5

La quantizzazione

Quantizzazione del segnale campionato

L'insieme dei numeri rappresentabili con un calcolatore è finito: con n quantità

$$\{x_0, x_1, \ldots, x_{n-1}\}$$

che possono assumere i valori binari 0 e 1, il massimo numero di configurazioni diverse ottenibili è 2^n. Ad esempio, con $n = 3$ si hanno le otto configurazioni seguenti:

000, 001, 010, 011, 100, 101, 110, 111.

I numeri reali invece non sono rappresentabili con una quantità finita di risorse, pertanto:

> **una volta campionato un segnale analogico per poterlo trattare come segnale digitale è necessario adottare una rappresentazione finita dei campioni**

questo procedimento prende il nome di **quantizzazione**.

Delle varie possibilità che si presentano, quella più usata è la rappresentazione lineare con numeri interi.

Sia $s(t)$ il segnale continuo da campionare e siano

$$\{ \ldots, s(-2\,\Delta), s(-\Delta), s(0), s(\Delta), s(2\,\Delta), \ldots \} = \{ \ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots\}$$

i campioni da quantizzare, fissati un numero di bit n e una soglia di accettazione A (il massimo modulo del valore in ingresso che non causa troncamento) il generico campione s_i viene quantizzato nel modo seguente

$$Q[s_i] = \max\left(-2^{n-1}, \min\left(2^{n-1}, \left\langle 2^{n-1}\frac{s_i}{A}\right\rangle\right)\right)$$

dove con < . > si è indicata l'operazione di arrotondamento all'intero più vicino. In questo modo un segnale continuo compreso tra $-A$ e A viene trasformato in un segnale discreto compreso tra -2^{n-1} e 2^{n-1} rappresentato con n bit. Si notino i seguenti casi particolari

$$Q[s_k] = \begin{cases} 2^n, & \text{se} \quad s_k \geq A \\ 0, & \text{se} \quad s_k \leq -A \end{cases}$$

Se si desidera confrontare il segnale quantizzato con quello campionato si può fare la normalizzazione

$$Q_A[s_i] = \frac{A}{2^{n-1}} Q[s_i] - A$$

che, si noti bene, è una **pura astrazione matematica**: quelli che viaggiano nelle apparecchiature sono i bit di $Q[s_i]$.

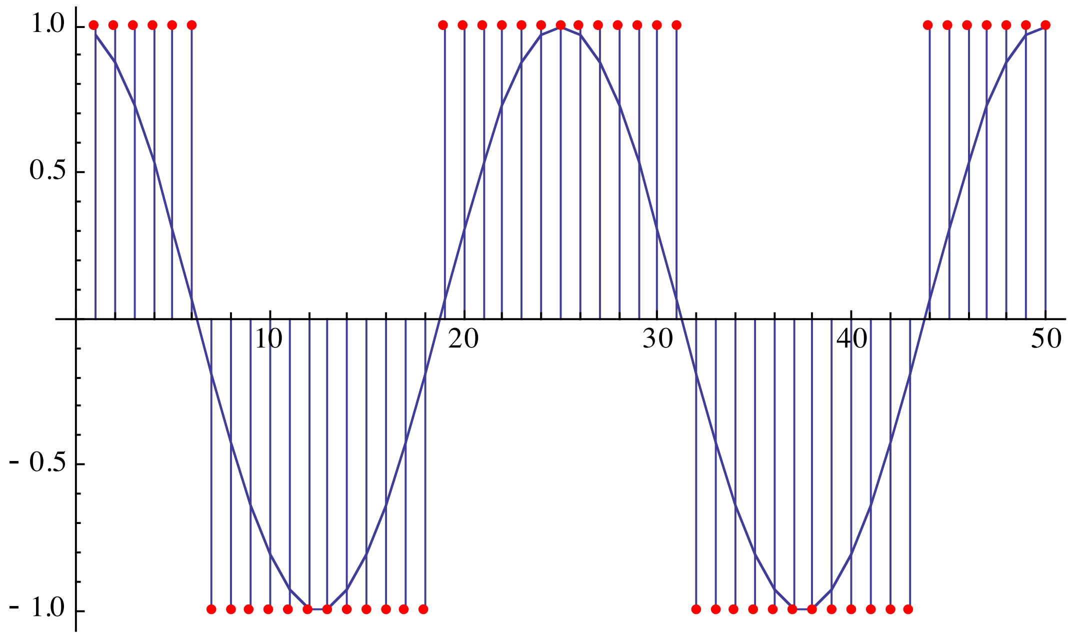
Un descrizione alternativa del processo di quantizzazione consiste nel considerare i bit del segnale quantizzato come un numero intero con segno rappresentato in complemento a due ma anche in questo caso la trattazione successiva rimane concettualmente equivalente.

Ai fini dello studio dell'errore è utile la quantità:

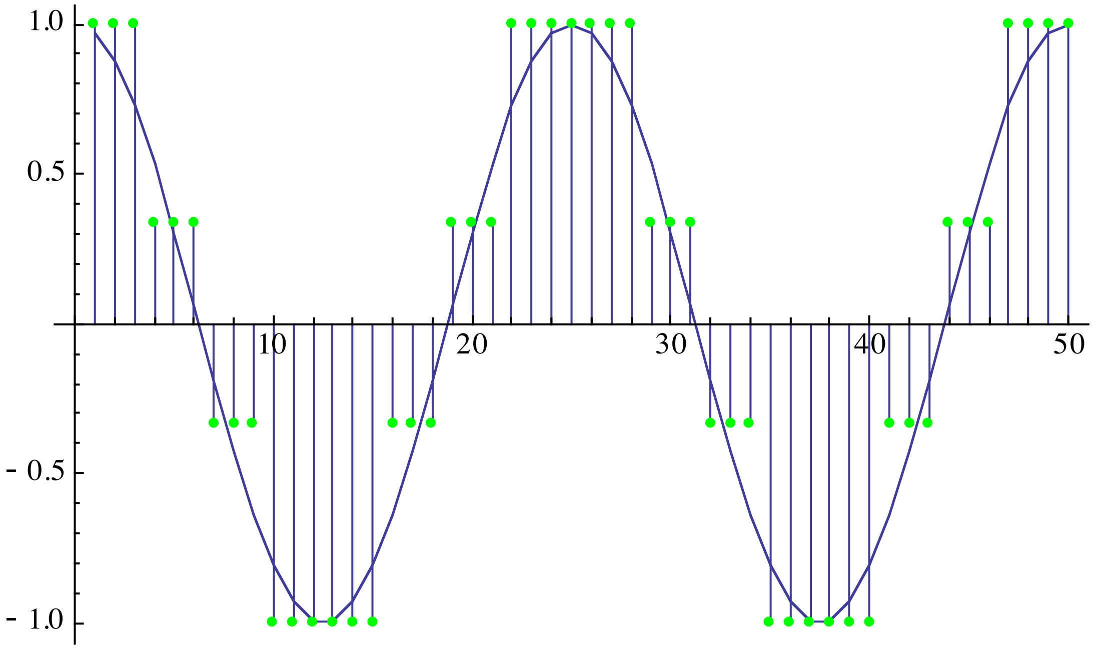
$$q = \frac{A}{2^{n-1}}$$

che rappresenta il valore di un **passo di quantizzazione** (ovvero il **quanto** elementare)

Nella figure seguenti, nel caso che *s(t)* sia una sinusoide campionata in 25 punti per periodo sono mostrati i valori di $Q_A[s_i]$ per A=1 e n=1



e per A=1 e n=4



1.0
0.5
- 0.5
- 1.0
10
20
30
40
50

Errore di Quantizzazione

Per poter valutare la bontà del processo di quantizzazione è necessario studiare con cura l'errore che si commette approssimando il segnale continuo con quello discreto. Si tratta di una analisi delicata che comporta la discussione di sottili questioni matematiche.

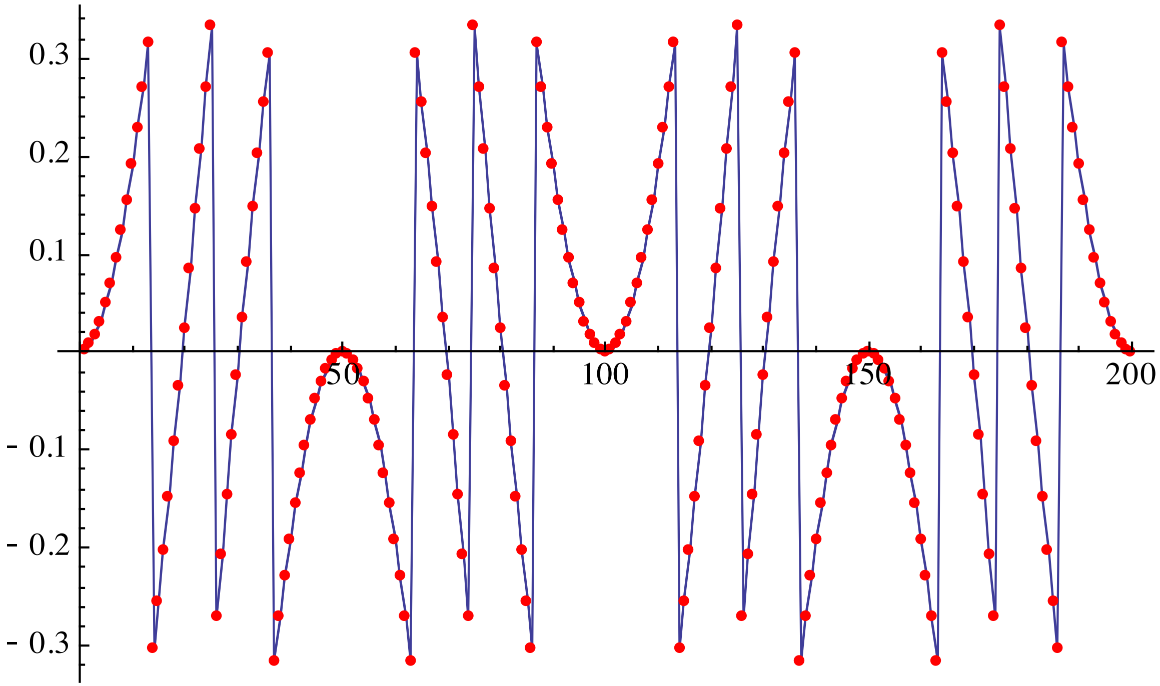
La prima considerazione da fare è che l'errore di quantizzazione

$$E_i = s_i - Q_A[s_i], \quad i = \ldots,-2,-1,0,1,2,\ldots$$

è una quantità puramente matematica come il segnale digitale stesso $Q[s_i]$: non si può parlare di suono di un segnale digitale, quello che suona è il segnale analogico dopo una **conversione digitale/analogica** e ai fini della qualità non si può non tenere conto di questa seconda fase.

La seconda considerazione è che da un punto di vista matematico qualunque tipo di deterioramento del segnale può essere visto come **rumore additivo**, si preferisce fare però la distinzione tra rumore e distorsione considerando il primo come il risultato di un processo stocastico scorrelato con il segnale di partenza (per esempio il rumore termico) mentre la seconda è il risultato di una operazione matematica correlata con il segnale di partenza.

Nel caso dell'errore di quantizzazione E_i si possono verificare entrambe le situazioni, per esempio nel caso della sinusoide visto prima. L'errore è fortemente correlato con il segnale. La figura seguente mostra l'errore nel caso di un campionamento a 2 bit con 100 punti per periodo, i punti dell'errore sono stati uniti per rendere visibile la regolarità dell'andamento dell'errore.



Si verifica immediatamente che vale sempre la relazione.

$$-\frac{q}{2} \le E_i = s_i - Q_A[s_i] \le \frac{q}{2}$$

ma una analisi accurata dell'errore dipende dalle assunzioni che possono essere fatte.

Tipicamente si suppone che

> la sequenza degli errori sia un processo stocastico stazionario non correlato con la sequenza dei campioni e i valori siano indipendenti tra loro e uniformemente distribuiti tra $-q/2$ e $q/2$.

Con questa ipotesi (verosimile nel caso di segnali musicali complessi di livello sufficientemente elevato ma non verosimile nel caso di segnali particolari o molto deboli rispetto alla soglia *A*) si può allora calcolare la dinamica del processo di quantizzazione in un intervallo di tempo T.

La potenza del segnale è

$$S = \frac{1}{T}\int_0^T s^2(t)dt$$

mentre quella del rumore è

$$N = \frac{2}{q}\int_0^{q/2} x^2 dx = \frac{q^2}{12}$$

Nel caso che il segnale di ingresso sia sinusoidale di ampiezza massima si ottiene la ben nota equazione per il **rapporto segnale-rumore** ed espresso in dB:

$$\mathrm{SNR} \approx 6.02\, n + 1.76$$

Che per n=16 (il valore scelto per i CD audio) da un SNR di circa 98 dB.

Questa formula, che spesso viene applicata sistematicamente in modo ingenuo impone invece alcune considerazioni:

1. Il termine 6.02 n mostra che per ogni bit in più nella parola di quantizzazione si guadagnano circa 6 dB di rapporto segnale-rumore. Non si deve però credere che n possa essere aumentato a piacere in quanto ogni apparecchiatura realizzata fisicamente ha dei limiti invalicabili legati al rumore termico introdotto nelle prime lezioni.
2. Il termine 1.76 è legato al fatto che si è utilizzato un segnale di ingresso sinusoidale, in genere un segnale musicale reale, a parità di valore massimo ha una potenza molto minore e si possono perdere anche più di 10 dB di rapporto segnale-rumore.
3. Tutta l'analisi viene fatta supponendo che il massimo livello del segnale sia esattamente l'accettazione limite A. In pratica, poiché nel caso che l'accettazione limite viene superata si verifica una inaccettabile distorsione, viene sempre lasciato un margine di tolleranza di qualche dB che va anch'esso a scapito del rapporto segnale-rumore.
4. In ogni caso anche se il massimo livello del segnale fosse esattamente l'accettazione limite A, un segnale musicale ha in genere momenti in cui il livello è elevato e altri in cui questo è anche molto basso (i fortissimi e i pianissimi delle esecuzioni sinfoniche) e nei pianissimi il rapporto segnale-rumore può essere molto piccolo.

ATTENZIONE: le problematiche ai punti 2, 3 e 4 valgono anche per l'audio analogico

Il Dither

Abbiamo già sottolineato come l'assunzione che la sequenza degli errori sia un processo stocastico stazionario non correlato con la sequenza dei campioni perda ogni validità per valori piccoli del segnale da campionare, in questo caso l'errore di campionamento viene avvertito come una fastidiosa distorsione. Esiste però una tecnica di facile applicazione che permette di risolvere completamente questo problema: se si aggiunge al segnale da campionare un rumore stocastico di piccolo valore questi fa sì che l'errore di quantizzazione diventi anch'esso un rumore con buone proprietà statistiche.

- Una prima possibilità consiste nell'aggiungere al segnale da quantizzare dei valori estratti a sorte in modo uniforme tra -q/2 e q/2, (**rectangular dither**) in questo caso il rapporto segnale rumore viene raddoppiato (ovvero si perdono 3 dB) ma la media dell'errore di quantizzazione viene forzata a zero.

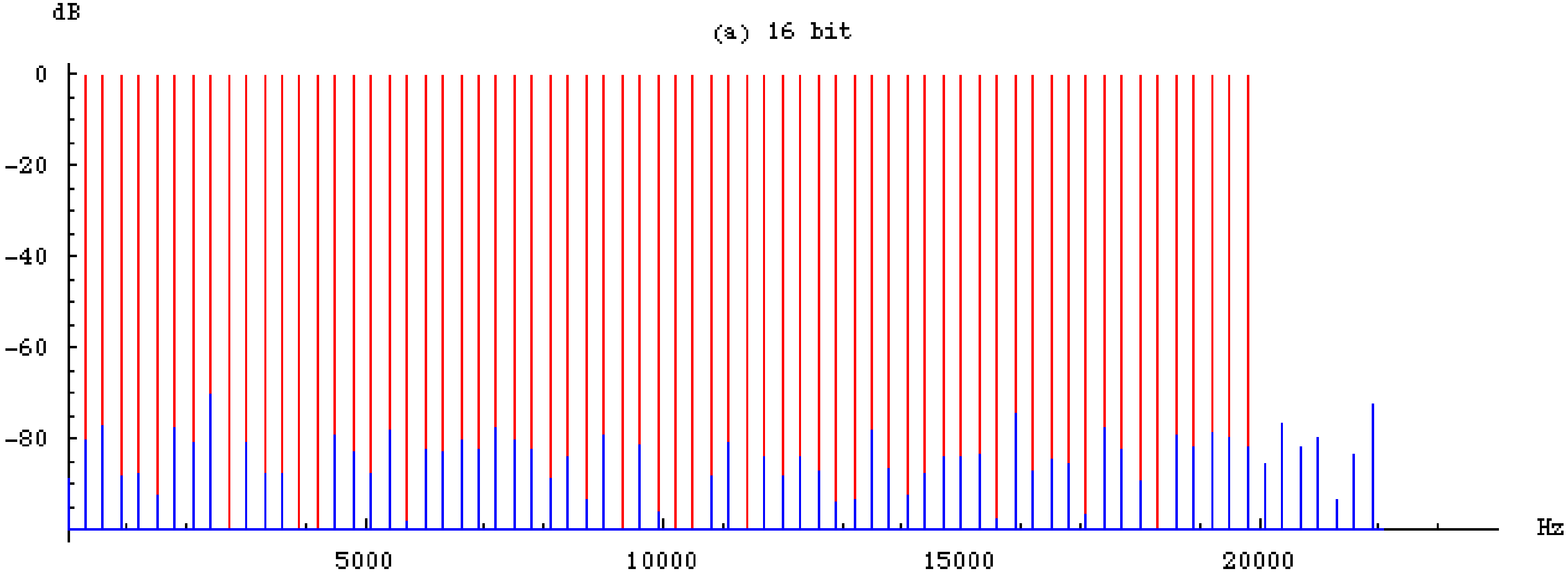
- Un'altra possibilità e di aggiungere al segnale da quantizzare la somma di due dei valori estratti a sorte in modo uniforme tra -q/2 e q/2, (**triangular dither**) in questo caso il rapporto segnale rumore viene triplicato (si perdono 4.8 dB), la media dell'errore di quantizzazione viene forzata a zero e la varianza assume un valore costante.

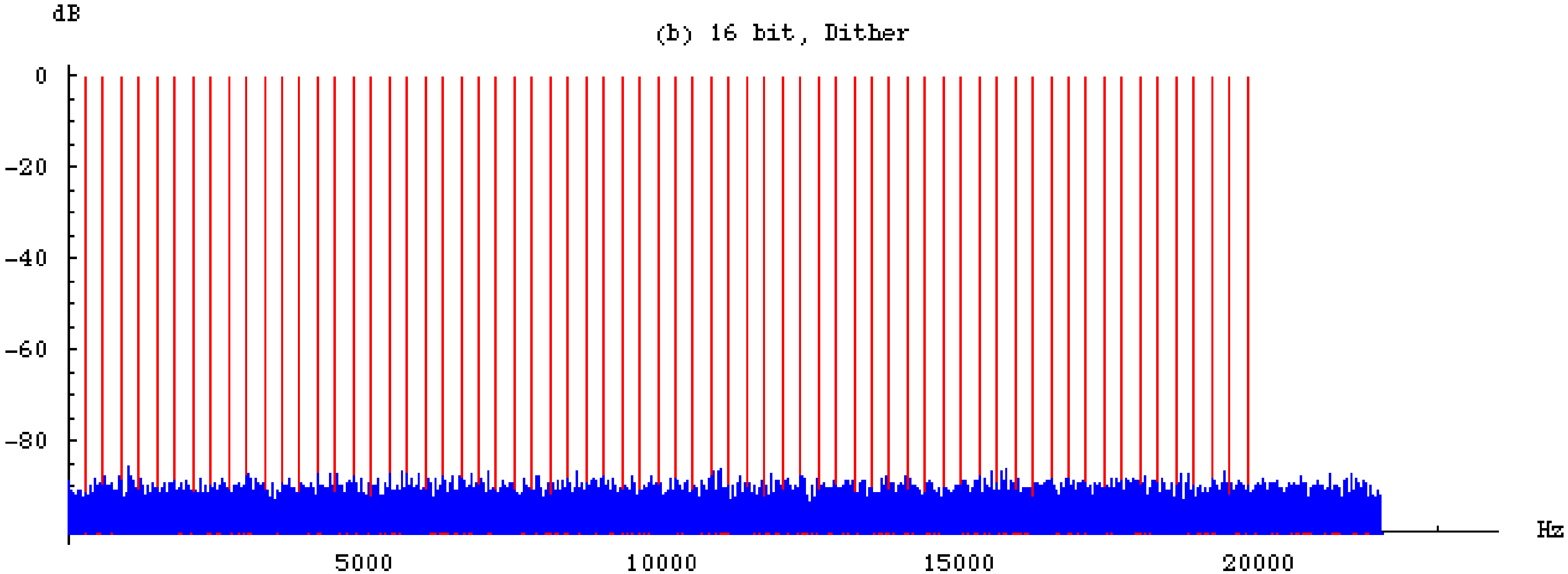
- Infine se si aggiunge al segnale da quantizzare dei valori estratti a sorte con una distribuzione normale (**gaussian dither**) con valore efficace q/2 il rapporto segnale rumore viene quadruplicato (si perdono 6 dB) e l'errore di quantizzazione tende comportarsi come un rumore bianco. Questa soluzione è la più semplice da applicare generando il dither per via elettronica sotto forma di rumore termico per esempio ai capi di un diodo.

La scoperta dell'utilità del dither avvenne in tempo di guerra quando si scoprì che i calcolatori meccanici usati nei bombardieri per stimare la trattoria delle bombe, funzionavano meglio in volo, quando erano soggetti a forti vibrazioni che a terra. Si è anche dimostrato che l'orecchio umano, per le sue intrinseche proprietà di filtraggio è in grado di distinguere segnali a banda stretta ben ditherati anche ben sotto il livello minimo di quantizzazione. Infine alcuni autori fanno notare come il rumore inevitabile nel segnale di partenza (rumore termico, fruscio del nastro, del LP, ecc.) possa fungere da dither rendendo accettabile l'errore di quantizzazione anche in assenza di un dither artificiale. In questo caso però il livello del rumore non è stabilito in base alle proprietà del processo di quantizzazione.

Un tipo di segnale artificiale che permette di esemplificare bene l'effetto del dither è il **segnale multitono**, ovvero una somma di sinusoidi dello stesso livello, equispaziate in frequenza, che nel dominio della frequenza appare come une specie di pettine.

Nella figura seguente vediamo, il segnale di prova in rosso e l'errore di quantizzazione in blu per un campionamento a 44100 Hz seguito da una quantizzazione a 16 bit (Grafico a). Si nota come il rumore di quantizzazione sia fortemente correlato con il segnale originario. Se lo stesso segnale viene discretizzato con l'aggiunta di dither rettangolare il risultato è una è una *spalmatura* del rumore che diviene scorrelato (Grafico b). Dalle figure si nota bene che, anche se l'effetto del dither causa un aumento della potenza del rumore la sua diffusione su tutte le frequenze ne causa un abbassamento del livello medio.

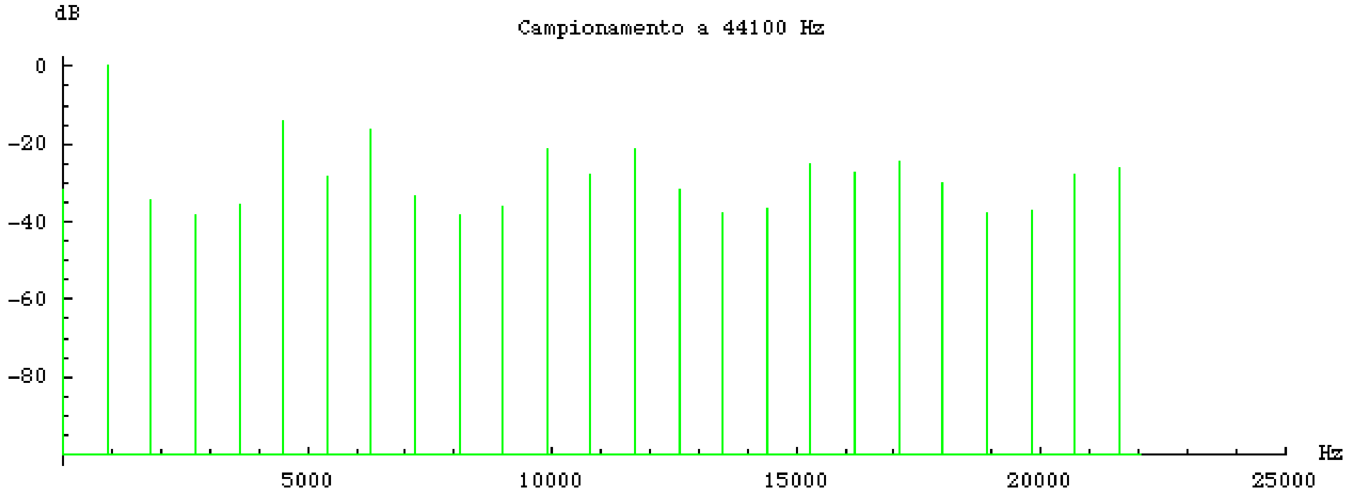

dB
(a) 16 bit
0
-20
-40
-60
-80
Hz
5000
10000
15000
20000

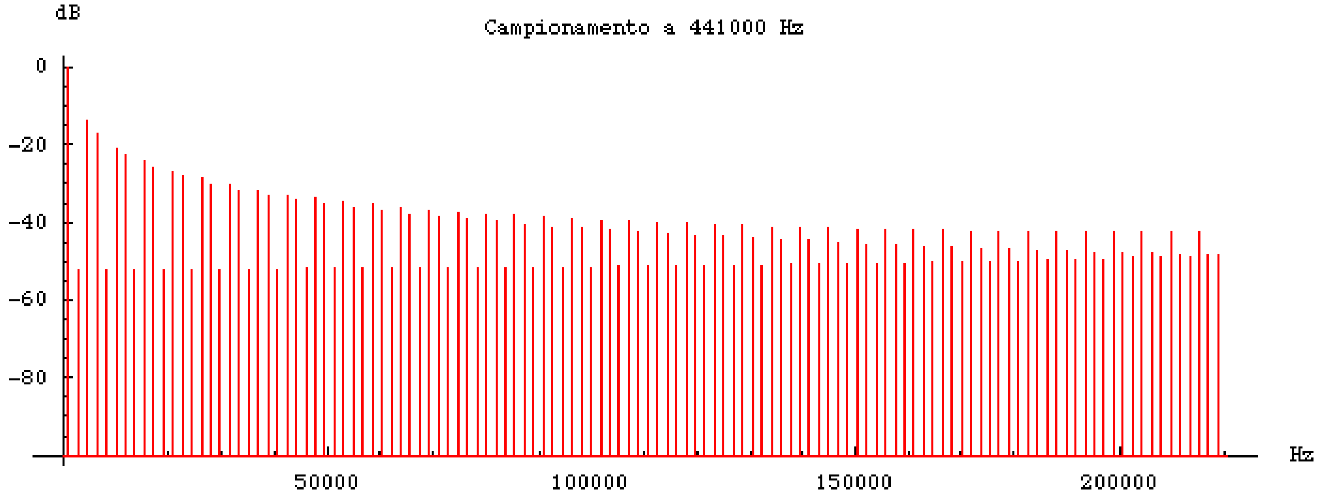

dB
(b) 16 bit, Dither
0
-20
-40
-60
-80
Hz
5000
10000
15000
20000

Relazioni tra errore di quantizzazione e frequenza di campionamento

È fondamentale a questo punto notare che anche se il segnale da rappresentare è stato filtrato passa basso prima della digitalizzazione lo spettro della funzione quantizzata non può essere limitato in frequenza perché gli errori non lo sono. Quindi, per il teorema del Campionamento, queste componenti ad alta frequenza vengono ribattute sotto la frequenza di Nyquist sotto forma di aliasing.

Passando nel dominio della frequenza si riesce a vedere meglio come stanno le cose. Se facciamo una rappresentazione del segnale quantizzato ad un bit dopo un campionamento a 44100 Hz e a 441000 Hz (esempi puramente didattici) nel primo caso il rumore viene spalmato solo in banda audio mentre nel secondo è distribuito fino a 220500 Hz, in questo secondo esempio è anche evidente come lo spettro dell'errore venga ribattuto alla frequenza di Nyquist.





Se il rumore di quantizzazione ha un buon comportamento (per esempio si può assimilare al rumore bianco) allora la distribuzione conseguente all'aumento della frequenza di campionamento causa un diminuzione della potenza di rumore in banda audio. In pratica ogni raddoppio della frequenza di campionamento causa una riduzione di 3 dB della potenza di rumore in banda audio. Questa proprietà è interessante di per sé, ma sfruttando la possibilità di agire sul segnale dopo la digitalizzazione si può fare molto di meglio modellando anche la forma dello spettro dell'errore di quantizzazione.

Il Noise Shaping

Come abbiamo già visto 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. 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 ininfluente all'ascolto.
- In genere il rumore digitale è correlato con il segnale.

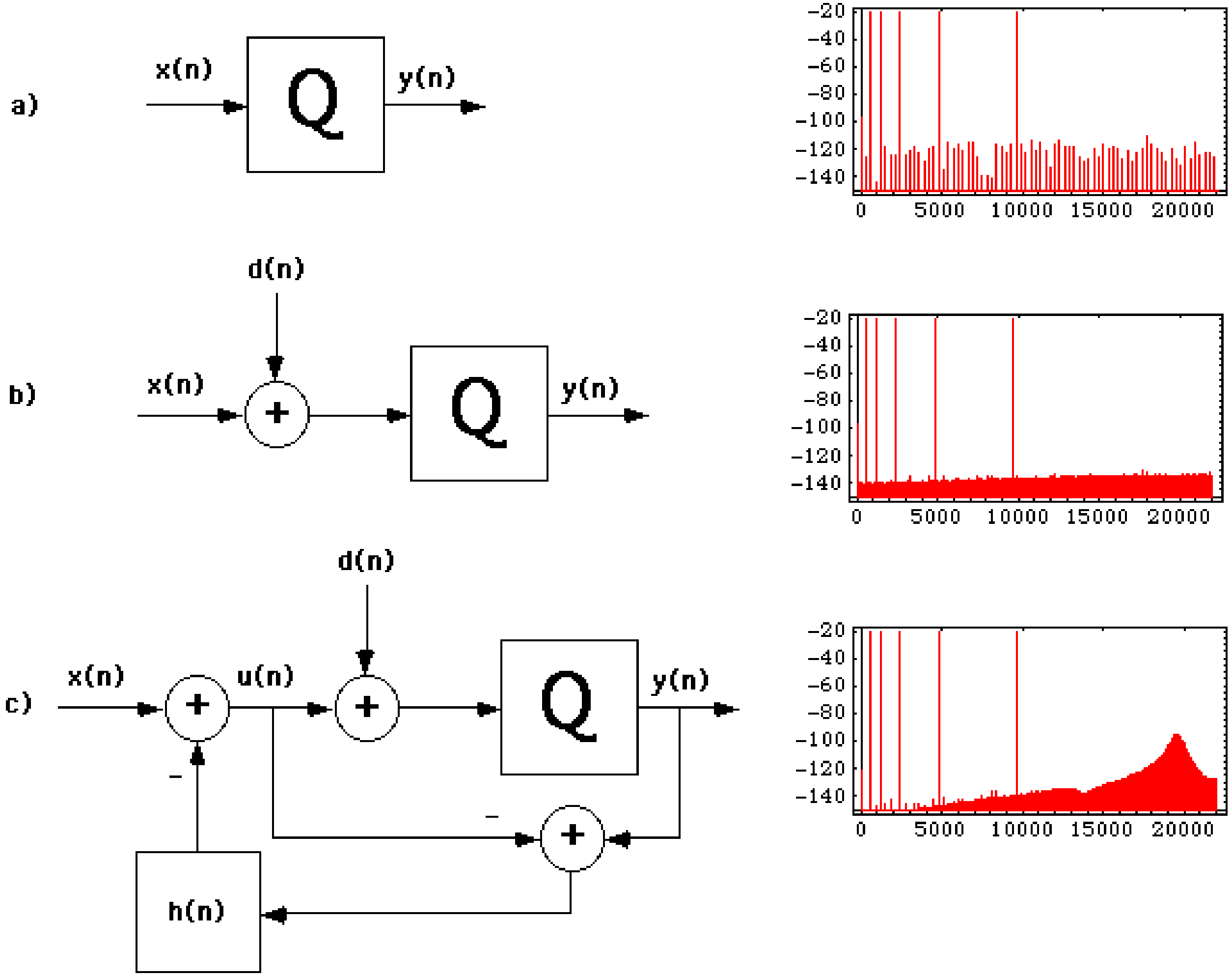
In pratica l'ultimo problema viene risolto aggiungendo al segnale un rumore casuale di piccola ampiezza (il dither) che se, da un lato, aumenta la potenza totale del rumore, dall'altro lo rende meno fastidioso all'ascolto. Dopo un buon "ditheraggio" il rumore digitale è con buona approssimazione distribuito in maniera uniforme su tutta la banda audio.

Il rimedio più semplice al problema del livello del rumore consiste ovviamente nell'usare più bit (diminuendo quindi il rumore di quantizzazione) e questa soluzione è stata presto adottata nelle apparecchiature professionali e ha poi dato vita alla famiglia dei **formati ad alta risoluzione**. 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:

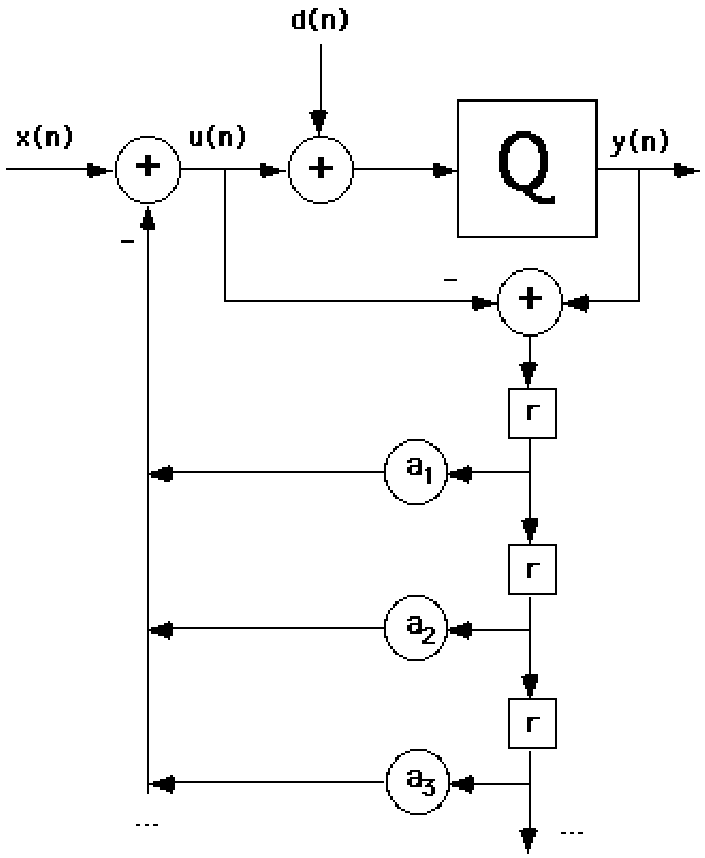
- permette di utilizzare rappresentazioni con un solo bit di profondità ed elevata frequenza di campionamento.
- permette, dopo una memorizzazione e/o una elaborazione ad alta risoluzione, di convertire il segnale digitale in una risoluzione più bassa..

Lo stesso problema della quantizzazione si presenta quando un segnale digitale con una risoluzione elevata (per esempio 24 bit) viene riquantizzato (per esempio passandolo a 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, ottenendo un rumore scorrelato e con distribuzione uniforme. Se però si introduce un anello di reazione con un filtro digitale è possibile modellare la forma dello spettro il rumore digitale.

Queste tre possibilità sono illustrate in figura: **a)** Solo riquantizzazione a 16 bit, **b)** riquantizzazione con dither, **c)** riquantizzazione con dither e Noise Shaping.



Prendiamo in considerazione il terzo tipo di riquantizzazione visto sopra,entrando 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 k coefficienti $\{a_1, a_2, a_3, \ldots a_k\}$ e sottratto al segnale in ingresso, 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:

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

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

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 $\{a_1, a_2, a_3, \dots a_k\}$, 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.

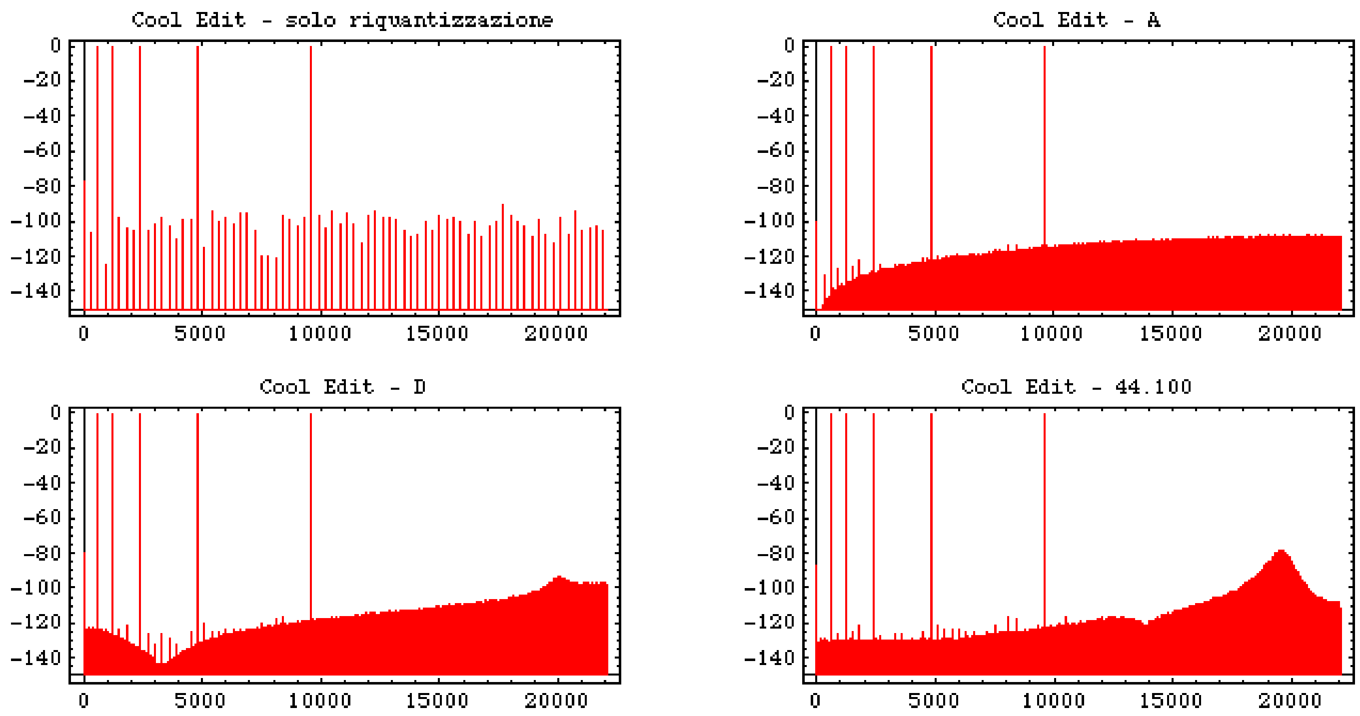
Progettare il filtro FIR, da cui dipende la bontà dell'operazione di rimodellamento, è un compito delicato perché bisogna diminuire il rumore udibile mentre contemporaneamente aumenta la quantità del rumore totale, è necessario un compromesso e la scelta migliore può essere verificata solo all'ascolto.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

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 un vecchio programma di editing: *Cool Edit 2000*. Il segnale originale è formato da 5 toni puri a 0 dB. 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.



Ora vediamo alcune delle possibilità offerte da Wavelab. Il primo esempio è il famoso Apogee UV22 usato da tempo dai 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.