

Virtù, vizi, pregi e difetti nella **percezione** della musica

(Seconda parte)

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

L'analisi spettrale sottopone a FFT l'intera traccia audio in una singola "finestra di acquisizione". Per analizzare un singolo canale di una traccia che dura tre minuti si devono processare oltre otto milioni di campioni: il tempo richiesto dal calcolo è incompatibile con un'analisi "in tempo reale", per contro i risultati sono attendibili e ripetibili. Nelle appendici si trovano diversi capitoli dedicati ai segnali, all'analisi spettrale e agli errori introdotti dalla FFT. L'analisi spettrale viene realizzata

Nella prima puntata di questa serie di articoli abbiamo spiegato come il "fattore di cresta" rappresenti un elemento essenziale da esaminare per valutare la qualità di una registrazione: in tal senso sono state presentate le **Analisi Statiche della Distribuzione dell'Ampiezza (ASDA)** di due tracce (classica e pop) caratterizzate da fattori di cresta pari a 3 e oltre 24. È noto che le registrazioni siano elaborate con compressori di dinamica, limitatori, filtri ecc. (addirittura indispensabili per realizzare un LP). In questa seconda parte viene introdotta l'analisi spettrale, ulteriore elemento di discriminare della qualità della musica registrata.

attivando l'opzione FFT del wav_stat (lo stesso programma, realizzato in Visual Basic 6 da chi scrive, che calcola l'ASDA). Appare così un ulteriore grafico con lo spettro (fig.1). Anche lo spettro rappresenta quantità relative: il livello 0 dB è normalizzato rispetto all'ampiezza della riga spettrale di

ampiezza massima. Per questo motivo nelle tracce registrate a livello più basso il rumore di fondo è più evidente. L'analisi spettrale viene fatta su un solo canale in quanto le frequenze basse sono sostanzialmente monofoniche e le differenze spettrali tra i canali destro e sinistro, tranne casi partico-

lari, sono modeste. Il grafico dello spettro abbraccia sei decadi da 0.02 a 20000 Hz per mostrare anche le componenti subsoniche più profonde. Dato che il numero di campioni da analizzare non è mai pari ad una potenza di 2, è necessario "allungare il segnale" aggiungendo un opportuno numero

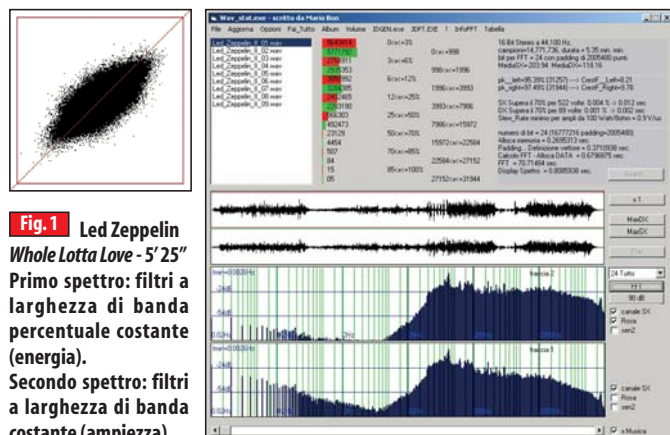


Fig. 1 Led Zeppelin *Whole Lotta Love - 5' 25"*
Primo spettro: filtri a larghezza di banda percentuale costante (energia).
Secondo spettro: filtri a larghezza di banda costante (ampiezza).

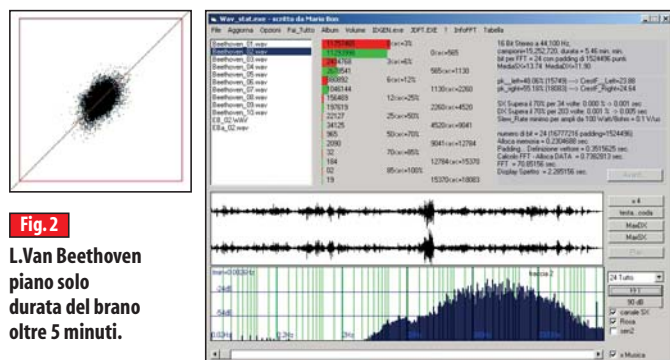


Fig. 2 L. Van Beethoven piano solo
durata del brano oltre 5 minuti.



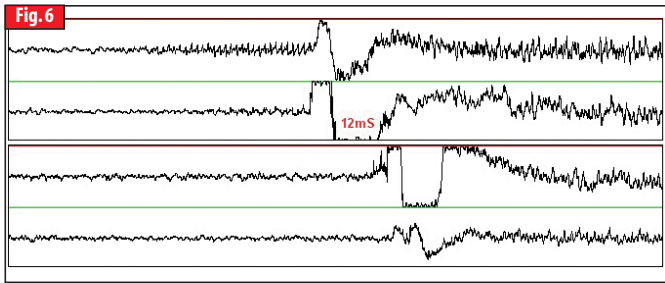
Fig. 3 Segmento temporale di 4092 campioni (Alicia Keys, *Songs in A Minor* traccia 4). Le saturazioni durano 1.5-2.5 millisecondi.



Fig. 4 Segmento temporale di 1024 campioni con saturazioni. (L. Pausini: *La mia banda suona il rock da Io canto*). Queste "tosature" si ripetono lungo tutto il brano come mostrato dalle figure di Lissajous.



Fig. 5 Figure di Lissajous: Alicia Keys, Laura Pausini, Celine Dion e Norah Jones da sinistra a destra.



di zeri in testa e in coda fino a raggiungere il numero richiesto dalla FFT. Questa operazione si chiama *padding* e, come insegna la teoria, non altera l'involuppo dello spettro; tuttavia questa operazione interferisce con un eventuale offset presente nel segnale e altera lo spettro a bassissima frequenza.

L'unica finestra di pesatura applicata prima del padding altera il valore dei primi quattro campioni in testa e degli ultimi 4 in coda per azzerare il segnale (0, 1/4, 2/4, 3/4... 3/4, 2/4, 1/4, 0). Quando la traccia inizia e finisce nel silenzio, come dovrebbe essere, questa "pesatura" cade su valori già nulli ed è ininfluente. Lo spettro ottenuto viene rappresentato così come sarebbe visto attraverso un banco di filtri a larghezza di banda percentuale costante. In altre parole se all'ingresso di questo analizzatore fosse applicato rumore rosa lo spettro apparirebbe piatto. In altre parole ancora questo è un analizzatore a "ennesimi di ottava". Questa scelta dipende dal fatto che siamo più interessati all'energia che all'ampiezza: bande di frequenza (della stessa larghezza percentuale) che appaiono allo stesso livello, trasportano la stessa quantità di energia e richiedono la stessa potenza all'amplificatore. Per rendere più chiaro questo concetto la figura 2 mostra lo stesso spettro nelle due diverse rappresentazioni Energia (sopra) e Ampiezza (sotto).

Cominciamo dunque con l'analisi della traccia *Whole Lotta Love* dal CD *Led Zeppelin II* (riversato da master analogico). I Led Zeppelin sono un gruppo rock composto da quattro elementi: chitarra, basso, batteria e voce solista. Osservando il risultato dell'ASDA di **figura 1**, si nota che il massimo assoluto, sia sul canale destro che sinistro, è molto vicino alla Massima Modulazione Assoluta (32767) e che il valore di cresta raggiunge 9.78. Dal grafico a barre (distribuzione dell'ampiezza) si evince che

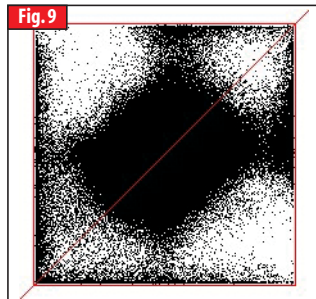
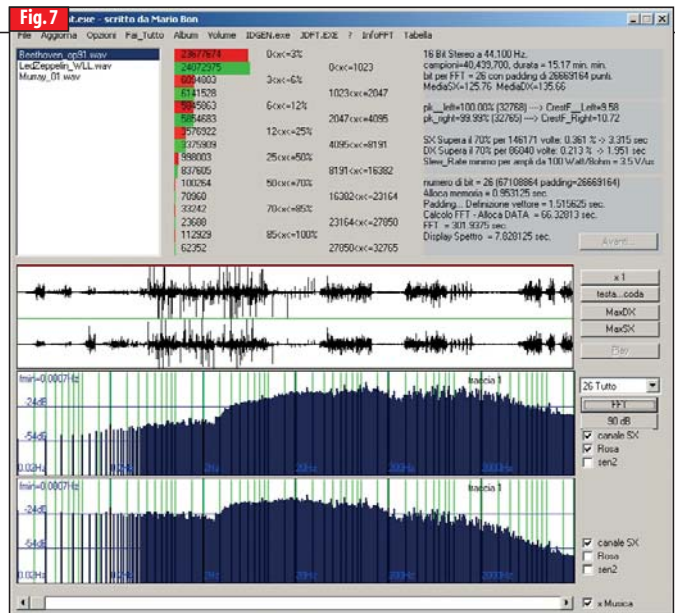


Figura 6 • Beethoven, La vittoria di Wellington, Opera 91 (Telarc). Il "suono" prodotto da una saturazione così profonda non può essere distinto dal clipping dell'amplificatore o dell'altoparlante e si è quindi indotti ad attribuire alla catena di riproduzione un difetto della registrazione.

Figura 7 • Beethoven, Opera 91 (Telarc). Primo spettro (alto): filtri a larghezza di banda percentuale costante (energia). Secondo spettro (basso): filtri a larghezza di banda costante (ampiezza).

Figura 8 • Rappresentazione di Lissajous per l'Opera 91.

solo qualche centinaio di campioni (per un tempo totale inferiore a 16 millisecondi) supera il 70% della MMR e solo 20 vanno oltre l'85% della MMR. In questa traccia non sono presenti saturazioni. In generale si tratta di una registrazione ben fatta (almeno per quanto riguarda la regolazione dei livelli ed il riversamento) come confermato dalla figura di Lissajous (**fig.1**) dove la "macchia nera" appare regolare e ben distante dai limiti di saturazione (quadrato rosso). Il fatto che l'ellissoide non sia perfettamente allineato lungo la diagonale indica un leggero sbilanciamento tra i canali (il canale centrale virtuale non è esattamente al centro). Anche la compressione dinamica appare contenuta. Per quanto riguarda lo spettro si nota la concentrazione attorno a 80 Hz e quindi un progressivo calo delle basse frequenze sotto i 40 Hz. Del resto la nota più grave



I Doors a 24 bit?

Non è raro che vecchie registrazioni siano riproposte in versione "rimasterizzata" magari a 24 bit. Il caso vuole che esistano due versioni, anzi due riversamenti su CD, dello stesso LP dei Doors *The Doors*: la prima è stata acquistata in negozio ed è presentata come "rimasterizzata a 24 bit"; la seconda invece è stata acquistata in edicola come allegato a *Sorrisi e Canzoni TV* (dicembre 2007). La versione rimasterizzata è anche più costosa. I due CD contengono gli stessi brani nello stesso or-

dine: confrontiamo la seconda traccia. La versione rimasterizzata a 24 bit presenta il fattore di cresta più basso (passato da 8.97 a 4.61) e un numero di saturazioni decisamente superiore. Come si vede lo spettro appare diverso specie agli estremi della banda audio. Anche il volume della traccia destra è stato alterato come risulta anche dalle figure di Lissajous. Entrambi i CD sono stati analizzati per intero. Non servono ulteriori commenti se non uno: al peggio non c'è limite!

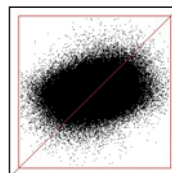


Fig. E
The Doors
traccia 2
versione commerciale

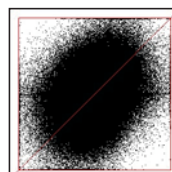
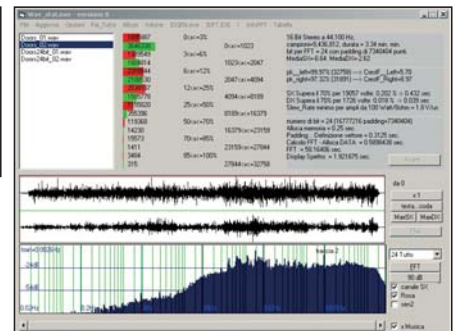
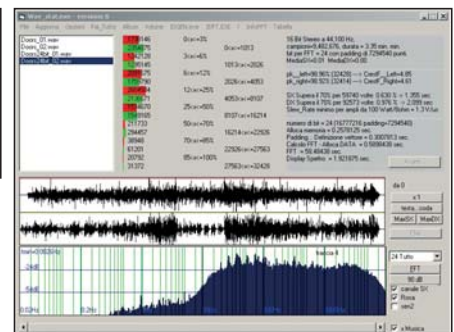


Fig. F
The Doors
traccia 2
versione rimasterizzata
a 24 bit 96 kHz.



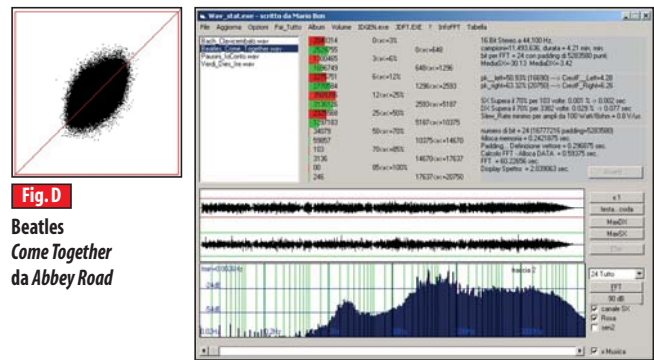
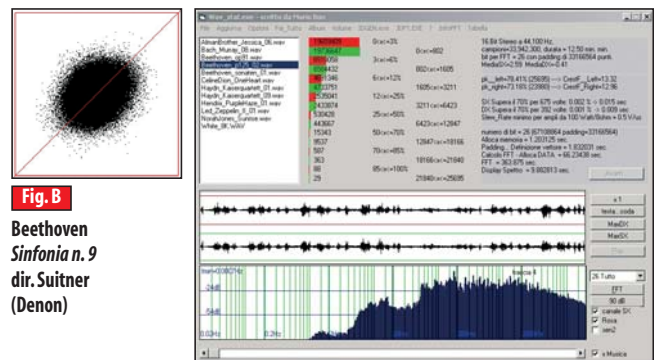
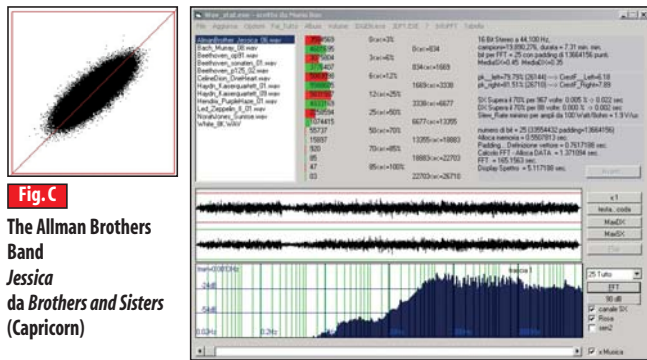
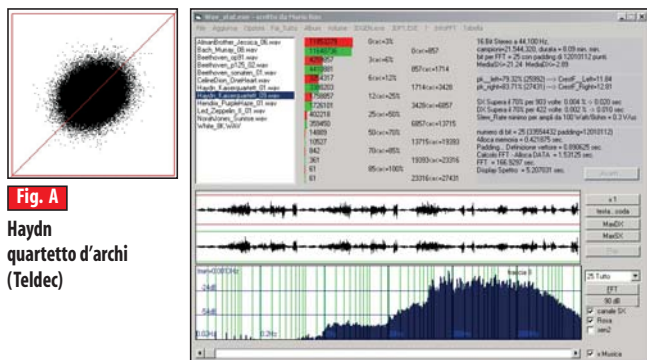
Alcuni esempi di ASDA

Nelle parti precedenti sono stati presentati alcuni esempi commentati di ASDA e di analisi spettrale. Dato che questo vuole essere uno studio statistico vi si devono analizzare un congruo numero di tracce. Qui di seguito vengo-

no proposte una serie di ASDA che il lettore sarà in grado di commentare autonomamente. Il fattore di cresta è riportato nella parte destra della figura (quinta e sesta riga CrestF_Right CrestF_Left). Si noti che, quando il valo-

re medio del segnale è prossimo a zero, non sono presenti componenti sub-soniche. Va rimarcato che quelli riportati sono gli spettri di tracce intere e quindi uno scarso contenuto di basse frequenze non significa che "i bassi non ci sono" ma che la partitura non prevede una linea di basso continua (tipica della musica jazz, rock o pop). Una parte delle tracce scelte, specie quelle di musica

moderna, sono datate per dimostrare che le buone registrazioni esistono da sempre. Molte registrazioni recenti non si sono dimostrate all'altezza dei vecchi master analogici riversati in digitale. Come atteso, i fattori di cresta misurati per la musica classica sono mediamente superiori rispetto alla musica moderna. Nelle appendici sono riportate le analisi di oltre 600 tracce.



prodotta dal basso elettrico è il Mi a circa 40 Hz. L'energia rimane abbastanza costante da 100 a 4000 Hz per poi calare più decisamente verso l'estremo alto. Si tenga conto che la dinamica in ascissa vale 90 dB (con la sommità del grafico a +6 e il fondo a -84 dB). Questa è una traccia non banale da riprodurre sia come dinamica che come estensione: in particolare la gamma media richiede altoparlanti decisamente robusti. Sotto i 6 Hertz sono presenti sub-sonici con contenuto energetico ridotto. Si noti la differenza tra la rappresentazione proporzionale all'energia e quella proporzionale all'ampiezza: l'ampiezza delle componenti sub-soniche è alta ma trasporta poca energia. Le componenti a bassissima frequenza sono presenti perché non è

Tabella A	Celine Dion	Led Zeppelin	Beethoven	Stravinsky
CF = Vpico/RMS	2,92	9,78	24,64	26,28
V RMS sul carico	2,83	2,83	2,83	2,83
Vpico sul carico	8,26	27,67	69,73	74,37
Vpico = CF x 2.83 (Volt)				
Potenza in Watt RMS su 8 Ohm	4,26	47,87	303,9	345,7
SPL a 1 m. con diffusore da 90 dB 2.83V/1m	90	90	90	90
SPL di picco a 1 m. con diffusore da 90 dB 2.83V/1m	99,3	109,8	117,8	118,4

stato corretto, in sede di pre-trattamento dei dati, un qualche offset lentamente variabile presente nel segnale (il valore medio non è nullo...). In effetti il segnale non è sottoposto ad alcun pre-trattamento prima della FFT (a parte la blanda pesatura descritta prima).

Il secondo esempio riguarda il brano per pianoforte già analizzato nella prima puntata (Beethoven, *Sonaten*, eseguite da Emil Giles su Deutsche Grammophon - traccia 2). Si vedano le **figure 1 e 2**. Questo brano evidenzia la dinamica del pianoforte registrato,

in questo caso, sfruttando poco più del 50% della MMA per non incappare in saturazioni. Il fattore di cresta è decisamente alto (23.88 e 24.64 sui due canali). La distribuzione dell'ampiezza è tipica delle registrazioni di ottima qualità. Per quanto riguarda lo spettro questo trova il suo massimo attorno all'ottava centrale del pianoforte (La centrale = 440 Hz) e si attenua sotto i 100 Hz e oltre i 2000 Hz. Ciò non significa che le frequenze basse siano assenti ma che questa partitura utilizza prevalentemente le ottave centrali della tastiera (cosa vera in generale per tutte le tastiere). Nella musica pop e rock la linea del basso è continua e persistente (e manifesta la sua energia nello spettro) mentre nella partitura del pianoforte non lo è ed il contenuto

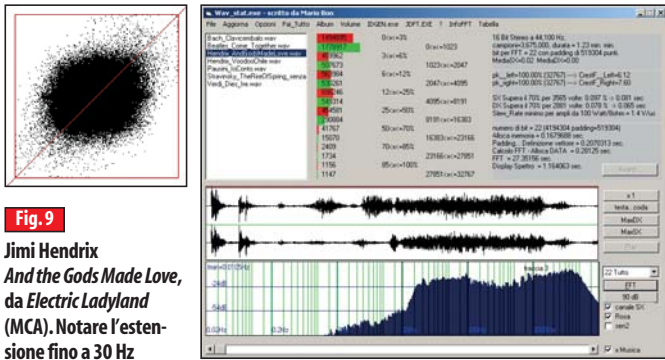


Fig. 9
Jimi Hendrix
And the Gods Made Love, da Electric Ladyland (MCA). Notare l'estensione fino a 30 Hz

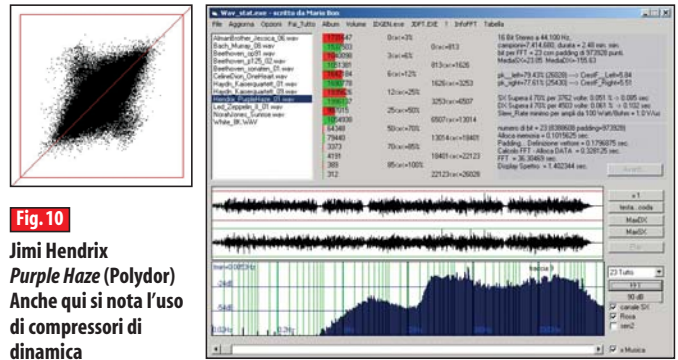


Fig. 10
Jimi Hendrix
Purple Haze (Polydor)
 Anche qui si nota l'uso di compressori di dinamica

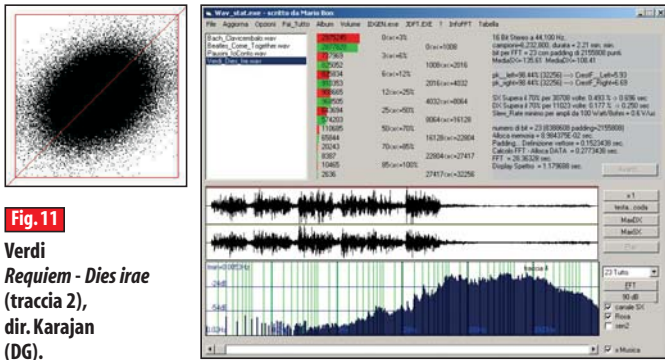


Fig. 11
Verdi
Requiem - Dies irae (traccia 2), dir. Karajan (DG).

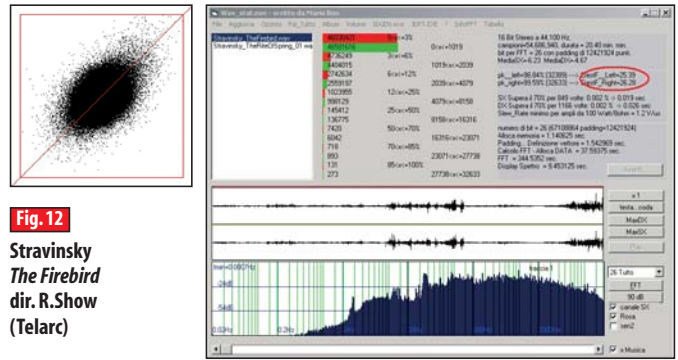


Fig. 12
Stravinsky
The Firebird dir. R. Show (Telarc)

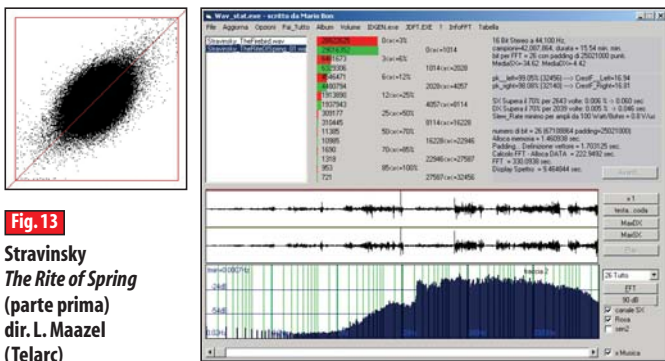


Fig. 13
Stravinsky
The Rite of Spring (parte prima) dir. L. Maazel (Telarc)

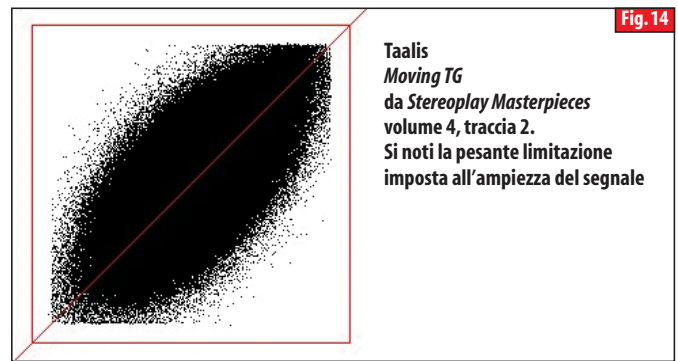


Fig. 14
Taalis
Moving TG da Stereoplay Masterpieces volume 4, traccia 2.
 Si noti la pesante limitazione imposta all'ampiezza del segnale

energetico alle basse frequenze, distribuito su tutta la durata del brano, risulta inferiore. La figura di Lissajous relativa a questa traccia appare come una macchia ellissoidale ben lontana dai bordi (fig.1). La **tabella A** pone a confronto 4 tracce: le due appena viste più una traccia con CF molto basso e la traccia con il CF più alto riscontrato. Supponendo di voler ottenere 90 dB SPL a un metro (senza saturare l'amplificatore) la traccia dei Led Zeppelin richiede poco meno di 50 Watt mentre la traccia di Beethoven (analizzata nella scorsa puntata) ne esige più di 300 (oltre 6 volte in più) e Stravinsky 345. La traccia di Celine Dion (anch'essa analizzata nella scorsa puntata) si accontenta di 5 Watt scarsi.

Ancora una volta i "numeri" dicono che, per riprodurre le tracce ad alta dinamica mantenendo l'amplificatore lontano dal clipping, servono potenze notevoli.

Se solo volessimo aumentare il livello della riproduzione di 3 dB (o utilizzare un diffusore da 87 dB), la potenza andrebbe raddoppiata per arrivare a 600-700 Watt RMS per canale. Andando avanti capiremo perché riusciamo ad ascoltare la musica anche con amplificatori molto meno potenti.

TRACCE CON SATURAZIONI

Finora, abbiamo visto una serie di tracce "di buona qualità" nel senso che, al di là delle qualità artistiche o delle tecniche di ripresa del suono, quanto meno non contengono saturazioni; veniamo ora a tracce

che contengono delle saturazioni evidenti... Per trovare le saturazioni è necessario osservare il segnale nel tempo, un'operazione lunga e noiosa anche se, fortunatamente, queste saturazioni sono ben evidenziate dalle figure di Lissajous (in rosso i limiti di saturazione, la diagonale identifica il canale centrale). Cominciamo con le tracce estratte dai CD di Alicia Keys, Laura Pausini, Celine Dion (già vista nella prima parte) e Norah Jones. Sono tutti CD abbastanza recenti e di successo. A dispetto di ciò sono caratterizzate da numerose e ricorrenti saturazioni come dimostrano le **figure 3, 4 e 5**. Le figure di Lissajous (**fig.5**) parlano da sole: è evidente la tendenza ad uscire dai limiti di MMA. Le saturazioni d'altronde sono presenti anche su CD la cui quali-

tà dovrebbe essere al di sopra di ogni sospetto. Si veda la traccia *La vittoria di Wellington* (Opera 91 di Beethoven - Telarc). Questa traccia (**fig.6 e 7**) è famosa per le cannonate e le raffiche di moschetto. La figura 4.4 mostra alcune delle severe saturazioni presenti (una in particolare dura 12 millisecondi ed è immediatamente preceduta da un'altra di circa 4 millisecondi) ed è proprio una delle "cannonate". Sembrerà strano ma questi segnali rischiano di schiantare i tweeter dei diffusori acustici! La figura di Lissajous per l'Opera 91 (**fig.8**) è quasi irriconoscibile. Malgrado tutto il fattore di cresta vale circa 10 per i due canali. C'è da chiedersi a quanto sarebbe arrivato senza le saturazioni... Quando sono presenti saturazioni il

fattore di cresta misurato è corretto (dal punto di vista del calcolo) ma minore del valore che avrebbe raggiunto se il segnale avesse potuto “continuare la sua corsa”. L’ASDA relativa all’Opera 91 è quella di una traccia ad alta dinamica ma con picchi persistenti (sul canale sinistro il segnale supera per circa 3 secondi complessivi l’85% della MMR). Il contenuto spettrale a bassa frequenza è molto intenso ma è difficile dire quanto sia condizionato dalle saturazioni (cose ancora peggiori si incontrano nelle tracce delle colonne sonore dei film che contengono effetti speciali: astronavi, dinosauri, ecc.). Cambiando completamente genere musicale, l’esame di un paio di brani di Jimi Hendrix (fig.9 e 10) dimostra come anche la musica rock possa essere prodotta con uno standard qualitativo dignitoso: le saturazioni, pur presenti, interessano complessivamente meno di 50 millisecondi.

Poi, sempre a titolo di esempio, le analisi di tre brani: Il *Dies irae* di Verdi (fig.11) e due famose com-

posizioni di Stravinsky (fig. 12 e 13). La traccia di Verdi presenta una modulazione molto alta con una limitazione all’ampiezza del segnale, il fattore di cresta è limitato a 6 (questo spiega come mai, chi scrive, non sia mai riuscito a sentire questo *Dies irae* decentemente riprodotto). Stravinsky presenta i fattori di cresta più alti riscontrati (oltre 26) ma anche qualche saturazione. Dopo le registrazioni “ben fatte” e le registrazioni “mal fatte”, abbiamo così individuato altri due tipi di registrazioni “al limite” (Verdi) e “ben fatte ma con qualche saturazione”(Stravinsky). In effetti le tracce di Stravinsky, se non fosse per le poche saturazioni presenti, dovrebbero essere classificate come ottime: suonano comunque in modo ottimo. Per quanto riguarda il *Dies irae* di Verdi siamo “al limite”: all’ascolto la registrazione appare buona ma, anche migliorando la qualità della catena di riproduzione, non si riesce a migliorare la qualità oltre un certo limite e rimane sempre un senso di insoddisfazione: il problema non è

nell’impianto ma nella registrazione! Queste sono le tracce più pericolose perché inducono ad attribuire i difetti della riproduzione all’amplificatore o ai diffusori acustici che, almeno in questo caso, non hanno colpe.

MOMENTANEE CONCLUSIONI

Errare humanum est... anche quando si prepara un CD? Le saturazioni sono tante e si trovano dappertutto ma *perseverare diabolicum* (si veda “Questione di mastering” – SUONO n. 409 pagg. 46-47). I CD di Alicia Keys, Norah Jones e Laura Pausini (ma anche Madonna e Diana Krall) hanno ottenuto premi e riconoscimenti, sono stati preparati negli studi di registrazione di importanti case di produzione da tecnici sicuramente competenti, ascoltati da moltitudini di “addetti ai lavori” (critici, giornalisti, ecc.); è difficile credere che il risultato ottenuto non sia quello voluto e la stessa cosa si può dire per le tracce di musica classica: Karajan partecipava attivamente alle fasi di post produzione...

A tutto però c’è un limite; osserviamo la figura 14: riporta la figura di Lissajous relativa alla traccia numero 2 del CD *Stereoplay Masterpieces* - volume 4 (prodotto dalla rivista tedesca *Stereoplay*) e presentato come “45 minuti di musica audiophile”. Come si vede l’ampiezza del segnale è pesantemente limitata e, all’ascolto, la qualità del suono si presenta per quello che è: eufemisticamente non buona! Resta il fatto che un’autorevole rivista tedesca ha inserito questa traccia in un CD di riferimento. Un errore? Una svista? Se per la musica pop possiamo supporre che si sia voluto privilegiare la riproduzione sui lettori MP3 portatili (svilendo consapevolmente la qualità delle registrazioni), per le tracce di musica classica dobbiamo supporre che le (poche) saturazioni presenti siano passate inosservate all’ascolto. Registriamo poi che la sensibilità alle saturazioni cambia da nazione a nazione: in Germania, per esempio, vale tutto... in Italia no. (continua)

