Politecnico di Torino

Politecnico di Torino

Ingegneria del Cinema e dei Mezzi di Comunicazione
A.a. 2020/2021
Sessione di Laurea Luglio 2021

Audio Analogico e Digitale

Analisi e confronto tra le tecniche di registrazione e riproduzione musicale in analogico e in digitale.

Relatori:
Marco Carlo Masoero
Renato Capajola

Candidati:
Giorgio Antonio Lepori
s255462

Politecnico di Torino

Indice

Abstract **1**
1 - Storia dei Supporti **2**

1.1 Le origini 2
1.2 Vinile 5
1.3 Tape 7
1.4 Digitale fisico 8
1.5 Streaming e il vinyl revival 9

2 - Analisi del Supporto Analogico **10**

2.1 Sistema di funzionamento di un giradischi 10
2.1.1 Trasmissione meccanica della rotazione 11
2.1.2 Lettura 11
2.1.3 Testina 12
2.2 Processo di fabbricazione del vinile 14
2.2.1 Incisione 14
2.2.2 Equalizzazione 16
2.2.3 Centratura e pressaggio 18

3 - Analisi del Supporto Digitale **23**

3.1 Intro 23
3.2 CD 24
3.2.1 Caratteristiche 25
3.2.2 Lettura Laser 25
3.2.3 CAV e CLV 26

3.2.4 Codice di rilevamento....26
3.2.5 Codifica CIRC....27
3.2.6 Codifica Efm....28
3.2.7 Red book....30
3.2.8 Yellow book....31
3.3 DVD....32
3.3.1 DVD Audio....33
3.4 SuperAudio CD....34
3.4.1 DSD....36
3.5 DAT....37

4 - Conversione Analogico Digitale....38

4.1 Introduzione....38
4.2 Catena Audio Digitale....39
4.3 Campionamento....42
4.3.1 Teorema di Nyquist e Aliasing....43
4.4 Quantizzazione....46
4.4.1 Quantizzazione lineare....46
4.4.2 Dithering....47
4.5 Codifica PCM....48
4.6 ECC – Codici Correzione Errore....50
4.7 Sample Rates....51
4.7.1 Oversampling....54
4.8 Noise shaping....56
4.9 Modulazione sigma delta....57
4.10 Quantizzazione non lineare....58
4.11 Quantizzazione logaritmica....59
4.12 Quantizzazione virgola mobile....61
4.13 Ricostruzione del segnale....63
4.14 Jitter....65
4.14.1 Jitter nei Processori....68

Politecnico di Torino

4.14.2 Misurazione jitter....71
4.15 Errore di framing....74
4.15.1 Reclocking....76

5 - Confronto tra segnale musicale digitale e segnale musicale analogico....79

5.1 Separazione dei Canali....79
5.2 Perdita di informazione....80
5.3 Feedback....82
5.3.1 Rumble Feedback....82
5.3.2 Tip Mass Resonance....84
5.4 Rumore....85
5.5 Distorsione....85
5.6 Elaborazione del Segnale....86
5.6.1 Degradazione oggettiva e miglioramento soggettivo....86
5.6.2 Strumenti d'analisi....88
5.6.3 Mixing Analogico e Digitale....91
5.6.4 Equalizzatori....93
5.6.5 Processori di dinamica....96
5.6.6 Simulatori analogici....98
5.6.7 Limiter....102
5.6.8 Conclusioni....105
5.7 Gamma Dinamica....106

6 - Loudness War....107

6.1 Introduzione....107
6.2 Storia....109
6.2.1 Escalation....109
6.2.2 Streaming e LUFS....116
6.3 Ristampe....119

7 – Conclusioni.....**123**

Bibliografia e Sitografia.....**124**

Ringraziamenti.....**126**

Politecnico di Torino

Abstract

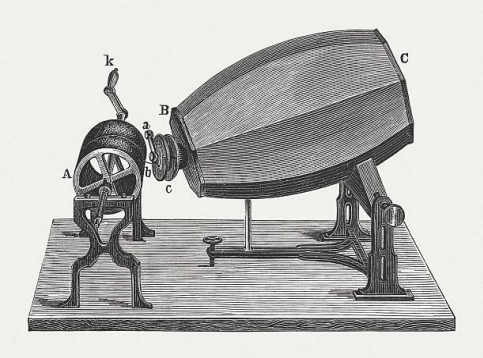
La musica registrata è stata in costante evoluzione fin da quando è stato possibile l'incisione del suono su un supporto che fosse riproducibile e diffondibile. La diffusione delle tecnologie di registrazione e dei medium contenenti audio nel mercato di massa è paragonabile all'invenzione dei caratteri mobili di Gutenberg, quando l'invenzione della tecnica di stampa permise la diffusione del pensiero critico occidentale e pose le basi per il mondo moderno; allo stesso modo la musica incisa su un supporto che potesse essere riprodotto ha avuto un ruolo importante nella diffusione di culture, movimenti ed idee su scala globale, accelerando la globalizzazione e la creazione del villaggio globale.
I primi formati che vennero utilizzati furono quelli analogici: essi si basano sul fatto che l'audio in essi contenuto sia memorizzato sotto forma di segnale analogico, cioè senza discretizzazione e conversione in numeri tipica del segnale digitale. Il principale medium analogico è il vinile. Esso viene considerato dagli audiofili come un formato superiore sia per la riproduzione e l'ascolto e, negli ultimi tempi, ha vissuto una rinascita nel mercato musicale superando in vendite il CD, che lo aveva soppiantato come supporto predominante. Lo scopo di questa tesi è mettere al confronto le due tecnologie, attraverso l'analisi dei processi di fabbricazione e le tecniche utilizzate in fase di registrazione dell'informazione per mettere in luce quale delle due sia, dal punto di vista scientifico, più fedele al segnale originale, approfondendo anche le differenze tra gli elaboratori di segnale analogici e digitali.
Nel Capitolo 1 sarà trattata la storia dei supporti e la loro evoluzione nel tempo, illustrando i principali formati utilizzati nella storia della musica. Nel Capitolo 2 sarà trattato il supporto analogico, prendendo in esame il sistema di funzionamento di un giradischi e il processo di produzione di un disco in vinile. Nel Capitolo 3 verranno considerati i principali formati dell'audio digitale, concentrandosi in particolare sul CD e il processo di memorizzazione dei codici binari che rappresentano l'audio. Nel Capitolo 4 si analizzerà la conversione audio digitale, mostrando sia i principi di teoria del segnale che ne permettono la corretta riuscita e sia le principali problematiche che si possono creare. Nel Capitolo 5 si metteranno a confronto i due medium in base ai criteri della riproduzione audio di qualità, dimostrando che alcuni fenomeni possono essere descritti sia in maniera soggettiva e sia in maniera oggettiva; in più verranno confrontati gli elaboratori dell'audio analogici e digitali normalmente utilizzati nella fase di lavorazione del segnale per identificare quale sia la loro influenza sul risultato finale. Nel Capitolo 6 sarà esposta la loudness war, un fenomeno indipendente dalle tecnologie utilizzate che spiegherà il motivo per cui i due medium suonano diversamente.

Politecnico di Torino

1 - Storia dei Supporti

1.1 Le origini

Il primo apparecchio creato per la registrazione dell'audio, in questo caso della voce umana, fu il fonoautografo. Esso fu brevettato da Léon Scott nel 1857 e si basava su un sistema composto da un diaframma vibrante e da una puntina di gesso. Il suo scopo era registrare in maniera grafica le onde sonore sotto forma di linee su dei fogli di carta, ma solamente per analisi visiva e senza nessun intento di riproduzione. In pratica l'esecutore si collocava nei pressi della membrana di registrazione e, mentre eseguiva la propria perfomance, metteva in moto con una manovella un cilindro su cui veniva disegnato dal gessetto il tracciato della variazione in ampiezza della membrana: esso era un corrispettivo grafico dell'onda sonora generata e concettualmente è paragonabile alla traccia audio che viene rappresentata in un editor digitale.
Nel 2008, queste tracce furono scannerizzate per la prima volta da degli ingegneri audio e convertiti digitalmente in un suono udibile: questi fonoautografi si scoprirono di essere filastrocche e discorsi incisi da Scott nel 1860, insieme ai suoni di diapason e di frammenti vocali incomprensibili; queste sono le più antiche registrazioni del suono conosciute nella storia dell'uomo.



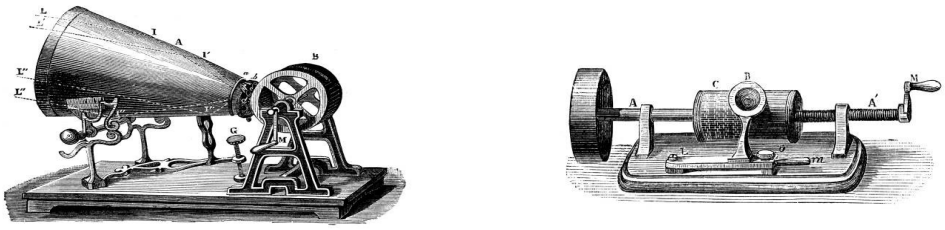
Fonoautografo di Scott
fonte: gettyimages

Politecnico di Torino

Tracce del fonoautografo registrate da Scott su carta carbone
fonte: gettyimages

Nel 1877, Thomas Edison inventò il fonografo. Al contrario del fonoautografo, esso poteva sia registrare che riprodurre il suono. Nonostante la similarità nel nome, non ci sono prove documentate che il fonografo di Edison fosse basato sul fonoautografo di Scott. Edison tentò prima di registrare il suono su un nastro di carta impregnato di cera, con l'idea di creare un registratore per un'altra sua invenzione, cioè il telegrafo, facendolo in pratica un antesignano della nostra segreteria telefonica.
Nonostante i risultati lo resero confidente sul fatto che il suono potesse essere registrato su supporto fisico e riprodotto, la fragilità del nastro non permise di riprodurne il suono; ebbe invece successo nel suo esperimento successivo, nel quale utilizzò una pellicola di stagno come medium di registrazione. Questa pellicola di stagno era arrotolata intorno ad un cilindro di metallo il quale era in contatto con una puntina di metallo che, messa in vibrazione dal suono, incideva la pellicola mentre il cilindro veniva rotato. La registrazione poteva essere riprodotta immediatamente cambiando la puntina di registrazione con una puntina di riproduzione e sostituendo il cono di registrazione con un cono di riproduzione, che era più ampio. Il fonografo di Edison si basava su un sistema di riproduzione e registrazione del suono completamente meccanico, per cui la qualità del audio era molto bassa e i cilindri avevano la brutta abitudine di consumarsi dopo una decina d'ascolti.

Politecnico di Torino



Nell'immagine a sinistra il fonografo di Edison in visione laterale con il cono di riproduzione e a destra in visione frontale
fonte: gettyimages

Il metodo di incisione dei cilindri brevettato da Edison è detto a modulazione verticale, dove la puntina di registrazione scava i solchi secondo l'asse perpendicolare rispetto al medium; ciò aveva una serie di limitazioni: una di queste risiede nel fatto che se l'onda sonora è troppo ampia, il solco che viene scavato è troppo profondo e la puntina non riesce a leggere correttamente il solco, nei migliori dei casi introducendo distorsione e nel peggiore uscendo dal tracciato.
Con l'introduzione dei primi dischi di gommalacca, il cilindro iniziò a scomparire.
I primi dischi potevano suonare più di tre o quattro minuti mentre i cilindri loro contemporanei potevano suonare solo per due minuti. In un tentativo di recuperare il vantaggio del disco piatto, Edison introdusse il cilindro di Amberol nel 1909, con un tempo di riproduzione di quattro minuti e mezzo, ad una velocità di rotazione di 160 rpm, che furono a loro volta sorpassati dai cilindri di Blue Amberol, i quali avevano una superficie di riproduzione fatta di celluloide che era molto meno fragile. Nonostante questi miglioramenti, durante il primo decennio del 1910, i dischi vinsero in maniera decisiva questa primordiale guerra tra formati; ciò nonostante Edison continuò a produrre cilindri fino al 1929 per una clientela sempre più in calo.
I dischi non sono completamente migliori dei cilindri nel fornire una fedeltà audio maggiore; le registrazioni fatte su cilindro mantengono una velocità lineare costante per l'interezza della registrazione, mentre quelle incise su disco piatto hanno una velocità lineare maggiore nella zona del solco più esterna rispetto a quella del solco più interno. Invece, i vantaggi derivati dal formato vanno ricercati nel processo di manifattura: i dischi potevano essere stampati facilmente, mentre il processo per stampare i cilindri venne scoperto più tardi.
Le prime registrazioni si basavano completamente sull'energia meccanica, dove il suono veniva ripreso da un corno e indirizzato verso un diaframma, il quale avrebbe fatto vibrare la puntina di taglio. La sensibilità e il range di frequenza erano abbastanza scarsi e la risposta in frequenza era molto irregolare, dando a queste registrazioni acustiche una qualità tonale subito riconoscibile e una qualità sonora che spesso rasentava l'inascoltabile.

Politecnico di Torino

Un cantante, per ottenere una registrazione dove si sentisse qualcosa, doveva praticamente collocare la sua faccia dentro il corno di registrazione. Una tecnica utilizzata per diminuire le risonanze che si potevano creare all'interno della camera di registrazione era quella di avvolgere il corno di registrazione con del nastro.
Gli strumenti d'orchestra con uno spettro di frequenze basso, come il violoncello e il contrabbasso, venivano doppiati o rimpiazzati da strumenti più forti come la tuba. I violini standard degli ensemble orchestrali venivano comunemente rimpiazzati dai violini di Stroh, i quali divennero popolari negli studi di registrazione.

Violino di Stroh. Fonte: gettyimages

Anche le batterie, se rinforzate e microfonate in maniera appropriata, potevano essere effettivamente registrate e infatti sono presenti nelle primissime registrazioni jazz e nelle registrazioni di bande militari.

1.2 Il vinile

I dischi registrati con modulazione orizzontale furono sviluppati nel 1887 da Emile Berliner, il quale chiamò il suo sistema col nome di grammofono, distinguendosi dal cilindro di cera di Edison. Berliner utilizzò la modulazione orizzontale, dove l'incisione avviene grazie al movimento della puntina parallelo al piano d'incisione, con un solco a spirale che tendeva verso il centro del disco; questa tecnica eliminava i problemi della modulazione verticale relativi alla profondità dei solchi.
Nel 1901 furono introdotti i dischi da 10 pollici, seguiti nel 1903 dai dischi da 12 pollici.
Durante la prima metà degli anni '20, gli ingegneri alla Western Electric svilupparono la tecnologia per catturare un suono sotto forma di corrente elettrica, utilizzando un microfono amplificato da valvole termoioniche che amplificano il segnale per poi pilotare una puntina di registrazione elettromagnetica. Le innovazioni della Western Electric produssero una registrazione con una risposta in frequenza molto più ampia e omogenea, la quale si tradusse in una sonorità

Politecnico di Torino

drammaticamente più piena, più chiara e che suonava molto più naturale. Suoni distanti o attenuati che prima erano impossibili da registrare adesso potevano essere catturati; ora il volume era solo limitato dalla larghezza dei solchi su disco e dall'amplificazione del dispositivo di riproduzione.
La Victor e la Columbia comprarono il brevetto del sistema elettrico dalla Western Electric e iniziarono a registrare dischi durante la primavera del 1925. I primi dischi registrati furono “Impromptus” di Chopin e i “Litanei” di Schubert.
Nel 1930 il vinile, originariamente conosciuto come vinilite, fu introdotto come materiale di registrazione per la radio broadcast e le pubblicità in radio. Al tempo, praticamente nessun disco per il consumo privato venne fatto con questo materiale, poiché i giradischi casalinghi ai tempi erano dotati di puntine di metallo molto pesanti che ne avrebbero subito danneggiato la superficie, e quindi venne mantenuta la stessa formula delle origini, cioè di utilizzare dischi fatti di gommalacca, i quali però erano molto fragili e avevano come periodo di rotazione i 78 giri al minuto
Nel 1940 la Colombia sviluppò il 33 giri che aveva come diametro 12 pollici e poteva contenere 33 minuti di musica per ogni lato. La sua rivale, la RCA Records invece immise nel mercato i 45 giri da 7 pollici.
Nel 1955 la Philco sviluppò e produsse i primi modelli di giradischi completamente a transistor. Negli anni '60 e '70, si ebbe un'ampia diffusione del giradischi, divisi principalmente in due categorie: quelli portatili a basso costo con bracci di lettura pesanti e rozzi e quelli con meccanismi cambia-dischi automatici, i quali di solito erano posti all'interno di cabine di legno ed a volte dotati di console di comando. Queste console erano spesso equipaggiate con puntine a pick-up di qualità migliore. Le copertine e i supporti stessi divennero una forma d'arte, grazie alla superficie del disco su cui grafiche e testi potevano essere stampati, e i dischi venivano plasmati in forme e colori inusitati oppure stampati ad immagini, creando la corrente dei cosiddetti picture disc.

Picture Disc di Sgt. Pepper dei Beatles

Il giradischi rimase un elemento comune nei sistemi audio casalinghi ben dopo l'introduzione del nastro magnetico e dei primi CD. Nonostante il prezzo di produzione di quest'ultimi scendesse sotto a quello dei dischi, essi rimasero un formato musicale costoso rispetto alle cassette e ai dischi in vinile. Di conseguenza non era difficile trovare il vinile nelle case degli ascoltatori fino all'inizio degli anni '90.

1.3 Nastro Magnetico

L'introduzione della musicassetta, sviluppata dalla Philips nel 1963, cambiò la percezione della musica registrata, permettendo agli ascoltatori di fare le proprie registrazioni di album e live show, e fare mixtape e bootleg. Da qui arrivò un esplosione di musica piratata che le case discografiche tentarono di ridurre con l'iconica campagna “Home Taping is Killing Music”. Le musicassette erano in competizione negli anni '60 e '70 con la cassetta ad 8 tracce, detta “Stereo 8” la quale era diventata lo standard della musica ascoltata in macchina, ma le musicassette le sopravvisse e divenne il primo formato completamente portatile grazie all'invenzione del Walkman della Sony.

Politecnico di Torino

8 Tracce fonte: This Day In Music

1.4 Digitale Fisico

Nel 1974 la Philips ebbe l'idea iniziale per il CD come formato che potesse rimpiazzare i tape e i dischi in vinile, ma fu la Sony che lanciò nel 1976 il primo CD e nel 1982 il primo lettore CD. Con il CD arrivarono anche i CD-ROM, i CD-RW e lo standard del 16 bit e 44,1 kHz; il CD racchiuse il meglio di ogni formato precedente: audio di alta qualità, compatto, portabile, scrivibile e poco costoso. Esso può essere considerato come estremamente importante per lo sviluppo dell'industria musicale, diventando di fatto il formato base per le decadi successive.
Ma d'altro canto il CD fu l'inizio della fine per i formati fisici. Il computer e l'MP3 presero il sopravento sulle abitudini d'ascolto; appena l'accesso alla musica cominciò ad essere associata al personal computer, la maggior parte delle persone non volle più delle copie fisiche della musica, poiché potevano mettere tutto dentro la cartella di un PC.
Con l'arrivo del ventunesimo secolo, il giradischi divenne un prodotto di nicchia, in concomitanza col fatto che i lettori CD, che potevano riprodurre la musica senza scoppiettii e graffi, iniziarono a diventare molto meno costosi e molto più comodi dei riproduttori ad alta fedeltà dei nastri magnetici e dei giradischi. Ciononostante, vi fu un aumento di interesse per la musica su supporto analogico: molte catene di elettronica continuarono a proporre giradischi, spesso in cabinet in cui si trovavano contemporaneamente lettori CD e mangianastri, mentre i negozi specializzati in equipaggiamento per DJ continuarono a vendere giradischi e puntine che non persero mai la loro attrattiva come strumento musicale professionale. Attualmente la maggior parte degli

amplificatori di fascia bassa e media omettono l'input fono; ma d'altra parte sono largamente diffusi i giradischi di fascia bassa con un pre amplificatore integrato. I dischi e i giradischi continuano a essere prodotti e venduti, anche se in quantità minori rispetto ai tempi d'oro del disco fonografico.
La qualità dei riproduttori, dei bracci di lettura e delle puntine ha continuato a migliorare, nonostante la domanda sia diminuita, permettendo ai giradischi di rimanere competitivi nel mercato dell'audio ad alta fedeltà; gli amanti del vinile hanno spesso richiesto e sono spesso coinvolti nella ristrutturazione e nel ritocco dei sistemi vintage.
L'arrivo del Mp3, originariamente sviluppato nei primi anni '80 dal ricercatore Karlheinz Brandenburg nel 1987, ha stravolto definitivamente l'industria discografica. Nel 1999 Napster permise, gratuitamente, il file sharing peer-to-peer di brani musicali anche protetti da copyright, e ciò risultò in un crollo delle vendite dei formati fisici e aprì le porte a servizi come Itunes, che permisero all'utente di cercare, comprare e riprodurre istantaneamente la musica con un paio di click. L'introduzione del Ipod della Apple ha definitivamente sancito l'utilizzo del Mp3 e del download digitale.

1.5 Streaming e Vinyl Revival

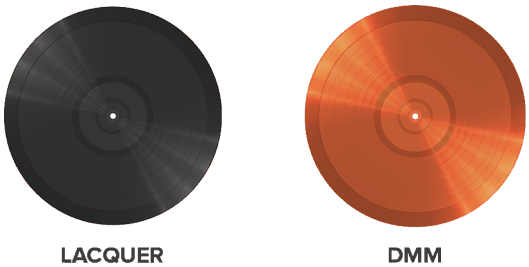
Con l'accesso ad Internet 24/7, alcuni imprenditori hanno visto l'opportunità di ascoltare e scoprire nuova musica senza la necessità di scaricare file e comprare canzoni; in più queste piattaforme puntavano idealmente a rendere la musica digitale un modello di business sostenibile.
La release dell'Iphone e degli smartphone in generale nel 2007, ha aumentato vertiginosamente la popolarità dei servizi di streaming e di radio online come Spotify e Tidal.
La vendita dei vinili nel 2015 ha subito un aumento del 49-50% rispetto all'anno precedente, rimanendo comunque un fenomeno di nicchia rispetto alle vendite digitali e ai CD, i quali però stanno subendo un declino nelle vendite. Nel Marzo 2020, come viene riportato da Billboard, il formato in vinile è stato il salvatore dell'industria dei media musicali durante l'epoca Covid, ed è cresciuto del 11,2 % rispetto allo stesso periodo del 2019, passando da una vendita di 8,3 milioni a 9,2 milioni di copie ed è l'unico formato fisico che sta aumentando le vendite, mentre tutti gli altri formati sono in declino.
Ciò è collegato alla nascita dei servizi di streaming digitale, dove il consumatore medio, secondo una ricerca pubblicata dalla BBC, tende ad utilizzare le piattaforme digitali basate sull'advertisement come SoundCloud e Youtube, per scoprire nuova musica, e poi talvolta comprare il vinile, evitando di acquistare la copia fisica digitale che ha quindi perso la sua attrattiva.

Politecnico di Torino

2 - Analisi dei Supporti Analogici

2.1 Sistema di funzionamento di un giradischi

Una rappresentazione tridimensionale dell'onda sonora viene incisa sul vinile attraverso una testina di taglio la cui traccia sul disco è, in teoria, un diretto corrispettivo del segnale analogico.
In realtà non si incide su vinile direttamente, ma si incide su un disco di acetato da cui, tramite processi galvanici, si ottiene lo stampatore di nickel che stamperà a caldo il vinile. Quando il disco verrà riprodotto le onde sonore saranno dei corrispettivi fisici dell'onda originaria e il volume dipenderà dalla ampiezza del solco, quindi dall'ampiezza del segnale originario. Mentre la puntina di riproduzione segue i solchi del disco, esso muove un magnete inserito in una spirale di rame, il quale movimento genera una corrente elettrica che corrisponde all'onda sonora incisa sul vinile. La corrente elettrica viene inviata ad un pre amplificatore che aumenta l'ampiezza della tensione e la trasmette agli altoparlanti che la traducono in vibrazione meccanica. I sostenitori del vinile indicano che questa riproduzione del suono attraverso una catena elettroacustica analogica, dove il segnale rimane una grandezza continua e non viene campionata e quantizzata come l'audio digitale, sia più fedele e migliore di una riproduzione in formato digitale. Bisogna aggiungere però che al giorno d'oggi molto spesso vengono incisi vinili da master digitali, e quindi in questo caso non si può parlare di un segnale completamente analogico e la diatriba tra la presunta superiorità di un formato rispetto all'altro diventa futile. Una delle tecniche ritornate in auge con la rinascita del vinile è il DMM (Direct Metal Mastering), cioè una pratica di fabbricazione del vinile dove si salta il processo di taglio della lacca in acetato e della successiva argentatura e si incide il disco direttamente su un disco di rame, che sarà utilizzato per produrre tutti gli stampatori successivi. I modelli più recenti di torni di incisione utilizzano, per il pilotaggio della testina, una linea di ritardo digitale quindi, i vinili di produzione attuale spesso non sono prodotti con una catena completamente analogica ma si passa per un processo A/D – D/A.



Confronto tra la lacca usata nel processo di fabbricazione del vinile standard e il disco in rame usato nella tecnica DMM

Politecnico di Torino

Per riprodurre un disco lo si pone su un piatto rotante fatto girare a velocità costante e predefinita. Inizialmente, al tempo dei cilindri fonografici e dei primi grammofoni, il piatto era messo in movimento da una molla caricata con una manovella, che poi venne sostituita da un motore elettrico. Affinché il fonorivelatore possa avere qualche probabilità di estrarre una quantità ragionevole delle informazioni contenute nel solco, è necessario che la relazione geometrica tra gli assi su cui giacciono il perno del piatto e quello del braccio sia della massima stabilità. Qualora esista un grado di libertà anche minimo tra questi elementi, la capacità di recupero delle informazioni dal solco ne verrà irrimediabilmente compromessa.

2.1.1 Trasmissione meccanica della rotazione

Nel tempo si ebbero diversi tipi di trasmissione del moto dal motore al piatto: dapprima vennero utilizzate delle pulegge, rimpiazzate poi da cinghie in gomma che abbracciano il piatto principale e lo fanno girare alla stessa maniera di una puleggia: questa tecnica viene anche detta *trazione a cinghia* o *belt drive;* si giunse poi a far girare il piatto collegandolo direttamente al motore e al suo albero attraverso la *trazione diretta,* dove il motore è posto direttamente sotto il piatto ed ha la stessa velocità di rotazione; quest'ultimo è meno influenzabile da forze esterne, come può essere l'attrito della puntina sul disco, ed è più veloce nel raggiungimento della velocità nominale, mantenendo stabile la velocità di rotazione, di conseguenza è il più utilizzato dai DJ. Alcuni sofisticati modelli prevedono anche che il motore e il braccio di lettura siano separati fisicamente dal piatto rotante; queste modifiche concettuali nella costruzione dei giradischi hanno favorito una qualità di riproduzione man mano più accurata e priva di disturbi derivanti dall'apparecchio stesso, riducendo praticamente a zero l'apporto di rumore da parte del motore; raffinati accorgimenti tecnici sono stati introdotti anche per minimizzare il rumore generato dal perno in rotazione solidale col piatto, annullando perlomeno l'attrito verticale, adottando soluzioni come la sospensione magnetica od ad aria in pressione. Attualmente, per garantire la massima regolarità nella rotazione del piatto, il motore è azionato da un circuito elettronico pilotato da un oscillatore al quarzo.

2.1.2 Lettura

La riproduzione vera e propria si ottiene appoggiando sul disco inciso la puntina di lettura, che ha la funzione di pick-up; essa è sostenuta da un braccio munito di un apposito contrappeso con lo scopo di ottenere una pressione della testina sul disco costante e predefinita.
Il braccio è collegato ad uno snodo posto sulla base del giradischi in modo da fare seguire alla testina il percorso determinato dal solco del disco.

Politecnico di Torino

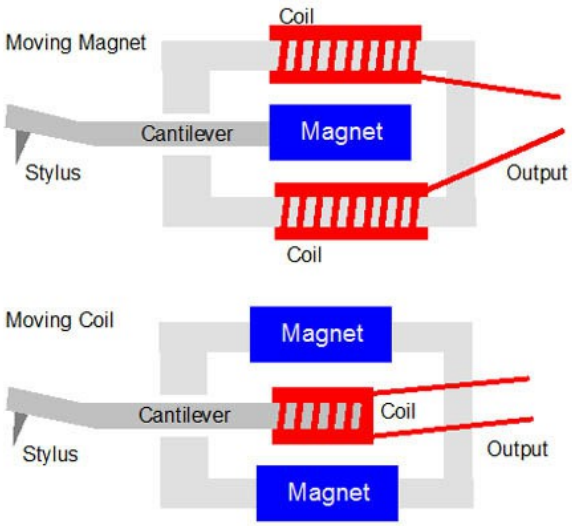
A partire dagli anni sessanta si sono affermati giradischi muniti di *braccio tangenziale*, ossia bracci che supportano l'andamento della testina tangenzialmente al diametro del disco, evitando così differenti modalità di lettura derivanti dalla curvatura del percorso, detto *errore tangenziale*, quasi inevitabile con bracci a snodo. Nei bracci a snodo all'errore tangenziale si somma l'errore di skating: il braccio di lettura subisce la forza di attrito radente che si sviluppa tra la puntina e i solchi del disco, provocando una forza che ha una componente perpendicolare rispetto ad essa, e tende a spostarla verso il centro del giradischi. Per controbilanciarlo viene installato su alcuni turntable un dispositivo meccanico (di solito un contrappeso) detto antiskating, che applica un leggera forza al braccio di lettura per controbilanciare la spinta verso il centro del disco.
I giradischi equipaggiati con braccio tangenziale hanno sempre costo elevato, per l'alto grado di ingegnerizzazione e precisione per la loro realizzazione ma sono esenti da lo skating.

2.1.3 Testina

Quando l'apparato puntina e bobina è posato sul disco, il profilo irregolare del solco provoca la vibrazione della puntina; tale vibrazione ai tempi del fonografo era trasmessa direttamente ad una membrana che riproduceva il suono in maniera esclusivamente meccanica; in seguito le vibrazioni sono state impiegate per generare deboli segnali elettrici da una bobina o un cristallo piezoelettrico inclusi nel corpo della testina: quelle di buona qualità sono costituite da una leva chiamata cantilever (un sottilissimo tubetto in metallo leggero), fulcrata al centro con silicone e pertanto libera di oscillare in ogni direzione; ad un suo estremo è fissata la puntina, mentre all'altro estremo può essere fissato un magnete o una bobina; nel primo caso, il fonorivelatore prende il nome di MM (moving magnet, cioè magnete mobile); nel secondo caso prende il nome di MC (moving coil, bobina mobile).
Le testine MM sono basate su un magnete permanente fissato all'estremità del cantilever, che viene posizionato in mezzo a due bobine che costituiscono i canali dello audio stereofonico.
Il movimento del magnete tra le bobine, dovuto alla vibrazione della puntina che segue il solco del vinile, crea una fluttuazione nel campo elettromagnetico, che a sua volta genera per induzione una corrente elettrica di pochi millivolt nelle bobine.
Le testine MC si basano sempre sull'induzione elettromagnetica ma con le bobine e il magnete invertiti. In questo caso, un magnete è attaccato al corpo della testina, mentre due bobine sono fissate al cantilever. Il vantaggio di questo sistema è la leggerezza del sistema cantilever-bobine che si muove in sospensione ed è molto più leggero rispetto al sistema cantilever-magnete delle testine MM; ciò permette alla puntina di reagire più velocemente al tracciato del solco, risultando in un suono più dettagliato e in più si consuma meno la superficie del disco perchè spesso appoggia con un minor peso.

Politecnico di Torino

Le bobine in questo caso sono più piccole rispetto al sistema a magnete mobile, generando una tensione di uscita ridotta ma che produce un segnale che risponde meglio ai transienti ed ha una migliore risposta in frequenza. La necessità di un design più raffinato e quindi più costoso ne fa levitare il prezzo rispetto alle testine a magnete mobile.
Essendo il sistema a bobina mobile più preciso ma circa 10 volte meno efficiente, poiché genera una tensione 10 volte inferiore dovuta alla dimensione minore della bobina, i deboli segnali rivelati devono essere prima pre amplificati da un apposito circuito chiamato per consuetudine "pre-MC", che, in alcuni casi, è fornito come elemento opzionale dal costruttore dell'amplificatore, dopodiché i segnali sono utilizzabili dal circuito di pre amplificazione vera e propria; da qui sono inviati all'amplificatore finale, il quale trasferisce il segnale agli altoparlanti montati nei diffusori, trasformandolo finalmente in suono.



Comparazione tra sistema MM e MC

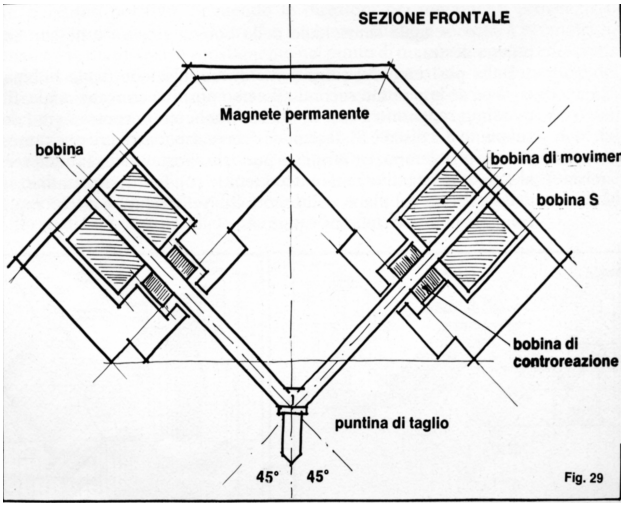
La velocità angolare fissa del piatto (33, 45 o 78 giri) implica un degrado costante nella fedeltà di riproduzione: nelle tracce esterne del disco si ha una maggiore fedeltà di riproduzione rispetto alle tracce interne. Poiché la distanza percorsa dalla puntina in un intervallo di tempo è molto superiore sulle prime tracce che nelle successive, i dati contenuti sulle tracce più esterne hanno più spazio a disposizione di quelli contenuti nelle tracce più interne; in altre parole, la densità di informazione per unità di spazio aumenta dall'esterno verso il centro del disco in modo inversamente proporzionale al raggio; in ambito analogico la densità di informazione penalizza la qualità anche in altre applicazioni (film fotografici, nastri magnetici, frequenze modulate sulle portanti delle onde elettromagnetiche).

Politecnico di Torino

2.2 Processo di fabbricazione del vinile

2.2.1 Incisione

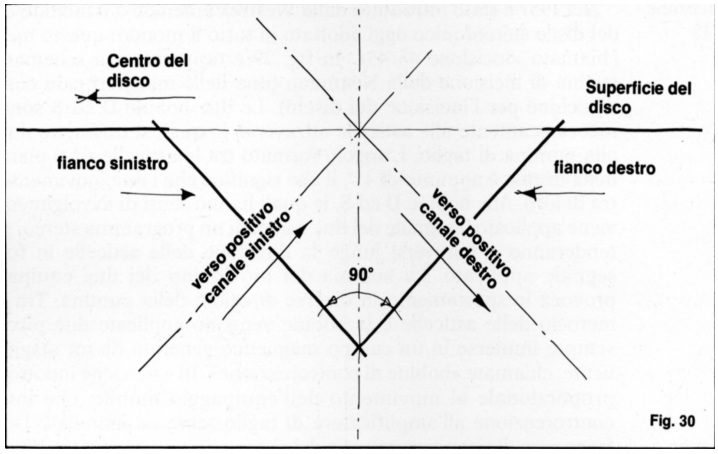
L'informazione audio analogica viene incisa sull'acetato attraverso una modulazione della puntina del tornio. La modulazione può essere orizzontale, verticale oppure a 45 gradi. I primi cilindri di Edison utilizzavano una modulazione di tipo verticale, ma con l'avvento del disco piatto nell'industria discografica la modulazione che prese il sopravvento fu quella di tipo orizzontale. Sia quella orizzontale che quella verticale creavano un canale mono. Poi con l'avvento dello stereo si è pensato di utilizzare la modulazione orizzontale per il canale destro e la modulazione verticale per il canale sinistro. Ma la modulazione verticale presentava dei problemi perché introduceva distorsione e generava un volume più basso dovuto all'incapacità della puntina di riproduzione di seguire un solco più profondo e aveva la tendenza a scivolare via dal solco se l'ampiezza diventava eccessiva. Quindi si è optato per un compromesso: la modulazione a 45 gradi.
Essa venne introdotta nel 1957 dalla Westrex ed è diventato il metodo standard di incisione del disco stereofonico adottato in tutto il mondo.
In figura è rappresentata schematizzata una testina di incisione della Neumann, una delle maggiori case costruttrici di macchine per l'incisione dei dischi.



Testina d'incisione della Neumann fonte: il disco microsolco

Politecnico di Torino

La bobina destra e quella sinistra sono collegate meccanicamente a delle asticelle attraverso le quali si trasmette il movimento alla puntina di taglio. L'angolo formato tra le asticelle ed il piano verticale della testina è appunto di 45 gradi, il che significa che i due movimenti sono a 90 gradi tra di loro. Alle bobine destra e di sinistra, le quali hanno sensi di avvolgimento opposti, viene applicato il segnale dei due canali di un programma stereo, per cui esse tenderanno a muoversi lungo la direzione delle asticelle in funzione del segnale applicato. La somma dei movimenti dei due equipaggi mobili provoca lo spostamento in diverse direzioni della puntina. Tra il punto di congiunzione tra le asticelle e le bobine vengono applicate due altre piccole bobine, sempre immerse in un campo magnetico generato da un magnete permanente, che vengono dette bobine di controreazione. In esse viene indotto un segnale proporzionale al movimento dell'apparato mobile, che inviato come controreazione all'amplificatore di taglio serve ad annullare l'effetto della frequenza di risonanza propria del sistema in movimento ed a ridurre la distorsione.
Nella figura successiva viene rappresentata una sezione trasversale di un solco in cui si possono vederne i fianchi, gli assi su cui avviene il movimento ed i versi definiti positivi dalla normativa internazionale.



Rappresentazione grafica del solco di un vinile fonte: il disco microsolco

Questo compromesso indica che nessuno dei due canali suona bene come un canale mono, ma almeno suonano uguali tra di loro.
Le informazioni dei due canali saranno quindi contenute in parte nella modulazione verticale e in parte in quella orizzontale. In linea di massima maggiormente in quella orizzontale, soprattutto nei dischi stereo compatibili dove si cerca di ridurre al massimo i segnali verticali provocati dai segnali in opposizione di fase. Si fissa un valore massimo di profondità p_{max} sotto al quale non si può scendere poiché la puntina di lettura uscirebbe dal solco. La profondità massima dipende dalle

caratteristiche della macchina di incisione e non ha un valore fisso. É chiaro come elevati livelli sonori di registrazione provochino un solco molto ampio e ciò provoca delle problematiche riguardanti la quantità di dinamica che può essere contenuta all'interno di un disco di vinile.

2.2.2 Equalizzazione

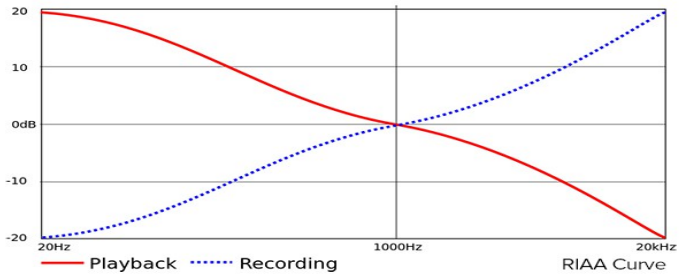
Prima di essere inciso il segnale segue una procedura di mastering e ottimizzazione per il supporto vinilico; si fa passare il segnale attraverso una serie di compressori, limiter ed equalizzatori dotati di filtri passa alto e passa basso, che vengono utilizzati per attenuare le altissime e le bassissime che rovinerebbero l'incisione; il segnale quindi viene limitato in banda e non viene incisa una rappresentazione completa dell'onda sonora originaria.

Esempio di attrezzatura destinata al mastering. Fonte: Sage Music

Dopo queste manipolazioni viene applicata l'equalizzazione RIAA.
Infatti i segnali che devono essere incisi su disco, una volta provenienti dal nastro magnetico ma al giorno d'oggi completamente digitali, vengono equalizzati secondo questa determinata caratteristica. Questa equalizzazione ha il duplice scopo di economizzare la superficie del disco stesso e di ridurre il rumore di fondo presenti in fase di riascolto. Se noi riproducessimo il segnale inciso sul disco con un sistema bobina-magnete, che produce un segnale elettrico proporzionale alla velocità di incisione e con tensione di uscita costante per tutte le frequenze, dovremmo incidere un segnale a velocità costante e quindi con ampiezza inversamente proporzionale alla frequenza; in altre parole le basse frequenze occuperebbero uno spazio immenso e le alte uno spazio di taglio talmente piccolo da essere inferiore alla ruvidità media del disco. Questa equalizzazione, introdotta negli States in comune accordo tra le case discografiche a fine anni '50

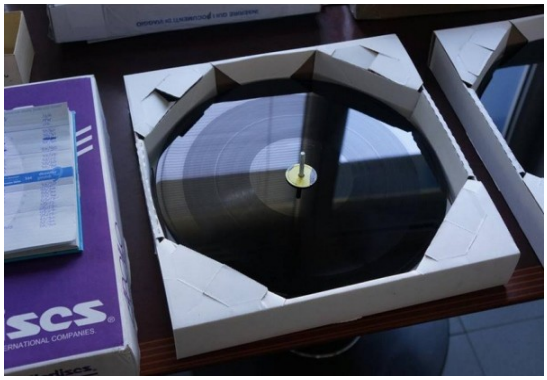
Politecnico di Torino

con l'avvento dei 33 giri, viene detta RIAA (Record Industries Association of America), che permette di ottimizzare lo spazio sul disco ed è usato come standard di mastering sul vinile. La sua trasformata inversa verrà poi applicata alla parte di trasduzione elettrica della puntina per ricostruire il segnale originale.



Curva RIAA

Il segnale audio viene quindi inciso con un tornio su una lacca che sarà la madre di tutti i dischi che verranno riprodotti. La lacca è costituita da un disco con un'anima di alluminio su cui è stato versato il composto di nitrocellulosa, cotto ad altissime temperature per eliminare ogni tipo di grumo, la cui composizione precisa è tenuta celata dai produttori di lacche e sono frutto di ricerche che i fabbricanti fanno di continuo per migliorarne le caratteristiche. Infatti il materiale deve rispondere ad un serie di requisiti molto spesso in contrasto tra di loro: infatti deve essere tenero per permettere allo stilo di tagliare con facilità, ma al tempo stesso deve essere rigido per impedire la deformazione del disco. Inoltre deve essere antipolvere, antistatico e resistere alle radiazioni ultraviolette. Dopo aver inciso una facciata del disco si utilizza un microscopio per un controllo visivo dei livelli più elevati del solco.

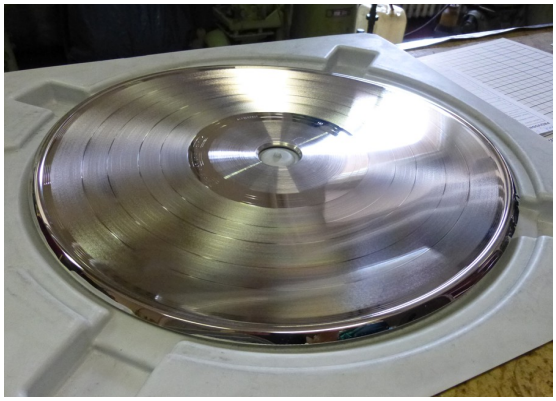


Acetato in attesa di essere inciso fonte: gzvinyl

Politecnico di Torino

2.2.3 Elettro placcaggio

L'elettroplaccaggio è la prima parte del processo di creazione del vinile vero e proprio. Dalla lacca si ricavano gli stampatori che andranno montati sulle presse dei dischi; questi stampatori si ottengono per via elettrolitica attraverso una serie di passaggi successivi perfettamente coordinati tra di loro.
Il problema è ottenere uno stampo negativo, cioè con i solchi di fuori, che possa essere montato sul piatto della pressa e che resista alla enorme pressione di pressaggio, oltre che al continuo passaggio della temperatura caldo-freddo che compone il normale ciclo di stampa di un disco. Il materiale che si presta al processo è il nichel, il quale ha la proprietà di resistere a queste sollecitazioni e soprattutto di adattarsi al processo elettrolitico. Si inizia la lavorazione con un lavaggio della lacca con acqua demineralizzata per eliminare ogni traccia di sedimenti e impurità e poi si passa ad un bagno nell' idrossido di potassio per togliere i residui di impronte digitali e sporco. La lacca viene fatta passare attraverso altri bagni in apposite soluzioni, per prepararla al deposito d'argento: in questo caso la tempistica di questi ultimi bagni deve essere precisa, poiché se tenuto troppo a lungo si potrebbe aumentare il rumore di fondo o viceversa se tenuto troppo poco il disco potrebbe creparsi. E' molto importante il costante lavaggio in acqua demineralizzata per eliminare ogni tipo di residuo che potrebbe rimanere nella traccia della lacca, che si tradurrebbe in vari tipi di rumore e persisterebbe in tutte le copie successive. Infine viene deposto sulla lacca una sottile pellicola d'argento (< 1 micron) con una pistola a spruzzo o in una vasca contenente argento in sospensione, in modo tale che l'argento faccia da conduttore elettrico per la fase successiva. Senza le fasi di preparazione e lavaggio lo strato di argento non si attaccherebbe in maniera uniforme.



Disco dopo il processo di silvering fonte: gzvinyl

Politecnico di Torino

Dopo un' ulteriore lavaggio, si mette poi la lacca argentata in un bagno elettrolitico composto da acqua e nichel, con intensità di corrente di 30 A/cm^2; il nichel presente si attacca all'argento e forma una pellicola metallica che a poco a poco diventa sempre più spessa fino ad arrivare al valore ottimale di circa 0,2 millimetri ed un peso di 150 grammi.
Finita l'operazione di deposito, dalla durata di circa un'ora e mezza, il meccanismo che comanda il bagno stacca la corrente e alza il disco fuori dal bagno in modo che l'operatore possa vederlo ed accorgersi che il bagno è così ultimato.

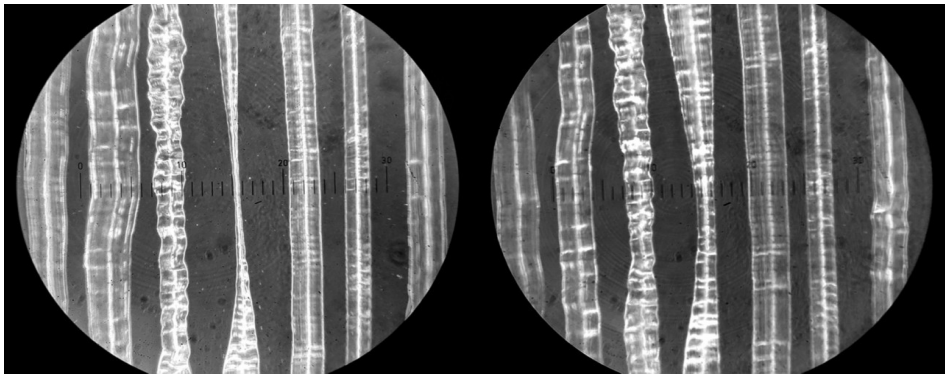
Vasche per il bagno galvanico fonte: gzvinyl

Si stacca il disco dal supporto e poi, aiutandosi con un coltellino, si stacca la pellicola di nichel-argento dalla lacca. Abbiamo ottenuto quindi un negativo che però non è molto resistente e, se usato per creare delle copie del disco, dopo un po' di operazioni di stampa inizia a perdere qualità quindi inizia a deformarsi oppure i solchi perdono definizioni alle alte frequenze. Di conseguenza viene riprodotto il processo elettrolitico visto poc'anzi, a temperature più alte e con tempi minori, per formare una copia positiva che viene detta madre. La madre viene ascoltata e catalogata poiché da essa si ricaveranno tutte le altre copie degli stampatori anche in tempi successivi, come ad esempio in caso di ristampe. L'ascolto della madre è un ascolto di qualità, per vedere che non vi siano delle imperfezioni di natura varia alla fonte del disco. La madre va ascoltata per intero, da personale specializzato appositamente istruito per questo scopo che deve essere sensibile ai tic più impercettibili. Se tutto va bene e non vi sono rumori o altri disturbi, si passa alla fase conclusiva del reparto galvanica, cioè il bagno finale per ottenere gli stampatori.

Politecnico di Torino

2.2.4 Centratura e pressaggio

Prima di montare gli stampatori sulle presse, deve essere fatta un'importantissima operazione: la centratura. Bisogna trovare perfettamente il centro dei solchi e far sì che il foro centrale sia perfettamente il centro del disco. In caso contrario, invece di essere circolare, il solco è praticamente ellittico, con conseguente aumento e diminuzione della velocità sotto la puntina una volta messo sul piatto del giradischi. La centratura si fa appoggiando il disco su di un piatto simile a quello del giradischi privo di perno centrale; al posto della puntina, su una specie di braccio, c'è un microscopio che inquadra i solchi del disco: facendo girare il disco si vede se i solchi si spostano dalla tacca centrale del microscopio oppure rimangono sempre allineati a questa. Lo spostamento deve essere ridotto a zero facendo scivolare con dei piccolissimi movimenti il disco sul piatto. Quando, girando il piatto, tutti i solchi appaiono fermi, o meglio come una spirale convergente al centro del disco, allora a questo centro corrisponde il centro del piatto. C'è una piccola trancia del diametro del perno del giradischi che taglia il centro dello stampatore e crea un foro dell'esatta dimensione e perfettamente centrato rispetto ai solchi.

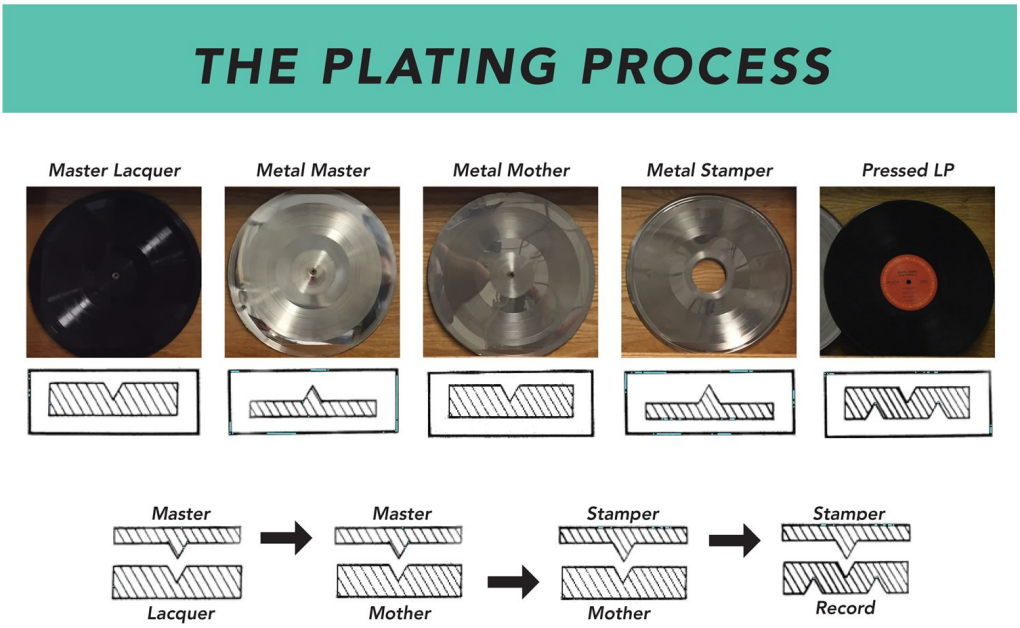


Solco vinilico al microscopio fonte: gzvinyl

Dopo questa fase lo stampatore viene alloggiato nello stampo, che è una piastra metallica perfettamente liscia con una cornice rotonda per trattenerlo. Lo stampo ha, nella parte posteriore, una serie di alettature per favorire lo scambio termico dello stampo stesso ed accelerare il processo: nella parte posteriore infatti scorre sia il vapore per il riscaldamento che l'acqua per il raffreddamento. Nella maggior parte dei casi gli stampi sono in acciaio, più raramente in rame al berillio. Lo stampo viene poi montato sulla pressa.

Politecnico di Torino

Viene iniettata la palla di pasta mentre la pressa è aperta, poi la pressa si chiude ed inizia il pressaggio vero e proprio: dopo qualche secondo il pressaggio è ultimato e la pressa si apre. Al posto della palla di pasta c'è il disco. Attraverso una serie di binari, passa prima per l'etichettatrice e poi per la lama rifilatrice che taglia l'eccedenza di materiale attaccato al disco stesso. Dopo un giro completo il bordo eccedente è distaccato e, dopo essere imbustato, va a finire in una scatola raccoglitrice. Ogni 100-150 dischi prodotti viene effettuato un controllo di qualità per vedere che non vi siano difetti nella fabbricazione del disco. Il difetto più frequente è il classico tic provocato da un piccolissimo grumo di pasta non ben amalgamato. Dopo il controllo, i dischi passano all'imbustamento (talvolta manuale) e alla cellofanatura. Infine, ultima fase, i dischi vengono confezionati in scatole e inviati in magazzino.



Schema riassuntivo della produzione di un disco in vinile raffigurante il processo di copia negativa e copia positiva
fonte: gzvinyl

I dischi per la maggior parte sono composti da PVC (cloruro di polivinile) e di PVA (acetato di polivinile) e di una serie di componenti che portano la plastica ad avere le caratteristiche ottimali per lo stampaggio. Ogni fabbrica ha la sua pasta particolare, frutto di una serie di ricerche e di esperienze fatte nel corso degli anni e delle quali è molto gelosa. Non si tratta di veri e propri segreti, ma, cambiando anche di poco una delle percentuali e abbinando a ciò una modifica della

velocità del ciclo di stampaggio, si può ridurre di molto il rumore di fondo e la distorsione del disco.
Alla resina base, che è incolore, viene aggiunta un copolimero più duro, che ha lo scopo di creare contrasto nella fase di pressaggio. Vengono anche aggiunti degli stabilizzanti, che hanno lo scopo di tenere unita la resina che si sgretolerebbe con il calore della pressa, dei coloranti, che ultimamente vengono utilizzati per creare delle versioni speciali e limitate delle uscite discografiche e infine dei lubrificanti per far scorrere meglio la pasta anche nei solchi più microscopici.
Si può intuire quindi che il processo di fabbricazione di un disco in vinile, oltre ad essere un processo complesso in cui è richiesta grande precisione ed esperienza per ottenere un risultato almeno sufficiente, non è una copia perfetta del segnale audio originale e anzi, ad ogni passaggio nel processo elettrolitico e copia dello stampatore, si perde in qualità. Inoltre le lavorazioni sono in buona parte manuali e quindi risentono della capacità dell'operatore.

Politecnico di Torino

3 - Analisi del Supporto Digitale

3.1 Introduzione

I campioni del segnale audio digitale possono essere registrati su una moltitudine di supporti, con caratteristiche fisiche e logiche diverse. Da un punto di vista fisico i supporti possono essere magnetici ottici oppure elettronici; da un punto di vista logico, ciascun supporto è caratterizzato da uno standard proprio per la struttura dei dati, similmente al formato dei file.
Due sono le caratteristiche più rilevanti dal punto di vista dell'audio digitale: la capacità di contenere i dati e la velocità di scrittura e di lettura di essi, in modo da garantire una corretta registrazione e riproduzione. Si consideri un frammento audio stereo di 1 minuto, con un sample rate di 44,1 kHz e una quantizzazione a 16 bit, cioè a 2 byte. Il segnale che presenta tali parametri è quello presente su un CD.
Poiché il segnale è stereo, si hanno due canali; i campioni totali in un secondo sono quindi 88.200 (44.100 * 2). L'espressione aritmetica totale è:

44.100 campioni/s * 2 canali * 2 byte/campione * 60 s = 10 MB

Poiché i campioni totali di 1 secondo sono 88.200 e ciascuno occupa 2 byte, i bit totali che occorre inviare ai convertitori in 1 secondo sono 176.400. Per cui, per ascoltare della musica con i parametri indicati precedentemente, occorrono almeno 10 MB per minuto di musica e una velocità di lettura dal supporto di 176 Kb/s. Ovviamente, se i parametri hanno dei valori inferiori, le richieste sono inferiori. Con un segnale audio a 22.050 campioni/s, con 8 bit di quantizzazione, in mono, si scende a 1,3 Megabyte/minuto con una velocità trasferimento dati di 22 Kb/s, riducendo le necessità di un fattore 8 rispetto all'esempio precedente. Con questi parametri, non tutti i supporti sono idonei alla memorizzazione dell'audio digitale. Ad esempio il vecchio floppy disk, con la sua capacità di 1,44 MB, non può contenere che brevissimi brani audio e di qualità scadente. Ora verranno analizzati i principali supporti adatti a registrare l'audio digitale.

Politecnico di Torino

3.2 CD

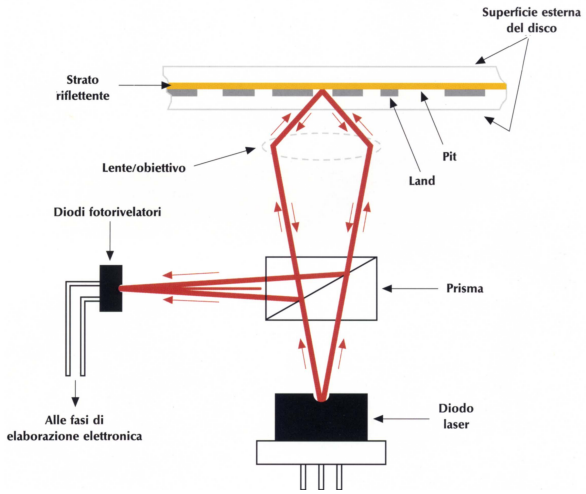
3.2.1 Caratteristiche

Il supporto nato con la diffusione dell'audio digitale è il compact disc (CD), che si è evoluto in una serie di standard che hanno consentito prima la memorizzazione dell'audio e poi la memorizzazione contemporanea di audio e dati di diversa natura.
Nato all'inizio degli anni '80 da una collaborazione tra Sony e Philips, è il supporto tipico dell'audio digitale e ha eliminato, dal mercato di massa, quasi del tutto l'audio analogico, i cui supporti di largo consumo tipici erano il disco in vinile e la Compact Cassette.
Il CD è un disco in policarbonato di 12 cm di diametro, con un foro centrale di 1,5 cm.
La parte dì più interna del disco, quella intorno al foro, e la parte più esterna, quella vicino al bordo non contengono dati: essi infatti partono da 2,5 cm e finiscono a 11,6 cm rispetto al centro e sono scritti dal centro verso l'esterno, seguendo un percorso a spirale. Esternamente all'area dati vi sono due aree di informazione aggiuntive: nei dintorni del centro c'è l'area di entrata detta lead-in, e verso il bordo c'è l'area di uscita che viene detta lead-out. L'entrata e l'uscita si riferiscono alla direzione in cui il disco viene letto, in questo caso dal centro verso l'esterno. La memorizzazione dei dati si basa sulle proprietà di riflessione della luce sulla superficie metallizzata del disco.
Per poter rappresentare gli 0 e gli 1 occorre modificare delle zone della superficie in modo che la riflessione in tali zone sia differente dal resto della superficie. Queste zone assomigliano a degli avvallamenti, detti pit, e vengono prodotti in due diverse maniere in base ai quantitativi che devono essere prodotti. La prima modalità è detta replicazione, ed è tipica dei CD di produzione industriale dove si utilizza un master di vetro ricoperto da un layer fotosensibile, che viene impressionato tramite un laser, montato su una meccanica di altissima precisione. Dopo lo sviluppo fotografico del layer si procede con un processo galvanico per ottenere lo stampatore di nickel/cromo che, inserito in una pressa per il policarbonato produrrà le copie successive (in maniera simile alla fabbricazione del vinile). Per piccole tirature si procede invece con la duplicazione: si utilizzano dei CD stampati come sopra, che contengono però solo i dati necessari a guidare il pickup ottico (CDR), ed hanno in più uno strato di un materiale che ha la proprietà di cambiare indice di rifrazione quando scaldato oltre una certa temperatura. Le informazioni vengono scritte in questo layer per mezzo di un intenso raggio laser che provoca il riscaldamento dei soli punti corrispondenti ai pit.
I pit sono disposti sulle tracce a spirale dal centro verso l'esterno. La superficie tra i pit è detta land. I pit sono larghi circa 0,5 µm e lunghi da 0,83 µm a 3,56 µm, con una profondità tra 0,1 e 0,2 µm e con una distanza tra le due tracce di 1,6 µm. Per farsene un'idea rispetto ad altri supporti si può usare la misura tpi (tracks per inch – tracce per pollice): un CD ha una densità di 16.000 tpi,

contro i 96 del floppy disk e i 400 in media per l'hard disk; la lunghezza della spirale dei pit si allunga per oltre 5 Km. Il policarbonato su cui sono impressi i pit è ricoperto da un sottile strato di alluminio e al di sopra di esso viene posta una pellicola protettiva, su cui viene poi applicata l'etichetta. La lettura dei pit avviene in realtà dal lato opposto al quale sono stati impressi e quindi si rilevano delle gobbe più che degli avvallamenti e, al contrario di quanto si ritiene comunemente, il lato più delicato del CD risulta la parte dell'etichetta, in quanto lo strato protettivo è di soli 2 micron e, se si crea una crepa, l'aria può ossidare l'alluminio e rendere il disco inutilizzabile; invece il lato della lettura è molto più spesso (1,2 millimetri), e quindi più resistente.

3.2.2 Lettura Laser

In lettura un raggio laser, focalizzato mediante una lente, percorre la traccia determinata dai pit e dalle land: la luce riflessa passa attraverso un prisma che la deflette verso un fotosensore, generando degli impulsi di ampiezza proporzionale. È importante dire che gli 1 vengono segnalati dalla lettura del bordo di un pit, cioè quando viene rivelata una differenza di riflessione. La parte più delicata del meccanismo è mantenere il disco nel fuoco del raggio laser e su un piano che permetta al raggio di seguire la traccia che sta leggendo. Il metodo più comune è l'utilizzo di tre raggi, di cui due prodotti per diffrazione dall'unico generato: uno centrale che segue la traccia principale e due laterali che percorrono lo spazio tra le tracce. Tale spazio infatti ha coefficienti di riflessione costanti: se si verifica qualche alterazione, un meccanismo riposiziona la lente. Per i movimenti verticali del disco, si verifica che il fascio laser riflesso abbia una forma circolare; se questa diventa ellittica, si opera di nuovo sulla lente per riportare il fascio a una sezione circolare.



Sistema ottico di lettura di un CD

Politecnico
di Torino

3.2.3 CLV e CAV

La rotazione dei dischi può avvenire in due modalità: a velocità lineare costante, detta CLV (Constant Linear Velocity), o a velocità angolare costante detta CAV (Constant Angular Velocity). I lettori CLV sono dedicati al CD-Audio e al CD-ROM di prima generazione, che assicurano un tasso di trasferimento dati di 150 kB/s o 1x per i lettori CD-ROM. Per mantenere costante la velocità lineare, e quindi la velocità di lettura dei dati, il disco rallenta gradualmente la rotazione man mano che il pick-up ottico si muove dall'interno verso l'esterno. I dati rilevati vengono inviati ad un buffer che li memorizza e li restituisce nello stesso ordine di arrivo, spaziandoli a intervalli di tempo precisi secondo il clock interno. Imprecisioni della velocità del disco non influenzano così l'afflusso dei campioni al convertitore DAC per la riproduzione, il cosiddetto jitter, che verrà analizzato successivamente. Un controller cerca di mantenere le informazioni contenute nel buffer sotto la metà della sua capacità, rallentando la lettura quando i dati si accumulano oltre questa soglia.
I lettori CAV fecero la loro comparsa a metà degli anni '90, quando la Pioneer lanciò il primo lettore che operava non solo col sistema CLV ma anche con il CAV, il quale legge i dati ad un tasso variabile perché il disco ruota a velocità costante, con un meccanismo simile alla lettura di un hard disk. Con il sistema CLV si ha una riduzione dei tempi di accesso a parti diverse del disco, effetto che non è particolarmente importante per un CD-Audio, in quanto è inutile velocizzare la lettura oltre il bit-rate necessario per la riproduzione. Per il CD-ROM, che può contenere dati di diversa natura, tra cui dati in input a programmi per l'elaborazione, è invece importante ridurre il più possibile il tempo di accesso ai dati e la velocità di lettura. Con l'avvento di elaboratori del segnale molto veloci, i processori DSP, la tecnica CAV è diventata la più diffusa, aumentando quindi la velocità di trasferimento dati. Ad esempio, la velocità 48x corrisponde a 7,2 MB/s (= 150 kB/s x 48), con velocità angolare di 12.000 rpm (round per minute), quasi comparabile alla prestazione di un hard disk.

3.2.4 Codice di rilevamento

Il formato di scrittura dei dati sul CD segue una struttura molto articolata. Oltre ai campioni audio così come arrivano dalla conversione analogico-digitale, il formato dei dati sul CD include i codici di rilevamento e correzione dell'errore. Inoltre i dati vengono rimescolati e sottoposti alla codifica EFM.
L'idea è di aggiungere ai bit di informazione dei bit di controllo che sono utili per rilevare se sono occorsi degli errori, i cosiddetti bit “flippati”, ed eventualmente correggerli.
Il rimescolamento dei dati segue una tecnica di interfogliamento, per cui i campioni in un gruppo fissato vengono disposti in modo differente dall'ordine cronologico originale. In tal modo, in caso

di errori di lettura, che sono probabilmente dei burst di errori sequenziali, non verranno a mancare una serie di campioni vicini e sarà possibile ricostruire correttamente i dati originali.
Le tecniche di prevenzione degli errori nella lettura partono dal segnale RF che esce dalla testina laser di lettura, la quale ha delle oscillazioni che sono prossime a quelle di un' onda sinusoidale. Per semplificarne la lettura, viene applicata una codifica preliminare nella realizzazione dei pit. Infatti la lunghezza dei pit non può essere superiore alle 11 unità e inferiore alle 3 unità, quindi non vi possono essere più di due 1 consecutivi e il numero di zeri consecutivi è compreso tra 2 e 10. Ciò permette una lettura semplificata nel caso di errori o di danni sulla superficie del disco: il segnale RF viene filtrato da circuiti elettronici che eliminano le frequenze che non possono essere state scritte sul disco, in questo caso imperfezioni e polvere, e viene trasformato in un'onda quadra in cui i fronti di salita e discesa rappresentano i bit a 1, ovvero il cosiddetto segnale NRZ1. Da questo si ottiene il segnale NRZ che è una vera onda quadra costituita da valori di 0 e 1, da cui vengono prelevati i sincronismi che controllano la velocità e i bit validi che sono stati trasmessi.

3.2.5 Codifica CIRC

Le tecniche di controllo degli errori e rimescolamento nella scrittura su disco sono denominate Codifiche di Reed-Solomon di Interfogliamento Incrociato (Cross Interleave Reed-Solomon Code detta anche CIRC).
Il processo totale di memorizzazione dei dati sul CD segue i seguenti passi. Si parte dalla forma d'onda digitalizzata in parole binarie da 16 bit per sample, i quali vengono interfogliati tra il canale sinistro e il canale destro. Un gruppo di due campioni consecutivi potrebbe quindi essere:

$$L_1 = 0111\ 0000\ 1010\ 1000$$
$$R_1 = 1100\ 0111\ 1010\ 1000$$

Si raggruppano i sample consecutivi in frame, in questo caso prendendo sei sample dal canale sinistro e sei sample dal canale destro, per un totale di 12.

$$L_1R_1L_2R_2L_3R_3L_4R_4L_5R_5L_6R_6$$

Poiché ogni sample è di 16 bit, tutti gli L_i e gli R_i sono composti da due byte. Quindi il frame è composto di fatto da 24 unità da 1 byte, che costituiscono il blocco base dei dati di informazione vera e propria.

$$L_{1a}L_{1b}R_{1a}R_{1b}L_{2a}L_{2b}R_{2a}R_{2b}L_{3a}L_{3b}R_{3a}R_{3b}L_{4a}L_{4b}R_{4a}R_{4b}L_{5a}L_{5b}R_{5a}R_{5b}L_{6a}L_{6b}R_{6a}R_{6b}$$

Politecnico di Torino

Ora si iniziano ad aggiungere i codici descritti in precedenza per arrivare alla struttura completa del CD. Innanzitutto si fa il rimescolamento: i byte di numero pari vengono ritardati di due blocchi, e ciò che risulta viene rimescolato.

$$R_{6b}L_{5a}L_{4a}\ L_{3a}L_{5b}R_{5a}R_{1b}L_{4b}R_{2a}R_{6a}L_{6b}R_{4a}R_{1a}R_{2b}R_{3b}R_{5b}L_{1b}L_{2a}L_{2b}R_{4b}L_{6a}R_{3a}L_{3b}L_{1a}$$

I byte in corsivo sono quelli derivanti dai due blocchi che precedono il blocco in questione nell'ordine cronologico.
A questo punto si aggiungono 4 byte di parità, con una media di un bit di parità ogni 6 bit di dati. Si tratta di una prima assegnazione mediante un criterio detto di tipo Q. Il risultato è ora un blocco di 28 byte (24 + 4), che dovrebbe assomigliare alla seguente catena di byte:

$$R_{6b}L_{5a}L_{4a}L_{3a}L_{5b}R_{5a}R_{1b}L_{4b}R_{2a}R_{6a}L_{6b}R_{4a}R_{1a}R_{2b}R_{3b}R_{5b}L_{1b}L_{2a}L_{2b}R_{4b}L_{6a}R_{3a}L_{3b}L_{1a}Q_1Q_2Q_3Q_4$$

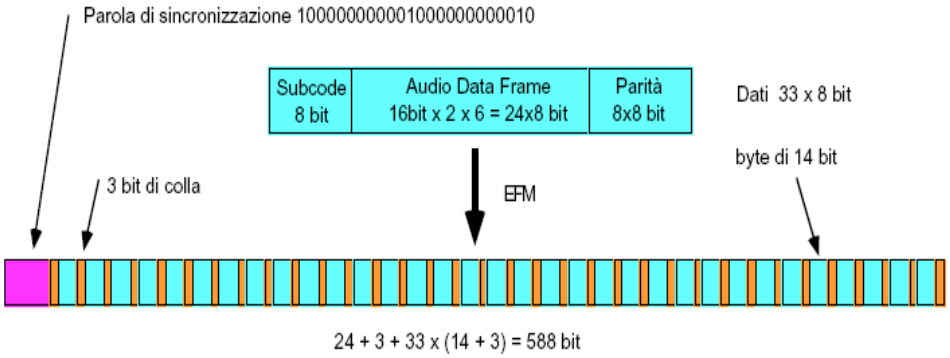
Ora si opera un nuovo rimescolamento in un modo più sofisticato. Ciascuno dei 28 byte viene riposizionato in un nuovo blocco, che viene calcolato moltiplicando per 4 la posizione del byte. Quindi il primo byte finisce 4 blocchi più avanti, il secondo byte finisce 8 blocchi più avanti, il terzo byte finisce 12 blocchi più avanti e così via fino all'ultimo byte, il 24esimo, che finisce 112 blocchi più avanti. I nuovi blocchi così costituiti, sempre di 28 byte, vengono di nuovo sottoposti a una codifica di parità, in questo caso di tipo P. Il risultato è l'aggiunta di ancora 4 byte di parità e il blocco ha ora assunto la dimensione di 32 byte. Le due operazioni di parità contribuiscono al rilevamento e alla correzione di errori di diverso tipo. Per ridurre ancora le possibilità di errore, si fa un ulteriore spostamento dei byte dispari di un singolo blocco.
Ora si aggiungono altri 8 bit, cioè un byte all'inizio del gruppo di 32 byte. Questi bit sono di informazioni aggiuntive che possono essere il numero di brani presenti nel disco, la loro lunghezza, il tempo totale di audio presente sul disco, il controllo sull'accesso casuale ai brani eccetera.
Ovviamente tutte queste informazioni sono distribuite su tutti i gruppi di 8 bit che accompagnano i frame, si tratta quindi di una sorta di sotto-codice: i diversi bit, dal primo all'ottavo, formano come dei canali di informazione che sono distribuiti per tutti i frame presenti sul CD.
Il totale risulta in 33 byte per un frame. Infine, viene aggiunta la codifica EFM.

3.2.6 Codifica EFM

La codifica EFM è un modo molto efficace di ridurre gli errori.
EFM sta per Eight to Fourteen Modulation, cioè Modulazione da 8 a 14: l'idea è di minimizzare il numero di transizioni 0-1 e 1-0 in lettura, evitando quindi la scrittura di pit di piccole dimensioni, che provocano transizioni veloci e che possono essere causa di letture errate; quindi i codici da 8

bit diventano da 14, da cui deriva il nome della codifica. Per esempio, la configurazione di 8 bit 11011111 viene trasformata nella configurazione di 14 bit 00100000010001. Esistono delle tabelle complete per tutti i 256 codici EFM. La codifica EFM comporta delle scritture di pit che coprono da 3 a 11 bit e gli 0, cioè la dimensione dell'avvallamento, sono tra due e dieci 0 consecutivi.
Seguendo le tabelle che codificano tutti i 256 byte possibili in parole da 14 bit, tutto il nostro frame attuale viene convertito in formato EFM. Mentre prima avevamo 33 byte da 8 bit, ora ci ritroviamo 462 bit (33*14). Per completare il frame, si aggiungono ancora 24 bit di sincronizzazione all'inizio del frame (essi sono 100000000001000000000010), che sono due sequenze di dieci 0 consecutivi, il massimo permesso dalla codifica EFM. I bit del frame sono diventati 486.
Infine ad ogni sequenza di 14 bit vengono agganciati 3 bit di intercapedine tra una sequenza di 14 bit e la successiva: lo scopo è di tenere distinti i codici EFM per non far occorrere due 1 consecutivi, che si possono trovare alla fine di un codice e all'inizio del successivo rispettivamente. Questo è il motivo per cui è stata pensata la codifica EFM. I tre bit vengono scelti in modo da avere il minimo della somma digitale. Di codici EFM all'interno di un frame ce ne sono 33, per cui 32 gruppi di 3 bit in mezzo e 2 agli estremi per un totale di 34 gruppi di 3 bit. Il totale dei bit di intercapedine fa 102, che sommati ai 486 bit già presenti nel frame risulta in 588. Si ricordi che i dati audio effettivi da cui siamo partiti erano 12 campioni (6 per canale) di 16 bit ciascuno, per un totale di 192 bit.
Le informazioni extra, che sono contenute nelle aree di entrata (lead-in area) e di uscita (lead-out area), oltre all'area di dati (program area), seguono due modalità.
Nella modalità 1, l'informazione dell'area di entrata consiste nel numero di tracce più il tempo di inizio assoluto di ogni traccia, questa informazione è ripetuta, e permette al lettore CD di costruirsi un indice; nelle aree di programma e di uscita, l'informazione consiste di numero di traccia, numeri di indice all'interno di una traccia, tempo della traccia e tempo assoluto.
La modalità 2 permette di includere anche il numero di catalogo del disco, oltre ad altre informazioni. Esiste anche una modalità 3 non utilizzata.

Il lettore CD legge ogni frame da 588 bit in 136 ms, il che vuol dire 4.323529 bit/s, di cui 1.411765 bit/s circa di puri dati audio (192/588), cioè 177,471 byte al secondo (176 kbyte/s).
Ricordiamo che il tasso di campionamento del segnale registrato su CD è di 44,100 campioni/s, che la quantizzazione è di 16 bit, e che il segnale è stereo. Per poter ascoltare la musica tranquillamente, la velocità di trasferimento dei bit dal CD audio deve essere almeno di 1.411200 bit/s (44.100 x 2 x 16). La velocità del lettore CD è quindi sufficiente a mantenere il tasso per la riproduzione, e inoltre il lettore dispone di una memoria del buffer, che funge da tampone.



Struttura del frame EFM

Ciò che è stato illustrato finora è il CD-Audio, detto anche CD-DA cioè CD Digital Audio, ovvero il primo standard della famiglia dei CD, ma che non aveva questa nomea all'inizio degli anni '80, quando venne lanciato sul mercato. Le sue caratteristiche di durata nel tempo, la capacità estesa rispetto ai mezzi del tempo, in questo caso i dischi in vinile, le caratteristiche di accesso casuale rispetto ad altri mezzi di archiviazione capienti come i nastri, ai quali si accede in modo sequenziale, e la qualità del segnale audio furono i fattori che permisero di conquistare il mercato. In realtà il CD era destinato ad andare oltre la mera funzione l'audio, verso una versatilità inaspettata per l'archiviazione delle informazioni. Infatti, dopo un avvio lento (la standardizzazione è avvenuta nel 1984), il CD-ROM è diventato per un certo periodo di tempo il supporto privilegiato per applicazioni software, giochi, enciclopedie, presentazioni e tutto ciò che la multimedialità è riuscita a produrre.
Ora diamo un'occhiata agli standard successivi al CD-Audio, soffermandoci maggiormente su quelli di maggiore interesse, come il CD-ROM.
Gli standard successivi al CD-Audio, di fondamentale importanza nell'ambito multimediale, sono stati pubblicati in una serie di libri, che hanno assunto i nomi dal colore delle copertine: i più famosi solo il Red Book, che contiene le specifiche del CD-DA, e il Yellow Book, che contiene specifiche del CD-ROM.
Di fatto i lettori di CD-ROM sono compatibili sia con le specifiche del Red Book che con le specifiche del Yellow Book, rendendo così possibile l'ascolto del CD Audio sul PC.

3.2.7 Red Book

Esso contiene le specifiche descritte in precedenza: dalle proprietà fisiche del supporto (dimensioni del disco, parametri del laser) alla codifica (PCM lineare a 16 bit, tasso di

Politecnico di Torino

campionamento a 44,1 kHz, codici di rilevamento e correzione errori, informazioni aggiuntive, sotto-codici). Delle aggiunte successive permisero di sfruttare fino a 4 degli 8 canali di sotto-codici previsti (all'inizio ne erano usati solo due). Facendo i calcoli e prendendo in considerazione le dimensioni del disco, dei pit e delle tracce, e i parametri del segnale, la capacità del CD-DA è di 74 minuti di audio, oggi estesa ad 80 minuti.

3.2.8 Yellow Book

È l'altro libro molto noto della saga CD. Contiene le specifiche del CD-ROM (Read Only Memory), in parte riprese dal Red Book. La principale novità è la struttura dei dati. Anche se è difficile da credere dopo tutto ciò che è stato detto per la codifica del CD-DA, il sistema di rilevamento e correzione degli errori impiegato dal CD-ROM è molto più sofisticato di quello del CD-DA. Il problema è che, se la correzione dell'errore non riesce in un CD musicale, il peggio che può capitare è di udire dei click in riproduzione; viceversa, se il flip di un bit occorre indisturbato in un CD-ROM, il software che è stato acquistato non parte neanche. È stato quindi necessario introdurre un secondo livello di rilevamento e correzione errori (Error Detection Codes, o EDC, e Error Correction Codes, o ECC). La struttura di base del CD-ROM è differente dal CD-DA: mentre nel CD-DA si parlava di frame di 588 bit (con 192 bit di dati), nel caso del CD-ROM è stata imposta una sovrastruttura di 98 frame, detta settore, che permette più spazio ai dati dell'utente. Nel settore vengono accomodati 24 byte per frame e quindi 2352 byte in totale (98*24 = 2352). Di questi 2352 byte, almeno 2048 sono riservati ai dati (in particolare 2048 in Modalità 1 e 2336 Modalità 2), mentre i rimanenti sono riservati ad altre funzioni di sistema.
I primi dodici byte servono per la sincronizzazione: si tratta di configurazioni fisse (dieci byte 11111111, racchiusi in due byte 00000000).
I quattro byte successivi servono da intestazione (header) del settore. I primi 3 byte sono di indirizzamento dei dati: il primo indica i minuti (fino a un massimo di 74), il secondo indica i secondi (fino a un massimo di 59) il terzo indica il settore all'interno dei secondi (fino a un massimo di 75). Questa informazione è ripetuta in un canale del sotto-codice. Il quarto byte indica la modalità: modalità 0 per i dati nulli, modalità 1 con un controllo efficace dell'errore che presenta un livello in più rispetto all'audio (utile per la memorizzazione di software), modalità 2 con un controllo dell'errore pari alla tecnica CIRC vista per l'audio (utile per la memorizzazione di dati multimediali).
I successivi byte contengono i dati effettivi, 2048 byte per la modalità 1 e 2336 per la modalità 2. Quindi, nel caso della modalità 1, il settore prosegue con 4 byte addizionali di un codice per il rilevamento degli errori del tipo controllo della somma applicato ai byte 0-2063 (EDC – il controllo in più rispetto all'audio), 8 byte nulli “intermedi”, e quindi 276 byte di parità (ECC - quelli dell'audio ma applicati al settore, il secondo livello di controllo). Nel caso della modalità 2, oltre ai 16 byte

Politecnico di Torino

iniziali comuni, il resto è tutto riservato per i dati. Non si hanno controlli degli errori oltre a quelli previsti per l'audio, al livello dei frame (non al livello dei settori). La modalità è adatta quando il CD-ROM deve contenere dati multimediali (audio, video,...) per i quali sono sufficienti i controlli di rilevamento e correzione degli errori già presenti nel CD-DA al livello dei frame.

3.3 DVD

Nel 1996 nacque il primo standard del DVD, il DVD-ROM. Sebbene l'acronimo iniziale di Digital Video Disc si sia trasformato in Digital Versatile Disc, l'industria dei computer, che spingeva per un raggiungimento rapido degli accordi sullo standard per immettere il prodotto sul mercato, vedeva l'opposizione dell'industria cinematografica, preoccupata riguardo all'introduzione di criteri di protezione contro la pirateria.
L'impiego attualmente più noto del DVD è nella distribuzione video, dovuto alla migliore qualità offerta rispetto alla cassetta VHS, grazie all'utilizzo della compressione MPEG-2 e, nel caso del suono multicanale, dall'uso congiunto del MPEG 2 e del Dolby-AC3, tutto ciò a costi di produzione decisamente inferiori.
Ma come è avvenuto per il CD, che ha subito un'evoluzione dall'uso esclusivo per l'audio a quello di contenitore di qualsiasi tipo di dati, il DVD-ROM subì la stessa evoluzione e i lettori di DVD cominciarono a sostituire o ad affiancare i lettori CD nei case dei PC, acquisendo per un certo periodo di tempo la nomea di supporto multimediale universale. La capacità elevata del DVD infatti permette ai produttori multimediali di esprimere la propria creatività senza problematiche di spazio, mantenendo alti livelli di qualità audio/video.
Sono quattro le caratteristiche fondamentali che distinguono un DVD da un CD, anche se l'aspetto esteriore è simile: entrambi hanno 12 cm di diametro e 1,2 mm di spessore, entrambi basati sulla tecnologia laser per la lettura dei pit su una traccia a spirale.
La prima differenza riguarda la densità dei pit e delle tracce. Le tracce del DVD infatti sono distanti 0,74 µm (rispetto a 1,6 del CD), e la minima dimensione del pit è di 0,4 µm (rispetto a 0,83 del CD). Grazie a questo accorgimento la capacità del DVD-ROM supera di quattro volte la capacità di un CD. Infatti si ha la metà della dimensione e della distanza tra i pit, per cui la capacità è aumentata di 4 volte.
La seconda differenza è che l'informazione contenuta sulla superficie del DVD può essere disposta su 2 livelli, i quali possono essere letti in modo distinto calibrando il fuoco del laser di lettura. Il primo livello quindi non può essere opaco, in quanto deve far passare la luce per il livello successivo; la capacità del secondo livello deve essere inferiore, perché esso non può essere denso come il primo per evitare sovrapposizioni con i pit. La traccia sul secondo livello viene scritta

dall'esterno verso il centro: arrivando al termine della traccia del primo livello sul bordo esterno, il laser deve solo aggiustare il fuoco senza neanche effettuare uno spostamento per accedere ai dati sul secondo livello. Queste due modalità di scrittura sono denominate PTP (Parallel Track Path) poiché le tracce su entrambi i livelli vanno in direzioni opposte. Il PTP è molto utile per la visione dei film, in cui non sono accettabili lunghe interruzioni.
La terza differenza è che i DVD possono avere due facce, e ciò ne raddoppia la capacità: per focalizzare il laser sui pit, che sono più piccoli rispetto al CD, viene utilizzato un substrato di plastica più sottile, riducendo così il percorso per la lettura dei pit. Essendo il risultato troppo sottile e delicato, i costruttori incollarono due strati, uno sull'altro, raggiungendo lo spessore descritto precedentemente di 1,2 mm. Il risultato è il DVD a doppia faccia.
Infine, la quarta differenza è che il DVD impiega tecniche più efficienti nella correzione degli errori , assicurando più spazio per i dati effettivi. I dischi DVD hanno quindi tempi di trasferimento dati superiori al CD-ROM e tempi di accesso ai dati simili al CD-ROM.

Esistono quattro versioni dei dischi DVD:

DVD-5 : un solo lato e un solo livello, con una capacità di 4,7 GB
DVD-9: un solo lato e due livelli, con una capacità di 8,5 GB
DVD-10: due lati e un solo livello, capacità 9,4 GB
DVD-18: due lati e due livelli, capacità 17 GB

Ora verranno descritti brevemente i formati fisici del DVD che, in molti casi, sono diventati standard internazionali approvati dall'organizzazione ISO, facendo particolare attenzione all'audio.

3.3.1 DVD-Audio

Il DVD-Audio è il formato DVD per l'audio, corrispondente al CD-DA. Tuttavia la storia non è così semplice come negli altri casi. Nel settore audio, le richieste di una fedeltà del digitale che andasse oltre ai parametri offerti dal CD, ha portato ad altri prodotti fuori dall'area del DVD. Il principale competitore del DVD-Audio era il Super Audio CD (SACD) che, progettato da Sony e Philips, si annunciava come il seguito del CD.
La prima specifica del DVD-Audio, che non includeva aspetti di protezione dalla pirateria, è apparsa nella primavera del 1999. Quindi dovettero passare sei anni dall'apparizione del DVD-Audio, tempo che è stato necessario per le varie trattative con l'industria discografica, le quali riguardavano soprattutto la definizione di adeguate tecniche di protezione contro la pirateria.

Le principali caratteristiche del DVD-Audio sono:

Politecnico
di Torino

- audio di qualità multicanale
- compatibilità di lettura tra DVD-Audio e CD
- un' ampia gamma di scelte per i produttori di contenuto riguardanti la qualità e i canali
- apertura verso nuove tecnologie
- contenuti extra, come video, immagini statiche, testo e menù: di fatto sono state rilasciate due versioni, una audio, con testo, immagini e menù ma senza la possibilità di inserire video, che non può essere riprodotta dai lettori video; l'altra versione è audio/video, che contiene una parte di video che può essere riprodotta dai lettori video, ma non dai lettori solo audio.
- sistema di navigazione dei contenuti del disco con interfaccia user-friendly
- connessione a internet per informazioni aggiornate sul titolo

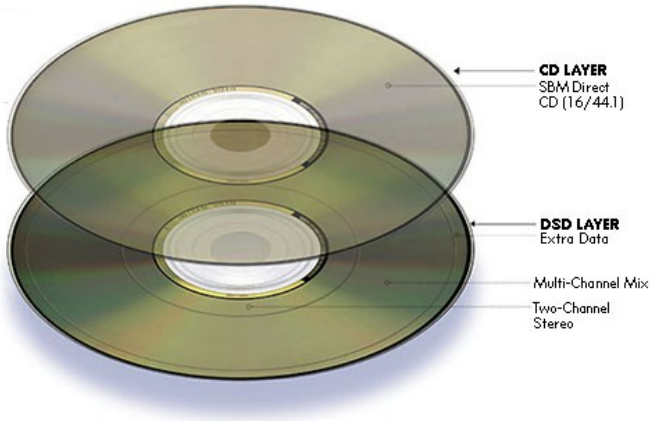
La specifica audio del DVD include alcuni parametri: la codifica (PCM lineare o compressa con la modalità MLP, Meridian Lossless packing), il tasso di campionamento (i più comuni sono 44,1, 48, 96, 192 kHz), la quantizzazione (16, 20, 24, per quest'ultimo il massimo rapporto SNR è di 144 dB) e i canali (6 canali fino a un massimo di 96 kHz o 2 canali fino a 176/192 kHz).
Il massimo tasso di lettura dei campioni permesso è posto a 9,6 Mbps, con differenti tassi di campionamento e quantizzazione. Il formato include anche tabelle di coefficienti per il down-mixing, per ottimizzare il passaggio da audio multicanale a audio stereo quando vi è necessità; ovviamente il down-mixing deve essere supportato dai lettori.
La struttura logica di un DVD-Audio contiene l'informazione (audio, video, …) sotto forma di file.
I dati audio e la loro gestione sono collocati nella cartella AUDIO_TS; gli eventuali dati video, che soddisfano una parte del formato DVD-Video, sono collocati nella cartella VIDEO_TS. Per problemi di velocità di trasferimento dati, l'audio associato al video è sempre in formato compresso.
Oggi il DVD Audio è quasi scomparso dal mercato ed ufficialmente non viene più prodotto industrialmente.

3.4 Super Audio CD

La questione della compatibilità, insieme con la lentezza delle decisioni intorno ai formati DVD per colpa degli innumerevoli settori coinvolti (industria cinematografica, computer, musica), portò Sony e Philips a pensare in proprio alla seguente generazione di supporti per l'audio di qualità. Il progetto del Super Audio CD prende forma con l'obiettivo di unificare la tecnologia DVD e la retro-compatibilità verso il CD tradizionale, inteso come quello standardizzato dal Red Book.
Anche il SACD è un disco a due livelli, fatto più o meno come il DVD. I due livelli sono costituiti da due substrati di policarbonato di circa 0,6 mm, quindi lo spessore totale è lo stesso del DVD, cioè

Politecnico di Torino

di 1,2 mm. Il livello più alto, quello vicino all'etichetta, è a bassa densità, con pit impressi della stessa dimensione di un CD tradizionale. Il livello più basso è ad alta densità, con pit molto più ridotti. I due dischi, una volta stampati, vengono incollati assieme, in modo tale che il livello ad alta densità venga incollato al di sotto del livello a bassa densità, cioè il livello CD tradizionale. Il laser nello spettro dell'infrarosso deve arrivare al livello più alto per leggere i dati del CD; il laser nello spettro del rosso deve stare al livello più basso per leggere i dati ad alta densità. Questo obiettivo viene raggiunto coprendo il livello più basso con un materiale semi-trasmissivo a seconda della lunghezza d'onda del raggio laser: il comportamento di questa pellicola, che conferisce un colore dorato, è di far passare quasi completamente il laser nell'infrarosso, che è quello del lettore CD tradizionale, permettendogli quindi di leggere i dati del layer del CD tradizionale, e di riflettere parzialmente (25%) il laser nel rosso visibile, consentendogli così di vedere solamente il livello ad alta densità che gli pertiene. In questo modo si è ottenuta la compatibilità tra i vari tipi di riproduttori: inserendo un SACD all'interno di un normale lettore CD, esso leggerà senza problemi il layer CD, ignorando la parte ad alta densità; utilizzando un lettore apposito si potrà leggere anche la parte del layer SACD, risultando quindi in un formato molto versatile. Il costo di manifattura di questi dischi ibridi è solo leggermente superiore al DVD a due livelli che però non è retro compatibile.



Super Audio CD con i due layer CD e DSD messi in risalto

3.4.1 DSD

Una delle maggiori innovazioni risiede nel fatto che la parte del disco ad alta densità viene codificata con il formato Direct Stream Digital (DSD). Esso è' un sistema proprietario della Sony che usa flussi di dati di 1 bit con un elevato tasso di campionamento di 2822,4 kHz, 64 volte più veloce del 44,1 kHz. L'idea è di rivoluzionare il sistema audio digitale così come è stato conosciuto finora, dove il DSD può risultare un sostituto del PCM. Di fatto, anche il DSD è un PCM, ma speciale, in quanto non si tratta dello standard PCM multi bit ma di un PCM a un solo bit. La tecnica DSD è un'evoluzione del sovracampionamento e si basa completamente su un flusso di dati composto da un solo bit, eliminando le operazioni di filtraggio del sovracampionamento, cioè di selezione di un campione ogni n, e di decimazione per arrivare ai bit della codifica PCM richiesta. Poiché i campioni sono di un solo bit, il sistema DSD attuale ha un bitrate di 2.822.400 bit/s, e quindi presenta un bitrate solo quattro volte superiore al CD (4 x 16 bit x 44,1 kHz), sebbene il numero di campioni al secondo sia 64 volte superiore. Il bitrate è leggermente superiore al sistema ad alta risoluzione a 96 kHz e 24 bit (2,304 Mb/s) e la gamma dinamica che ne risulta è di circa 120 dB, cioè pari a un sistema multibit di 20 bit, ma la qualità giudicata in modo soggettivo da molti artisti, produttori, ingegneri del suono e critici è decisamente superiore, ed è probabilmente pari se non superiore ai sistemi a 96 kHz e 24 bit.
Il SACD è ancora in produzione, seppure con quantitativi ridottissimi. La codifica DSD non ha mai riscosso, dalla sua introduzione nel 1991, un grande successo nel settore dell'audio professionale, per una serie di gravi problemi tecnici, e i SACD prodotti al giorno d'oggi vengono quasi sempre registrati a 24 bit/96 kHz e poi convertiti in un master DSD.

3.5 DAT

Il DAT nasce come formato audio di qualità alla metà degli anni '80 e nel 1987 viene trasformato da Sony e HP in un formato ad uso generico per l'archiviazione di dati digitali. Sebbene non sia riuscito nel compito di sostituire sul mercato la cassetta analogica e non abbia mai guadagnato il mercato consumer, diventò uno standard negli ambienti di registrazione sia professionale sia amatoriale. Una motivazione sul mancato successo del formato fu dovuto alla RIAA, che fece da muro contro il formato DAT, tentando di imporre delle restrizioni e della tecnologia anti-pirateria sui riproduttori, poiché aveva paura che questo formato incrementasse le copie piratate, vista la facilità di copia e l'alta qualità che si poteva ottenere. La tecnologia prevedeva un nastro magnetico largo 4 mm e con una registrazione magnetica a scansione elicoidale (cioè la stessa usata nei nastri video, e usata in generale per le grandi capacità su nastro). Trattandosi di una

registrazione di dati digitali, il nastro veniva registrato “in saturazione” (così come accade negli HD) eliminando quindi la testina di cancellazione ed il bias, presenti invece nei registratori video. Inoltre le tracce erano scritte con un parziale overlap, aumentando di molto la densità di informazione su nastro Il DAT supportava i formati di registrazione 32 kHz a 12 bit su due tracce (in modalità cosiddetta LP – Long Playing), 32 kHz a 16 bit su due o quattro tracce, 44,1 o 48 kHz a 16 bit su 2 tracce, solo i registratori per il mercato professionale potevano registrare a 44,1 kHz mentre quelli consumer registravano solamente a 48 kHz e riproducevano a 44,1 e 48 kHz, per impedire la copia digitale dei CD. Furono disponibili sul mercato anche registratori a 96 kHz/24 bit, ma il materiale registrato non era poi riproducibile su altre macchine.
Il DAT ha avuto notevole fortuna solo nell'archiviazione di dati digitali generici (sistemi di backup), ma occorre ricordare che, al contrario di quanto avviene per i lettori CD, le macchine che leggono i dati su DAT non riproducono l'audio (a parte qualche eccezione ormai obsoleta). Esistono più formati implementati sulla tecnologia DAT, che vanno dal solo audio ai formati più recenti come il DDS (Digital Data Storage). Questi ultimi hanno capacità variabili da 2 GB fino a 40 GB, e tasso di trasferimento dati tra 55 kB/s e 4,8 MB/s. Quindi sufficienti per l'audio di alta qualità.

Una cassetta DAT

Politecnico di Torino

4 – Conversione Analogico Digitale

4.1 Introduzione

Come accennato nel capitolo dei supporti analogici, una rappresentazione analogica è una rappresentazione che lavora per analogia: la curva continua nel tempo delle variazioni di ampiezza viene rappresentata da una curva continua nel tempo delle variazioni di tensione elettrica, la quale viene memorizzata nei solchi del disco in vinile o nel campo magnetico di un nastro. Il tracciato dei solchi del vinile segue l'andamento della curva dell'ampiezza, dopo che questa è stata convertita in un segnale elettrico; al contrario una rappresentazione digitale non cerca di imitare la curva continua di ampiezza con una curva analoga ad essa, ma piuttosto assegna dei numeri che rappresentano di volta in volta il valore dell'ampiezza in istanti successivi di tempo; sarà questa successione di numeri a rappresentare il segnale audio.
La differenza balza subito all'occhio: da una parte si ha una curva continua che approssima una curva continua; dall'altra si ha una successione di numeri. La rappresentazione digitale non è continua ma discreta, cioè esistono degli eventi ben definiti che sono i valori dell'ampiezza in precisi istanti di tempo. Essa si avvale di un codice simbolico preciso dato dalle cifre che compongono i numeri: fissate le cifre a disposizione e la lunghezza dei numeri binari, la rappresentazione avrà a disposizione un insieme finito di simboli preciso ed univoco.
I vantaggi di un codice simbolico preciso sono innumerevoli. Uno dei vantaggi più evidenti si ha nelle operazioni di copia del segnale. E' noto infatti come copie ripetute dello stesso segnale analogico producano una degradazione sempre maggiore della qualità del segnale: ciò è intuitivamente comprensibile, in quanto l'operazione di copia deve di nuovo approssimare con una curva continua la curva che era stata memorizzata in precedenza, introducendo sempre più errori. Al contrario, fare una copia di numeri non comporta alcun errore: la curva memorizzata in precedenza rimane sempre la stessa e non subisce alterazione nella clonazione.
Lavorando con i numeri è anche immediato constatare quanto debba essere facile manipolare il segnale: le operazioni di manipolazione altro non sono che operazioni aritmetiche su dei numeri binari, le quali, la maggior parte delle volte, sono reversibili. Ad esempio, se si vuole aumentare l'intensità di un frammento di un suono, è sufficiente moltiplicare i numeri che corrispondono a quel frammento per il fattore desiderato. Nel caso analogico, invece, occorre applicare un dispositivo elettronico che accresce l'ampiezza della curva in questione, che introduce distorsione.
Un altro aspetto totalmente nuovo introdotto dai supporti dalla rappresentazione digitale è la possibilità di correzione degli errori introdotti dai supporti per la memorizzazione, come visto per i CD, e dai dispositivi di trasmissione.

La rappresentazione digitale del segnale porta anche degli svantaggi. Innanzitutto occorre realizzare un sistema di comunicazione tra l'interno e l'esterno dell'elaboratore: nel mondo esterno, quello in cui ascoltiamo i suoni, i numeri devono essere convertiti in pressione sonora; nella rappresentazione analogica un'onda sonora manteneva la sua natura di forma d'onda per tutto il percorso della catena elettroacustica; questo processo di conversione implica una certa perdita di qualità. Inoltre la qualità del segnale, affinché sia comparabile con i migliori apparati analogici, necessita di grandi capacità di memoria e velocità di trasmissione per la memorizzazione e il trasferimento del segnale, comportando costi a volte eccessivi in termini di potenza di calcolo. Servono quindi grandi capacità della CPU, e di spazio di memorizzazione, poiché i file audio digitali non compressi occupano molto spazio.
La digitalizzazione del suono ha comportato una serie di trasformazioni a partire dagli anni '80, le quali hanno interessato sia i professionisti del settore nelle pratiche di elaborazione del suono, sia l'evoluzione della musica.
Per comprendere come le variazioni di pressione dell'aria vengano convertite in sequenze di 0 e 1, occorre seguire la catena dell'audio digitale.

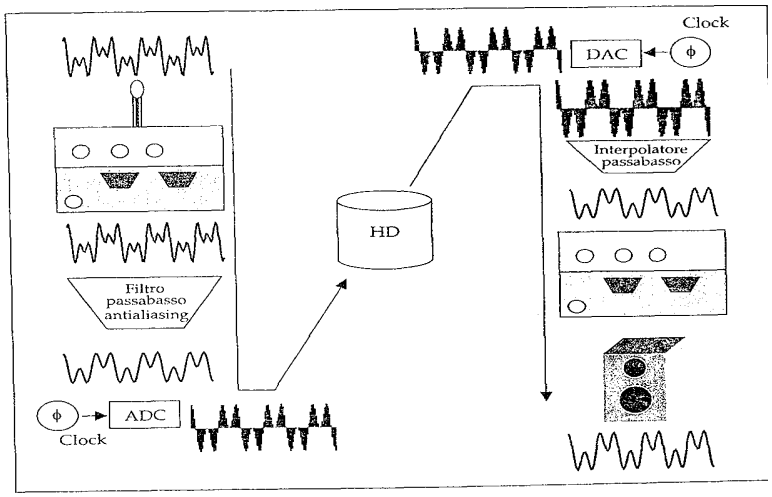
4.2 Catena Audio Digitale

La catena dell'audio digitale parte dalla semi-catena di registrazione. Il segnale analogico della pressione sonora viene convertito in un segnale elettrico. A questo punto occorre convertire questo segnale in formato digitale: lo strumento fondamentale per la conversione è il convertitore analogico-digitale (analog to digital converter – ADC, comunemente chiamato A/D) che, in base a un suo orologio interno, detto clock, preleva i campioni dell'ampiezza dal segnale analogico.

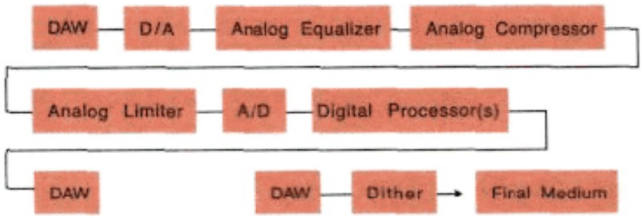
Il convertitore A/D Focusrite Saffire Pro 26

Politecnico di Torino

Il clock fissa la frequenza di campionamento, detto sample rate, che rappresenta l'intervallo di tempo nel quale vengono prelevati i campioni. I campioni sono numeri espressi in bit, che possono essere memorizzati su un qualsiasi dispositivo digitale, simbolizzato in figura dall'icona cilindro che rappresenta il disco rigido.
Prima di convertire il segnale occorre però filtrarlo in modo da eliminare una parte del suo spettro che potrebbe causare il cosiddetto problema dell'aliasing. Infatti, secondo il teorema di Nyquist che verrà approfondito nel paragrafo 4.3.1, per digitalizzare un segnale che contiene componenti di frequenza fino a x Hertz e ricostruirlo fedelmente occorre campionare ad un sample rate f_c di almeno 2x campioni per secondo. Stabilito il sample rate occorre eliminare dal segnale tutte quelle frequenze che sono oltre la metà di f_c. Per far ciò si usa un filtro che elimina tutte le frequenze sopra una certa soglia, detta frequenza di taglio o di cut-off. Tale filtro è detto passa basso in quando fa passare solo frequenze sotto una determinata soglia ed è detto antialiasing in quanto appunto evita il problema dell'aliasing.

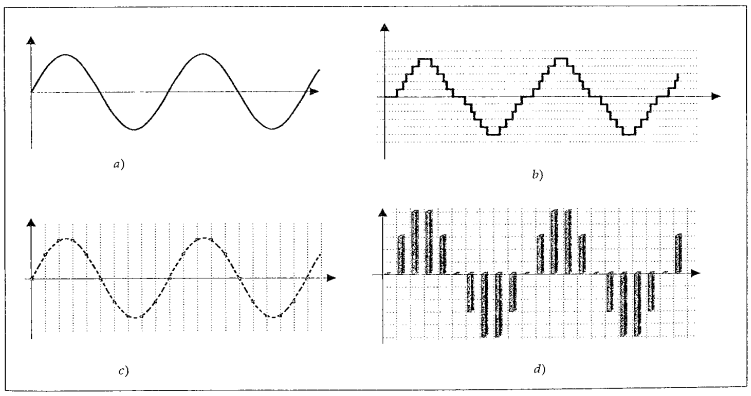


Catena elettroacustica della conversione digitale
fonte:Audio e Multimedia



Tipica catena di elaborazione di un segnale audio musicale in uno studio professionista
fonte: Mastering Audio: the Art and the Science

Il passaggio da analogico a digitale è caratterizzato da due operazioni di discretizzazione, in cui delle grandezze continue vengono ad assumere valori discreti: la prima grandezza discretizzata è il tempo, dal momento che il clock dei convertitori stabilisce quali sono gli istanti da considerare per il campionamento; la seconda è l'ampiezza, dal momento che i numeri che è possibile usare nella rappresentazione appartengono a un codice simbolico definito e preciso. L'operazione che discretizza il tempo è detta campionamento, l'operazione che discretizza l'ampiezza è detta quantizzazione. Nella figura sottostante si possono osservare gli effetti separati e combinati delle due operazioni: in (a) si ha il segnale analogico di partenza; in (b) il segnale in forma quantizzata, dove l'ampiezza assume solo alcuni valori e non tutte le infinite possibilità offerte dal continuo attraverso una circuiteria di campionamento detta Sample/Hold che incamera e misura il livello analogico per la durata del periodo di campionamento mentre il tempo rimane una grandezza continua; in (c) il tempo esiste solo in istanti discreti, quindi anche il segnale esiste solo in quegli istanti ma i valori che può assumere l'ampiezza sono infiniti; in d) il segnale è discreto sia nel tempo che nell'ampiezza cioè è digitale e numerico.



Operazione di campionamento e quantizzazione fonte:Audio e Multimedia

Politecnico di Torino

4.3 Campionamento

Il campionamento è la discretizzazione del segnale analogico nel tempo. Visto che il segnale audio è rappresentato dalla variazione dell'ampiezza nel tempo, il campionamento corrisponde alla individuazione dei valori discreti sull'asse delle ascisse. Intuitivamente, affinché il segnale digitale rappresenti in modo efficace il segnale analogico di partenza, occorre che i campioni siano prelevati ad una velocità abbastanza elevata da riuscire a tener conto di tutte le variazioni del segnale. Se il segnale passa da una fase ascendente a una fase discendente e viceversa in un lasso di tempo ristretto, gli istanti di tempo considerati devono appartenere a tutte le tre fasi; se una della fasi non viene considerata, il segnale digitale non potrà rendere conto di tale variazione. Di conseguenza, la velocità di acquisizione dei campioni, cioè il sample rate, deve essere dipendente dalla velocità di variazione del segnale. L'intervallo di tempo tra un campione e il successivo è detto periodo di campionamento; l'inverso del periodo, e quindi il numero di campioni in un secondo, è detto sample rate e viene espresso in Hz (s^{-1}). Esso dipende a sua volta dalla parziale con frequenza più alta: sarà questa frequenza a determinare la velocità massima di variazione del segnale e quindi anche il sample rate, affinché la digitalizzazione non provochi degli errori di rappresentazione. In modo ancora più intuitivo, più brevi sono gli intervalli di tempo tra un campione e il successivo, più simile al segnale analogico originario sarà il segnale a gradini: se si porta all'estremo questo ragionamento per intervalli di tempo infinitamente brevi, cioè per sample rate altissimi, i due segnali analogico e digitale coincideranno. Ad ogni periodo di campionamento si preleva un campione dell'ampiezza del segnale, cioè il valore del segnale analogico in quell'istante: il risultato è una sequenza di valori che corrispondono all'andamento del segnale. Più sono ravvicinati i valori, più efficacemente sarà descritto l'andamento del segnale. Poiché man mano che si riduce la dimensione dell'intervallo di campionamento aumenta l'accuratezza della descrizione del segnale, è cruciale scoprire qual'è la dimensione migliore dell'intervallo: i due parametri per i quali occorre trovare un equo compromesso sono la qualità del suono e il numero di campioni. Una migliore qualità del segnale corrisponde a una maggiore quantità di informazione, cioè a un maggior numero di campioni nell'unità di tempo; un maggior numero di campioni però richiede un maggiore spazio in memoria e una velocità superiore di trasferimento dati. Nella digitalizzazione di un segnale i due parametri rispettivamente di qualità del suono e di quantità dei campioni sono in contrasto e la scelta del sample rate dipende dalle esigenze delle applicazioni individuali. Partendo tuttavia dall'idea di avere il massimo della qualità, cioè di riuscire a non percepire la differenza tra il segnale analogico e il segnale digitale, bisogna stabilire il minimo valore di velocità di campionamento in modo tale che non si abbia della perdita d'informazione. Infatti anche nella digitalizzazione occorre tenere a bada i due nemici di sempre della telecomunicazione, la distorsione e il rumore, rivisitati in senso digitale.

Politecnico di Torino

4.3.1 Nyquist e Aliasing

La nostra percezione rimane invariata dal segnale analogico al segnale digitale se la forma d'onda, e quindi lo spettro del segnale, non subisce alcuna alterazione quindi se le parziali del segnale in ingresso sono identiche al segnale in uscita oppure se le alterazioni subite non sono percepibili. Nella digitalizzazione di un segnale si ha un fenomeno di distorsione dovuto tipicamente all'operazione di campionamento che ci è utile per fissare dei limiti inferiori al sample rate, il cosiddetto aliasing. L'aliasing è un fenomeno che introduce in un segnale digitale delle frequenze spurie, che non erano presenti nel segnale analogico in input. A partire dall'aliasing, si possono formulare dei principi generali sulla velocità di campionamento.
Il principio su cui si basa l'annullamento della distorsione è il teorema di Nyquist, formulato nel 1928 da Harry Nyquist e da cui deriva la teoria dell'informazione di Claude Shannon del 1949. Secondo il teorema, il sample rate deve essere almeno il doppio della frequenza massima presente nel segnale in input: questo tasso di campionamento è detto tasso di Nyquist.

$$f_{nyquist} \geq 2f_{max}$$

Se il sample rate è al di sotto del tasso di Nyquist, si verifica il fenomeno dell'aliasing. In generale, supponiamo che f_c sia il tasso di campionamento e f sia una frequenza presente nel segnale di ingresso tale che:

$$f > f_c/2$$

cioè f è superiore alla metà del sample rate, mentre per rispettare il teorema di Nyquist dovrebbe essere il contrario ($f \leq f_c/2$). La frequenza ricostruita f_r sarà nell'intervallo tra $-f_c/2$ e $f_c/2$ cioè:

$$-f_c/2 < f_r < f_c/2 < f$$

Poiché la frequenza ricostruita f_r viene a cadere in un intervallo di frequenze più basso della frequenza originale f, avviene il fenomeno di aliasing, detto anche foldover.
Esiste una formula che permette di calcolare la frequenza ricostruita f_r a partire dalla frequenza originale f e il sample rate f_c:

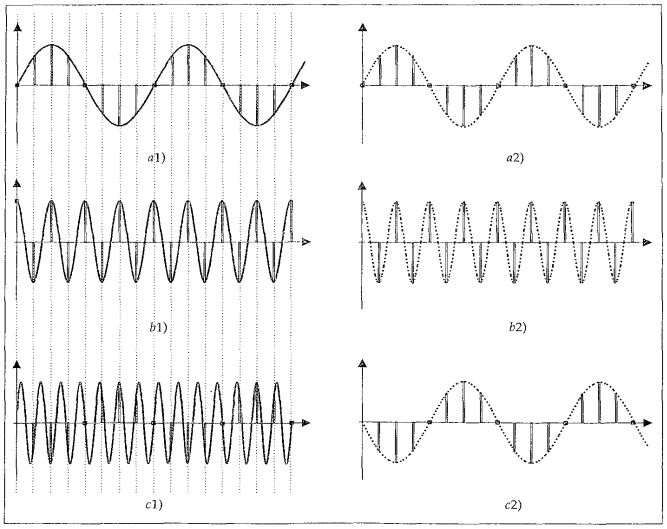
$$f_r = f + kf_c$$

Politecnico di Torino

cioè la frequenza ricostruita f_r è data dalla somma della frequenza originale f e del sample rate f_c moltiplicato per un numero k, il quale è un numero intero positivo o negativo che soddisfa la relazione:

$$-f_c/2 < f + kf_c < f_c/2$$

cioè k fa in modo che la frequenza ricostruita cada nell'intervallo $(-f_c/2, f_c/2)$.
La verifica immediata della relazione si ha per le frequenze appartenenti all'intervallo $(-f_c/2, f_c/2)$: in questo caso la formula è soddisfatta per k=0 descrivendo un comportamento corretto, senza aliasing.
Per poter calcolare in modo immediato la frequenza ricostruita basta calcolare k a partire dalla disuguaglianza $-f_c/2 < f + kf_c < f_c/2$ e poi applicare la $f_r = f + kf_c$.
In generale si ha che $k = -2f/f_c$, per cui il calcolo di f_r risulta essere $f_r = f – (2f/f_c)f_c$.
Le conseguenze dell'aliasing possono essere disastrose dipendendo dai rapporti che le frequenze spurie, introdotte ex-novo nella conversione, hanno con le altre frequenze presenti nel segnale. Ad esempio, nel caso un cui la frequenza spuria sia molto vicina a una frequenza già presente nel segnale, si potrebbe verificare il fenomeno dei battimenti. Si osserva che per descrivere correttamente il segnale occorre ricordarsi di ogni singola variazione del segnale. Se l'operazione di campionamento manca di memorizzare qualcuna delle variazioni, si ha un sotto campionamento e il fenomeno di aliasing. Poiché in un segnale complesso sono presenti più parziali, occorre che il campionamento prelevi almeno due campioni per la parziale di frequenza massima presente nel segnale; le altre parziali saranno tutte sovracampionate, ma questo non è certo un problema.



In figura troviamo tre casi di campionamento con differente sample rate e relativa ricostruzione del segnale.
fonte:Audio e Multimedia

Nella figura soprastante si ha nella colonna di sinistra il campionamento del segnale e a destra il segnale ricostruito. Nel caso a) si ha il sovracampionamento, per cui la frequenza di campionamento f è maggiore della frequenza di Nyquist, permettendo di ricostruire il segnale in maniera corretta e, come si vedrà più avanti, permette di ridurne gli artefatti; nel caso b) si ha campionamento critico dove la frequenza di campionamento è uguale a quella di Nyquist ed è sufficiente a ricostruire il segnale; nel caso c) si ha un caso di sotto campionamento, che provoca un errore nella ricostruzione, introducendo aliasing.
Stabilito un tasso di campionamento, per campionare correttamente il segnale in input occorre che le frequenze che sono la metà della frequenza di campionamento vengano eliminate. Con l'eliminazione di tali frequenze il segnale risulterà impoverito rispetto al segnale originale, ma non è detto che le conseguenze di questa riduzione del contenuto armonico siano percepibili. Ad esempio, poiché la percezione uditiva di un uomo in salute arriva a segnali di circa 20 kHz, è sufficiente che il tasso di campionamento sia fissato ad oltre 40 kHz per non avere perdite significative nella qualità del segnale ricostruito. Il tasso di campionamento dei CD è di 44,1 kHz, quindi sufficiente a rappresentare correttamente il segnale audio di partenza. Per eliminare le frequenze dal segnale in ingresso occorre filtrare il segnale in modo da far passare solo le frequenze al di sotto della metà del tasso di campionamento. Il dispositivo che permette di fare questa operazione è il filtro passa-basso. L'eliminazione delle frequenze più elevate è una semplificazione: in realtà il filtro riesce solo ad attenuare le frequenze indesiderate. La soglia oltre la quale le frequenze vengono attenuate è detta frequenza di taglio o di cut-off. Le stesse tematiche che riguardano il filtro passa-basso riappaiono nella ricostruzione del segnale analogico, quando il filtro deve eliminare i cambi repentini di intensità dovuti alla successione di tensione generate dal DAC in corrispondenza della letture dei singoli campioni dal segnale digitalizzato. Nel caso del campionamento il segnale di partenza era il segnale analogico originale; nel caso della ricostruzione il segnale di partenza è il segnale a scalini generato dal DAC e dal circuito di Sample/Hold che è stato memorizzato su un supporto digitale. I cambi di intensità del segnale a scalini sono repentini, in quanto la tensione rimane costante fino alla conversione del campione successivo, ma il segnale analogico di partenza non può avere questi cambi repentini di intensità, in quanto era stato filtrato mediante un filtro passa-basso sotto la metà del tasso di campionamento. Per ricostruire in modo corretto il segnale è sufficiente quindi che il segnale a scalini sia filtrato di nuovo mediante un filtro passa-basso, tarato come il filtro usato nella fase di campionamento. Il segnale digitalizzato è quindi a banda limitata, ma arriva agli altoparlanti come il corrispettivo sonoro e grafico del segnale originale.

Politecnico di Torino

4.4 Quantizzazione

L'operazione di quantizzazione sul singolo campione di ampiezza di tensione elettrica ha l'obbiettivo di assegnare al campione uno dei valori numerici che sono consentiti nella codifica digitale. La codifica digitale divide la gamma delle ampiezze possibili in intervalli, o regioni, in cui viene assegnato il valore di ogni campione in base al proprio valore d'ampiezza. Il numero di regioni in cui è suddivisa la gamma delle ampiezze dipende dai bit a disposizione per la codifica: la rappresentazione digitale, infatti, si basa sulle sequenze di bit. Il bit è l'unità minima di informazione, e può valere 0 o 1; una sequenza di bit è detta parola binaria. Una parola binaria di N bit può assumere 2^N valori o configurazioni differenti, cioè si tratta di 2^N possibili sequenze differenti di 0 o 1. Valori tipici per la lunghezza di una parola binaria nell'audio sono 8, 16 e 24 bit, per cui si hanno un numero di valori rappresentabili di 256 (2^8), 65.535 (2^{16}) e 16.777.215 (2^{24}) rispettivamente. Fissato il numero di regioni possibili mediante i bit a disposizione, occorre stabilire un metodo per associare una configurazione di bit a una regione della gamma delle ampiezze.

4.4.1 Quantizzazione lineare ed errore di quantizzazione

Un metodo comune è la quantizzazione lineare che divide la gamma delle ampiezze in 2^N regioni di ampiezza identica e associa a ciascuna delle regioni uno dei 2^N- valori numerici. Ciascuno dei valori numerici è un codice binario a N bit. Nell'operazione di quantizzazione, tutti i valori di tensione che finiscono in una certa regione vengono convertiti nel codice binario che è assegnato a quella regione. Nel processo di quantizzazione si perde dell'informazione, in quanto tutti i valori che rientrano in una regione verranno convertiti nello stesso codice binario, riducendo quindi le differenze tra i valori di tensione come se il codice corrispondesse a un valore di tensione che potrebbe essere il valore medio della regione. Rispetto al valore medio, quando viene quantizzato un altro valore della regione, si introduce un errore che è detto errore di quantizzazione. Man mano che i bit della codifica aumentano, le regioni si riducono di ampiezza e il codice binario viene quindi associato a una regione più piccola, riducendo l'errore di quantizzazione. Il massimo errore di quantizzazione è dato dalla metà dell'ampiezza di una regione, in quanto assumendo che il codice corrisponda al valore di mezzo, al più il valore di tensione per un certo campione sarà agli estremi della regione.
Si assuma che la tensione del segnale abbia come ampiezza V, cioè +V e -V sono rispettivamente i valori di massima e minima ampiezza del segnale. Se si hanno a disposizione N bit, la dimensione delle regioni di quantizzazione sarà di $2V/2^N$ e la massima ampiezza dell'errore di quantizzazione sarà la metà della regione di quantizzazione, cioè $V/2^N$. Si assuma inoltre che il segnale sia

Politecnico di Torino

sinusoidale. In teoria dei segnali il valore efficace (in inglese Root Mean Square, abbreviato con "RMS") di una funzione periodica è il valore che avrebbe un segnale costante di pari potenza media. L'RMS di un segnale sinusoidale è il valore di picco V diviso la radice di 2, cioè $V/\sqrt{2}$. L'RMS dell'errore di quantizzazione è $V/2^{N-1} * 1/\sqrt{12}$, tipica per un'onda a dente di sega.
Il rapporto SNR, cioè il rapporto tra l'RMS del segnale e il suo errore di quantizzazione, si può allora scrivere nel modo seguente:

$$SNR = (V/\sqrt{2}) / (V/2^{N-1} * 1/\sqrt{12}) = V/\sqrt{2} * [(2^{N-1} * \sqrt{12})/V] = 2^N * \sqrt{(3/2)}$$

$$SNR_{dB} = 20 \log_{10}(2^N * \sqrt{(3/2)}) = 20 (N\log_{10} + \log_{10} \sqrt{(3/2)}) = 6{,}02N + 1{,}7609$$

N è il numero di bit della rappresentazione digitale e SNR è espresso in funzione di N: quest'ultima relazione stabilisce che ogni bit contribuisce per circa 6 dB al rapporto tra il segnale e l'errore di quantizzazione SNR.
Per avere un' idea del contributo dato da un singolo bit al rapporto segnale/rumore, si consideri che con 8 bit l'SNR è uguale a 48 dB, mentre con 16 bit l'SNR risulta 96 dB, che è un valore molto alto anche rispetto ai migliori registratori analogici. Questo ultimo rapporto è quello utilizzato dallo standard del CD, ma le tecniche di registrazione adottate nell'industria discografica si basano anche su una quantizzazione a 24 bit e godono di un rapporto SNR di circa 144 dB.

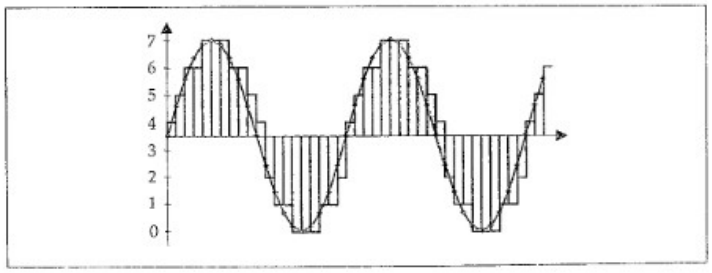
4.4.2 Dithering

Occorre notare che i valori del SNR riguardano un segnale alla massima ampiezza; per ampiezze inferiori, il rapporto SNR scende. Infatti le regioni interessate sono in numero inferiore, per cui si ha un numeratore inferiore. Nel caso di ampiezze minime il segnale varia solo su due regioni; quando il segnale varia all'interno di una singola regione esso è nullo per il sistema digitale e ciò risulta in silenzio; in tal caso si ha una variazione dell'ultimo bit della parola binaria e tale variazione è detta flip, dal momento che la variazione interessa solo i valori 0 e 1. Se si verifica una regolarità del segnale per cui si alternano con una certa periodicità campioni in una regione e campioni nell'altra regione, il segnale ricostruito sarà un'onda quadra o un'onda a impulsi. Il fenomeno è comunemente denominato “low level quantization noise”, in quanto causato dall'errore di quantizzazione alle piccole ampiezze. La presenza di quest'onda introduce quindi nel segnale ricostruito delle armoniche indesiderate. La presenza di armoniche introdotte nel segnale ricostruito causa alle ampiezze minime degli effetti sgradevoli, simile ad una distorsione “metallica” nella fase di estinzione di una nota. La soluzione in questi casi, soprattutto per convertitori non particolarmente sofisticati, è una correzione particolare, detta dithering. Prima della conversione A/D si introduce nel segnale del rumore bianco. Il livello di introduzione del

rumore è molto ridotto, pari a ⅓ del valore del bit meno significativo. L'andamento casuale del rumore, che è indipendente dal segnale in ingresso, provoca delle variazioni casuali nei valori quantizzati alle ampiezze minime, trasformando l'energia delle armoniche in un rumore bianco, molto meno udibile di una distorsione armonica. Il dithering è una soluzione appropriata soprattutto per digitalizzazioni a basso costo, in quanto l'aumento dei bit di quantizzazione, che migliorano il rapporto SNR, fa aumentare i costi delle apparecchiature di registrazione. Nelle applicazioni audio è preferibile utilizzare del rumore dither con densità di probabilità triangolare (detto anche dither TPDF, cioè Triangular Probability Distribution Function).
Infine, si prenda in considerazione la gamma dinamica di un segnale digitale, che è la differenza in dB tra l'ampiezza massima e l'ampiezza minima del segnale. Anche la gamma dinamica aumenta con il numero di bit: ogni bit raddoppia la gamma dinamica della rappresentazione digitale (con 3 bit 8 valori, con 4 bit 16 valori). Con ragionamenti simili al rapporto SNR, si ha che la gamma dinamica per un sistema digitale a N bit è 6N dB.
Quindi, come si può notare, nel caso della quantizzazione lineare la gamma dinamica ha lo stesso valore del rapporto SNR e con 16 bit si ha che la gamma dinamica è di 96 dB.

4.5 Codifica PCM

Il passo finale della digitalizzazione è la generazione del codice associato al campione e che identifica tutta la regione. In genere, valori in successione della codifica sono associati a regioni successive di quantizzazione. I codici sono parole binarie che possono rappresentare valori interi e sono questi che vengono memorizzati sui supporti per il digitale o vengono trasmessi per via telematica. Esistono molti modi per codificare un segnale. Il modo forse più immediato è la modulazione dell'ampiezza dell'impulso (Pulse Amplitude Modulation – PAM), dove un impulso occorre ad ogni intervallo di campionamento e l'ampiezza della forma d'onda è un valore digitale che corrisponde all'ampiezza analogica.

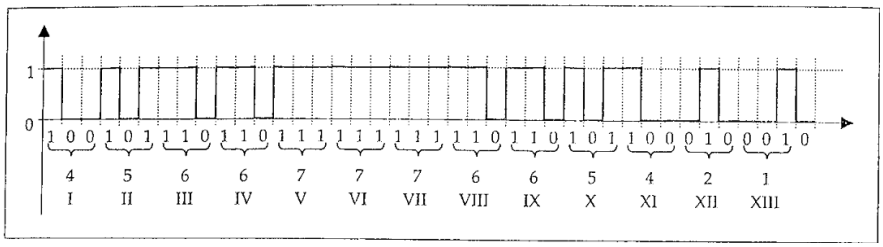


Modulazione dell'ampiezza d'impulso (PAM) fonte:Audio e Multimedia

Il segnale modulato PAM è ancora analogico: per immagazzinare le informazioni in esso contenuto su un supporto digitale occorre quindi convertire ogni singolo impulso in un numero binario.

Politecnico di Torino

La forma più diffusa di codifica digitale avviene mediante la modulazione del codice dell'impulso (Pulse Code Modulation – PCM): l'informazione digitale viene codificata in modo seriale con l'impulso rappresenta l'1 e l'assenza di impulso rappresenta lo 0.



Modulazione dell'ampiezza d'impulso (PCM) fonte:Audio e Multimedia

Con una parola digitale ad N bit si possono esprimere 2^N-1 valori diversi, dove il valore sottratto rappresenta lo 0 digitale, cioè il valore centrale della rappresentazione. Nel campo audio viene utilizzata la codifica denominata a complemento di due, dove i valori positivi vengono rappresentati in binario naturale e quelli negativi con tutti i bit invertiti e l'aggiunta di un 1 rispetto ai corrispondenti positivi. Alla base della scelta del codice a complemento di due c'è una ragione ben precisa: è infatti l'unico in grado di operare simmetricamente rispetto al valore centrale, caratteristica che risulta fondamentale nell'elaborazione digitale di segnali audio che hanno appunto come riferimento un valore centrale: lo zero analogico, o, nel caso che il segnale rappresenti una tensione, 0 Volt.
Poiché la spaziatura delle regioni di ampiezza è uniforme, la codifica PCM descritta è detta PCM lineare.
Il rapporto SNR, che vale per l'ampiezza massima del segnale, va corretto con un fattore S che lo fa decrescere con il diminuire dell'ampiezza. Un' espressione più accurata del SNR può essere:

$$SNR_{dB} = 6N + S$$

dove S è il fattore di correzione che vale 0 dB per l'ampiezza massima del segnale, ritornando così all'espressione originaria; vale -6 dB quando il segnale decresce a metà dell'ampiezza massima; -12 dB quando il segnale va a un quarto dell'ampiezza massima e così via. Quindi l' SNR diminuisce con l'ampiezza del segnale; nella codifica PCM lineare i segnali più deboli subiscono una degradazione di qualità rispetto ai segnali forti.

Politecnico di Torino

Per mantenere un livello elevato del rapporto SNR è sufficiente usare un numero elevato di bit, in modo che, anche se il rapporto decresce con il diminuire dell'ampiezza, esso continua a godere di valori adeguati alle ampiezze deboli. Per esempio, lo standard CD usa 16 bit per la quantizzazione, con un rapporto SNR di 96 dB. Se il livello di pressione sonora massimo SPL_{max} in riproduzione nel nostro impianto hi-fi è di 100 dB, allora i campioni con questo SPL_{max} godono del rapporto SNR di 96 dB. Con porzioni del segnale più deboli il rapporto SNR scenderà, ma sarà ancora accettabile sulla soglia minima di udibilità, all'incirca oltre i 60 dB. Questi valori permettono di mantenere elevato il rapporto SNR, pure con una quantizzazione lineare.

4.6 ECC – Codici Correzione Errore

Quando si deve trasmettere o memorizzare su supporti fisici un segnale PCM, per irrobustire la codifica si aggiungono dei bit extra che sono utili per il controllo di eventuali errori nella trasmissione. Un esempio molto semplice di controllo degli errori è un bit di parità: in questo caso si aggiunge un bit ad una sequenza fissata di bit, che assume i seguenti valori: se nella sequenza il numero di bit a 1 è pari, allora il bit verrà posto a 0; se nella sequenza il numero di bit a 1 è dispari, allora il bit verrà posto a 1. Nel caso in cui un bit per errore subisce un flip, cioè varia da 0 a 1 o viceversa, il bit di parità permette di scoprirlo e intervenire per correggere l'errore.
Ad esempio, supponiamo che ci siano 5 bit a 1 nella sequenza. Il bit di parità viene posto a 1. Se uno qualsiasi dei bit cambia polarizzazione, da 1 a 0 e viceversa, allora si ha un numero pari di 1 e si ha un'inconsistenza tra la sequenza di bit e il bit di parità. Ovviamente non si sa dove è occorso l'errore, ma il sistema può richiedere al dispositivo di interfaccia una nuova lettura del dato o una ritrasmissione. Il problema con il bit di parità è l'impossibilità di rilevare i casi in cui si verifichino due errori o in generale quando questi sono in numero pari; inoltre, il bit di parità non distingue tra uno, tre, cinque o n+1 errori (con n numero pari), cioè quando gli errori sono in numero dispari. Aumentando il numero di bit di controllo è possibile rilevare un numero maggiore di errori e in alcuni casi di correggerli.
Si consideri l'esempio di quattro bit di parità su un campione di 8 bit. Il primo bit di parità codifica i primi quattro bit, il secondo bit di parità codifica gli ultimi quattro bit, il terzo bit di parità controlla i bit 1, 2, 5, 6 e il quarto bit controlla i bit 2, 3, 6, 7. Un esempio è il seguente:

1 1 0 1 0 0 0 0 (1 0 0 1)

In parentesi sono indicati i bit di parità: il primo è a 1 (ci sono tre 1); il secondo è a 0 (non ci sono 1); il terzo è a 0 (ci sono due 1); il quarto è a 1 (c'è un solo 1). Supponiamo che ci sia un errore nel quarto bit. La nuova configurazione sarà:

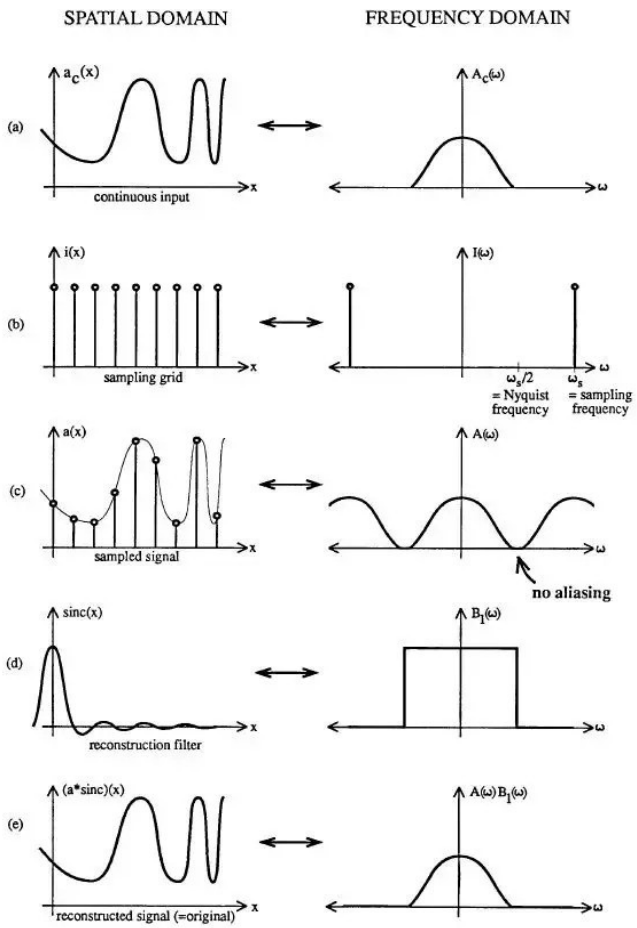
Politecnico di Torino

1 1 0 0 0 0 0 0 (1 0 0 1)

Viene sottolineato il quarto bit per indicare l'avvenuto flip.
Si analizzi la nuova situazione, assumendo che ci sia un solo errore e sapendo che ne possiamo scoprire solo uno: il primo bit di parità non è più corretto (quindi si ha un errore nell'insieme di bit [1, 2, 3, 4]); il secondo bit va ancora bene (quindi l'errore non è nell'insieme [5, 6, 7, 8], ma questo era ovvio dal dato precedente); il terzo bit va ancora bene (quindi l'errore non è nell'insieme [1, 2, 5, 6]), per cui l'insieme con un errore si riduce a [3,4]; il quarto bit va ancora bene, e quindi si esclude il bit 3. Di conseguenza l'errore è nel bit 4, che deve essere flippato a 1 per correggere l'errore.
Le configurazioni con i bit di parità possono essere anche più complesse; infatti aggiungendo 5 bit di parità a un campione di 16 bit e con una codifica appropriata, si possono rilevare errori occorsi su un solo bit. Aggiungendo ulteriori bit, si riescono a migliorare le prestazioni di rilevamento e correzione degli errori: questi codici binari aggiuntivi sono detti codici di correzione degli errori (Error Correcting Codes – ECC) e sono presenti in tutte le forme di archiviazione digitale.

4.7 Sample Rate

Le tecniche illustrate nei paragrafi precedenti sono considerate standard nella rappresentazione e conversione in digitale del segnale. Tuttavia dietro un'apparente semplicità si nascondono metodi e approcci molto complessi che verranno illustrati nei prossimi paragrafi.
Il processo di campionamento si può vedere come il prodotto di due segnali: il primo segnale è la forma d'onda da campionare (a), che modula in ampiezza il secondo segnale, un treno di impulsi di periodo e ampiezza costanti (b). Guardando a cosa succede nel dominio della frequenza, si ha che lo spettro del treno di impulsi è dato da una serie infinita di repliche del sample rate f_c.

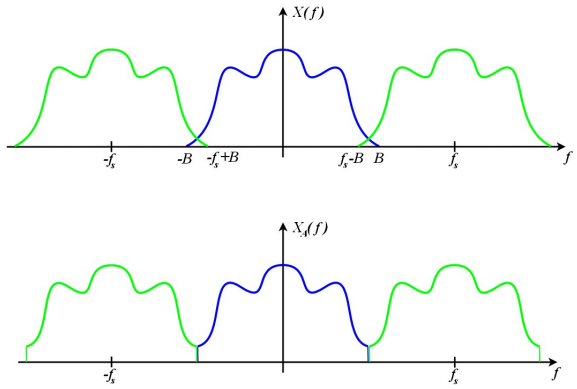


Processo di campionamento senza aliasing

Lo spettro della forma d'onda da campionare occuperà una certa lunghezza di banda che ha il massimo nella frequenza f_{max}; lo spettro del segnale campionato sarà replicato per ogni replica della frequenza di campionamento f_c (c). Quando il segnale verrà convertito in analogico, il passaggio attraverso il filtro passa-basso assicurerà che tali bande in eccesso vengano escluse, in modo da fare emergere solo la banda del segnale originale (d). Questa situazione ideale si ha quando si rispetta il teorema di Nyquist: ad esempio se il treno di impulsi ha un tasso di 48 kHz e il segnale da campionare ha una frequenza massima di 20 kHz; le bande laterali sulle armoniche di f_c non si sovrappongono alla banda originale ed il taglio operato dal filtro è sufficiente a garantire la ricostruzione fedele del segnale.

Politecnico di Torino

Nel caso della seconda figura invece il sample rate non rispetta il teorema di Nyquist e le bande di frequenza laterali alla prima armonica si sovrappongono allo spettro del segnale originale e si ha aliasing. Nell'esempio precedente, se il tasso di campionamento scendesse a 32 kHz, si avrebbe una sovrapposizione nella banda compresa tra 12 kHz e 20 kHz e le frequenze ripiegate provocherebbero problematiche udibili e fastidiose. I diagrammi in figura assumono inoltre che il filtro passa-basso abbia un comportamento ideale, tale cioè da lasciare intatte tutte le frequenze sotto la frequenza di taglio e azzerare l'ampiezza delle frequenze superiori ad essa, con una pendenza infinita alla frequenza di taglio. In realtà i filtri reali non possono avere una pendenza di taglio verticale. I filtri hanno tutti quindi una pendenza finita che inizia al bordo della banda richiesta. Di conseguenza il tasso di campionamento deve essere fissato ad un tasso superiore al doppio della massima frequenza presente nel segnale, in modo da rendere innocua la porzione di segnale che non viene eliminata dal filtro passa-basso.



Processo di campionamento con aliasing e relativo errore di ricostruzione

Per questi motivi il tasso di campionamento selezionato agli albori dell'audio digitale fu di circa 44 kHz. In tal modo, i filtri garantiscono un'attenuazione sufficiente anche per le frequenze intorno ai 20 kHz, impedendo una sovrapposizione delle bande di frequenza.
Inoltre, poiché il numero di dati da immagazzinare era di circa 1 MB per canale (44 kHz * 16 bit = 704 kbit/s), si aveva bisogno di supporti che riuscissero a memorizzare i dati con capacità a velocità sufficiente. Ad esempio, se si usa un nastro magnetico, occorre che il nastro in movimento, che è dotato di una certa velocità determinata dal sistema di trascinamento, faccia passare sotto la testina di registrazione una porzione sufficiente di tape per la registrazione, ovvero che la lunghezza d'onda registrata sul nastro sia superiore alla dimensione del traferro

della testina, condizione difficilmente ottenibile con testine stazionarie, ma usuale nei registratori analogici a testine rotanti per il video. Il tasso di campionamento di 44,1 kHz fu il risultato dell'uso dei registratori video per le prime registrazioni: solo in questo caso infatti la testina mobile dei videoregistratori e la scansione elicoidale del nastro riuscivano a garantire le condizioni per la memorizzazione. Considerato il formato di registrazione del video strutturato per "fotogrammi", cioè diviso in campi e linee, occorreva inserire un numero intero di campioni per linea. Il tasso di 44,1 kHz rispondeva alla necessità di garantire che un numero intero di campioni entrasse in una linea del segnale video nei due standard principali: nel sistema PAL (Phase Alternating Line), diffuso in Europa, si hanno 625 linee che vengono tracciate 25 volte al secondo (25 kHz), per cui il tasso di linee al secondo è di 15.625; nel sistema NTSC, diffuso negli USA, si hanno 525 linee che vengono tracciate 30 volte al secondo, con tassi di 15.750.
Nel sistema PAL, delle 625 linee, 588 sono linee attive e 37 sono linee vuote, per cui la frazione 588/625 è quella che viene utilizzata; nel sistema NTSC la frazione è 490/525. Si può osservare che si arriva allo stesso numero di campioni memorizzati in un secondo se si inseriscono 3 sample per linea. Infatti:

$$f_{PAL} = 588/625 * 15.625 \text{ linee/s} * 3 \text{ sample/linea} = 44,1 \text{ kHz}$$

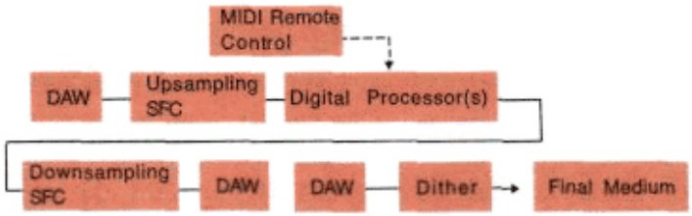
$$f_{NTSC} = 490/525 * 15.750 \text{ linee/s} * 3 \text{ sample/linea} = 44,1 \text{ kHz}$$

Un altro tasso di campionamento molto diffuso fin dall'inizio fu ottenuto raddoppiando il numero di linee in un frame video. Infatti 32.000 kHz, sebbene non siano utili per tutta la banda percepibile dell'audio, riescono a rappresentare i 15 kHz comunemente utilizzati per la trasmissione FM. Inoltre, poiché la conversione da 44,1 kHz a 32 kHz non produce un rapporto semplice (44,1 kHz / 32 kHz = 1,378125), si adottò il tasso di campionamento di 48 kHz per le applicazioni audio professionali (48 kHz = 1,5*32 kHz).

4.7.1 Oversampling

Come visto in precedenza, la conversione AD tradizionale è limitata dal fatto che è impossibile costruire filtri passa-basso che applichino in maniera netta il filtraggio delle frequenze superiori alla metà del sample rate. La soluzione di continuità tra la banda passante, che dovrebbe essere la banda base del segnale, e la banda rigettata, che dovrebbe essere tutto ciò che si trova oltre alla frequenza di taglio, dovrebbe essere infinitesima, cioè la pendenza dovrebbe essere infinita. Ma i filtri effettivi presentano piuttosto una graduale attenuazione delle frequenze che si estende nell'intorno della frequenza di taglio: con un tasso di 44,1 kHz, la frequenza di taglio è fissata a 22,050 kHz, che è molto vicina al limite superiore di udibilità. Ciò vuol dire che parte delle

frequenze attenuate dal filtro hanno ancora un'intensità rilevante e rischiano di provocare aliasing; un modo per aggirare questo problema è il sovracampionamento, detto anche oversampling.
L'oversampling consiste nell'utilizzare una velocità di campionamento molto più elevata di quella stabilita dal teorema di Nyquist permettendo ai filtri di utilizzare una pendenza meno inclinata vista la distanza tra la banda del segnale e le bande laterali. L'oversampling porta anche un migliore sfruttamento delle capacità di rappresentazione del digitale.
Si ha infatti che la banda del segnale digitalizzato è talmente ampia che il contributo spettrale del rumore è per la maggior parte sulle frequenze altissime, ben distanti dalla banda udibile e quindi dalla banda iniziale del segnale originale. Nei convertitori analogico-digitale vengono utilizzati sia altissimi tassi di campionamento e sia codifiche con parole binarie standard mediante un processo successivo di ri-quantizzazione. Oltre alla ri-quantizzazione si ha un'operazione di decimazione, dove si selezionano solo alcuni dei campioni: in particolare, se il sovracampionamento è di un fattore n oltre il tasso di Nyquist, si selezionerà un campione su n. Ad esempio, se l'oversampling è di 176,4 kHz, cioè 4* 44,1 kHz, si selezionerà un campione su quattro.
Se si aumenta la velocità di conversione, la potenza totale del rumore rimane la stessa, ma esso è distribuito su una larghezza di banda maggiore, secondo quanto previsto dalla teoria dell'informazione. Ad esempio, se il tasso di campionamento raddoppia, la potenza del rumore di quantizzazione si allarga fino a 44,1 kHz. Tuttavia, il segnale utile è sempre rimasto nella banda fino a 22,050 kHz e quindi in questa banda la potenza del rumore si è dimezzata. Il filtro passa-basso digitale con una frequenza di taglio a 22,050 kHz eliminerà la metà del rumore di quantizzazione dalla banda base del segnale, aumentando il rapporto SNR. Iterando il processo ogni volta che si raddoppia il sample rate, il rumore di quantizzazione nella banda base si abbassa di 3 dB: a questo punto con un sovracampionamento di 4 volte il rumore di quantizzazione sarà inferiore di 6 dB rispetto allo stesso convertitore senza sovracampionamento. Ricordando che un bit di quantizzazione aumenta l'SNR di 6 dB, si ha che un convertitore a 15 bit con oversampling 4x avrà le stesse prestazioni nella banda udibile di un convertitore a 16 bit e 44,1 kHz.
In generale, si ha che ogni volta che si incrementa il sample rate di un fattore 4 si può eliminare un bit di quantizzazione, mantenendo lo stesso livello di SNR. Teoricamente, sovracampionando a 4^{15}x possiamo disfarci di 15 su 16 bit e usare un convertitore ad un solo bit. Il sample rate equivarrebbe a circa $50*10^{12}$ Hertz (Tera Hertz).
Vi sono sul mercato dei convertitori A/D che possono operare fino a frequenze di 192 kHz. Il fatto di utilizzare delle frequenze maggiori di campionamento permette di migliorare la definizione della traccia audio ed aumentare il bit depth equivale ad avere più headroom, cioè più gamma dinamica in fase di mixing. Questi parametri permettono di creare una traccia audio con maggiore qualità che, quando viene processata con equalizzatori ed effettistica, produrrà dei risultati più accurati ed è per questo che alcune DAW permettono di utilizzare i plug-in in modalità “up sampling”, cioè con frequenza di campionamento maggiore.



Catena del segnale completamente digitale dove viene utilizzato il sovracampionamento in fase di elaborazione
fonte: Mastering Audio: the Art and the Science

Frequenze di sampling più alte sono utili anche quando si lavora con dei suoni che devono essere modificati nel dominio del tempo o nel dominio della frequenza. Bisogna infine considerare il medium digitale su cui si andrà a masterizzare e incidere il nostro audio; per cui se si andrà a incidere su DVD, che utilizza la frequenza di campionamento di 192 kHz, sarà necessario tenere la frequenza di campionamento durante la fase di lavorazione ad un valore maggiore o uguale.
In vari test condotti da Bob Katz nel 1996 si dimostrò che, utilizzando dell'attrezzatura ad alto sample rate, i soggetti non riuscivano a distinguere due registrazioni identiche ma rispettivamente filtrate intorno ai 20 kHz e ai 40 kHz, dimostrando quindi che la necessità di campionare anche le frequenze al di fuori della banda dell'udibile con il mero scopo di migliorare l'ascolto non fosse necessario. I maggiori artefatti sonori si creano nell'utilizzo di convertitori A/D e D/A dotati di filtri passa basso e passa alto di bassa qualità, i quali introducono distorsione di fase o introducono degli artefatti che il nostro orecchio percepisce come perdita di chiarezza nei transienti; quindi l'utilizzo dell'upsampling si rivela necessario per “rilassare” i filtri e spingere gli artefatti nel range del non udibile.
Per ridurre il tasso di campionamento a un livello di praticità per i convertitori con quantizzazione molto bassa si può lavorare con la tecnica della sagomazione del rumore (noise shaping).

4.8 Noise shaping

Il noise shaping è una tecnica di feedback in cui il processo di quantizzazione del campione corrente è influenzato dall'errore di quantizzazione del campione precedente attraverso un circuito di feedback. Supponiamo di avere un campione di x bit e la possibilità di memorizzare solo y bit. Si acquisiscono gli y bit di ordine superiore, cioè i più significativi, lasciando la differenza dei bit (x – y), che costituiscono l'errore di quantizzazione. Questi bit rimanenti vengono riemessi in un circuito di feedback con il calcolo del campione successivo; quindi la reazione del sistema all'errore

Politecnico
di Torino

avviene con il ritardo di un campione. Se l'input fosse costante, si avrebbe che l'errore di quantizzazione medio andrebbe a zero dopo un certo numero di campioni.
Consideriamo ad esempio la sequenza di campioni di quattro bit:

0001 0001 0001 0001 0001 0001 0001 0001

Supponiamo di considerare solo i due bit più significativi e rimuovere gli altri due come errore di quantizzazione:

Sequenza originale	0001	0001	0001	0001	0001	0001	0001	0001
Errore quantizzazione	+00	+01	+10	+11	+00	+01	+10	+11
Somma	0001	0010	0011	0100	0001	0010	0011	0100
Output a due bit	00	00	00	01	00	00	00	01

Nella prima riga abbiamo la sequenza originale di campioni, nella seconda l'errore di quantizzazione del campione precedente, cioè i due bit meno significativi; nella terza la somma tra il campione attuale e l'errore precedente, cioè la somma tra la prima e la seconda riga; nella quarta l'output effettivo dei due bit più significativi. Come si nota, dopo un certo numero di passi si ha un ritorno a 0 dell'errore e si è mantenuto il livello di quantizzazione nonostante la rimozione di due bit. Ovviamente, man mano che un segnale contiene variazioni più veloci, il feedback dell'errore non è così efficace e l'effetto di ritardo prende il sopravvento. Con questo trattamento l'errore di quantizzazione viene ad assumere uno spettro che non è uniforme ma aumenta nelle vicinanze delle alte frequenze con una forma sinusoidale (da qui il nome di noise shaping) con un picco sulla frequenza di Nyquist, ad un livello determinato dalla dimensione della regione di quantizzazione. Quando è accoppiato al sovracampionamento, il noise shaping sposta il picco sempre più verso le altissime frequenze, fuori dalla banda dell'udibile.

4.9 Modulazione Delta-Sigma

Vi sono molti metodi effettivi per applicare le tecniche di oversampling e noise shaping nella conversione D/A e A/D. Le tecniche si possono classificare mediante due parametri: uno riguarda il numero di bit della codifica, dal momento che si hanno conversioni ad uno o a più bit rispettivamente. Nella rappresentazione PCM standard si hanno due codifiche che si differenziano per l'uso di bit multipli o di un solo bit: nel primo caso si parla di Sigma DPCM e nel secondo caso si parla di Modulazione Sigma-Delta. Quest'ultimo metodo di conversione, che usa un bit di rappresentazione del campione, implementa il noise shaping e l'oversampling. In questo caso il DAC ha veramente un solo bit di quantizzazione e il rapporto della densità del segno dell'impulso,

Politecnico di Torino

che può essere 0 o 1, è correlata ai dati originali a 16 bit; quindi viene codificato non il segnale ma i suoi cambiamenti, i cosiddetti delta, creando una sequenza di impulsi invece che una sequenza di numeri. Data questa sequenza si può ricostruire il segnale originario senza l'utilizzo di un codice multi-bit come quello PCM e si evita il relativo errore di quantizzazione. Questa tecnica viene utilizzata dal Super Audio CD, che memorizza i dati uscenti dal modulatore delta-sigma direttamente sul supporto fisico.

4.10 Quantizzazione non lineare

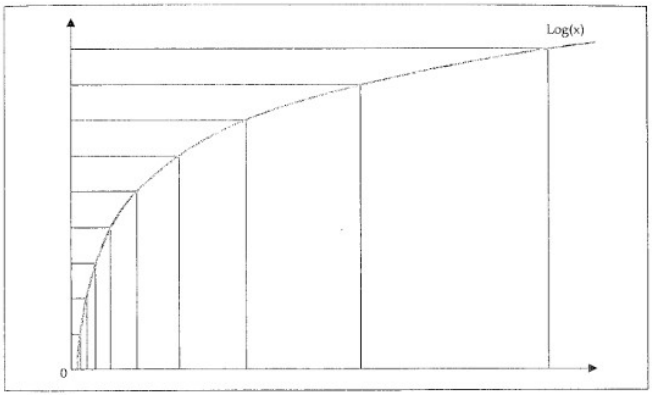
Una soluzione alternativa per evitare la degradazione dei segnali deboli è quella di dare una spaziatura non uniforme alle regioni di quantizzazione attraverso una quantizzazione non lineare. Benché il numero delle regioni continui a rimanere lo stesso (in quando dipende dal numero di bit), la dimensione delle regioni di quantizzazione diminuisce per le regioni delle ampiezze deboli e aumenta per le regioni delle ampiezze forti. Di conseguenza l'errore di quantizzazione decresce per le ampiezze deboli e aumenta per le ampiezze forti: il guadagno in qualità alle ampiezze deboli è compensato da una perdita di qualità alle ampiezze forti.
Il primo effetto della quantizzazione non lineare è la separazione tra il rapporto SNR e la gamma dinamica. Per la quantizzazione non lineare il rapporto SNR può assumere forme diverse, in quanto l'ampiezza dell'errore varia con l'ampiezza del segnale e con il tipo di spaziatura scelto. L'idea è essenzialmente quella di ridurre il rumore di quantizzazione alle ampiezze deboli, permettendo invece che aumenti alle ampiezze forti, dove comunque viene mascherato. Questo risultato viene raggiunto mediante una spaziatura non uniforme delle regioni di quantizzazione: le regioni vicine all'ampiezza nulla vengono quantizzate in modo molto più fine rispetto alle regioni alle ampiezze elevate. Quantizzando in modo non lineare, e quindi spaziando in modo non uniforme le regioni di quantizzazione, il rapporto SNR viene a dipendere dall'errore variabile nelle varie regioni, mentre la quantità di gamma dinamica resta invariata.
Di seguito vengono illustrate brevemente due tipi di quantizzazione non lineare, quelle logaritmica e a virgola mobile, che ci permettono di comprendere come avviene la spaziatura non uniforme degli intervalli con modifica del rapporto SNR e l'espansione della gamma dinamica.

Politecnico di Torino

Quantizzazione lineare a) e quantizzazione non lineare b)
fonte: Audio e Multimedia

4.11 Quantizzazione logaritmica

La quantizzazione logaritmica si ritrova in alcuni comuni formati audio.
La caratteristica fondamentale è che la dimensione delle regioni di quantizzazione cresce con l'ampiezza del segnale secondo una curva logaritmica.



Curva logaritmica e regioni di quantizzazione

Politecnico di Torino

fonte: Audio e Multimedia

Nella figura soprastante viene illustrata una curva logaritmica e come essa opera sulla dimensione delle regioni di quantizzazione. Sull'asse delle ordinate sono state indicate le regioni di quantizzazione uniformi secondo una scala logaritmica; sull'asse delle ascisse sono invece indicate le corrispondenti regioni di quantizzazione sulla scala lineare: come si può notare le regioni di quantizzazione vicine allo 0 sono molto più ravvicinate delle regioni di quantizzazione distanti dall'origine. Quindi, quantizzare in modo logaritmico vuol dire assegnare le regioni di quantizzazione in modo uniforme rispetto ad una scala logaritmica, risultando in una non uniformità sulla scala lineare. I codici digitali vengono poi assegnati ai campioni secondo tale suddivisione.
Il rapporto SNR è migliore alle ampiezze deboli rispetto alla corrispondente quantizzazione lineare con lo stesso numero di bit; tale miglioramento viene pagato con un peggiore SNR alle ampiezze elevate, ma i fenomeni di mascheramento permettono che questa situazione risulti tollerabile.
Nonostante questi vantaggi la quantizzazione di tipo logaritmico non è molto efficiente quando il segnale quantizzato deve essere in seguito sottoposto a elaborazione : ad esempio, la somma di due valori su una scala logaritmica corrisponde al logaritmo del prodotto dei due valori e il prodotto è un'operazione molto più costosa in termini di tempo di elaborazione. Quando l'elaborazione diventa molto complessa, anche il minimo risparmio su una qualsiasi delle operazioni diventa molto importante. La soluzione adottata in genere in questi casi è di convertire il segnale digitale in forma lineare e applicare quindi le operazioni di elaborazione del segnale. É consigliabile adottare la forma logaritmica in applicazioni in cui non ci si può permettere un numero elevato di bit di quantizzazione, soprattutto nella fase finale di archiviazione di un segnale quando non verrà più elaborato. Oltre al miglioramento del rapporto SNR alle ampiezze deboli, una quantizzazione logaritmica permette di migliorare notevolmente la gamma dinamica. Empiricamente si può verificare che una quantizzazione logaritmica a 8 bit produce una gamma dinamica pari a una quantizzazione lineare a 13 o 14 bit.
Per comprendere più nel dettaglio come avviene il miglioramento della gamma dinamica, viene introdotto il secondo tipo di quantizzazione non lineare, la quantizzazione con virgola mobile.

4.12 Quantizzazione in virgola mobile

La quantizzazione con virgola mobile si ottiene aggiungendo dei bit extra a quelli presenti nella quantizzazione lineare; mantenendo il rapporto SNR, la gamma dinamica viene estesa proporzionalmente alla somma del numero di bit di quantizzazione e dei bit extra.

Politecnico di Torino

Si consideri un sistema di quantizzazione lineare a 8 bit, con un rapporto SNR e una gamma dinamica di circa 48 dB. Il segnale verrà quantizzato in 256 regioni uniformi di quantizzazione, con codifiche che vanno da 00000000 a 11111111.
Aggiungiamo ora un bit extra al sistema lineare, e convertiamo il valore a 8 bit in un valore a 9 bit con la seguente regola: se il bit extra è a 0, gli 8 bit sono i bit meno significativi dei 9 totali e il primo bit, quello più significativo, diventa uno 0; se il bit extra è a 1, gli 8 bit sono i bit più significativi, cioè più a sinistra dei 9 e l'ultimo bit, quello meno significativo, diventa uno 0.
In poche parole, il bit extra funziona da traslatore (o shift) dei bit lineari (0 = no shift, 1 = shift di 1).

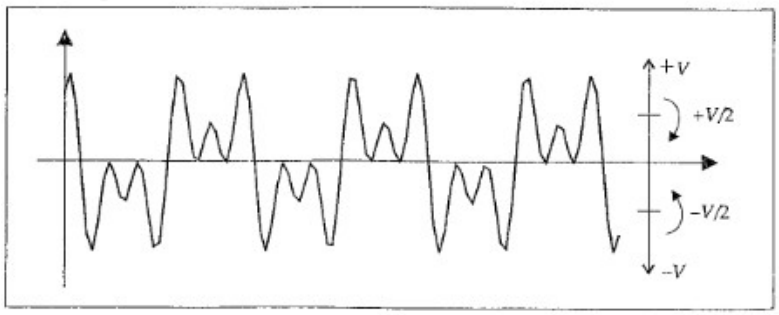
L'esempio di seguito mette in pratica ciò che è stato appena descritto:

00001101 (bit lineari) + 0 (bit extra) → 000001101 (9 bit)
(traslazione di 0 posti)

00001101 (bit lineari) + 1 (bit extra) → 000011010 (9 bit)
(traslazione di 1 posto)

Si supponga ora che il valore del bit extra venga deciso in base all'ampiezza del segnale: sapendo che l'ampiezza massima del segnale è 2V volt, cioè il segnale varia tra ±V volt, la gamma delle ampiezze viene divisa in due parti uguali, con valori di limite superiore e inferiore di ±V/2. La quantizzazione opera nel seguente modo: se l'ampiezza del campione è dentro l'intervallo, la quantizzazione avviene in modo lineare sui bit e il bit extra viene posto a 0, per cui non ci sarà nessun shift; se l'ampiezza del campione è fuori dall'intervallo, essa viene divisa per 2, cioè ricondotta nell'intervallo e quantizzata su 8 bit lineari e il bit extra viene posto a 1; avviene quindi una traslazione di un posto. I campioni a 8 bit sono diventati a 9 bit. Il nuovo sistema a 9 bit quantizza i campioni a 8 bit (cioè su 256 regioni) ma cambia poi la codifica con la traslazione. In particolare, i campioni delle ampiezze dentro l'intervallo, cioè delle ampiezze deboli, avranno la stessa codifica prevista dal sistema a 8 bit anche nel sistema a 9 bit derivato, con l'aggiunta dello 0 iniziale; al contrario, i campioni delle ampiezze fuori dall'intervallo, cioè delle ampiezze forti, avranno una nuova codifica ottenuta traslando i bit di un posto a sinistra e inserendo lo 0 in fondo. La differenza subito evidente rispetto a un sistema lineare a 9 bit è che nel nuovo sistema non sarà mai possibile assegnare delle codifiche che iniziano e finiscono con 1, in quanto il modo di ottenere i campioni da 8 bit è solo mediante traslazione e concatenazione dello 0. E' come lavorare con un sistema a 9 bit in cui la spaziatura delle regioni alle ampiezze elevate è doppia rispetto alle ampiezze deboli, in quanto mancano tutti i codici che finiscono con 1. Dato che la quantizzazione opera ancora su 8 bit il rapporto SNR non è variato ma i segnali rappresentabili nel sistema si estendono sui 9 bit, per cui la gamma dinamica risulta estesa: in questo caso è di 54 dB

invece di 48 dB. Ovviamente con una quantizzazione lineare a 9 bit si migliorerebbe sia il rapporto SNR che la gamma dinamica.
Il sistema può aumentare la sua non linearità e la sua gamma dinamica incrementando il numero di bit extra, che sono anche detti bit esponente, da cui il nome di virgola mobile per questo tipo di rappresentazione. Infatti, se i bit extra diventano 2 i possibili shift diventano 4, con 3 bit extra gli shift diventano 8, e così via. In tal modo, la spaziatura non uniforme viene caratterizzata da 4più possibili regioni di quantizzazione. Se il sistema consiste di N bit di quantizzazione e P bit di traslazione, la gamma dinamica sarà 6N + 6P dB, invece dei 6N dB del sistema lineare a N bit.
In riproduzione, occorre amplificare i valori attenuati durante la conversione analogico-digitale. L'amplificazione deve essere di un fattore proporzionale alla loro attenuazione ed è determinata dal numero di bit extra. Si noti come il metodo assomigli alla compansion, anche se la prospettiva risulta invertita: nel caso del tape, nella fase di codifica, erano i segnali deboli a essere amplificati, e nella fase di decodifica, la gamma dinamica veniva estesa. Nel caso della quantizzazione in virgola mobile, invece, sono i segnali forti a essere attenuati nella fase di codifica e vengono di nuovo amplificati nella fase di decodifica.



Quantizzazione con virgola mobile
fonte: Audio e Multimedia

4.13 Ricostruzione del Segnale

La catena di riproduzione del segnale digitale inizia con la conversione in segnale analogico, mediante il convertitore digitale analogico (DAC).
Il DAC genera delle tensioni elettriche che sono proporzionali ai campioni del segnale digitale, in corrispondenza degli istanti di tempo stabiliti dal clock interno. Un interpolatore si occupa di smussare i gradini dovuti al salto da un valore di tensione al valore vicino. Il modo in cui

Politecnico di Torino

l'interpolatore svolge il suo compito necessita di nuovo di un filtro passa-basso: i gradini ripidi tra un valore e l'altro corrispondono alla presenza di frequenze alte nel segnale analogico generato; tali frequenze devono essere eliminate dal filtro passa basso, la cui frequenza di taglio è pari alla metà del sample rate del clock.
Supponendo di avere dei campioni x[nT] di un segnale continuo x(t), dove t è il tempo in secondi, n il numero di campioni e T il periodo di campionamento, con x(t) un segnale a banda limitata tra + $f_c/2$ e - $f_c/2$, dove $f_c = 1/T$ è il sample rate e $X(\omega)$ è la trasformata di Fourier di x(t) che vale:

$$X(\omega) = \int x(t)e^{-j\omega t}$$

con l'integrale calcolato nell'intervallo [-∞,+∞].
Assumendo che $X(\omega)$ sia uguale a zero per :

$$|\omega| \geq \pi f_s$$

allora il teorema di Shannon ci assicura che x(t) può essere ricostruito attraverso i sample di x[nT] attraverso la sommatoria compresa tra [-∞,+∞]:

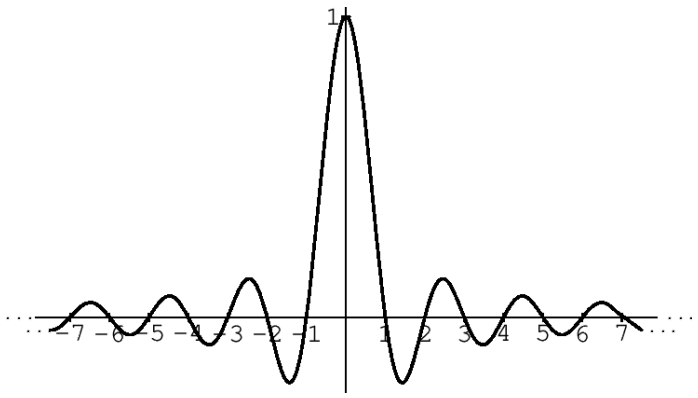
$$x'(t) = \sum x(nT)h_s(t - nT) = x(t)$$

dove :

$$h_s(t) = \sin(\pi f_s t)/(\pi f_s t) = \mathrm{sinc}(t)$$

Quindi, per ricampionare x(t) a un nuovo sample rate f ' = 1/T', bisogna solo ricalcolare l'equazione della ricostruzione con valori multipli di T', usando come finestra di ricalcolo una funzione sinc valutata per tutti i campioni del segnale digitale.
Nel grafico sottostante è raffigurata la funzione rispetto all'origine t=0:

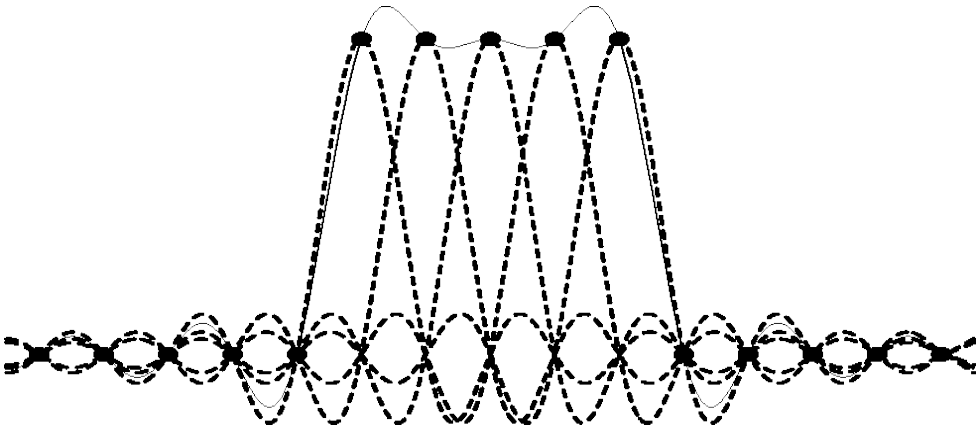
$$\mathrm{sinc}(t) = \sin(\pi t)/(\pi t)$$



Politecnico di Torino

Funzione sinc

La ricostruzione può essere reinterpretata come una convoluzione tra il segnale digitale in ingresso e la finestra data dai segnali sinc [$h_s(t)$] traslati e scalati per ogni campione, in cui ogni valore della funzione sinc è calcolata per ogni singolo sample e questi valori vengono poi sommati tra di loro.
Si noti come i passaggi allo zero della funzione sinc occorrono in tutti gli interi tranne che in t=0. Ciò vuol dire che al momento t=nT, cioè nell'istante del sample, l'unica contribuzione alla somma è il sample x[nT] e tutti gli altri sample contribuiscono con una funzione sinc che vale zero al momento t=nT. Di conseguenza, l'interpolazione riesce a ricalcolare i sample esistenti.
Il grafico sottostante indica come le funzioni sinc prese in considerazione vengano sommate per ricostruire il segnale in banda limitata. Esso mostra la sovrapposizione di 5 funzioni sinc, ognuna con la propria unità di ampiezza e dislocate in intervalli da un sample.
Infine il segnale viene amplificato e mandato ai diffusori per la traduzione in segnale acustico.

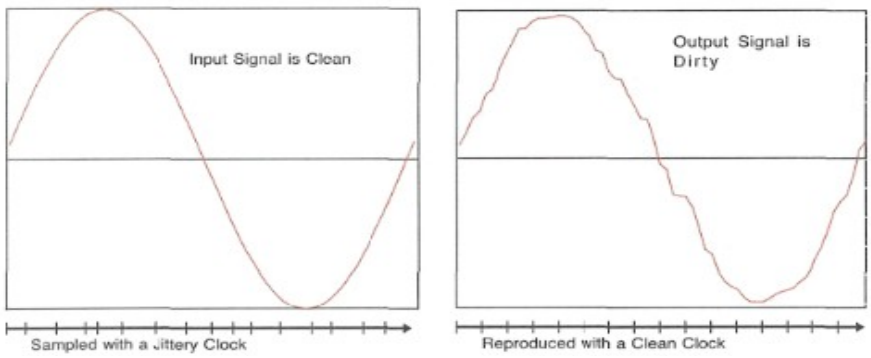


Ricostruzione del segnale a banda limitata. I punti indicano i sample del segnale, le linee tratteggiate mostrano i componenti della funzione sinc e la linea continua mostra la ricostruzione univoca dei sample ottenuta dalla somma delle componenti della funzione sinc.

4.14 Jitter

Uno dei fenomeni meno conosciuti ma importante nel mondo dell'audio digitale è il jitter.
Negli anni '80, poiché la maggior parte dei processori di A/D e di D/A era dotata di una bassa risoluzione, gli errori di jitter non erano nella lista delle priorità. Il troncamento, il dithering non eseguito correttamente e l'aliasing creavano degli artifizi sonori più problematici, ma ora che la

performance audio raggiunge frequentemente il livello dei 20 bit e oltre, il jitter è diventato importante nella conversione A/D-D/A. I sintomi del jitter sono percepibili nella fase di conversione: il suono risulta confuso, sfuocato e vi è un'apparente perdita di livello sonoro, che porta l'ascoltatore ad aumentare il livello dei monitor anche quando i segnali sono riprodotti con guadagno unitario; in più vi è una mancanza di profondità, che spesso è difficile da identificare anche per un orecchio critico.
Ciò è dovuto al fatto che l'audio digitale si basa sul campionamento di un segnale a intervalli regolari e, per mantenere questi intervalli costanti, vi è la necessità di un clock molto stabile. Se la frequenza del clock varia durante la conversione analogica-digitale, questa causerà una distorsione permanente nel segnale acquisito, poiché la forma d'onda originaria è stata campionata con dei valori d'ampiezza scorretti perchè rilevati istanti di tempo scorretti. La stessa cosa avviene con un segnale campionato correttamente ma riprodotto con un clock affetto da jitter.
Di conseguenza, un clock instabile provocherà distorsione durante la conversione sia in A/D che in D/A.

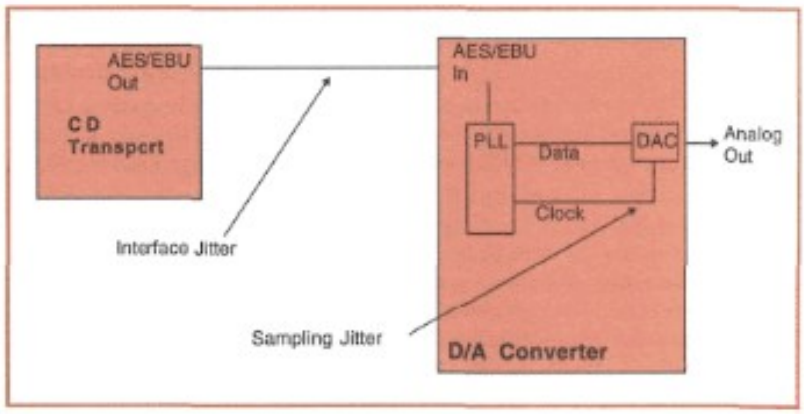


Jitter durante la conversione A/D che crea distorsione permanente
fonte: Mastering Audio: the Art and the Science

Il periodo di un clock con frequenza 44,1 kHz è di 22,7 µs; variazioni nel clock di questo ordine di grandezza e addirittura fino ai 10 picosecondi potrebbero causare degli artefatti udibili, dipendendo dalla lunghezza della parola binaria. Il jitter deve essere mantenuto ad un livello basso per mantenere la qualità sonora, poiché esso va ad influenzare il livello di rumore totale producendo delle bande di frequenze che vanno poi a mascherare i dettagli fini di una registrazione e possono far degradare un convertitore perdendo anche parecchi bit di risoluzione reale.

Politecnico di Torino

Un oscillatore quarzato standard è sufficiente ad un computer per processare i dati ma i convertitori audio richiedono un oscillatore che sia estremamente stabile. Per ottenere una performance da 20 bit a 44,1 kHz serve un oscillatore con una stabilità, ossia un jitter, che sia minore o uguale a 25 picosecondi, in termini di distanza temporale tra picco e picco. Per fare un esempio, un nanosecondo (1000 picosecondi) equivale alla frequenza di 1 Ghz, che inquinerebbe in maniera decisa qualunque frequenza di clock ed è per questo che il circuito di un convertitore deve essere schermato e isolato dai disturbi elettromagnetici, che possono provenire dall'alimentazione o dai circuiti logici di controllo.
Il jitter può essere trovato nella catena dell'audio digitale in due modalità differenti:
1) il jitter d'interfaccia, cioè il jitter prodotto nelle interconnessioni tra le apparecchiature audio
2) il jitter di campionamento, cioè il jitter nel clock che pilota la conversione analogico/ digitale o digitale/analogica.
Se un convertitore è dotato di un' ottima repulsione al jitter di campionamento allora il jitter d'interfaccia, anche se di livello elevato, non sarà rilevante. In altre parole, si può anche avere jitter prodotto nell'interfaccia o nel cavo di collegamento ma questo non influenzerà un convertitore progettato bene.



Struttura interna di una DAC
fonte: Mastering Audio: the Art and the Science

In figura viene mostrato come il clock di campionamento viene estratto dal PLL (Phase Locked Loop) all'interno del convertitore. Se un PLL è di livello superbo (cosa che accade raramente) nessuno degli artifizi provenienti dall'interfaccia affetta da jitter verrà trasmesso al clock di campionamento. Questa confusione tra le due tipologie di jitter ha generato una vera e propria sottocultura audiofila imperniata sulla scelta dei cavi di collegamento e la proliferazione di unità di

riduzione del jitter nel tentativo di ottenere una riproduzione migliore, con la speranza che una sostituzione dei cavi o l'uso di generatori di clock esterni “estremamente stabili” avrebbe eliminato il jitter.
In generale, cavi e generatori di clock esterni sono solo dei paliativi per i problemi di jitter, i quali possono essere risolti solo all'interno dei convertitori. Di conseguenza i generatori di clock esterni saranno sempre limitati nella loro efficacia poiché si può generare del jitter d'interfaccia tra l'output del generatore di clock e il convertitore.
Molte specifiche dei produttori nascondono questi artefatti poiché non si hanno ancora delle misurazioni standard per i loro effetti. Per esempio, certi convertitori A/D e D/A riportano un rapporto SNR maggiore di 120 dB, che teoricamente equivale ad una performance di 20 bit, ma ciò non corrisponde alla pratica. Infatti questi numeri vengono ottenuti con il metodo tradizionale per calcolare i rapporti SNR: prima viene misurato un segnale Full Scale, poi viene rimosso il segnale e viene misurato il rumore analogico residuo. Ciò non tiene conto del rumore di quantizzazione (o distorsione) che si genera solo quando il segnale è presente. Questo è il motivo per cui si percepiscono delle differenze sonore tra i convertitori nonostante sulla carta abbiano le stesse caratteristiche tecniche. La maggior parte della misurazioni citate nei manuali sono di conseguenza irrilevanti e la maggior parte dei convertitori non raggiunge una risoluzione reale di 20 o 24 bit.
Idealmente, il PLL del convertitore dovrebbe respingere completamente il jitter in ingresso con il suo filtro passa basso, ma se il PLL ha un'attenuazione del jitter inadeguata esso trasmetterà un po' del jitter d'interfaccia al clock di conversione. Uno dei più eminenti esempi di jitter è quello segnale-dipendente, causato dalle interfacce esterne come l'AES/EBU, un protocollo di trasmissione dell'audio basato sull'uso di un cavo bilanciato con connettore XLR, e il S/PDIF, un protocollo basato sul connettore RCA. Un convertitore il quale respinge con successo il jitter può suonare molto più pulito che un altro con un rumore di fondo assoluto più basso.
Il jitter può produrre diversi tipi di problematiche: effetti segnale-dipendenti (i quali apportano distorsione dovuta al battimento tra il sample rate e il segnale audio); effetti casuali (che fanno alzare il noise floor, rumore di quantizzazione) ed effetti di frequenza discreti (interferenza di altri dispositivi contenuti nella DAC che producono toni casuali o spuri).
Alcuni di questi effetti sono meno fastidiosi da ascoltare rispetto ad altri; per questo motivo è difficile dare un' interpretazione numerica univoca al jitter.

4.14.1 Jitter nei Processori Audio

Durante la conversione l'ammontare di jitter è determinato dai seguenti meccanismi: la qualità del motore servo, la qualità del sistema di buffering e al circuito di clock che guida l'informazione.
I produttori differiscono ampiamente nella loro abilità di mantenere il clock sotto controllo; la stabilità del clock non è importante per le tecnologie utilizzate dal PC, le quali vengono riadattate

per l'audio digitale. Infatti le interfacce standard degli hard disk sono asincrone, cioè senza clock, e quindi hanno un output completamente irregolare. Il jitter equivalente di un' interfaccia USB, che è un interfaccia di trasmissioni dati asincrona, è enorme e si ha un oscillazione tra lo stato in cui non viene trasmessa nessuna informazione ad uno stato in cui si hanno più streaming contemporanei. Quando queste interfacce asincrone vengono utilizzate è compito del circuito di ricezione fare in modo che l'informazione si conformi ad un clock regolare.
Il jitter di interfaccia non influenza l'informazione quando si rimane nel dominio digitale, quindi quando si lavora tutto in digitale e si utilizzano solo apparecchiature digitali. Dopo l'eventuale conversione A/D, l'informazione può passare da un processore all'altro e da un medium all'altro nonostante ci sia jitter, almeno finché esso sia abbastanza basso da permettere un trasferimento privo di errori. Similmente, il jitter del clock non ha effetti sulla performance della maggior parte degli equalizzatori digitali out of the box, come limiter e compressori, i quali sono praticamente macchine a stati.
Una macchina a stati è definita come un processore il quale produce un output identico per lo stesso input di dati e non considera il tempo o la velocità dell'informazione ma solo lo stato precedente dell'informazione. Con una macchina a stati si può avere un clock completamente irregolare o anche rallentarlo fino a un sample per secondo e il processore riuscirebbe a produrre tutte le informazioni in maniera corretta. Quando queste informazioni verranno riprodotte alla velocità corretta e con un clock regolare, non si avranno errori.
Tutti i processori professionali che fanno uso dell' oversampling, come equalizzatori e compressori, sono macchine a stati. Essi usano dei convertitori sincroni per raddoppiare il sample rate interno.
Con i processori digitali di pitch, come l'Autotune, la spiegazione è un po' più complessa, ma esso non è affetto da jitter. I processori di pitch non sono delle macchine a stati: ciò è dovuto al loro algoritmo casuale, dove vengono prodotti degli output differenti dallo stesso frammento musicale ogni volta che viene utilizzato. Ma essi prendono in considerazione i sample in arrivo, uno alla volta, indipendentemente dalla regolarità del clock che alimenta il processore: questi repitchers infatti possono funzionare offline in una DAW, a qualunque velocità, senza aver bisogno di un clock. In altre parole, la maggior parte dei processori digitali sono completamente immuni al jitter.
Il jitter di solito diventa significativo in un mix digitale solo durante la riproduzione, quando l'informazione è trasmessa all'esterno ad un convertitore D/A. In questo caso il jitter influenza solo l'esperienza di ascolto singola e non ha nessuno effetto sul dato digitale archiviato.
Il dato rimane intatto; quindi bisogna preoccuparsi del messaggero, cioè il clock che pilota il DAC. Il jitter influenza l'informazione durante un mix digitale anche quando si lascia il dominio digitale e si usano processori analogici out of the box e di conseguenza devono essere utilizzati dei convertitori di qualità.
Qualche console digitale contiene un convertitore con sample rate asincroni detto ASRC (Asynchronous Sample Rate Converter). Questi convertitori usano filtri variabili basati su una stima

continua nel tempo del sample rate del dato in entrata e quindi sono sensibili al jitter del clock. Un ASRC non è una macchina a stati e riprodurrà un output differente ogni volta che viene utilizzato.
Per creare dell'audio di qualità bisognerebbe testare la qualità di ogni ASRC e tentare di alimentarle con un clock di qualità maggiore. Questa è una questione importante nelle console low cost, dove i clock sono spesso scadenti per motivi economici e specialmente per le console pilotate da un clock esterno globale (Master word clock), che può peggiorare il jitter di un PLL interno di qualità mediocre. Le console che contengono ASRC sono sconsigliabili per le applicazioni audiofile, a meno che possano essere completamente bypassate quando non necessarie anche se i chip ASRC al giorno d'oggi contengono algoritmi di riduzione del jitter sofisticati e hanno una relativa bassa distorsione, quindi la performance della console è solo leggermente degradata. L'effetto udibile è un leggero smorzamento o diminuzione della stabilità dell'immagine stereo del 10% rispetto alla qualità sonora originaria.
É meglio utilizzare il registratore multitrack e i convertitori D/A con un clock interno. In teoria il convertitore interno dovrebbe maneggiare ogni buon generatore di clock o cavo dell'interfaccia senza influenzare il suono. Ma bisogna notare che è molto più facile ed economico progettare un oscillatore clock stabile che un PLL, il quale deve performare il doppio compito di oscillatore e di barriera per il jitter. Di conseguenza, ogni convertitore o registratore multitraccia performa al meglio con il clock interno e in un convertitore la performance con il un clock esterno può al massimo essere comparabile con l'oscillatore incorporato. Se un convertitore performa meglio con il clock esterno, allora questa deve essere vista come una critica alla qualità del clock. Ciò nonostante, i wordclock out-of-the-box sono necessari per sincronizzare fra loro varie apparecchiature digitali indipendenti.
Riassumendo, ci sono due maniere per impostare il clock ad un convertitore:

1) attraverso il Sync Interno, dove un oscillatore clock stabile collocato all'interno del convertitore pilota direttamente il circuito.

2) attraverso il Sync Esterno, il quale di solito richiede un PLL che, come detto in precedenza, è un circuito critico spesso causa dei problemi di jitter e degli artefatti dovuti alla conversione. Il PLL deve filtrare il jitter causato dai clock della fonte originaria e dalle interferenze tra i cavi che poi vanno a sommarsi al clock.

Esempi di Sync Esterno:

i) Sincronizzazione con lo standard AES/EBU, il quale è predisposto al jitter segnale-dipendente cioè il collegamento AES/EBU riprodurrà un clock più pulito in assenza del segnale rispetto a quando il segnale è presente. Ciò dipende dal PLL, che se di buona qualità non trasmetterà il jitter.

ii) Sincronizzazione con un generatore di wordclock, il quale può portare jitter estremamente basso perché il carico sul PLL è minore. Ciò nonostante solo pochi convertitori hanno una degradazione inudibile. Quindi quando si lavora con attrezzatura che richiede una sincronizzazione esterna, come il video, bisogna calcolare bene la scelta del convertitore se si vuole ottenere una performance di livello audiofilo.

iii) Altre interfacce. L'industria del computer sta reinventando continuamente il mercato e l'industria dell'audio è pronta ad adottare l'ultima novità, in questo caso un' interfaccia per computer affetta da molto jitter conosciuta come Firewire o MLAN, che dovrebbe essere utilizzato solo per trasmettere i dati, non il clock.

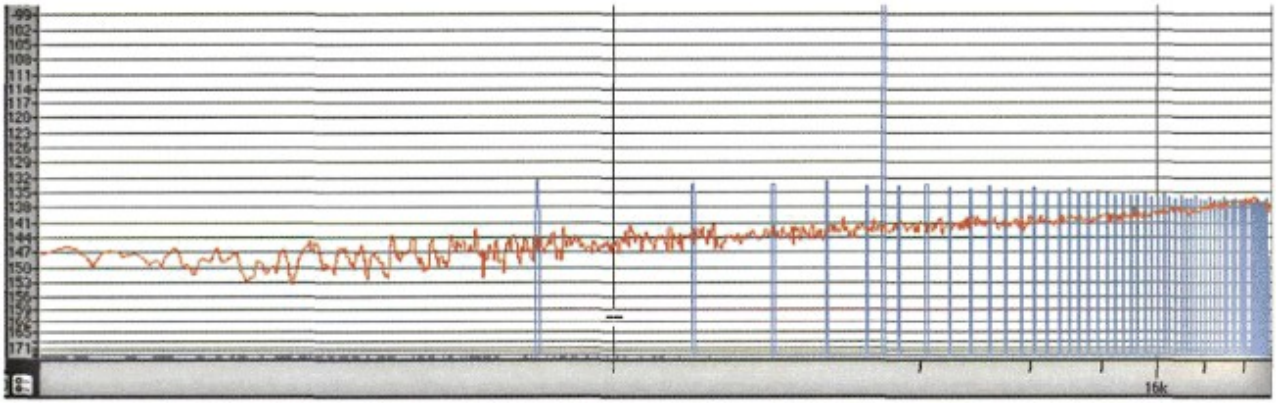
In conclusione, bisognerebbe utilizzare sempre A/D progettati bene con sync interno per la migliore performance a meno che non si stia lavorando con il video, con il quale gli A/D devono essere per forza clockati esternamente. L'A/D dovrebbe essere il master clock in qualunque sistema di registrazione e il D/A il master quando si sta riproducendo, ricordandosi che il jitter nell' A/D si traduce in distorsione la quale non potrà essere più rimossa.
Per costruire un convertitore immune al jitter serve tempo, ricerca e l'implementazione di una progettazione raffinata. Gli ingegneri che producono i convertitori hanno speso un anno-uomo solo sul PLL e un anno aggiuntivo sui dettagli del convertitore. I produttori di convertitori di successo devono essere abilissimi nelle tecniche di layout delle schede del PC, dei circuiti di massa, della distribuzione del clock interno e della separazione estrema tra i segnali digitali e quelli analogici.
Nel caso dell'ingegnere del suono presente nello studio di registrazione, bisogna porre l'attenzione nel ripulire i percorsi dei cavi usando quelli con la giusta impedenza, evitando loop di massa e interferenze elettromagnetiche: ciò risulterà in un monitoraggio migliorato. Comunque, a sample rate elevati, le differenze di impedenza degradano la trasmissione del segnale e di conseguenza producono un alto jitter d'interfaccia, risultando in artefatti e interferenze; è saggio quindi configurare il cablaggio in maniera appropriata. Le connessioni digitali bilanciate possono anche ridurre la radiazione RF e migliorare la performance nei convertitori sensibili al jitter.

4.14.2 Misurazione Jitter

Di seguito vengono riportate alcune misurazioni del jitter eseguite sui convertitori D/A. I convertitori A/D e D/A possono essere testati utilizzando come segnale di test un' onda sinusoidale a frequenza molta alta, che deve essere molto pura e a frequenza stabile; di solito viene generata digitalmente.

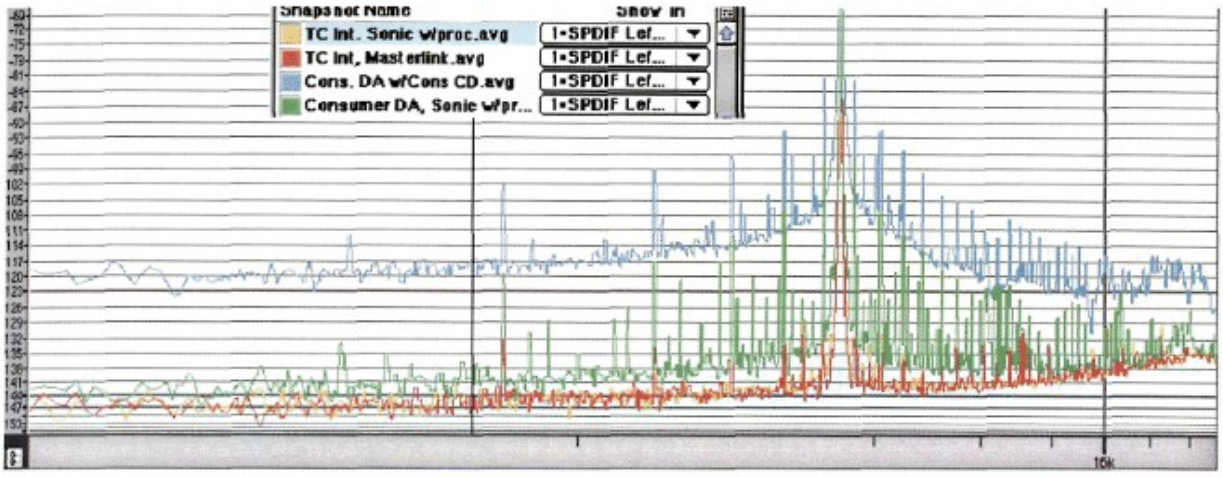
Politecnico di Torino

In questo caso viene utilizzato un segnale di test a 16 bit. Più basso il rumore di fondo, meno è il jitter. Non si è ancora compreso completamente quali picchi di jitter siano importanti dal punto di vista psicoacustico ma, come è stato detto prima, i test d'ascolto dimostrano che il jitter deve essere molto basso e vicino al rumore di fondo del sistema per risultare inudibile.



Test di jitter sul A/D UltraAnalog; in rosso il rumore di fondo e in blu l'analisi in frequenza del segnale
fonte Mastering Audio: the Art and the Science

La figura mostra il rumore di fondo della DAC usata nelle misurazioni, l'UltraAnalog A/D, la quale viene usata per campionare gli output di vari D/A sotto test; in blu sono rappresentati gli artefatti del segnale di test a 16 bit, i quali sono compresi tra -132 e -135 dBFS.

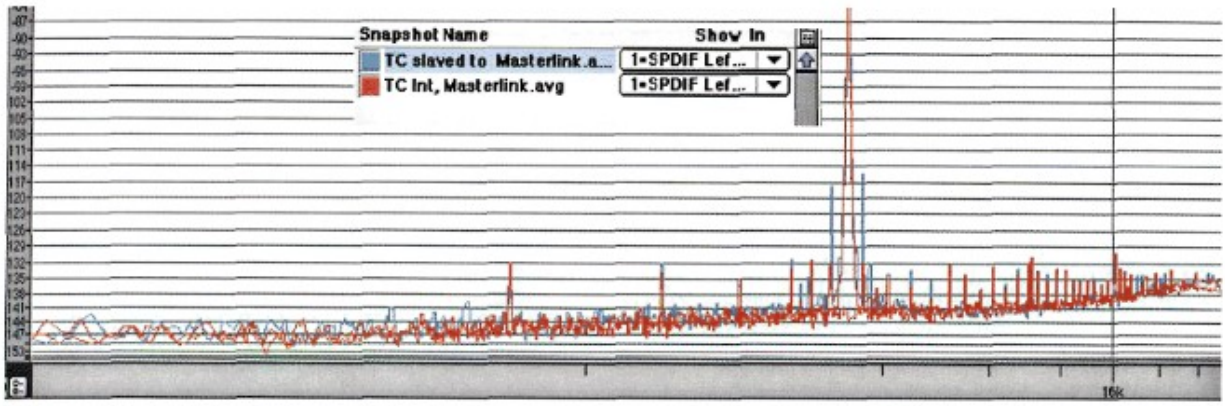


Misurazione del jitter con un segnale di test.

Politecnico di Torino

Traccia blu: DAC di fascia bassa, alimentato da un CD player livello consumatore
Traccia verde: DAC di fascia bassa, alimentato dalla workstation Sonic Solutions
Traccia arancione: jitter del TC con clock interno, alimentato dalla workstation Sonic Solutions
Traccia rossa: jitter del TC con clock interno, alimentato dal Alesis MasterLink
fonte Mastering Audio: the Art and the Science

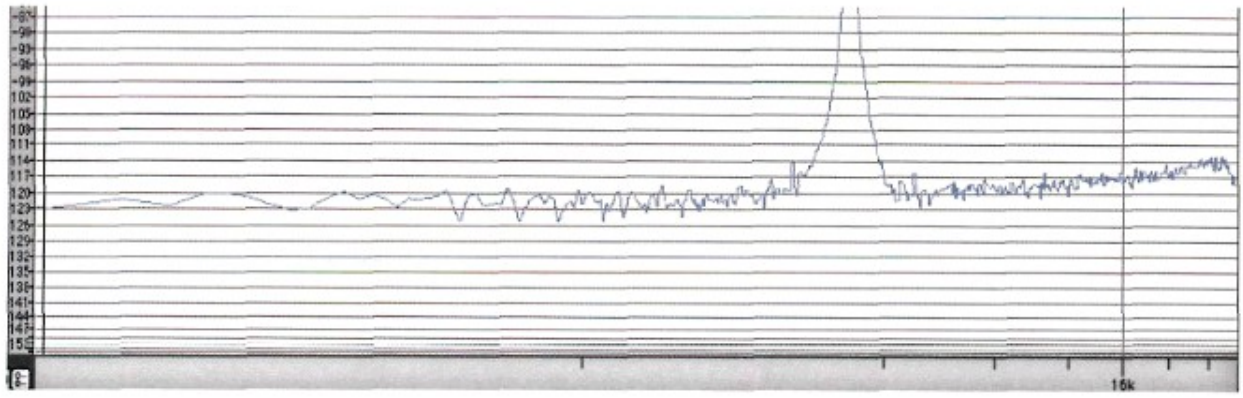
La figura soprastante compara il jitter di due DAC di qualità differenti, una di fascia bassa e una di fascia alta, il TC Electronic System 6000 D/A. A loro volta esse sono alimentate da tre sorgenti differenti, un CD player a buon mercato e due workstation digitali di alta qualità, rispettivamente della Sonic Solutions e dell'Alesis. Si può notare come vi sia nel convertitore economico una differenza sostanziale nella performance del jitter quando è alimentato dalle due sorgenti. Infatti utilizzando come sorgente un CD player a buon mercato si ha una quantità maggiore di jitter, mentre con l'input della workstation della Sonic Solutions se ne ha una quantità molto minore. Se questo fosse un grafico lineare invece di uno semi-logaritmico, sarebbe più evidente che il jitter produce artefatti concentrati intorno alla frequenza centrale del segnale di test, di solito con una deviazione simmetrica rispetto ad essa. Si osservi la performance del D/A di fascia bassa rispetto a quella del TC: quando alimentato da una delle due fonti, il jitter del TC è effettivamente identico al rumore di fondo a riposo.



Misurazione di jitter che dimostra come metodi differenti di clocking possano provocare risultati diversi, anche con la stessa sorgente Masterlink, che alimenta un TC D/A con un segnale a 16 bit
Traccia blu: TC D/A clockato esternamente al Masterlink attraverso l'AES/EBU
Traccia rossa: TC D/A con il clock interno
fonte Mastering Audio: the Art and the Science

Politecnico di Torino

La figura soprastante dimostra che il cambio nella tipologia di sincronizzazione difficilmente influenza la performance del jitter nel TC Electronic System 6000, con misurazioni molto buone con il clock interno e leggere differenze quando clockato esternamente all'AES/EBU. In quest'ultimo caso esso produce una quantità aggiuntiva di jitter ma molto leggera, che si traduce in due linee blu all'incirca a -117 dB, che rappresentano delle frequenze discrete vicine alla frequenza centrale. Si riesce a percepire una leggera degradazione nella chiarezza sonica, con un' immagine sonora più sporca e meno brillante nel caso del clock esterno, il quale implica che i picchi degli artefatti di jitter possono essere significativamente udibili. Quando è impostato sul clock interno (traccia rossa), il jitter è quasi al livello del rumore di fondo della sorgente e la maggior parte del tappeto di frequenze è composto dal segnale di test medesimo.



Segnale di test a 16 bit che alimenta una DAC Weiss con sync AES/EBU
fonte Mastering Audio: the Art and the Science

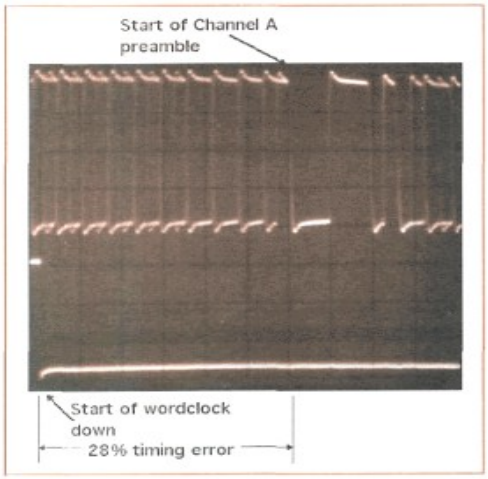
Il Weiss è il primo DAC misurato che, quando connesso all'AES/EBU, non mostra traccia apparente di jitter e risultanti frequenze spurie nel suo output. Al contrario il suo rumore di fondo si innalza con il segnale di test e tutti i jitter in arrivo sono stati convertiti in rumore casuale. Ciò può essere considerato come un miglioramento sonoro dovuto alla colorazione eufonica; dove in pratica il rumore di fondo maschera il jitter rendendolo inudibile.
Questo porterebbe a delle preoccupazioni riguardanti il rumore di fondo del convertitore, il quale maschererebbe i segnali di basso livello o dettagli fini come la riverberazione. In realtà, la bassa

Politecnico
di Torino

amplitudine la casualità del rumore sono la firma sonora più benigna che si possa desiderare in una DAC che suoni bene.

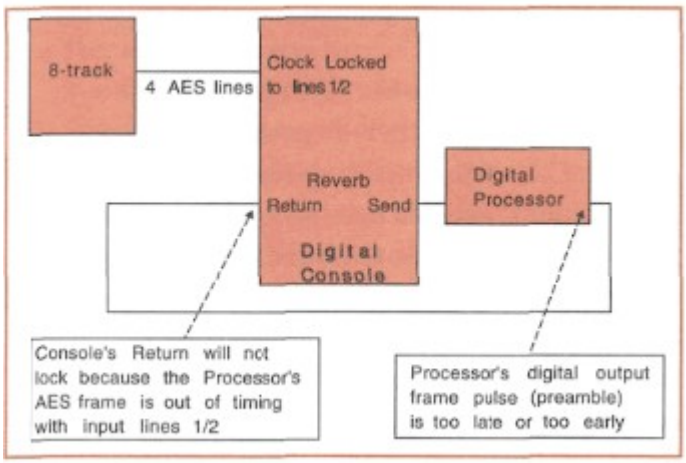
4.15 Errore di Framing

Sebbene il jitter sia il capro espiatorio per la maggior parte dei problemi nell'audio digitale, la maggior parte degli errori nell'elaborazione digitale sono problemi di framing. Il problema di framing causa degli errori di sincronizzazione nella comunicazione seriale e potrebbe creare delle problematiche quando si utilizzano dei processori digitali. Esso è il risultato della lettura errata del data frame, cioè quella stringa di simboli raggruppati in blocchi necessaria alla sincronizzazione, che viene letta nel punto d'inizio errato. I simboli sono costituiti da bit e i blocchi da byte, con l'uso di dieci bit nella trasmissione asincrona e otto in quella sincrona. L'errore di framing in uno stream di dati asincrono si può risolvere facilmente, ma un errore nello stream sincrono invece provoca degli errori nella comunicazione dei dati audio.
Uno dei principali segnali critici utilizzati nella sincronizzazione tra apparecchiature audio professionali è lo standard AES. Il segnale è costituito da blocchi audio, ognuno suddiviso in 192 frame, suddivisi a loro volta in due subframe, uno per il canale destro e uno per il canale sinistro.
Ogni subframe è diviso in 32 slot temporali ed ogni slot è riservato ai bit che portano le informazioni riguardanti il segnale audio tra cui: se il segnale è mono o stereo, il sample rate, il bit depth, l'origine e la destinazione all'interno della catena audio e il controllo di parità. I primi quattro spazi dello slot sono riservati ai bit di sincronizzazione, necessari per la comunicazione seriale e questa sequenza di bit viene detta preambolo. Si possono creare dei problemi tra sorgenti utilizzate contemporaneamente se la loro sincronizzazione non è perfetta e, nel caso del AES, si possono avere dei problemi nella ricezione del preambolo.



Rappresentazione dell'errore di framing
fonte Mastering Audio: the Art and the Science

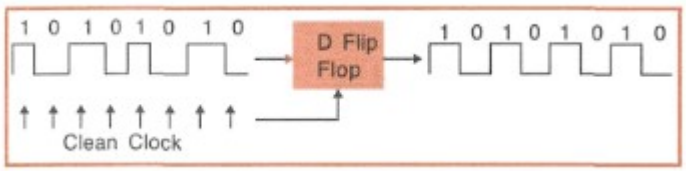
Sopra vi è la foto di un oscilloscopio: in alto vi è l'inizio del preambolo dell'AES e sotto il punto dove il wordclock cambia il suo stato da alto a basso.
Non vi è uno standard che definisce a quale cambio di stato del wordclock (cioè da basso ad alto o da alto a basso) il preambolo dell'AES dovrebbe allinearsi. Si ha quindi una differenza temporale di 11 µs a 44,1 kHz, che è sufficiente per depistare workstation, processori e console mediocri, producendo degli errori o addirittura interrompendo il segnale.
Molte console digitali e DAW sono ipersensibili ai problemi di framing. Poiché alcuni processori audio producono un output AES in ritardo rispetto allo stesso segnale d'input AES, esse hanno dei problemi nell'interfacciarsi a queste unità aggiuntive, risultando poco tolleranti all'errore di framing.
Ad esempio, si può essere costretti a introdurre una semplice unità di riverbero attraverso l'analogico invece che attraverso il digitale, poiché la workstation, sincronizzata ad un clock esterno AES, non è riuscita a sincronizzarsi al percorso di send/return del riverbero. La causa di questa falla è l'intolleranza della console agli errori di framing dello standard AES, utilizzato come clock esterno, in aggiunta al fatto che l'output del processore digitale è leggermente in ritardo, come illustrato nella figura seguente.



Errore di framing in una workstation durante l'utilizzo di un'unità di riverbero
fonte Mastering Audio: the Art and the Science

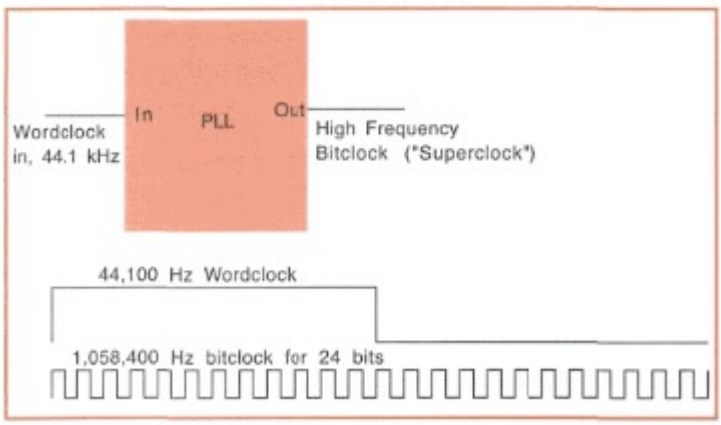
4.15.1 Reclocking

Se si hanno dei problemi di sincronismo non dovuti ad errori di framing, bisogna controllare che la frequenza della sorgente sia corretta o altrimenti bisogna regolare l'oscillatore che svolge il ruolo di master. I dati all'interno del tipico processore audio viaggiano da chip a chip in maniera seriale, cioè bit a bit. Un clock dà l'impulso che muove l'informazione lungo l'asse positivo della catena elettroacustica. Il bus del clock è distribuito a tutti i chip critici all'interno del box. Come visto prima, non importa se il clock è impreciso, il dato riesce ancora a passare al chip successivo nella catena elettroacustica. Ma qualche volta il dato ha bisogno di essere riclockato, per esempio quando si va a pilotare un convertitore D/A.



Circuito di reclocking
fonte Mastering Audio: the Art and the Science

Sopra è rappresentato un semplice circuito di reclocking: sul lato sinistro vi è una parola dati in arrivo che è stata clockata da un oscillatore affetto da jitter e quindi la sequenza temporale dei bit è leggermente fuori fase. Questa parola binaria viene fatta passare, un bit alla volta, in un circuito logico chiamato Flip Flop di tipo D, il quale è alimentato da un clock pulito. Quando il dato esce dal Flip Flop, in teoria, tutto il jitter è scomparso e il dato è pronto per pilotare la DAC. Si noti ora che la pulsazione del clock pulito permette di campionare in maniera appropriata ogni valore del dato input, ma solo se la pulsazione avviene all'interno del tempo d'accettazione di ogni bit in entrata. In questa illustrazione, il quarto e l'ottavo bit del dato rischiano di essere mancati se arrivano un attimo in ritardo; nel qual caso il clock pulito acquisirebbe il bit precedente e l'output restituirebbe un dato errato. Fortunatamente le fonti audio di solito hanno molto meno jitter di quelle illustrate nell'esempio, altrimenti il sistema collasserebbe e si avrebbe una elevata perdita di qualità, perchè i dati dell'output sarebbero completamente cambiati.



Reclocking di un PLL
fonte Mastering Audio: the Art and the Science

La figura illustra il motivo per cui serve un phase-locked loop (PLL). Se si sta facendo passare all'interno della DAC un audio da 24 bit, bit a bit, allora ci serve un clock ad una frequenza pari a 24 volte la frequenza del wordclock globale. Il wordclock entra nel device e deve essere moltiplicato alla frequenza più alta per controllare questi bit, in questo caso prendendo il nome di bitclock. E' facile dividere un segnale senza creare jitter, ma è molto difficile utilizzare il PLL per creare frequenze più alte e nel frattempo ridurre il jitter in uscita. La cosa che rende questi circuiti difficili da disegnare bene è il fatto che le alte frequenze, che si irradiano dalla porzione del circuito affetta da jitter, possono creare delle correnti parassite che ne contaminano la porzione pulita. Questi percorsi possono includere l'alimentatore e la massa. Accoppiando interferenze esterne e i loop di massa si ha l'incubo di un designer di macchine audio digitali: un errore di 10 picosecondi può fare la differenza tra il noise floor di 18 o 20 bit. Alcuni produttori usano un PLL duale, dove la prima parte è costituita da un circuito analogico e la seconda da un oscillatore a quarzo controllato in tensione (VCXO), in un tentativo di rendere il jitter simile a quello del cristallo di quarzo. Sfortunatamente i design che usano i VCXO hanno poca tolleranza alle variazioni di frequenza e quindi non possono essere usati quando si vuole variare la velocità di riproduzione variando il clock.

E' difficile, ma possibile, disegnare un PLL immune al jitter che sia preciso come il quarzo e che abbia anche un' ampia tolleranza alle frequenze e una sincronizzazione rapida. Rimane sempre il discorso che non importa quanti PLL vengano indicati sulle specifiche tecniche del prodotto, l'importante è la qualità del lavoro del designer, che dovrebbe essere testato in maniera oggettiva.

Politecnico di Torino

5 - Confronto tra segnale musicale digitale e segnale musicale analogico

5.1 Separazione dei Canali

Lo standard della musica odierna è lo stereo, cioè la traccia audio viene riprodotta in due canali separati, destro e sinistro, che insieme formano l'immagine sonora. Questa immagine crea un spazio sonoro tridimensionale dove vengono disposti gli strumenti in una maniera simile ad un palcoscenico e le nostre orecchie discriminano uno strumento in base alla posizione virtuale in questo spazio sonoro. Questa tecnica viene utilizzata nella fase di mixaggio per distribuire gli strumenti nell'immagine stereo, collocandoli all'interno di essa con vari gradi di posizionamento rispetto all'ascoltatore. Lo strumento usato nella console per regolare il posizionamento è il pan-pot (panoramic potenziometer), il quale sposta l'immagine stereo a destra o a sinistra, variando il livello relativo fra i 2 canali. L'immagine stereo non è la stessa per tutti i supporti.
Nella registrazione digitale, quando il segnale viene indirizzato al canale sinistro, il canale destro rimane completamente silenzioso. Ciò è dovuto al fatto che si utilizzano dei segnali sonori memorizzati come numeri e quindi con delle variabili completamente discrete, cioè con delle sequenze di 0 e di 1 che possono essere facilmente spente o accese; al contrario nel vinile abbiamo a che fare con delle grandezze continue. La puntina che segue il solco del vinile, legge i due lati del solco che rappresentano rispettivamente il canale destro e il canale sinistro, in virtù della modulazione a 45 gradi con cui è stato inciso. Quello che succede è che un lato della puntina, che dovrebbe captare solo le vibrazioni del proprio canale, in realtà cattura anche le vibrazioni parassite provenienti dall'altro lato e viceversa, per cui l'immagine stereo risultante non è altrettanto precisa. Questo fenomeno viene chiamato channel bleeding ed è una caratteristica che indica la fedeltà di riproduzione. La quantità di suono percepito dal lato non coinvolto nella riproduzione viene calcolato in decibel e varia in base alla frequenza ed è un parametro che indica la qualità di una testina: ad esempio per la Shure M97xE la separazione stereo fornita dal costruttore è di -25 dB a 1 kHz .
Questo fenomeno può essere verificato, ad esempio confrontando le copie in digitale e in vinile di Shook Me All Night Long degli ACDC, dove il riff iniziale della chitarra è panpottato completamente

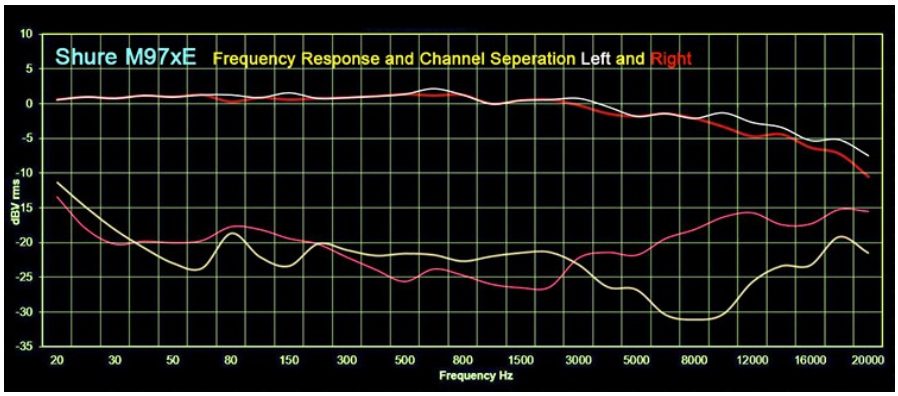
a destra, mentre all'estrema sinistra vi è la batteria. Nella versione digitale, ascoltando separatamente il canale destro e canale sinistro, le immagini stereo della chitarra e della batteria risultano completamente separate e non vi è alcuna intersezione tra le due; nella versione in vinile, al contrario, nell'ascolto del canale della batteria si può sentire l'interferenza della chitarra proveniente dall'altro canale. Quindi, in questo caso, si può dire che il medium più fedele alla registrazione originaria è il digitale.

5.2 Perdita di informazione

In natura non esistono segnali veramente analogici: le varie forme di trasmissione comunemente chiamate analogiche in realtà si basano su una discretizzazione più o meno fine dovuta alle varie soluzioni fisiche adottate. La risoluzione di un LP è limitata dalle dimensioni delle asperità su cui striscia la puntina e, siccome queste vengono erose durante l'ascolto, esso non suonerà mai due volte nello stesso modo; in più è suscettibile a temperatura, umidità e qualità dell'aria che cambiano in maniera irreversibile il contenuto d'informazione. Il vinile è talmente suscettibile alla temperatura che,ad esempio, non può essere appoggiato su un termosifone, pena una deformazione irreversibile.
Il nastro magnetico possiede dei domini di magnetizzazione di dimensione finita e la loro struttura determina la risposta in frequenza del nastro e la quantità di rumore. La caratteristica di queste discretizzazioni è che sono completamente scorrelate dal segnale, dando origine ad un rumore di fondo che in alcuni casi è quasi inavvertibile e in altri casi, come gli LP e i nastri analogici, è avvertibile ma può essere soggettivamente meno fastidioso delle alterazioni introdotte dal digitale, come accennato in precedenza.
Il segnale digitale, se rispetta i requisiti del teorema di Nyquist, riesce a ricreare in maniera corretta il segnale analogico; come detto in precedenza, considerando che il suono più alto percepibile da un essere umano sano è di 20 kHz, la frequenza di campionamento di 44,1 kHz è più che sufficiente per ricreare tutte le onde sonore percepite dall'uomo.
Al contrario, nella fase di incisione della lacca dell'acetato, il segnale analogico subisce già una equalizzazione e un filtraggio di alcune frequenze: le bassissime, che creerebbero delle difficoltà alla puntina nel seguire i solchi dovuti alla eccessiva larghezza di essi, e le altissime, dove la puntina non riesce ad incidere dei solchi troppo ristretti e ravvicinati. Questo processo di mastering del vinile è necessario alla corretta stampa dei solchi, ma inevitabilmente si converte in una perdita di informazioni.
Inoltre, quando viene letta l'informazione audio viene stampata vicino al centro del disco, la velocità di lettura che diminuisce influendo sulla risposta alle alte frequenze, influenzando negativamente l'abilità della testina nel seguirne il tracciato. Di conseguenza, un album inizia a perdere alcune delle alte frequenze già a partire dalla metà del disco, creando distorsione. Un

Politecnico di Torino

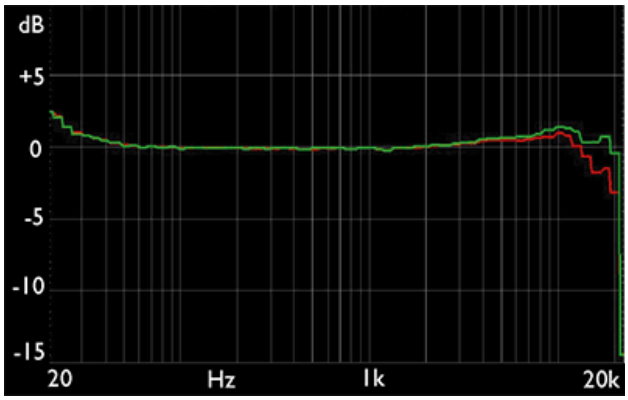
modo per compensare questa degradazione di qualità è il posizionamento dei brani nella fase di mastering finale: si può massimizzare la qualità del suono collocando le tracce che sono più dinamiche e intense all'inizio del LP e le tracce più tranquille (come le ballate) nella “zona di pericolo” intorno al centro.
Un altro aspetto da tenere in considerazione è la lunghezza dei lati del disco: più è corto un lato del disco (inteso come minutaggio), più spazio è disponibile per i solchi e, di conseguenza, un ingegnere del mastering può incidere i suoni con volume maggiore e con più dinamiche, senza preoccuparsi della collisione tra i solchi. Questo è il motivo per cui gli album odierni vengono distribuiti su due o più LP, in modo tale da conservarne la dinamica. Infatti più lunga durata di un disco, minore dovrà essere il passo dei solchi e di conseguenza le tracce dovranno essere incise ad un volume minore. Purtroppo più un disco diventa basso in termini di volume, più diventa suscettibile al rumore di superficie e quindi si tratta di trovare un giusto equilibrio tra spazio disponibile e perdita d'informazione. Quest'ultima è anche presente nei processi di creazione degli stampatori, che sono processi di clonazione in cui ad ogni copia si perde in qualità.
Un ultimo aspetto da considerare è la risposta in frequenza delle testine, la quale varia in base alle caratteristiche costruttive. Ciò significa che le testine agiscono come degli equalizzatori, attenuando in genere le alte e, come si vedrà dopo, enfatizzando le basse per effetto del feedback prodotto dal giradischi. Le vecchie testine a bobina mobile, come la Shure M97xE, hanno una risposta in frequenza che decade nella vicinanza delle alte, producendo un timbro tonale che viene percepito come “caldo”. Questa lacuna nella risposta in frequenza viene controbilanciata dal feedback alle alte frequenze.



Confronto tra canale destro e canale sinistro della Shure M97xE, sopra la risposta in frequenza e sotto la separazione di canale.

Politecnico di Torino

Nel grafico soprastante vengono messi a confronto il canale destro e il canale sinistro della Shure M97xE, prendendo in considerazione la risposta in frequenza e il channel bleeding.
Il grafico sottostante mostra come le testine moderne, come la Ortofon 2M Black, una testina a magnete mobile, abbiano una risposta in frequenza migliore nelle medio-alte e ciò risulta in un suono più preciso e pulito. Si consideri anche che la qualità e la condizione delle puntine ne cambia pesantemente la risposta in frequenza.

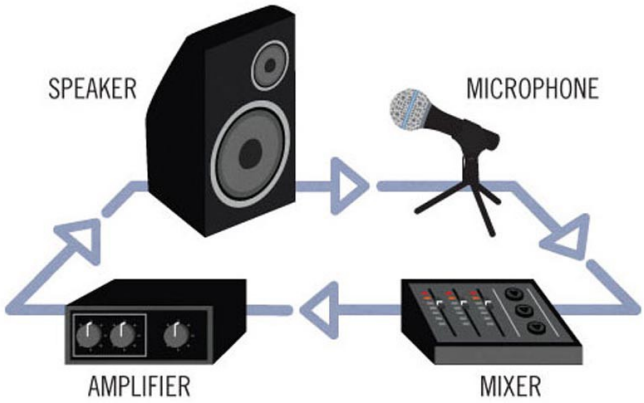


Risposta in frequenza della Ortofon 2M Black (traccia verde)

5.3 Feedback

5.3.1 Rumble Feedback

Il rumble feedback, o feedback di rimbombo, è una caratteristica forma di rumore causata dalle imperfezioni nei cuscinetti del motore dei giradischi. Il piatto tende ad avere una leggera quantità di movimento oltre alla rotazione desiderata. Il movimento addizionale è aggiunto al segnale desiderato come rumore, di solito alle basse frequenze, creando un suono di rimbombo durante i passaggi a volume basso. Ciò che può succedere è un cortocircuito sonoro detto feedback dove le vibrazioni che sono riprodotte dagli speakers vengono ricatturate dalla testina e le frequenze che si trovano in simpatia si vanno a sommare in ampiezza, in un circolo in teoria infinito e portando quindi il segnale a superare la soglia di saturazione e a provocare distorsione. I giradischi di fascia bassa usano cuscinetti poco stabili, i quali provocano feedback più facilmente, mentre quelli di fascia alta utilizzano cuscinetti costruiti in maniera più solida, i quali sono meno suscettibili al fenomeno.

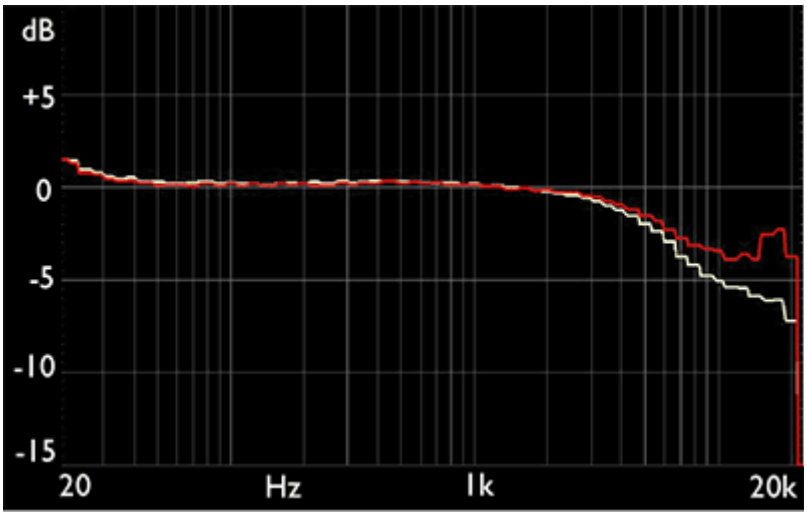


Circuito di Feedback

Il feedback avviene soprattutto durante la riproduzione di suoni con una decisa componente di bassi, come ad esempio un kick della batteria o una linea di basso. Nel caso di un locale al chiuso, dove di solito vi è la presenza di un sistema di amplificazione dotato di subwoofer, i quali sono altoparlanti specifici per le basse frequenze, il fenomeno risulta accentuato.
La distorsione introdotta però, se rimane all'interno di certi limiti, può risultare soggettivamente non fastidiosa e anzi desiderabile, dando l'illusione di una maggiore intensità e colore alle basse frequenze. I DJ e gli amanti del vinile di solito definiscono il suono del vinile come più “caldo” proprio in virtù di questo fenomeno.

5.3.2 Tip Mass Resonance

Il corrispettivo alle alte frequenze del low rumble feedback viene detto “tip mass resonance” ed è dovuto ad una risonanza meccanica tra la testina e il vinile. Essa avviene intorno ai 18 kHz e fa aumentare in maniera decisa l'intensità delle alte. Nelle vecchie testine, come ad esempio nella Shure M97xE, ciò permetteva di recuperare un po' delle alte frequenze che venivano smorzate dalla risposta in frequenza, ripristinando il suono originario. Nelle testine moderne, principalmente quelle a magnete mobile, ciò genera un suono più brillante che può soggettivamente risultare piacevole, ma che enfatizza la distorsione presente in vinili incisi male oppure danneggiati, evidenziando imperfezioni come graffi e grumi nella pasta del disco. Per controllare la risposta in frequenza della testina è anche molto importante utilizzare una corretta impedenza di ingresso del preamplificatore. I preamplificatori migliori consentono di variare la resistenza e la capacità dell'ingresso phono.



Risposta in frequenza della testina a bobina mobile Audio Technica AT-F3, con un picco nella alte frequenze causato dalla tip mass resonance

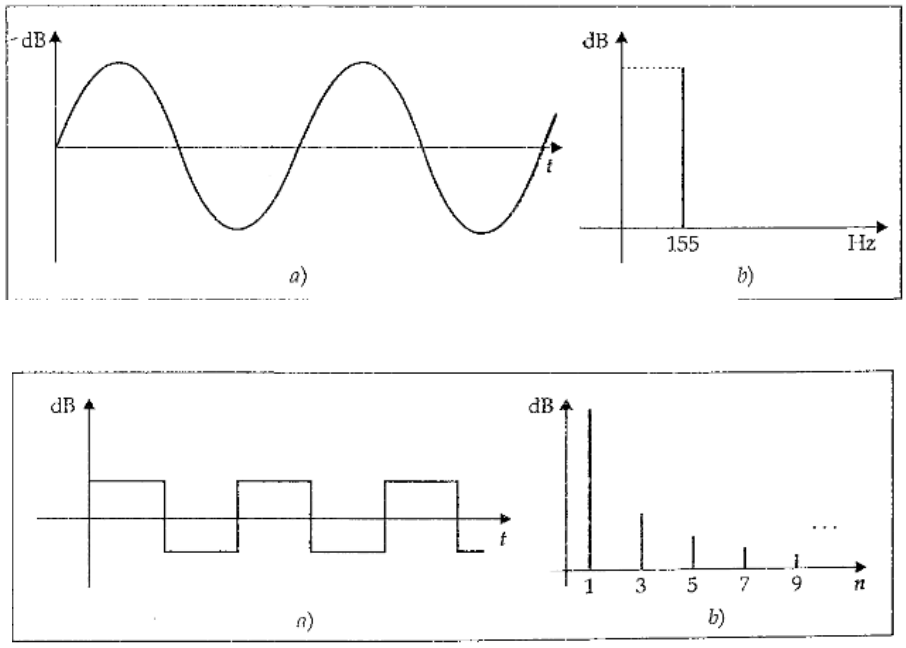
5.4 Rumore

I rumori provengono dalle imperfezioni e depositi di vario genere presenti sui solchi del vinile; essi si traducono in ronzii, sibili, scoppiettii o crepitii e sono in genere causati dall'usura e dalla polvere. Il segnale digitale, al contrario, è completamente immune ai rumori causati dalla polvere e dalla elettricità statica presenti sul vinile; queste imperfezioni possono essere soggettivamente desiderabili e vengono emulate da plug-in ed emulatori commerciali che riescono a riprodurre in maniera più o meno convincente il tipico fruscio del vinile. Nel caso del digitale la principale fonte di rumore risiede nella conversione A/D che, nel processo di quantizzazione, introduce del rumore bianco che è inversamente proporzionale al bit depth. Ma nel processo di mastering di un file digitale il rumore di fondo può essere cancellato in maniera arbitraria attraverso processori di dinamica, come il gate, o mascherati in frequenza attraverso la quantizzazione non lineare e il sovracampionamento, diventando perciò impercettibile.

Politecnico di Torino

5.5 Distorsione

La distorsione digitale è diversa da quella analogica.
Nel caso del vinile, la distorsione avviene quando l'ampiezza del segnale catturato dalla puntina satura il livello d'ingresso del pre-amplificatore. In essa le frequenze non vengono immediatamente clippate e tagliate completamente ma vengono flesse: la forma dell'onda sonora risultante è una sinusoide schiacciata in corrispondenza dei picchi. La distorsione analogica viene anche detta soft clipping. Guardando allo spettro del segnale le frequenze che vengono generate in questa fase sono delle armoniche pari, cioè armoniche che hanno un rapporto intero rispetto alla fondamentale e quindi gli artefatti possiedono una natura musicale. Per la sua caratteristica musicale, la distorsione analogica permetteva agli ingegneri di registrare un segnale che presentava in alcune parti distorsione; il passaggio del segnale saturo attraverso l'hardware analogico, composto da valvole e trasformatori, ne cambiava leggermente il tono e aggiungeva delle componenti armoniche pari al segnale, dando una tipica colorazione sonora che viene considerata come “vintage analogico”: una colorazione sonora che al giorno d'oggi è ritornata in auge ed molto desiderata in ambito artistico. L'indicatore utilizzato per misurare il segnale analogico è il VU Meter: lo 0 dBVU, non va considerato come asintoto superiore ma permette di superarne il limite e di utilizzare segnali più intensi con picchi fino a12-18 dB oltre lo zero. Oltre questo limite i picchi dell'onda vengono tagliati e si ha una distorsione armonica con componenti dispari in maniera crescente. Può anche accadere che la puntina si surriscaldi durante la riproduzione (e pure nel processo d'incisione della lacca d'acetato) e di conseguenza, a causa dell'eccitazione termodinamica, inizi a vibrare secondo dei modi che non sono quelli dovuti al solco sul disco, provocando distorsione e introducendo frequenze spurie.
Al contrario la distorsione digitale può risultare molto fastidiosa, poiché il cut-off è molto preciso e tagliente ed è quindi mandatorio evitare il clipping tenendo la traccia audio normalizzata sotto i 0 dBFS durante la registrazione. Infatti i picchi del segnale vengono tagliati di netto, formando delle onde quadre: ciò è dovuto al fatto che nel digitale ci troviamo di fronte ad un livello massimo del segnale, detto livello 0 nel decibel Full Scale, oltre al quale, per definizione del segnale digitale che è determinato da una sequenza finita di codici binari, non è possibile la rappresentazione. Tutte le informazioni al di là del range di quantizzazione vengono brutalmente eliminate, formando un onda quadra il cui spettro di frequenze è una serie di armoniche dispari. Questo tipo di distorsione viene detta hard clipping.



Nella prima figura sinusoide nel dominio del tempo a) e frequenza b).
Nella seconda figura onda quadra nel dominio del tempo a) e frequenza b).
fonte: Audio e Multimedia

5.6 Elaborazione del Segnale

5.6.1 Degradazione oggettiva e miglioramento soggettivo

In questo paragrafo verrà misurata e interpretata la performance di alcuni processori audio analogici e digitali, dimostrando che vi è un' interazione tra degradazione oggettiva e miglioramento soggettivo.
Non esiste un processore audio trasparente e le misurazioni utilizzate provvedono soltanto ad una piccola parte del quadro; un processore audio è come un oggetto all'interno di una casa senza porte, ma dotata solo di un certo numero di piccole finestre da cui ci si può affacciare. Osservando l'oggetto attraverso la visione di ogni singola finestra si possono interpretare più dati e farsi un'idea migliore, ma non si è mai completamente sicuri di quello che si sta guardando e bisogna sempre lasciare la possibilità che vi siano degli aspetti che non possano essere visti, un qualche mistero sul perché un processore suoni bene oppure suoni male.

Politecnico di Torino

Per esempio, ecco un paio di misurazioni oggettive che semplicemente non funzionano, dove un filtro digitale che dovrebbe suonare “smorzato” risulta “chiaro”.
Il TC Electronic System 6000, già citato nel Capitolo 4, permette all'utente di scegliere tra dei filtri passa-basso per la conversione A/D e D/A con caratteristiche differenti. Uno di questi filtri, chiamato Lineare, ha una frequenza di cut-off oltre i 16 kHz e, considerando un sample rate a 44,1 kHz, si è portati a pensare che suoni attenuato sulle alte frequenze. Ma invece suona più aperto e chiaro rispetto al filtro a 20 kHz chiamato Vintage e, rispetto ai filtri di altri convertitori che si estendono anche loro fino ai 20 kHz, suona più attenuato. Quindi la larghezza di banda non tiene conto della psicoacustica.

Il TC Electronic System 6000
fonte Mastering Audio: the Art and the Science

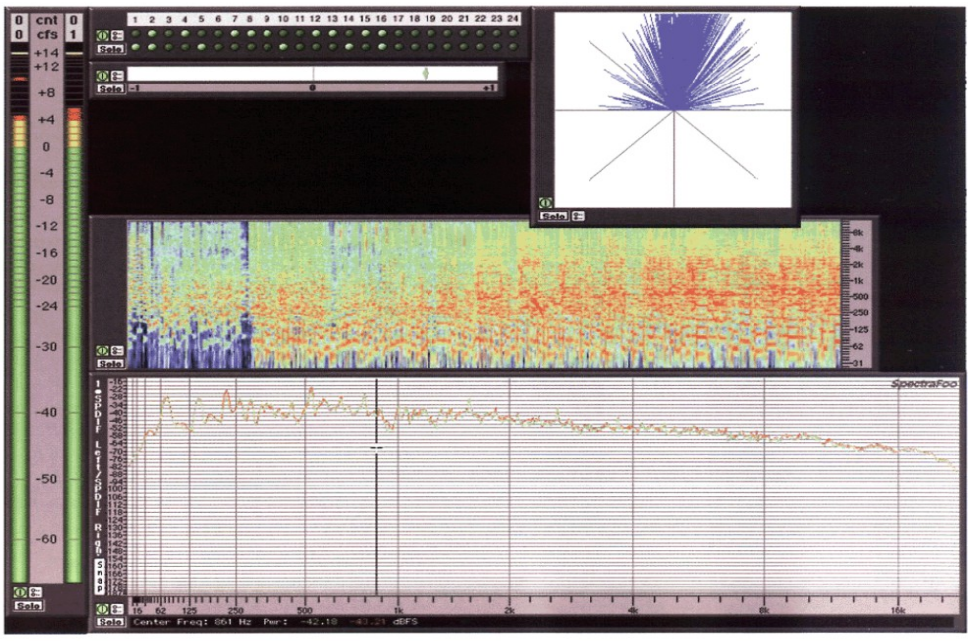
Infatti si ha attrezzatura da studio di registrazione il cui rumore di fondo, dopo la conversione A/D, si trova nel range tra -120 dBFS e i -50 dBFS. Ma la maggior parte di essa viene percepita comunque come silenziosa: se si deve mettere l'orecchio attaccato all'altoparlante per ascoltare l'hiss, allora il rumore può essere considerato insignificante.
Questa discrepanza è dovuta ai metodi di misura con i quali i produttori valutano il noise floor, che hanno poco a che fare con la percezione umana. Un convertitore, che ha un rumore di fondo calcolato con la curva di pesatura A di -108 dBFS, può apparire meno rumoroso di un altro convertitore con un rumore di fondo di -155 dBFS, sempre calcolato con la curva di pesatura A.
Il motivo è che la citata curva di pesatura A non considera in maniera adeguata la sensibilità dell'orecchio nelle bande critiche. Ciò vuol dire che un convertitore, con il rumore di fondo misurato più basso ma che produce più energia sui 3 kHz dove l'orecchio, secondo la curva di Fletcher-Munson, è più sensibile, sarà percepito come più rumoroso rispetto ad un convertitore

Politecnico di Torino

con rumore di fondo più alto. Per essere accurati dal punto di vista psicoacustico, gli standard di misurazione del rumore dovrebbero adottare una curva più vicina alla sensibilità in frequenza dell'orecchio umano, ad esempio utilizzando curve di nono ordine, dette di pesatura F, che sono utilizzate dai dither di alta qualità.
Ci sono molte altre aree nelle quali le misurazioni tradizionali non sono correlate a quello che ci dicono le nostre orecchie, in particolare nella valutazione dei sistemi di codifica con un basso bit rate: questi sistemi numericamente funzionano abbastanza bene ma, poiché l'orecchio è stato sensibilizzato a sentirne gli errori, riesce facilmente a identificarne gli artefatti.
Questo discorso preliminare è importante per scoprire perchè alcuni processori analogici e alcuni digitali suonino meglio degli altri.

5.6.2 Strumenti d'analisi

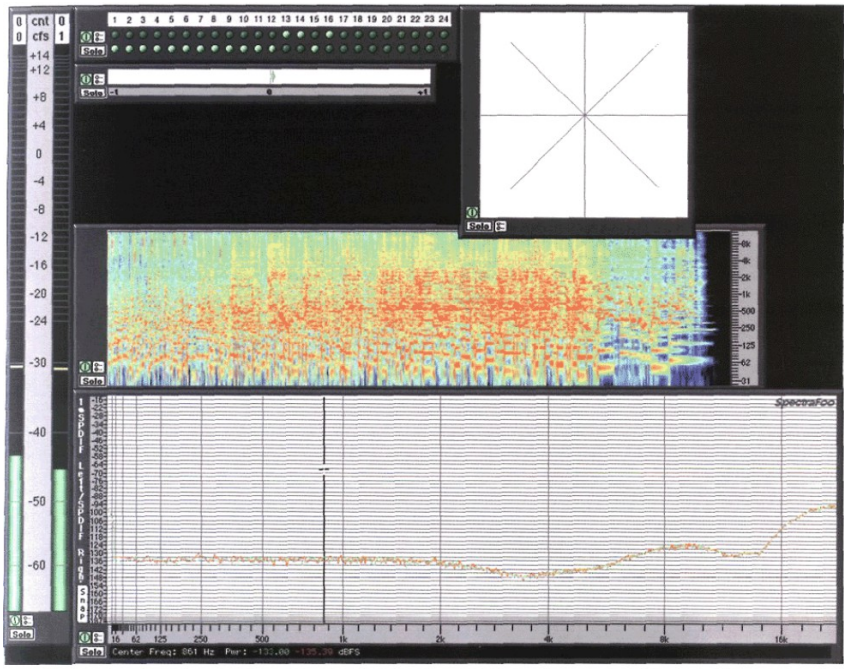
Il principale strumento utilizzato è la FFT (Fast Fourier Transform). Gli analizzatori FFT sono utili per analizzare e comparare le prestazioni dei processori audio analogici e digitali.
È possibile misurare in maniera oggettiva il segnale sotto la soglia del rumore con un analizzatore FFT.
Anche se l'analizzatore utilizzato negli esempi successivi può esaminare solo 24 bit (a causa della limitazione dell'interfaccia AES/EBU, utilizzata come interfaccia d'input), esso può analizzare distorsioni al di sotto del rumore di fondo fino a -40 dB sotto il noise floor del sistema a 24 bit. Ma non tutte le sottigliezze possono essere percepite e alcuni fenomeni non sono analizzati in maniera efficace. E' stato dimostrato che si può percepire un troncamento a 24 bit in un sistema a 18 bit e che il range dinamico dell'orecchio è approssimativamente di 20 bit, cioè ha una gamma dinamica all'incirca di 129 dB sempre in base alla frequenza e a certe frequenze possiamo addirittura sentire sotto lo 0 dB SPL. Quindi non è detto che l'analisi con la FFT riesca a spiegare in maniera esaustiva le qualità sonore di un processore audio.



Politecnico di Torino

La figura mostra una sessione di mastering di un CD.
Da sinistra a destra in alto: il display del volume sonoro, il bitscope e un indicatore della posizione stereo. Direttamente sotto il bitscope vi è un misuratore di correlazione. A metà dello schermo vi è uno spettrogramma, con la parte calma nella sezione sinistra prima dell'inizio della canzone. In alto a destra vi è un indicatore della posizione stereo e in fondo lo spettrografo con il canale sinistro in verde e il canale destro in rosso.
fonte Mastering Audio: the Art and the Science

In figura viene mostrato una sessione di mastering di un CD, dove vengono indicati gli strumenti utilizzati per l'analisi del segnale audio.
Nella parte superiore vi è il bitscope che in questo caso mostra 16 bit attivi, un' indicazione che il generatore di dither sta facendo il proprio lavoro attivando le regioni di quantizzazione di basso livello.
In alto a destra è collocato l'indicatore della posizione stereo, il quale al momento dell'istantanea mostra che l'informazione è leggermente bilanciata a destra.
Appena sotto il bitscope vi è l'indicatore di correlazione. Le deviazioni del misuratore nelle vicinanze del centro della scala indicano meno correlazione tra i canali e quindi una immagine stereo più spaziosa e larga.
A metà dello schermo vi è lo spettrogramma, il quale mostra l'intensità spettrale nel tempo. Questo può essere utile per identificare la banda di frequenza delle problematiche sonore.
Al fondo vi è lo spettrografo, la cui forma dà una vaga idea sul timbro del brano.



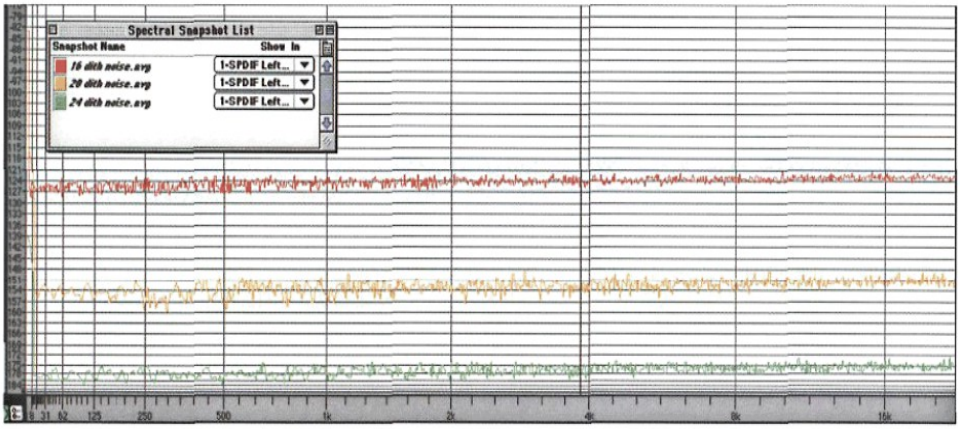
Istantanea degli strumenti d'analisi del segnale audio durante un momento di pausa nel brano musicale. Si noti come nel bitscope sono in funzione solo gli ultimi quattro bit, e lo spettrografo rileva la curva caratteristica di un dither del nono ordine. Si può vedere la fine del brano musicale alla destra della timeline dello spettrogramma, dove inizia la banda nera.
fonte Mastering Audio: the Art and the Science

La figura soprastante mostra lo spettrografo durante una pausa nel programma musicale, con solo gli ultimi quattro bit funzionanti, confermando che il dither sta lavorando correttamente.
Bisogna notare che il bitscope mostra quattro bit in funzione (poiché il dither è casuale, in questa istantanea il bit 15 è a zero) e che lo spettrogramma mostra la curva del rumore di dither, il quale può essere identificata dalla sua forma come una curva del nono ordine. Usando questo strumento, si può determinare il tipo di dither utilizzato dall'ingegnere del mastering nella realizzazione di un brano.
Il livello sonoro non era ancora decaduto quando questa istantanea fu scattata. Il metro di correlazione fluttua molto leggermente vicino al centro del contatore, mostrando che il dither non è correlato ai canali e la fase è casuale, quindi l'immagine stereo è corretta.

5.6.3 Mixing Analogico e Digitale

Politecnico di Torino

Nell'audio, sia digitale che analogico, il percorso del segnale più pulito è quello che contiene il numero minori di componenti. Il convertitore è la componente che degrada di più il segnale e, nonostante negli ultimi anni la qualità sia migliorata ampliamente, si dovrebbe evitare ogni volta una conversione A/D non necessaria.
Nel caso della elaborazione completamente analogica che verrà trasferita sul digitale, è meglio fare la lavorazione analogica prima della conversione A/D ma, poiché molto spesso i mix sono basati completamente sul digitale e spesso un ingegnere del mastering deve utilizzare dei processori analogici poiché suonano più organici dei loro equivalenti digitali, si possono usare i benefici tonali dell'elaborazione analogica abbinati all'uso del sovracampionamento per controbilanciare le perdite di trasparenza dovute a conversioni multiple.

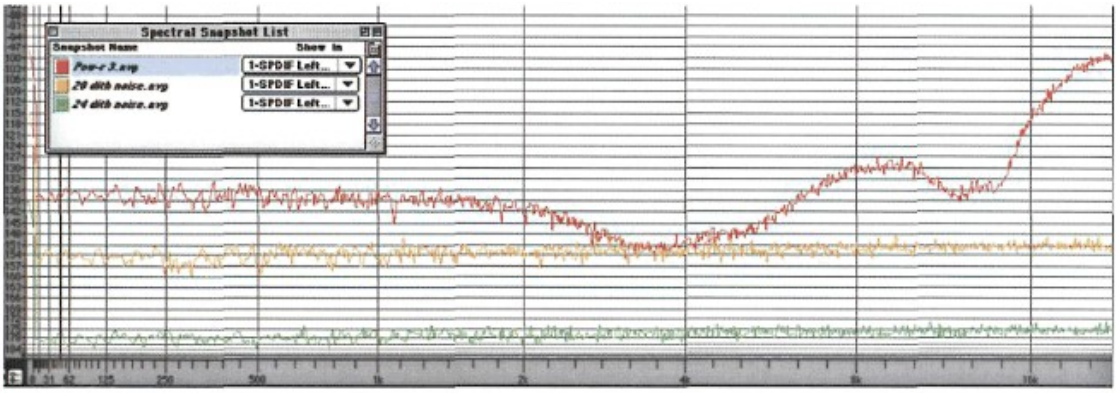


In rosso, dither a 16 bit; in giallo dither a 20 bit; in verde dither a 24 bit
fonte Mastering Audio: the Art and the Science

Nella figura soprastante, vengono comparati il rumore di dithering lineare, rispettivamente a 16, a 20 e a 24 bit.
Dal punto di vista psicoacustico, per la maggior parte della musica i 16 bit, utilizzati in maniera corretta, sono sufficienti. Gli psicoacustici che studiano i limiti dell'orecchio umano hanno determinato che 20 bit è sufficiente per una buona performance di conversione A/D e D/A. Ogni altra cosa funge da palliativo ed è molto raro trovare un convertitore a 24 bit che abbia un livello di rumore migliore dei 18-20 bit. Ma quando si post-produce il segnale si accumulano rumori di quantizzazione e troncamenti e, vista l'esigenza di headroom professionale, vi è la necessità di una

Politecnico di Torino

lunghezza di parola superiore ai 24 bit, per evitare l'accumulo di distorsione digitale e avere più margine di rappresentazione. Per l'elaborazione da purista servirebbero quindi fino a 48 o 72 bit, specialmente con cambiamenti estremi di gain, l'uso di filtraggio complesso, l'uso di compressioni o di processi a cascata. E' un mito che non ci sia una perdita ad ogni passaggio della lavorazione digitale; passaggio dopo passaggio, bit dopo bit, il suono soffre con ogni operazione nell'audio digitale. Un corollario importante a questa discussione sui processori è la regola della qualità della sorgente: le registrazioni e master sorgenti dovrebbero avere una risoluzione più alta che l'eventuale release sul medium. Bisogna sempre partire dalla sorgente alla massima risoluzione possibile e mantenere la risoluzione per il maggior tempo possibile durante l'elaborazione. Quando viene fatto il mastering, una conseguenza di questo regola è di ridurre il numero di generazioni e copie, e se possibile, tornare indietro di una o più generazioni quando un nuovo processo deve essere aggiunto o applicato. Questa regola viene applicata anche quando si sta producendo per il formato MP3 o un altro risultato finale dove i dati vengono compressi.
Si consideri un medium lossy come la cassetta analogica. Si copi su cassetta da una sorgente ad alta qualità, come un CD, e si vedrà che suona molto meglio di una copia da sorgente inferiore, come la radio FM, poiché sono stati evitate delle perdite cumulative di banda. In altre parole, più alta la qualità dell'audio da cui si inizia, migliore sarà il prodotto finale, sia che si tratti di un CD per audiofili, un CD Rom multimedia, o un OGG Vorbis per Spotify. Può sembrare banale ma non ci si può sbagliare iniziando a 88,2 kHz con una bit depth di 24 bit, soprattutto se il prodotto finale sarà a 44,1 kHz e 16 bit. La conversione del sample rate dovrebbe essere il penultimo processo, seguito poi dal dithering per poi arrivare al medium finale.

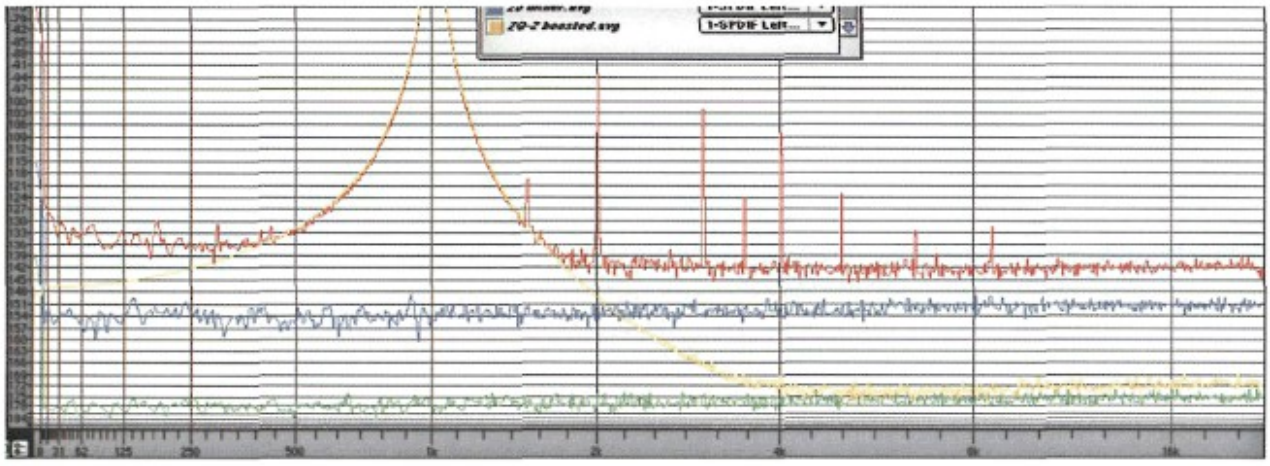


Spettrografo del rumore di fondo di un dither del nono ordine a 16 bit (traccia rossa), confrontato con un dither lineare a 20 bit (arancione) e un dither lineare a 24 bit (verde).
fonte Mastering Audio: the Art and the Science

Politecnico di Torino

In figura viene confrontato il rumore di fondo di un dither del nono ordine a 16 bit (traccia rossa) con il rumore di un dither lineare a 20 bit (arancione) e un dither a 24 bit (verde). La forma del nono ordine è strutturata in modo tale da massimizzarne la performance, tenendo il rumore nei pressi della zona delle frequenze dove l'orecchio è meno sensibile. Questo dither raggiunge la performance a 20 bit nella banda critica superiore (circa 3,5 kHz). Di conseguenza, la maggior parte del rumore di basso livello e della riverberazione, che sarebbe stata mascherata dal rumore di quantizzazione, viene preservato pure con la riproduzione a 16 bit. Questa performance può essere ottenuta registrando all'inizio con una lunghezza di parola maggiore, lasciando che il rumore si accumuli e l'SNR diventi leggermente peggiore, per poi aggiungere alla fonte processata il dither finale.

5.6.4 Equalizzatori



Distorsione e rumore del Millenia Media NSEQ-2, un equalizzatore analogico (rosso), confrontato col rumore di fondo casuale a 20 bit (blu), il rumore di fondo a 24 bit (verde) e l'equalizzatore digitale Z-Systems ZQ-2 (giallo)
fonte Mastering Audio: the Art and the Science

In figura viene mostrata la bassa distorsione e la ottima performance del rumore del Millenia NSEQ-2 (traccia rossa), il quale è conosciuto e apprezzato per il suo design e rappresenta lo stato dell'arte nell'equalizzazione analogico-valvolare; in questo caso è stato eseguito un boost alla frequenza di 1 kHz .

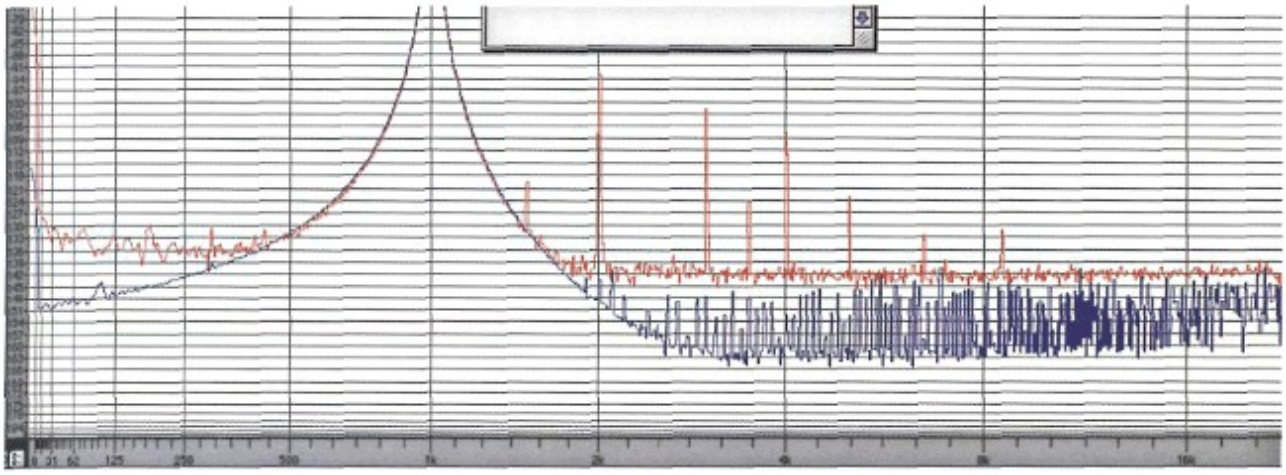
Millenia NSEQ-2

Per questa analisi in frequenza viene utilizzata una sinusoide a 24 bit ditherata, con un livello di -6 dBFS ad 1 kHz. Essa viene utilizzata per alimentare una DAC. Il segnale viene processato poi dal Millenia NSEQ-2 per poi essere trasmesso ad un convertitore A/D e infine mandato alla FFT.
Si noti che la distorsione dell'equalizzatore è dominata da armoniche di seconda, di terza e di quarta, le quali tendono ad addolcire il suono.
Per confronto sono mostrati rispettivamente in blu e verde il rumore a 20 e a 24 bit: si noti come il rumore analogico del NSEQ è maggiore di 10 dB rispetto al rumore a 20 bit, rendendolo virtualmente un equalizzatore a 18 bit.
Il risultato è che questo equalizzatore è estremamente silenzioso. Il suo rumore è inudibile su un monitor settato ad un gain ragionevole, a meno che non si mettano le orecchie attaccate allo speaker, dimostrando che il rumore di fondo è probabilmente l'ultima delle preoccupazioni. In confronto il nastro analogico è ancora più rumoroso, ma nessuno se ne lamenta nel suo utilizzo nella produzione musicale.
In giallo vi è la performance di un equalizzatore digitale di alta qualità, lo ZQ-2, ditherato a 24 bit, su cui è applicato un boost di 5,8 dB a 1 kHz con un Q factor di 0,7.
La performance della distorsione armonica è da manuale e non ci sono armoniche visibili nella FFT. Qualche ingegnere del suono usa la parola “secco” per descrivere il suono di un componente che ha poca o nessuna distorsione.
Le armoniche sono molto più che mere differenze soniche tra diversi elaboratori audio: le frequenze spurie prodotte dalla circuitazione possono far perdere definizione ai suoni. Questo equalizzatore digitale, al contrario, mantiene il mordente anche alle basse frequenze che sono più suscettibili alle distorsioni. La curva dell'equalizzatore digitale e quella dell'analogico sono differenti, sebbene lo ZQ-2 fa un buon lavoro nel simulare le forme di un filtro analogico dolce. Si può concludere in questo caso che gli equalizzatori raggiungono la stessa performance, nonostante quello analogico sia, sulla carta, più rumoroso. Ciò va a supporto della teoria per cui il

Politecnico di Torino

fattore importante è la qualità e la cura utilizzata nel creare processori audio, confermato dal prezzo elitario di queste due apparecchiature.

Z Systems Z-Q2

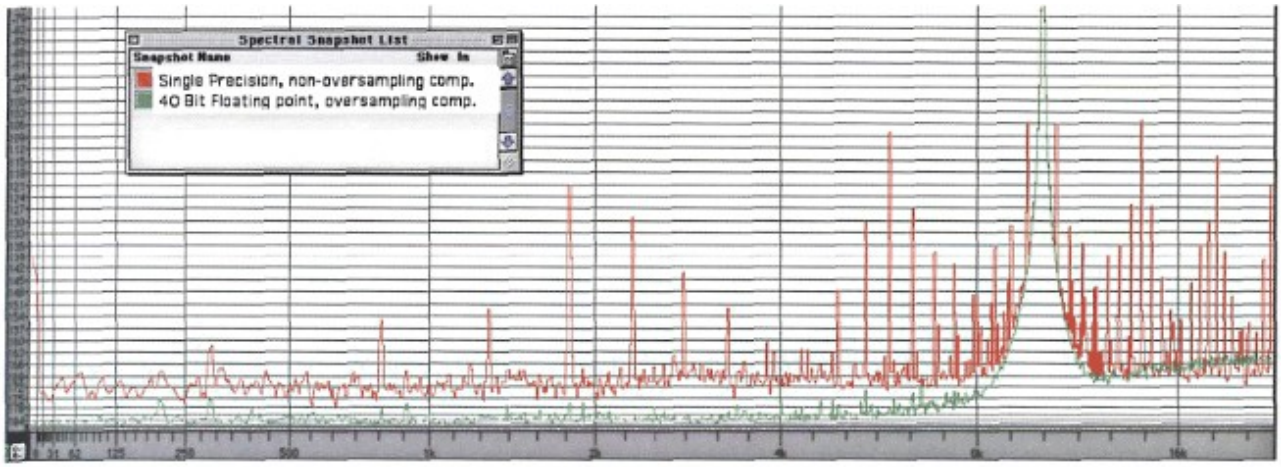


Distorsione e performance del rumore dell'equalizzatore analogico Millenia Media NSEQ-2 (traccia rossa) confrontato con il Z-Q2, settato per troncare a 20 bit senza dither (traccia blu).
fonte Mastering Audio: the Art and the Science

La distorsione per troncamento può essere definita brutta. Nella figura soprastante, vengono messi di nuovo a confronto il Millenia NSEQ e il Z-Q2; in questo caso quest'ultimo viene settato per troncare a 20 bit senza dither. È una pratica altamente sconsigliata e ci sono dei modi migliori per aggiungere “sporco” all'audio invece che spegnere il dither. Infatti la maggior parte dell'ambiente, dello spazio e del calore della sorgente originaria è stato troncato, perso completamente e convertito a rumore di basso livello, provocando una grave distorsione inarmonica che può essere

identificata nel grafico come la sequenza ravvicinata di frequenze spurie presenti al di sopra del boost; anche una piccola parte della distorsione non armonica può essere fastidiosa.
In questo caso si giunge alla conclusione che suona molto meglio un processore analogico con un rumore di fondo più alto nella scala dei decibel ma con un rumore più lineare, cioè senza frequenze spurie inarmoniche e con l'aggiunta di distorsioni armoniche di seconda e terza, rispetto ad un processore digitale senza dither con un rumore di fondo mediamente più basso ma che soffre di distorsione digitale.

5.6.5 Processori di dinamica



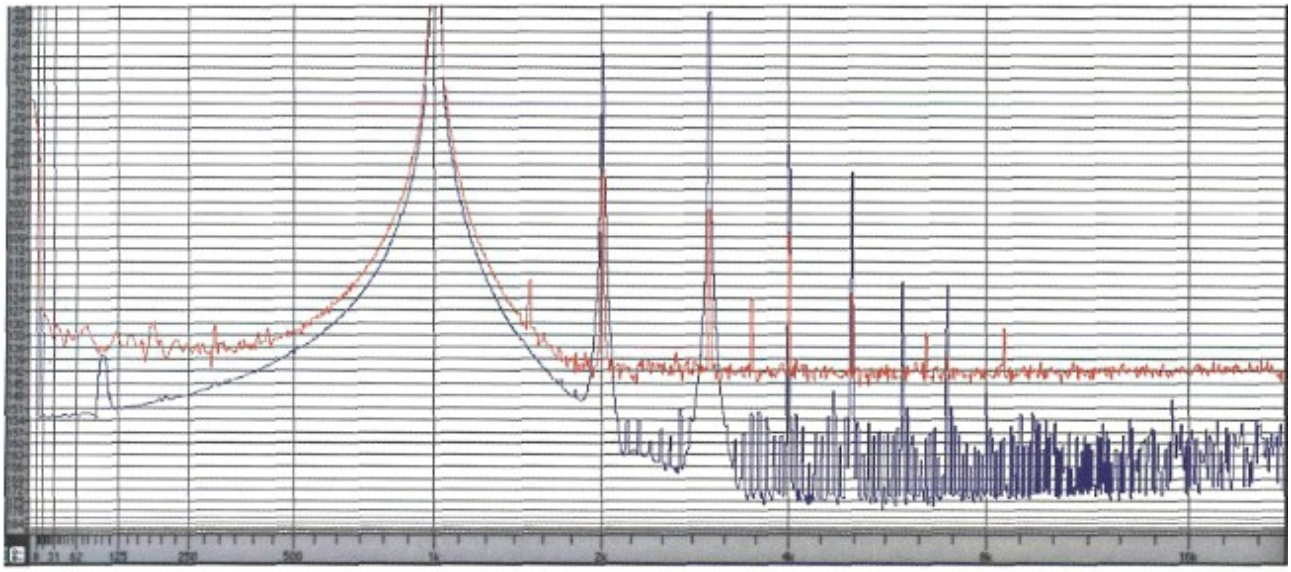
Confronto tra due compressori digitali, entrambi con una compressione di 5 dB di un segnale a 10 kHz.
Traccia rossa: compressore con precisione a cifra singola senza oversampling.
Traccia verde: compressore a 40 bit a virgola mobile, con doppio campionamento e dither a 24 bit.
fonte Mastering Audio: the Art and the Science

I compressori digitali implementati in maniera economica producono una seria distorsione inarmonica. In questo caso vengono messi a confronto due processori di dinamica digitali, entrambi con una compressione di 5 dB ad un segnale a 10 kHz.
In rosso vi è il compressore a precisione singola senza sovracampionamento e in verde il processore a doppio campionamento, implementato con 40 bit a virgola mobile. Si noti come la compressione a precisione singola produce molti artefatti di aliasing non armonico, specialmente nella banda critica delle frequenze mediane.
I compressori che suonano male sono comuni nelle console digitali a basso costo e nei plug-in delle DAW, poiché serve molta potenza di calcolo per fare il doppio campionamento. Ciò è un

punto molto importante nel confronto tra l'elaborazione analogica e quella digitale: la proliferazione e il cattivo uso del digitale a buon mercato, il cosiddetto digital cheap, ha degradato la qualità della musica registrata nell'era del personal computer e ha creato una cattiva nomea nei confronti dell'elaborazione digitale.
Le misurazioni sulla distorsione statica non spiegano comunque in maniera esaustiva i motivi per cui alcuni compressori suonino benissimo mentre altri risultino dolorosi per le orecchie.
Ci sono dei compressori analogici che, sebbene non siano trasparenti e cioè hanno una loro influenza arbitraria sulla componente tonale presente nel segnale in output, aggiungono del carattere sonoro eccitante ed interessante. Le registrazioni su nastro analogico sono un esempio perfetto di questo tipo di processo: misurate oggettivamente sono rumorose e distorte, ma soggettivamente sono belle da ascoltare.
La ricerca psicoacustica è più avanzata sugli effetti della distorsione di mascheramento e il rumore rispetto ai componenti tonali e ai rumori di fondo. Per esempio, il rumore di fondo del sistema DSD, utilizzato nel Super CD e illustrato nel Paragrafo 3.4.1, non è particolarmente silenzioso (infatti è intorno ai -120 dB nella banda dell'udibile) ma ha un suono eccellente, indicando che il rumore basso non deve essere il solo parametro di giudizio. Si potrebbe addirittura concludere che una caratteristica di un processore che suona bene è il mascheramento che si ottiene con il rumore di fondo: nel caso del DSD forse -120 dB è un livello abbastanza alto da coprire le parti fastidiose della distorsione delle attrezzature analogiche e digitali. In aggiunta, le registrazioni libere da rumore possono suonare molto sterili poiché tutti le imperfezioni e gli errori causati dai musicisti e dalla loro amplificazione vengono completamente captate dai media silenziosi. Quindi, qualche volta, aggiungere rumore extra può essere più benefico alla musica rispetto al lavorare senza rumore.
Una delle ragioni per cui il nastro analogico suona più musicale è dovuto al fatto che il rumore può essere molto eufonico. Le firme sonore di molti compressori analogici sono considerate dei classici, perchè sono “fat” ma allo stesso tempo chiare e sono frutto di una combinazione unica di attacco e release, che in teoria potrebbero essere emulati da un processore digitale. Ma i designer dei plug-in sono obbligati a minimizzare i loro processori per alleggerire il carico di lavoro sulla CPU e permettere di inserire più plug-in all'interno della DAW senza che si abbia un crash del sistema. Certamente alcuni compressori digitali sviluppati come plug-in, come il Weiss, non “suonano” digitali e quindi la qualità sonora analogica può essere ottenuta anche con delle skill di programmazione.

5.6.6 Simulatori analogici

Cranesong Hedd 192 Fonte: Cranesong.com

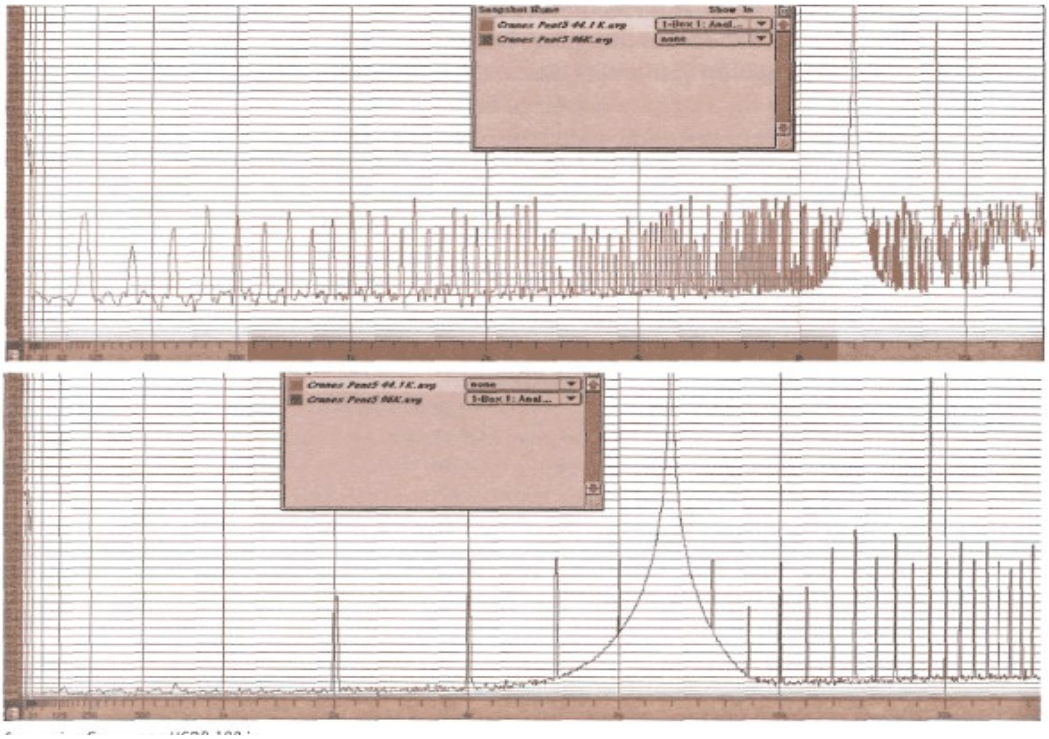


Comparazione tra il simulatore digitale analogico Cranesong HEDD-192 (traccia blu) e il NSEQ (traccia rossa). Fonte: Mastering Audio: the Art and the Science

Il grafico soprastante mette a confronto il NSEQ e il Cranesong HEDD-192, un convertitore digitale-analogico che simula le caratteristiche del nastro analogico ed è dotato di una qualità sonora eccellente. Il Cranesong (traccia blu) è stato impostato per produrre una struttura armonica simile a quella del NSEQ e i livelli del HEDD sono stati settati di proposito per produrre più distorsione rispetto a quella prodotta dal Millenia.
All'orecchio giunge comunque l'illusione che si stia ascoltando un processore analogico eccellente, senza nessuna perdita di risoluzione o di rappresentazione. Le frequenze spurie nella parte bassa del grafico potrebbero destare sospetti e, guardando attraverso questa finestra, si potrebbe

pensare che il Cranesong stia troncando delle informazioni importanti. Ma ci sono due importanti fattori da considerare per comprendere meglio la situazione: la prima è che lo “sporco” del Cranesong è più basso di 12 dB rispetto ad un processore digitale con troncamento (si veda il caso del Millennia confrontato con lo ZQ-2 settato per troncare a 20 bit) e quindi le frequenze spurie vengono mascherate dal rumore e dalle sue armoniche eufoniche. Il secondo punto è che l'Heed è dotato di un' architettura di somma interna unica nel suo genere che non altera, tronca o ricalcola il segnale della sorgente originale. Infatti, esso clona il segnale originale e lo indirizza verso l' output, mixandolo nel frattempo con la distorsione che viene calcolata a parte; di conseguenza preservando l'ambiente e i livelli bassi del segnale originale. La distorsione di basso livello presente nello spettrogramma è frutto quindi di una distorsione additiva e non il risultato di ricalcoli della sorgente. In altre parole, solo la distorsione viene distorta.
Le misurazioni vengono effettuate secondo diverse frequenze di campionamento: prima a 44,1 kHz, poi a 88,2 kHz e infine a 96 kHz. Come si può vedere nelle due figure successive, a 96 kHz le frequenze spurie sono virtualmente scomparse e la distorsione del Cranesong è ancora più pulita.
Il Cranesong è utilizzato per aggiungere della distorsione analogica, la quale può essere molto funzionale se usata per produrre il giusto risultato estetico. Infatti, utilizzando la registrazione digitale ad alta risoluzione, l'attrezzatura dal suono analogico può aggiungere una gradevole colorazione o può essere utilizzata come un' utile copertura, una sorta di cerotto di “fuzz”. Infatti la lavorazione ad alta risoluzione in digitale rivela i dettagli più fini della sorgente, con il rischio che tutti i nei presenti nella registrazione vengano a galla. Un modo per risolvere il problema è aggiungere con un leggero “fuzz” delle delicate armoniche in stile nastro magnetico. Infatti, nel caso in cui la performance non sia la migliore in assoluto, il mix non sia meraviglioso oppure il suono semplicemente è migliore quando non è perfettamente chiaro, allora l'aggiunta di “fatness” e di fuzz mascherante, che a questo punto possono essere definiti come gli ingredienti magici della distorsione analogica, può essere il giusto approccio nella fase di elaborazione. L'ingegnere del mastering Bob Katz preferisce applicare questo processo prima facendo passare il segnale attraverso dell'elettronica ad alta risoluzione, la quale aggiunge poca o nessuna distorsione, e poi inserire selettivamente un tocco di fuzz con un componente analogico o in alternativa con un dither rumoroso. Questo approccio è metodico, controllabile e reversibile. Chiaramente, l'uso artistico del rumore può mascherare e quindi attenuare la distorsione di basso livello. Ironicamente, quindi, il livello molto basso di rumore nella registrazione digitale può essere il suo più grande nemico.
Il Cranesong può mimare tre tipi di distorsione analogica, che vengono utilizzati per aggiungere il “fuzz” che attenua la freddezza digitale: essi sono chiamati Triode, Pentode e Tape. Il controllo detto Triode aggiunge frequenze di seconda armonica pura ed è utilizzata per ripulire le basse frequenze, aggiungendo definizione al basso.
Il Pentode ai livelli più bassi aggiunge armoniche di terza e di quinta, producendo un boost nella presenza e nella brillantezza inserendo pochi artefatti digitali, specialmente a 96 kHz.

Politecnico di Torino



Confronto del Cranesong HEDD 192 in modalità Pentode a due sample rate differenti, con un segnale di test a 10 kHz e con un livello di -15 dBFS. Sopra il sample rate è di 44,1 kHz, sotto è di 96 kHz. Si noti la differenza nella scala delle frequenze, poiché il sample rate maggiore mostra le frequenze armoniche fino a 48 kHz. Fonte: Mastering Audio: the Art and the Science

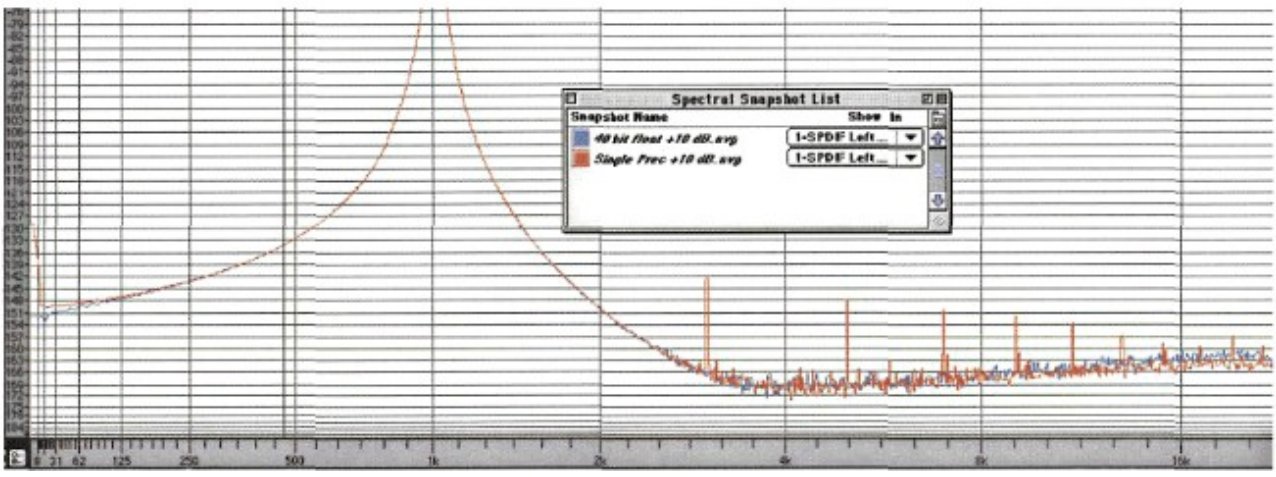
Alle frequenze più alte le armoniche dispari addizionali aggiungono un po' di fatness. Oltre alla quinta, piccole quantità di armoniche di settima e nona aggiungono un' acutezza desiderabile.
La modalità Tape del Cranesong inserisce un timbro che può essere definito dolce e, quando utilizzata in combinazione al Pentode, ne può addolcire la brillantezza. La terza armonica (un' ottava più una quinta) addolcisce e ingrassa il suono emulando il nastro analogico e ciò viene apprezzato in fase di mixing per riuscire ad “incollare” tra di loro gli strumenti che compongono un brano. Il Tape può aiutare le sorgenti mixate digitalmente, che possono essere anche registrate bene ma a cui manca un po' di fatness “rock 'n roll”. Questa modalità produce distorsioni di seconda e terza armonica, ma anche armoniche di grado più alto, emulando la performance del nastro analogico. Un uso eccessivo può rendere il risultato finale confuso e pesante per le orecchie, una qualità raramente desiderabile. Solo una leggera quantità può funzionare come

Politecnico di Torino

attenuante per le registrazioni troncate e taglienti, considerando che se c'è stato un troncamento prima dell'uso del Cranesong, lo spazio e la profondità del segnale sono stati permanentemente persi.

5.6.7 Limiter

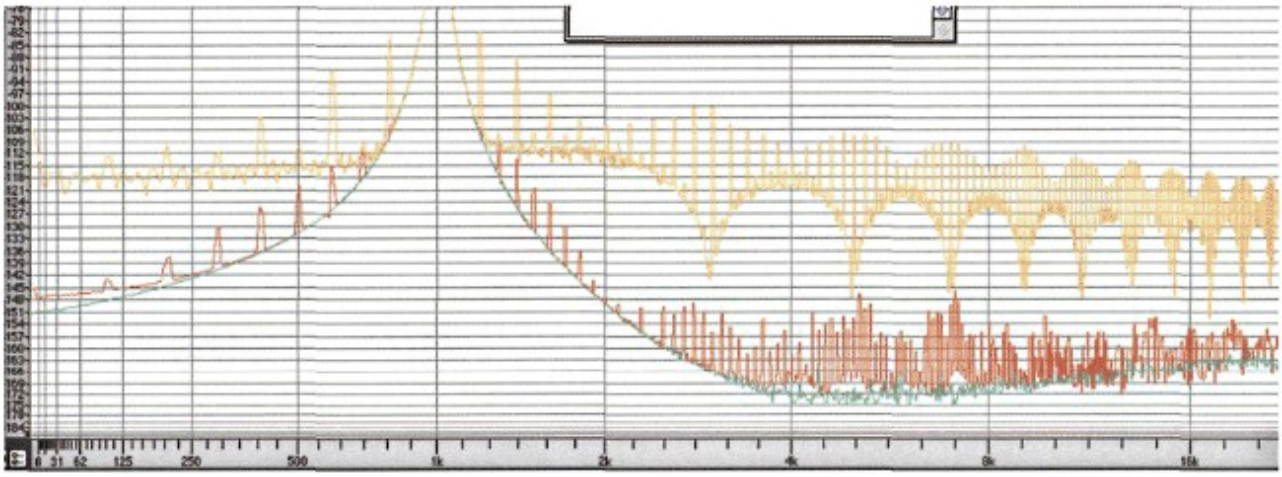
I processori digitali a precisione singola hanno dato all'elaborazione digitale una brutta nomea, per questo i processori a 24 bit a precisione singola stanno scomparendo dagli studi professionali mentre proliferano nel mercato consumer. I processori a 32 bit a virgola mobile sono generalmente considerati inferiori dal punto di vista del suono rispetto ai 48 bit a doppia precisione fissa e i 40 bit a virgola mobile. Qualche device a virgola mobile, come il programma software ChannelStrip della Metric Halo, lavora sui 64 bit e ha una distorsione incredibilmente bassa. Lo stesso designer ha prodotto equalizzatori digitali a 32 bit in virgola mobile usando tecniche di riduzione della distorsione che suonano molto bene e ottengono le stesse performance degli equalizzatori che usano lunghezze di parola più lunghe. Si può concludere che la skill del designer è quella che determina quanto bene suonerà l'apparecchio. La matematica utilizzata non è banale e la scelta da parte del designer dei coefficienti per i filtri passa basso può fare la differenza, allo stesso livello della scelta della lunghezza di parola.



Un semplice boost di 10 dB applicato a due differenti tipi di processori.
In rosso, un processore a precisione singola, la quale distorsione è il risultato del troncamento di tutti i prodotti sotto il 24esimo bit; in blu l'output del processore a 40 bit a virgola mobile con un dither di 24 bit applicato all'output.
Fonte: Mastering Audio: the Art and the Science

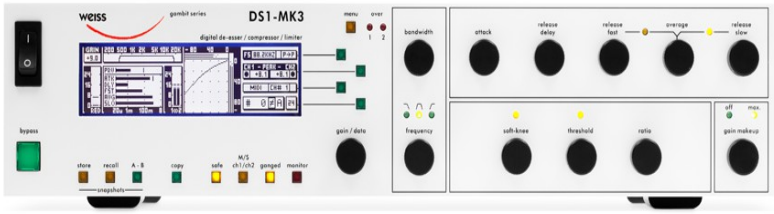
Politecnico di Torino

In figura viene mostrato come, con un processore a precisione singola, anche un semplice boost del gain può generare grave distorsione.
Un tono a 1 kHz, ditherato a 24 bit e al livello di -11 dBFS, viene elaborato attraverso due tipi di processori; entrambi applicano un boost di 10 dB. La distorsione del processore a precisione singola (traccia rossa) è il risultato del troncamento sotto il 24esimo bit., che crea aliasing. Il prodotto più alto della distorsione si trova a -142 dBFS e quindi è estremamente basso nella scala dei decibel; ciò nonostante questa distorsione non armonica è sufficiente da diventare irritante all'orecchio sensibile.
In blu è rappresentato l'output perfettamente pulito di un processore a 40 bit a virgola mobile il quale applica il dither al suo output a 24 bit. Performance simili possono essere misurate con un processore a 48 bit a doppia precisione e un processore a 32 bit a virgola mobile, ai quali viene poi applicato il dither a 24 bit.
Gli equalizzatori e i processori di dinamica digitali più avanzati utilizzano la tecnologia del doppio campionamento, cioè il rate di sampling interno viene raddoppiato per ridurre la distorsione da aliasing, come visto nel Capitolo 3.



Confronto tra due processori di dinamica digitali di qualità sonora eccelsa, il Weiss DS1-MK2 in sovracampionamento (traccia verde), il quale usa calcoli a 40 bit in virgola mobile e il Waves L2 con il sampling standard (rosso), il quale usa 48 bit a precisione fissa. Il Weiss, utilizzato nella modalità Safe limiter, la quale non è sovracampionata, è mostrato in arancione.
Fonte: Mastering Audio: the Art and the Science

La figura soprastante compara due eccellenti processori di dinamica digitali, il Weiss DS1-MK2, il quale usa calcoli a 40 bit a virgola mobile e in sovracampionamento, e il Wave L2, il quale usa 40 bit a precisione fissa con il campionamento standard.



Weiss DS1-MK2

Waves L2

Si noti come il processore in sovracampionamento (il Weiss) esibisce una distorsione di quantizzazione considerevolmente più bassa.
A 88,2 kHz e oltre (non raffigurato), il limiter Weiss e il Waves raggiungono performance migliori e il doppio campionamento potrebbe non essere necessario. Ciononostante c'è sempre un vantaggio considerabile nel fare l'elaborazione a frequenza più alta, poiché si muovono i prodotti della distorsione nello spettro del non udibile (oltre i 20 kHz) per poi convertire, durante l'ultimo passaggio nella catena elettroacustica, il sample rate a 44,1 kHz e filtrare via la maggior parte del rumore che è stato distribuito alle alte frequenze.
Nonostante le distorsioni misurate, la finestra di osservazione scelta probabilmente ha poco a che fare con le performance percepite da questi due limiter, dotati di un suono eccellente. Infatti questo tipo di misurazioni non hanno relazione con la performance udibile dei limiter: infatti il fattore più importante per l'orecchio è la durata dell'azione del limiter.
Nel suo uso tipico, i limiter vanno nella fase di riduzione del gain per un periodo molto breve. Al ratio di limiting di 1000:1, con attacco istantaneo e release veloce, questi processori producono una distorsione momentanea che è più corta della sensibilità dell'orecchio umano (intorno ai 6 ms). Ovviamente se si sovraccarica il limiter in modo tale che lavori sia ai livelli RMS (cioè ai volumi

medi) così come con i picchi, allora la distorsione della sinusoide presa in considerazione diventa significante.
I compressori sono differenti e il campionamento doppio è critico per essi, perché il compressore può andare nella fase di riduzione del gain per una buona percentuale di tempo: in questo caso il doppio campionamento contribuisce al suono robusto e caldo del Weiss quando viene usato come compressore. Ciò spiega perché di solito nella registrazione dei dischi Heavy Metal, che usano come effetto sonoro una quantità considerevole di distorsione, vengano utilizzati dei compressori analogici, proprio con lo scopo di evitare aliasing inarmonici che sono tipici dei processori digitali.

5.6.8 Conclusioni

Fino ad un paio d'anni fa, il suono che si otteneva dal passaggio tra vari apparecchi digitali non era molto apprezzato, a causa dell'immaturità nel design dei filtri e dei circuiti che accumulavano errori e distorsioni inarmoniche; erano tollerabili solo un paio delle migliori unità a precisione singola in serie. Dopo ciò, si passava il segnale di nuovo all'analogico.
Se si processa in digitale bisogna essere consapevoli delle sue debolezze e, finché non si adottano processori prestanti con una certa potenza di elaborazione, bisogna limitare il numero di passaggi attraverso qualunque sistema digitale. Ogni passaggio infatti suonerà un po' più freddo anche usando la quantizzazione di livello alto. Un mix fatto attraverso una console digitale può suonare meglio di uno fatto attraverso una console analogica ad alta qualità in base a vari fattori: il numero di passaggi e rimbalzi nella catena elettroacustica, il numero di tracce che vengono mixate, la qualità dei convertitori che vengono utilizzati, l'utilizzo di attrezzatura out of box, la qualità del mixing interno e gli algoritmi di equalizzazione nella console digitale. Mentre nessuna console ha un equalizzatore che ha la potenza di un Weiss da 6000 €, economicamente è molto più semplice replicare un buon algoritmo di equalizzazione per 144 canali di una DAW che implementarne uno equivalente con del hardware analogico.
Comunque, gli apparecchi digitali dotati di sample rate elevati stanno prendendo il sopravvento e continueranno a suonare meglio.
Quindi per l'elaborazione audio si può optare per tre configurazioni, ognuna con i suoi pro e i suoi contro:

1) processori digitali molto trasparenti, con rumore basso e con un suono puro
2) processori analogici di buona qualità, abbastanza trasparenti, con rumore basso che vengono utilizzati per aggiungere un po' di colorazione o semplicemente per prevenire che il suono diventi troppo “freddo”.
3) processori digitali che simulano la distorsione analogica e il suo calore.

Politecnico di Torino

Gli elaboratori di audio digitale di buona qualità comunque rimangono molto cari, perchè possono servire fino a 5 anni uomo per produrre del buon software audio e c'è la necessità di molta esperienza e sperimentazione.

5.7 Gamma Dinamica

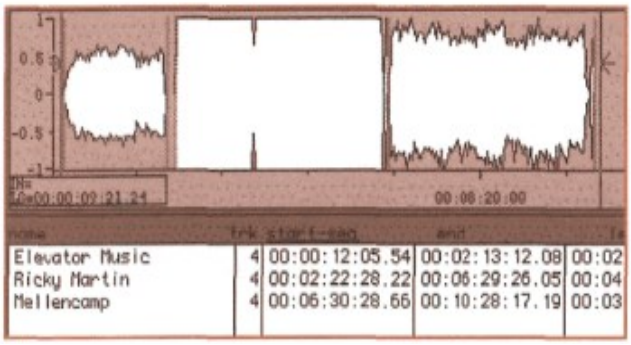
Il segnale digitale ha molto più potenziale rispetto all'analogico per quanto riguarda la gamma dinamica. Infatti l'SNR del medium analogico è all'incirca di 70 dB, considerando sempre un disco tenuto bene e con un buon sistema giradischi-puntina, mentre per un CD a 16 bit con sample rate di 44,1 kHz e quantizzazione lineare si ha una gamma di 96 dB. Ciò permette di collocare più informazione utile in termini di volume all'interno del CD. Considerando che il decibel è una grandezza logaritmica si ha che la dinamica del vinile è 10 volte minore di quella digitale. Ma ciò ha portato all'abuso di compressione per sfruttare il volume maggiore del digitale.
La compressione consiste nel prendere le parti a volume più basso di una traccia audio e farle diventare più forti e, viceversa, prendere le parti a volume più alto e diminuirle d'intensità. In tal modo, appiattendo la dinamica del segnale e incrementandone il volume complessivo, si aumenta il volume percepito. Infatti, se si confrontano due brani con gamme dinamiche differenti, entrambi possono avere lo stesso livello di picco nel mixer in termini di decibel, ma uno può suonare all'orecchio come più forte rispetto all'altro, proprio grazie all'effetto di compressione dei suoni. Una traccia audio compressa risulta quindi più forte rispetto ad una con una gamma dinamica più ampia. Con l'arrivo del digitale, che ha meno limiti di dinamica rispetto al vinile e ha un noise floor molto basso, si ebbe una escalation nell'utilizzo della compressione durante la fase di mastering, con lo scopo di sfruttare al massimo la gamma dinamica e creare brani che potessero competere e superare il volume degli altri brani in circolazione nelle stazioni radio: questa corsa verso “il volume più forte” viene detta loudness war. Questa corsa iniziò con l'avvento del CD negli anni '80 e incrementò il divario percettivo tra la musica su CD e la musica in vinile, poiché il master per CD e il master per vinile potevano divergere proprio per l'utilizzo di questa tecnica, contribuendo alla convinzione che il vinile suonasse meglio rispetto al corrispettivo digitale.

Politecnico di Torino

6 Loudness War

6.1 Introduzione

La loudness war è un trend che riguarda l'aumento dei livelli di volume percepito nella musica registrata la quale riduce la fedeltà audio e, secondo molti, riduce l'apprezzabilità della musica da parte dell'ascoltatore. Il picco del livello massimo di intensità delle registrazioni analogiche è limitato dalle specifiche tecniche del vinile, come la larghezza dei solchi e il rumore di superficie, e dalle dispersioni della catena elettroacustica tra la sorgente e l'ascoltatore, in questo caso del sistema giradischi e puntina. Con l'avvento negli anni '90 della elaborazione del segnale digitale e del Compact Disc, la musica poteva essere codificata in un formato senza perdita d'informazione che permetteva un volume molto più alto rispetto all'analogico.

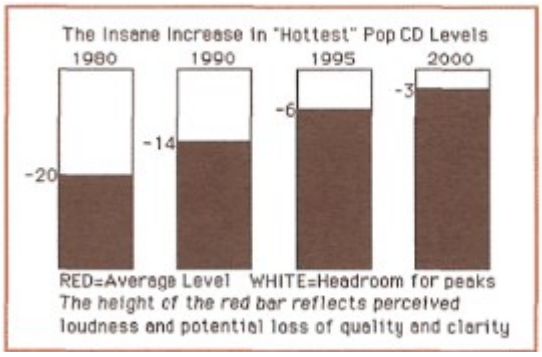


Confronto del volume e della compressione presenti in diversi brani musicali: il primo è musica di sottofondo, il secondo è un singolo di Ricky Martin, il terzo un brano rock.
Fonte: Mastering Audio: the Art and the Science

Una volta che la massima ampiezza del CD viene raggiunta, cioè i picchi del segnale audio raggiungono il limite superiore della rappresentazione digitale, detto Full Scale, la loudness può essere ancora aumentata attraverso tecniche di elaborazione del segnale come la compressione dinamica e la equalizzazione multibanda.
La compressione dinamica consiste nel diminuire i picchi presenti nel segnale, cioè diminuire la variazione dinamica tra il volume più alto e quello più basso all'interno di una traccia musicale;

Politecnico di Torino

facendo così, si può innalzare il volume del segnale rimanente, aumentando il volume percepito ma con meno escursioni tra le parti tranquille e le parti forti.
In figura viene comparata la gamma dinamica di tre brani: il primo è un brano di sottofondo per ascensori; il secondo è un singolo pop dei primi anni 2000, dove si può notare un eminente clipping per tutta la durata del brano e dove non c'è quasi nessuna escursione dinamica; il terzo è un brano rock che non ha subito una compressione eccessivamente aggressiva.
Il problema psicoacustico risiede nel fatto che quando due programmi musicali identici sono presentati con due loudness leggermente diverse, la più forte tra le due viene spesso definita "quella che suona meglio", ma solo nei termini di un ascolto a breve termine. Fin dalla loro introduzione, i livelli di loudness dei CD sono aumentati in maniera progressiva fino a quando la qualità sonora è diventata così bassa che tutti poterono percepirla; i risultati di questa escalation furono dei dischi di musica pop-mainstream con una escursione dinamica massima di 15 dB e dell'audio che, nonostante fosse registrato a 96 kHz e con profondità di 24 bit, hanno come gamma dinamica 1 bit.



Aumento dagli anni '80 ai primi 2000 del livello medio RMS e la conseguente diminuzione di headroom.
Fonte: Mastering Audio: the Art and the Science

Gli ingegneri del suono possono applicare alla registrazione una quantità di compressione sempre maggiore e alzare il volume finché tutta la traccia musicale non raggiunge il Full Scale, creando un brano con volume medio molto alto e con poca escursione tra i picchi e le valli del segnale. In casi estremi, gli sforzi utilizzati per aumentare la loudness possono trasformarsi in una prominente distorsione digitale, il cosiddetto hard clipping, generando una distorsione udibile che risulta fastidiosa e inaccettabile nell'ascolto a lungo termine.
Questo sfruttamento dell'elaborazione dell'audio è dovuta al mercato discografico, dove si svolse una crescente competizione nel realizzare brani che suonassero forti come gli altri brani presenti

sul mercato, scatenando una vera e propria guerra di volumi. Essa ha influenzato la produzione di molti album musicali, considerati delle vittime della loudness war.

6.2 Storia della Loudness War

6.2.1 Escalation

Aumenti nella loudness possono essere rintracciati fin dagli anni '40 con i singoli da 45 giri utilizzati nei jukebox: infatti un produttore discografico aveva la preoccupazione che il suo prodotto non suonasse altrettanto forte rispetto agli altri dischi presenti nell'apparecchio, esigendo quindi che il master inciso sull'acetato venisse ritoccato per renderlo più appetibile e catturare l'attenzione dell'ascoltatore. Questa tecnica veniva utilizzata anche per stimolare l'attenzione dei direttori delle radio e permettere alla canzone di competere con gli altri brani presenti nella programmazione.
Un altro antesignano della loudness war è la formula del Muro del Suono: essa è una tecnica di mixaggio inventata da Phil Spector negli anni '60, dove gli strumenti vengono mixati in modo tale da occupare tutta l'immagine sonora disponibile con lo scopo di creare una traccia di maggior impatto sonoro, utilizzando in maniera creativa una serie di tecniche di mixing come il dubbing (cioè il raddoppio) degli strumenti, la riverberazione e la compressione.
Un altro predecessore si può trovare negli studi della casa discografica della Motown, che produceva delle uscite discografiche che spingevano sempre più in là i limiti di volume che potevano essere incisi su vinile; essa era conosciuta per tagliare i 45 giri più “caldi” dell'industria attraverso una tecnica detta “Loud and Clear”, che permetteva di aumentare la loudness apparente senza creare distorsione sul supporto vinilico. La tecnica consisteva nel mantenere il minutaggio dei brani sotto i 3 minuti, in modo tale che si potesse incidere sul vinile dei solchi più larghi e meglio distanziati per evitare di incorrere in distorsioni; venivano poi applicate delle curve di compressione in base al crescendo dei volumi, abbassando il livello massimo dei picchi di circa 0,5 dB, permettendo di aumentare leggermente il volume totale del brano. Durante il mastering venivano anche applicate delle equalizzazioni intorno agli 8 kHz-10 kHz per compensare la distorsione di tracing durante il mastering. Le frequenze basse venivano tagliate a 70 Hz da un filtro passa alto, come da norma nell'incisione della lacca, ma per recuperarne la percezione venivano aumentate le seconde armoniche dei bassi: ciò era molto funzionale al loro stile musicale Soul e RnB, dove le batterie e i bassi sono molto rilevanti, creando dei prodotti che erano unici nel loro genere. Ciononostante, a causa delle limitazioni del formato vinilico, l'abilità di manipolare la loudness rimaneva limitata. Tentativi di ottenere una loudness estrema potevano rendere il medium irriproducibile a causa della larghezza dei solchi che non potevano superare una certa

soglia d'ampiezza, pena il salto della puntina o la produzione di distorsione. I media digitali rimossero queste restrizioni e come risultato si poté ottenere un volume percepito più alto.
Gli effetti utilizzati nel audio digitale per l'elaborazione del segnale permisero agli ingegneri del mastering di avere un controllo diretto e quasi illimitato sulla loudness di una canzone: per esempio, un limiter, che è un compressore con una ratio di compressione infinita, può ridurre la gamma dinamica in maniera estrema ed arbitraria, permettendo poi di aumentare il livello sonoro nel mix e farlo arrivare il più possibile vicino al limite digitale.
Gli stadi dell'aumento della loudness dei CD sono spesso divisi in base alle decadi dell'esistenza del supporto. Fino alla fine degli anni '80, poiché il CD non era il supporto primario per la musica popolare, vi era poca motivazione per la competizione in termini di volume e guerra di loudness.
La pratica più comune per masterizzare musica per CD consisteva nel far combaciare il picco massimo della registrazione digitale con il livello di saturazione di una registrazione analogica; infatti la Digital Full Scale ai tempi era poco familiare agli ingegneri del mastering e si lavorava utilizzando i sistemi VU.

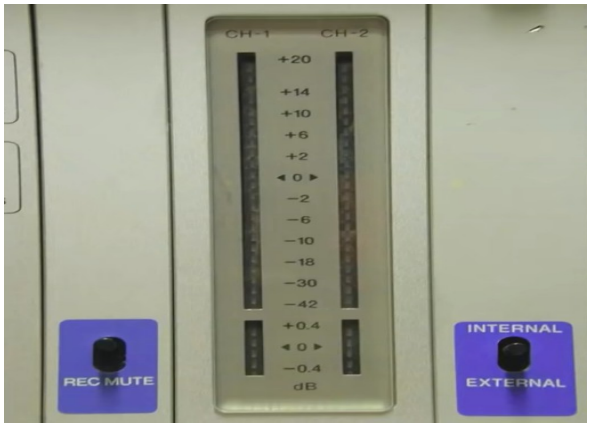
VU Meter
Fonte: Mastering Audio: the Art and the Science

Usando i VU meter nella masterizzazione nel digitale, il punto indicato come 0 VU nella scala dei valori corrispondeva all'incirca a -14 dB rispetto alla massima amplitudine del CD. Visto che lo 0 VU veniva usato come punto di saturazione, ciò lasciava della headroom di svariati decibel che permetteva di gestire gli eccessi oltre il punto di saturazione (spesso riferita come zona rossa, rappresentata dalla barra rossa nella scala del VU meter). Rimane comunque da considerare che il VU meter è un metodo di valutazione della loudness abbastanza primitivo, principalmente per due motivi: il primo è che la lancetta è stata progettata per funzionare in maniera corretta solo con la

Politecnico di Torino

voce, a causa del suo tempo di risposta di 300 ms che la fa reagire bene alle sillabe ma non è accurata per tutti i tipi di suoni; di conseguenza gli utilizzatori esperti sapevano che la maggior parte delle volte se vedevano, ad esempio, un salto repentino e fulmineo da 0 a +3 VU probabilmente non ci sarebbe stata alcuna distorsione e la variazione in loudness sarebbe stata impercettibile; in più la risposta in frequenza del VU è piatta, quindi reagisce a tutte le frequenze indipendentemente dalle loudness relative, determinate dai diagrammi di Fletcher e Munson: un basso di un pezzo reggae può saturare nel VU Meter, ma il volume percepito può essere molto minore. Di conseguenza non si ragionava per livelli massimi e picchi ma per livelli medi e si aveva la tendenza a lasciare dello spazio al di sopra del segnale, per evitare di incorrere in distorsioni impossibili da prevedere. Il livello RMS medio di una canzone rock durante la maggior parte della decade era intorno ai -16,8 dBFS. Negli anni '80 i primi CD venivano registrati a -20 dB sotto il Full Scale e suonavano molto tranquilli e anche molto naturali.

La console Sony Pcm fonte:Loudness: War & Peace



Dettaglio del meter della Sony Pcm. Fonte: Loudness: War & Peace

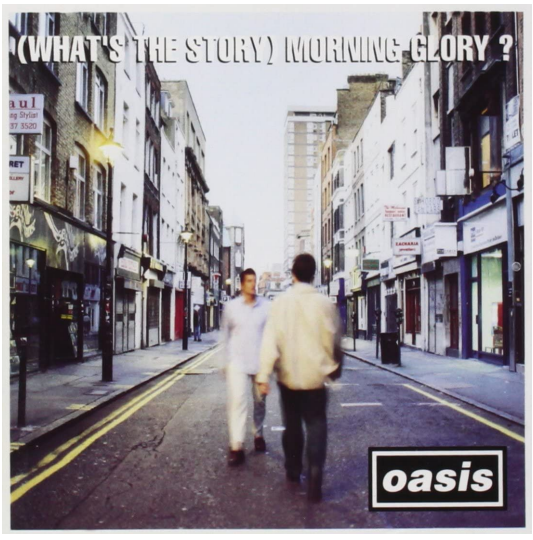
Politecnico di Torino

Si osservi la parte frontale del Sony PCM 1810, una delle prime console digitali. A quei tempi non c'erano le DAW e neanche i Digital Signal Processor e tutto doveva essere registrato in presa diretta sul video tape; i primi PC che si affacciavano sul mercato avevano una quantità di RAM che oscillava tra i 4 kB e i 64 kB e nessuno di essi poteva registrare audio di alta qualità per più di qualche secondo; gli hard disk capienti non erano disponibili.
Il meter del volume di questa console è la cosa più interessante. Ha una scala molto strana, diversa da quella dell'audio digitale. L'asintoto superiore 0 dB si trova a -20 dB sotto il picco della scala. Vi è poi un meter separato destinato al settaggio fine, utilizzato per calibrare la console con un tono di test mandato dalla console al valore nominale di 0 VU. Con questo prodotto la Sony sottintendeva che il livello medio RMS è quello di cui bisognerebbe preoccuparsi, lasciando che i picchi cadano dove vogliono nello spazio lasciato per la headroom. Dalla console successiva, la Sony cambiò la scala, forse perché stavano già cambiando le abitudini di mixing.
Ben presto infatti gli ingegneri del suono scoprirono la normalizzazione dei picchi, con la nascita di registrazioni dove vi erano un paio di transienti che raggiungevano il Full Scale. Il prossimo passo fu la scoperta che registrando su nastro analogico si avrebbe avuto una compressione dei livelli sopra i +6 dBVU e quindi registrando su nastro si avrebbe avuto un volume percepito maggiore. Di conseguenza, per competere con la loudness percepita, gli ingegneri iniziarono ad applicare nel digitale una leggera compressione ai loro CD, in modo tale da competere con il nastro magnetico.

Un catalizzatore della loudness war fu l'invenzione dei CD Player per le auto. Infatti, per sentire i passaggi più quieti di un CD nello stereo della macchina, che è sottoposta ai vari rumori di fondo durante la guida, si dovettero alzare le parti più quiete attraverso l'aumento di compressione e, di conseguenza, l'aumento del livello medio della traccia e della loudness percepita. Ciò faceva raggiungere il Full Scale ai picchi del segnale, e quindi si iniziò ad utilizzare un limiter analogico per normalizzare in maniera delicata i picchi del segnale.
Nei primi anni '90, la loudness war non era ancora del tutto cominciata, ma oramai gli ingegneri del mastering avevano imparato ad ottimizzare la gamma dinamica per il medium digitale. Iniziarono ad affacciarsi dei progetti musicali con un volume percepito mediamente più alto e con dei livelli medi che diventavano sempre più prossimi al limite digitale, col risultato che i picchi nella dinamica di una canzone rock o di una canzone pop ritmicamente ricca raggiungevano sempre più di frequente il Full Scale. Il concetto di realizzare delle uscite discografiche “più calde” iniziò dunque ad affascinare i professionisti all'interno dell'industria: ciò era dovuto al fatto che alcune uscite discografiche erano notevolmente più forti in termine di volume, creando uno standard competitivo di intensità sonora. Vi era anche la convinzione che i consumatori avrebbero preferito dei CD con dei volumi più consistenti che, nei contesti di centri commerciali e di radio commerciali, dove l'ambiente sonoro risulta saturo e vi è la necessità di catturare l'attenzione dell'ascoltatore-consumatore. Era considerata una buona tecnica di “guerilla marketing” puntare su un volume percepito molto alto, anche correndo il rischio di risultare fastidiosi. Di conseguenza ingegneri,

Politecnico di Torino

musicisti ed etichette discografiche svilupparono in maniera separata la propria idea su come rendere un CD più intenso in volume. Nei primi anni '90 arrivarono i primi DSP a buon mercato, che permettevano di applicare compressioni digitali molto aggressive, accelerando la loudness war. Nel 1994, il primo limiter a “muro di mattoni” con il look-ahead, il Waves L1, fu immesso sul mercato: da allora questo processore di dinamica divenne lo standard dei limiter digitali per il mastering. In pratica si tratta di uno dei primi limiter che riusciva a comprimere il suono praticamente all'istante, senza alcuna latenza: ciò è dovuto al fatto che lavora con una duplicazione del segnale che viene letta in anticipo per permettere di attivare la compressione al massimo delle proprie capacità su tutto il segnale audio, a prescindere da quanto siano lunghi i transienti. In più si iniziò ad applicare una equalizzazione multi banda aggressiva sullo spettro superiore del segnale, poiché, se si rende una traccia stereo più brillante sulle alte frequenze, dal punto di vista psicoacustico essa suonerà più forte rispetto ad una senza questo tipo di equalizzazione. Ma non è detto che il risultato sonoro sia migliore e accettabile per un ascolto a lungo termine. Un altro catalizzatore in questa gara di volumi furono i CD dedicati alle Greatest Hits poiché se i vecchi brani che venivano ripresi suonavano meno forti delle hit più moderne essi venivano compressi in maniera decisa. Uno degli album che diede una brusca impennata al volume contenuto all'interno di un CD, innescando definitivamente la loudness war, fu (Whats's the Story) Morning Glory? degli Oasis, il cui livello RMS era di -8 dBFS nella maggior parte delle tracce, che ai tempi era una cosa abbastanza rara, considerando che Appetite for Destruction dei Guns N' Roses aveva una RMS di -14 dB ed era già di per sé considerato molto rumoroso. L'album degli Oasis ottenne un successo clamoroso negli UK poiché la sua loudness così prominente lo rendeva perfetto per l'ascolto in auto e nei pub, dove il suono rockeggiante e chiassoso riusciva ad emergere in mezzo al rumore di fondo.



Politecnico di Torino

Cover di What's The Story Morning Glory

Cover di Californication dei Red Hot Chili Peppers

Un'altra pietra miliare della musica e della loudness war è rappresentato da Californication dei Red Hot Chili Peppers, dove vi è addirittura un prominente clipping durante tutta la durata dell'album.
Nei primi anni 2000, la loudness war ormai si era diffusa capillarmente e poteva essere percepita specialmente nelle release rimasterizzate.
L'ultimo acceleratore nella loudness war fu l'Ipod. A causa del suo uso basato sulla mobilità e l'uso di cuffiette poco isolanti, aveva la necessità di musica che suonasse forte e molto compressa per coprire il rumore dell'ambiente circostante, stimolando l'utilizzo di tecniche ancora più aggressive di compressione. Si iniziò ad utilizzare anche il limiter digitale per i picchi, riducendo incredibilmente la gamma dinamica, creando dei brani praticamente senza distinzioni tra parti tranquille e parti forti e caratterizzati da una seria distorsione digitale. Ma di fatto questo era ancora del "leggero" clipping digitale, che comunque infrangeva le leggi dell'audio digitale registrando oltre la Full Scale.
Dal 2005 fino all'avvento dello streaming divenne prassi comune applicare il clipping analogico, che consiste nell'alimentare un convertitore A/D con una sorgente già in clipping e poi farla ripassare attraverso un D/A giusto per ottenere ancora più volume percepito e di conseguenza distorsione. Ciononostante l'uso della distorsione aggressiva creò dei nuovi generi musicali basati su di essa: l'Alternative rock, il Grunge, l'Hip Hop Hardcore e l'Heavy Metal sono tutte espressioni musicali che utilizzano la distorsione in maniera creativa. Va comunque detto che l'uso della distorsione è diverso dalla ipercompressione e dell'abbattimento della dinamica, brani dotati di

ampia distorsione, come “Smells Like Ten Spirit” dei Nirvana, che ha come spinta sonica la differenza di intensità tra la tranquillità catartica delle strofe e la esplosività del ritornello, hanno una gamma dinamica più ampia rispetto ad un brano radiofonico commerciale.
A questo punto della loudness war la differenza in volume tra un brano degli anni '80 e uno dei 2000 era di 20 dB.
Il 2000 fu la fine dell'era CD e l'inizio dell'era del download, del MP3 e delle casse per computer, che sono dotate di scarsa gamma dinamica. Nella prima decade dell'era internet si trova una distorsione che incrementa in maniera esponenziale e ancor meno gamma dinamica; nei generi pop praticamente ogni singolo transiente raggiunge o supera il Full Scale. Per ottenere questo effetto di loudness estrema si fa clippare digitalmente il segnale per poi applicare una compressione parallela aggressiva. Si converte poi in analogico e si esegue un nuovo clipping con una compressione analogica aggressiva. Si fa ripassare in un A/D eseguendo un altro clipping di almeno di 5 dB e poi si usa un limiter per tagliare via i picchi.
La loudness diventa un problema di gusto estetico poiché ogni artista vuole che il suo brano suoni forte come quello in radio, trasformandosi in un gusto musicale e in un'abitudine operativa. Di conseguenza, il mastering viene considerato solamente come il processo in cui si rende più forte un brano.
Nel 2008, le pratiche di mastering a volume alto ricevettero l'attenzione dei media mainstream con l'uscita di Death Magnetic dei Metallica. La versione CD dell'album aveva una loudness media molto alta che spingeva i picchi oltre il punto della rappresentazione digitale, causando distorsione inarmonica. Questo venne evidenziato dai consumatori e dai professionisti dell'industria musicale e fu un tema trattato da molte pubblicazioni internazionali, scatenando un primo dibattito sul tema della loudness. Ted Jensen, uno degli ingegneri del mastering coinvolto nelle registrazioni di Death Magnetic, criticò l'approccio utilizzato durante il processo di mastering, alla cui direzione c'era Rick Rubin. Una versione dell'album senza la compressione del range dinamico venne inclusa nei contenuti scaricabili del videogioco di simulazione chitarristica Guitar Hero III. Infatti, dopo un primo no comment da parte della band e del produttore Rick Rubin, venne riproposta per la distribuzione online la versione “Mastered for Itunes”, che molto probabilmente è quella che si può trovare sulle piattaforme di streaming al giorno d'oggi, dove la gamma dinamica venne parzialmente ripristinata.

Cover dell'album Death Magnetic

6.2.2 Streaming e LUFS

Nel 2008, l'ingegnere del mastering Bob Ludwig offrì tre versioni dell'album dei Guns N'Roses Chinese Democracy per l'approvazione del co-produttore Axl Rose e Caram Costanzo, che selezionarono quella con minore compressione, mantenendo la gamma dinamica intatta e accantonando la versione con la maggiore loudness. Secondo il produttore, ciò era dovuto alla rivolta dei fan e della stampa di settore nei confronti delle registrazioni pesantemente compresse, che erano funzionali ad un ascolto a breve termine, ma erano inutilizzabili per quelle a lungo termine, ponendo finalmente le basi per una presa di posizione e un ritorno a creare musica con le dinamiche sotto il livello di picco.
Infatti uno degli album di maggiori successo del 2013 fu Random Access Memories dei Daft Punk, con la stampa di settore che ne applaudiva il suono fantastico, che si rifaceva alle atmosfere analogiche della disco anni '70. Mick Guzauski, l'ingegnere responsabile del mixing, scelse deliberatamente di usare poca compressione e distorsione nel progetto, commentando che l'obbiettivo della crew artistica, composta dai Daft Punk e da vari musicisti dell'era disco, non era quello di fare musica alta di volume ed ipercompressa; per questo motivo l'album suona più naturale, mantenendo comunque un vantaggio nei passaggi radiofonici poiché i brani suonano più pieni ed orchestrali e la compressione utilizzata dalla trasmissione radiofonica compensa un po' la loudness naturale. Nel Gennaio 2014 l'album vinse 5 Grammy Awards, incluso il premio per l'album registrato meglio (Best Engineered Album (Non-Classical)).

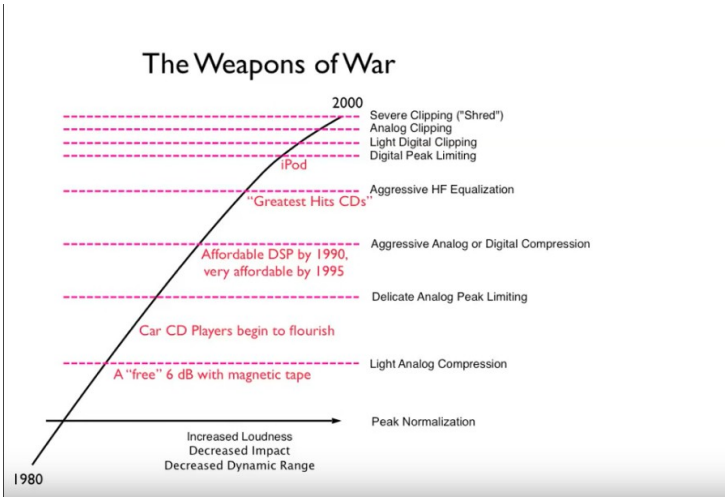
Politecnico di Torino

Cover dell'album Random Access Memories

Le analisi del primo decennio del 2000 suggeriscono che il trend della loudness ha raggiunto un picco intorno al 2005 e successivamente si sia ridotto, con un aumento consistente del range dinamico degli album. Ciò è anche dovuto alla nascita delle piattaforme di download digitale e streaming, che hanno introdotto delle limitazioni alla loudness. Infatti un brano sulla piattaforma può avere solo una certa quantità di volume e tutte le eccedenze vengono limitate e riportate ad un valore medio. Ciò probabilmente spingerà produttori ed ingegneri a creare dei master con all'interno più dinamica, visto che la loudness verrà decisa in maniera arbitraria e non ci sarà più spazio per la competizione tra due brani per vedere quale suona più forte. Dall'avvento di Itunes, i i media e lo streaming hanno iniziato ad utilizzare la loudness normalizzata, portando a pensare che ci possa essere un ritorno alla qualità sonora della musica registrata. Spotify e Youtube eseguono lo streaming intorno ai -14 LUFS (loudness units relative to full scale). Se una traccia che viene caricata su queste piattaforme supera il livello di LUFS, essa viene automaticamente abbassata di volume.
Ora come ora i media con la loudness normalizzata hanno in parte frenato il vantaggio competitivo della compressione. Essi rivelano che i master supercompressi suonano piccoli, deboli e distorti.
Alcuni artisti hanno optato per un ritorno ad uno stile con più gamma dinamica, ed il media permette di far coesistere registrazioni compresse con altre con più gamma dinamica. Di conseguenza il limiter applicato ai picchi è diventato meno necessario.

Politecnico di Torino

Il problema che rimane è il cosiddetto vantaggio acustico: l'orecchio vuole che il materiale considerato “soft” suoni soft, vicino quindi alla sua loudness percepita come naturale, cioè le orecchie non vogliono la stessa loudness per piccoli e larghi gruppi di strumenti. La normalizzazione dei picchi esagera il vantaggio acustico. I gruppi piccoli senza batteria, quindi senza picchi, possono essere alzati di 10-14 dB rispetto ai gruppi grandi/rumorosi. I gruppi grandi sono costretti ad utilizzare tutta la headroom necessaria, incoraggiando la compressione. Senza la compressione si otterrebbe che Mozart suona più forte di Beethoven e Carol King suona più forte dei Metallica.
La loudness normalizzata non risolverà questo problema, ma migliorerà la situazione poiché ci sarà solo una differenza di 6 dB tra gli album percepiti come troppo forti o troppo piano, ritornando alle dinamiche dei tempi del LP.

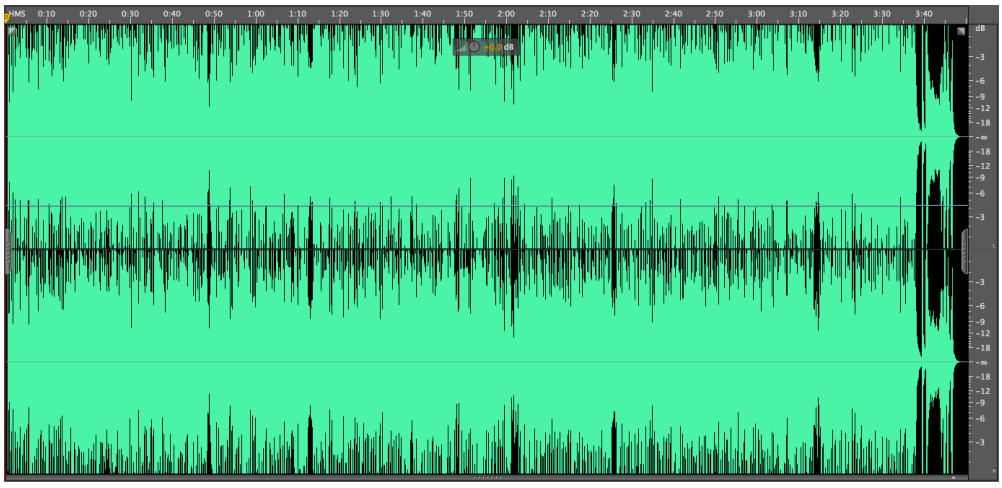


Schema riassuntivo della escalation di loudness.
Fonte: Loudness: War & Peace

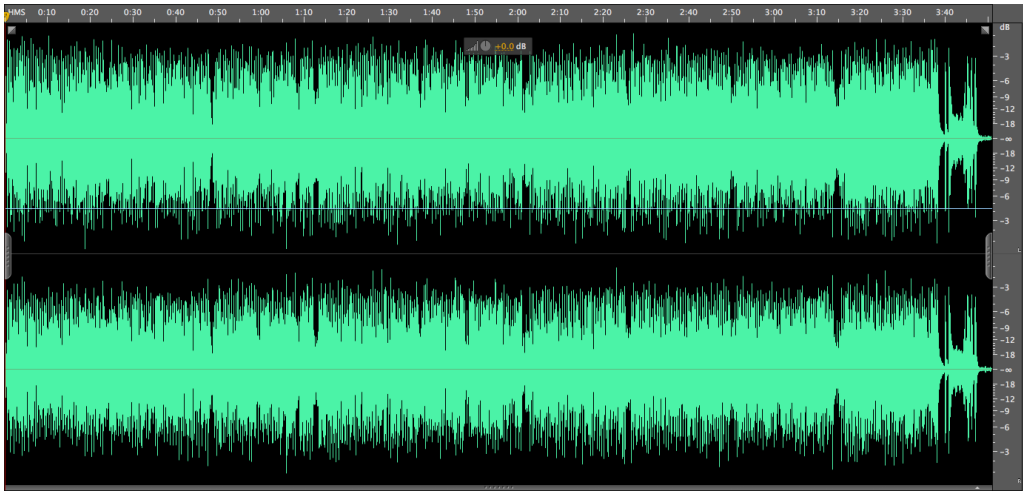
La loudness war può essere considerata come il punto cardine di confronto tra vinile e musica digitale. Infatti lo sfruttamento intenso della gamma dinamica del CD, utilizzando i processori di dinamica, ha portato a creare una differenza sostanziale con la musica incisa su vinile. Ciò è ancora

Politecnico di Torino

più evidente nel caso della stampa dello stesso programma musicale su i due formati, dove le pratiche diverse di mastering e gestione della loudness portano a dei prodotti che, oltre alle differenze intrinseche tra i due supporti, suonano molto diversi.
Se si comparano ad esempio le forme d'onda riguardanti le due versioni dell'album di Bob Dylan del 2006, Together Through Life, ed in particolare la canzone “Beyond Here Lies Nothing”, si può notare che la versione LP ha un range dinamico maggiore di 4 dB rispetto al formato in CD, nonostante quest'ultimo abbia in teoria più capacità nel contenere gamma dinamica.
La media RMS del CD è intorno ai -11 dB, il quale risulta molto alto in volume e molto compresso, provocando un certo fastidio negli ascolti ripetuti e prolungati. Invece la versione su LP, con i picchi normalizzati a 0 dB, ha una RMS media di -15 dB, il che risulta in un ascolto più piacevole e meno soggetto alla fatica d'ascolto.

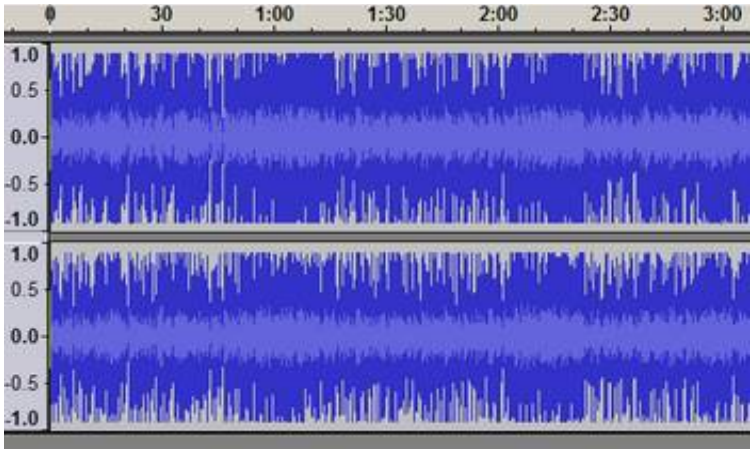


Versione CD del brano “Beyond Here Lies Nothing” di Bob Dylan



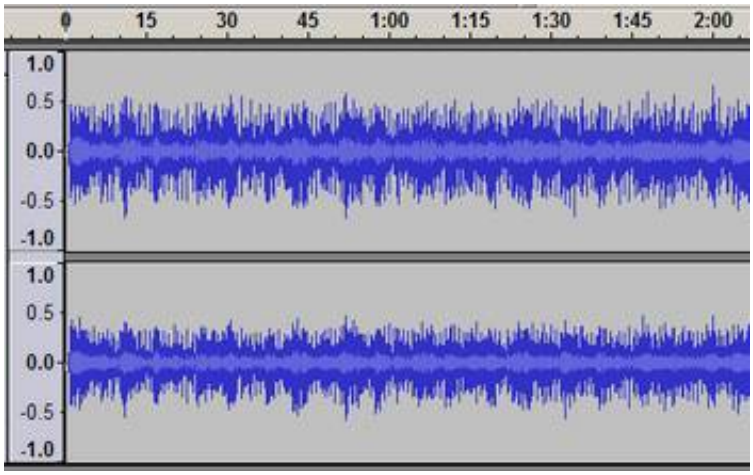
Versione LP del brano "Beyond Here Lies Nothing" di Bob Dylan

Di seguito vengono mostrate le differenze presenti tra la versione LP e la versione CD dell'album Modern Times, con particolare riferimento ai brani "Spirit On The Water" e "I Feel A Change Comin' On", in cui è ancora più evidente la differenza tra le onde sonore del CD e del vinile, con le prime molto più compresse e schiacciate verso il limite Full Scale.

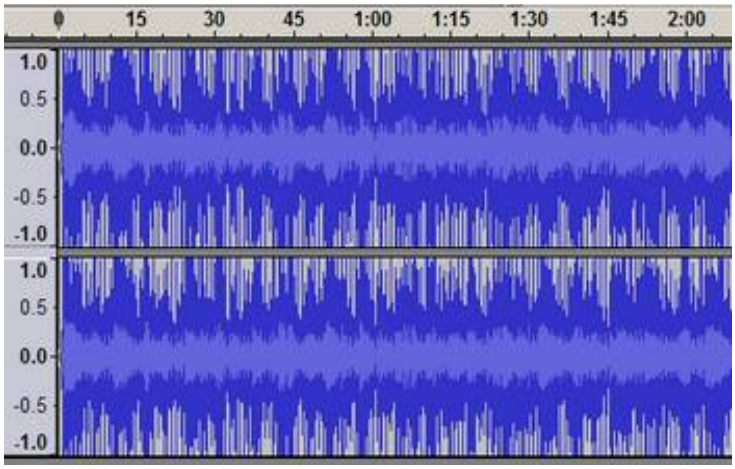


Versione CD del brano "Spirit On The Water" di Bob Dylan

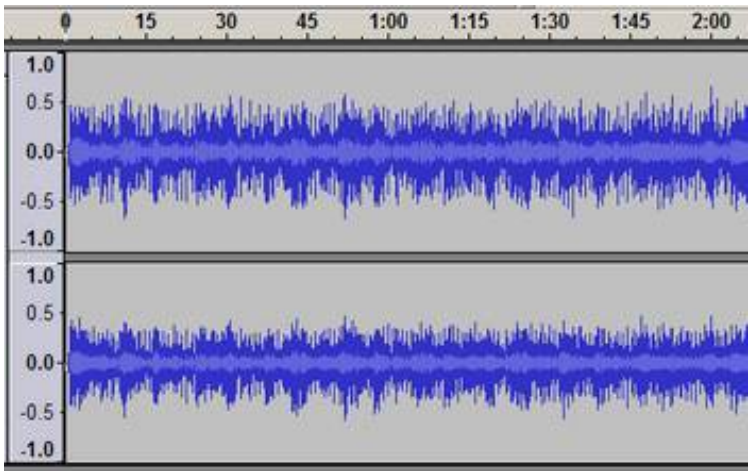
Politecnico di Torino



Versione LP del brano "Spirit On The Water" di Bob Dylan



Versione CD del brano "I Feel A Change Comin' On" di Bob Dylan



Versione LP del brano "I Feel A Change Comin' On" di Bob Dylan

Ciò è un punto cardine nel confronto tra il supporto vinilico e quello digitale: l'ascoltatore medio, che non è al corrente delle differenti tecniche di mastering utilizzate, percepirà i CD, cioè la musica digitale masterizzata con una loudness sproporzionata, come più faticosa e innaturale all'orecchio e sposterà la propria preferenza sul vinile e sul suo apparente calore, dovuto in realtà ad un'appropriata masterizzazione, giudicandolo come un formato superiore, nonostante le differenze sonore in questo caso non siano dovute alle qualità tecniche del formato ma dalle pratiche artistiche e di marketing che vengono applicate in maniera arbitraria dal produttore.
Si può concludere che la loudness war ha portato al netto distacco tra l'esperienza d'ascolto della musica su formato digitale e analogico, influenzando l'opinione del pubblico sulla qualità del digitale e, senza le giuste chiavi di lettura, porta ad una errata interpretazione della disputa tra musica in analogico e musica in digitale.

Politecnico di Torino

Conclusioni

Si può giungere alla conclusione che il formato digitale nella riproduzione è oggettivamente superiore rispetto al formato analogico. Le analisi effettuate nel Capitolo 5 dimostrano che il vinile soffre delle proprie limitazioni tecniche dovute al taglio della lacca e la larghezza dei solchi, le quali sono illustrate nel Capitolo 2, dimostrandosi oggettivamente meno fedele rispetto al corrispettivo digitale.
Il segnale digitale, come dimostrato nel Capitolo 4, riesce a riprodurre in maniera corretta il segnale analogico grazie al teorema di Nyquist e, attraverso le tecniche di oversampling e noise shaping, riesce a cancellare completamente gli artefatti dell'errore di quantizzazione presenti nella conversione.
Si può concludere che il vinile è ritornato in auge per una motivazione soggettiva e un conseguente exploit di mercato, dove l'effetto nostalgia e il romanticismo che si è formato intorno ad esso è stato utilizzato e sfruttato come trend per tenere in vita l'industria discografica, che dall'avvento del digitale e dello sharing peer-to-peer ha iniziato a disintegrarsi. Bisogna aggiungere chee probabilmente non tutta la catena audio utilizzata nei vinili moderni nella loro creazione è completamente analogica, come nel caso dei vinili incisi con la tecnica DMM, e quindi il confronto tra analogico e digitale diventa confusionario; molto spesso le uscite viniliche odierne vengono masterizzate da fonti digitali e non seguono quindi una catena puramente analogica.
La vera differenza sonora tra i due formati è da ricercare nelle pratiche di mastering utilizzate: come spiegato nel Capitolo 6, la loudness war ha cambiato in maniera definitiva il modo in cui vengono prodotte le registrazioni per il digitale e, nonostante vi sia stata un'inversione di rotta con l'introduzione dello streaming, l'abuso di compressione ha portato il digitale a suonare peggio dei vinili, poiché venivano masterizzate due edizioni differenti per i due supporti, alimentando la convizione che l'analogico avesse proprietà sonore che lo rendessero un formato superiore. Lo sfruttamento della gamma dinamica ha portato quindi all'estinzione del CD, facendo preferire in maniera soggettiva il vinile.
L'impiego dell'attrezzatura analogica è ancora parzialmente valido nel campo dell'elaborazione audio di alto livello , visto il design artigianale e accurato utilizzato nel progettare alcuni processori analogici di qualità. Il digitale sta prendendo il sopravvento anche in questo ambito ma l'attrezzatura deve essere di alta qualità per competere con il corrispettivo analogico; il prodotto di registrazione digitale a livello consumatore, visto il design industriale e spesso superficiale del circuito interno, non è comparabile. Nel caso della registrazione e dell'elaborazione audio quindi entrambe le tecnologie risultano valide, analogico e digitale, ma dipende molto dalla qualità dell'attrezzatura utilizzata.

Politecnico di Torino

Bibliografia e Sitografia

[1] Lombardo V. , Valle A. (2008)- *Audio e Multimedia*, terza edizione, Apogeo

[2] Katz B. (2002)– *Mastering Audio: the Art and the Science*, Focal Press

[3] Il disco microsolco

[4] Smith, J.O.(2020) - *Digital Audio Resampling Home Page*
http://www-ccrma.stanford.edu/~jos/resample/.

[5] Home theater hifi - *Shure M97xE Phono Cartridge Review*
https://hometheaterhifi.com/reviews/vinyl/turntable-accessories/shure-m97xe-phono-cartridge-review/

[6] Tears of Rage: the Great Bob Dylan Audio Scandal

[7] Nielsen Music/MRC Data (2020) - *Midyear Report US*

[8] https://www.redbull.com/int-en/music-formats-from-thomas-edison-phonograph-to-spotify

[9] https://blog.landr.com/music-formats-history/

[10] https://www.altriprodotti.com/testina-giradischi-a-bobina-mobile-e-magnete-mobile/

[11] Katz B. - *Loudness: War & Peace*

[12] Vickers E. (2010) - *The Loudness War: Background, Speculation and Recommendations* - 129th Audio Engineering Society Convention. San Francisco: Audio Engineering Society. 8175.

[13] http://floweringtoilet.blogspot.com/2009/04/bob-dylan-together-through-life.html

[14] Devine, K. (2013). Imperfect Sound Forever: Loudness Wars, Listening Formations and the History of Sound Reproduction.

[15] Huber D., Runstein R. (2007) - *Manuale di Registrazione Sonora,* Hoepli

[16]https://www.hi-fiworld.co.uk/vinyl-lp/70-tests/103-cartridge-tests.html?showall=1

[18] https://en.wikipedia.org/wiki/AES3, consultato il 25/05/2021

[19] https://it.wikipedia.org/wiki/Antiskating, consultato il 25/06/2021

[20]https://www.audio-technica.com/en-us/support/audio-solutions-question-week-differences-moving-magnet-moving-coil-phono-cartridges/

[21] https://replicat.com.au/pages/replication-or-duplication

Politecnico di Torino

Ringraziamenti

Ringrazio i miei relatori Marco Masoero e Renato Campajola per la disponibilità e i preziosi suggerimenti.

Ringrazio i miei genitori che mi hanno sempre sopportato (e supportato) durante questi anni di studio, fatti di sudore e pazienza; sono anche riusciti a superare le difficoltà della malattia senza perdere il loro spirito e il loro carisma. Senza di loro non sarei mai arrivato fino a qui.
Grazie.
Ringrazio tutta la gente di Santa Giulia per il calore e l'affetto dimostrato, i miei coinquilini Brownie e Ricky, la Saletta che mi salvato durante questa pandemia, tutti i Misfits là fuori, Nina, la Papaya, Egreen, il Danno, Conway, J Cole, Kendrick, Mr Simpatia, Tyler, YBN Cordae, Evidence, SxM, Mac Miller.