

A.P.A.T. - Linea Guida introduttiva:

la dinamica dei segnali audio

L'argomento specifico della dinamica dei segnali audio, naturali e registrati, non ha probabilmente eguali per la molteplicità di possibili approcci, che è tale da confondere le idee letteralmente a chiunque. In un tentativo di evidenziare, come premessa, proprio questa complessità culturale, verrà dapprima proposto un caso tipico, a grandi linee, esaminato nelle sue più diverse connotazioni. Gli approfondimenti seguiranno con il dovuto spazio e con i riferimenti alle più autorevoli trattazioni.

Nove versioni di una stessa realtà

- Un gruppo di musicisti esegue un brano, prima in una sala prove, poi nell'Auditorium dinanzi al pubblico: gli strumenti possono essere suonati, seguendo le indicazioni contenute negli spartiti, ad un numero di diversi livelli, denominati “*pianissimo*”, “*mezzo-forte*”, “*fortissimo*”, per citarne alcuni. A seconda dello strumento, a queste indicazioni corrisponderanno livelli di pressione sonora differenti. Per esempio, a 10 metri di distanza, un violino emetterà dai circa 45 deciBel di un “*pianissimo*” ai 60 dB di un “*fortissimo*”, un trombone dai 57 ai 70 deciBel¹.

L'intervallo dinamico emesso dai singoli strumenti è quindi piuttosto limitato: molti strumenti passano dal livello minimo al massimo entro un intervallo compreso tra i 5 ed i 17 deciBel, ma la presenza nell'orchestra di più strumenti della stessa famiglia consentirà un sostanziale ampliamento dell'intervallo. Il compositore (lo scopriremo più avanti esaminando il risultato di nostri rilevamenti) tenderà a concentrare il livello dell'orchestra entro tre “intervalli” dinamici di valore semantico ben determinato: i passaggi in pianissimo², il fraseggio intermedio³, i “pieni orchestrali”⁴.

- Il fonico che stia operando l'eventuale registrazione del brano musicale sarà interessato a due livelli specifici: quello del rumore di fondo del sistema di registrazione ed il livello massimo lineare registrabile senza distorsione. Il livello di rumore di fondo sarà deciso dalla scelta del tipo di microfono e/o del preamplificatore⁵, oltre che dalla risoluzione effettiva del convertitore analogico-digitale⁶. Il livello massimo sarà appena più difficile da definire: qui sarà importante la tensione di alimentazione dei microfoni e la massima tensione accettata dal preamplificatore⁷. Un fonico della

¹ M. Clarke e D. Luce – *Intensities of Orchestral Instruments Scales Played at Prescribed Dynamic Markings* – Journal of Audio Engineering Society, Vol. 13, n.2, April 1965.

² Tipicamente a livelli inferiori a 60 dB”A” Slow, ma circa 15-20 dB superiori al valore del rumore di fondo ambientale.

³ Con livelli tipici situati attorno a 70 dB”A” Slow: questo livello appare particolarmente significativo e frequente.

⁴ Con valori compresi tra gli 83 ed i 96 dB”A” Slow, in funzione della compagine orchestrale e della vastità della sala.

⁵ I microfoni a condensatore possono avere un livello di rumorosità intrinseca inferiore a 17 dB”A” (Slow) ed una sensibilità di 20-50 mV per 94 deciBel, che è tale da superare quasi sempre il livello di rumore termico tipico degli stadi di preamplificazione. Diverso il caso dei microfoni dinamici, che hanno sensibilità dell'ordine di 1-2 mV per 94 dB: per essi il livello minimo registrabile è più spesso in funzione della configurazione dello stadio di preamplificazione.

⁶ Esistono convertitori a 16 e 18 Bit con prestazioni corrispondenti alla teoria. Quelli a 20 e –ancor più– quelli a 24 Bit dovrebbero avere livelli minimi quantizzabili sostanzialmente inferiori a quanto lo stato attuale della tecnologia dei semiconduttori consenta in pratica di realizzare: le loro specifiche sono –quindi– il più delle volte assai ottimistiche.

⁷ I microfoni a condensatore necessitano di una tensione di alimentazione (detta “Phantom”), che può variare dai 5 Volt delle realizzazioni economiche, ai 200 Volt dei microfoni di misura (tipicamente è di 48 Volt DC): a parità circa di sensibilità è proprio la maggiore tensione di alimentazione a consentire di operare senza distorsione ai più alti livelli, ammesso che gli stadi di preamplificazione e di conversione li possano poi effettivamente trattare.

generazione che ha operato con i registratori a nastro sapeva di poter contare su un certo margine di manovra: chi invece opera, oggi, con gli apparati digitali, sa che il livello massimo registrabile è perfettamente determinato⁸.

- Il fonico di Studio, che riceverà il nastro DAT o il CD della registrazione per “*masterizzarlo*”, cioè per renderlo adatto alla distribuzione commerciale, sarà interessato ad un livello assai diverso da quelli individuati dal suo collega in sede di ripresa: egli dovrà curare che il “prodotto”, cioè il CD commerciale, suoni ad un livello sostanzialmente simile a quello delle altre realizzazioni distribuite dalle varie case discografiche. Esiste (anzi esisteva) uno “*Standard*”, che prescriveva un valore medio di 18 deciBel inferiore al livello massimo registrato⁹ ed anche uno strumento di misura specifico assai meno ingenuo di quanto sia opinione comune: il *VU Meter*. Dunque il fonico di Studio doveva preoccuparsi di contenere i picchi appunto non oltre i 16-20 deciBel al di sopra del livello medio¹⁰, operando correzioni manuali, oppure avvalendosi di vari apparati elettronici, come i *limitatori* ed i *compressori*. La tendenza attuale è quella di situare il livello medio quanto più vicino possibile al livello massimo quantizzabile, anche appena 5-6 dB al di sotto¹¹.

- Il fonico che operasse la trasmissione dello stesso brano musicale dalla sala di controllo di una stazione radio in FM stereo sarebbe anch'egli attento a scegliere un valore “medio” consistente con le tipiche condizioni di ascolto domestiche o in auto: a limitare con assoluta precisione i picchi ci penserebbe, però, un *limitatore* di grande precisione, imposto dai regolamenti di concessione degli intervalli di frequenza di trasmissione, che prevedono un preciso limite di modulazione. Anche qui emergono ulteriori requisiti specifici: il pubblico domestico ascolta tipicamente a livelli piuttosto contenuti, mentre quello in auto ha il problema della forte rumorosità di fondo. In entrambi i casi vi sarebbe la penalità di perdere i passaggi in “*pianissimo*” se venissero riprodotti nelle corrette proporzioni: perciò si impiegano dei *compressori* con ampio, graduale intervallo di azione e precise costanti di tempo, che innalzano il livello nei passaggi più delicati, con il minimo di compromissione della qualità di ricezione.

- Quando lo stesso brano musicale viene trasmesso da una emittente in AM, ad onde medie, ogni parametro qualitativo assume connotati del tutto diversi: qui è il regno dei *compressori-limitatori* multibanda¹², che possono ridurre l'intervallo dinamico ad anche meno di 4-5 deciBel, con un sacrificio della qualità all'ascolto che è comunque minimo, rispetto a quanto altrimenti causerebbero le saturazioni dei mini-amplificatori ed altoparlanti delle radio. Naturalmente anche il fonico della radio AM ha, bene in vista, uno strumento indicatore di livello, la cui scala di lettura riflette bene il contesto assolutamente particolare¹³.

- Diverso e speciale anche il contesto in cui andrebbe ad operare un fonico che impieghi il brano musicale come colonna sonora di un film: qui il livello di riferimento è ancora differente, ed è, logicamente, quello del parlato. Ma la situazione non è affatto semplice come sembra: nei cinema esistono impianti audio con potenzialità ben diverse, che vanno da quelle in auge nelle arene estive, non molto dissimili da quelle dei televisori domestici, a quelle, spettacolari, dei cinema digitali, in grado di riprodurre un concerto rock esattamente agli stessi livelli delle esecuzioni dal vivo. Il parlato sarà rigorosamente allo stesso livello, ma per la musica e per gli “effetti” sono aperte

⁸ E' proprio il livello di riferimento, detto “0 deciBel”, cioè il massimo livello quantizzabile.

⁹ Bob Katz – *Integrated Approach to Metering, Monitoring and Leveling Practices, Part.1* – Journal of Audio Engineering Society, Vol. 48, n.9, September 2000.

¹⁰ Ma qual è il livello “medio” di un brano per grande orchestra ?

¹¹ Negli approfondimenti successivi scopriremo che questo è un aspetto carico di conseguenze a livello generale, incluso il problema specifico del contenimento delle emissioni da impianti audio.

¹² R. Orban – *Increasing coverage of international shortwave broadcast through improved audio processing techniques* – Preprint n.2789, 86th Convention of Audio Engineering Society, Hamburg, March 1989.

¹³ S. Klar, G. Spikofski – *On levelling and loudness problems at television and radio broadcast studios* – Preprint n.5538, 112th Convention of Audio Engineering Society, 10-13 May 2002, Munich (Ge).

opzioni assai diverse, anche in funzione del messaggio emozionale da trasmettere¹⁴.

- Il progettista o l'installatore che dovessero dimensionare un grande impianto audio per amplificare un concerto dal vivo o per riprodurre musica in una grande discoteca, sarebbero entrambi interessati a conoscere ancora una diversa grandezza, di interesse specifico: lo spettro energetico. Nel caso degli impianti audio è infatti relativamente facile reperire trasduttori efficientissimi e potenti, che emettano le frequenze medie e medio-alte, voce inclusa. Diverso è il caso se le richieste prevedono che la risposta in frequenza dell'impianto si estenda sino alle più alte frequenze udibili¹⁵. Ancora più impegnative le richieste di potenza per la riproduzione delle frequenze basse inferiori ai 60-80 Hz., che in pratica determinano gran parte del progetto e della configurazione dell'impianto¹⁶.

- Sia per l'impianto da concerto che per quello della discoteca sarà necessaria una Valutazione d'Impatto Acustico, che verrà redatta da un *tecnico competente*. Ancora una volta lo stesso brano musicale sarà esaminato, e misurato, secondo parametri assolutamente diversi da quelli sinora sottoposti all'attenzione di tutti gli altri tecnici operanti nel mondo dell'audio professionale. Il *tecnico competente* impiegherà un fonometro di Classe-1, che rileverà valori di Livello Equivalente ponderato "A", cioè con una curva di compensazione che riproduce il tipico andamento della sensibilità fisiologica ai bassissimi livelli, mentre i livelli effettivamente rilevati potranno superare i 100-109 dB"A" Leq. Il *tecnico competente* sarà interessato ai valori a lungo termine.

- Infine vi è un nuovo profilo professionale che si affaccia al mondo dell'audio, ed è quello del Consulente Tecnico di Parte (C.T.P.) cui il cittadino sottoposto alle immissioni ricorrerà a tutela del suo riposo e del rispetto dei vigenti limiti di immissione, in occasione di vertenze civili e penali. Le grandezze che il C.T.P. rileverà corrisponderanno solo in parte con quelle d'interesse del *tecnico competente*, in quanto gli attuali orientamenti giudiziari rendono possibile il caso in cui la Valutazione d'Impatto Acustico affermi la liceità di un determinato livello d'immissione acustica, mentre lo stesso livello potrà essere oggetto di sanzione in sede di vertenza civile ex Art. 844 C.C.¹⁷

In sintesi, abbiamo percorso una sequenza di differenti comparti dell'audio, ove sono effettuati "rilevamenti" ed operati interventi anche cospicui sul livello e sulla dinamica di qualsiasi tipo di registrazione o emissione audio, seguendo criteri e perseguendo obiettivi del tutto eterogenei e talvolta ampiamente conflittuali. Un approfondimento è evidentemente indispensabile¹⁸.

Il problema fondamentale: le unità di misura

Una cospicua parte dei problemi di reciproca comprensione in materia di dinamica tra i diversi profili professionali operanti nell'audio professionale deriva, prevedibilmente, dal diffuso impiego di unità di misura assai diverse, ma accomunate dal suffisso "dB".

Leggendo i resoconti di quanti abbiano tentato di affiancare i dati raccolti da differenti studiosi¹⁹, è

¹⁴ Un esempio: in alcune sequenze del film "Shine" –nelle sale digitali- il livello del pianoforte era riprodotto a livelli centinaia di volte superiori a quelli reali, senza che alcuno tra il pubblico avvertisse l'artefatto.

¹⁵ Esiste infatti un limite fisico al rendimento dei "drivers a compressione", come già accennato nell'*Introduzione tecnica agli impianti audio*, ed occorre dunque prestare molta attenzione al dimensionamento dell'impianto in questa gamma.

¹⁶ A queste frequenze, l'affiancamento di diffusori identici produce effetti di incremento del rendimento non solo per via dell'aumento della direttività di emissione, ma anche per l'aumento dell'impedenza di radiazione: il tutto rende assai critico il dimensionamento degli impianti audio di grande potenza, per questa specifica gamma di frequenze.

¹⁷ Vedi alla **Linea Guida n.16**.

¹⁸ Per dare una semplice idea del livello di eterogeneità dei criteri e degli apparati di "misura", basti pensare che l'indicazione del massimo livello campionato nel corso di una registrazione digitale (come tutte le odierne), segnala appunto il livello massimo tra tutti i campionamenti, rilevati ad intervalli pari alla frequenza di campionamento adottata, che nel campo professionale può essere di 48-96-192 KHz. Non viene impiegata alcuna pesatura nella risposta. Per confronto, la rilevazione del livello di Rumore di Fondo (ponderato "A"), per una vertenza, può richiedere ore – anche una notte intera- per ricavare un singolo valore statistico (L-95) superato appunto per il 95% del tempo di misura.

¹⁹ Due letture rendono bene l'idea della possibile confusione creata nel settore: il Preprint n.3869 – *Hearing Loss &*

inevitabile essere colti dal dubbio sul fatto che valori diversi (di oltre mille volte) possano rappresentare una medesima realtà fisica. Così è.

L'odierno diffuso impiego di apparati digitali, cioè "numerici", aiuta forse a comprendere meglio il contesto, che resta in gran parte complesso ed ambiguo, come sopra delineato. Anche i fonometri odierni sono tutti *digitali*, ma riportano al termine della misura risultati fortemente pre-elaborati, proprio per descrivere i livelli in termini più omogenei con quelli in uso nell'era dei rilevamenti analogici. Questo tipo di dati era, e resta, significativamente descrittivo del livello di sensazione.

Un sistema di registrazione digitale comporta necessariamente uno stadio di conversione, che trasforma il segnale elettrico, in uscita dal microfono o dal preamplificatore, in un segnale numerico binario, con una ben determinata risoluzione e frequenza di campionamento. All'inizio degli anni '80, all'introduzione dell'audio digitale negli impianti domestici, il formato tipico era quello che prevedeva 16 Bit per i livelli di campionamento ed una frequenza di 44.100 Hz per l'audio domestico (48 KHz nel professionale). A tutt'oggi questo è ancora il formato più diffuso.

Se il segnale emesso dal microfono è lineare, rispetto al livello di pressione sonora incidente, allora ad ogni "campionamento", espresso in un numero binario di 16 Bit, corrisponderà esattamente²⁰ il valore di pressione sonora al microfono, rilevato (in un apparato professionale) ogni 1/48.000° di secondo, senza alcuna ponderazione in frequenza.

Questo segnale *numerico* è facilmente analizzabile per via statistica, ottenendo un risultato di precisione e ripetibilità pari soltanto alla sua relativa inutilità...²¹

Il *valore massimo* campionato potrebbe sembrare di un certo interesse, senonché è del tutto possibile ipotizzare il confronto tra due registrazioni, di cui la prima contiene rarissimi campionamenti di valore prossimo al limite massimo registrabile²², mentre la seconda non contiene alcun campione che superi il valore di -10 dB rispetto al massimo, (p.es. perché il microfono o il preamplificatore a monte del convertitore saturano ad una tensione inferiore alla massima campionabile). Nonostante ci si aspetti il contrario, la seconda registrazione potrà essere facilmente quella con il valore medio più elevato, oltre ad essere senz'altro la più distorta e sgradevole all'ascolto. Il *valore medio* di una registrazione bipolare è, ovviamente, pari a zero o equivalente ad un piccolo *offset* di corrente continua... Il *valore minimo* è, inevitabilmente, una sequenza di 16 zeri...²³ Il *valore RMS*, cioè la radice quadrata di ogni singolo valore campionato elevato al quadrato, è indicativo del livello energetico: il problema è che ne abbiamo 44-48.000 al secondo, per canale. Approfondiremo più avanti il possibile utilizzo dell'analisi statistica di questo immenso numero di campionamenti, ma è già intuibile la difficoltà ad operare la valutazione in tempo reale...

In sintesi, l'audio digitale permette di ipotizzare una perfetta corrispondenza tra un livello di quantizzazione digitale ed un corrispondente livello di pressione sonora al microfono: la quantità ed il tipo di dati che si raccolgono sono ridondanti e tendenzialmente poco significativi, sia nella prospettiva di una semplice analisi energetica che, ancor più, come strumento per caratterizzare e,

Music – di Ken Dibble, 96th Convention of Audio Engineering Society (1994, Amsterdam) ed il Simposio: "Am I Too Loud", pubblicato sul Journal of Audio Engineering Society, Vol. 25, n.3, March 1977.

²⁰ Lo stato attuale della tecnologia permette ampiamente il rispetto di questa condizione, per un intervallo di 16-18 Bit, tant'è che i registratori digitali DAT sono considerati alla stregua degli apparati di misura di Classe-1 ed ampiamente utilizzati in ambito peritale.

²¹ Un caso in cui, invece, questo tipo di informazioni si è rivelato diagnostico è descritto nella **Linea Guida n.11**.

²² P.es. 10 campioni in un brano di tre minuti (cioè 10 campioni su 10 milioni circa, a 48 KHz). Del tutto inavvertibili all'ascolto...

²³ Se invece lo consideriamo come valore RMS, allora scopriremo che in tutti i convertitori digitali di accettabile qualità viene aggiunto appositamente un certo quantitativo di "rumore di fondo", detto "*Dither*", che ha lo scopo di impedire il crearsi di artefatti particolarmente spiacevoli all'ascolto: distorsioni e/o rumori a frequenza fissa. Vedi J. Vanderkooy e S. Lipshitz – *Dither in Digital Audio*- Journal of Audio Engineering Society, Vol.35, n.12, December 1987.

quindi, regolare i livelli, a disposizione dei diversi operatori dell'audio professionale.

Nella precedente era dell'audio analogico la situazione era possibilmente ancora peggiore.

Le variazioni della grandezza nota come “*livello di pressione sonora*” sono infatti subitane e quanto mai irregolari: due strumenti (analogici) che siano caratterizzati da costanti d'integrazione e da velocità di scrittura (o di lettura) differenti, leggeranno livelli diversi anche se collegati al medesimo microfono di misura... Quale sarà il valore significativo, rappresentativo quanto meno del livello di sensazione d'ascolto corrispondente ?

Il punto di vista del tecnico competente

Chi opera correntemente con strumenti di misura calibrati e certificati, sa che per ogni tipo di rilevamento esistono uno o più “parametri” che meglio descrivono la grandezza indagata. I fonometri sono in parte analogici: hanno un microfono a condensatore ed uno stadio di preamplificazione: tipicamente questi componenti sono realizzati con grande precisione e dispendio, al punto di superare facilmente le specifiche del migliore microfono da registrazione esistente.

Il fonometro ha il vantaggio non indifferente di poter disporre di differenti scale di misura, cioè di più rapporti fissi di amplificazione o attenuazione del segnale analogico, prima della conversione. La scala più sensibile permetterà, tipicamente, di rilevare livelli anche inferiori a 20 dB”A”. La scala con la maggiore attenuazione consentirà di rilevare livelli di picco anche superiori a 140 dB.

Per una banda di frequenze simile a quella registrata (20 KHz massimi), il fonometro avrà una frequenza di campionamento minima di 45-50 KHz., e dunque raccoglierà anch'esso una quantità enorme di dati, anche nel corso di un breve intervallo di misura. Diversamente dai sistemi di registrazione professionale, i fonometri possono però manipolare in tempo reale i dati che raccolgono, effettuando medie energetiche per intervalli successivi di campionamento (per esempio ogni secondo), oltre a memorizzare i livelli minimi e massimi nel medesimo intervallo, con diverse curve di ponderazione e costanti d'integrazione. Il tutto è operato nella prospettiva di duplicare, per quanto possibile, i meccanismi percettivi e di elaborazione del sistema uditivo umano: il risultato è spesso assai verosimile.

Il *tecnico competente* che vorrà individuare il livello sonoro di un particolare tipo di sorgente sarà coadiuvato, oltre che dalle capacità di calcolo dello strumento di misura, anche da un numero di semplificazioni ormai entrate nella pratica. La più importante è quella che vede l'impiego pressoché universale della curva di ponderazione “A”, che attenua alquanto le altissime frequenze, ma assai di più le basse al di sotto dei 500 Hz, imitando assai da vicino la normale sensibilità fisiologica dell'orecchio umano ai bassi livelli d'ascolto. Agli alti livelli questa curva di ponderazione (“A”) riesce ad imitare in modo convincente il livello di fastidio (e di danno all'udito) corrispondente ai valori di pressione sonora continui superiori ai 95-100 dB”A”: per questo viene ancora utilizzata per rilevare e valutare i livelli di esposizione agli alti livelli di rumore, nei luoghi di lavoro. L'unico tipo di rilevamento fonometrico operato normalmente senza ponderazione è quello del Livello di Picco Lineare, che i migliori strumenti rilevano con una costante di tempo di alcuni microsecondi.

I fonici non dispongono, ad oggi, di strumenti così sofisticati e ripetibili²⁴, ma, trovandosi comunque nella necessità di rilevare i livelli operativi, per poterli in qualche modo regolare, hanno dovuto elaborare strategie assai sofisticate ed interessanti²⁵.

Il punto di vista tecnologico

Vi è una sola condizione in cui il fonico potrebbe fare a meno di dover effettuare regolazioni di sensibilità, ottenendo tuttavia una registrazione perfetta: è il caso in cui la dinamica della

²⁴ Non dimentichiamo che i fonometri segnalano e memorizzano persino l'incidenza percentuale delle condizioni di sovraccarico, che possono condurre ad invalidare i rilevamenti.

²⁵ vedi ad esempio l'articolo di Bob Katz sul Journal of AES, Sept. 2000, citato alla nota n.9.

catena di registrazione sia talmente ampia da contenere tutto l'intervallo di valori di tensione corrispondenti al livello di pressione sonora al microfono.

Sin dagli albori dell'audio digitale, vi è chi ha intuito che questo sarebbe divenuto presto possibile²⁶, e ne ha studiato a fondo le implicazioni²⁷.

Alcune deduzioni sono particolarmente interessanti: nei normali ambienti domestici, come negli Studi di registrazione, il livello di rumorosità di fondo è tale da non mascherare affatto il minimo livello sonoro udibile nella banda di frequenze più critica, cioè quella centrata sui 4 KHz. Se anche infatti esiste un minimo di rumore, lo spettro di questo è tipicamente composto da componenti situate a frequenze inferiori a 200-300 Hz, con un'elevata pendenza di attenuazione verso le più alte frequenze; questa è tale da comportare livelli inferiori al valore di soglia di 0 (zero) deciBel anche a frequenze appena inferiori alla banda critica dei 3-6 KHz. Operando una misura con la ponderazione "A", si finisce dunque per rilevare solo il contributo preponderante della basse frequenze. Dunque la soglia minima di sensibilità è indicata (da Fielder) in un valore tipico di appena 4 (quattro) deciBel, per una banda di misura di 20 KHz.

Quanto al valore massimo, lo stesso Fielder ipotizza un valore di 128 deciBel (frutto sostanzialmente di estrapolazioni, più che di effettivi rilevamenti), per poi formulare una più ragionevole richiesta di appena 122 deciBel (questa volta, lo vedremo, piuttosto conservativa).

In realtà il punto è meritevole di più di un approfondimento, se non altro per il fatto che i valori ipotizzati dai diversi studiosi che hanno pubblicato in materia appaiono compresi in un intervallo ampio migliaia di volte e, per deduzione logica, del tutto inverosimile.

Anticipando appena alcune delle evidenze che appariranno più chiare di seguito, quel che occorre individuare preventivamente è il cosiddetto "*fattore di cresta*", cioè il rapporto tra il valore di picco ed il valor medio della forma d'onda su cui si indaga. Ebbene, in campo musicale esiste una conclamata e sostanziale evidenza circa il fatto che *fattori di cresta* superiori a 23-24 deciBel siano assolutamente infrequenti ed innaturali, specialmente negli strumenti acustici²⁸.

Si può arrivare alle medesime conclusioni anche partendo da un approccio leggermente diverso: ipotizziamo uno strumento (inesistente, forse assimilabile ad una pistola) che emetta soltanto un brevissimo impulso, caratterizzato da un fattore di cresta elevatissimo. Ebbene, se quello "strumento" è suonato e registrato in un qualsiasi ambiente, anche piccolo e relativamente poco riverberante, comunque la sua energia verrà diluita rispetto al tempo, con un andamento matematicamente deducibile, in funzione della densità temporale delle riflessioni. Se la sorgente sonora sarà situata, come sempre, ad una distanza superiore alla distanza critica²⁹, allora è ben difficile che alcun fattore di cresta possa superare il valore di 20-24 dB, specialmente nel caso, rappresentativo della pratica totalità degli strumenti acustici, in cui la sorgente sonora è pressoché omnidirezionale e vi sono oggetti e pareti riflettenti nelle immediate vicinanze sia della sorgente che dell'ascoltatore (o del microfono che registra).

Una semplice prova vale a fugare ogni dubbio: con i più recenti programmi di editing digitale è facile registrare un paio di secondi di silenzio, modificare tre campioni in modo di portarli su un valore elevatissimo (p.es. -3 dB rispetto al massimo campionabile), poi equalizzarne lo spettro fino

²⁶ Louis D. Fielder – *Dynamic-Range Requirements for Subjectively Noise-Free Reproduction of Music* – Journal of Audio Engineering Society, Vol.30, n.7/8, July/Aug. 1982.

²⁷ Louis D. Fielder – *Dynamic-Range Issues in the Modern Digital Audio Environment* – Journal of Audio Engineering Society, Vol.43, n.5, May 1995.

²⁸ Non è un caso che i valori più alti di pressione sonora di picco siano stati rilevati (e pubblicati) per posizioni di misura assai vicine, (pochi cm.) da strumenti a percussione: una condizione in cui l'uso di limitatori di picco è praticamente indispensabile (ove non provvedano in tal senso le limitazioni dinamiche di microfoni e preamplificatori). Gli stessi strumenti a percussione, ascoltati o registrati ad una normale, seppur minima, distanza, mostrano fattori di cresta ampiamente inferiori, grazie all'intervento delle riflessioni normalmente presenti in qualsiasi ambiente.

²⁹ Distanza alla quale l'energia emessa direttamente dalla sorgente eguaglia quella del campo riverberato.

a ricondurlo a quello tipico di un brano musicale, ed infine aggiungere mezzo secondo di riverbero (un valore tipico per un ambiente domestico)... Il risultato sarà il passaggio da un *fattore di cresta* di oltre 80 deciBel ad un valore anche inferiore a 30 deciBel, in funzione del rapporto energia diretta/riflessa (cioè come accade realmente, in ambiente...).

Individuato un fattore di cresta massimo di 20-24 dB, resta quindi da individuare un valore massimo del *Livello Medio*, che sia tollerabile all'ascolto: esiste una diffusa e ragionevole convergenza d'opinione su valori dell'ordine di 110-112 deciBel (rilevati con la costante "Slow"). Al di sopra di questi livelli medi, infatti, praticamente tutti gli ascoltatori percepiscono dapprima fastidio e poi dolore³⁰. Se, a questi, aggiungiamo un normale fattore di cresta di 20-24 deciBel o poco più, ecco venir fuori dai calcoli un valore di Picco Lineare di poco superiore a 130-136 deciBel, che, per la cronaca, è il valore tendenziale di molti dei sistemi di monitor da palco impiegati in occasione dei grandi concerti Rock degli anni '80³¹, prima che subentrasse la tendenza ad impiegare sistemi di monitoraggio personali, collegati via radio ed emittenti in auricolari appositamente sagomati.

Esistono microfoni che rilevano (e possono registrare, indistorti) livelli di picco anche di 140 deciBel: sono microfoni che possono avere anche un livello di rumorosità termica intrinseca pari a meno di 17 dB"A"³². Il problema non è nemmeno quello di progettare e realizzare preamplificatori che non compromettano quest'intervallo dinamico: di fatto ogni microfono a condensatore ha almeno uno stadio di amplificazione di corrente, per adattarne l'altissima impedenza elettroacustica³³ alla ben più bassa impedenza trasmissibile sui normali cavi di segnale, senza captare eccessivi disturbi.

Il problema maggiore è quello di realizzare un convertitore analogico-digitale che accetti le alte tensioni in arrivo dal microfono, corrispondenti ai livelli di picco prima ipotizzati. In teoria i convertitori a 20 e 24 Bit dovrebbero già oggi soddisfare queste specifiche³⁴, ma ben pochi, se non proprio nessuno di essi passerebbe indenne un test effettivo. Non va dimenticato che uno dei motivi "tecnici" alla base della diffusione dei formati di registrazione a 20-24 Bit è stato quello di tentare di limitare i problemi di compromissione della qualità del segnale in cui si cade ogni volta che esso viene trattato, senza aggiungervi la corretta dose di "*Dither*", come avviene per tutti gli interventi di equalizzazione, modifica del livello, ecc. Prima dell'introduzione del "*Dither*" era frequente ascoltare registrazioni digitali con un sottofondo di rumore ad alta frequenza, simile al segnale di trasmissione di un Fax. Risolto, con il "*Dither*" il problema della risoluzione ai più bassi livelli, emerge di recente la consapevolezza tecnica del fatto che la saturazione (o la limitazione) di segnali audio, se operata nel dominio digitale, comporta la possibile creazione di distorsioni da *aliasing* delle componenti ultrasoniche³⁵. Dunque il problema dell'intervallo dinamico registrabile resta ancora aperto alla necessità di approfondimenti.

Una volta conseguito il risultato tecnologico di poter registrare la completa dinamica udibile,

³⁰ Vi è, peraltro, una vasta documentazione relativa all'instaurarsi di cospicui spostamenti temporanei della soglia di udibilità, per esposizioni anche di pochi minuti a questi livelli sonori. Per esposizioni anche leggermente più protratte sono frequenti disturbi temporanei e –in prospettiva– danni anche permanenti alle Cellule del Corti, nell'orecchio interno.

³¹ Naturalmente questi sistemi di diffusori venivano operati molto spesso in regime di sovraccarico, caso in cui il fattore di cresta veniva sostanzialmente diminuito dall'insorgere del *clipping* degli amplificatori e dalla non-linearità dei trasduttori e dell'aria nei condotti risonanti.

³² Lo spettro del rumore di fondo di questi microfoni è tale da comportare solo un marginale superamento della soglia di udibilità, per un ristretto intervallo di frequenze centrato sui 4 KHz.

³³ Tipicamente una capsula di microfono a condensatore da 12 mm. di diametro ha un'impedenza capacitiva di appena 12-16 PicoFarad, per cui occorre un carico resistivo di centinaia di MegaOhm, pena la perdita di risposta alle basse frequenze.

³⁴ J. Vanderkooy e S. Lipshitz – *Resolution Below the Least Significant Bit in Digital Systems with Dither* – Journal of Audio Engineering Society, Vol.32, n.3, March 1984.

³⁵ Paul Kragh – *Aliasing in Digital Clippers and Compressors* – Journal of AES, Vol.48, n.11, Nov. 2000.

resterebbe tuttavia il non piccolo problema di utilizzare queste registrazioni: chi ha avuto in casa un impianto audio in grado di raggiungere o superare i 130/140 dB di picco³⁶ sa che è letteralmente impossibile ascoltarlo senza intervenire di continuo sulla regolazione del volume, anche, e soprattutto, nel caso di registrazioni con dinamica tutt'altro che ampia.

Tuttavia esiste già un formato di registrazione e distribuzione in grado di veicolare una dinamica ben superiore alle possibilità di riproduzione in pressoché qualsiasi possibile contesto: per la cronaca, le specifiche del formato di registrazione degli attuali DVD audio prevedono sia il formato stereo con 24 Bit e 192 KHz di frequenza di campionamento, sia il formato a 5-6 canali (20-24 Bit, 96 KHz.)³⁷. Esamineremo più avanti il problema della compressione mediante “*metadata*”, da porre necessariamente in atto per rendere utilizzabili questi sistemi.

Il punto di vista fisiologico

La registrazione di un flusso di dati numerici corrispondente a due canali di audio, campionato a 24 Bit e 192.000 volte al secondo³⁸, costituisce di per sé una sfida alle più sofisticate tecnologie esistenti: oggi siamo anche certi della grande ridondanza di dati raccolti.

Il problema è stato infatti affrontato al massimo livello culturale e di ricerca, sotto la pressante spinta di consentire la trasmissione via etere dei segnali audio digitali per radio e TV. E' infatti noto che anche la semplice trasmissione radio di un segnale numerico a 16 Bit/44.1 KHz porterebbe ad eccedere tutti i limiti di occupazione dello spettro delle radiofrequenze operanti a livello internazionale. Eguali, se non più gravi, sarebbero i problemi di occupazione di banda per la trasmissione in rete di segnali digitali nel formato di codifica lineare impiegato per i CD e DVD.

Il risultato, spettacolare, di queste ricerche, è stato quello di consentire l'elaborazione di sistemi di codifica “*percettiva*”, caratterizzati dalla capacità di eliminare dall'immenso flusso di dati numerici quella parte di informazione verosimilmente ridondante, rispetto all'ascolto finale, riducendo il flusso di informazione trasmessa di più del 90 % .

La storia del “*perceptual coding*” resta una delle pagine più affascinanti e meno note della storia dell'elettroacustica, alla quale hanno contribuito i massimi esperti in tutte le nazioni più avanzate³⁹.

Per il grande pubblico, il risultato del “*perceptual coding*” ha un nome a tutti noto: MP3⁴⁰. Questa sigla è l'acronimo per MPEG 1/2 Layer 3, cioè per la specifica caratterizzata dal maggiore livello di compressione (Layer 3) contenuta nel protocollo MPEG-1 del 1991 (ed MPEG-2, 1994): MPEG sta per Moving Pictures Expert Group, cioè il gruppo di esperti internazionali mobilitati per risolvere il particolare quanto impegnativo incarico.

Il primo e più importante aspetto dei sistemi di codifica percettiva è quello di operare una reale scrematura del flusso di dati, che non potrà mai più essere ricostruito pari all'originale. In questo delicato aspetto è presente la sostanziale differenza con tutti i precedenti sistemi di compressione, analogica e digitale⁴¹.

La necessità di questo grave sacrificio è stata imposta dal fortissimo coefficiente di riduzione del flusso di dati (tipicamente compreso tra 10 e 12 volte), nella conversione in MP3 di normali *files* PCM 16Bit/44,1 KHz.

³⁶ Stereoplay –*Il Mostro*- Anno X, nn. 87 ed 88, Aprile e Maggio 1981

³⁷ Bike H. Suzuki – *DVD AUDIO SPECIFICATIONS* – Paper 17-004, 17th International Conference of Audio Engineering Society, September 1999, Firenze (It.)

³⁸ O –peggio- sei canali a 20-24 Bit campionati a 96 KHz...

³⁹ Per un compendio, vedi la raccolta degli interventi alla 17th *International Conference of Audio Engineering Society*, 2-9 September 1999, Firenze (It.).

⁴⁰ K. Brandenburg – *MP3 AND AC3 EXPLAINED* - Paper 17-009, 17th International Conference of Audio Engineering Society, September 1999, Firenze (It.).

⁴¹ I sistemi di compressione-espansione impiegati dagli anni '70 per ridurre il livello di rumore di fondo nei registratori e nelle colonne sonore analogiche dei film, erano caratterizzati da interventi complementari in codifica e decodifica, appunto per consentire quanto più possibile il ripristino del dato originario.

In pratica il *Perceptual Coder* opera una sequenza continua di analisi spettrali, confrontandone in tempo reale i risultati con quelli di una simulazione della soglia di mascheramento; questa corrisponde ad un modello psicoacustico che riproduce quanto più verosimilmente possibile i meccanismi interpretativi del sistema uditivo umano: tutte le informazioni situate al di sotto di questa soglia **non** vengono codificate nei blocchi di dati destinati alla registrazione.

Esiste una grande e deliberata asimmetria tra la complessità degli apparati di codifica e la semplicità dei decoder: il protocollo di codifica è lasciato alla libera scelta di chi lo sviluppa e realizza, mentre lo *Standard* prescrive in pratica il solo tipo di rappresentazione dei dati e la rappresentazione matematica della decodifica, ivi incluse le tolleranze numeriche.

Quest'ultima caratteristica dei sistemi a codifica percettiva lascia ampio spazio per incorporare, in futuro, le più recenti acquisizioni sia tecnologiche che culturali.

Il limite nella qualità finale, al termine del ciclo di codifica/decodifica, non sarà commisurato alla finezza del campionamento iniziale (per numero di livelli o per la frequenza scelta), quanto alla sofisticazione del modello matematico alla base dell'elaborazione, operata ogni frazione di secondo, della curva di mascheramento da applicare (per la specifica frazione di secondo) per rigettare i dati ridondanti. Dunque un sistema di codifica/decodifica percettiva non ha, in linea di principio, limitazioni nella riproduzione di una dinamica in ingresso anche dell'ordine di 20-24 Bit ed oltre.

Esaminando in dettaglio i meccanismi di codifica, si scopre però che le informazioni ad alta frequenza ed alto livello, contenute nelle fasi iniziali dei transienti, sono quelle che creano le maggiori difficoltà in sede di codifica, per cui è comune la strategia di preferire il sacrificio delle frequenze comprese tra 16 KHz e 20 KHz, all'eventuale comparsa di artefatti di codifica avvertibili. La risposta alle più alte frequenze viene comunque recuperata nei segmenti codificati successivamente⁴². Altri artefatti tipici delle condizioni di sovraccarico dei *Coders* sono la presenza di pre-echi e doppi campionamenti, oppure la comparsa di rumori di fondo in bande ristrette, anche diverse da quella del segnale: sono tutti eventi la cui sensazione all'ascolto è del tutto differente rispetto a quella delle tipiche distorsioni dei precedenti sistemi di registrazione e riproduzione, sia analogici che digitali.

L'argomento è tra i più interessanti, ma le implicazioni fondamentali che se ne possono trarre, per quanto concerne il nostro interesse specifico ai problemi d'immissione, sono due.

In primo luogo i sistemi a codifica percettiva consentono, in linea di principio, di registrare e riprodurre l'intero intervallo dinamico udibile (e tollerabile): non a caso la loro prima applicazione è stata quella nei cinema digitali, ove hanno permesso di elevare di quasi dieci volte il limite superiore della dinamica disponibile. Applicazioni più moderne e sofisticate, già accessibili, promettono ulteriori ampliamenti verso l'alto nei livelli di picco registrabili senza distorsione.

La seconda implicazione, più sottile, sta nel diverso tipo di distorsione.

E' infatti del tutto possibile che più di una generazione di ascoltatori abbia maturato una particolare sensibilità percettiva verso il tipo di distorsione tipico della entrata in saturazione dei sistemi di registrazione/riproduzione sia analogici che digitali (tradizionali): come questi, anche i più comuni ed antichi strumenti musicali condividono la caratteristica di aumentare il contenuto di armoniche di alto ordine all'elevarsi del livello a cui vengono suonati⁴³.

Dunque un improvviso incremento di armoniche di alto ordine viene interpretato facilmente, e logicamente, come il segnale del raggiunto limite superiore della dinamica: spesso questo è l'indicatore maggiormente affidabile ed usato per il livello d'ascolto.

Prendiamo ad esempio tre tipici impianti audio: quelli domestici, per auto, o per discoteca: sorprendentemente in tutti e tre i casi è del tutto comune che si giudichi "alto" il volume d'ascolto al quale corrisponde appunto l'entrata in saturazione dell'impianto audio, nei picchi occasionali più

⁴² Questo è uno degli artefatti più facilmente individuabili all'ascolto, sebbene prestandovi molta attenzione.

⁴³ David A. Luce – *Dynamic Spectrum Changes of Orchestral Instruments* – Journal of A.E.S., Vol. 23, n.7, Sept. 1975

elevati. Questo nonostante i livelli sonori (sia medi che di picco) siano assolutamente differenti, nei tre casi esaminati. Mediante lo stesso criterio interpretativo si riconosce come un *pieno orchestrale* un livello di 83-86 dB”A” in un grande Auditorium da 2-3000 posti, ma anche il livello di 93 dB”A” emesso dalla stessa orchestra in una sala di capienza minore⁴⁴.

I meccanismi di codifica/decodifica percettiva alterano il contesto della dinamica audio in una maniera del tutto particolare: all'aumentare del livello d'ascolto corrisponde dapprima una sostanziale riduzione del livello di distorsione percepita. Questo avviene per via del fatto che la soglia calcolata in sede di codifica, per un livello d'ascolto arbitrario ma definito, risulterà più benevola se il livello d'ascolto a valle della decodifica sarà scelto più alto di quello di riferimento. Infatti in queste condizioni la soglia di mascheramento tipica del sistema uditivo umano si amplia sostanzialmente, attenuando o eliminando del tutto la percezione degli eventuali artefatti. Diversamente, qualora l'ascoltatore abbia la disaccortezza di preferire un livello d'ascolto inferiore a quello scelto come riferimento in sede di codifica, allora sarà del tutto verosimile che la più ristretta soglia di mascheramento presente in queste condizioni lasci emergere un numero di artefatti e distorsioni altrimenti impercettibili

Stiamo in pratica prevedendo l'esistenza di una sinergia tecnologico-psicoacustica che induca ad un allineamento verso l'alto dei livelli d'ascolto⁴⁵...

L'importanza del modello percettivo

L'affermazione del fatto che sia oggi possibile, ed ampiamente praticata, un'analisi strumentale, in tempo reale, tale da riprodurre fedelmente i meccanismi interpretativi dell'apparato uditivo umano, merita la massima attenzione. Esiste infatti un interesse del tutto parallelo a quello, tecnologico e commerciale, che ha fatto da propulsore agli avanzatissimi studi che hanno condotto al “*perceptual coding*” ed ai sistemi di trasmissione e registrazione digitale più recenti, come MP3, AAC⁴⁶, etc. Se infatti si esaminano con attenzione molti degli aspetti più di recente approfonditi tra quelli che sottendono il normale operare dell'apparato uditivo umano, non si potrà fare a meno di trovare proprio lì la spiegazione di un numero di deduzioni, sinora supportate assai più dall'esperienza e dal comune sentire, piuttosto che da obiettivi e scientifici riscontri.

Una pubblicazione in particolare merita un serissimo approfondimento, ed è l'enunciazione del più recente e completo modello percettivo ad opera di Brian C. Moore, B.Glasberg e T. Baer (1997)⁴⁷.

La prima parte di questa pubblicazione è dedicata alla descrizione del modello fisico e matematico, che approssima il comportamento dell'udito umano ponendo in sequenza dapprima le due funzioni di trasferimento dell'orecchio esterno e dell'orecchio medio, quindi ipotizza l'esistenza di un meccanismo di *amplificazione cocleare*⁴⁸, che in pratica opera un'attenuazione selettiva e graduale delle frequenze situate al di sotto di circa 500 Hz ai livelli inferiori a 100 deciBel. Quindi sono individuate le bande di frequenza critiche per l'analisi spettrale ed i coefficienti di mascheramento.

La seconda parte provvede a validare la teoria in un modo assai originale ed ineccepibile: in pratica vengono applicati i coefficienti del modello ad un numero di situazioni sperimentali già sottoposte al vaglio dai massimi esperti del campo, confrontando i risultati della simulazione calcolata con quelli sperimentali. Le implicazioni di questo studio, d'importanza storica, sono tali da interessare senz'altro anche gli ambiti di competenza specifici di chi si occupa del contenimento delle emissioni da impianti audio.

⁴⁴ All'epoca in cui furono composte ed eseguite le grandi Sinfonie del '700/'800, praticamente nessun Auditorium superava la capienza di 800/1000 posti. Oggi è normale eseguire gli stessi brani in Auditori con capienze di oltre 2200.

⁴⁵ Ancora una volta interverranno le distorsioni dei diffusori ed il *clipping* degli amplificatori a segnare il limite...

⁴⁶ AAC sta per Advanced Audio Coding, ed è il formato destinato probabilmente a succedere all'MP3.

⁴⁷ B.C. Moore, B.R. Glasberg, T. Baer – *A Model for the Prediction of Thresholds, Loudness, and Partial Loudness* – Journal of Audio Engineering Society, Vol.45, n.4, April 1997 (con 75 titoli in bibliografia).

⁴⁸ Già ipotizzato da K. Kryter nel 1970, sebbene in una prospettiva meno definita.

Due punti in particolare meritano attenzione.

Il primo e più evidente –oltre che carico di conseguenze- è il palese andamento convesso di tutte le curve che rappresentano il livello di sensazione sonora corrispondente all'incremento della pressione sonora, nella prima parte della curva, quella cioè che corrisponde ai livelli sonori più contenuti (indicativamente da 0 a 40 dB). Ai livelli più elevati l'andamento diviene lineare, salvo un ulteriore incremento di pendenza al di sopra dei 100 dB, che va a costituire il secondo aspetto.

La conseguenza fondamentale di questo primo assunto (teorico, ma assolutamente verificato), è che esiste un ampio intervallo critico, in cui ad un lieve incremento della pressione sonora rilevato dal microfono del fonometro corrisponde una sensazione di incremento della percezione del rumore che è assai più accentuata, super-proporzionale. Questo spiega assai bene la forte conflittualità che si determina allorché vengono subite immissioni anche di livello energetico minimo (p.es. 25-30 dB”A”) ma all'interno di abitazioni in cui, in assenza delle lamentate immissioni, il livello di rumore di fondo sarebbe altrimenti ben inferiore⁴⁹.

Il secondo punto è connesso con la presenza ed il ruolo dell'*amplificatore cocleare*, la cui presenza è stata dedotta per via dell'assai maggiore pendenza nell'attenuazione della sensibilità alle più basse frequenze, rispetto a quella fisicamente creata dalla trasmissione attraverso l'orecchio medio. La soglia dei 100 deciBel, alla quale questo meccanismo (di attenuazione) cessa di operare, assume quindi non più un valore psicologico, cioè di soglia emotiva, ma altresì quello del livello oltre il quale il meccanismo uditivo umano opera una sostanziale variazione dei suoi parametri⁵⁰.

La pratica effettiva: gli anni delle valvole

Esiste oggi, come si è già visto, una possibilità tecnologica concreta di registrare e riprodurre materiale audio caratterizzato da una dinamica del tutto pari a quella del sistema uditivo⁵¹.

Lo stato attuale della discografia internazionale è, però, in totale ed assoluta controtendenza⁵²... Gran parte delle incisioni più recenti mostrano, infatti, un intervallo dinamico sostanzialmente equivalente a quello tipico di una trasmissione A.M., ottenuto, peraltro, impiegando tecniche di limitazione multibanda non dissimili da quelle da tempo in uso in ambito Radio e TV.

La spiegazione di questo fenomeno, tuttavia razionale, è tutt'altro che semplice: occorre tornare indietro sino agli albori dell'audio, all'epoca delle valvole. Se infatti l'audio professionale fosse comparso all'epoca dei semiconduttori, tutto l'approccio ai problemi della dinamica di registrazione e riproduzione ne sarebbe stato profondamente ed irreversibilmente influenzato.

Senza che alcuno se ne potesse render conto, fino a data recente⁵³, gli stadi di preamplificazione (e di amplificazione di potenza) che impiegavano tubi a vuoto operavano un ruolo fondamentale nel contenere il livello dei picchi di più breve durata, senza distorsioni avvertibili, consentendo quindi di elevare il livello medio di registrazione e riproduzione. Il meccanismo è stato spiegato in parte

⁴⁹ Non bisogna mai trascurare il fatto che i comuni fonometri di Classe-1 raramente misurano a livelli inferiori a 20 deciBel, mentre l'orecchio umano risulta anche cento volte più sensibile. Quando si riscontra un'immissione, in presenza di un livello di Rumore di Fondo di 20 dB”A” o inferiore, non vi è praticamente alcun modo di escludere che il *differenziale* non sia di eguale misura rispetto al livello dell'immissione.

⁵⁰ L'andamento più che proporzionale che caratterizza la sensazione a livelli superiori ai 100 deciBel può essere invece bene ascritto come causa della soglia psicoacustica così ben nota ai fonici che operano nei concerti Rock dal vivo o dagli installatori che dimensionano gli impianti audio da discoteca.

⁵¹ Occorrono microfoni del tipo analogo a quelli di misura, convertitori a 24 Bit (veri), registratori in formato DVD (o analogo), ed elettroniche di amplificazione multiviva, con filtraggio digitale: quest'ultima accortezza è indispensabile per consentire alla ridotta estensione della risposta alle alte frequenze dei trasduttori dei bassi e medi di “filtrare” -all'ascolto- il rumore di fondo dei loro ampli specifici, i più potenti e quindi tendenzialmente meno silenziosi.

⁵² S.H. Nielsen e T. Lund – *Level Control in Digital Mastering* – Preprint n.5019, 107th Convention of Audio Engineering Society, 24-27 Sept. 1999, New York.

⁵³ vedi R.O. Hamm – *Tubes versus Transistors, is There an Audible Difference ?* – Journal of Audio Engineering Society, Vol.21, n.4, May 1973.

da Hamm⁵¹, in parte da molti altri studiosi che hanno indagato a posteriori un fenomeno emerso all'attenzione, nelle sue vere proporzioni, solo decenni dopo.

In pratica la maggior parte degli stadi di preamplificazione a valvole è realizzata seguendo configurazioni assai simili tra loro ed assolutamente semplici: si impiega il minimo indispensabile di *controreazione*⁵⁴ e le impedenze, sia in ingresso che tra i diversi stadi, sono solitamente altissime. Diversamente dai transistor e da tutti gli altri dispositivi di amplificazione, le valvole hanno un modo tutto particolare di entrare in saturazione: il loro terminale d'ingresso (cioè la “*griglia*”), ha un'impedenza altissima (cioè non assorbe alcuna corrente), fino a che riceve segnali di polarità negativa. Appena superata una determinata soglia, la *griglia* della valvola inizia a condurre, cioè ad assorbire correnti sempre maggiori, naturalmente a condizione che la sorgente di segnale a monte sia in grado di erogarle, il che è esattamente l'opposto della normale situazione d'impiego.

I preamplificatori microfonic professionali avevano tutti dei trasformatori d'ingresso, per consentire collegamenti bilanciati: ebbene l'avvolgimento che alimentava la *griglia* era, all'epoca, avvolto con migliaia di spire di filo finissimo, di enorme resistenza. Anche i *pick-up* delle chitarre elettriche erano e sono ancora avvolti utilizzando migliaia di spire di filo di rame di minima sezione. Naturalmente, stadio dopo stadio, le stesse valvole operavano a tensioni d'alimentazione anche assai elevate (3-400 Volt), ma con minime correnti (1-5 mA) e dunque ancora una volta con minima disponibilità a far fronte alle richieste di corrente delle *griglie* degli stadi d'amplificazione successivi⁵⁵, una volta condotti in saturazione.

In una catena di preamplificazione, missaggio, registrazione e riproduzione, tutta caratterizzata da stadi di amplificazione a valvole, si creano innumerevoli occasioni in cui una forte resistenza a monte interviene a limitare la corrente di *griglia* dello stadio di amplificazione immediatamente successivo. In questi casi il tipo di distorsione che si produce è asimmetrico e graduale, cioè ricco di armoniche di basso ordine, ivi inclusa la seconda armonica (il doppio della frequenza fondamentale).

Questo è esattamente il tipo di distorsione che ha le minime conseguenze negative all'ascolto, addirittura ne può avere di assai positive: due in particolare.

Un lieve e graduale incremento delle prime armoniche riproduce esattamente il normale comportamento della maggior parte degli strumenti acustici, che, quando sono suonati ai livelli più elevati, mostrano un inviluppo delle armoniche totalmente diverso da quello tipico a basso livello⁵⁶. Chiunque abbia esperienza di ascolto musicale dal vivo, acquisisce presto la capacità di riconoscere la modificazione degli spettri di emissione nei “*fortissimo*”, riscontrandola anche nell'ascolto di materiale riprodotto, dove contribuisce ad un'inconsapevole sensazione di maggiore livello. Anche nei piccoli impianti audio (radio, TV, ecc.) sono gli altoparlanti a distorcere gradualmente e con armoniche di basso ordine.

Il secondo effetto, particolarissimo, causato dalle distorsioni di basso ordine, è quello di attivare il meccanismo con il quale l'apparato uditivo sente perfettamente una fondamentale anche in assenza di energia a quella particolare frequenza: è sufficiente che le prime armoniche siano rappresentate tutte e con livelli gradualmente decrescenti, ed il fenomeno si attiva, inevitabilmente⁵⁷. Pochi, tra i non esperti nel campo specifico della progettazione di diffusori professionali,

⁵⁴ Negli stadi di amplificazione si usa riportare all'ingresso una quota del segnale in uscita, invertito in fase, in modo di stabilizzare i parametri operativi del circuito, diminuendo la distorsione ed aumentandone la banda passante. Tutto questo comporta una notevole perdita nel fattore di amplificazione, cui si sopperisce facilmente nei circuiti a transistor, aumentando il numero di stadi di amplificazione. Fare lo stesso con ampli a valvole è solitamente inutile –dato che le prestazioni di base sono già ottime- e, soprattutto, ha costi del tutto ingiustificabili.

⁵⁵ Una *griglia* condotta anche pochi Volt oltre la soglia di conduzione può assorbire facilmente anche 5-10 milliAmpere.

⁵⁶ vedi D.A. Luce – *Dynamic Spectrum Changes of Orchestral Instruments* – Journal of AES, Vol.23, n.7, Sept. 1975.

⁵⁷ L.L. Beranek – *ACOUSTICS* – American Institute of Physics 1954/1986, Pag.404.

immaginano quanto siano diffusamente impiegati questi ed altri simili meccanismi psicoacustici...

Anche i registratori professionali degli anni anteriori al 1970 avevano prestazioni sostanzialmente accettabili all'ascolto anche se operati in regime di parziale saturazione, sia per le amplificazioni interne a valvole, che per il comportamento di testine e nastri.

In sintesi, accadeva allora che il fonico addetto alla regolazione del livello di registrazione avesse ben pochi problemi a contenere il livello dei picchi entro il massimo consentito dalla catena di registrazione (e riproduzione): in suo favore operava una sequenza di possibili limitazioni da entrata in conduzione delle griglie di vari stadi attivi, assai poco avvertibili all'ascolto e, soprattutto, ben poco indagate a livello teorico. Il risultato di una indagine condotta con apparati e tecniche di misura moderni è illustrato in dettaglio nella **Linea Guida n.11**.

I transistor e l'era digitale

Il contesto operativo muta assai gradualmente, nel corso degli anni '70: non dobbiamo dimenticare che gli apparati audio professionali erano costosissimi, e dunque poco soggetti a rapide sostituzioni.

L'introduzione dei transistor, dalla fine degli anni '60, porta con sé un modo completamente diverso di progettare tutti i circuiti, dai preamplificatori agli amplificatori di potenza. Visto infatti il basso costo ed ingombro dei transistor, diveniva assai più semplice impiegare tassi elevatissimi di controreazione, tanto più che, di per sé, i transistor erano e restano tutt'altro che intrinsecamente lineari, e quindi ve ne era ben necessità. Così accade, ancor oggi, che un normale stadio di preamplificazione, o un finale di potenza, a transistor, siano entrambi realizzati mettendo in cascata 4-6 stadi ad elevato guadagno⁵⁸, per cui la quasi totalità del segnale d'uscita viene riportata in ingresso, fino a riportare il guadagno dalle centinaia di migliaia di volte al normale fattore, p.es. di 30-40 volte. I problemi di stabilità di queste configurazioni circuitali costituiscono di per sé già una specie di sfida, per il progettista, ma quel che è veramente spiacevole all'ascolto è il modo di saturare tipico di questi circuiti.

Immaginiamo un preamplificatore. Con pochissimi microVolt la sua tensione in uscita può raggiungere facilmente valori anche elevati, ma questa tensione viene riportata all'ingresso in fase opposta, stabilizzando il guadagno sul valore desiderato: questo meccanismo lo possiamo descrivere come pressoché istantaneo, al punto che non vi è effettivamente modo di separare temporalmente il primo "passaggio" di amplificazione dal sovrapporsi della correzione operata dalla controreazione. Il problema si verifica al momento in cui la tensione in uscita raggiunge il valore massimo, di poco inferiore alla tensione di alimentazione in continua: il segnale che ritorna all'ingresso resta lo stesso, piuttosto squadrato, e non può più controllare il fortissimo guadagno naturale del circuito. La forma dell'onda in uscita rimarrà ancorata sulla tensione massima, con una precisione estrema, fino a che il segnale in ingresso non sia sceso di quella minima frazione di millivolt al di sotto del valore cui corrisponde la saturazione, riprendendo il controllo del servomeccanismo.

Ecco spiegata la differenza: per tutto il tempo in cui un circuito a transistor satura, la sua tensione d'uscita rimarrà rigidamente ancorata alla tensione massima erogabile, che a sua volta sarà pari a quella di alimentazione, sottratte le minime inevitabili perdite. Le componenti spettrali di questo tipo di segnali saturati sono di ordine elevato, spessissimo solo di ordine dispari (per cui viene a mancare l'effetto della "fondamentale mancante"), taglienti ed avvertibilissime all'ascolto. Soprattutto la saturazione è improvvisa e netta, diversa da quella di qualsiasi strumento musicale.

I circuiti integrati hanno ripreso in pieno, anzi accentuato, le caratteristiche di artificiosità e sgradevolezza all'ascolto in condizioni di sovraccarico anche minimo. L'audio digitale ha aggiunto, senza nulla perdere (vi sono transistor dappertutto, anche nei convertitori...), tutta una serie di ulteriori problematiche⁵⁹⁻⁶⁰.

⁵⁸ Una valvola amplificava dalle 8 alle 50 volte, in tensione; uno singolo stadio a transistor può amplificare 3000 Volte...

⁵⁹ S.H. Nielsen e T. Lund – *0dBFS+ Levels in Digital Mastering* – Preprint n.5251, 109th Convention

La dinamica delle incisioni moderne

Sia i registratori a nastro professionali che i torni d'incisione per le matrici dei dischi in vinile avevano caratteristiche dinamiche assai prevedibili sul versante del minimo livello registrabile, individuabile (e misurabile) come livello di rumore di fondo. Assai meno definito era il margine superiore dell'intervallo dinamico disponibile: i nastri e le testine saturavano in misura assai variabile, mentre i solchi nei dischi non erano affatto limitati da un preciso valore. Diverso il caso delle trasmissioni in A.M. ed F.M., dove i limiti sono stabiliti da regolamenti internazionali e sono da rispettare imperativamente e con precisione.

Il fatto che sia i nastri che i dischi potessero essere registrati con valori di picco assai variabili ha condotto a sottovalutare, anzi praticamente ad ignorare, per decenni, il ruolo di *limitatori* operato dagli stadi di amplificazione a valvole. A questo si aggiunge che gli strumenti di misura del livello praticamente accessibili al fonico sono stati costantemente inabili a rilevare i picchi più brevi, per tutta l'era dei circuiti con tubi a vuoto.

Lo strumento tipicamente impiegato per allineare i livelli d'incisione (oltre al proprio giudizio d'ascolto) era il "VU-meter", uno strumento ad ago, le cui caratteristiche di sensibilità e di velocità di integrazione erano stabilite da un preciso "Standard" internazionale, che data al 1939 ed è rimasto intatto dal 1961⁶¹ ad oggi. Il VU-Meter era comunque uno strumento costoso, se realizzato conformemente alle specifiche: il suo ago doveva raggiungere il valore della tensione di riferimento entro 0,3 secondi dall'improvvisa applicazione della stessa ai terminali, con una sovraoscillazione non inferiore all'un per cento e non inferiore all'1,5%. Queste caratteristiche erano ottimali quando questo tipo di strumenti veniva impiegato per valutare il livello di una voce, assai meno per forme d'onda più complesse, come quelle tipiche dei segnali musicali.

La costante di tempo di un VU-Meter è infatti decisamente lenta, trattandosi di uno strumento in cui l'integrazione è ottenuta meccanicamente.

Il fonico che effettuava una registrazione analogica operava, il più delle volte, tentando il rispetto di **tre condizioni** di riferimento. Il livello di ascolto del sistema di monitoraggio era spesso regolato in modo di tenere il livello del rumore di fondo all'incirca sulla soglia di avvertibilità, mentre il livello di picco era dedotto tipicamente dalla sensazione di maggiore o minore distorsione avvertibile nei picchi, all'ascolto: se, ottenuto un ragionevole equilibrio tra questi due confini, il livello medio, letto dai VU-Meter era anche tale da individuare il livello medio all'interno della ristretta gamma di valori visibili, centrata sul livello di riferimento, allora tutto bene... Fino a che gli stadi di amplificazione sono stati del tipo a valvole, questo modo di procedere ha consentito di ottenere registrazioni qualitativamente eccelse⁶², senza ricorrere ad alcun sistema di limitazione che non fosse qualche raro ritocco manuale del controllo di livello in occasione di qualche passaggio particolarmente impegnativo, sia per il livello troppo alto che, all'opposto, per i *pianissimo*. La limitazione dei picchi operata dalle valvole era, come già detto, inconspicua, e dunque ignorata.

I problemi si sono presentati improvvisamente, al momento del diffondersi dei circuiti a transistor. Il periodo dell'introduzione dei transistor negli apparati audio professionali è finito per coincidere con un'ampia diffusione, presso le abitazioni, di impianti audio di progetto e realizzazione assai differenti da quelli degli omologhi professionali, diversamente da come era accaduto sino ad allora. Facile attribuire a questi apparati, dalle caratteristiche qualitative e dinamiche piuttosto limitate, la causa del brusco calo della qualità d'ascolto⁶³.

dell'Audio Engineering Society, Sept 2000, Los Angeles

⁶⁰ D. Mapes-Riordan – *A Worst Case Analysis for Analog-Quality (Alias Free) Digital Dynamics Processing* – Journal of Audio Engineering Society, Vol.47, n.11, November 1999.

⁶¹ ANSI C16.5-1961

⁶² Specialmente quelle su nastro magnetico 35 mm. perforato o su nastro da 1".

⁶³ I primi amplificatori domestici a transistor soffrivano quasi sempre della cosiddetta "distorsione d'incrocio", che aumenta al diminuire del livello d'ascolto ed è causata dall'errata polarizzazione degli stadi d'uscita in Classe B.

Il problema era però altrove: nella brusca e spiacevole saturazione degli stadi di preamplificazione a transistor, nelle instabilità di circuiti complessi ma ancora poco compresi, nella tendenza ad impiegare in ambienti domestici diffusori non più ingombranti ed efficienti, ma piccoli e di rendimento minimo (a *sospensione pneumatica*).

La soluzione è apparsa subito quella di ridurre il più possibile l'intervallo dinamico delle registrazioni, in modo di adeguarsi non solo ai limiti dell'ascolto domestico, ma anche di "unificare" il tipo di registrazione musicale commercializzato su supporto, con quello trasmesso⁶⁴.

Un altro aspetto innovativo nella tecnica di registrazione interviene ad accentuare la tendenza alla compressione dinamica: la registrazione multipista. Prima della metà degli anni '60 le registrazioni erano sempre effettuate in tempo reale, in mono o in stereo, ma con un singolo apparato di registrazione. La disponibilità di più tracce registrate sullo stesso nastro apre nuovi orizzonti creativi per i musicisti, che possono sovrapporre, per esempio, le registrazioni di due diversi strumenti, suonati dallo stesso esecutore ma in tempi diversi. Le possibilità di controllo dinamico sulla singola pista sono evidentemente assai maggiori, ed inavvertibili all'ascolto.

Il fenomeno della registrazione multitraccia esploderà presto, portando ad impiegare 48 o 64 canali anche per brani musicali assai semplici. I musicisti non devono più nemmeno incontrarsi fisicamente nello stesso Studio di registrazione, ma possono incidere le loro singole parti in sequenza, sincronizzandosi con un metronomo o con una traccia di percussioni. Le conseguenze sulla dinamica delle incisioni sono enormi. Basti pensare che la prima traccia, quella della batteria, viene incisa a livello praticamente costante (il massimo consentito...): ogni successiva traccia è registrata da musicisti che indossano cuffie, e che dunque devono suonare il loro strumento ad un livello non minore di quello al di sotto del quale l'isolamento della cuffia ed il mascheramento causato dalla traccia di sincronizzazione porterebbero a non sentire letteralmente nulla.

Gli andamenti dinamici delle singole tracce tendono quindi ad essere praticamente costanti: quando si "*missano*" tutte insieme basta un minimo disallineamento nel livello a rendere quasi inaudibile un particolare strumento⁶⁵.

Tante tracce miscelate comportano anche un probabile altissimo livello di rumore di fondo... Se le ampiezze che si sommano sono tutte elevate⁶⁶, allora lo stadio di amplificazione che (nel mixer) opera la sommatoria, sarà condotto facilmente in saturazione, a meno che non sia alimentato a tensioni altissime (il che non si è mai visto in alcun mixer...).

In tutte queste occorrenze l'impiego estensivo di *compressori* e *limitatori* appare tanto inevitabile quanto giustificato dalle nuove tipologie di ascolto, specie in auto, ma anche in abitazioni adiacenti altre e con un minimo di fonoisolamento tra loro. Resistono tuttavia ancora alcuni particolari comparti, come i dischi in vinile per uso in discoteca o gli stessi "nastri master" su formati grandi (p.es. 1") e costosi, ove sono registrabili intervalli dinamici assai ampi.

La rivoluzione degli anni '90

La registrazione multipista è sempre stata assai impegnativa, in termini di risorse tecnologiche e, quindi, economiche. L'introduzione dell'audio digitale, negli anni '80, ha visto l'impiego di mixer e

L'effetto all'ascolto di questo tipo di distorsione è senz'altro il più spiacevole in assoluto, poiché ad essa corrisponde un forte incremento delle armoniche dispari e di alto ordine, peraltro presente in ogni momento (diversamente dalle altre distorsioni, che intervengono solo nei picchi...).

⁶⁴ Esisteva già una rilevante cultura ed esperienza nella progettazione e realizzazione di *compressori* e *limitatori* per impiego radiofonico e -poi- televisivo, si è trattato soltanto di rivederne verso l'alto alcune caratteristiche qualitative, specie le soglie di rumorosità di fondo, per adattarli all'impiego musicale generale.

⁶⁵ E' in questo periodo che emerge fortemente all'attenzione la necessità di ascoltare con cuffie o diffusori monitor di grande linearità nella risposta. Una minima enfasi può infatti condurre a valutare come più alto il livello dello strumento che in quella banda di frequenze emette preferenzialmente, inducendo il fonico a miscelare la traccia ad un livello più basso, che però in tutte le altre condizioni d'ascolto risulterà inadeguato.

⁶⁶ Proprio per porsi al di sopra del livello di rumore di fondo...

di registratori multipista di costo paragonabile a quello di un'imbarcazione d'altura. Il noleggino degli Studi di registrazione era allora sostenibile soltanto dalle grandi case discografiche, che recuperavano l'investimento concentrando l'attenzione su un numero ristretto di artisti, di successo.

La situazione muta improvvisamente grazie all'evoluzione delle capacità di calcolo e, soprattutto, di memoria dei personal computers: il passaggio dai costosissimi registratori digitali 48 piste su nastro alla registrazione su hard-disk ha creato una riduzione di costi dell'ordine delle centinaia di volte. Tutto il contesto tecnico relativo alla registrazione audio è improvvisamente e radicalmente mutato.

I computers, per quanto potenti, lavorano preferenzialmente su pochi canali in contemporanea: non vi è però alcun problema ad "assemblare" anche centinaia di piste diverse, localizzate in segmenti diversi di memoria. Questo è detto "*editing non-lineare*" ed è il cuore della rivoluzione operata nell'audio professionale degli anni '90.

I primi sistemi di "*editing non-lineare*" erano costosi e di difficile impiego, ma hanno trovato immediata diffusione presso un numero di musicisti e produttori, che hanno potuto trasferire nei loro ambienti domestici l'intero "Studio di registrazione", con ulteriori e prevedibili conseguenze. Gli Studi di registrazione degli anni '70 ed '80 erano costosi da realizzare e da gestire: a quei livelli il costo di trattamenti acustici degli ambienti e di sistemi d'ascolto (monitoraggio) di qualità estrema era una voce di spesa sempre inferiore a quella dei mixer e dei registratori, di 10 volte. Un tecnico in sede, esperto dei complessi apparati e dei collegamenti, era un'ulteriore risorsa su cui investire: questi era spesso un fonico, ma a volte anche un esperto di manutenzione e quindi a conoscenza della configurazione e delle prestazioni dei circuiti.

Nulla di tutto questo era più indispensabile, negli Studi domestici: per effettuare il controllo qualitativo si impiegarono diffusori del tutto analoghi a quelli hi-fi di medio-bassa levatura⁶⁷. L'esperienza e la cultura del fonico, di solito lo stesso artista, potevano così limitarsi ai soli manuali operativi dei programmi e dei rari apparati esterni al PC, come i convertitori d'ingresso e d'uscita.

Il primo e più importante punto di discontinuità nell'evoluzione dell'audio professionale si situa proprio in concomitanza con la scomparsa dei grandi sistemi di diffusori monitor da Studio, capaci di risposte in frequenza estese anche al di sotto dei 30 Hz e, soprattutto, in grado di emettere livelli di picco indistorti anche dell'ordine di 130 deciBel e più. Questi diffusori, ingombranti e costosi, erano in precedenza realizzati e tarati con grande cura, in ambienti dall'acustica assai ben studiata.

Per confronto, il tipico studio domestico era basato su una coppia di monitor con un woofer da 20 cm. di diametro ed un tweeter a cupola morbida da 1", con un rendimento di poco inferiore a 90 dB per 1 watt ad un metro ed una risposta alle basse frequenze tagliata tra i 50 ed i 60 Hz. Soprattutto, questo tipo di diffusori non sopportava affatto gli alti livelli energetici a frequenze vicine al taglio inferiore che sono assolutamente tipici dei normali strumenti a percussione...

Anche accettando la tipica configurazione di ascolto a distanza ravvicinata, questi limiti dinamici erano ben avvertibili e lontanissimi da quelli dei grandi monitor precedentemente impiegati.

Il risultato è stato, inevitabilmente, quello di un'ulteriore fortissima spinta a comprimere la dinamica delle registrazioni, per renderla compatibile con quella riproducibile accettabilmente dai monitor.

Il secondo punto di discontinuità si verifica con la diffusione dell'audio nei PC portatili e nei sistemi "*Home-Theatre*", cioè nella televisione multicanale, imitazione domestica dell'audio cinematografico, incluso il "*surround*": in entrambi i casi i diffusori acustici sono di proporzioni irrisorie, così penalizzanti⁶⁸ da non potersi nemmeno definire, per essi, un vero e proprio "intervallo dinamico". Nessuna sorpresa che l'audio dalla fine anni '90 sia tendenzialmente limitato ad una dinamica di appena qualche deciBel.

⁶⁷ Un diffusore in particolare, le Yamaha NS-10, finì per monopolizzare letteralmente il mercato del monitoraggio.

⁶⁸ Utile la lettura di J.R. Ashley e T. Saponas – *Wisdom and Withcraft of Old Wives' Tales About Woofer Baffles* – Journal of Audio Engineering Society, Vol.18, n.5, October 1970.

I risultati ottenibili mediante l'impiego di particolari programmi di compressione e limitazione⁶⁹ sono sconcertanti, sia alle analisi strumentali che all'ascolto: non è infrequente che il livello d'uscita da incisioni anche a 20-24 Bit sia stabilmente compresso entro pochi deciBel al di sotto del valore massimo quantizzabile. Questo argomento verrà riaffrontato in più d'una Linea Guida, poiché le conseguenze specifiche di ciò, in ordine al problema del contenimento delle emissioni da impianti audio, sono enormi e di ben difficile soluzione.

La qualità delle registrazioni compresse digitalmente può tuttavia restare elevata, diversamente da quanto accadeva all'epoca dei *compressori-limitatori* analogici, a transistor, le cui costanti di tempo dovevano essere il frutto di notevoli compromessi⁷⁰. Laddove gli effetti sulla qualità dell'ascolto delle diverse proporzioni di compressione e delle differenti costanti di tempo sono stati indagati con rigore⁷¹, essi sono emersi tutt'altro che trascurabili, salvo nel caso particolare in cui l'ascolto dovesse necessariamente effettuarsi in condizioni di elevato rumore di fondo ambientale, per esempio in auto.

Il cinema digitale e i “metadata”: l'audio del futuro

Mentre in ambito musicale l'intervallo dinamico delle incisioni si è allineato sul valore tipico delle emittenti A.M., nel mondo del cinema digitale è avvenuto esattamente il contrario. Come accennato, la prima applicazione dei sistemi di codifica percettiva è stata proprio nel cinema digitale, ove ha permesso di riprodurre, nelle sale più avanzate, livelli di picco anche ben superiori a 120 deciBel, senza rumore di fondo o compromissione della qualità.

Anzi, si è addirittura affermata una tendenza a situare il livello del parlato in corrispondenza di valori anche assai moderati (64-66 dB”A” Slow), mentre le corrispondenti scene d'azione oltrepassano i 90-93 deciBel. Ne deriva la possibilità, per alcuni gestori, di elevare il livello d'ascolto in sala senza che il parlato diventi fastidioso⁷², ma con livelli assolutamente emozionanti nelle scene d'azione⁷³.

La diffusione di film con “audio digitale”⁷⁴ anche in ambito domestico⁷⁵ (mediante i DVD), avrebbe di certo comportato enormi problemi di distorsione e/o di vera e propria “sopravvivenza” degli attuali e diffusi minuscoli diffusori per “*Home-Theatre*”, se non fosse avvenuta in modo assai cauto, oggi ancora parziale. Per il cinema, infatti, l'opzione di “abbassare il volume” non è normalmente praticabile come nell'ascolto di musica registrata: se infatti si regola il volume d'ascolto in modo che sia adeguato alle potenzialità (minime) dell'impianto audio nel corso delle scene d'azione, allora il livello dei dialoghi sarà talmente ridotto da renderli incomprensibili, anche ascoltando in ambienti più che silenziosi.

La soluzione per il prossimo futuro sarà quella di impiegare un flusso di informazioni aggiuntive, dette “*metadata*”⁷⁶, immesse nel flusso principale di dati al livello della produzione⁷⁷, ma destinate a

⁶⁹ J. Chalupper – *Aural Exciter and Loudness Maximizer: What's Psychoacoustic about “Psychoacoustic Processors”?* – Preprint n.5208, 109th Convention of Audio Engineering Society, September 2000, Los Angeles.

⁷⁰ F. Floru – *Attack and Release Time Constants in RMS-Based Feedback Compressors* – Journal of Audio Engineering Society, Vol.47, n.10, October 1999.

⁷¹ Wagenars, Houtsma e Van Lieshout – *Subjective Evaluation of Dynamic Compression in Music* – Journal of Audio Engineering Society, Vol. 34, n.1/2, Jan/Feb. 1986.

⁷² Questo avviene quando si superano gli 80-82 dB”A” Slow, caso in cui la voce appare “urlata”.

⁷³ Naturalmente a condizione che l'impianto audio sia in grado di emetterli: si approfondirà questo aspetto assai più in dettaglio nella Linea Guida n.15.

⁷⁴ S. Vernon – *DOLBY DIGITAL: Audio Coding for Digital Television and Storage Applications* – Paper 17-005, 17th International Conference of Audio Engineering Society, September 1999, Firenze (It.).

⁷⁵ T. Spath e J. Couling – *Loudness and Dynamic Range in the Home* - Paper-16, 17th UK Conference of Audio Engineering Society, April 2002.

⁷⁶ C.Q. Robinson e K. Gundry – *Dynamic Range Control via Metadata* – Preprint n.5028, 107th Convention of Audio Engineering Society, September 1999, New York.

⁷⁷ S.G. Cossette e N.G. Guillen – *New Techniques for Audio Metadata Use and Distribution* - Preprint

controllare a distanza il livello sonoro presso l'utente finale, riducendo la dinamica d'ascolto secondo un programma di opzioni prescelto da quest'ultimo (p.es. ascolto notturno). L'argomento è tra i più interessanti, ma un controllo sull'applicazione pratica eseguito su un impianto domestico non ha dato alcun esito rilevabile, nonostante un numero di diverse combinazioni attivate, ascolto notturno incluso. Diversamente i tre diversi formati in uso nell'audio delle sale cinematografiche producono risultati assolutamente diversi sia all'ascolto che ai rilevamenti fonometrici, tant'è che è assolutamente facile identificare, anche dai soli rilevamenti, il tipo di impianto dietro lo schermo.

Una fase di transizione

E' particolarmente difficile trarre alcuna conclusione sulle implicazioni, a livello di contenimento delle emissioni, degli attuali sviluppi tecnologici, in presenza di tendenze attive così divergenti ed in evidente e rapida evoluzione. E' elevata la probabilità di essere smentiti assai presto. Che la dinamica delle incisioni musicali resti compressa entro pochi deciBel, quando il formato di registrazione e di distribuzione (DVD) ne consente ben 144, è, per esempio, solo il frutto di una concomitanza tra l'esigenza di far entrare nelle case impianti "Home-Theatre" di ingombro minimale, e la tendenza di chi registra a impiegare estensivamente i programmi di *massimizzazione*.

Di converso è del tutto probabile che, una volta entrata a regime la tecnica dei "metadata", si proceda all'unificazione degli intervalli dinamici dapprima delle colonne sonore dei film, poi anche dell'audio musicale, allineandole finalmente sul pieno intervallo della dinamica reale, perfettamente accettato già dalle attuali specifiche DVD.

Tutto quello che è stato realizzato, studiato, affermato in tanti anni di storia dell'audio è oggi avviato in una prospettiva di totale ripensamento, ove le possibilità del mezzo di registrazione non rappresentano più un limite, alla luce di una nuova cultura della psicoacustica⁷⁸.

Le singole **Linee Guida** sul *Contenimento delle emissioni ambientali da impianti audio* approfondiranno, caso per caso, sia le implicazioni attuali che le prospettive aperte dalle nuove tecnologie. Questa e l'altra Linea Guida introduttiva, sugli impianti audio, costituiscono sia una base tecnica, sia, soprattutto, **un compendio della bibliografia** disponibile sugli argomenti basilari.

FINE

n.5044, 107th Convention of Audio Engineering Society, September 1999, New York

⁷⁸ C'è chi sostiene, con eccellenti ragioni, che l'introduzione della stereofonia multicanale otterrà il risultato di far sembrare più alto il livello apparente d'ascolto a parità di livello di pressione sonora rilevabile strumentalmente. Questo venne ben notato nel passaggio dalle registrazioni mono a quelle stereofoniche...