

Virtù, vizi, pregi e difetti nella percezione della musica

di **Mario Bon**
(resp. tecnico Opera Loudspeakers)

Rispondere all'intrigante domanda "perché un amplificatore a valvole sembra suonare più forte di (buona parte) degli amplificatori allo stato solido di pari potenza?" potrebbe già essere un buono scopo per un lavoro di ricerca... In realtà, non sarà questo il tema della serie di articoli di cui fa parte questa prima puntata ma l'interrogativo e le considerazioni che ne conseguono tornano utili per introdurre un discorso più ampio che riguarda l'incerta corrispondenza tra i risultati dei test di laboratorio e la qualità soggettiva della riproduzione. Ciò avviene anche perché da una parte i test non rispecchiano pienamente le "effettive condizioni d'uso" delle apparecchiature e dall'altra perché i risultati non sempre tengono conto delle caratteristiche dell'apparato uditivo (cosa che si fa in una certa misura solo per i diffusori acustici): nelle "effettive condizioni d'uso" l'amplificatore è collegato ad una sorgente ed ai diffusori acustici e tratta, con alterne fortune, segnali musicali.

Per tentare di capirne di più conviene partire proprio dai segnali musicali, anzi, visto che il CD è ancora il supporto più utilizzato, partiremo direttamente dai "segnali musicali" che il CD contiene. Per qualsiasi altro tipo di supporto valgono, *mutatis mutandis*, considerazioni analoghe. Sicuramente parleremo di alta fedeltà, di CD, di amplificatori e diffusori acustici, di misure, di distorsione... insomma della riproduzione ma in particolare del segnale musicale.

La validità statistica dei risultati ottenuti dall'analisi delle tracce di un certo numero di CD, e illustrati nel seguito, dipende naturalmente dal numero di tracce analizzate: ne sono state analizzate oltre 600 (tutte estratte da CD audio) valu-

Pesa più un chilo di paglia o un chilo di piombo? In meccanica un chilo pesa sempre un chilo; in hi-fi incontriamo delle varianti interessanti: suona più forte, ad esempio, un amplificatore a valvole da 14 Watt o un amplificatore allo stato solido di eguale potenza? Un diffuso luogo comune vuole che l'amplificatore a valvole suoni più forte: in pratica è come dire che un Watt di vetro pesa più di un Watt di silicio... (1 parte)

tandone il contenuto nel dominio del tempo e, per una parte di esse, anche nel dominio della frequenza. Rispetto al totale delle tracce reperibili sul mercato, 600 sono decisamente poche quindi non si può pretendere di aver stilato una statistica pienamente rappresentativa! Tuttavia sono state analizzate tracce con caratteristiche estreme (le migliori e le peggiori). Analizzandone altre 600, o 6000, probabilmente si otterrebbe uno spostamento dei valori medi ma difficilmente cambierebbero i massimi ed i minimi. Per la finalità di questo studio era interessante stabilire i valori massimi (di cosa sarà chiarito nel seguito). Nella trattazione saranno utilizzati concetti e terminologie proprie della teoria dei segnali, dell'elettroacustica e dell'elettronica (segnali, spettri, valore RMS, fattore di cresta, FFT, ecc.); per facilitare il lettore, e non appesantire il testo, i "dettagli tecnici" sono trattati nelle appendici e nel glossario.

IL SEGNALE MUSICALE: COSE NOTE E MENO NOTE

Oggi, grazie al lavoro di molti ricercatori, conosciamo lo "spettro di emissione probabile" caratteristico sia dei singoli strumenti che dei diversi generi musicali. Le informazioni sui valori assunti dall'ampiezza (di picco) del segnale musicale nel tempo sono invece meno precise. Ma a cosa serve conoscere l'andamento "nel tempo" del segnale musicale? Serve per dimensionare l'amplificatore ed i diffusori in funzione del livello di

pressione che si desidera ottenere nell'ambiente d'ascolto. Oggi, se è vero, per tornare al quesito iniziale, che un amplificatore a valvole suona soggettivamente più forte di uno allo stato solido di pari potenza, la questione si risolve solo... "in base all'esperienza" (quest'ultima affermazione ha carattere provocatorio!). Non che questo sia un male ma è fastidioso dal punto di vista scientifico e metodologico: la mancanza di una correlazione certa tra le misure di laboratorio e la resa sonora degli apparecchi hi-fi ci autorizza a ridiscutere qualsiasi argomento (specie quelli meno presenti in letteratura).

Proviamo a ridurre quest'area di aleatorietà partendo dal segnale musicale; lo scopo dell'hi-fi è riprodurre la musica: se miglioriamo la conoscenza dei segnali musicali probabilmente riusciremo anche a riprodurli meglio! Prendiamo allora un amplificatore e immaginiamo di applicargli in ingresso un segnale sinusoidale. Se si "alza il volume" ad un certo punto l'ampiezza del segnale eccede la massima escursione di tensione d'uscita e l'amplificatore "clippa" (dall'inglese *to clip* = to-sare, limare - in italiano: "satura"). Il segnale clippato si presenta con i picchi appiattiti. Qualsiasi apparecchiatura elettronica ha un limite fisico superato il quale clippa: può clippare l'amplificatore, il pre amplificatore, lo stadio fono RIAA, il filtro crossover del diffusore acustico, un altoparlante, il lettore CD e, peggio del peggio, le

saturationi possono essere già belle e registrate nei CD! Nella figura 1 si vede un segnale sinusoidale come dovrebbe essere e come diventa a causa del clipping: in regime di saturazione i picchi del segnale sono appiattiti (clipping ideale). Nel caso illustrato la distorsione della forma d'onda dura circa mezzo periodo (un quarto di periodo nella semionda positiva e un quarto di periodo nella semionda negativa). Questa distorsione, con un segnale sinusoidale, è chiaramente udibile. Qualcuno si sarà chiesto cosa sia il "clipping ideale": quando un amplificatore satura, la forma d'onda non è sempre così pulita come appare nella figura 1; possono intervenire altri fenomeni (per esempio l'intervento delle protezioni o il cedimento dell'alimentazione, intermodulazione con i residui di rete, ecc.) e quindi le cose non possono che andare peggio della situazione rappresentata dal "clipping ideale" illustrato nella figura.

Facciamo un attimo un passo indietro e riconsideriamo quel che ho detto: "peggio del peggio, le saturazioni possono essere già belle e registrate nei CD!"...

I CD musicali sono in commercio, ad oggi, da 25 anni e sembra strano occuparsi dei CD audio dopo tutto questo tempo e quando sono già disponibili dei validi sostituti. Quello che interessa non è il CD ma la musica che c'è dentro ovvero il "segnale musicale registrato". Le considerazioni che seguono valgono, *mutatis mutandis*, anche per i formati digitali più

avanzati presenti e futuri. La lettura laser del CD ha eliminato il contatto meccanico tra il vinile del disco e il fonorivelatore o tra il nastro e le testine del registratore. Grazie a ciò sono scomparsi l'usura del supporto, il rumore di fondo, crepitii, distorsioni varie, gli errori di decodifica RIAA, i problemi di interfaccia dei fonorivelatori, il cross-talk tra i canali... e ne sono arrivati altri di nuovi (per esempio il jitter, il rumore di quantizzazione). La lettura laser è stata resa possibile dal passaggio dal formato analogico a quello digitale dei dati. Questo ha richiesto il campionamento (a 44100 Hz) e la quantizzazione (a 16 bit) del segnale musicale, diventato una sequenza di numeri (con valori compresi tra -32767 e +32767) che, riconvertita, restituisce il segnale musicale originale. Va anche detto che, nei 25 anni trascorsi, la qualità di lettori, di amplificatori e diffusori acustici è sensibilmente migliorata mentre la qualità dei CD è rimasta potenzialmente la stessa (stesso campionamento e stessa quantizzazione). Possiamo schematizzare il processo di registrazione del suono, dalla ripresa alla lettura del CD (fig.2), con una serie di "blocchi funzionali" collegati in cascata.

All'ingresso c'è la musica che viene ripresa dal microfono e trasformata in un segnale elettrico (analogico). Dopo vari passaggi, ritroviamo il segnale elettrico di uscita ottenuto dalla decodifica del CD (ancora analogico) e che sarà trasmesso all'amplificatore e quindi ai diffusori acustici. Le trasformazioni più rilevanti avvengono nel primo tratto (microfono, pre e filtro anti-alias). Il microfono trasforma il suono in un segnale elettrico ma soprattutto riduce il campo sonoro tridimensionale in un segnale elettrico monodimensionale: anche usando più microfoni il risultato è comunque un segnale elettrico monodimensionale. Il segnale proveniente dal microfono viene trattato (amplificato, mixato, compresso, ecc.) dal pre e limitato in banda dal filtro anti-alias che è indispensabile per la successiva conversione digitale. Il "segnale elettrico di ingresso" è quindi il segnale presente a valle del filtro anti-alias ed è diverso dalla "musica" non fosse altro per il fatto che, strada facendo, ha perso due dimensioni spaziali su tre. Questa precisazione ci consente di affermare che, se i blocchi funzio-

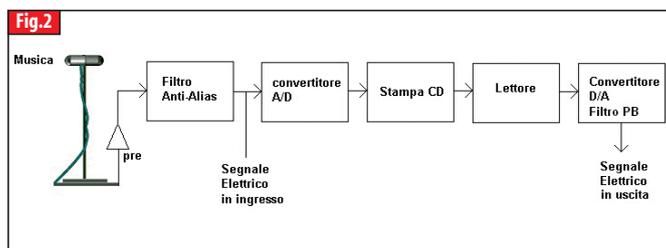
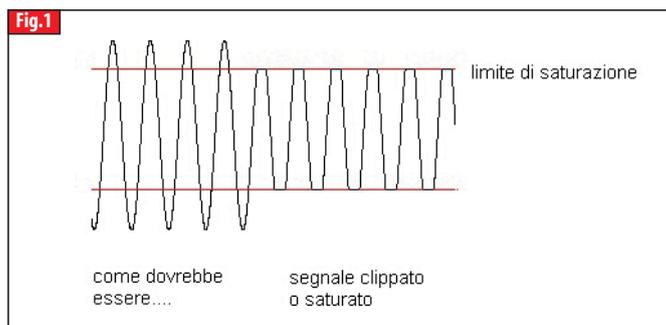


Fig.3

		dB	dB
pianississimo	PPP	0	-17.5
pianissimo	PP	+5	-12.5
piano	P	+10	-7.5
mezzo piano	MP	+15	-2.5
mezzo forte	MF	+20	+2.5
forte	F	+25	+7.5
fortissimo	FF	+30	+12.5
fortississimo	FFF	+35	+17.5

nali funzionano a dovere, l'unica differenza tra il "segnale elettrico di ingresso" ed il "segnale elettrico di uscita" è il "rumore di quantizzazione" dovuto dal numero finito di bit (16 per il CD) utilizzati per la conversione A/D (analogico/digitale). I formati a 24 bit "godono" di un rumore di quantizzazione inferiore; i formati con frequenza di campionamento più elevata "godono" di una banda passante più estesa. Non trattiamo la ricostruzione spaziale dovuta alla stereofonia (che ricostruisce un fronte bidimensionale arbitrario) e le caratteristiche di dispersione degli altoparlanti (che ricostruiscono un campo tridimensionale arbitrario); continuiamo invece a fare riferimento al formato CD e al "segnale elettrico di ingresso" che nel seguito sarà chiamato "segnale musicale" o anche semplicemente "segnale". Quando il segnale porta il convertitore A/D alla massima ampiezza di 32767 (-32767) diciamo che è stata raggiunta la "Massima Modulazione Assoluta" (MMA). Se invece il convertitore si ferma a

16384 si dice che ha raggiunto il 50% della modulazione, e via di questo passo. Ma cosa succede se si cerca di portare il convertitore A/D oltre il limite di MMA? Ogni valore superiore a 32767 (minore di -32767) viene troncato e avremo registrato nel CD un segnale "saturato" o "clippato" introducendo una distorsione nella forma del segnale. Quanto dura questa distorsione? Fin quando il segnale non riporta al convertitore A/D nella regione di funzionamento lineare (da un minimo di circa 22 microsecondi in su). Una delle funzioni dell'ingegnere del suono, che sovrintende alla realizzazione del CD, è quella di ottimizzare il livello del segnale da trasferire nel CD per sfruttare al massimo la modulazione disponibile senza produrre saturazioni. Non sempre i CD sono realizzati come dovrebbero. La qualità della registrazione è una delle caratteristiche immediatamente avvertite all'ascolto anche con sistemi di riproduzione di qualità non eccelsa: una buona registrazione fa suonare meglio qualsiasi impianto.

LA DINAMICA

Un generico segnale è caratterizzato dal "fattore di cresta" (CF dall'inglese *Crest Factor*) il cui valore adimensionale esprime il rapporto tra il massimo valore di picco ed il valore RMS rilevato in un intervallo di tempo specificato (picco = quantità istantanea; RMS = quantità integrale). Se il fattore di cresta vale 20 significa che l'ampiezza del picco massimo è

20 volte superiore al valore RMS. Affinché il fattore di cresta abbia significato deve essere specificata la durata dell'intervallo di tempo sul quale è stato misurato.

La dinamica è invece, in acustica, la differenza di intensità (media) tra il suono più forte e il suono più debole percepito e si esprime in deciBel (dB). La dinamica dell'apparato uditivo umano si estende per circa 120-130 dB e copre la differenza di intensità sonora prodotta dal ticchettio di un orologio in una stanza buia fino alla "soglia del dolore" rappresentata dal frastuono del tuono o del martello pneumatico (a distanza ravvicinata). La dinamica realmente percepita dall'orecchio non è così ampia: essa è limitata dal rumore ambientale, dalla riverberazione, dal mascheramento e anche da cause fisiologiche quali il "tempo di salita" dell'apparato uditivo (circa 2 millisecondi) ed i fenomeni di ipoacusia temporanea che interessano l'orecchio quando viene colpito da un suono inteso. L'esposizione prolungata a suoni di intensità superiore a 85 dB provoca danni permanenti all'udito. Con la musica i massimi livelli sopportabili sono più alti (il livello sonoro ad un concerto rock o in discoteca raggiungono 100 dB).

Nei CD la massima variazione di ampiezza, ovvero la differenza tra il valore di ampiezza massima e minima, vale $2^{16} = 65536$ (circa 96 dB) la dinamica, invece, raggiunge 90 dB circa. Per quanto riguarda il rumore di fondo presente negli ambienti questo può variare:

- 20 dB (case in aperta campagna, di notte)
- 40 dB (appartamento urbano tranquillo)
- 45-50 dB (appartamento urbano tipico)
- 70 dB (ufficio rumoroso, ristorante)

Ne segue che, in ambiente domestico, la dinamica fruibile offerta dal CD si riduce a valori compresi tra 45 e 70 dB (tipicamente 50-60 di dB). Per i dischi LP in vinile la massima ampiezza del segnale è limitata dalle dimensioni fisiche del solco e la dinamica ottenibile risulta di una trentina di dB più bassa rispetto al CD (20-30 dB). Per quanto riguarda la percezione soggettiva della dinamica nella musica Blake R. Patterson dei Laboratori Bell ha chiesto ad un certo numero di professori di orchestra di

suonare il proprio strumento riproducendo vari livelli: pianissimo (PPP), piano, forte fortissimo (FFF). Ne è risultata, mediamente, una differenza di circa 5 dB tra due gradi successivi. In questo test un singolo esecutore suona un singolo strumento. Su questa base la dinamica del singolo strumento, intesa come differenza di livello sonoro medio da PPP a FFF, vale circa 35 dB (fig.3). L'orchestra è composta da un numero variabile di elementi e non tutti partecipano ai pianissimo o ai fortissimo. In pratica la dinamica dell'orchestra viene valutata, a seconda degli autori, tra 60 e 90 dB. È evidente che, se si considera la pressione sonora prodotta da 100 elementi che eseguono un fortissimo e il singolo flauto che emette una nota al più basso livello di cui è capace, si raggiungono valori di dinamica impressionanti. Leo Beranek (*Acoustical Measurements*, 1988, p. 690) riporta i dati illustrati in fig.4. Si può stimare che i valori di picco corrispondono a un CF compreso tra 2 e 4. La fig.5 – ripresa da: Cingolati, Spagnolo (a cura di) *Acustica musicale e architettonica*, UTET, a pag. 682 – riporta i livelli SPL attesi per l'esecuzione delle opere di Haydn in quattro sale dove le opere furono effettivamente rappresentate. Il livello, in platea, va da 84 a 92 dB. Questi sono probabilmente valori medi (nel tempo e nello spazio) mentre il fattore di cresta del segnale non è specificato.

La dinamica dell'orchestra è condizionata dal tempo di riverberazione della sala e dalla posizione dell'ascoltatore. La dinamica percepita dall'ascoltatore dipende dalla distanza: il direttore d'orchestra percepisce la dinamica più elevata (con suono diretto predominante) mentre gli spettatori del loggione percepiscono la dinamica più bassa (con campo riverberato predominante, il fattore di cresta si stima compreso tra 3 e 4). Qualche cosa di simile si riscontra anche nella musica registrata: ci sono registrazioni fatte con i microfoni molto vicini agli strumenti e altre fatte con i microfoni più lontani, il che si traduce in un diverso rapporto tra suono diretto e riflesso (una diversa dinamica e una diversa prospettiva spaziale). Una certa quantità di riverberazione naturale, pur limitando la dinamica, rende l'ascolto più gradevole e meno affaticante (addirittura più dettagliato); un eccessivo riverbero limita l'intelligibilità e diventa controprodu-

Fig.4

SPL calcolato su un quarto di secondo, il fattore di cresta assume valori compresi tra 2 e 4 con i quali è possibile calcolare le presunte pressioni di picco.

Sorgente	Max SPL RMS per intervalli di 1/4 di sec. (dB)	Picco (stimato) per CF da 2 a 4 (dB)	Distanza in campo libero (m)
conversazione uomo	56.3	-	10
conversazione donna	54.5	-	10
conversazione normale	40-60	46-66	1 (*)
conversazione sostenuta	70	76	1 (*)
musica: una voce	87	93	10
musica: 100 voci	107	113-119	10
orchestra di 75 elementi	106	112-118	10
orchestra di 18 elementi	96	102-110	10

(*) questo valore è stato aggiunto.

Fig.5

	Castello Eisenstadt	Castello Esterhaza	Hanover square Rooms, Londra	King's Theatre, Londra
Numero di posti (circa)	6800	1530	1875	4550
Volume metri cubi	400	200	800	1050
Altezza del soffitto in m.	12.4	9.2	8.5	10.7
Tempo di riverberazione T60	1.7	1.2	1.0	1.5
Dimensione dell'orchestra (senza percussioni)	16	22	36	57
Numero di violini	6	11	14	24
SPL stimato in un forte tutti	84 dB	90 dB	91 dB	92 dB
SPL RMS stimato (CF=14.14)	-	96 dB a 1 m	-	-

cente. La musica per organo è accompagnata da un forte riverbero (gli organi si trovano nelle chiese e vengono registrati con i microfoni ad una certa distanza); per contro la musica da camera viene eseguita in ambienti più raccolti e meno riverberanti. È più probabile incontrare maggiore dinamica nella registrazione ravvicinata di un unico strumento che in un concerto per organo (non per nulla il fattore di cresta più elevato è stato trovato nella registrazione di un singolo colpo di piatto in ambiente anecoico). Nelle riprese ravvicinate vengono perse le informazioni ambientali che concorrono a determinare le corrette dimensioni fisiche e la collocazione spaziale dell'esecutore. Tutto dipende dal rapporto tra suono diretto e suono riflesso contenuto nella registrazione e dalla riverberazione aggiunta dall'ambiente dove il programma viene riprodotto. Ne segue che la dinamica non può essere assunta come unico indicatore della qualità di una registrazione: un'ottima registrazione può presentare una dinamica modesta, al contrario, se una registrazione possiede una dinamica eleva-

ta, è probabile che sia anche di buona qualità. Esiste una relazione tra il fattore di cresta e la dinamica? In altre parole: ha senso descrivere la dinamica di un brano musicale con:

$$CF=20 \text{ vale } 26 \text{ dB } (20 \times \log_{10}(20))$$

In linea di principio la risposta è negativa perché dinamica e CF misurano due quantità diverse. Tuttavia il fattore di cresta è una quantità cui corrisponde un metodo di misura standardizzato ed appositi strumenti per misurarla. Dato che il valore RMS fa riferimento ad un certo intervallo temporale è evidente che, se questo cambia, anche il fattore di cresta cambia.

IL SEGNALE MUSICALE

Il segnale musicale è il "segnale elettrico di ingresso" che troviamo a valle del microfono e del filtro anti-alias (fig.2). La musica produce emozioni ma non abbiamo strumenti semplici per misurare le emozioni. Quando la musica viene raccolta da un microfono si trasforma in segnale elettrico (per il quale

sono state elaborate teorie matematiche, metodi di misura, ecc.); del segnale elettrico possiamo conoscere anche la forma analitica; un'emozione, spesso, non è descrivibile nemmeno con una combinazione di gesti e di parole: la capacità di esprimere le emozioni è sostanzialmente un fatto culturale! Sempre per i segnali, utilizzando gli opportuni strumenti di misura, possiamo ottenere informazioni relative al dominio del tempo (durata, valori minimo e massimo, valore medio, valore RMS, fattore di cresta, distribuzione dell'ampiezza, autocorrelazione) e al dominio della frequenza (le quantità deducibili dallo spettro). Nel seguito parleremo di segnali e non di emozioni. Ciò a cui siamo interessati sono, in particolare, il fattore di cresta e la distribuzione dell'ampiezza di segnali musicali estratti da CD audio. L'allestimento di un laboratorio per misure audio richiede un investimento rilevante. Fortunatamente molte misure in banda audio oggi possono essere fatte utilizzando il pc e qualche software dedicato. In rete si trovano programmi gratuiti che eseguono l'analisi spettrale e sono disponibili schede di acquisizione, per misure audio, a costi abbordabili anche se non proprio popolari. Oggi l'analisi spettrale è una metodologia di misura alla portata di chiunque posseda un pc e un software adatto. Con questi strumenti è possibile eseguire l'analisi spettrale di una traccia riprodotta dal lettore CD presente nello spesso pc. Anzi la facilità di accesso all'analisi spettrale ha portato a trascurare altri tipi di analisi altrettanto importanti e tra questi l'Analisi Statistica della Distribuzione dell'Ampiezza (abbreviato: ASDA). L'ASDA viene utilizzata, in particolare, per l'analisi nel tempo di segnali non stazionari. Prendiamo, per esempio, il rumore prodotto da un trapano elettrico. Questo dipende dal numero di giri, dalla presenza della percussione, dal tipo di materiale che si va a forare... Per proteggere l'utilizzatore da livelli di rumori potenzialmente dannosi si devono indagare le diverse condizioni di utilizzo, valutare il livello di "rumore" prodotto ed i tempi di esposizione. In pratica il rumore del trapano viene registrato e analizzato per vedere per quanto tempo supera determinate soglie di intensità. L'ASDA dice quante volte l'ampiezza del segnale analizzato transita attraverso un certo numero di soglie predefinite

(fig.6). Il risultato dell'ASDA si riassume in una serie di quantità numeriche (media, picco, RMS ecc.) e in un grafico a barre che rappresenta la distribuzione dell'ampiezza (normalizzata rispetto ad un valore di riferimento). Per esempio potremo conoscere quante volte il segnale assume valori prossimi al valore massimo o per quanto tempo si trova al di sopra o al di sotto del 50% del massimo... e così via. L'ASDA "conta" quante volte il segnale attraversa un certo numero di soglie prefissate. La musica ha poco a che vedere con il rumore prodotto da un trapano, tuttavia, per la teoria dei segnali, la musica e il rumore del trapano vanno trattati allo stesso modo perché appartengono alla stessa categoria (segnali non stazionari). Per l'analisi dei segnali musicali sono state fissate sette soglie come riportato in figura 7. Le soglie sono fissate come percentuale rispetto al picco massimo del segnale (Massima Modulazione Relativa, MMR) per poter confrontare distribuzioni diverse. Dato che il segnale presente nei CD è campionato con passo costante "il numero di volte" rappresenta anche "per quanto tempo" l'ampiezza del segnale cade all'interno in un certo intervallo. Il CD audio contiene un segnale stereofonico, torna quindi utile anche un altro tipo di rappresentazione: le figure di Lissajous che permettono di visualizzare la distribuzione spaziale del segnale stereofonico ma soprattutto evidenziano la presenza di eventuali limitazioni di ampiezza imposte al segnale. Le figure di Lissajous si studiano al liceo (combinazione di moti armonici) e poi si dimenticano... invece sono molto utili! Praticamente, per eseguire l'ASDA di un brano musicale presente su un CD, la traccia viene estratta (via software) e salvata in formato .wav sul disco fisso del computer per essere analizzata con un programma ad hoc che esegue poche e semplici operazioni:

- 1 raddrizza il segnale (prende il valore assoluto)
- 2 trova la massima ampiezza del modulo del segnale (Massima Modulazione Relativa o MMR)
- 3 calcola i valori di soglia in base alle percentuali prestabilite
- 4 valuta l'ampiezza del segnale campione per campione e la confronta con i valori di soglia per incrementare i corrispondenti contatori
- 5 presenta i risultati

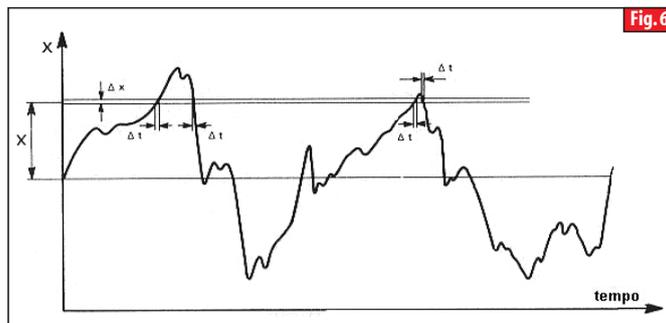
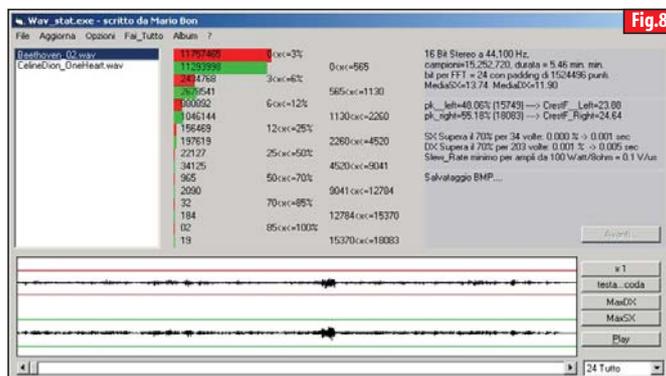


Figura 6: analisi della distribuzione dell'ampiezza: in questo esempio il segnale attraversa 4 volte la soglia prestabilita.



La scelta di raddrizzare il segnale, in linea di principio, è criticabile: si tratta di una scelta "pratica" che risparmia molto lavoro e semplifica la presentazione dei risultati. È giustificata dal fatto che la pressione media di qualsiasi segnale sonoro è nulla e quindi la distribuzione dell'ampiezza si presenta simmetrica rispetto allo zero. Lo stesso dovrebbe essere anche per il segnale musicale. La normalizzazione rispetto alla MMR è necessaria perché i CD sono registrati con livelli di segnale medio variabile: la normalizzazione consente di confrontare distribuzioni diverse. Dato che l'analisi avviene nel dominio del tempo non è necessario applicare alcun tipo di filtro o di pesatura. Nel seguito sarà utilizzato il programma wav_stat (realizzato in Visual Basic 6 da chi scrive). L'ASDA fornisce il numero esatto di volte in cui il segnale analizzato attraversa le soglie predefinite. I dati si possono interpretare anche in senso probabilistico. Supponiamo che una traccia duri 100 secondi e ammettiamo che 441 campioni (pari a un centesimo di secondo) superino la soglia dell'85% della Massima Modulazione Relativa (MMR). Ciò significa che, se osserviamo un secon-

do di segnale a caso, avremo una probabilità su cento di osservare un'ampiezza maggiore all'85% della MMR. Quando si ascolta un CD il volume dovrebbe essere regolato all'inizio della prima traccia per poi rimanere sempre nella stessa posizione. Se il volume rimane fisso il valore RMS del segnale (moltiplicato per il guadagno dell'amplificatore) rappresenta "l'equivalente termico" trasferito ai diffusori (a meno che l'amplificatore non abbia una risposta in frequenza particolarmente limitata o irregolare). Il valore RMS consente quindi di calcolare (nota la sensibilità dei diffusori acustici) il livello SPL RMS e, noto il fattore di cresta, si può prevedere il massimo livello SPL istantaneo (con tutte le opportune ipotesi del caso). Alcuni programmi commerciali (per es. Adobe Audition) analizzano i segnali musicali e propongono 3 diversi valori RMS: minimo, medio e massimo. Evidentemente questi programmi analizzano il segnale "a pezzi", il che avrebbe senso se durante l'ascolto l'utente modificasse il volume alzandolo quando è basso e abbassandolo quando è alto. I valori ottenuti poi dipendono da come sono scelti gli intervalli temporali. Non è questo il

Definizione delle soglie per l'ASDA

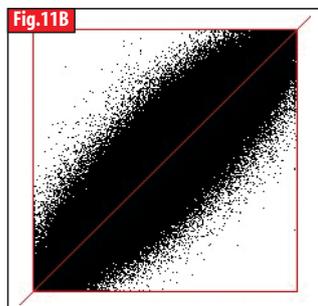
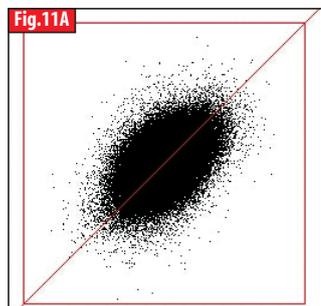
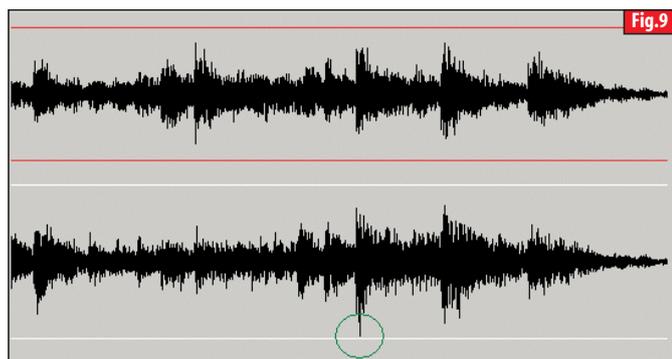
	Modulazione in %	in dB
1	da 0 a 3.125	fino a -24
2	da 6.25 a 12.5	-24 a -18
3	da 12.5 a 25	-18 a -12
4	da 25 a 50	-12 a -6
5	da 50 a 70	-6 a -3
6	da 70 a 85	-3 a -1.5
7	da 85 a 100	-1.5 a 0

caso del programma wav_stat che analizza il segnale in un'unica finestra che comprende l'intera traccia e quindi calcola correttamente un solo valore RMS.

PRIMI ESEMPI

Che cosa succede quando si esamina una registrazione molto buona e una di qualità scarsa? La figura 8 mostra l'ASDA della traccia n. 2 estratta dal CD *Beethoven - Sonate* eseguite da Emil Giles per Deutsche Grammophon (traccia "classica"). La qualità di questa traccia va classificata come "molto buona". Le barre rosse e verdi distinguono i due canali sinistro e destro. I numeri in corrispondenza delle barre indicano quante volte il segnale è "caduto" in quella determinata fascia di ampiezza. Per esempio, in questa traccia, per 11/15 del tempo complessivo il segnale rimane sotto il 3% della Massima Modulazione Relativa (MMR) e solo due campioni del canale sinistro e 19 del canale destro, sono caduti tra l'85 ed il 100% della MMR per un tempo complessivo inferiore a mezzo millisecondo (fig.9). Mezzo millisecondo è un intervallo di tempo più breve del tempo di integrazione dell'orecchio (circa 50 millisecondi) e anche più breve del "tempo di salita" del sistema uditivo (stimato in 2 millisecondi). Ma in questo momento non ci chiediamo se questi picchi siano o meno udibili. Il brano dura oltre 5 minuti, i campioni sono 44100 per ogni secondo, in totale sono stati analizzati oltre quindici milioni di campioni per canale, di questi 20 superano l'85% della MMR. Venti su quindici milioni sono decisamente pochi sia dal punto di vista statistico che dal punto di vista energetico.

La seconda traccia invece (illustrata nella fig.10) è tratta dal CD di Celine Dion *One Heart* ed è il brano che dà titolo al CD. In questo secondo caso il fattore di



cresta è decisamente basso. In questa traccia ("pop") il maggior numero di occorrenze è concentrato tra il 12 ed il 50% della MMR che in questo caso è molto prossima alla Massima Modulazione Assoluta (MMA). I segnali nel tempo sono rappresentati con la stessa scala e, dal confronto, appare molto più intenso per la traccia "pop". Nella traccia "classica", a causa della scarsa risoluzione grafica, i picchi (molto brevi) non sono visibili: un segmento temporale significativo è riportato in **figura 9** dove il picco massimo è evidenziato.

Le figure di Lissajous relative alle due tracce appena analizzate (rispettivamente **fig. 11A e 11B**) mostrano la tendenza della traccia "pop" (a destra) a superare i limiti del grafico che fissano la MMA. L'analisi più accurata della traccia "pop" (**figura 12**) mostra molti tratti dove il segnale è "clippato": si tratta di saturazioni direttamente registrate nel CD. Queste non avvengono in corrispondenza della MMA (pari a 32767) ma ad un livello leggermente più basso (32432). Ciò significa che la saturazione del segnale è avvenuta prima della conversione A/D. Ma questo è solo un dettaglio... poco importa "chi ha saturato": le saturazioni ci sono! Benché le saturazioni siano assolutamente evidenti, all'ascolto non si traducono in un suono "gracchiante" come ci si potrebbe aspettare: evidentemente

l'apparato uditivo "integra" il segnale e riesce comunque ad interpretarlo. In linea di principio più la MMR è prossima alla MMA tanto più elevati possono essere il rapporto segnale rumore ed il fattore di cresta; al contrario una modulazione molto elevata, unita ad un fattore di cresta basso, indica una registrazione compressa o saturata (di bassa qualità, come in questo caso). L'esperienza mostra che quando il segnale soggiorna sui valori massimi della MMA sono presenti saturazioni o limitazioni che vengono ben evidenziate dalle figure di Lissajous.

La qualità di una registrazione dipende da molti fattori a cominciare dalla perizia degli esecutori, il numero e la disposizione dei microfoni, il rumore di fondo, il mixaggio, equalizzazione, compressione, fino all'ultimo passaggio: il riversamento su due tracce stereo e la scrittura del master (equivalente alla preparazione della lacca di un LP). L'ASDA contiene informazioni sulla dinamica (attraverso il fattore di cresta) e sulle (eventuali) saturazioni. Va rimarcato con forza che già prima dell'invenzione del CD era possibile ottenere registrazioni di ottima qualità: oggi sembra che la qualità delle registrazioni stia subendo una involuzione, come risulta dal confronto delle analisi relative a vecchi CD riversati da master analogici comparate con quelle di CD recenti.

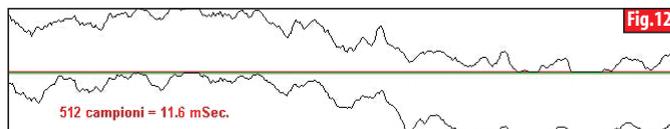
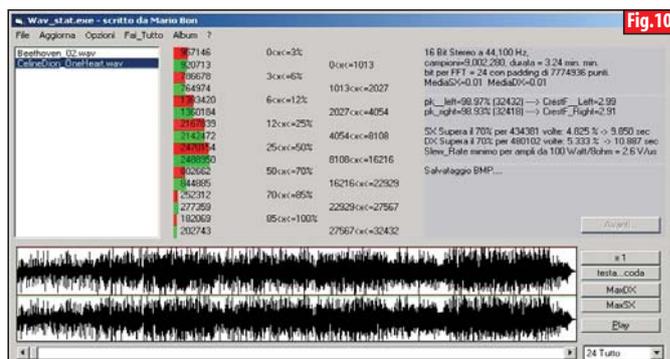


Figura 9: Beethoven, Traccia 2. Il segnale raggiunge la MMR con un unico picco isolato. **Figura 11:** Figure di Lissajous delle due tracce (A = "classica" a sinistra che appare completa, B = "pop" a destra decisamente troncata). Il rettangolo rosso rappresenta il limite di Massima Modulazione Assoluta; la diagonale rappresenta il canale centrale virtuale. **Figura 12:** Traccia "pop". Canali DX e SX, segmento di 512 campioni con evidenti saturazioni.

FATTORE DI CRESTA E MASSIMO SPL RIPRODUCIBILE

Dato un determinato amplificatore e una coppia di diffusori, quale livello SPL saremo in grado di produrre? Oltre che dalle dimensioni dell'ambiente e dalla distanza del punto di ascolto, molto dipende dal fattore di cresta del segnale. Cominciamo da qui: prendiamo un amplificatore ideale da 100 Watt RMS (40 Volt di picco sul carico) e un diffusore ideale da 90 dB @ 2.83 V/1 m. La **figura 13** elenca i livelli SPL (efficaci e di picco) ottenuti a 1 metro al variare del fattore di cresta del segnale in condizioni anecoiche.

Come si vede, più il fattore di cresta è alto, tanto più basso risulta l'SPL efficace ottenibile a parità di tensione di picco sul carico, anche la pressione di picco rimane costante. Questo succede perché l'escursione in tensione dell'amplificatore va "sacrificata" per riprodurre i picchi del segnale e a fattori di cresta crescenti corrispondono livelli RMS più bassi. Se ci atteniamo ai numeri della **figura 13** risulta che per ottenere poco più di 85 dB di SPL a 1 metro con un brano dove CF = 24 servono 100 Watt RMS e un diffusore da 90 dB. Per riportare i dati di SPL appena visti ad un tipico ambiente domestico si consideri che, con il punto di ascolto a due metri dal diffusore, si perdono 6 dB per raddoppio della distanza ma se ne guadagnano 6-7 per ef-

fetto della riverberazione dell'ambiente (supponendo T60 = 0.5 secondi e Q=2). Quindi 90 dB a un metro in camera anecoica corrispondono ancora a circa 90-91 dB (in gamma media) a due metri in un tipico ambiente da 52 metri cubi (5 x 4 x 2.6 metri). In ambienti più grandi, a distanze maggiori o con diffusori più direttivi, il livello SPL nel punto di ascolto diminuisce. Per compensare una diminuzione di 3 dB SPL si deve raddoppiare la potenza dell'amplificatore. Per interpretare i dati della **figura 13** si deve stabilire il livello SPL minimo per la riproduzione musicale in ambito domestico. La risposta non è univoca perché dipende da diversi fattori ampiamente variabili: la forma ed il volume della stanza, il suo tempo di riverberazione, la distanza del punto di ascolto, il rumore ambientale... e dal fattore di cresta del particolare programma musicale. Come possiamo dare una valutazione tenendo conto di tutte le possibili varianti? Invece di formulare ipotesi più o meno plausibili, consideriamo un prodotto commerciale come il diffusore elettrostatico Quad ESL 2905, unanimemente considerato un riferimento per la qualità della riproduzione, sottoposto a test da numerose riviste nazionali e internazionali, in ambienti e condizioni diverse, e ritenuto idoneo a sonorizzare adeguatamente l'ambiente domestico. Questo diffusore

possiede una sensibilità piuttosto bassa di circa 82 dB @ 2.83 Vrms/1 metro e può essere abbinato ad un amplificatore da 100 Watt (40 Volt di picco sul carico). Nella **figura 14** sono elencati i valori di SPL prodotti dai Quad ESL 2905, abbinati ad un amplificatore da 100 Watt, per diversi valori del fattore di cresta del programma musicale. Queste condizioni corrispondono al reale utilizzo del diffusore.

I livelli SPL riportati nelle **figure 13 e 14** vanno confrontati con i dati esposti nelle **figure 4 e 5**.

Dal confronto risulta che il diffusore Quad ESL 2905 è in grado di riprodurre un livello di 90 dB SPL pressione adeguato solo se il fattore di cresta del programma musicale da riprodurre è minore di 6. In tal caso risulta disponibile un livello di circa 90 dB SPL (con picchi di 105 a un metro o di 99 dB a 2 metri). Un CF = 6 è caratteristico di tracce non proprio brillanti mentre i Quad sono consigliati per la musica classica dove il fattore di cresta è anche quattro volte maggiore. Per esempio il brano di Beethoven analizzato poco fa presenta CF = 24 (**figura 8**) che non è il valore massimo riscontrato. Con segnali caratterizzati da CF elevati i Quad possono offrire un livello di pressione paragonabile ad una conversazione percepita alla distanza di uno o due metri (vd **fig.4**). Non sembra abbastanza anche in ambienti particolarmente silenziosi... eppure funzionano!

La **figura 15** mostra la potenza necessaria al variare della sensibilità dei diffusori acustici per CF pari a 12 e 24 quando si richiedono poco più di 91 dB (a un metro in ambiente anecoico). La potenza indicata nella quarta colonna è quella minima necessaria per ottenere l'SPL voluto e per riprodurre i picchi del segnale senza che l'amplificatore clippi. Per ogni ulteriore incremento di 3 dB di SPL la potenza va raddoppiata.

La corretta riproduzione dei picchi di segnale, anche disponendo di diffusori da 90 dB, richiede amplificatori di una certa potenza. La potenza richiesta cresce in progressione geometrica al diminuire della sensibilità del diffusore o incrementando il livello richiesto. Con i diffusori a bassa sensibilità (82 dB) e fattori di cresta elevati, sembrerebbero servire amplificatori con potenze proibitive solo per arrivare a 91 dB a un metro.

Fig. 13

SPL ottenibile in funzione del fattore di cresta (CF) per 40 Volt di tensione di picco sul carico CF=1.414 corrisponde al segnale sinusoidale usato per misurare la potenza dell'amplificatore.

CF	CF in dB	Volt RMS = 40/CF	dB SPL RMS a 1 m con diff. da 90 dB ampli da 100 Watt	dB SPL di picco
28.28	29	1.414	84	113
24	27.6	1.66	85.4	113
20	26	2	87	113
14.14	23	2.83	90	113
12	21.58	3.33	91.4	113
10	20	4	93	113
7.07	17	5.656	96	113
6	15.56	6.66	97.4	113
5	14	8	99	113
1.414	3	28.28	110	113

Fig. 14

SPL ottenibile con Quad ESL 2905 in funzione del fattore di cresta del segnale musicale. I dati sono gli stessi di fig.13 ma con 8 dB in meno (per compensare 8 dB si deve moltiplicare la potenza per 6.3 volte).

V picco (100W/8 Ohm)	Fattore di cresta del segnale	Vrms	SPL prodotto a 1 m	SPL di picco
40	28.28	0.7	70.0 dB	105 dB
40	24	1.66	77.3 dB	105 dB
40	14.14	2.83	82.0 dB	105 dB
40	12	3.33	83.4 dB	105 dB
40	6	6.66	89.4 dB	105 dB
40	1.414	28.28	102 dB	105 dB

Fig. 15

SPL RMS in dB richiesto a 1 metro	Sensibilità dei diffusori in dB se CF = 24 (27.6 dB)	Sensibilità dei diffusori in dB se CF = 12 (21.6 dB)	Potenza richiesta in W (RMS su 8 Ohm) senza clipping
91.4	102	96	25
91.4	99	93	50
91.4	96	90	100
91.4	93	87	200
91.4	90	84	400
91.4	87	81	800
91.4	84	78	1600
91.4	81	75	3200

In conclusione abbiamo un diffusore che, in base all'esperienza, è adeguato per ricostruire il livello e la dinamica per la corretta riproduzione dell'evento musicale ma che, in base ai calcoli, non dovrebbe essere in grado di farlo (anche se si dovesse considerare il contributo della dispersione dipolare che incrementa il livello in ambiente di un paio di dB).

Prima di abbandonare questo argomento va fatta una considerazione in merito all'impedenza elettrica dei diffusori acustici commerciali. Secondo le norme DIN, l'impedenza non dovrebbe

scendere sotto 3.2 Ohm (4 Ohm nominali) anche se oggi si tende ad accettare, con un eccesso di tolleranza, valori anche molto inferiori. Non è raro incontrare diffusori acustici con minimi di impedenza attorno a due Ohm e qualche modello anche sotto 1 Ohm (!). Quello che conta per riprodurre i picchi del segnale è la massima tensione prodotta sul carico: tanto più basso è il carico tanta più corrente dovrà erogare l'amplificatore. Quando l'amplificatore deve erogare corrente, la massima tensione sul carico si riduce (in misura diversa per ogni

amplificatore...). Da questo punto di vista anche l'impedenza dei diffusori ha un effetto sulla qualità del suono riprodotto perché induce l'amplificatore a saturare prematuramente (non tutti e non nella stessa misura...).

PRIME CONCLUSIONI

I contenuti di questa prima parte suggeriscono alcuni spunti di riflessione:

- 1 il fattore di cresta del segnale musicale raggiunge e supera il valore 24 (27.6 dB)
- 2 la differenza di qualità tra registrazioni diverse è molto marcata (per non dire abissale)
- 3 le tracce dei CD possono contenere saturazioni
- 4 per riprodurre un livello SPL adeguato, senza saturare l'amplificatore, sono necessari amplificatori molto potenti e/o diffusori molto sensibili
- 5 un diffusore da 82 dB e un ampli da 100 Watt sono adeguati per la riproduzione hi-fi domestica
- 6 i punti 4 e 5 sono in contraddizione.

All'ascolto la differenza "sonora" tra le due tracce "pop" e "classica" utilizzate come esempio è molto evidente: mentre il pianoforte di Emil Giles è preciso e dinamico la traccia di Celine Dion suona piatta e abbastanza confusa. Per quanto riguarda le saturazioni di percepirle dipende da quanto durano e da quanto sono ravvicinate tra loro. Ma si è sempre supposto che il principale responsabile (delle saturazioni) fosse il "clipping" dell'amplificatore e non quello del CD! All'ascolto si può essere indotti a confondere le saturazioni presenti nel CD con un difetto della catena di riproduzione e la colpa di solito viene attribuita ai diffusori acustici. Sempre a seguito dell'analisi di queste due prime tracce è lecito sospettare che, in generale, la dinamica dei segnali musicali sia sottostimata; contrariamente in commercio si troverebbero solo amplificatori da oltre 100 Watt e diffusori con sensibilità minima di 90 dB @ 2.83 V/1 m. L'esempio dei Quad ESL 2905 indica che un diffusore (molto) poco sensibile, abbinato ad un amplificatore da 100 Watt, riproduce un livello di ascolto adeguato. La contraddizione con i dati calcolati è evidente. (continua nel prossimo numero) 🐣