

# ANALISI STATISTICA DELL'ERRORE DI QUANTIZZAZIONE

## CONVERTITORI SIGMA-DELTA

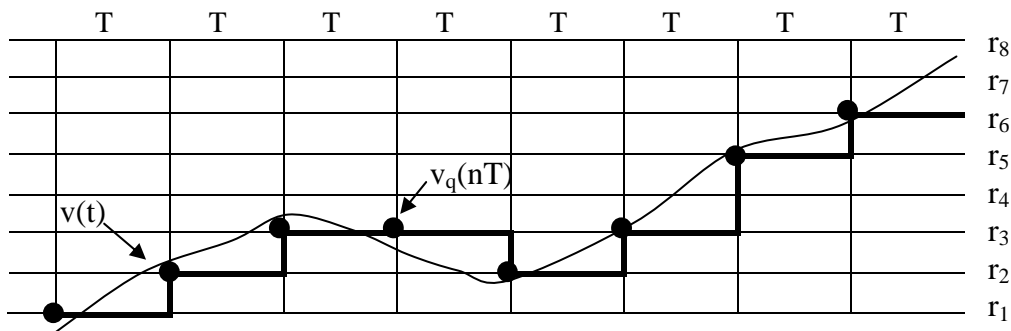
Consideriamo un segnale  $v(t)$  a banda limitata ( $-f_m \text{---} f_m$ ). Esso potrà essere campionato rispettando il teorema del campionamento se per la frequenza di campionamento vale la relazione  $f_c = \frac{1}{T} > 2f_m$ . Attraverso l'operazione di campionamento passiamo da un segnale continuo all'insieme dei suoi campioni:

$$v(t) \rightarrow v(nT)$$

Se quantizziamo i campioni ottenuti abbiamo:

$$v_q(nT) = v(nT) + e_q(nT)$$

dove gli  $e_q(nT)$  rappresentano i campioni dell'errore di quantizzazione. Assumiamo di utilizzare un quantizzatore uniforme.



Gli  $r_1 \text{---} r_8$  rappresentano i livelli di restituzione del quantizzatore.

Vogliamo studiare le caratteristiche statistiche dell'insieme dei campioni dell'errore di quantizzazione e del segnale (processo) che può essere ricostruito da tali campioni. Per fare questo è comodo rendere stazionario il processo stocastico in analisi.

In particolare consideriamo il processo rappresentante l'insieme degli errori di quantizzazione definito nel modo seguente:

$$x(t) = \sum_{n=-\infty}^{\infty} e_q(nT + \gamma) \cdot T \cdot \delta(t - nT - \gamma)$$

dove  $\gamma$  è una variabile aleatoria uniformemente distribuita fra  $0 \text{---} T$ . Ciò equivale ad assumere di non conoscere quale sia la fase del "pettine" di campionamento, non è una limitazione molto stringente. Il fattore  $T$  è introdotto per mantenere congruenza con quanto sempre fatto nella rappresentazione delle operazioni di campionamento.

Poiché nell'operazione di campionamento rispettiamo la condizione  $f_c > 2f_m$  possiamo anche scrivere, senza cambiare il problema:

$$x(t) = \sum_{n=-\infty}^{\infty} e_q(nT) \cdot T \cdot \delta(t - nT - \gamma) \quad (1)$$

L'informazione dell'insieme dei campioni  $e_q(nT + \gamma)$  è la stessa di quella contenuta nei campioni  $e_q(nT)$ .

Assumiamo come ragionevole:

$$E[e_q(nT)] = 0$$

$$R_{e_q}(k) = E[e_q(nT)e_q((n-k)T)] = \begin{cases} \sigma_q^2 & \text{per } k = 0 \\ 0 & \text{per } k \neq 0 \end{cases}$$

$$\sigma_q^2 = \frac{\Delta^2}{12} \text{ con } \Delta = \text{passo di quantizzazione}$$

Come possiamo giustificare queste ipotesi? Se i livelli del quantizzatore sono molti è ragionevole assumere che il segnale, per istanti di campionamento diversi, interessi intervalli di quantizzazione diversi, i valori dell'errore di quantizzazione sono quindi ragionevolmente incorrelati fra loro ed il loro valor medio sarà nullo (la densità di probabilità del segnale di ingresso, condizionata al fatto di ricadere in uno specifico intervallo di quantizzazione, è sostanzialmente costante).

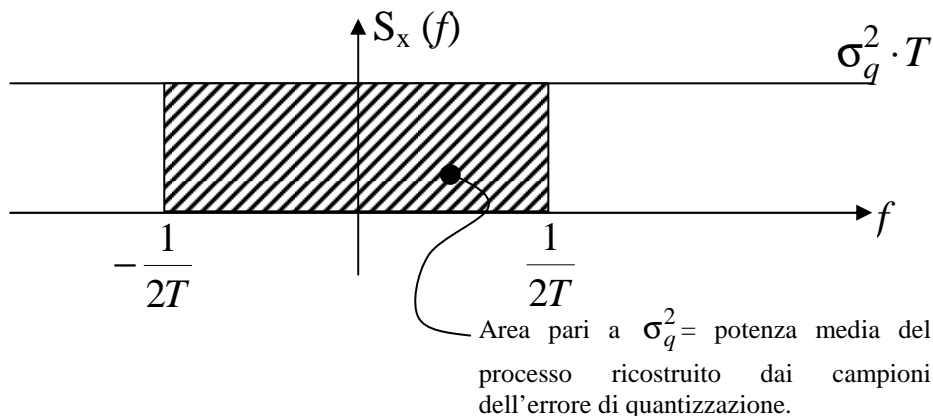
Il processo rappresentato in (1) è come già detto stazionario e caratterizzato dalla seguente funzione di autocorrelazione (attenzione a non dimenticare il fattore  $T$  presente nella definizione del processo  $x(t)$ ):

$$R_x(\tau) = \frac{1}{T} \sigma_q^2 \cdot T^2 \cdot \delta(\tau) = \sigma_q^2 \cdot T \cdot \delta(\tau) \xleftrightarrow{F} S_x(f) = \sigma_q^2 \cdot T$$

dove  $S_x(f)$  rappresenta la densità spettrale di potenza del processo in analisi.

Il segnale  $v(t)$  è assunto come un processo stazionario avente densità di probabilità delle ampiezze uniforme fra  $-V$  e  $V$ . Quindi i campioni del segnale saranno associabili a variabili casuali con la stessa densità di probabilità delle ampiezze. La potenza del segnale (e dei suoi campioni) sarà quindi data da  $V^2/3$ .

Nel seguito viene indicata con  $S$  la potenza di segnale ed  $N_q$  la potenza dell'errore di quantizzazione.



$$\sigma^2 = \frac{\Delta^2}{12} = \frac{V^2}{3} \cdot 2^{-2N}$$

$$\frac{S}{N_q} = 2^{2N} \quad \left. \frac{S}{N_q} \right|_{dB} \approx 6N$$

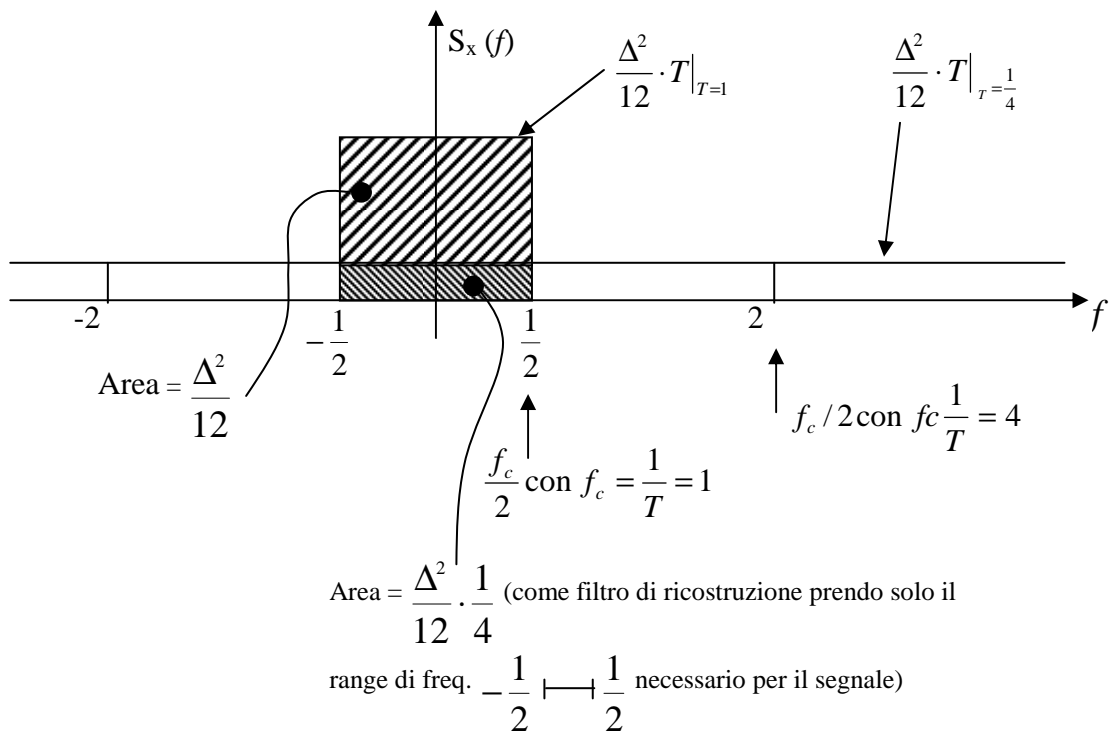
+ V |-----| - V escursione segnale  
N = numero bit quantizzatore

Va notato come  $\sigma_q^2 = \frac{\Delta^2}{12}$  non dipenda dalla frequenza di campionamento. Posso variare  $f_c$ ,

ma la potenza del rumore di quantizzazione rimane costante.

Ricordiamo inoltre come  $\sigma_q^2$  corrisponda sia al valore quadratico medio dei campioni dell'errore di quantizzazione, sia alla potenza media del segnale continuo ricostruito dagli stessi campioni (dell'errore di quantizzazione) attraverso un filtro ideale di banda  $[-\frac{f_c}{2}, \frac{f_c}{2}]$ .

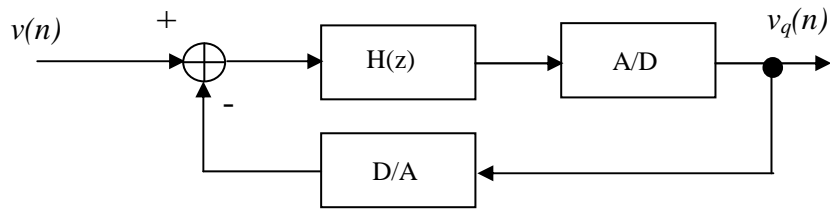
Assumiamo che per  $T=1$  il segnale sia ben campionato, se  $T < 1$  ancora meglio. Cosa succede se sovracampiono? Ad esempio in che situazione ci troviamo se usiamo una frequenza di campionamento caratterizzata da  $T'=1/4$  (al posto di  $T=1$ )? Il risultato è illustrato nella figura seguente.



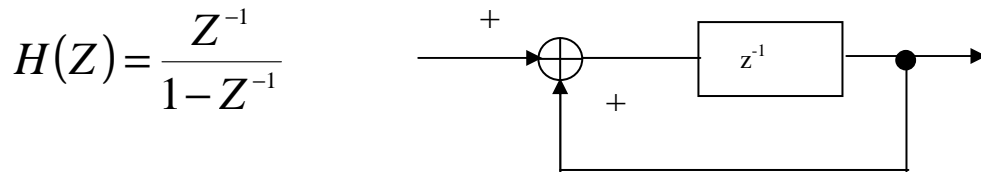
In altre parole se quadruplico la frequenza di campionamento, a parità di livelli di quantizzazione, guadagno un fattore 4 sull'S/N<sub>q</sub> (se ricostruisco il segnale nella banda minima). È come se guadagnassi un bit nel quantizzatore (cioè è come se si avesse un quantizzatore con un bit in più).

A partire da un quantizzatore a pochi bit è quindi possibile realizzarne uno con caratteristiche migliori (con più bit) sovracampionando il segnale. È una operazione molto costosa: per guadagnare un bit devo sovracampionare di un fattore 4 (e per guadagnare 4 bit devo sovracampionare di un fattore  $4^4 = 256$ ). Si può fare di meglio?

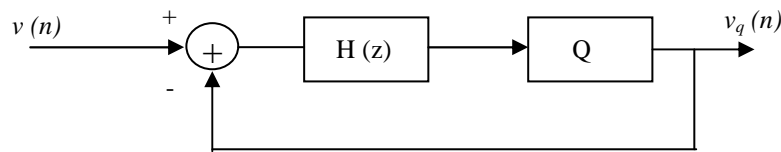
Consideriamo la seguente struttura (assumiamo per semplicità  $T=1$ ).



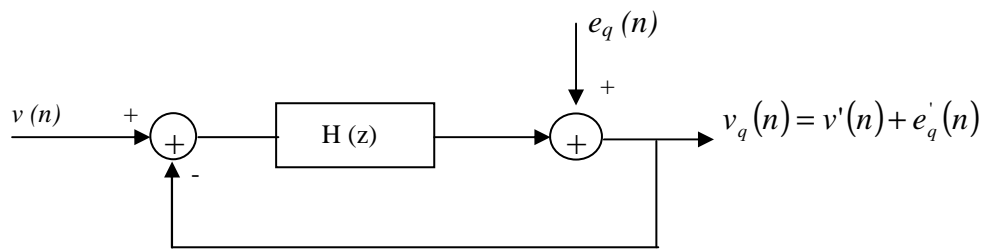
Dove per  $H(z)$  svolge la funzione di integratore per i segnali campionati, si ha quindi:



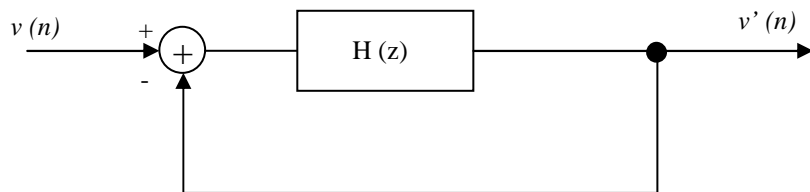
Il blocco A/D contiene oltre al quantizzatore, anche un codificatore *livello*  $\rightarrow$  *codice numerico*. Il blocco inverso *codice numerico*  $\rightarrow$  *livello* è il D/A. Possiamo, per i nostri scopi, lasciar perdere i blocchi di codifica e decodifica. Lo schema a blocchi del sistema diventa



Il blocco di quantizzazione può essere visto come un nodo di somma nel quale si aggiunge, al segnale utile, il rumore di quantizzazione.



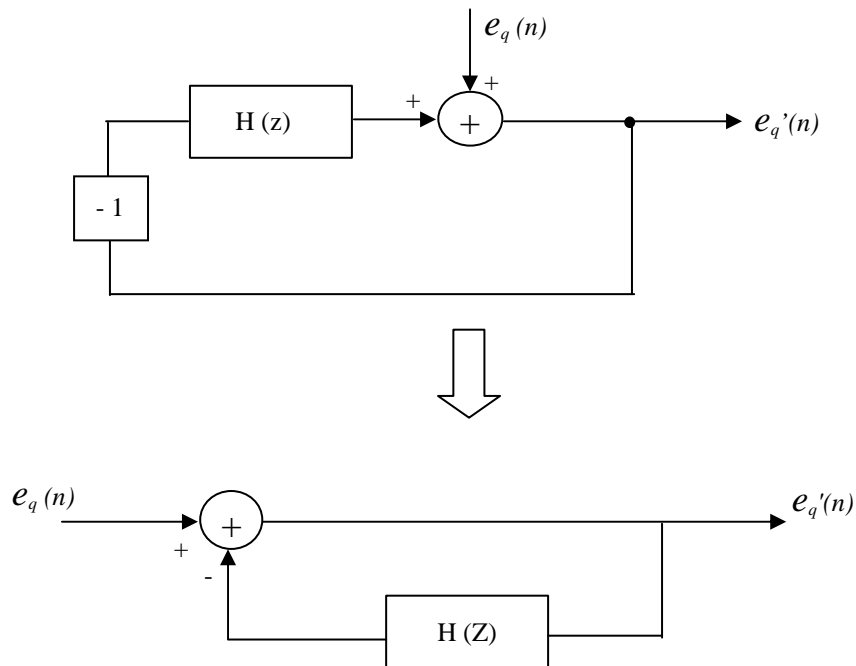
Analizziamo, prima, la funzione di trasferimento fra  $v(n)$  ed  $v'(n)$ , e poi quella fra  $e_q(n)$  e  $e_q'(n)$ .



$$V'(z) = V(z) \cdot \frac{H(z)}{1+H(z)} = z^{-1} \cdot V(z); \quad H_v(z) = \frac{V'(z)}{V(z)} = z^{-1}; \quad H_v(f) = e^{-j2\pi fT}$$

Se consideriamo tempi non normalizzati ( $T \neq 1$ ) abbiamo  $H_{v'}(f) = e^{-j2\pi fT}$ . In ogni caso il rapporto fra  $v'(n)$  ed  $v(n)$  è un semplice ritardo.

Per quanto riguarda l'errore di quantizzazione si ha:

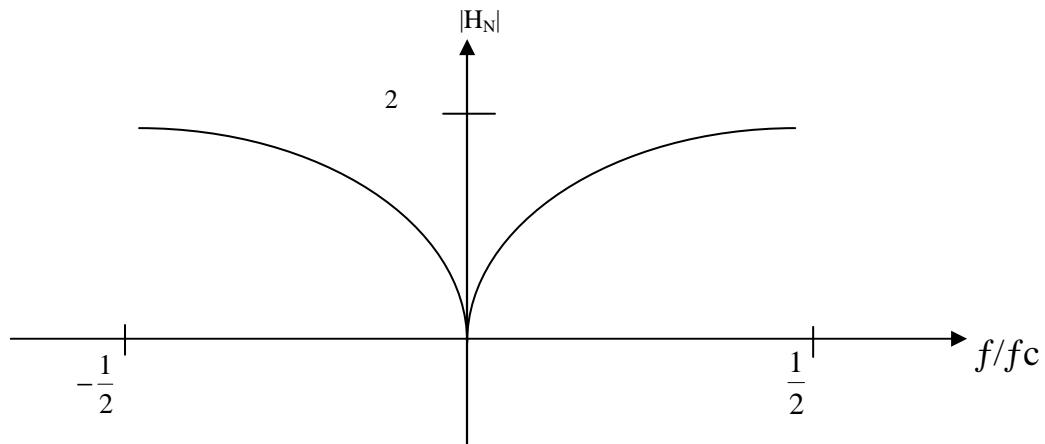


$$\frac{E_q'(z)}{E_q(z)} = \frac{1}{1+H(z)} = H_N(z) = (1-z^{-1})$$

$$H_N(z) = 1-z^{-1}$$

$$H_N(f) = 1-e^{-j\omega T} = 1-e^{-j2\pi fT} = 2je^{-j\frac{2\pi fT}{2}} \cdot \sin(\pi fT)$$

$$|H_N(f)| = |2\sin(\pi fT)|$$



$H_N(f)$  tende a concentrare il rumore di quantizzazione alle alte frequenze.

$S_q(f)$  = densità spettrale di potenza rumore generato dal quantizzatore =  $\sigma_q^2 \cdot T$ .

La densità spettrale del processo  $e_q'(nT)$  sarà invece data da:

$$S_q'(f) = \sigma^2 \cdot T \cdot 4 \cdot \sin^2(\pi f T)$$

Mentre la potenza del processo sarà:

$$P_q' = \int_{-f_m}^{f_m} S_q'(f) df = \int_{-f_m}^{f_m} \sigma^2 \cdot T \cdot 4 \sin^2(\pi f T) df$$

dove  $f_m$  è la reale frequenza massima del segnale che interessa campionare e quantizzare. Se come accade nella realtà  $f_m \ll 1/T$  abbiamo

$$\begin{aligned} P_q' &\cong \int_{-f_m}^{f_m} \sigma_q^2 \cdot T \cdot 4(\pi f T)^2 df = \sigma_q^2 T^3 \pi^2 4 \left[ \frac{f^3}{3} \right]_{-f_m}^{f_m} = \sigma_q^2 T^3 \pi^2 8 \frac{f_m^3}{3} = \\ &= \sigma_q^2 \pi^2 \frac{8}{3} \left( \frac{f_m}{\frac{1}{T}} \right)^3 = \sigma_q^2 \pi^2 \frac{1}{3} \left( \frac{2f_m}{f_c} \right)^3 = \sigma_q^2 \cdot 3,3 \cdot R^{-3} \end{aligned}$$

con  $R$  = fattore di sovracampionamento =  $\left( \frac{f_c}{2} \frac{1}{f_m} \right)$

$\frac{P_q'}{\sigma_q^2}$  rappresenta il rapporto fra rumore di quantizzazione ottenuto da strutture come quella

analizzata e da quantizzatori "classici". Nel caso analizzato si ha  $\frac{P_q'}{\sigma_q^2} = \frac{3,3}{R^3}$ .

Quindi per ridurre il rumore di quantizzazione di un fattore  $2^{2M}$  con  $M = 4$  (guadagnare 4 bit sul quantizzatore) bisogna considerare un fattore di sovracampionamento  $R \approx 9,5$ , da confrontarsi con un  $R=256$  che si ottiene nel caso di semplice sovracampionamento.

Le strutture di quantizzazione analizzate prendono il nome di strutture (quantizzatori) Sigma-Delta. Il nome deriva dal fatto che includono sia una operazione di accumulo (Sigma) che di quantizzazione, spesso ad un bit (che genera due soli valori +Delta, -Delta).

Il quantizzatore Sigma-Delta analizzato è quello caratterizzato dalla struttura topologica più semplice. I dispositivi reali utilizzano architetture un po' più complesse. Considerando fattori di sovracampionamento molto elevati dell'ordine del migliaio ed anche superiori si riescono a realizzare A/D a 16 bit partendo da un quantizzatore ad 1 bit.

Con fattori di sovracampionamento così elevati, l'analisi statistica precedentemente effettuata risulta un po' semplificata poiché non si può assumere che l'errore di quantizzazione generato dal quantizzatore ( $e_q(n)$ ) sia indipendente fra campioni successivi. E' altresì vero che utilizzando strutture retroazionate più complesse di quella analizzata si compensa, almeno in parte, questo fenomeno.

Un'altra prerogativa interessante degli A/D Sigma-Delta è rappresentata dal fatto che campionando a frequenze molto elevate non necessitano di filtri anti-aliasing.

Nelle pagine precedenti abbiamo visto la configurazione dei convertitori A/D a sovracampionamento (Sigma-Delta). Sulla base degli stessi principi si possono anche realizzare D/A a sovracampionamento. Questi dispositivi sono molto utilizzati negli apparecchi di tipo consumer (CD player ad esempio) poiché un D/A tradizionale sarebbe molto costoso richiedendo tipicamente la realizzazione di reti di resistenze di notevole precisione.

Schema a blocchi D/A Sigma-Delta

