartizione dell'energia nei due casi risulta più immediato: possiamo infatti affermare che la densità di energia cinetica W_K è maggiore nel caso dell'evento registrato realmente nell'ambiente. Questo effetto è causato dal rumore di fondo presente nei segnali di velocità, già discusso nei capitoli precedenti. Nel calcolo dell'energia cinetica il rumore viene elevato al quadrato prima dell'operazione di media, pertanto

apporta un contributo positivo, che si riflette poi anche nel valore della densità totale di energia. I due valori di η , alla luce delle incertezze descritte nel capitolo 2, sono al limite della compatibilità. Un valore inferiore di η nel caso dell'evento sonoro reale è in parte causato dalla crescita del termine W_K che compare al denominatore nella sua espressione, anche se entra in gioco il valore di A presente al numeratore.

Risultati analoghi sono stati ottenuti applicando la stessa metodologia in altri ambienti.

La convoluzione quadrifonica può essere impiegata per diversi scopi. Una naturale applicazione consiste nel suo utilizzo nella post-produzione musicale e cinematografica. Nel cinema infatti la maggior parte dei dialoghi e dei suoni viene registrata in studi quasi anecoici e l'ambienza viene aggiunta in seguito mediante riverberazione artificiale. Inoltre, quando viene effettuato il doppiaggio in un'altra lingua, ci si trova alle prese col problema di dover aggiungere una riverberazione che sia simile il più possibile a quella contenuta nell'opera in lingua originale. Se in fase di ripresa si misurasse per ogni ambiente la risposta all'impulso e poi la si rendesse disponibile tra i *metadati* assieme ai dati video, sarebbe possibile in ogni parte del mondo effettuare un doppiaggio anecoico e rendere il suono uguale all'originale mediante la convoluzione con i metadati. Questa operazione è stata sviluppata e presentata con successo nell'ambito del progetto europeo *IP-Racine*.

Un'altra applicazione è la spazializzazione sonora, che permette di ricostruire un campo sonoro virtuale in cui le sorgenti siano collocate nello spazio secondo criteri stabiliti dal compositore: è sufficiente registrare le risposte all'impulso negli ambienti e nelle posizioni desiderate e poi effettuare le convoluzioni con segnali anecoici delle sorgenti; questo permette tra l'altro di fondere in un unico evento sonoro ambienti dalle caratteristiche notevolmente diverse. Tra l'altro, applicando una matrice di rotazione ai segnali di velocità, possiamo operare la virtuale rotazione del sistema di riferimento, che equivale di fatto a ruotare la sonda intensimetrica al momento della misura. La rotazione di un angolo φ in senso antiorario sul piano orizzontale si ottiene trasformando i segnali $\{p, v_x, v_y, v_z\}$ in un nuovo set $\{p', v'_x, v'_y, v'_z\}$, con

$$\begin{cases} p' &= p \\ v'_x &= v_x \cos \varphi - v_y \sin \varphi \\ v'_y &= v_x \sin \varphi + v_y \cos \varphi \\ v'_z &= v_z \end{cases} \quad (4.1)$$

Ciò consente di poter manipolare a piacere la direzione di arrivo del suono per una data sorgente a partire da una singola misura di risposta all'impulso¹.

Recentemente la convoluzione quadrifonica è stata applicata ad ambienti virtuali, cioè a modelli di ambienti realizzati al calcolatore. Il lavoro, attualmente in fase di sviluppo per conto del gruppo *BarcelonaMedia* nell'ambito di *IP-Racine*, prevede il calcolo numerico della risposta quadrifonica all'impulso di

¹Purchè tale misura contenga informazioni direzionali significative a livello percettivo: ad esempio, una misura effettuata in campo diffuso e priva del contributo diretto della sorgente non renderebbe possibile la localizzazione della sorgente, pertanto sarebbe inutile operare una rotazione.

un ambiente mediante tecniche di soluzione numerica dell'equazione d'onda e *Ray-Tracing* e la successiva *auralizzazione*² per l'ascolto.

4.2 Il problema della transcodifica e della riproduzione

Avendo dimostrato che l'insieme dei dati quadrifonici relativi ad un evento sonoro costituisce un set completo per la sua identificazione, sorge il problema, una volta registrato correttamente³ il campo acustico in un punto, di come riuscire a riprodurlo in un altro ambiente mantenendone le caratteristiche originali. Per quanto riguarda l'approccio da seguire, siamo convinti che un approccio mirato alla ricostruzione fedele delle grandezze fisiche e dei parametri energetici sia da preferirsi ad uno di tipo psicoacustico. L'ideale sarebbe riuscire a ricostruire correttamente il campo di pressione e velocità nel punto d'ascolto, cosa che garantirebbe automaticamente il rispetto dei parametri energetici originali [28]. Questo obiettivo è tuttavia di difficile realizzazione pratica, e perciò anche per la restituzione dei dati quadrifonici acustici nel presente lavoro sono stati seguiti i criteri attualmente in uso per i sistemi tradizionali.

Tradizionalmente infatti i problemi che si pongono nell'affrontare la questione della riproduzione del suono sono due:

1. Qual'è la configurazione di altoparlanti che meglio si presta alla ricostruzione del campo?
2. Una volta scelta la configurazione di altoparlanti, quale combinazione di segnali devo inviare a ciascuno di essi per ottimizzare la fedeltà di riproduzione dell'evento sonoro?

La risposta alla prima domanda dipende dalle caratteristiche del campo sonoro da ricostruire. Ad esempio, per restituire in maniera soddisfacente un dialogo cinematografico, un sistema monofonico potrebbe risultare sufficiente. Tuttavia, per l'ascolto della musica sarebbe opportuno poter offrire un ascolto "surround" in cui l'ascoltatore possa percepire la posizione delle sorgenti sonore in un piano e le caratteristiche dell'ambiente in cui è stata effettuata la ripresa⁴. Vediamo quindi come i tradizionali sistemi di restituzione del suono sono stati utilizzati nell'ambito della Quadrifonia Acustica. Ciò verrà fatto partendo dalla decodifica monofonica e stereofonica dei dati quadrifonici, fino ad arrivare alla ricostruzione di un evento sonoro mediante un sistema 5.1 surround.

Per quanto riguarda l'ascolto monofonico, è sufficiente inviare all'altoparlante il segnale di pressione, che garantisce la corretta ricostruzione spettrale dell'evento sonoro. Qualora si vogliano ridurre suoni o rumori d'ambiente indesiderati o attenuare parte delle riflessioni dell'ambiente per avere una maggior chiarezza, è sufficiente sommare al segnale di pressione i segnali di velocità in

²Il termine *auralizzazione*, traduzione della parola inglese *auralization*, si riferisce in generale alle procedure che permettono di tradurre il risultato di una simulazione acustica, per esempio la risposta all'impulso di un ambiente ricavata numericamente da un modello fisico, in un prodotto ascoltabile quale ad esempio un file audio ottenuto mediante convoluzione della risposta con un segnale anecoico.

³Qui col termine "correttamente" si intende in maniera consistente dal punto di vista fisico.

⁴Ovvero le caratteristiche di ambienta create artificialmente in fase di post-produzione.

modo da ottenere, secondo l'equazione 3.8, il segnale equivalente ad un microfono virtuale direzionale con la massima sensibilità diretta verso la sorgente sonora.

Per la decodifica stereofonica, il metodo migliore consiste nel ricostruire i segnali *Left* e *Right* corrispondenti ad una coppia di microfoni direzionali in configurazione coincidente. Ciò è suggerito dall'analogia delle caratteristiche polari dei sensori di pressione e velocità con quelle dei microfoni omnidirezionali e a figura di otto. Ad esempio, per una decodifica equivalente alla tecnica Blumlein si ha $Left = v_x + v_y$ e $Right = v_x - v_y$: questa operazione equivale infatti alla rotazione di 45° dei due segnali di velocità, che genera due segnali a figura di otto orientati a $\pm 45^\circ$ rispetto alla direzione dell'asse x , come previsto dalla configurazione Blumlein. Qualora si volesse ottenere maggior ambienta o variare le dimensioni dell'immagine stereofonica, si può ricorrere ad una diversa scelta per i pattern e gli angoli dei microfoni virtuali. Un'applicazione ancora più raffinata consiste nel suddividere i segnali in bande di frequenza ed applicare matrici di decodifica diverse per ottimizzare la restituzione sonora nel modo desiderato su tutto lo spettro. Il grosso vantaggio rispetto alle tecniche convenzionali è che tutte queste scelte vengono fatte in fase di post-produzione, dove la presenza di un sistema di monitoraggio accurato e la possibilità di confronti comparativi possono guidare facilmente il tecnico esperto verso la decodifica migliore, cosa non possibile utilizzando le tradizionali tecniche microfoniche. Inoltre, i segnali ricostruiti a partire da combinazioni lineari dei dati quadrifonici non sono soggetti a *comb-filtering*⁵ e ad artefatti dovuti allo sfasamento tra le capsule microfoniche quasi coincidenti, in quanto, come visto, i sensori di pressione e velocità sono di fatto coincidenti e "puntiformi" rispetto alle lunghezze d'onda del suono.

Per quanto riguarda la restituzione di registrazioni quadrifoniche con un impianto 5.1 Surround, la tecnica utilizzata in questo lavoro si basa sull'implementazione di un processo già disponibile sul mercato sotto forma di un apparecchio hardware per la codifica surround del B-Format, che prevede la creazione dei segnali L, C, R, RL, RR a partire da W, X e Y, mediante ricostruzione di cinque microfoni virtuali di pattern cardioide orientati nelle direzioni 0° , $\pm 45^\circ$, $\pm 135^\circ$. L'apparecchio in questione è il processore surround *SoundField SP451*, mostrato in figura 4.2, da noi utilizzato in principio come riferimento per lo sviluppo di un software analogo.

Il processore, che opera interamente nel dominio analogico, presenta quattro ingressi per i segnali B-Format, controlli per il guadagno complessivo del segnale in ingresso e per i livelli singoli dei segnali in uscita, due controlli per variare l'angolo tra i due microfoni virtuali L ed R e/o tra i due posteriori RL e RR, ed un controllo per variare il pattern polare dei microfoni che generano i segnali dei canali surround. Basandoci su questo, abbiamo sviluppato un software di transcodifica surround su piattaforma *Matlab* in grado di eseguire le stesse

⁵Il *comb-filtering* o "filtro a pettine" è la caratteristica curva di risposta di un sistema in cui un segnale relativo ad un fenomeno oscillatorio viene misurato simultaneamente in due o più punti diversi e poi sommato. La distanza spaziale tra i due punti di ricezione determina un ritardo di fase che si traduce in un effetto di interferenza che porta all'amplificazione delle componenti la cui lunghezza d'onda è contenuta un numero intero di volte nella distanza che separa i sensori e alla soppressione delle componenti la cui lunghezza d'onda è contenuta un numero semintero di volte. Disegnando la curva di risposta in frequenza del filtro, si vede un andamento che richiama la forma di un pettine, da cui il nome dato all'effetto. Nel caso delle onde sonore, in presenza di *comb-filtering* il timbro del suono viene drammaticamente falsato.

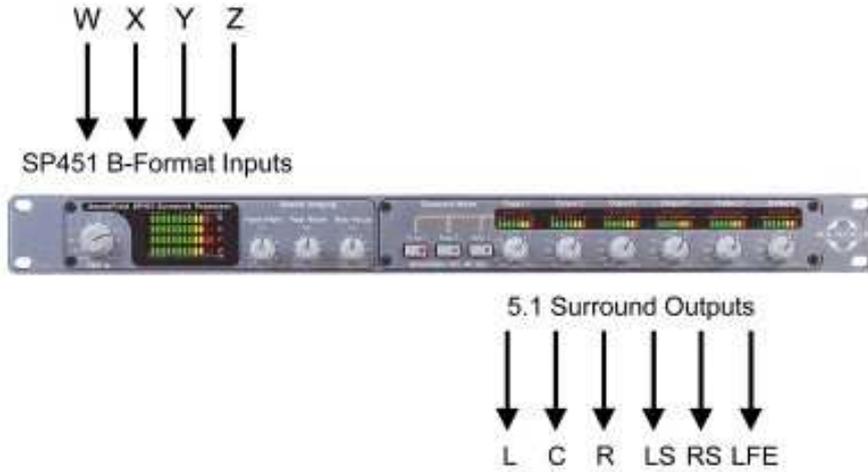


Figura 4.2: Pannello frontale del processore SoundField SP451.

operazioni nel dominio digitale, evitando così le fasi di conversione D/A e A/D necessarie per l'utilizzo del processore hardware⁶.

A partire dai dati quadrifonici $\{p, v_x, v_y, v_z\}$, il nostro software applica un filtro crossover splittando i segnali originali mediante l'applicazione di filtri Butterworth passa alto e passa basso con pendenza e frequenza di *cutoff* selezionabili (in genere usiamo una pendenza di $18dB/oct$ e una frequenza di taglio di $100Hz$), anch'essi implementati in *Matlab*. Dal segnale di pressione filtrato col passa basso viene creato il file audio *wav* che costituisce il canale LFE per gli effetti a bassa frequenza. I quattro segnali filtrati col passa alto vengono utilizzati per derivare i cinque canali effettuando le seguenti operazioni:

$$\begin{cases} L &= (1 - A_F)p - A_F(v_x \cos(\theta_F) + v_y \sin(\theta_F)) \\ C &= (1 - A_F)p - A_F v_x \\ R &= (1 - A_F)p - A_F(v_x \cos(\theta_F) - v_y \sin(\theta_F)) \\ RL &= (1 - A_R)p - A_R(v_x \cos(\theta_R) + v_y \sin(\theta_R)) \\ RR &= (1 - A_R)p - A_R(v_x \cos(\theta_R) - v_y \sin(\theta_R)) \end{cases} \quad (4.2)$$

A_F e A_R sono i coefficienti che determinano rispettivamente il pattern dei microfoni virtuali frontali e posteriori. θ_F e θ_R sono invece gli angoli corrispondenti alle direzioni nelle quali vengono orientati tali microfoni. Notiamo, rispetto all'equazione 3.8, la definizione diversa del parametro del pattern, per il quale si ha $A = 1 - B$, con A variabile da zero (omnidirezionale) a uno (figura di otto). Osserviamo inoltre che, sempre rispetto all'equazione 3.8, i segnali di velocità vengono qui cambiati di segno rispetto ai corrispondenti gradienti di pressione. Questo perchè mentre il segnale di gradiente di pressione risulta in fase con la pressione per un'onda piana progressiva proveniente dalla direzione positiva dell'asse x , la stessa onda provoca una velocità di segno invertito rispetto alla pressione (e dà luogo infatti ad un vettore intensità orientato nel verso negativo

⁶In realtà la casa produttrice del SP451 ha recentemente commercializzato una versione software dello stesso processore sotto forma di plug-in per tre delle più diffuse applicazioni audio professionali.

dell'asse) in quanto la misura viene effettuata con l'asse x del sensore orientato al contrario rispetto alla direzione di propagazione dell'onda. Lo stesso discorso vale per le altre componenti della velocità.

I parametri A_F e A_R assumono di default il valore 0.5, che corrisponde ad un pattern di tipo cardioide, mentre, sempre di default, si ha $\theta_F = 45^\circ$ e $\theta_R = 135^\circ$. Questi valori costituiscono un sicuro punto di partenza per una buona transcodifica, tuttavia il software permette all'utente di variarli a proprio piacimento a seconda delle necessità e situazioni. I segnali così generati vengono salvati come file audio in formato *wav* in una apposita cartella, pronti ad essere masterizzati su DVD per l'utilizzo sui sistemi consumer. Tutte le operazioni vengono eseguite in doppia precisione per garantire la massima accuratezza del risultato.

Il processo appena descritto costituisce uno dei blocchi all'interno di un software che permette all'operatore di eseguire tutte le operazioni che portano alla realizzazione dei file per l'authoring di un DVD a partire da una cartella di file anecoici e una libreria di risposte all'impulso quadrifoniche, il tutto dando in input una sola riga di comando. I blocchi che compongono questo software sono illustrati in figura 4.3. L'utente specifica un file anecoico e una cartella contenente le risposte quadrifoniche all'impulso. Il software provvede ad eseguire le convoluzioni del file anecoico con le risposte all'impulso, creando così le auralizzazioni quadrifoniche, che vengono salvate in una cartella specifica sull'hard disk. Successivamente, lo stesso software accede alla cartella delle auralizzazioni e ne opera la transcodifica nel formato surround, applicando il metodo sopra descritto, tenendo conto di vari parametri, quali pattern, angoli, frequenza di taglio dei filtri, che possono essere definiti dall'utente. I risultati della transcodifica vengono poi salvati su disco in una apposita cartella, come file audio in formato *wav* con dei nomi che ne consentono l'immediata identificazione da parte del software di authoring. Al momento il software è sprovvisto di un'interfaccia grafica, la cui creazione costituisce uno dei possibili sviluppi futuri di questo lavoro. Il diagramma a blocchi dell'intero processo di convoluzione e transcodifica è riportato in figura 4.3. In tale figura viene riportato anche un metodo di transcodifica alternativo, basato sul metodo delle matrici di Gerzon, che verrà discusso nel paragrafo seguente.

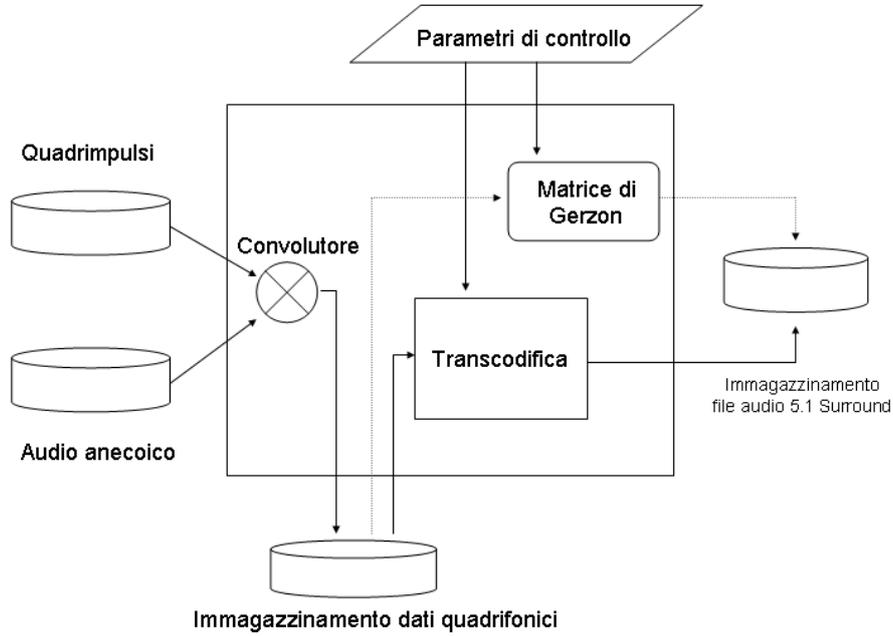


Figura 4.3: Diagramma a blocchi dell'intero processo di convoluzione e transcodifica.

4.2.1 Le Matrici di Gerzon

Abbiamo visto nella sezione 3.2 che la restituzione del B-Format può essere effettuata seguendo criteri psicoacustici per ottimizzare la ricostruzione di parametri diversi del campo in diverse regioni dello spettro. In particolare, abbiamo visto come un criterio di ottimizzazione consiste nel massimizzare il parametro r_V per la ricostruzione del vettore velocità alle basse frequenze ($f \lesssim 700Hz$) e il parametro r_E per la ricostruzione dei flussi di energia alle alte frequenze ($f \gtrsim 700Hz$). Fu Michael Gerzon ad occuparsi del problema di trovare i coefficienti per le combinazioni lineari dei segnali B-Format da inviare agli altoparlanti disposti in varie configurazioni per soddisfare i criteri psicoacustici [25]. Le Matrici di Gerzon per una data configurazione di altoparlanti esprimono i coefficienti di tali combinazioni lineari nelle due regioni dello spettro. Nel caso di un sistema d'ascolto 5.1 Surround, egli trovò per le basse frequenze:

$$\begin{cases} C &= 0.3419W + 0.23322X \\ L &= M_F + S_F \\ R &= M_F - S_F \\ RL &= M_R + S_R \\ RR &= M_R - S_R \end{cases} \quad (4.3)$$

con

$$\begin{cases} M_F &= 0.26813W + 0.38191X \\ S_F &= 0.50527Y \\ M_R &= 0.56092W - 0.49852X \\ S_R &= 0.45666Y \end{cases} \quad (4.4)$$

e per le alte frequenze

$$\begin{cases} C &= 0.38324W + 0.37228X \\ L &= M_F + S_F \\ R &= M_F - S_F \\ RL &= M_R + S_R \\ RR &= M_R - S_R \end{cases} \quad (4.5)$$

con

$$\begin{cases} M_F &= 0.44022W + 0.37228X \\ S_F &= 0.54094Y \\ M_R &= 0.78238W - 0.55322X \\ S_R &= 0.42374Y \end{cases} \quad (4.6)$$

Gerzon propone per il B-Format l'utilizzo di una trasformazione di coordinate che prende il nome di *dominance* e ha la seguente formulazione:

$$\begin{cases} W' &= \frac{1}{2}(\lambda + 1/\lambda)W + \frac{1}{\sqrt{8}}(\lambda - 1/\lambda)X \\ X' &= \frac{1}{2}(\lambda + 1/\lambda)X + \frac{1}{\sqrt{2}}(\lambda - 1/\lambda)W \\ Y' &= Y \\ Z' &= Z \end{cases} \quad (4.7)$$

Si osserva che nel caso di un'onda piana progressiva proveniente dalla direzione dell'asse x con verso positivo, i segnali del B-Format $\{W, X, Y, Z\}$, che stanno in rapporto come i numeri $\{1, \sqrt{2}, 0, 0\}$ ⁷, vengono trasformati nei segnali di un'onda piana progressiva con un fattore di amplificazione λ . Al contrario, nel caso di un'onda piana progressiva proveniente da dietro, per la quale i rapporti relativi sono $\{1, -\sqrt{2}, 0, 0\}$, i segnali vengono trasformati nei segnali di un'onda piana progressiva proveniente sempre da dietro, ma amplificata di un fattore $1/\lambda$, cioè attenuata di un fattore λ . Il guadagno relativo tra i segnali anteriori e posteriori risulta quindi alterato di un fattore λ^2 . Una trasformazione di dominance con $\lambda > 1$ produce quindi un'attenuazione dei segnali posteriori e una amplificazione di quelli anteriori, con l'effetto percettivo di uno zoom del microfono nella direzione anteriore. Il contrario accade con $\lambda < 1$. Per diversi sistemi di riproduzione Gerzon calcolò i valori di λ da inserire nelle trasformazioni di dominance da effettuare prima della transcodifica, per l'ottimizzazione dell'ascolto.

Abbiamo presentato questi dettagli perchè abbiamo scelto di implementare nel nostro software anche la transcodifica basata sul metodo delle matrici di Gerzon applicata ai dati quadrifonici, in alternativa al metodo precedentemente descritto. Per completezza abbiamo implementato anche la dominance, dando all'utente la possibilità di inserire il parametro *dominance gain* D , che è il rapporto λ^2 tra i fattori di amplificazione frontale e posteriore, espresso in decibel tramite la relazione $D = 20\text{Log}(\lambda^2)$. Il valore di default è $D = 0$, che mantiene inalterati i segnali.

4.3 DVD-Audio e DVD-Video Authoring

Per quanto riguarda il riversamento dell'audio multicanale su un supporto, per l'ascolto e la distribuzione, la nostra scelta è caduta quasi obbligatoriamente sul

⁷Ricordiamo che W è attenuato di $3dB$, cioè di un fattore $\sqrt{2}$ rispetto agli altri segnali.

DVD. Il DVD (*Digital Versatile Disc*) è un supporto ottico per la memorizzazione di dati digitali. Attualmente vengono prodotti quattro tipi di supporti che si differenziano per la possibilità di scrivere dati su uno strato singolo o doppio, su un lato solo o su entrambi. La capacità di memoria varia da 4.7GB per il disco a singolo strato su un lato, fino ai 17GB del disco doppio strato scrivibile su entrambi i lati. Vista l'elevata capacità e le dimensioni contenute, pari a quelle del Compact Disc, questo supporto è la soluzione più adatta per la memorizzazione e diffusione di contenuti multimediali quali video digitali e audio multicanale⁸. Il DVD viene applicato nei settori del video, dell'audio e dell'archiviazione dei dati. I formati più utilizzati sono:

- DVD-Video, per la diffusione di film in sostituzione del VHS
- DVD-ROM, per l'archiviazione dei dati in sostituzione del CD-ROM
- DVD-Audio, capace di contenere fino a otto canali di audio non compresso ad alta risoluzione

Il formato DVD-Audio permette la memorizzazione su un DVD di tracce audio offrendo, grazie all'incremento di capacità rispetto al cd, la possibilità di una lunghezza di contenuti maggiore e/o di una qualità audio maggiore (teoricamente); non solo su un DVD-Audio è possibile inserire audio multicanale (fino a otto canali), ma finalmente si può disporre dell'audio ad alta risoluzione a 24bit, con frequenze di campionamento fino a 192kHz, cioè si può avere la stessa risoluzione alla quale viene registrato e prodotto l'audio negli studi professionali. Tracce audio con risoluzione diversa possono coesistere sullo stesso disco. Il DVD-Audio prevede l'utilizzo del formato lineare PCM o del formato MLP (Meridian Lossless Packing) compresso senza perdita di informazione. Inoltre, ad ogni traccia audio possono essere associate fino a 99 immagini, che vengono mostrate ciclicamente dai lettori dotati di uscita video contemporaneamente alla riproduzione della traccia audio. Il formato DVD-Audio, pur presentando rispetto al CD-Audio i vantaggi qualitativi sopra descritti, non ha raggiunto una larga diffusione nel mercato: dati relativi all'anno 2004 riportano uno share di mercato del 2.7% su tutta la musica acquistata negli Stati Uniti.

Ben più diffuso risulta invece il DVD-Video, formato in grado di contenere fino a 240 minuti di video con compressione MPEG-2. L'audio di un DVD-Video può essere in formato Stereo PCM (non compresso), Dolby Digital AC3 (da uno a sei canali), DTS (fino a sette canali) o MPEG-2. I formati di compressione utilizzati per l'audio multicanale nel DVD-Video sono l'AC3 della Dolby e il DTS (*Digital Theatre Surround*), non a caso gli stessi formati che dominano il mercato cinematografico. Si tratta di formati proprietari, che vengono creati a partire dall'audio multicanale in formato lineare mediante appositi software (a volte incorporati in unità hardware dedicate) di terze parti, realizzati su licenza, che effettuano il processo di *encoding*. La successiva fase di riproduzione richiede un chip e un algoritmo di *decoding*, anch'esso soggetto a licenza e disponibile sulla quasi totalità dei lettori presenti sul mercato. Tutta la fase di lavorazione nella post-produzione viene solitamente effettuata utilizzando il formato lineare,

⁸Attualmente, lo sviluppo del video digitale ad alta definizione spinge l'industria verso l'impiego di supporti di capacità ancora maggiore, quali il Blue Ray e il DVD ad alta definizione (HD DVD). Questi formati, tra loro incompatibili, non hanno ancora raggiunto una larga diffusione.

fino a produrre l'audio definitivo nei suoi canali discreti (sei nel caso dei DVD-Video destinati all'Home Theater). A questo punto interviene l'algoritmo di encoding, che produce a partire da questi canali un unico file (con estensione *.ac3* nel caso del Dolby, *.dts* nel caso del DTS) pronto per l'inserimento sul supporto.

Sia nel caso del DVD-Audio che del DVD-Video, la creazione del supporto finale per la distribuzione e il playback richiede l'organizzazione dei file digitali audio e video secondo una struttura ben precisa, che risponda allo standard dei due formati. Pertanto, prima di eseguire la masterizzazione del supporto, è necessario disporre di una specifica applicazione in grado di organizzare i file e creare gli elementi necessari al corretto funzionamento del supporto negli apparecchi conformi allo standard. Questo processo prende il nome di *DVD Authoring* e prevede, tra l'altro, l'indicizzazione dei file, la creazione dei menù e sottomenù e l'eventuale associazione delle immagini alle tracce del DVD-Audio. L'*authoring* viene svolto mediante software ormai piuttosto diffusi e disponibili per tutte le piattaforme, con interfacce e livelli di complessità che variano da prodotti consumer a basso costo ad applicazioni professionali con controllo su tutte le opzioni previste dai formati.

Nel nostro caso, siamo stati in grado di produrre supporti nei formati DVD-Audio e DVD-Video sfruttando due diversi software a basso costo in grado di gestire l'*authoring* e la masterizzazione su pc: *Cirlinca DVD-Audio Solo* per l'*authoring* del DVD-Audio e *Nero 7 Premium* per l'*authoring* DVD-Video. I DVD-Audio da noi realizzati usano il formato audio PCM 16bit/44.1kHz, mentre nel caso dei DVD-Video abbiamo utilizzato, partendo dallo stesso formato PCM, la codifica in formato Dolby AC3.

4.4 Il confronto oggettivo tra evento reale e campo risintetizzato

4.4.1 Proposta oggettiva per il confronto

In questa sezione ci proponiamo di testare la bontà del metodo di transcodifica utilizzato per il playback delle registrazioni quadrifoniche descritto nella sezione 4.2 e di metterne in evidenza le limitazioni.

Dal punto di vista fisico, il sistema ideale di riproduzione dovrebbe essere in grado di restituire correttamente nel punto d'ascolto i segnali di pressione e velocità misurati in fase di registrazione; abbiamo visto infatti che tali dati caratterizzano in maniera completa il campo acustico, pertanto sembrerebbe ovvio concentrare l'analisi su di essi. Naturalmente, essendo alla ricerca di dati oggettivi, dobbiamo basarci su dati misurabili e, come abbiamo visto nel corso di questo lavoro, tutte le caratteristiche fisiche di un campo acustico derivano dai segnali di pressione e velocità.

L'idea alla base del nostro metodo è quella di misurare con la sonda intensimetrica posizionata nel punto d'ascolto il campo prodotto dagli altoparlanti in fase di riproduzione e di confrontare i dati quadrifonici così raccolti con quelli corrispondenti all'evento reale, ai quali sono stati applicati gli algoritmi di transcodifica. A questo punto abbiamo la possibilità di effettuare un confronto tra due set di dati quadrifonici, uno relativo ad un evento sonoro e l'altro relativo alla sua restituzione nell'ambiente d'ascolto. Come in ogni esperimento, ci si

pone il problema di come interpretare i dati raccolti e di quali metodi di analisi portino alle conclusioni più significative. In questo senso abbiamo scartato per il momento l'idea di un confronto diretto dei segnali, magari calcolando la correlazione tra i canali corrispondenti: questo perchè in realtà non sapremmo bene in che termini interpretare un'eventuale plausibile discrepanza e non saremmo in grado di metterla direttamente in relazione con la qualità dell'ascolto. Tale tipo di approccio potrebbe aver senso in una seconda fase, abbinato a dei test psicoacustici d'ascolto, cosa che può costituire uno degli sviluppi di questo lavoro. Essendo l'operazione di transcodifica e riproduzione finalizzata ad offrire un risultato la cui qualità verrà poi valutata soggettivamente da ogni singolo ascoltatore, siamo interessati ad ottenere dalla misura qualche indicatore che possa essere messo in relazione con la nostra percezione.

Naturalmente la cosa più importante di un sistema di registrazione e riproduzione è che vengano mantenute entro una buona approssimazione le caratteristiche timbriche (cioè il contenuto spettrale) dell'evento sonoro. La risposta in frequenza del nostro apparato di misura e di transcodifica garantisce buoni risultati da questo punto di vista.

Il metodo che proponiamo per valutare la fedeltà di restituzione del campo acustico prevede il confronto del vettore intensità sonora $\vec{i}(t)$, che fornisce la direzione del flusso di energia, e dell'indicatore η , che misura la frazione di energia irradiata attraverso il punto di misura.

L'intensità sonora risulta utile per valutare la bontà della ricostruzione spaziale, infatti la sua direzione fornisce la posizione apparente della sorgente nell'immagine sonora. Nel nostro caso, avendo a che fare con un sistema di riproduzione bidimensionale, andremo a valutare le proiezioni del vettore intensità istantanea sul piano orizzontale xy per tutta la durata di un evento sonoro, mediante l'analisi dei polar plot.

L'indicatore η si presta a fornire un'indicazione di quanto vengano rispettate le proprietà energetiche del campo originale, ovvero di quanta energia venga irradiata e quanta invece rimanga localizzata attorno al punto di misura.

4.4.2 Audio quality test

Per gli ascolti delle registrazioni quadrifoniche e delle convoluzioni abbiamo allestito un impianto di diffusione conforme allo standard 5.1 surround dell'Home Theater in due ambienti dalle diverse caratteristiche acustiche: un ampio salone moderatamente riverberante e una camera anecoica. Il sistema usato comprende cinque casse amplificate professionali Genelec due vie modello 1030A, con woofer da $17cm$ e tweeter da $2.5cm$, la cui risposta in frequenza si estende da $50Hz$ a $20kHz$, collocate su piedistalli metallici in modo da avere il tweeter all'altezza delle orecchie di un ascoltatore seduto. Per il canale LFE degli effetti a bassa frequenza è stato utilizzato un subwoofer Genelec modello 7060A. La sorgente usata è un lettore DVD multi-formato Samsung modello DVD-HD950. Un attenuatore multicanale SPL modello 2608 è stato interposto tra l'uscita del DVD player e le casse, per ottimizzare il volume d'ascolto e bilanciare il livello medio di pressione sonora tra registrazioni con un diverso livello medio.

Nel salone gli altoparlanti sono stati collocati ad una distanza di $2.5m$ dal punto d'ascolto; questa condizione permette un adeguato sviluppo del fronte d'onda proveniente dagli altoparlanti, che può essere considerato in buona ap-

prossimazione piano per frequenze superiori ai $350Hz$ ⁹. In questo ambiente il suono emesso dalle casse viene riflesso dalle superfici e ritorna all'ascoltatore sotto forma di campo riverberato. Le dimensioni del locale, quindi le distanze degli altoparlanti dalle pareti, sono tali da garantire che le prime riflessioni non interferiscano significativamente col suono diretto, assicurando così l'assenza di fenomeni di colorazione, comb-filtering e distorsione dell'immagine. Le riflessioni dell'ambiente, attenuate e ritardate, vengono percepite dall'ascoltatore come riverberazione e non influiscono significativamente sulla percezione dell'immagine sonora, essendo questa principalmente determinata dai primi fronti d'onda che raggiungono l'ascoltatore. Ci si aspetta tuttavia che il campo ricostruito dagli altoparlanti in questo ambiente sia caratterizzato da un indicatore di radiazione η più basso. Infatti, l'energia irradiata dagli altoparlanti, che ha le caratteristiche proprie dell'ambiente in cui è stata misurata, viene ulteriormente riflessa dalle pareti, di conseguenza una parte comparirà come energia oscillante localizzata nei modi propri dell'ambiente d'ascolto. Dal punto di vista dell'analisi del sistema, la ridiffusione di una registrazione in un ambiente equivale ad una ulteriore convoluzione con la risposta all'impulso dell'ambiente d'ascolto al quale corrisponde, in questo caso, un indicatore di radiazione minore di uno. In questo ambiente la diffusione è stata effettuata durante il testbed del progetto *IP-Racine* col solo scopo di presentare l'esperienza d'ascolto, di conseguenza non sono disponibili le misure relative alla diffusione.

L'altro ambiente utilizzato per la diffusione e per le misure è la camera fonoassorbente del Laboratorio di Acustica Musicale e Architettonica FSSG-CNR, mostrata in figura 4.4. La camera, di forma parallelepipedica, ha dimensioni $4.7m \times 3.1m \times 2.8m$, per un volume di $37m^3$ ed una superficie totale di $70m^2$. L'assorbimento sonoro è effettuato per mezzo di pannelli di *Fiber form*[®] di dimensioni $1.2m \times 0.6m \times 0.1m$ e densità di $40kg/m^3$. Il tempo di riverberazione $RT60$ misurato nella camera risulta inferiore a $0.1s$, con un corrispondente coefficiente medio di assorbimento $\bar{\alpha} \approx 0.75$. In questo ambiente gli altoparlanti sono stati disposti su una circonferenza di raggio $1.30m$. L'assenza di riverberazione e riflessioni garantita dal materiale fonoassorbente alle frequenze medio-alte rappresenta la condizione ideale per verificare la bontà dei metodi di riproduzione delle registrazioni quadrifoniche mediante confronto con l'evento reale, in quanto possiamo assumere che l'ambiente d'ascolto non apporti alcun contributo al campo sonoro emesso dagli altoparlanti.

Per effettuare le misure abbiamo posto la sonda intensimetrica Microflown USP nella posizione ideale d'ascolto, cioè al centro del cerchio delimitato dagli altoparlanti, con l'asse x diretto verso l'altoparlante centrale. Grazie ad un'interfaccia audio multicanale abbiamo configurato il pc con software Audition per inviare i sei segnali dell'audio 5.1 surround ai rispettivi altoparlanti, misurando simultaneamente su quattro tracce i segnali provenienti dalla sonda intensimetrica. Il confronto è stato poi effettuato con i dati quadrifonici dai quali siamo partiti per generare i sei canali surround.

Il primo test consiste in una verifica della correttezza della restituzione spaziale del campo sonoro. Per fare questo, abbiamo preso come sorgente uno degli altoparlanti usati nel sistema 5.1 e lo abbiamo posizionato in vari punti sulla circonferenza sulla quale sono disposti gli altoparlanti. Abbiamo poi inviato al-

⁹Si assume che il fronte d'onda possa essere approssimato a quello di un'onda piana quando vale la relazione $d \gtrsim 3\lambda$, dove d è la distanza da una sorgente puntiforme e λ è la lunghezza d'onda del suono emesso.



Figura 4.4: Il sistema di diffusione 5.1 nella camera anecoica.

l'altoparlante uno sweep sinusoidale di frequenza variabile da $100Hz$ a $20kHz$ e abbiamo misurato il segnale al centro. Tale segnale è stato poi transcodificato in surround col nostro software e ridiffuso dal sistema di altoparlanti, misurando di nuovo il campo prodotto al centro. Ricordiamo che nel nostro sistema gli altoparlanti sono collocati a 0° , $\pm 30^\circ$, $\pm 120^\circ$. Essendo in camera quasi anecoica, ci si aspetta un vettore intensità avente la direzione della congiungente tra la sorgente e il centro del cerchio e verso opposto alla sorgente. La figura 4.5 mostra i risultati ottenuti quando la sorgente è collocata in posizione centrale: la simmetria del sistema e l'algoritmo di conversione mantengono la direzione dell'intensità. Un leggero sparpagliamento del vettore intensità del campo risintetizzato è dovuto al fatto che un maggior numero di altoparlanti è coinvolto nella riproduzione del suono: infatti, per quanto visto, un suono proveniente dalla direzione frontale genera segnale in fase di playback anche nei canali sinistro e destro. Nel caso della restituzione di una sorgente collocata 40° a sinistra (figura 4.6) vediamo invece come il sistema, oltre a produrre un allargamento nel plot del vettore intensità, provochi una leggera rotazione del vettore intensità verso il centro. Ciò è dovuto al fatto che anche l'altoparlante centrale emette suono in maniera considerevole durante la riproduzione, perchè il segnale che lo pilota è ricavato da un microfono virtuale di pattern cardioide diretto verso il centro, che evidentemente non attenua a sufficienza un'onda piana proveniente da 40° a sinistra. L'assenza, rispetto al caso precedente, di un altoparlante simmetrico al centrale rispetto alla direzione della sorgente, determina la rotazione dell'immagine. L'effetto è ancor più evidente con una sorgente posta 60° a sinistra, che, oltre a subire un allargamento leggermente maggiore, tende ad essere ruotata in

modo che la sua direzione coincida con quella dell'altoparlante frontale sinistro (figura 4.7). La figura 4.8 mostra come si comporta il sistema per una sorgente collocata a 110° sulla sinistra. Anche in questo caso notiamo l'allargamento e la tendenza a ruotare l'immagine della sorgente verso il centro del sistema, dove è presente un numero maggiore di altoparlanti. Infine, la figura 4.9 mostra il campo ricostruito a partire da una sorgente collocata dietro l'ascoltatore: questo è il caso in cui il sistema 5.1 si comporta nel modo peggiore, non riuscendo a ricostruire correttamente l'intensità sonora. La mancanza di un altoparlante collocato dietro l'ascoltatore e l'eccessiva separazione tra i due canali posteriori impedisce la corretta ricostruzione del campo. Pur non avendo effettuato dei veri e propri test psicoacustici, in questo caso l'impressione dell'ascoltatore è quella di essere circondato dal suono e di non riuscire ad identificare la posizione della sorgente. Più che all'algoritmo di transcodifica, l'inaccuratezza della restituzione spaziale delle sorgenti posteriori è un limite intrinseco del sistema 5.1, dovuto al fatto che gli altoparlanti posteriori sono troppo separati; questo causa, all'ascolto, un effetto di "hole in the middle" posteriore.

La tabella 4.2 riporta i valori dell'indicatore η per il campo reale e quello risintetizzato. Nelle prime quattro posizioni i valori di η risultano compatibili, mentre nel caso della sorgente posta dietro l'ascoltatore η si abbassa considerevolmente, a riprova del fatto che il campo ricostruito assomiglia di più ad un campo diffuso che ad un'onda piana progressiva.

	front	40° Left	60° Left	110° Left	180° Rear
η evento reale	0.88	0.85	0.82	0.74	0.75
η campo risintetizzato	0.83	0.81	0.80	0.73	0.17

Tabella 4.2: Confronto dei valori dell'indicatore η per il campo reale e per quello risintetizzato dagli altoparlanti per diverse posizioni della sorgente sonora.

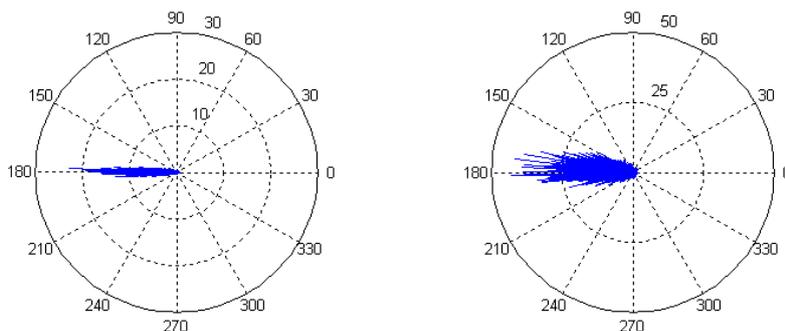


Figura 4.5: Sorgente in posizione frontale. A sinistra: campo reale; a destra: campo risintetizzato.

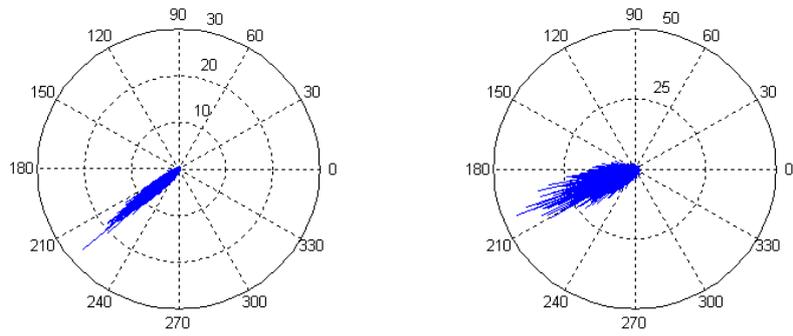


Figura 4.6: Sorgente in posizione 40° Left. A sinistra: campo reale; a destra: campo risintetizzato.

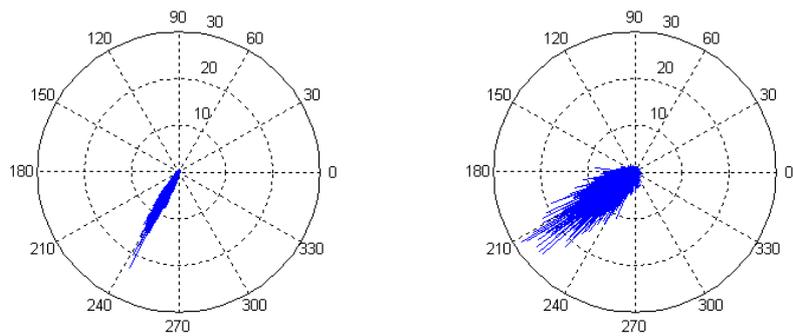


Figura 4.7: Sorgente in posizione 60° Left. A sinistra: campo reale; a destra: campo risintetizzato.

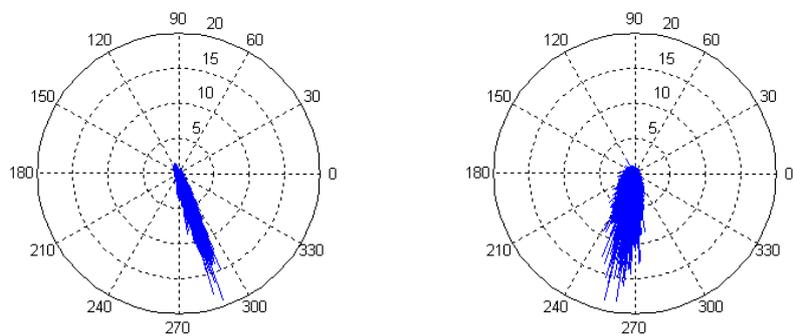


Figura 4.8: Sorgente in posizione 110° Left; A sinistra: campo reale; a destra: campo risintetizzato.

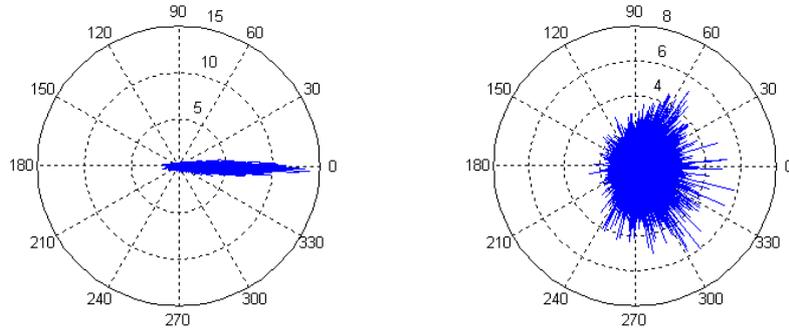


Figura 4.9: Sorgente in posizione 180° . A sinistra: campo reale; a destra: campo risintetizzato.

Vogliamo ora effettuare il confronto tra dati quadrifonici ottenuti mediante convoluzione di un segnale anecoico con le risposte all'impulso da noi misurate e il campo ricostruito dal sistema d'ascolto, per testare come il sistema 5.1 sia in grado di restituire il campo sonoro presente in un ambiente reale. Anche in questo caso siamo partiti da un parlato anecoico, che è stato convoluto con le risposte quadrifoniche all'impulso, transcodificato in 5.1, riprodotto in camera anecoica e contemporaneamente registrato quadrifonicamente.

La figura 4.10 mostra il confronto tra i vettori intensità del campo ottenuto mediante convoluzione quadrifonica con la risposta all'impulso misurata nel Dormitorio del Buora in posizione 1 e del campo ricostruito dal sistema 5.1: la direzionalità del campo viene rispettata con buona approssimazione. Calcolando i valori di η , riportati in tabella 4.3, si osserva un incremento nel valore corrispondente al campo ricostruito dagli altoparlanti, interpretabile come un difetto del sistema 5.1 nel restituire un campo caratterizzato da energia sonora prevalentemente oscillante. Questo effetto si può anche interpretare ammettendo che il campo ricostruito da un sistema convenzionale di altoparlanti ad una distanza superiore a poche lunghezze d'onda sia approssimabile ad una sovrapposizione di onde piane, caratterizzate da un valore di η prossimo ad uno, e che il sistema tenda a "tradurre" in onda piana progressiva anche un campo divergente o diffuso, perlomeno fintantochè la direzione di provenienza del suono è localizzata prevalentemente in avanti. Le stesse considerazioni fatte riguardo l'esempio del Dormitorio del Buora in posizione 1 valgono per il campo misurato nella Sala Carnelutti in posizione 8, mostrato in figura 4.11. In questo caso si osserva anche l'effetto precedentemente discusso di rotazione del campo verso l'altoparlante più vicino. Vogliamo ora discutere i casi di convoluzione con risposte misurate in direzione posteriore, come quelle misurate nel Dormitorio del Buora in posizione 4 e nella Sala Carnelutti in posizione 4, e i confronti con i campi ricostruiti in 5.1 mostrati rispettivamente nelle figure 4.12 e 4.13. Avendo già constatato l'incapacità del nostro sistema di ricostruire un'onda piana proveniente dalla direzione posteriore, ci aspettiamo in questi casi la restituzione di un campo diffuso, caratterizzato da un indice di radiazione basso e da vettori intensità distribuiti in tutte le direzioni. L'analisi effettuata conferma questa ipotesi e mostra anche che, nel caso di figura 4.12, il sistema d'ascolto

può addirittura invertire la direzione dell'intensità media, dando luogo ad un suono vagamente localizzato in avanti.

	Buora 1	Buora 4	Carnelutti 8	Carnelutti 4
η convoluzione quad.	0.45	0.32	0.34	0.56
η 5.1 playback	0.69	0.28	0.65	0.12
σ convoluzione quad.	0.99	0.99	0.99	0.99
σ playback	0.99	0.97	0.99	0.98

Tabella 4.3: Confronto degli indicatori η e σ nel caso di convoluzione quadrifonica e riproduzione mediante sistema 5.1.

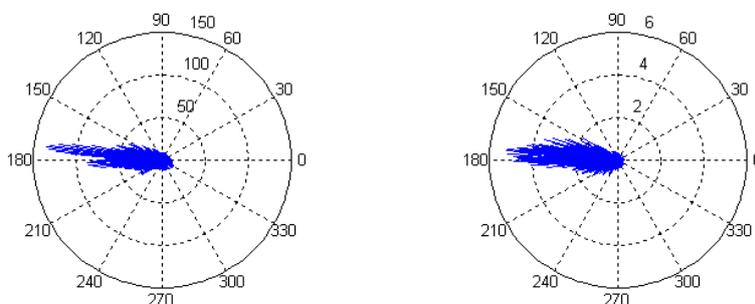


Figura 4.10: Confronto del vettore intensità, convoluzione con la risposta all'impulso misurata nel Dormitorio del Buora in posizione 1. A sinistra: vettore intensità del campo ottenuto mediante convoluzione. A destra: vettore intensità del campo risintetizzato.

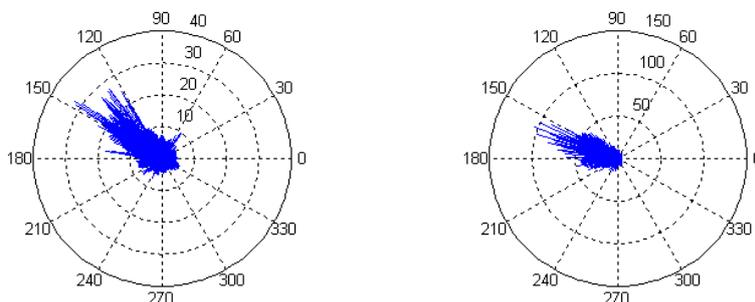


Figura 4.11: Confronto del vettore intensità, convoluzione con la risposta all'impulso misurata nella Sala Carnelutti in posizione 8. A sinistra: vettore intensità del campo ottenuto mediante convoluzione. A destra: vettore intensità del campo risintetizzato.

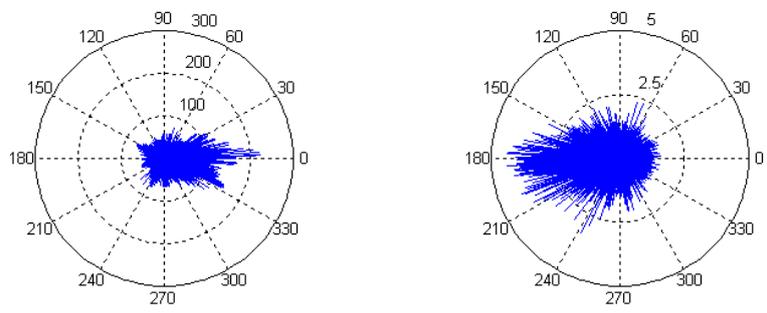


Figura 4.12: Confronto del vettore intensità, convoluzione con la risposta all'impulso misurata nel Dormitorio del Buora in posizione 4. A sinistra: vettore intensità del campo ottenuto mediante convoluzione. A destra: vettore intensità del campo risintetizzato.

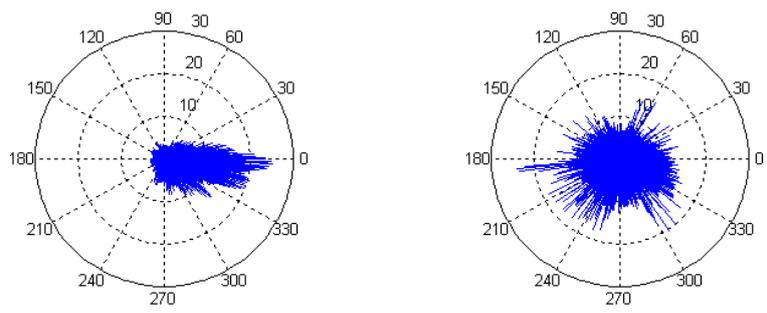


Figura 4.13: Confronto del vettore intensità, convoluzione con la risposta all'impulso misurata nella Sala Carnelutti in posizione 4. A sinistra: vettore intensità del campo ottenuto mediante convoluzione. A destra: vettore intensità del campo risintetizzato.

Conclusioni e sviluppi futuri

Conclusioni

In questo lavoro, partendo dall'analisi fisica del campo acustico, è stato definito l'evento sonoro mediante un set completo di dati, chiamati dati quadrifonici, che lo identificano dal punto di vista energetico. Inoltre abbiamo visto come sia possibile scomporre un evento sonoro nel contributo della sorgente ed in quello dell'ambiente. Grazie all'innovazione fornita dall'invenzione dei sensori per la misura diretta della velocità acustica dell'aria, le tecniche da noi utilizzate permettono sia di registrare l'evento sonoro che di misurare il contributo dell'ambiente, mediante la tecnica di misura della risposta quadrifonica all'impulso.

I dati sperimentali che abbiamo presentato convalidano i fondamenti teorici e dimostrano come la Quadrifonia Acustica risulti un metodo accurato per la misura, la sintesi e la ricostruzione di un evento sonoro, con tutte le sue proprietà energetiche e spaziali.

L'apparato sperimentale allestito si è dimostrato valido e adatto all'impiego in vari settori dell'acustica, da quello delle misure in ambito architettonico a quello della virtualizzazione di eventi sonori, al test di fedeltà di un sistema d'ascolto multicanale. Sottolineiamo che la tecnologia ed i suoi metodi di utilizzo proposti sono già all'attenzione dell'industria cinematografica, nella quale pare possano trovare un utilizzo concreto largamente diffuso grazie alla possibilità di integrarsi con il *workflow* dell'audio cinematografico. Resta ovviamente del lavoro da fare per il perfezionamento delle tecniche e per dare il via ad ulteriori sviluppi, cui verrà brevemente fatto cenno nella sezione successiva, tuttavia il fatto di avere finalmente un impianto che poggia su solide basi teoriche costituisce un'ottima premessa per il raggiungimento di importanti risultati.

Sviluppi futuri

Nel corso del lavoro è stato fatto cenno ad alcune delle potenzialità ancora inesplorate della Quadrifonia Acustica e ad aspetti che richiedono un ulteriore perfezionamento. I suoi sviluppi futuri pertanto riguardano sia la ricerca di nuove applicazioni che l'ottimizzazione delle procedure attualmente in uso.

Come si è visto, il processo di convoluzione quadrifonica si presta naturalmente all'utilizzo in fase di post-produzione cinematografica. Per facilitarne l'utilizzo e l'integrazione col flusso di produzione, sarebbe ottimale sviluppare un software con interfaccia grafica per la convoluzione, attraverso il quale l'operatore possa selezionare le risposte quadrifoniche all'impulso da una libreria

e abbia la possibilità di implementare con semplici controlli ulteriori operazioni quali, ad esempio, la rotazione dell'immagine sonora o l'aggiunta di segnale anecoico in proporzioni variabili per aumentare la definizione e la chiarezza di un dialogo mantenendo le caratteristiche di riverberazione di un particolare ambiente, e così via. La scelta a nostro parere più logica sarebbe di proporre tale software come *plug-in* multicanale in formato *TDM* e *VST*, in modo da renderlo direttamente disponibile all'interno delle applicazioni audio-video che attualmente rappresentano lo stato dell'arte per la produzione audio, quali *ProTools*, *Nuendo* e *SADiE*. Un'ulteriore funzionalità del *plug-in* potrebbe essere quella di operare la transcodifica dei dati quadrifonici nei formati stereo e surround.

Come si evince dai risultati presentati nel capitolo 4, il "collo di bottiglia" dell'intero processo è al momento rappresentato dal fatto di dover utilizzare un sistema 5.1 Surround per l'ascolto delle registrazioni quadrifoniche, che consente un limitato controllo sull'accuratezza dei risultati ottenuti. Un primo passo per il miglioramento di questo aspetto consiste nell'ottimizzazione dei parametri dell'algoritmo di transcodifica, quali i coefficienti che determinano il pattern e gli angoli dei microfoni virtuali, la frequenza di taglio e la pendenza dei filtri usati; inoltre, ci si potrebbe dedicare a cercare diverse configurazioni di altoparlanti (ad esempio 6.1 Surround o array di altoparlanti collocati ai vertici di un poligono regolare), ottimizzare i parametri di transcodifica per queste nuove configurazioni e condurre esperimenti per quantificare i loro pregi e difetti rispetto al sistema attuale. Tutte queste operazioni andrebbero affiancate a test psicoacustici d'ascolto effettuati su un campione di popolazione "allenata" nel campo della musica e dell'audio professionale.

Il passo innovativo per quel che riguarda la restituzione di una registrazione quadrifonica è rappresentato dalla realizzazione di un sistema di altoparlanti configurati e pilotati in modo da riprodurre correttamente i campi di pressione e velocità con le corrette relazioni di ampiezza e di fase, in grado cioè di restituire correttamente gli indicatori energetici per qualsiasi situazione di campo, dall'onda piana progressiva a quella stazionaria. Al momento lo studio di tale applicazione è in fase di completamento nell'ambito del progetto *IP-Racine* e ha portato all'implementazione di un controllo intensimetrico unidimensionale che può anche essere applicato al sistema 5.1 surround per l'ottimizzazione dei parametri energetici. Lo sviluppo di un sistema di diffusione quadrifonica¹⁰ può essere destinato all'affermazione di nuovi standard per gli impianti audio delle sale cinematografiche. Segnaliamo che è già stata avviata una collaborazione attiva con la *DTS* per la codifica dell'audio quadrifonico.

Altre applicazioni in fase di sviluppo riguardano la convoluzione quadrifonica di una sorgente estesa e la de-riverberazione del segnale audio di pressione. Per quanto riguarda la prima applicazione, si studierà il caso particolare di un coro mediante la registrazione anecoica della pressione sonora in vari punti vicino alla sorgente e la successiva convoluzione con risposte all'impulso misurate in diversi ambienti, disponendo la sorgente dodecaedrica in modo da rispecchiare la disposizione dei punti di misura. La seconda applicazione riguarda invece la codifica di algoritmi per la de-riverberazione del segnale audio. Questo importante argomento di ricerca applicata permetterebbe, nel caso della Quadrifonia

¹⁰Sottolineiamo di nuovo che l'aggettivo quadrifonico qui usato non ha nulla a che vedere con l'ascolto mediante i sistemi di quattro altoparlanti disposti ai vertici di un quadrato, il cui insuccesso fu decretato più di due decenni orsono.

Acustica, di ottenere il segnale pseudo-anechoico di pressione senza effettuare fisicamente registrazioni in camera anecoica [29]. Queste ricerche risulterebbero infine utili anche per l'impiego in applicazioni quali le teleconferenze, i sistemi di diffusione in ambienti pubblici (*public addressing* e *sound reinforcement*) e le protesi acustiche.

Appendice A - Segnali e sistemi lineari

Un *segnale* descrive la variazione di un parametro in funzione di un altro. Nel presente lavoro, i segnali trattati rappresentano la variazione nel tempo di grandezze fisiche quali, ad esempio, la pressione acustica dell'aria in un punto, la velocità acustica delle particelle d'aria o l'intensità acustica. Questi segnali sono dunque rappresentati da funzioni matematiche dalle quali possiamo calcolare particolari proprietà dei nostri segnali o ricavare nuovi segnali ad esse legati.

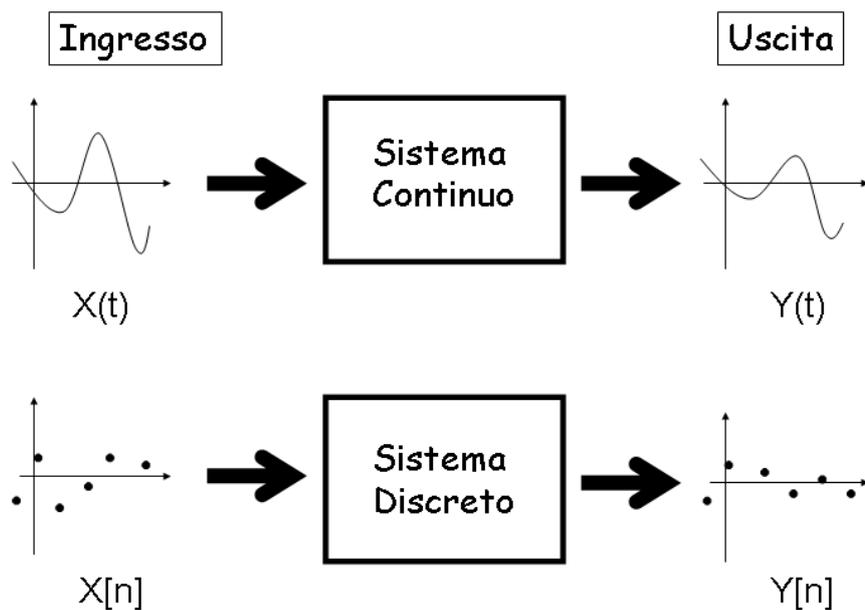


Figura 4.14: Rappresentazioni continua e discreta dei segnali e relativi sistemi.

I segnali possono essere continui o discreti. La maggior parte delle grandezze presenti in natura è rappresentata da segnali continui, tuttavia spesso si ricorre ad un campionamento delle variabili continue per permettere il loro trattamento

nel dominio digitale: in questo caso i segnali sono definiti solo per valori discreti numerabili della variabile indipendente.

Per *sistema* si intende qualunque processo in grado di produrre un segnale d'uscita in risposta ad un segnale di ingresso. Sistemi continui hanno segnali continui in ingresso e uscita, sistemi discreti agiscono su segnali discreti. La figura 4.14 riassume e schematizza questa classificazione di segnali e sistemi.

Un sistema viene definito *lineare* se gode delle proprietà di *omogeneità*, *additività* e *invarianza temporale*.

Omogeneità significa che un cambiamento in ampiezza del segnale in ingresso si traduce in un proporzionale cambiamento d'ampiezza del segnale d'uscita; ciò si riassume nella seguente formulazione matematica:

$$x(t) \rightarrow y(t) \implies kx(t) \rightarrow ky(t)$$

La formulazione deve valere per ogni segnale in ingresso e ogni costante k .

Consideriamo un sistema tale per cui si abbia $x_1(t) \rightarrow y_1(t)$ e $x_2(t) \rightarrow y_2(t)$. Il sistema gode della proprietà additiva se risponde ad un segnale di ingresso del tipo $x_1(t) + x_2(t)$ con un'uscita $y_1(t) + y_2(t)$.

La proprietà di invarianza temporale fa sì che un ritardo nel segnale in ingresso si traduca in un uguale ritardo sul segnale in uscita:

$$x(t) \rightarrow y(t) \implies x(t+s) \rightarrow y(t+s)$$

Questo equivale a dire che le proprietà del sistema non cambiano nel tempo.

Queste tre proprietà si possono riassumere nella seguente formulazione:

$$\begin{cases} x_1(t) \rightarrow y_1(t) \\ x_2(t) \rightarrow y_2(t) \end{cases} \implies k_1x_1(t+s) + k_2x_2(t+s) \rightarrow k_1y_1(t+s) + k_2y_2(t+s)$$

Un'importante caratteristica dei sistemi lineari è il fatto che, in corrispondenza di un segnale sinusoidale in ingresso, essi producono in uscita un segnale sinusoidale avente la stessa frequenza. Le sinusoidi in ingresso e uscita possono naturalmente differire per ampiezza e fase; un sistema lineare è pertanto descritto in maniera completa se vengono specificate le sue risposte in frequenza e in fase.

La linearità gode della proprietà commutativa rispetto all'azione combinata di più sistemi su un segnale. Questo significa che se due sistemi lineari vengono applicati in serie, cioè il segnale all'uscita del primo viene mandato all'ingresso del secondo, l'ordine in cui vengono applicati i due sistemi non cambia il risultato complessivo che producono.

La conseguenza più importante delle proprietà di linearità di un sistema è il fatto che la risposta ad un segnale complesso è esprimibile mediante la somma di risposte a segnali più semplici che sommati compongano il segnale di partenza. Le due scomposizioni più usate per lo studio dei sistemi lineari sono la scomposizione in segnali sinusoidali mediante l'analisi di Fourier e la scomposizione in impulsi.

La scomposizione in segnali sinusoidali si effettua per mezzo della trasformata di Fourier, che permette di esprimere una funzione $f(t)$ per mezzo delle sue componenti spettrali $F(\omega)$. Le relazioni che legano le due funzioni sono

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(\omega)e^{i\omega t} d\omega$$

Nel caso di segnali e sistemi discreti si ricorre alla trasformata di Fourier discreta *DFT*, spesso nella sua implementazione nota come *Fast Fourier Transform FFT*.

Nel caso di segnali discreti, è vantaggioso utilizzare la scomposizione in impulsi, come illustrato in figura 4.15. Per un segnale $y[n]$ si può scrivere

$$y[n] = \sum k_n \delta[n]$$

dove $\delta[n]$ è il segnale discreto in cui il campione n assume valore 1 e tutti gli altri campioni assumono valore nullo.

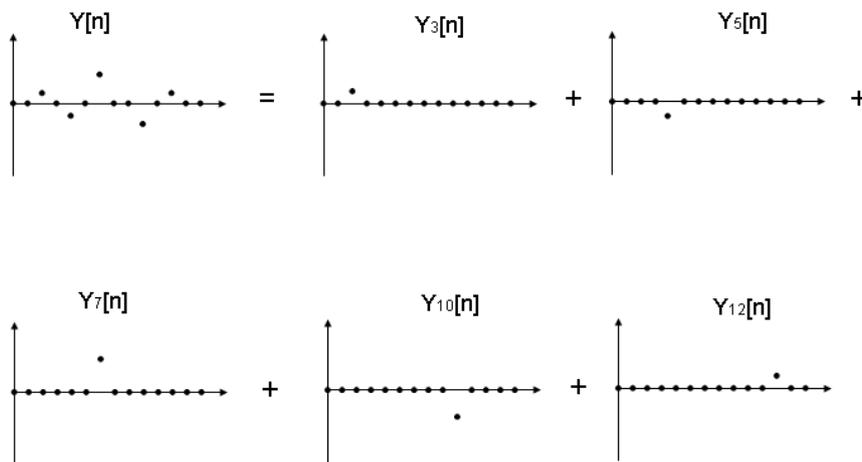


Figura 4.15: Un segnale discreto può essere scomposto in una successione di impulsi.

Conoscendo la risposta all'impulso $h[n]$ di un sistema, ovvero il segnale che il sistema produce in risposta ad un segnale $\delta[0]$, è possibile esprimere la risposta

del sistema ad un qualsiasi segnale $x[n]$, composto da M campioni, come [30]

$$y[n] = \sum_{j=0}^{M-1} h[j]x[n-j]$$

Questa relazione esprime l'operazione di convoluzione tra i due segnali h e x e si indica con $y = h \otimes x$.

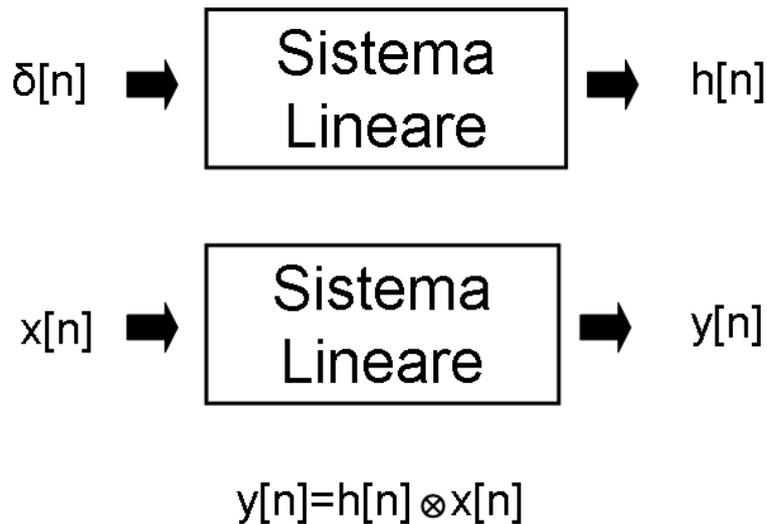


Figura 4.16: Il segnale in uscita da un sistema lineare è esprimibile come convoluzione del segnale in ingresso con la risposta all'impulso del sistema.

Nel caso continuo, la convoluzione si scrive come

$$y(t) = (h \otimes x)(t) = \int_{-\infty}^t h(\tau)x(t-\tau)d\tau$$

Il teorema della convoluzione afferma che il prodotto di convoluzione di due segnali $h(t)$ e $x(t)$ può essere espresso, nel dominio della frequenza, come prodotto algebrico complesso delle componenti di Fourier dei due segnali:

$$y(t) = (h \otimes x)(t) \Rightarrow Y(\omega) = H(\omega) \cdot X(\omega)$$

Grazie a questa relazione è possibile effettuare il prodotto di convoluzione di due segnali passando al dominio della frequenza mediante la trasformata di Fourier e poi ritornando al dominio del tempo mediante la trasformata di Fourier inversa:

$$y(t) = IFT[Y(\omega)] = IFT[H(\omega) \cdot X(\omega)] = IFT[FT[h(t)] \cdot FT[x(t)]]$$

Appendice B - Le grandezze acustiche

La *pressione sonora* è la differenza istantanea tra la pressione di una particella acustica in presenza di un'onda sonora e la pressione atmosferica di equilibrio. L'unità di misura della pressione è il Pascal $1Pa = 1N/m^2$. La pressione di equilibrio è $p_0 = 1atm = 101325Pa \sim 1 \times 10^5 Pa$. In presenza di una perturbazione di pressione dovuta ad un'onda sonora, si definisce il valore quadratico medio *rms* della pressione come $p = \sqrt{\langle p^2(t) \rangle}$. Il valore $p_0 = 20\mu Pa$ *rms* corrisponde allo stimolo sonoro più debole che l'orecchio umano sia in grado di percepire¹¹. Esso viene pertanto preso come valore di riferimento per esprimere l'intensità sonora.

Il *livello di pressione sonora* L_p è la misura, espressa in scala logaritmica, del rapporto tra la pressione rms di un suono e il valore di riferimento p_0 , dato dalla formula

$$L_p(dB SPL) = 10 \text{Log} \left(\frac{p^2}{p_0^2} \right) = 20 \text{Log} \left(\frac{p}{p_0} \right) \quad (4.8)$$

L'unità di misura di L_p è il *dB SPL* (*deciBel Sound Pressure Level*). Ad un suono alla soglia di udibilità corrisponde un livello di *0dB SPL*, mentre un suono alla soglia del dolore, corrispondente ad una pressione sonora di $20Pa$, ha un livello di *120dB SPL*. Il nostro orecchio risulta quindi in grado di percepire fenomeni sonori entro una variazione di ben sei ordini di grandezza della pressione sonora.

Analogamente alla pressione sonora, anche alla *velocità* del fluido si può associare un livello L_v espresso in scala logaritmica, con la relazione

$$L_v = 20 \text{Log} \left(\frac{v}{v_0} \right)$$

dove $v_0 = 5.0 \cdot 10^{-8} m/s$ è il valore *rms* di velocità in corrispondenza di un'onda piana alla soglia di udibilità.

La *velocità di propagazione del suono* in un gas è data da

$$c = \sqrt{\gamma \frac{p_0}{\rho_0}} \quad (4.9)$$

¹¹Il valore, ricavato da una media statistica, è riferito ad una frequenza di *1kHz*.

dove γ è il rapporto tra il calore specifico a pressione costante C_P e quello a volume costante C_V , p_0 è la pressione di equilibrio e ρ_0 la densità di equilibrio. Per l'aria la velocità di propagazione del suono in condizioni standard di pressione e temperatura è 343m/s .

L'*intensità sonora* $\vec{I}(\vec{x}) = \langle p(\vec{x}, t)\vec{v}(\vec{x}, t) \rangle$, dove le parentesi triangolari indicano l'operazione definita dalla relazione 1.59, è interpretabile come una potenza sonora per unità di area nella direzione delle linee di flusso energetiche e ha come unità di misura il W/m^2 .

La *densità di energia sonora* w esprime l'energia sonora per unità di volume e si misura in $\text{J/m}^3 = \text{W} \cdot \text{s/m}^3$.

L'impedenza acustica di un campo sonoro $Z(\omega)$ è il rapporto complesso delle componenti di Fourier della pressione e della velocità. Essa si misura in $\text{Pa} \cdot \text{s/m}$. Si distingue dall'*impedenza caratteristica* Z_0 , che è data da prodotto della densità dell'aria all'equilibrio e della velocità di propagazione del suono: $Z_0 = \rho_0 c$. Si può dimostrare che, per un campo di onde piane che si propagano in un gas, l'impedenza acustica non dipende dalla frequenza e coincide con l'impedenza caratteristica del gas. L'impedenza caratteristica dell'aria, in condizioni standard di pressione e temperatura, vale $Z_0 = 410\text{Pa} \cdot \text{s/m}$.

Bibliografia

- [1] L. Beranek: Acoustics. Acoustic Society of America 1954.
- [2] E. Torick: Highlights in the history of multichannel sound. *Journal of the Audio Engineering Society* 46 1/2 (1998) 27-31.
- [3] D. Stanzial. Appunti di acustica energetica. Venezia, dicembre 2006.
- [4] M. S. Howe: Theory of vortex sound. Cambridge University Press 2003.
- [5] L. D. Landau, E. M. Lifshitz: Fluid mechanics, Butterworth-Heinemann, 1987.
- [6] R. P. Feynman, R. B. Leighton, M. Sands: The Feynman lectures on physics. Addison - Wesley Pub. Company, 1977.
- [7] E. Meyer, E. G. Neumann: Physical and applied acoustics: an introduction. Academic Press, 1972.
- [8] D. Stanzial, D. Bonsi, G. Schiffrer: Four-dimensional treatment of linear acoustic fields and radiation pressure. *Acta Acustica united with Acustica* 89 (2002) 213-224.
- [9] F. J. Fahy: Sound intensity, Second Edition. E&FN SPON (1995).
- [10] J. A. Mann, J. Tichy: Acoustic intensity analysis: distinguishing energy propagation and waveform propagation. *Journal of the Acoustic Society of America* 90 (1) (1991) 20-25.
- [11] D. Stanzial, D. Bonsi, N. Prodi: Measurement of new energetic parameters for the objective characterization of an opera house. *Journal of Sound and Vibration* 232 (1) (2000), 193-211.
- [12] Equipment for collection and measurement of quadrasonic sound data and metadata and relative procedure. Domanda di Brevetto Europeo n°04425909.1 depositata in data 02/12/2003 Riferimento: C.N.R./VE Titolare: Fondazione Scuola di San Giorgio, Venezia.
- [13] D. Bonsi: Theoretical and experimental study of energetic properties of confined sound fields. Doctoral Thesis, University of Ferrara, 1997/98.
- [14] H. E. de Bree et al.: The Microflow; a novel device measuring acoustical flows. *Sensors and actuators: A, physical*, volume SNA 054/1-3 pp. 552-557, 1996.

- [15] D. Stanzial, D. Bonsi: Calibration of the p-v Microflown probe and some considerations on the physical nature of sound impedance. Euronoise, Naples 2003. Paper ID: 149.
- [16] A. Farina: Simultaneous measurement of impulse response and distortion with a swept-sine technique. Presented at the 108th convention of the Audio Engineering Society, February 2000, Paris.
- [17] <http://www.ipracine.org>
- [18] <http://www.settoreweb.com/camera-ve/index.php>
- [19] D. Howard, L. Moretti: Architettura e musica nella Venezia del Rinascimento. Ed. Bruno Mondadori (2006).
- [20] A. Abdou, R. W. Guy: Spatial information of sound fields for room-acoustics evaluation and diagnosis. Journal of the Acoustic Society of America, 100 (5), November 1996.
- [21] R. Nicol: Restitution sonore spatialisée sur une zone étendue: application à la téléprésence. Ph.D. Thèse, Université du Maine, 1999.
- [22] M. A. Gerzon: Periphony: with-height sound reproduction. Presented March, 1972, at the 2nd convention of the Audio Engineering Society, Munich, Germany.
- [23] M. A. Gerzon: Ambisonics in multichannel broadcasting and video. Presented at the 74th convention of the Audio Engineering Society, New York, October 1983.
- [24] M. A. Gerzon: The design of precisely coincident microphone arrays for stereo and surround sound. 1975.
- [25] M. A. Gerzon: The optimum choice of surround sound encoding specification. Presented at the 56th convention of the Audio Engineering Society, March 1977, Paris, France.
- [26] M. A. Gerzon, G. J. Barton: Ambisonics decoders for HDTV. Presented at the 96th convention of the Audio Engineering Society, March 1992, Vienna, Austria.
- [27] D. Bonsi, G. Cengarle, D. Gonzalez, D. Stanzial: Integration of acoustic quad audio technology in the digital cinema workflow and comparison with Wave Field Synthesis. IP-Racine Milestone, Dec. 2006.
- [28] D. Bonsi, D. Gonzalez, D. Stanzial: Quadraphonic impulse response for acoustic enhancement of audio tracks: measurement and analysis. Proceedings of Forum Acusticum, Budapest 2005, p. 335-340.
- [29] D. Bonsi, D. Stanzial: Acoustic test of a basic dereverberation process. Proceedings of the COST G-6 conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9, 2000.
- [30] S. W. Smith: The scientist and engineer's guide to digital signal processing, 2nd edition. California Technical Publishing. San Diego, California, 1999.

Ringraziamenti

Giunto a questo punto, finalmente mi si presenta una buona occasione per esprimere in maniera indelebile la mia gratitudine e la mia riconoscenza ai miei familiari e ai parenti tutti, per anni e anni di supporto incondizionato nel corso dei quali si sono sempre dimostrati il più sicuro punto d'appoggio desiderabile, sotto tutti i punti di vista.

Al mio relatore, prof. Domenico Stanzial, al correlatore dott. Davide Bonsi, al dott. Diego Gonzalez e a tutto il personale del laboratorio FSSG-CNR di Venezia va un sentito ringraziamento per tutto l'aiuto e il supporto che mi hanno offerto, sempre accompagnato da gentilezza e simpatia.

Infine, dopo quattro anni di attività nel settore della registrazione audio voglio ringraziare tutti quei musicisti, professionisti o dilettanti, che con la loro arte sono riusciti a regalarmi le migliori emozioni ed esprimere la mia stima verso coloro che lavorano con dedizione per far sì che tali emozioni possano essere registrate, ricordate e rivissute; mi auguro che i futuri sviluppi scientifici e tecnologici dell'acustica possano dimostrarsi un valido strumento al servizio vostro e di tutti coloro che hanno "orecchie per intendere".

Stop. □

Rewind... <<