

Percezione, invarianza e i 5 parametri

La mancanza di un completo modello della percezione del suono non impedisce di acquisire utili indicazioni sulla qualità dei sistemi elettroacustici e sul loro perfezionamento.

di Paolo Nuti

Secondo Daniel Raichel del New Jersey Institute of Technology, i soggetti affetti da ipoacusie di tipo ricettivo (sordità dovuta a danni delle terminazioni nervose dell'orecchio interno) sono più sensibili alla distorsione non lineare introdotta da un dispositivo di riproduzione del suono dei soggetti ad udito normale.

Se le argomentazioni portate da Raichel, che impropriamente parla di distorsione armonica, a sostegno della sua tesi sembrano discutibili, è stato invece provato sperimentalmente con metodologie psicoacustiche e neurofisiologiche che il comportamento non lineare dell'orecchio genera prodotti di distorsione di intermodulazione anche in soggetti affetti da ipoacusie stimolati con segnali di frequenza superiore a quelle per loro percepibili.

È sorprendente. Così come è sorprendente rilevare in soggetti a funzione uditiva normale la percezione, in esperimenti condotti con stimoli acustici costituiti da due segnali di frequenza 4 kHz e 5 kHz, del cosiddetto tono di combinazione a frequenza 3 kHz.

Percezione ed alta fedeltà

Di fatto l'orecchio umano è caratterizzato da rilevanti fenomeni di distorsione sia lineare (risposta in frequenza irregolare) sia non lineare (distorsione comunemente intesa). Nonostante queste sue imperfezioni intrinseche, l'elaborazione da parte del cervello dei dati trasmessi dagli organi recettori dell'udito in combinazione con le informazioni già memorizzate (esperienza) ed eventuali ulteriori dati percepiti per altra via (in certe situazioni si ascolta anche con il corpo) consente di estrarre dai suoni informazioni che i più moderni sistemi di misura non sono in grado di rilevare.

Si pensi al problema dell'identificazione di una persona attraverso la sua voce: la notevole concentrazione di studi condotti negli ultimi anni sul riconoscimento automatico del parlatore ha portato a risultati di grande interesse sul piano scientifico, ma tutt'altro che definitivi su quello pratico. Viceversa chiunque di noi è in grado non solo di riconoscere una persona conosciuta at-

traverso la sua voce, ma anche di percepirne lo stato d'animo.

Nel caso dell'alta fedeltà la situazione è simile: pur disponendo di raffinate tecniche di misura dei parametri atti a caratterizzare un sistema di riproduzione del suono, non siamo ancora in grado di predirne con totale certezza il maggiore o minore gradimento da parte dell'utenza attraverso rilevazioni oggettive (misure di laboratorio): in alcuni casi, per completare la valutazione di un componente elettroacustico è indispensabile ricorrere a prove soggettive di ascolto.

Sembrirebbe una argomentazione esclusivamente a favore dei sostenitori della superiorità delle prove di ascolto su quelle di laboratorio, del soggettivo sull'oggettivo, ma così non è; in realtà si tratta esclusivamente del riconoscimento di uno stato di fatto: non disponiamo di un modello sufficientemente completo della percezione del suono.

La mancanza di questo modello da un lato non ci consente di esaurire la valutazione di un sistema (dall'ambiente di ripresa a quello di riproduzione) con una serie di misure, dall'altro pone una seria ipotesi sulla validità delle stesse prove di ascolto dal momento che la incompleta conoscenza del ruolo e del numero dei parametri in gioco sia dal punto di vista elettroacustico che psicoacustico impedisce di definire esattamente le condizioni sperimentali di ascolto.

In pratica può accadere, ed accade molto frequentemente, che ascoltatori anche esperti emettano giudizi di ascolto su un impianto o parte di un impianto in perfetta buona fede, ma totalmente inutili perché legati a particolari condizioni sperimentali non completamente definite. Esempi classici sono le prove di ascolto dei fonorivelatori, dove i problemi di interfacciamento con il giradischi, il braccio e l'amplificatore, dipendenti dai possibili abbinamenti, prendono spesso il sopravvento sulle caratteristiche proprie dell'oggetto ascoltato, quelle degli amplificatori dove sempre i problemi di interfacciamento in ingresso ed in uscita possono alterare o sovvertire i giudizi soggettivi, quelle dei registratori e dei nastri.

Un patrimonio da non trascurare

Se è vero che la mancanza di un modello della percezione del suono ci impedisce di stabilire con certezza gli obiettivi di progetto dei componenti della catena di riproduzione, è altresì vero che quello che dovrebbe essere un patrimonio di conoscenze definitivamente acquisito, viene troppo spesso dimenticato o trascurato.

Senza alcuna pretesa di completezza, riassumiamo alcune nozioni di psicoacustica di particolare interesse nell'ambito dell'alta fedeltà.

Innanzitutto l'orecchio è sensibile alla intensità ed alla frequenza dei suoni. Secondo Ohm, che postulò la "legge della fase", la qualità dei suoni è unicamente legata al loro spettro di potenza. Se per quanto riguarda l'insensibilità dell'orecchio alla fase questa regola ammette delle eccezioni, la sensibilità a livello e frequenza è un dato certo ed acquisito. Helmholtz propose un modello dell'orecchio come risonatore; questo modello è superato, ma il principio della selettività in frequenza, che oggi sappiamo essere assicurata dalla diversa attenuazione in funzione della frequenza delle onde superficiali trasmesse alla membrana basilare dalla finestra ovale, resta.

La minima variazione di frequenza che l'orecchio è in grado di discriminare si chiama "soglia incrementale di frequenza". Secondo i classici dati di Shower e Biddulf, la soglia incrementale di frequenza è funzione della frequenza e del livello di pressione sonora. La figura 1 mette in evidenza come la selettività dell'orecchio cresca rapidamente al crescere del livello, sia massima tra i 500 e i 5000 Hz e per livelli (piuttosto bassi) di 60 dB a 1500 Hz sia dell'ordine dello 0.2%. Se si confrontano questi dati (per altro non aggiornati, riferentesi a toni puri e rilevati quasi sicuramente con un gruppo di ascolto meno allenato dell'audiofilo medio) con quelli relativi alle fluttuazioni di velocità dei registratori a cassette (tipicamente $0.2 \div 0.08\%$) si scopre che il margine per l'alta fedeltà è piuttosto ridotto.

Altra caratteristica psicofisiologica nota da tempo, ma troppo spesso dimenticata è

la "soglia incrementale di livello" che evidenzia la possibilità di percepire differenze di livello dell'ordine di 0.2 dB, una quantità al limite della risoluzione delle apparecchiature di misura impiegate in elettroacustica.

Risposta in frequenza e localizzazione della sorgente

La capacità dell'ascoltatore di localizzare la direzione di provenienza dei suoni è stata tradizionalmente attribuita alla elaborazione delle informazioni relative alla differenza di intensità, di tempo di arrivo, di fase e di spettro percepito dalle due orecchie. Per una sorgente che si sposta orizzontalmente attorno all'ascoltatore la spiegazione appare ragionevole. Supponiamo però che la sorgente si sposti sul piano verticale di simmetria della testa: comunque si muova, la distanza dalle orecchie, il tempo di arrivo, la fase, l'effetto della diffrazione sono assolutamente identici. Eppure se ad occhi bendati pregate un amico di far schioccare le dita davanti, sopra e dietro la testa, sarete sicuramente in grado di identificare la posizione della sorgente. Come è possibile? Ebbene, solo 8 anni fa, nel 1974, Blauert ha potuto dimostrare che questo tipo di localizzazione è legato alla risposta in frequenza delle orecchie al variare dell'elevazione della sorgente. Quando risulta impossibile localizzare l'emissione sonora in base alle differenze tra i suoni percepiti dalle due orecchie, il cervello elabora i dati relativi alla risposta in frequenza: "Apparentemente l'uomo ha imparato ad associare differenti caratteristiche spettrali a diversi angoli di elevazione" (Schroeder 1975).

Naturalmente occorre che la sorgente sia a largo spettro: un tono sinusoidale puro è difficilmente localizzabile di per sé, sul piano di simmetria è impossibile individuarne la direzione di provenienza; viceversa la parola, la musica, lo schiocco delle dita (assimilabile ad un impulso), sono segnali a largo spettro e possono essere localizzati. Come controprova si possono inviare da una sorgente posizionata sul piano verticale di simmetria dell'ascoltatore programmi filtrati secondo le diverse risposte dell'orecchio al variare dell'elevazione ed in effetti l'ascoltatore li identifica come provenienti da direzioni diverse (vedi figura 3).

L'applicazione di questa nozione all'alta fedeltà ci consente di trarre una conclusione importantissima: le differenze di risposta in frequenza dei sistemi di altoparlanti (volgarmente detti "casse") non solo influiscono sulla colorazione o tonalità dei suoni percepiti, ma anche sulla collocazione spaziale dell'immagine sonora.

Altra applicazione hi-fi della stessa nozione: è noto che ascoltando in cuffia si ha generalmente la sensazione che i suoni provengano da "dentro la testa" e risulta difficile pensare alle sorgenti come esterne. Per anni questo rompicapo è rimasto irrisolto: si era persino supposto che la localizzazio-

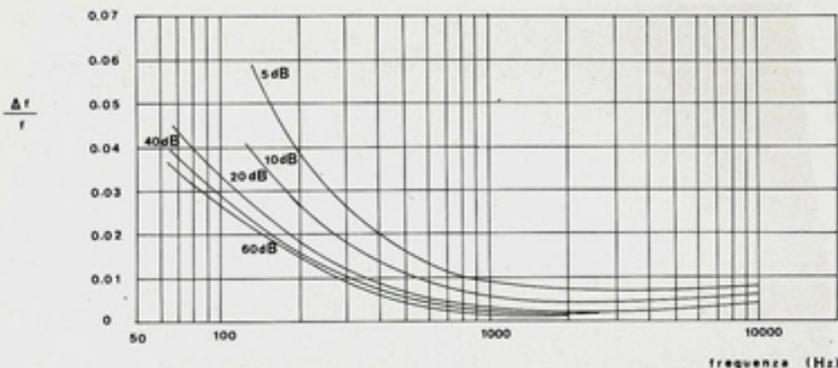


Figura 1 - Soglia incrementale di frequenza in funzione della frequenza per vari livelli di pressione sonora in dB riferiti a 20 μ Pascal. In gamma media sono percepibili variazioni di frequenza dello 0.2% (Shower e Biddulph)

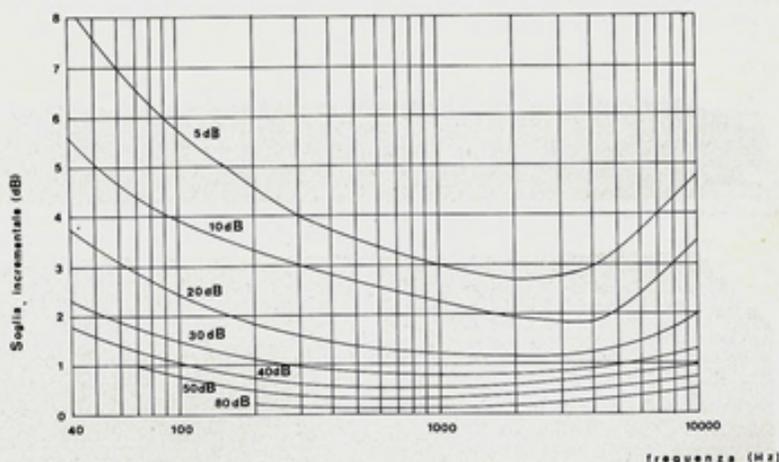


Figura 2 - Soglia incrementale di livello in funzione della frequenza e del livello di pressione sonora riferito a 20 μ Pascal. Per frequenze comprese tra i 150 e i 3000 Hz è possibile percepire variazioni di livello dell'ordine di 0.3-0.2 dB (Fletcher)

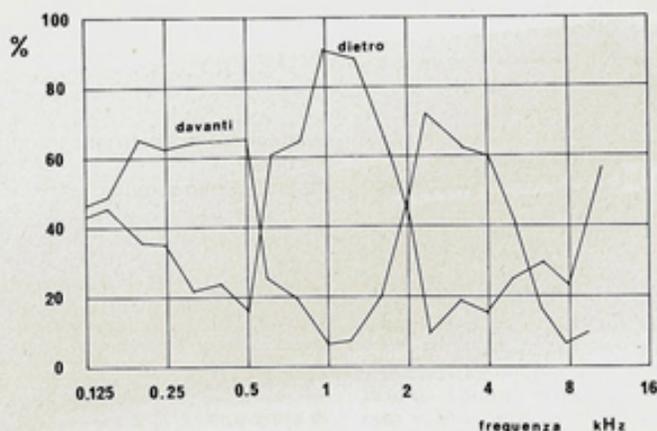


Figura 3 - La localizzazione di una sorgente sonora sul piano mediano (di fronte, sopra, dietro la testa) è determinata dallo spettro del segnale. In mancanza di altri riferimenti spaziali, suoni con forte contenuto energetico intorno a 1 kHz sono localizzati attorno all'ascoltatore. Se il contenuto energetico è concentrato sotto i 500 Hz e intorno ai 3 kHz, la sorgente viene considerata frontale. Sorgenti con elevato contenuto energetico attorno agli 8 kHz vengono localizzate sopra la testa. L'effetto filtrante della diffrazione causata dal padiglione auricolare esalta alcune frequenze e ne attenua altre. Attraverso questo meccanismo l'ascoltatore riesce a distinguere le diverse direzioni di provenienza dei suoni sul piano frontale (Blauert).



Figura 4 - Espansore di ambiente Koss K/4DS a linea di ritardo. L'interesse per i dispositivi di elaborazione di ambiente si sta rapidamente e giustamente diffondendo. Il cresciuto livello di integrazione delle memorie digitali e l'introduzione delle memorie a trasferimento di carica ha reso possibile la costruzione a costi accettabili per l'audiofilo di sistemi di espansione di ambiente a linea di ritardo fino a pochi anni orsono di costo proibitivo.

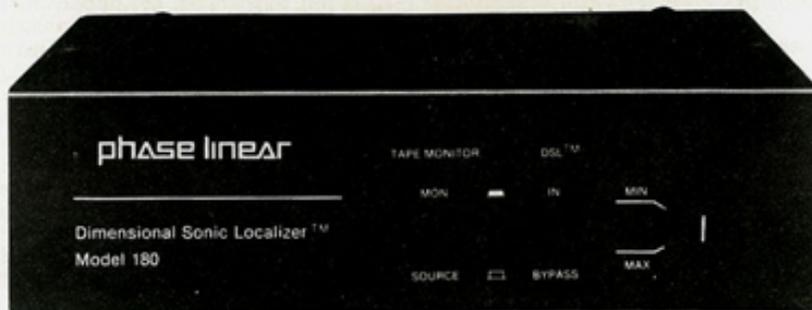


Figura 5 - Una seconda categoria di elaboratori di ambiente è costituita dai sistemi a matrice il cui principio di funzionamento si basa essenzialmente sulla miscelazione parziale dei segnali dopo opportuna filtratura ed eventuale ritardo. Anche in assenza di un modello completo della percezione dei suoni è possibile ottenere ottimi risultati sfruttando anche solo parzialmente le attuali conoscenze. In alto il "generatore di ologramma sonoro" C-9 di Carver, al centro il "localizzatore di suono dimensionale" modello 180 della Phase Linear, in basso l'"801 Omnisonic Imager" della OmniSonix

ne dei suoni interna alla testa fosse dovuta alla assenza dei sia pur minimi movimenti naturali della testa rispetto alla sorgente e alla pressione della cuffia sulle orecchie. Attualmente la mancata localizzazione esterna viene attribuita alle onde stazionarie che si vengono a formare tra la cuffia ed il timpano. L'effetto filtrante è diverso da quello della normale diffrazione dovuta ai padiglioni auricolari ed il cervello, non riuscendo a localizzare le sorgenti in base allo spettro, le associa all'unico punto per il quale non ha informazioni: il centro della testa. La fondatezza di questa teoria è stata provata con segnali filtrati per compensare l'effetto delle onde stazionarie e simulare la risposta dell'orecchio in capolibero: sotto queste condizioni la localizzazione è corretta.

Una applicazione intelligente

Negli ultimi due anni, principalmente in America, sono nati un ragguardevole numero di elaboratori di ambiente. Si dividono essenzialmente in due categorie: simulatori di ambiente a linea di ritardo ed espansori di immagine a matrice.

Tra i primi citiamo le "linee di ritardo" (Advent, Bozak, Phase Linear, Audio Pulse, Koss) che sia pure con implementazioni diverse del sistema di ritardo (digitale o a trasferimento di carica) e soprattutto maggiore o minore facilità di regolazione, hanno tutte lo scopo di simulare attraverso un paio di canali supplementari i fenomeni di riflessione in ambienti più o meno grandi. Già 5 anni orsono Robert Berkovitz della AR presentò pubblicamente un sistema del genere non commercializzabile sia per la complessità circuitale causata dal minor livello di integrazione della componentistica "di allora", sia per la difficoltà di impiego: per regolare il sistema (a 12 canali) era necessario calcolare esattamente i vari coefficienti con un computer.

Gli espansori di immagine a matrice funzionano su di un principio sostanzialmente diverso: preso atto che la localizzazione spaziale della sorgente dipende dalle differenze tra i suoni percepiti dalle due orecchie, si inviano ai due altoparlanti di un normale sistema stereo miscele varie dei segnali destro e sinistro opportunamente filtrati e/o ritardati.

Pur funzionando in base allo stesso principio, i vari sistemi sviluppati da Carver, da Cohen (Sound Concepts), dalla OmniSonix si differenziano tra di loro per il diverso peso dato ai vari elementi della somma: risposta in frequenza, ritardo, livello.

Riducendo il tutto ai minimi termini, si ottengono già dei risultati interessanti con una semplice miscelazione parziale dei due canali (riduzione della separazione) che i possessori del Galactron MK 16 hanno la possibilità di sperimentare facilmente. Già nel lontano 1962 Atal e Schroeder, dopo aver osservato che in un normale sistema stereofonico il campo sonoro appare generalmente appiattito ed i suoni alle spalle, di lato e sopra l'ascoltatore non vengono per-

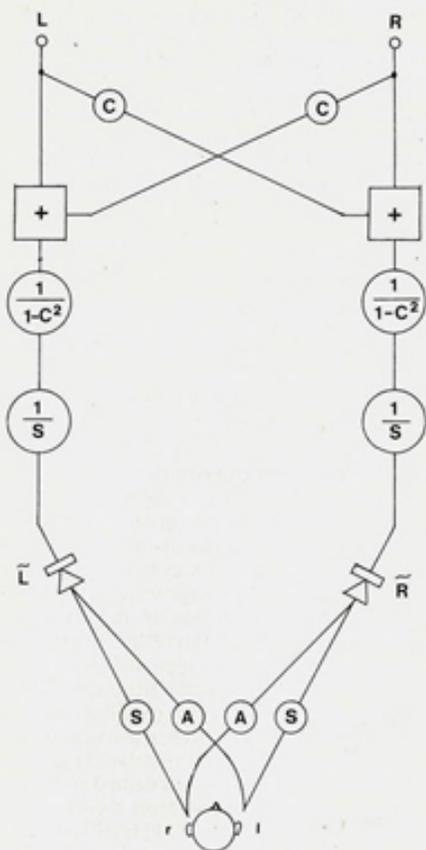


Figura 6 - Con una opportuna filtratura è possibile trasferire il segnale da una coppia di altoparlanti all'ascoltatore in modo tale che ciascun orecchio percepisca esclusivamente il suono proveniente dall'altoparlante corrispondente. S indica la funzione di trasferimento all'orecchio posto dallo stesso lato del diffusore, A la funzione di trasferimento all'orecchio posto dall'altro lato, C il rapporto A/S . Se i segnali R e L sono registrati con un testa artificiale corredata di microfoni di alta qualità, è possibile percepire una riproduzione altamente realistica con una perfetta localizzazione della sorgente per qualunque posizione attorno all'ascoltatore, compreso sopra e dietro la testa (Schroeder).

cepiti nella loro posizione, realizzarono il sistema di figura 5 valido per registrazioni effettuate con testa artificiale, che compensava per le differenze di livello e di risposta in frequenza trascurando il ritardo. Gli autori (Schroeder è considerato una delle massime autorità in tema di percezione e modelli dell'udito) riferiscono che il sistema era di fatto in grado di fornire la corretta immagine virtuale anche di sorgenti poste ai lati, dietro e persino sopra l'ascoltatore e definiscono i risultati ottenuti all'epoca "assolutamente entusiasmanti".

Sistemi invarianti e non invarianti

L'accurato ed intelligente controllo del parametro elettroacustico "risposta in fre-

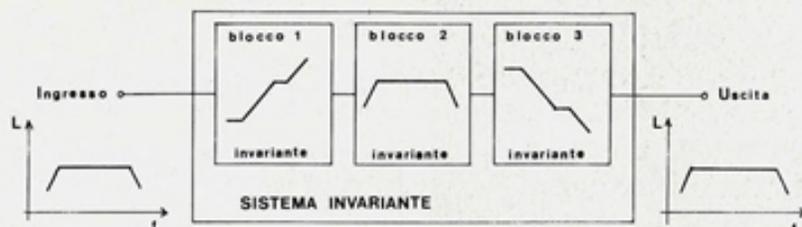


Figura 7 - Un sistema costituito da più blocchi è invariante se lo sono tutti i blocchi componenti. Se il sistema è invariante è possibile compensare le alterazioni della risposta in frequenza di un blocco con un'alterazione complementare di un altro blocco.

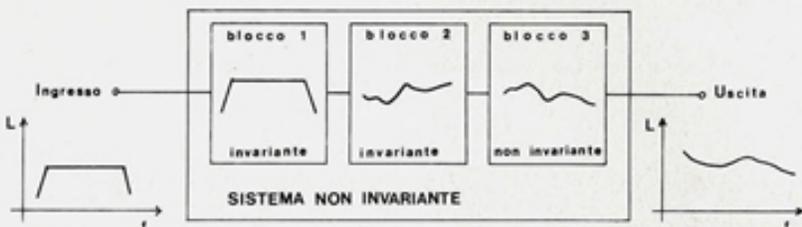


Figura 8 - Se anche uno solo dei blocchi del sistema è non invariante, tutto il sistema è non invariante. Non è lecito correggere alterazioni della risposta in frequenza introdotte da un blocco non invariante con alterazioni introdotte in un blocco invariante.



Figura 9 - Gli equalizzatori a ottave, a terzi di ottava e parametrici sono sistemi invarianti mentre a causa dei fenomeni connessi con la propagazione, l'ambiente di ascolto è un sistema non invariante. Dal punto di vista teorico l'equalizzazione ambientale non è completamente corretta e nella pratica se ne ha spesso la conferma. Gli equalizzatori trovano una loro valida applicazione solo nella correzione tonale del sistema e, al limite, ai fini di controllo della localizzazione spaziale.

quenza" a partire da poche nozioni di psicoacustica spesso dimenticate, può condurre a risultati di notevole rilievo ad onta delle semplificazioni introdotte. La risposta in frequenza del sistema deve dunque essere curata al massimo e non abbandonata al caso. Se alterazioni debbono essere introdotte, queste debbono essere volontarie e non involontarie, pena percepibili (ricordare la soglia di livello incrementale) mutazioni tonali senza la contropartita di alcun "beneficio spaziale".

Nel tenere sotto controllo la risposta in frequenza occorre però agire con cautela senza dimenticare alcune nozioni fondamentali di teoria dei sistemi.

Innanzitutto la catena di riproduzione del suono è a tutti gli effetti un "sistema" inteso come un dispositivo atto a trasformare

un segnale detto ingresso in un segnale detto uscita. Un sistema si dice lineare quando, se agli ingressi $x_1(t)$ e $x_2(t)$ corrispondono le uscite $y_1(t)$ e $y_2(t)$, applicando all'ingresso la sollecitazione $ax_1(t) + bx_2(t)$ si ha in uscita $ay_1(t) + by_2(t)$. Un sistema si dice invariante quando, se all'ingresso $x(t)$ corrisponde l'uscita $y(t)$, all'ingresso $x(t-t_0)$ corrisponde l'uscita $y(t-t_0)$. In termini pratici un sistema è lineare se per esso vale il principio di sovrapposizione degli effetti ovvero se l'uscita è rigorosamente proporzionale all'ingresso ed è invariante quando l'uscita è funzione solo del segnale di ingresso e non della sua "storia". La cosa più importante dal nostro punto di vista è che la funzione di trasferimento (risposta in frequenza e fase) di un sistema lineare invariante è data direttamente dal

INDIANA LINE Qx

computer aided design
loudspeaker system



Il diffusore "Indiana Line Qx" è la realizzazione di un sistema "vented" ottimizzato con l'aiuto del "computer" per ottenere un allineamento alla Butterworth, cioè con risposta massimamente piatta. Questo tipo di accordo rappresenta un ottimo compromesso tra estensione della risposta e smorzamento della risposta impulsiva.

Il woofer della Qx è un componente di qualità molto elevata con bobina mobile a 4 strati e magnete di grandi dimensioni in modo da ottenere un fattore di merito (Q) adeguato per la corretta realizzazione del progetto.

Il medio presenta caratteristiche di elevata efficienza e risposta estesa ad oltre 10 KHz, ben al di sopra della frequenza di taglio (circa 6 KHz). Il tweeter a cupola morbida da 1" è in grado di sopportare potenze elevate e di estendere la risposta a 30 KHz con buona dispersione angolare ed efficienza.

indiana line

SELETTOR

via Piria 3/h - 10144 Torino
tel. (011) 74.58.41

prodotto delle funzioni di trasferimento dei vari blocchi componenti il sistema e che, noto il comportamento del sistema nel dominio del tempo, è completamente noto anche il comportamento nel dominio della frequenza e viceversa. In altre parole per un sistema lineare invariante costituito da vari elementi o blocchi (p.e. fonorivelatore, preamplificatore, registratore, equalizzatore, finale), 1 - la risposta in frequenza complessiva è pari alla somma delle risposte in frequenza dei vari blocchi, 2 - nota la risposta impulsiva o quella al gradino è sempre possibile calcolarne la risposta in frequenza o viceversa.

Conseguenze pratiche: 1 - se un elemento della catena ha delle alterazioni della risposta in frequenza (p.e. equalizzazione RIAA del disco), posso compensarle con un elemento di risposta in frequenza complementare (preamplificatore fono), 2 - per valutare il comportamento del sistema posso indifferentemente servirmi di misure effettuate nel dominio del tempo (p.e. oscillogrammi di onde quadre) o della frequenza (p.e. risposta in frequenza) a seconda della mia comodità di rilevazione e di interpretazione.

Un errore eclatante

Dal microfono ai morsetti dell'altoparlante la catena di riproduzione del suono può essere considerata un sistema a tutti gli effetti lineare invariante, ma dall'altoparlante in poi il sistema diventa non invariante a causa degli effetti connessi con la propagazione del suono (ritardo, diffrazione, riflessione): le variazioni di risposta in frequenza introdotte dall'ambiente non sono più indipendenti dalla "storia" del segnale. In un punto dell'ambiente possono o meno stabilirsi onde stazionarie ed i relativi massimi e minimi a seconda della durata del segnale a quella frequenza: se il segnale è abbastanza lungo perché nel punto di ascolto giungano delle riflessioni, il tutto si sommerà dando luogo ad un valore maggiore o minore di quello che si avrebbe in campo libero (assenza di riflessioni).

Abbiamo dunque un sistema non invariante costituito in parte da blocchi invarianti (figura 8); le regole valide per i sistemi invarianti ora non lo sono più ed in particolare non possiamo correggere la risposta in frequenza del blocco non invariante agendo sul blocco invariante. In altri termini l'equalizzazione ambientale a ottave, terzi di ottava, parametrica, tesa a linearizzare la risposta nel punto di ascolto è un errore matematico: quel che si deve fare non è modificare il sistema invariante, ma quello non invariante spostando o riprogettando i diffusori, modificando le caratteristiche dell'ambiente, il punto di ascolto etc. Il bello è che ad orecchio si è sempre percepito benissimo che qualcosa non andava nell'equalizzazione ambientale con misura nel punto di ascolto tanto che sono state proposte varie curve "ottimali" di equalizzazione.

In definitiva l'applicazione di alcune consi-

derazioni di carattere matematico tutt'altro che astratte, apparentemente dimenticate in una sorta di lapsus collettivo, porta alle stesse conclusioni cui molti erano giunti praticamente: la risposta in frequenza di un sistema di riproduzione hi-fi può e deve essere manipolata solo a scopo creativo, di controllo della spazialità del campo sonoro, di correzione dei difetti della sola parte invariante.

I 5 parametri

Con alcuni esempi abbiamo voluto sottolineare l'assoluta necessità di mantenere sotto controllo i 2 parametri fisici risposta in frequenza e tempo.

Con esempi e conclusioni simili il discorso potrebbe essere esteso a tutti e 5 i parametri fisici che possiamo postulare definiscono completamente le caratteristiche di un sistema elettroacustico: risposta in frequenza, distorsione, massimo livello riprodotto, rumore di fondo, tempo.

Anche in assenza di un modello generale della percezione, l'impiego intelligente delle nozioni a noi note consente di pervenire a risultati di notevole interesse o di evitare errori grossolani: ad esempio nel rilevare la non linearità di un sistema attraverso delle misure di distorsione occorre utilizzare tecniche adatte al dispositivo in prova pena il raggiungimento di risultati privi di significato: ad esempio la misura della distorsione armonica di un registratore a cassette a frequenze superiori a 6-7 kHz dà risultati inconsistenti: apparentemente la distorsione scende al crescere della frequenza mentre in realtà sale. Alcune semplici considerazioni portano a concludere che il metodo di misura giusto è un altro, precisamente quello della distorsione per differenza di frequenze. Così nel chiederci quanto a distorsione è possibile percepire, è piuttosto semplicistico limitare la sperimentazione a toni puri e quindi alla distorsione armonica. I segnali musicali sono a larga banda e le non linearità del sistema (e dell'orecchio) producono consistenti tassi di intermodulazione. Non a caso uno dei più potenti mezzi di analisi impiegati dalla neurofisiologia e dalla psicofisiologia dell'udito è costituito da stimoli bitonali.

Conclusioni

La mancanza di un modello generale della percezione, sia pure limitato al solo campo della percezione acustica, non deve impedirci, anzi suggerisce, di sfruttare a fondo alcuni risultati sperimentali già acquisiti, ma talora inspiegabilmente dimenticati. Applicando con intelligenza lo strumento matematico ed i progressi compiuti nella rilevazione dei parametri fisici che caratterizzano i sistemi di ripresa e riproduzione del suono, si ottengono risultati in buon accordo e/o integrabili con i modelli parziali di percezione del suono già a nostra disposizione.

www.renatogiussani.it