

8

Formati e Conversioni

In un mondo ideale ci sarebbe un solo formato di audio digitale (ci sarebbero anche tante altre cose ma non è questa la sede per parlarne). Purtroppo, come avviene sempre quando la tecnologia si applica al mercato, le scelte tecniche vengono influenzate dai fattori più disparati e i risultati sono spesso sconfortanti.

Teoricamente è possibile digitalizzare un segnale analogico con qualunque combinazione di frequenza di campionamento f_c (che si misura in **Hz**) e di profondità di quantizzazione p (che si misura in **bit**). Il parametro f_c determina la banda passante utile (che è un po' meno di $f_c/2$) e p determina la dinamica (tradizionalmente stimata in **6 dB** per ogni bit utilizzato).

Vediamo ora una lista dei principali formati lineari multi bit, i formati compressi verranno trattati in una lezione a parte.

44100Hz/16bit (16/44)

È il formato del CD Audio che dà una banda passante di circa **20 KHz** con circa **96 dB** di dinamica. Il data rate è **0.7056 Mbit/s** per canale. Forse non tutti sanno perché è stato scelto proprio il valore magico **44100**.

Bisogna subito sgomberare il campo dalla leggenda che la moglie di uno dei progettisti voleva che la **Nona di Beethoven** entrasse tutta in un solo CD: è una balla!

Quando nacque l'audio digitale non esistevano apparecchiature ad hoc per registrare i dati e l'uso dei computer ancorché possibile (cosa dubbia per quei tempi) avrebbe avuto un costo spropositato, Era però ben noto che un segnale digitale serializzato non era altro che un onda quadra che poteva essere trasmessa e registrata trattandola come un segnale a radiofrequenza.

In particolare la banda passante dei registratori video era più che sufficiente a contenere un segnale digitale opportunamente trattato.

In altre parole il segnale digitale veniva trasformato in un segnale analogico con le stesse caratteristiche di una registrazione video e memorizzato su videocassette (**U-matic**, **Betamax** o **VHS**). Nella figura è mostrato l'adattatore Sony PCM F1 (sotto) insieme al registratore **Betamax** ad esso dedicato (sopra). È evidente dalla foto che il pannello di controllo del PCM F1 è simile a quello di un registratore con il controllo di livello e il display a segmenti.



Per ovvi motivi di semplicità di progetto era bene avere un numero intero di campioni per ogni linea di quadro.

Il formato televisivo americano **NTSC** usava **30 frame** di **525** linee al secondo (di cui **490** utilizzabili); con **3** campioni per linea si ottiene una f_c di **$30 \times 490 \times 3 = 44100$** .

D'altra parte il formato europeo **PAL/SECAM** usava **25 frame** di **625** linee al secondo (di cui **588** utilizzabili); con **3** campioni per linea si ottiene una f_c di **$25 \times 588 \times 3 = 44100$** .

Purtroppo **$44100/2 = 22050$** e la frequenza di Nyquist è troppo vicina al limite della banda audio per premettere un'agevole conversione DA e AD, e questo compromesso ha fortemente limitato la qualità della musica che poteva essere incisa su un CD.

48000Hz/16bit (16/48)

Quando, successivamente, venne progettato un sistema di registrazione dedicato all'audio (il **Digitale Audio Tape** abbreviato in **DAT**) la frequenza di campionamento preferenziale venne elevata a **48 KHz**. In questo caso la frequenza di Nyquist sale a **24 KHz** rendendo molto meno problematici i filtri precedenti e successivi alla digitalizzazione. Il data rate è **0.768 Mbit/s** per canale. Storicamente fu una sorpresa per i recensori scoprire la drammatica differenza di qualità sonora tra i due formati.

Con il miglioramento della tecnologia elettronica sono divenuti disponibili convertitori AD e DA con 18, 20 e 24 bit e aumentare la frequenza di campionamento non è più stato un

problema tecnico, d'altra parte lo sviluppo di **hard-disk** sempre più veloci, capienti ed economici ha risolto i problemi di memorizzazione del segnale digitalizzato.

Questo ha fatto nascere un certo numero di formati intermedi, utilizzati dapprima nelle apparecchiature da studio per sfruttare la maggiore risoluzione disponibile per ampliare la gamma dinamica almeno prima del riversamento su CD e divenuti formati poi di distribuzione del materiale musicale.

88200Hz/24bit (24/88)

È un formato esotico, usato da pochi puristi per le registrazioni destinate unicamente al riversamento su CD in quanto la riduzione di frequenza di campionamento e di profondità di quantizzazione si possono fare facilmente e in modo trasparente. Il data rate è **2.1168** Mbit/s per canale.

Questo sarebbe stato probabilmente lo standard ideale per il formato audio definitivo, se le apparecchiature a disposizione lo permettono è il formato giusto per i riversamenti domestici da LP o musicassetta su CD.

96000Hz/24bit (24/96)

Il costruttore giapponese **Pioneer** fu il primo ad inserire sui suoi registratori DAT questo formato che allontanando ancora la frequenza di Nyquist (**48 KHz**) dalla banda audio semplificava ulteriormente i problemi di filtraggio. Il data rate è **2.304** Mbit/s per canale.

Dopo molti anni dopo l'avvento del CD è stato possibile definire il formato fisico del successore del CD. Il **Digital Versatile Disc (DVD)** è un contenitore grezzo che nelle sue varie versioni può contenere da **5 a 18 GigaByte** di dati. Il fatto che sia **versatile** non è stato proprio un bene, visto che sono stati resi disponibili troppi formati DVD alternativi e la maggior parte dei consumatori se ne è stata alla finestra aspettando che i costruttori si decidessero.

Tra le varie possibilità di audio su DVD il formato 24/96 è il più diffuso, questo è stato a lungo il formato più ricco disponibile nelle apparecchiature commerciali a basso costo (per esempio nelle schede audio per PC).

I soliti spiritosi hanno pensato di interpretare i vantaggi di questa banda passante allargata e sono cominciate a spuntare casse dotate di super-tweeter ad ultrasuoni (fortunatamente poco diffuse nel mercato occidentale).

I 24 bit di profondità di quantizzazione forniscono una gamma dinamica teorica superiore al rumore termico delle normali apparecchiature. Alcuni recensori hanno criticato questa scelta giudicandola uno spreco. Bisogna notare che a fronte di un leggero incremento della occupazione in bit si riesce ad eliminare praticamente tutti i problemi legati alla quantizzazione: il fatto che **144 dB** non siano raggiungibili significa solamente che i problemi veri saranno legati alla qualità delle elettroniche è in particolare dei convertitori, lasciando molto spazio al miglioramento.

192000Hz/24bit (24/192)

È l'ultima moda in fatto di formato ad alta risoluzione, usato sia nei DVD che nella musica diffusa su Internet. La frequenza di Nyquist a **96 KHz** permette una banda passante utile fino a **80 KHz**, assolutamente inutile sotto ogni punto di vista ma, per fortuna, alcuni convertitori offrono di ridurre la banda passante fino a **40 KHz** utilizzando intelligentemente il margine per semplificare ulteriormente il filtraggio. Il data rate è **4.608 Mbit/s** per canale.

Una frequenza così elevata è un arma a doppio taglio in quanto, se da un lato la risoluzione teorica migliora rispetto al formato **24/96**, dall'altro la maggiore banda passante richiede apparecchiature di migliore qualità pena un avvertibile degrado musicale o peggio ancora l'impossibilità della riproduzione.

352800/24 kHz (24/352) DXD

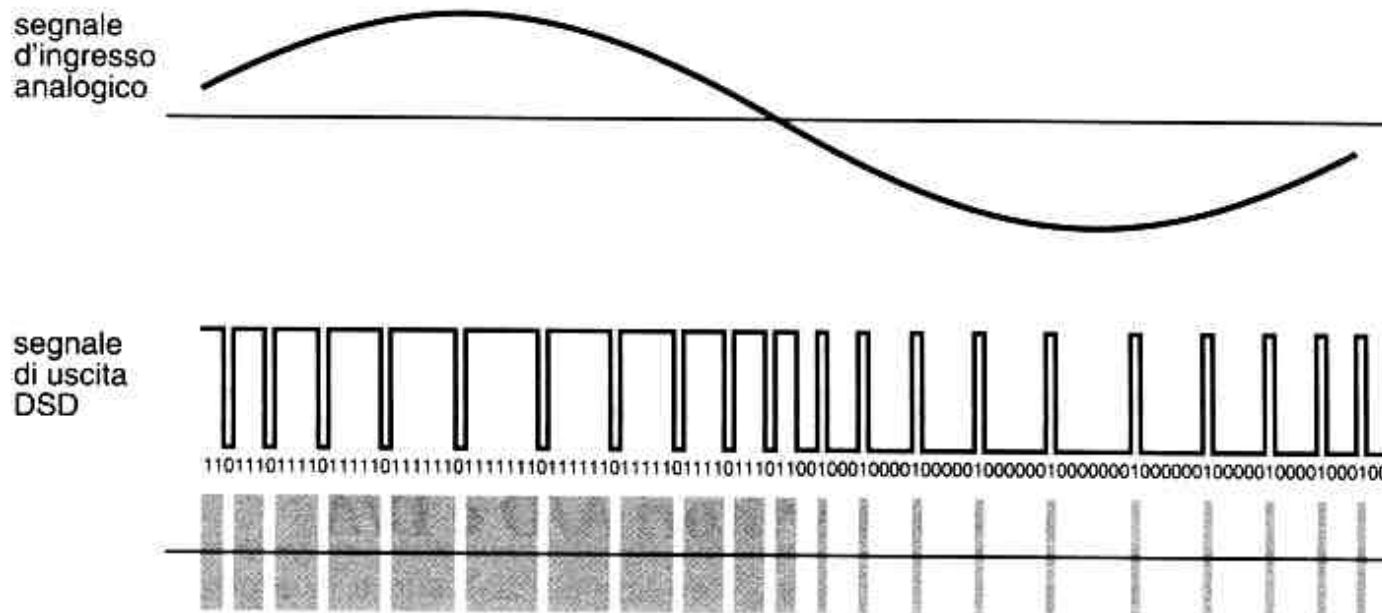
Il formato **Digital eXtreme Definition (DXD)** è stato introdotto per uso professionale allo scopo di permettere l'editing di materiale codificato **DSD**. DXD è un formato multibit con una frequenza di campionamento che è otto volte quella del CD e un ottavo di quella del DSD. I multipli interi rendono agevole la creazione dei master CD e SACD una volta finita la fase di editing. Il data rate è **8.4672 Mbit/s** per canale.

La codifica DSD

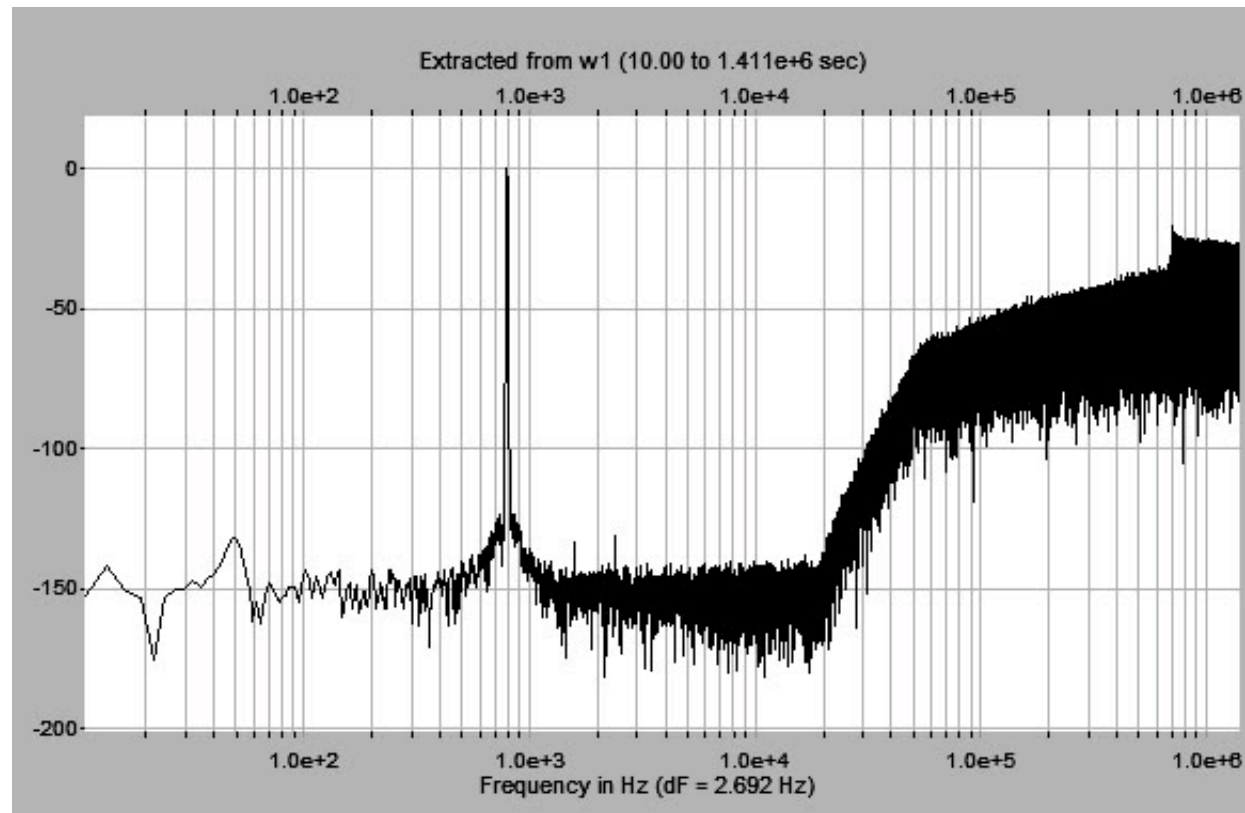
La tecnica del sovracampionamento con conseguente riduzione della profondità di quantizzazione unita ad un opportuno rimodellamento del rumore ha portato allo sviluppo del formato **DSD (Direct Stream Digital)** usato nel formato **SACD** introdotto dalla Sony su supporto DVD.

Il DSD prevede un solo bit di profondità di quantizzazione alla frequenza di **2.8224 MHz** (64 volte la frequenza del CD audio), il data rate è, ovviamente **2.8224 Mbit/s** per canale.

Ecco come viene rappresentata una sinusoide come flusso ad **1 bit** (la scala dei tempi è molto compressa rispetto alla realtà)



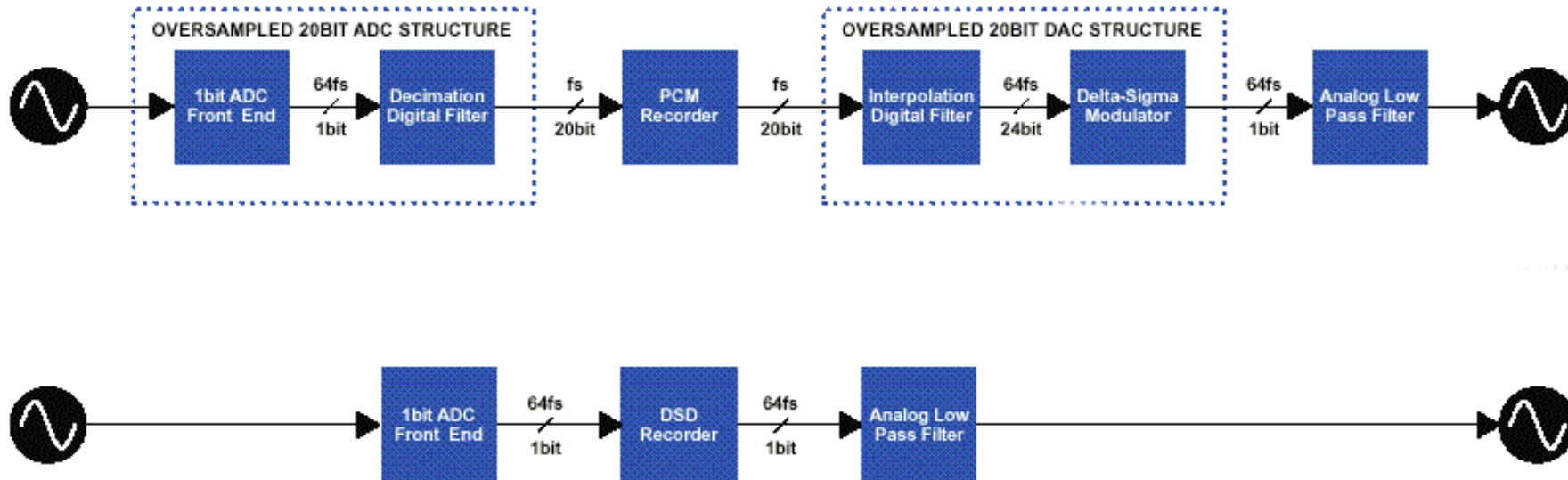
L'uso del **Noise Shaping** è parte integrante del formato. La risposta in frequenza dichiarata è di **100 kHz** e la gamma dinamica di **120 dB**. Vediamo un esempio del rimodellamento del rumore per un segnale test a **1000 Hz**.



All'inizio questo formato era stato introdotto per uso interno allo scopo di riversare in digitale i vecchi archivi su nastro. L'idea di fondo del formato DSD è la seguente

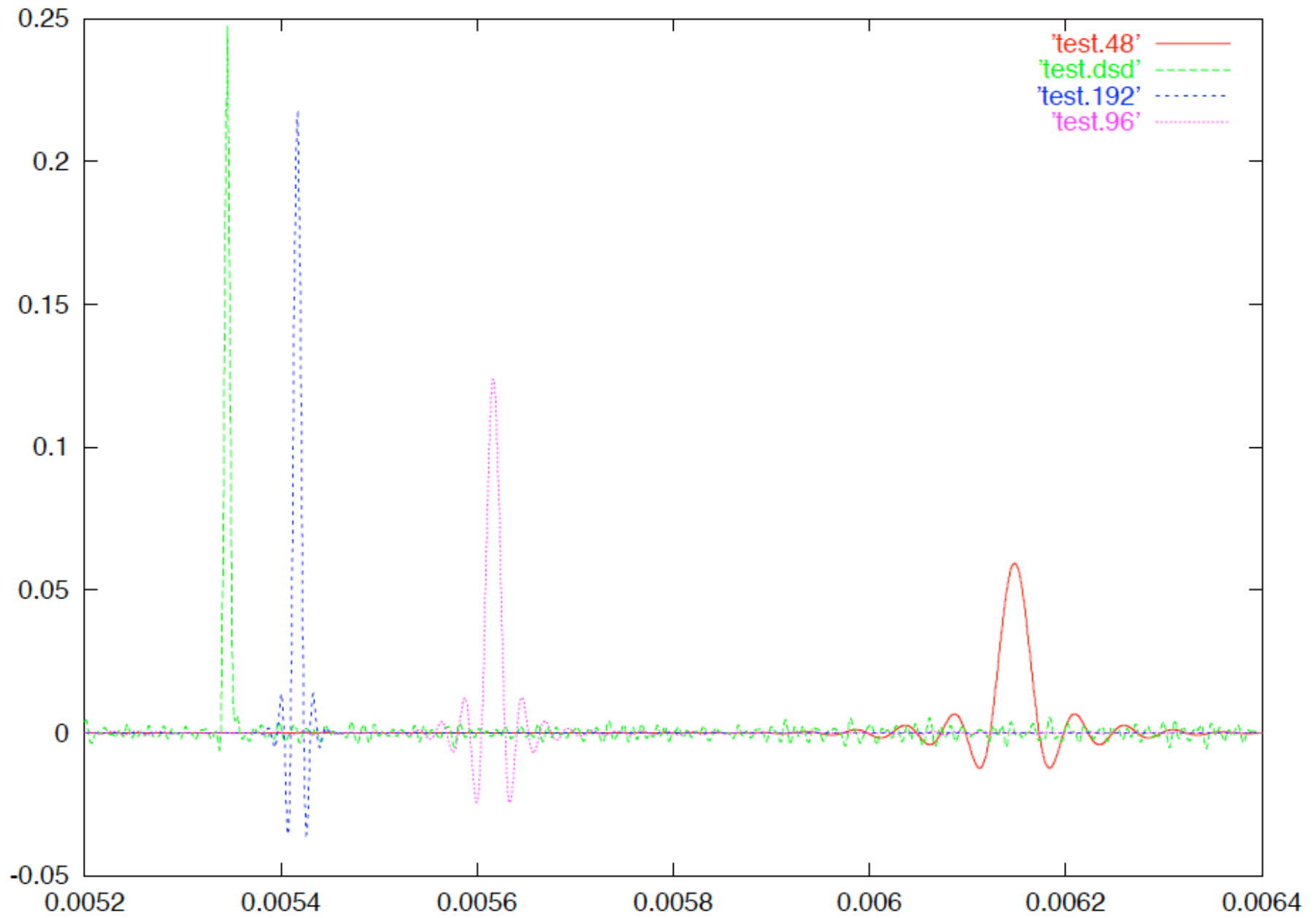
Se la conversione AD e DA sono implementate con convertitori a 1 bit è inutile degradare il segnale con due passaggi intermedi da 1 bit a multi bit e poi da multibit a 1 bit, tanto vale memorizzare direttamente lo stream digitale.

Il diagramma, rielaborato dalla pubblicità Sony dell'epoca, esemplifica questo concetto.



Purtroppo il formato ad 1 bit non permette editing e se si vuole elaborare il segnale (per esempio mixaggi ed equalizzazione, si deve passare ad un formato multibit, per esempio il **DXD**). Inoltre la alta risoluzione dei formati odierni viene ottenuta sempre attraverso il noise shaping ma usando rappresentazioni intermedie con un numero di bit maggiore di 1 (fatto che per sottili ragioni matematiche legate alle possibilità di **dithering** migliora grandemente la qualità del segnale). Una volta venute a mancare le motivazioni iniziali questo formato non ha avuto il successo sperato e al momento il suo uso è limitato soprattutto alla stampa dei **Super Audio CD**.

Per finire questa panoramica di formati, vediamo la risposta ad un impulso di 3 μ s a -6 dB, calcolata per alcuni di essi. La figura è tratta da D. Reefman, P. Nuijten. *Why Direct Stream Digital is the best choice as a digital audio format*, 110th AES Convention, Amsterdam, 2001



Conversioni

La presenza di più di un formato implica la necessità di passare dall'uno all'altro attraverso procedimenti matematici che devono essere implementati nelle apparecchiature elettroniche, le cose possono essere molto semplici o molto complicate. Vediamo dapprima quelle semplici

Conversioni di un fattore intero

Il caso più semplice se presenta quando si deve passare da una frequenza f_c ad un suo multiplo o sottomultiplo. Esempi tipici sono

$$48000 \Leftrightarrow 96000 \Leftrightarrow 192000$$

oppure

$$44100 \Leftrightarrow 88200 \Leftrightarrow 176400.$$

Studiamo separatamente il caso dell'*up-sampling* e del *down-sampling*.

- **Up-sampling intero**

Come abbiamo già visto la moltiplicazione della frequenza di campionamento di un fattore 2, 4 o anche 8 è un'azione indispensabile nella conversione DA multibit di qualità.

Il primo passo per moltiplicare la frequenza per n consiste nell'inserire $n-1$ zeri tra un campione e l'altro. Si ottiene un segnale campionato alla frequenza più alta che conserva **esattamente** la stessa informazione (**e lo stesso spettro**) del segnale di partenza. Bisogna quindi eliminare le frequenze immagine e questo si ottiene con un filtraggio digitale passa-basso.

NB il fatto che inserendo zeri lo spettro rimanga **esattamente lo stesso** e quindi non vi sia alcuna perdita di informazione è una conseguenza banale della linearità dell'operatore **Trasformata di Fourier**.

Un'altro modo di vedere la stessa cosa, stavolta nel dominio del tempo consiste nell'immaginare di costruire i campioni intermedi attraverso un processo di interpolazione.

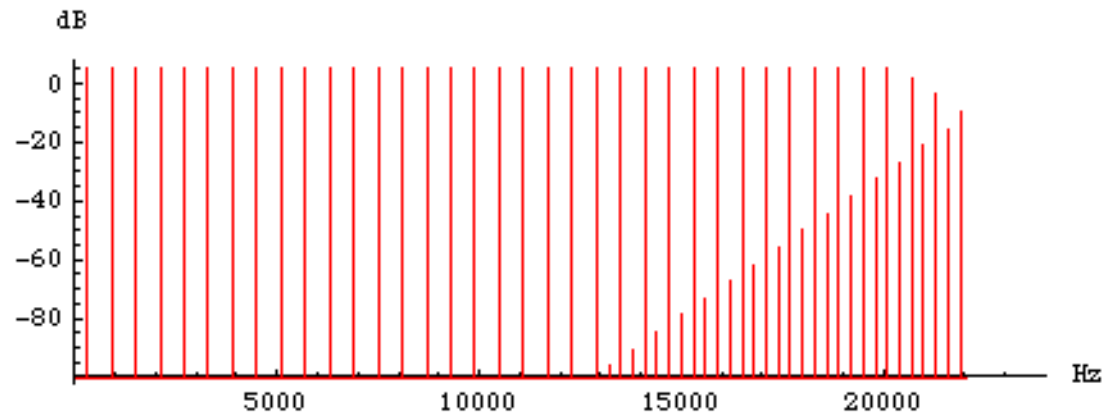
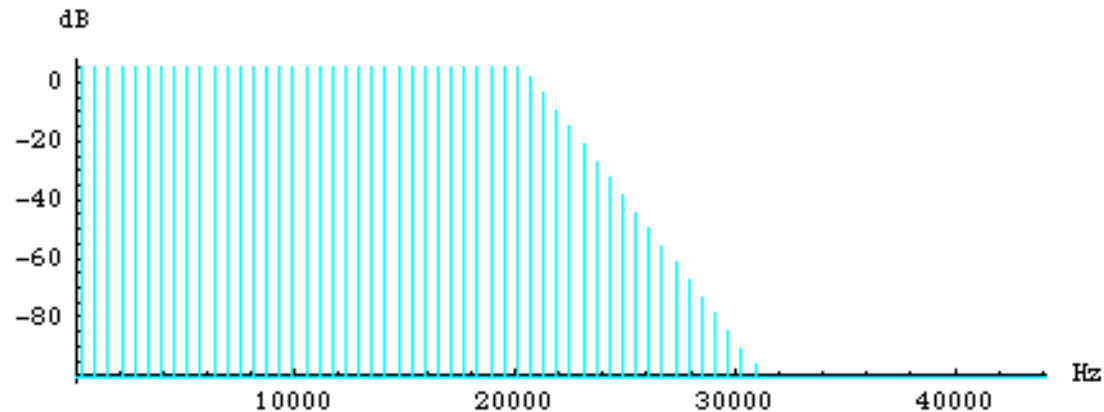
Nella pratica non si inseriscono davvero gli zeri ma si costruisce in modo opportuno il filtro digitale. Ricordo che altre tecniche come l'interpolazione **polinomiale** o l'interpolazione con **spline** sono solo un nome esotico per qualcosa che in questo contesto è resta sempre solo e soltanto un filtraggio digitale.

La qualità della conversione è legata alla bontà del filtro e non vi sono ostacoli ad ottenere risultati allo stato dell'arte come i migliori lettori CD possono testimoniare.

- **Down-sampling intero**

La divisione della frequenza di campionamento di un fattore n è un'operazione altrettanto semplice. Se si eliminassero semplicemente $n-1$ campioni consecutivi ogni n si otterrebbe un segnale campionato alla frequenza più bassa che ha come spettro quello originale tagliato e ribattuto.

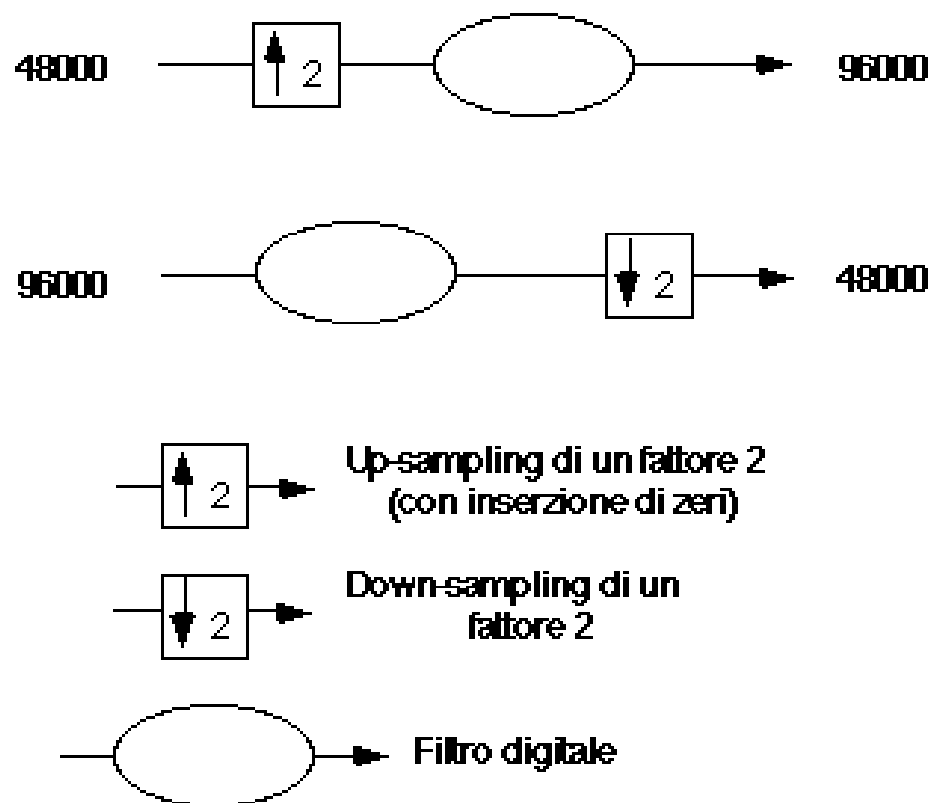
Per esempio, vediamo lo spettro di un segnale con componenti oltre i 30 KHz, campionato a 88200 (in celeste), lo spettro risultante dopo un sottocampionamento $88200 \Rightarrow 44100$ senza filtraggio preventivo. segnale (in rosso).



Occorre, quindi, fare precedere il sottocampionamento da un filtraggio passa basso. Anche in questo caso se si tiene conto dei campioni che verranno buttati dopo il filtraggio (e che quindi non vanno calcolati) il progetto si semplifica drasticamente.

L'uso tipico di questa tecnica si ha quando si vuole preparare un master a 44100 Hz partendo da un originale a 88200 Hz. Si noti che se si deve anche operare una riduzione della profondità di quantizzazione questa deve essere effettuata successivamente: infatti il **Noise Shaping** deve essere l'ultima operazione prima della stampa del CD.

Nella figura si vede lo schema del processo di conversione nei due casi appena visti



Conversioni di un fattore razionale

Se le frequenze da convertire stanno in un rapporto fisso non intero (per esempio 48 KHz e 32 KHz stanno come 3 a 2) si può effettuare la conversione effettuando prima un **up-sampling** (da 48 KHz a 96 KHz) e poi un **down-sampling** (da 96 KHz a 32 KHz).

Fare il contrario sarebbe drammatico oltre che idiota: alla fine avremmo solo 8 KHz di banda passante utile invece dei 16 KHz possibili.

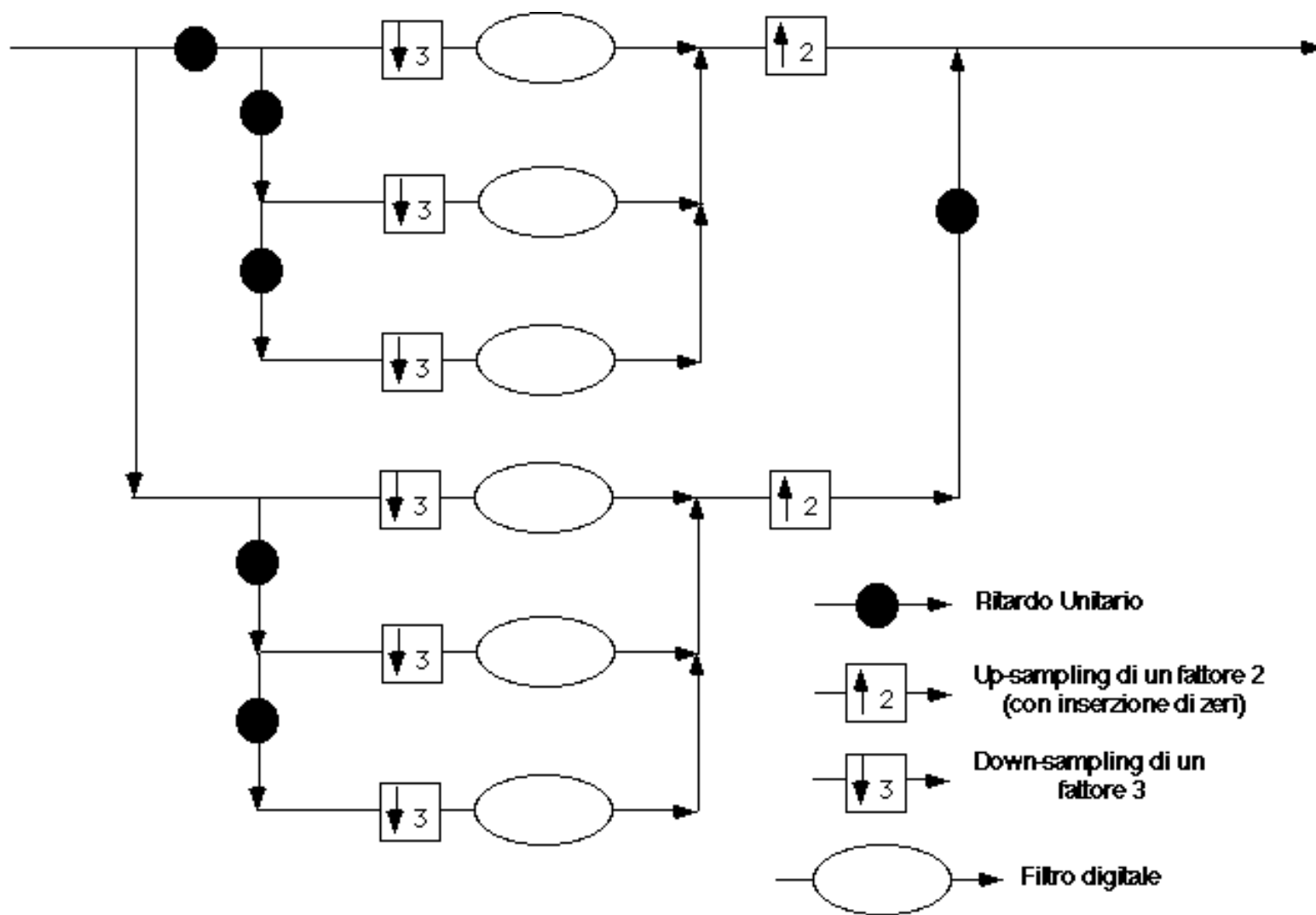
Poiché il sovracampionamento termina con un filtraggio passa basso e il sottocampionamento inizia con un filtraggio passa basso uno dei due filtri può essere eliminato utilizzando un filtro che operi a 96 KHz con frequenza di taglio sotto i 16 KHz.

Questo procedimento non presenta particolari inconvenienti fino a che il rapporto è formato da numeri piccoli, ma è possibile semplificarne l'implementazione con una tecnica alternativa (detta **polyphase filtering**).

Lavorando di carta e penna nel dominio della trasformata Z si riesce a riscrivere il processo di cui sopra in un altro equivalente in cui invece di un filtraggio alla frequenza più elevata si devono fare molti filtri alla frequenza più bassa, ognuno calcolato con i coefficienti opportuni. Nel caso della conversione da 48 KHz e 32 KHz abbiamo 6 filtri che operano direttamente a 16 KHz. Gli input dei filtri sono i campioni originali presi uno ogni tre, opportunamente sfalsati (ogni sottosequenza di campioni entra in due filtri).

I risultati dei filtri vengono opportunamente ricombinati. Se qualcuno vuole approfondire i dettagli e seguire la dimostrazione può consultare il capitolo sulle conversioni nel testo dello Zolzer. Nella figura si vede un esempio di conversione con rapporto 3/2 con filtri **polyphase** (Zolzer p. 225). Si noti che le operazioni iniziali di ritardo e sottocampionamento si risolvono

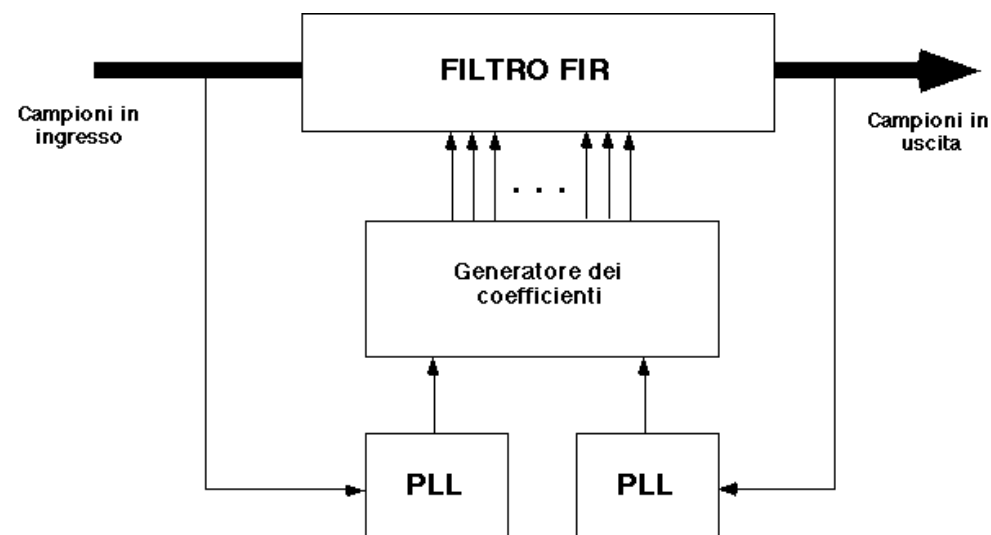
in un instradamento dei campioni; l'operazione finale equivale a prendere alternativamente un campione da una linea e uno dall'altra.



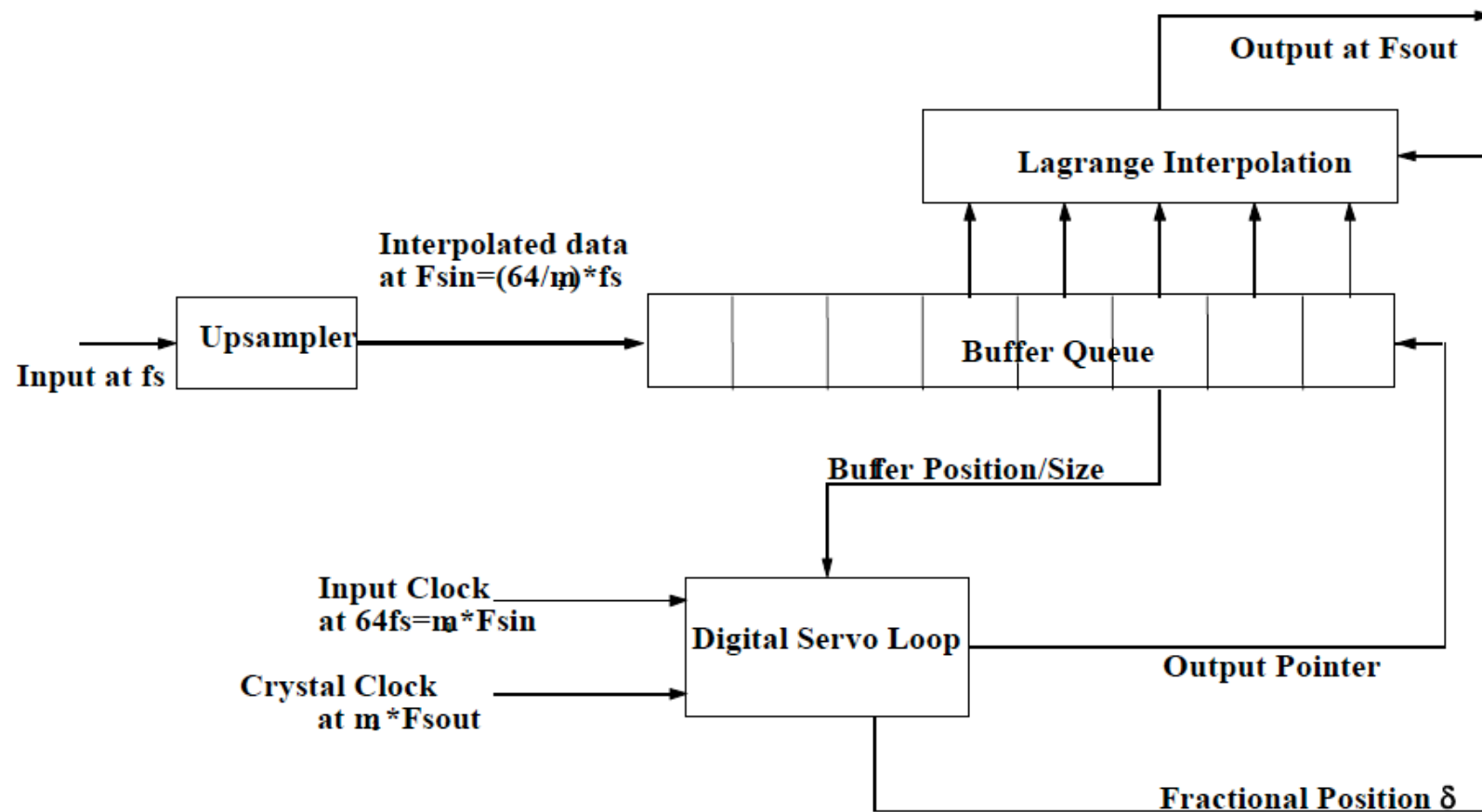
Conversioni asincrone

Il caso più sfavorevole si presenta quando le due frequenze non stanno un rapporto semplice oppure il rapporto di frequenza deve poter variare durante la conversione. Il modello da cui si parte per questa conversione è quello di un passaggio dal dominio digitale a quello analogico seguito da un nuovo campionamento. Questa operazione viene compiuto davvero tutte le volte che prendendo il segnale dall'uscita analogica di un apparecchio digitale si registra di nuovo in digitale. Ed è anche quello che fanno i pirati per fare le copie da vendere o da mettere in rete 5 minuti dopo l'uscita sul mercato dei cosiddetti CD "protetti". Quando la conversione deve avvenire all'interno di una apparecchiatura si cerca invece di eliminare il passaggio all'analogico rendendolo virtuale. In pratica ad ogni istante i campioni dell'ingresso vengono utilizzati per pilotare un filtro interpolatore a coefficienti variabili (determinati in base alle specifiche o ad un'analisi dei clock in ingresso e in uscita).

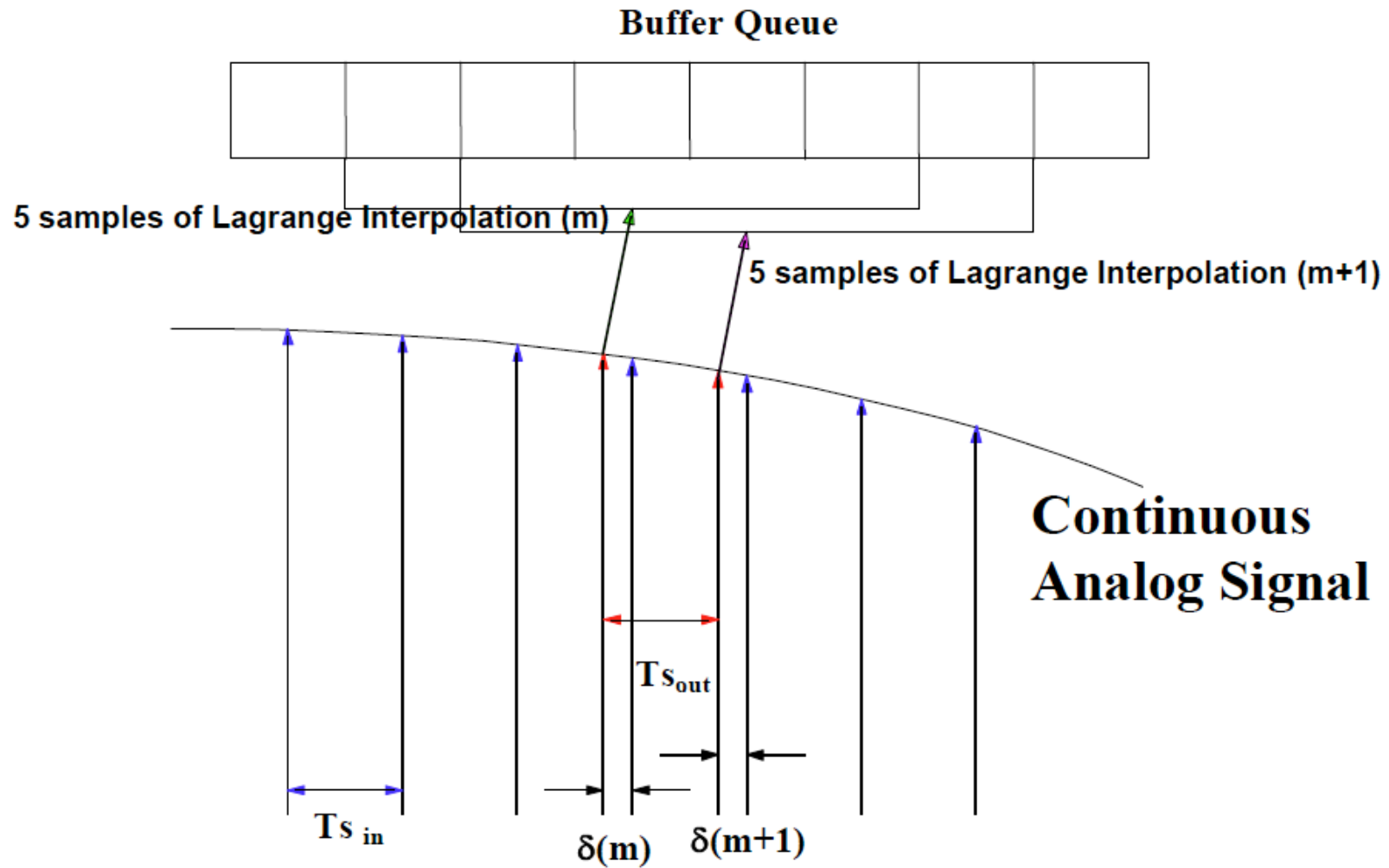
Vediamo la possibile struttura di un convertitore asincrono (figura ripresa dal testo di Watkinson p.148).



Un'altra possibilità è quella di interpolare il segnale in ingresso e valutare la funzione risultante negli istanti stabiliti dal clock di uscita.



La figura successiva mostra in maggiore dettaglio come avviene il processo di interpolazione



Un approccio del genere, come ogni computazione numerica presenta tre tipi di errore

- **Errore inerente**, dovuto alla applicazione della funzione da calcolare alla imprecisione dei dati di ingresso.
- **Errore algoritmico** dovuto alla finitezza della parola usata per effettuare i calcoli.
- **Errore analitico** o di **troncamento**, dovuto alla imprecisione della formula usata per il calcolo.

Inoltre è presente un'ulteriore errore, specifico al problema, dovuto all'inevitabile ritardo con cui i valori dei δ seguono le variazioni istantanee del clock in ingresso

Il primo tipo di errore implica un inevitabile lieve aumento del livello di rumore nel segnale trattato.

Il secondo e il terzo tipo di errore possono essere ridotti fino ad essere annullati aumentando la lunghezza della parola usata per fare i calcoli (per esempio usando 32 bit) e usando algoritmi più raffinati e costosi; entrambe le soluzioni rese possibili dal progredire della tecnologia.

Per l'ultimo errore non esiste una soluzione specifica.

A cosa serve e a cosa non serve la conversione asincrona

La conversione asincrona è ormai disponibile implementata su singoli chip a basso costo e viene spesso considerata la soluzione a molti dei problemi legati all'interfacciamento di segnali digitali. Anche se, grazie alla maggiore potenza di calcolo disponibile nei chip, non è più un'operazione dannosa come una volta è bene tuttavia avere chiaro cosa ci possiamo aspettare da una tale soluzione.

La conversione asincrona ...

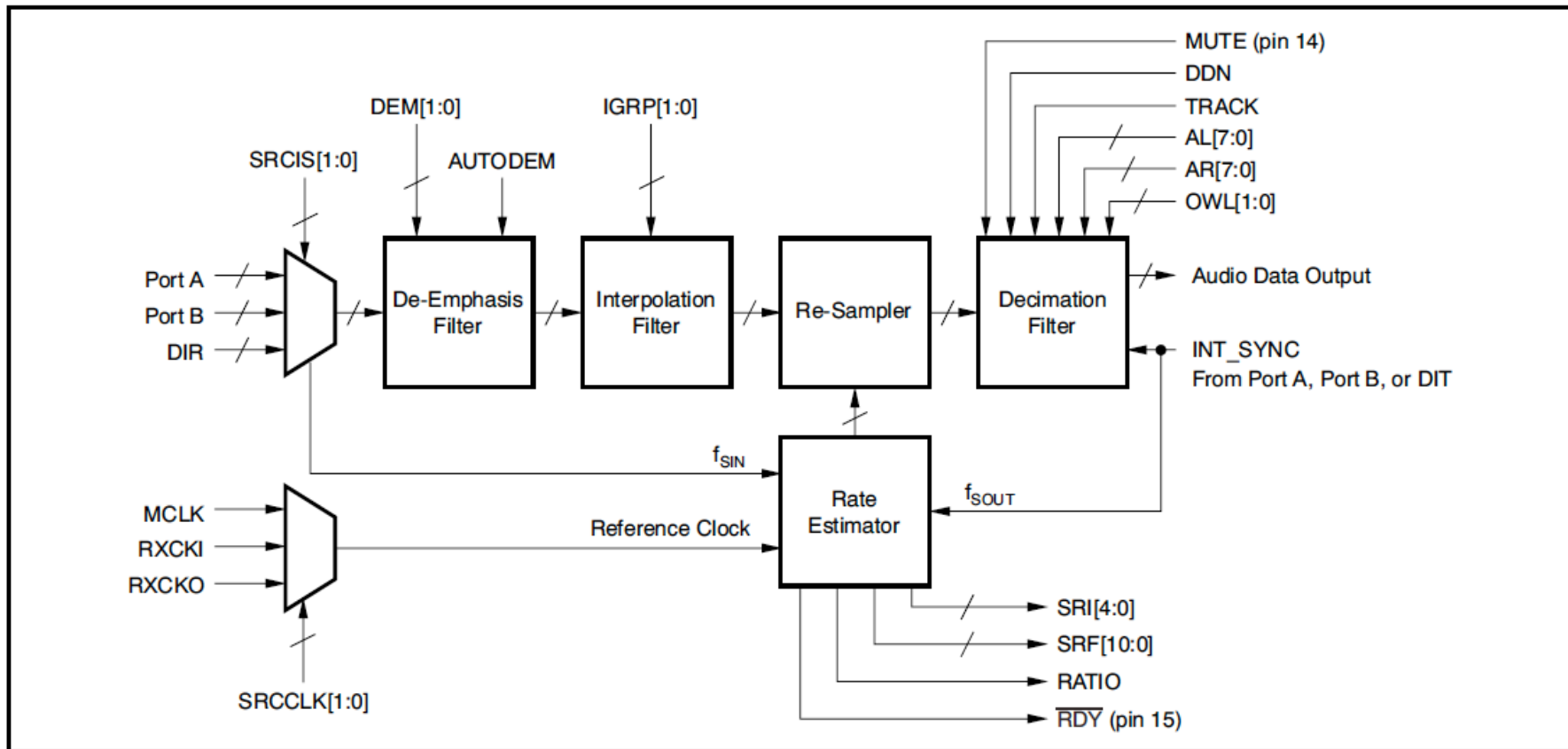
- È indispensabile per miscelare in tempo reale segnali provenienti da apparecchiature con clock non sincronizzati.
- È indispensabile per avere un segnale con clock stabile a partire da un segnale con clock altamente variabile.
- È utile per semplificare drasticamente gli stadi di ingresso delle apparecchiature digitali.
- Non serve a aumentare la qualità di un segnale audio grazie ad una maggiore frequenza di campionamento. Infatti anche se il **data rate** aumenta, la **quantità di informazione** resta quella del segnale originario e il suono non può che peggiorare o al limite rimanere della stessa qualità.
- Non serve a ridurre gli effetti del **jitter** sul suono, si può infatti dimostrare (e lo faremo nel capitolo dedicato al **jitter**) che la conversione asincrona stabilizza il clock congelando nel segnale gli effetti del **jitter** senza alcun miglioramento dal punto di vista della qualità sonora.

Vediamo ora lo schema e le specifiche di un ottimo convertitore asincrono commerciale.

SRC4392

Si tratta di un convertitore asincrono integrato che accetta in ingresso e fornisce in uscita frequenze di campionamento comprese tra 20 e 218 KHz.

- **Diagramma funzionale**



- **Specifiche**

| | |
|-------------------------------|-----------------|
| Frequenza di campionamento | da 20 a 216 KHz |
| Bit in ingresso | fino a 24 |
| Gamma Dinamica (44.1⇒192 KHz) | 138 dB |
| THD+N a 0 dB (44.1⇒192 KHz) | -137 dB |

Come si vede le specifiche sono strabilianti, molto migliori di quelle dei convertitori (non essendovi segnale analogico non c'è rumore termico). Bisogna però notare che le misure sono statiche, sarebbe stato interessante misurare l'effetto della modulazione del segnale in uscita

dovuta al ritardo nella variazione dei coefficienti dell'interpolazione in presenza di un clock di ingresso altamente instabile dato che questa è la situazione tipica in cui l'uso di un convertitore asincrono è realmente necessario.

Conversioni offline

Un tipo di conversione concettualmente simile a quella asincrona ma in pratica completamente diversa è la conversione di un rapporto **qualsiasi** effettuata numericamente sul segnale digitale trattato come **un insieme di dati numerici memorizzati in un computer**. La conversione non avviene in tempo reale ma ci si lita a costruire un nuovo file con i dati convertiti a partire da un file con i dati originari. In questo caso non vi sono problemi di clock perché i dati sono tutti memorizzati, si può usare un'aritmetica ad alta precisione (i **56+8** bit standard della doppia precisione) e non vi sono limiti alla potenza di calcolo.

NB ovviamente anche in questo caso non si può aumentare la qualità di un segnale audio con la maggiore frequenza di campionamento o la maggiore profondità di quantizzazione (in altre parole **non si può cavare sangue da una rapa**).