

5

Le meraviglie del Noise Shaping

Qualche lezione fa abbiamo parlato della quantizzazione, ovvero della rappresentazione dei campioni musicali con parole di lunghezza finita. Il problema da combattere quando si introduce questa inevitabile approssimazione è la comparsa del rumore di quantizzazione, un rumore non limitato in banda e correlato con il segnale che determina la massima gamma dinamica ottenibile con la nostra codifica.

Il rimedio più semplice consiste ovviamente nell'usare più bit (diminuendo il rumore di quantizzazione e aumentando la gamma dinamica) e questa soluzione è stata presto adottata nelle apparecchiature professionali e ha poi dato vita alla famiglia dei nuovi formati ad alta risoluzione usati per il DVD-audio. Esiste però un'altra strada percorribile, che sfrutta la non uniforme sensibilità dell'orecchio umano in banda audio per "rimodellare" il rumore digitale migliorando la qualità apparente del suono in barba alle leggi della fisica. Tale tecnica detta **Noise Shaping** ha due applicazioni principali:

- 1) permette di utilizzare rappresentazioni con un solo bit di profondità ed elevata frequenza di campionamento sia all'interno di alcuni modelli di convertitori analogico/digitale e digitale/analogico sia nella famiglia dei formati ad alta risoluzione usati per il SACD.
- 2) permette, dopo una memorizzazione e/o una elaborazione ad alta risoluzione, di convertire il segnale digitale in una risoluzione più bassa, per esempio per stampare un CD audio. Questo secondo tipo di procedimento, particolarmente interessante perché può essere sperimentato da chiunque sappia programmare un normale PC (mentre la costruzione dei convertitori A/D e D/A è riservata a pochi intimi).

Richiami sulla Quantizzazione

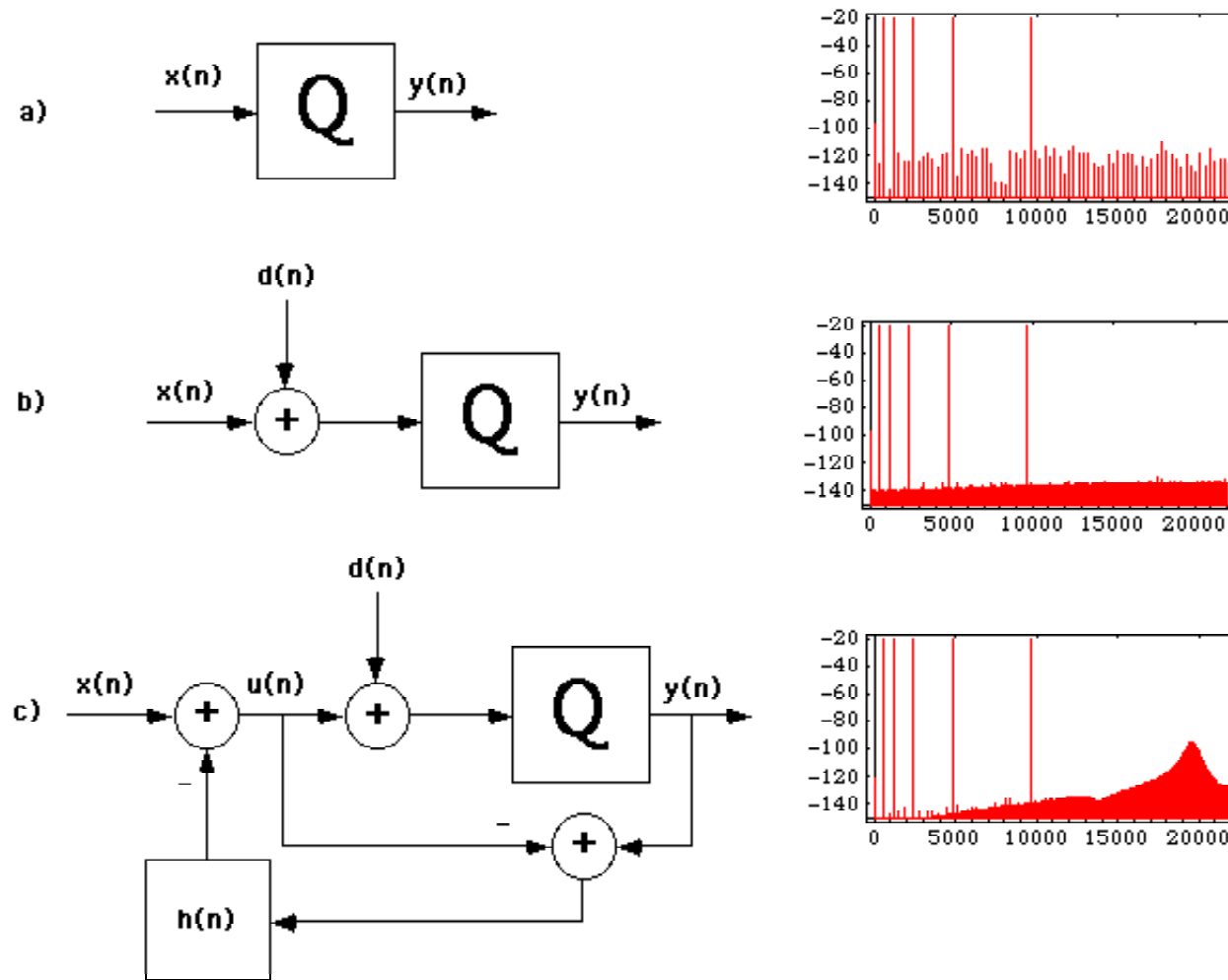
Il **campionamento** di un segnale continuo è una operazione matematica che associa al segnale una sequenza discreta di numeri reali. In pratica non è possibile evitare di far seguire al campionamento un fase di **quantizzazione**, ovvero la rappresentazione dei campioni con una quantità finita di informazione. Come abbiamo già visto si possono fare le seguenti considerazioni.

- Più alta la profondità di quantizzazione minore è la potenza totale del rumore digitale.
- Più alta è la frequenza di campionamento più il rumore si distribuisce ad alta frequenza. Il rumore fuori della banda audio può venire eliminato in un secondo tempo ed è influente all'ascolto.
- In genere il rumore digitale è correlato con il segnale.

In pratica l'ultimo problema viene risolto aggiungendo al segnale un rumore casuale di piccola ampiezza (il **dither**) che se, da un lato, aumenta la potenza totale del rumore, dall'altro lo rende meno fastidioso all'ascolto (è la stessa ragione per cui il rumore termico di un amplificatore a valvole, completamente scorrelato con il segnale, è meno fastidioso delle "interazioni deboli" dovute ad un cattivo cablaggio). Dopo un buon "ditheraggio" il rumore digitale è con buona approssimazione distribuito in maniera uniforme su tutta la banda audio.

Lo stesso problema della quantizzazione si presenta quando un segnale digitale con una risoluzione elevata (per esempio 24 bit) viene riquantizzato (per esempio passandolo ad 16 bit). Dal punto di vista matematico questo equivale ad una quantizzazione ex-novo e si produce un rumore digitale correlato con il segnale. Anche in questo caso si può applicare il dither (chi possiede Wavelab o Cool Edit può trovare le apposite opzioni nella finestra che controlla l'operazione), ottenendo un rumore scorrelato e con distribuzione uniforme. Se

però si introduce un anello di reazione con un filtro digitale è possibile rimodellare il rumore digitale. Le tre possibilità sono illustrate in figura.

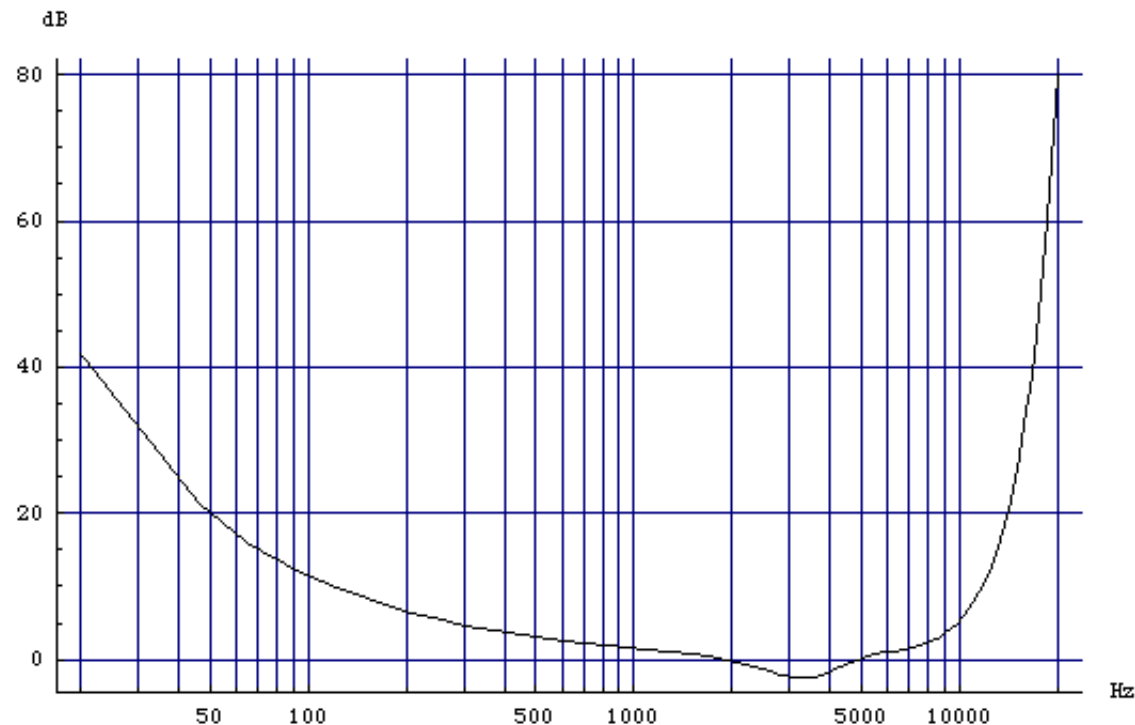


Tecniche di riquantizzazione e spettro del segnale digitale risultante; il segnale originale era costituito da 5 toni puri. a) Solo riquantizzazione a 16 bit, b) riquantizzazione con Dither, c) riquantizzazione con Dither e Noise Shaping.

La percezione uditiva

È ben noto che la sensibilità dell'orecchio umano non è la stessa a tutte le frequenze, le frequenze medie sono percepite a livelli molto inferiori rispetto a quelle basse e alte. Tale livello di percezione dipende pesantemente sia dalle caratteristiche individuali che dall'intensità del suono sotto esame.

Per fissare le idee consideriamo un'approssimazione di tale curva in scala bilogarithmica



La soglia assoluta di udibilità in quiete. Rappresenta il livello in dB SPL a cui un ascoltatore *medio* percepisce un tono puro in assenza di rumore ambientale.

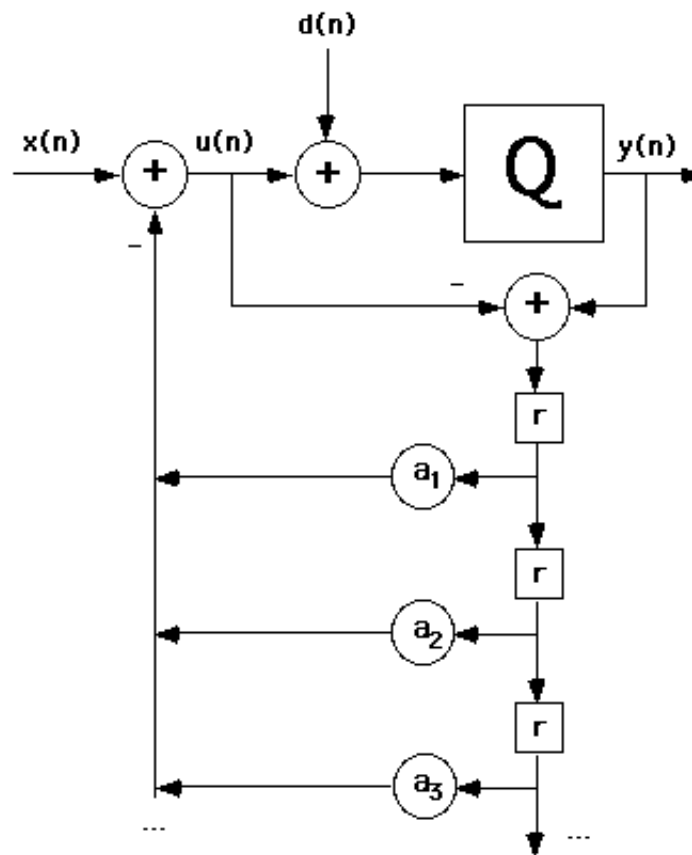
Analiticamente si può scrivere:

$$Q(f) = 3.64 \left(\frac{f}{1000} \right)^{-0.8} - 6.5 e^{\left[-0.6 \left(\frac{f}{1000} - 3.3 \right)^2 \right]} + \left(\frac{f}{1000} \right)^4 \text{ dB SPL}$$

Questa scelta è puramente indicativa, su testi diversi si trovano altre approssimazioni e solo un approfondito studio psicoacustico su numerose cavie umane può suggerire le approssimazioni migliori.

Il Noise Shaping

Prendiamo in considerazione il terzo tipo di riquantizzazione visto sopra ed entriamo in maggiore dettaglio . Il segnale digitale prima della riquantizzazione e del dithering è indicato con $u(n)$, il segnale dopo la riquantizzazione e il dithering con $y(n)$ la differenza tra questi due segnali è il rumore di riquantizzazione $e(n)$. Questo segnale viene filtrato con un filtro **FIR** a coefficienti $a_1, a_2, a_3 \dots$ e sottratto al segnale in ingresso.



Riquantizzazione con Noise Shaping non ricorsivo: all'interno della rete di reazione è usato un filtro FIR con coefficienti $1, -a_1, -a_2, -a_3, \dots$. I blocchi con il simbolo "r" rappresentano un ritardo unitario.

Con un po' di matematica (la teoria della trasformata Z) si ottiene la funzione di trasferimento del sistema $Y(z)$ e la potenza del rumore totale G :

$$Y(z) = X(z) + E(z)(1 - H(z))$$

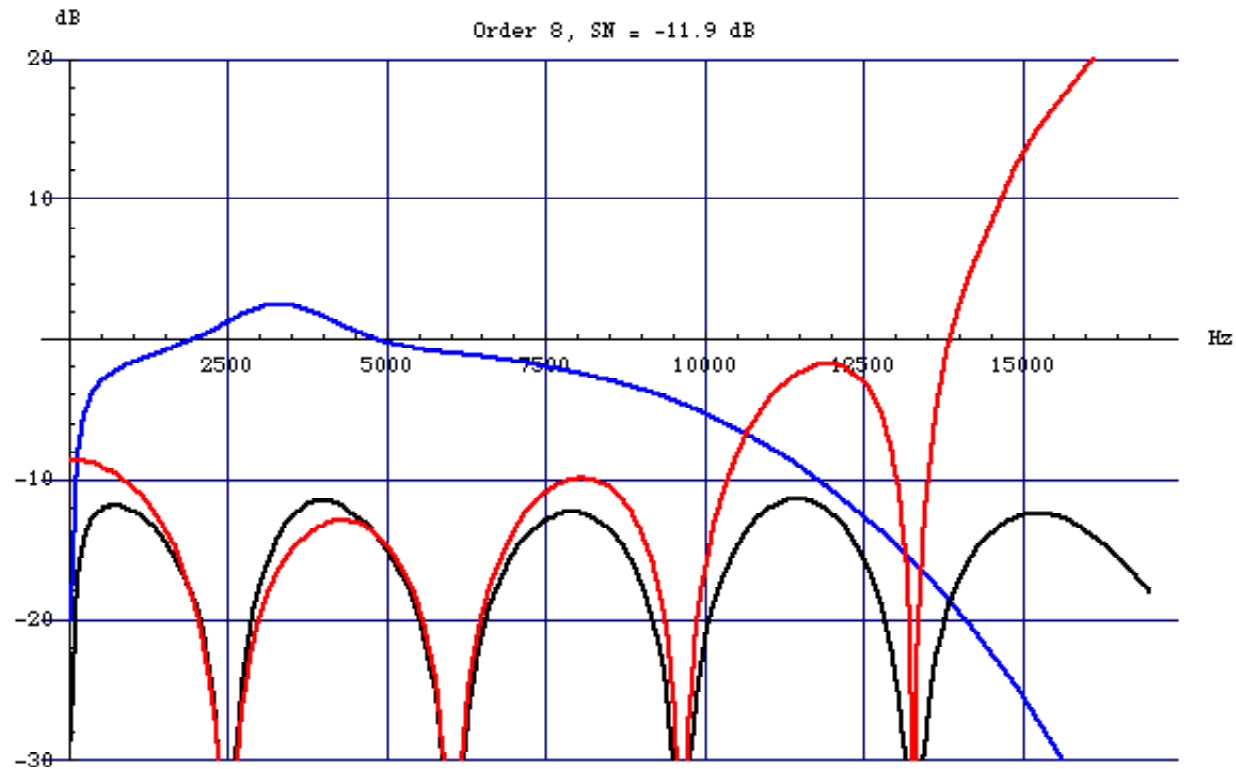
$$G = 1 + \sum_{p=1}^n a_p^2$$

La prima equazione significa che il rumore viene modellato come $1-H(z)$ dove $H(z)$ è la funzione di trasferimento del filtro FIR che ha per coefficienti $\{1, -a_1, -a_2, -a_3, \dots\}$, invece **il segnale buono passa inalterato**. La seconda equazione ci dice che il rumore totale “elettrico” non può che aumentare visto che le quantità che si vanno a sommare sono sempre positive.

Progettare il filtro **FIR**, da cui dipende la bontà dell'operazione di rimodellamento, è un compito delicato perché bisogna diminuire il rumore udibile mentre contemporaneamente aumenta la quantità del rumore totale, è necessario un compromesso e la scelta migliore può essere verificata solo all'ascolto. Per esempio si può procedere nel modo seguente.

- Si stabilisce una funzione che rappresenta la sensibilità dell'orecchio dell'ascoltatore tipo e la si inverte per ottenere una funzione di pesatura
- Si stabilisce un criterio per valutare la bontà del risultato, per esempio la potenza pesata del rumore nella banda di interesse oppure il massimo della stessa funzione in banda audio.
- Si decide il numero di coefficienti del filtro che si vuole costruire.
- Con un programma di ottimizzazione si trovano i coefficienti che soddisfano le specifiche di progetto
- Si ascolta il tutto, si cercano i difetti e si ricomincia da capo.

Ecco un esempio di cosa si ottiene con questo procedimento usando come funzione di pesatura quella vista sopra.

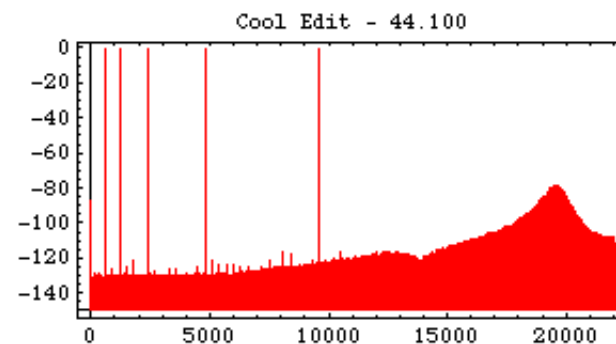
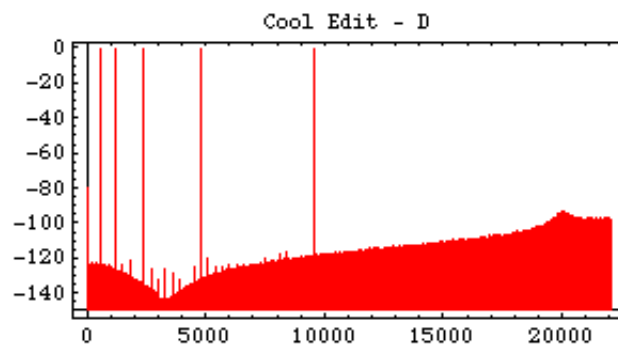
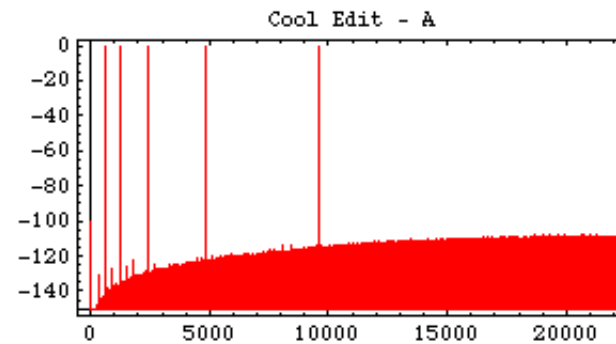
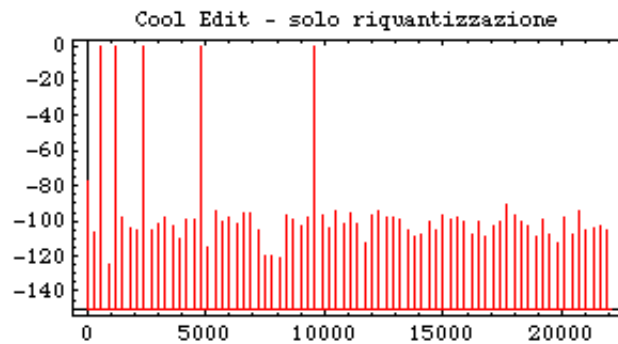


Esempio di progetto di un filtro di rimodellamento. In blu è rappresentata l'inversa della funzione di sensibilità, in rosso la curva di rimodellamento del rumore, in nero la curva del rumore pesato. Il programma di minimizzazione ha cercato di ridurre la potenza del rumore pesato. In questo caso con 8 coefficienti si guadagnano 11.9 dB di rapporto segnale/rumore.

Un fatto che deve essere assolutamente chiaro è che il segnale dopo la rimodellazione del rumore è molto fragile, qualunque operazione matematica sui bit del segnale effettua in pratica un nuova riquantizzazione e cancella totalmente gli effetti del rimodellamento precedente. **È quindi assolutamente necessario che la riquantizzazione, seguita dal rimodellamento, sia l'ultimo passo effettuato in digitale, dopo questo il file può solo essere inciso e suonato, pena la perdita di tutti i vantaggi.**

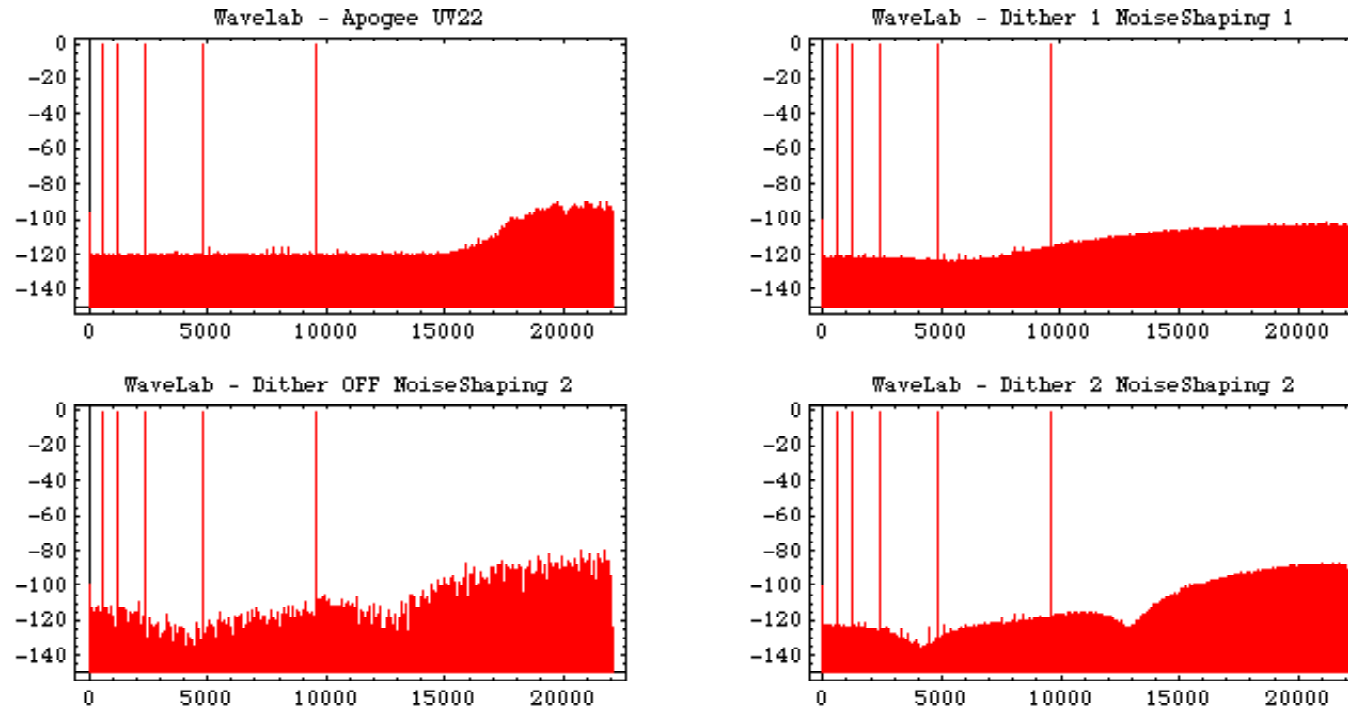
Esempi reali

Non è difficile usare un PC e un programma di editing digitale per vedere come viene applicato il Noise Shaping in pratica. Basta generarsi un segnale di prova a 24 bit e passarlo a 16 bit utilizzando una delle opzioni disponibili. L'analisi spettrale del risultato ci mostrerà la forma del rumore. In figura vediamo alcune delle possibilità offerte da Cool Edit. Se si effettua solo la riquantizzazione il rumore si presenta sotto forma di numerose componenti di frequenza discreta (perché strettamente correlate con i toni puri). Applicando il Dither e un Noise Shaping semplice (tipo A) si nota che il rumore si distribuisce su tutta la gamma (e anche se la potenza totale del rumore aumenta i picchi si smussano a causa della distribuzione). Nella parte inferiore vediamo altri due forme di Shaping più evolute.



Quattro possibilità offerte da Cool Edit, il segnale originale era formato da 5 toni puri a 0dB.

Ora vediamo alcune delle possibilità offerte da Wavelab. Il primo esempio è il famoso **Apogee UV22** usato ormai da quasi tutti i professionisti per le sue superiori qualità sonore. Il penultimo esempio è ottenuto con il solo Noise Shaping senza Dither, il vantaggio dell'uso del Dither è evidente dal raffronto con l'ultimo esempio.



Quattro possibilità offerte da Wavelab, il segnale originale era formato da 5 toni puri a 0dB.

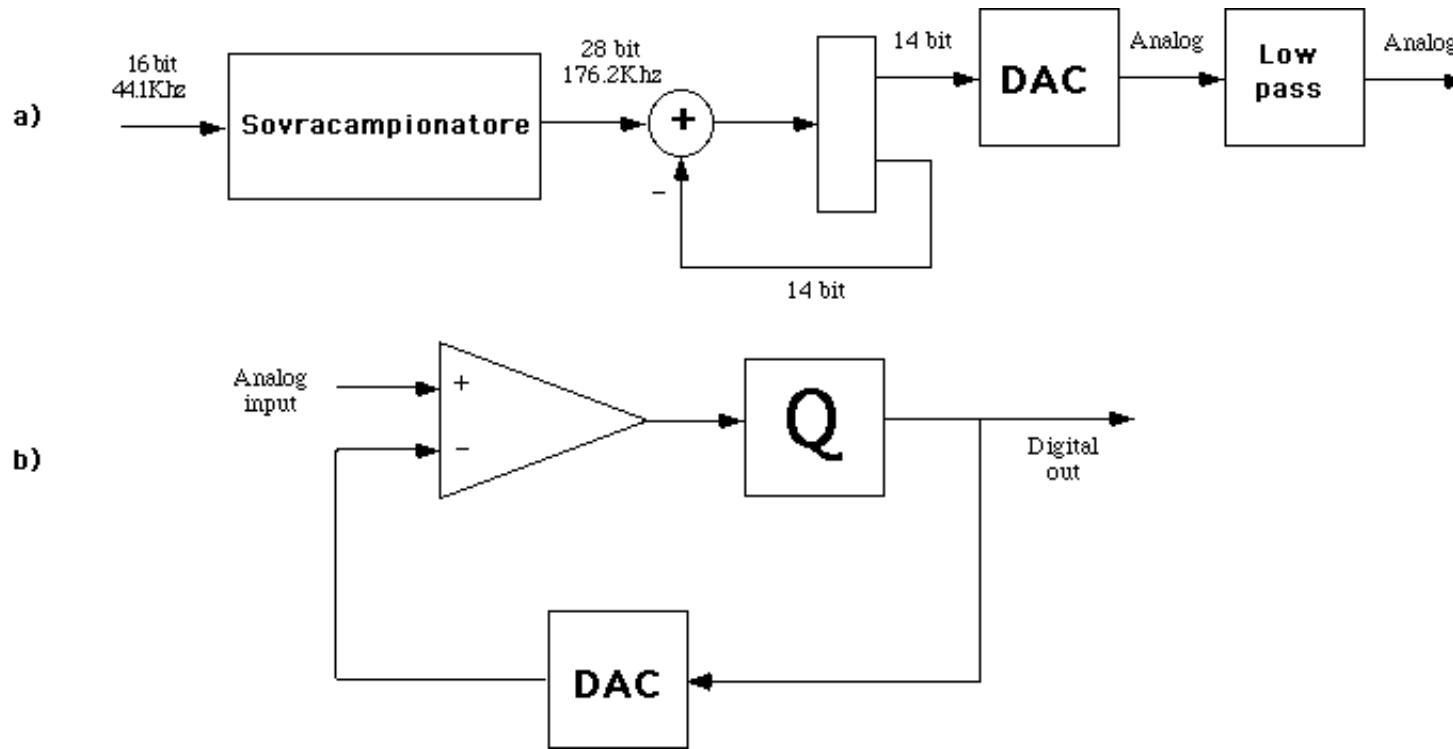
Il Noise Shaping nella conversione D/A e A/D

Finora abbiamo visto come si applica il Noise Shaping al processo di riquantizzazione. Ma la presenza di parole “lunghe” non avviene solo se si lavora con segnali ad alta risoluzione. Se si applica un’elaborazione matematica ad un segnale a **16 bit** conviene lavorare con una aritmetica maggiore per evitare che gli errori di calcolo influenzino la gamma dinamica reale, a questo punto dopo l’elaborazione ci si trova con un segnale a **24, 28 o 32 bit** che non conviene troncature direttamente, ma che si presta bene ad essere riquantizzato con l’uso del Noise Shaping. In particolare la fase di Oversampling nella conversione digitale analogica è spesso seguita da un Noise Shaping che ne migliora la risoluzione. La stessa Philips quando introdusse il procedimento (se non sbaglio con il lettore CD 104, cfr. Watkinson pag. 262) effettuava il filtraggio a **28 bit** e una successiva riquantizzazione a **14 bit** utilizzando i **14 bit** meno significativi per alimentare l’anello di reazione e i **14 bit** più significativi per alimentare un convertitore **D/A** a **14 bit**, a quei tempi più conveniente dei modelli a **16 bit** usati dai giapponesi.

Questo procedimento può essere portato al limite aumentando la frequenza di sovracampionamento e riducendo il numero dei bit in uscita fino a conservarne uno solo. Si ottengono così i vari tipo di conversione **D/A** monobit (**MASH, Pulseflow, Sigma-Delta**) in cui il rimodellamento del rumore svolge un ruolo essenziale.

Analogamente si può adattare lo schema del Noise Shaping al processo di conversione Analogico-Digitale utilizzando un convertitore **D/A** per alimentare il loop di reazione. In genere tutto il procedimento avviene all’interno di un circuito integrato. Anche questo processo, portato al limite, permette di generare un segnale digitale con parole di un solo bit e frequenza molto elevata.

Nella figura (riadattata dal testo di Wilkinson) si vede in modo schematico come la rimodellazione del rumore possa essere usata nei processi di conversione D/A e A/D,



Noise Shaping nella conversione D/A (a) e nella conversione A/D (b).

Noise Shaping e Controreazione

Una volta ho trovato scritto che data la somiglianza delle topologie il Noise Shaping dovrebbe soffrire degli stessi inconvenienti che la controreazione causa nei dispositivi analogici. Questa affermazione non mi trova molto d'accordo e cerco in questo spazio di chiarire un poco la questione.

La controreazione analogica permette di controllare il comportamento di un amplificatore, riducendone il guadagno e contemporaneamente rimodellando la distorsione. Il concetto chiave è questo, la distorsione non viene eliminata ma rimodellata, in particolare vengono tostate le componenti delle prime armoniche ma vengono introdotte numerose altre armoniche di ordine superiore.

L'effetto di questa operazione all'ascolto dipende dalla sensibilità dell'orecchio alle varie armoniche del segnale musicale, tipicamente le armoniche pari di basso ordine vengono percepite come una colorazione gradevole (al punto che certi costruttori esoterici le introducono a bella posta nei loro amplificatori). D'altra parte le armoniche di ordine elevato sono più sgradevoli e il risultato di una eccessiva controreazione è spesso la rimozione di una coloritura gradevole unita alla introduzione di distorsioni più lievi ma meno gradevoli. Bisogna dire anche che l'uso di poca o nessuna controreazione richiede apparecchiature che siano intrinsecamente di ottima qualità e questo conferma in pratica l'associazione: apparecchiature economiche con molta controreazione, apparecchiature di qualità con poca controreazione.

Un altro aspetto fondamentale è che se i segnali da trattare hanno una pendenza eccessiva rispetto alle capacità dell'amplificatore si possono avere istanti di saturazione con la comparsa della famigerata Intermodulazione Dinamica.

Il principale vantaggio del trattamento digitale dei segnali è che i fenomeni sono meno sfumati rispetto all'analogico. Il comportamento delle apparecchiature può essere previsto con maggiore dettaglio e controllato con maggiore precisione. In particolare per quello che riguarda il Noise Shaping bisogna considerare due fatti:

- Come si vede dalla formula $Y(z) = X(z) + E(z)(1-H(z))$ si rimodella solo il rumore, il segnale audio rimane inalterato.
- Se si usa un opportuno “ditheraggio” l'errore in gioco ha la forma di un rumore causale scorrelato non si ha alcuna concentrazione dell'errore in armoniche del segnale musicale.
- La pendenza dei segnali di ingresso è limitata dalla frequenza di campionamento e non vi possono essere fenomeni di saturazione transiente.

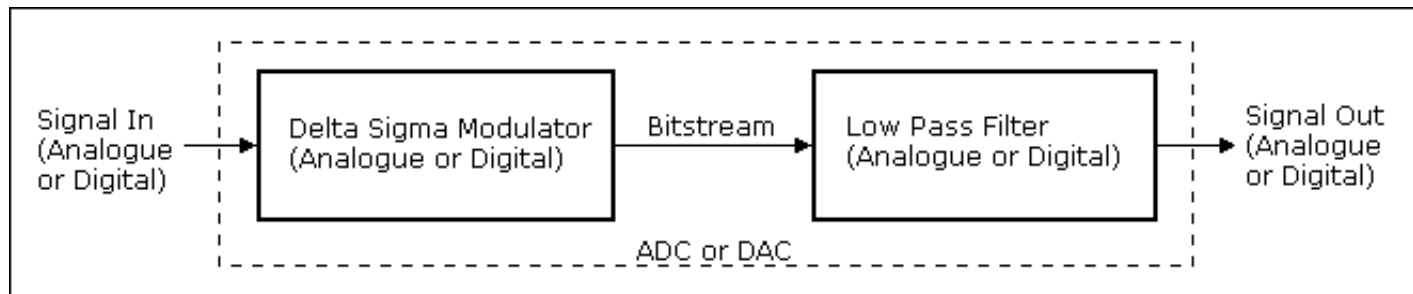
In conclusione a parte la somiglianza delle topologie sono i ben pochi i punti in comune tra le due tecniche.

La codifica DSD e lo standard SACD

La tecnica del sovracampionamento con conseguente riduzione della profondità di quantizzazione unita ad un opportuno rimodellamento del rumore ha portato allo **DSD (Direct Stream Digital)** usato nel formato **SACD** introdotto dalla Sony su supporto **DVD**.

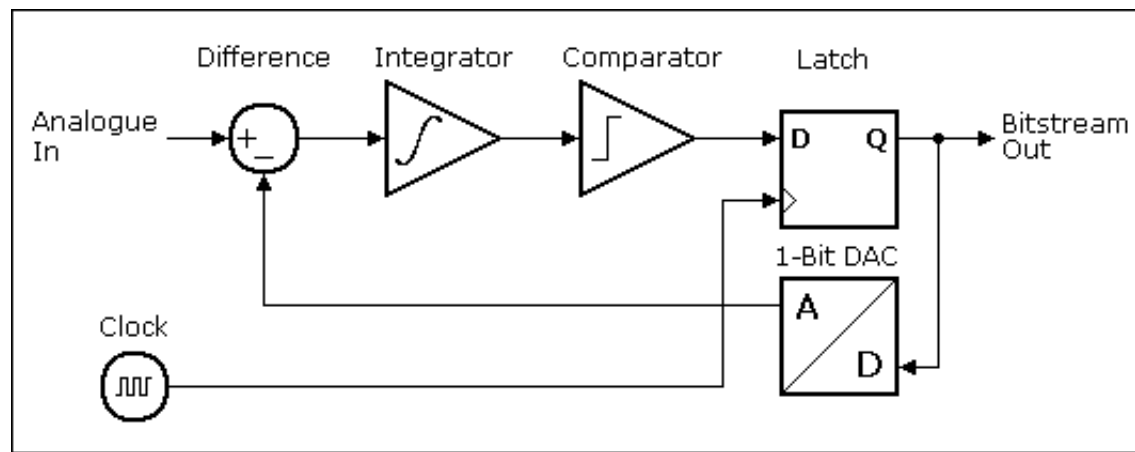
Il **DSD** prevede un bit di profondità alla frequenza di **2.8224 Mhz** (**64** volte la frequenza del **CD** audio) e l'uso del Noise Shaping è parte integrante del formato. La risposta in frequenza dichiarata è di 100 kHz e la gamma dinamica di 120 dB

Lo schema di gestione del segnale è molto semplice.

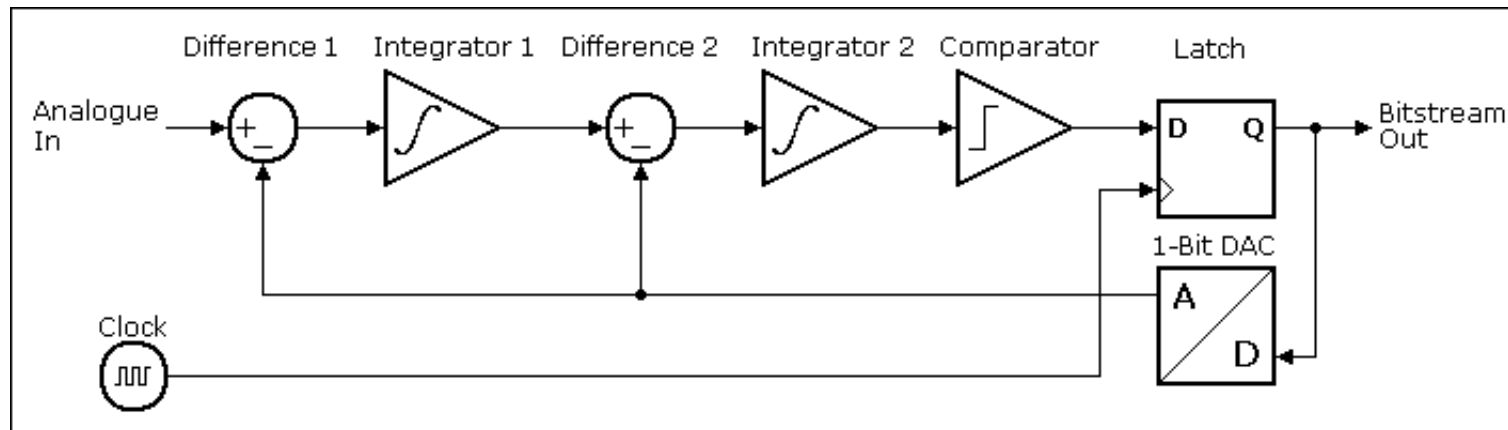


Schema del processo DSD.

Il convertitore A/D emette un bit per volta usando il segnale di errore come feedback per la generazione del bit successivo.



Modulatore Delta Sigma del primo ordine.



Modulatore Delta Sigma del secondo ordine.

Ed ecco come viene rappresentata una sinusoide come flusso ad 1 bit.

