

3.3. Constant-Q Transform

Una soluzione agli inconvenienti trattati nel paragrafo precedente è stata proposta da J.C. Brown nel 1991 ed è nota come Constant- Q transform [9]. Seppure sia stata formalizzata analiticamente, la sua implementazione come codice sorgente C non è di dominio pubblico. Per scrivere la *routine* che implementa l'analizzatore di spettro è stata sperimentata la teoria di J.C. Brown e si sono cercate delle semplificazioni.

Ciascun quarto di tono viene analizzato dalla trasformata discreta di Fourier della sequenza d'ingresso troncata ad una lunghezza diversa:

$$X_k = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k(n) \cdot x(n) e^{-j \frac{2\pi Q n}{N_k}} \quad (3.15)$$

dove $w_k(n)$ è una sequenza di finestra della durata opportuna di N_k campioni, per esempio la finestra di Hamming definita in (3.12); mentre il fattore di merito Q è definito come il rapporto tra il valore nominale della frequenza e il passo di campionamento in frequenza:

$$Q = \frac{f}{\delta f} \quad (3.16)$$

Il fattore di merito della progressione delle frequenze f_k (3.3) assume un valore costante indipendente dalla frequenza:

$$Q = \frac{f_k}{f_{k+1} - f_k} = \frac{1}{2^{1/24} - 1} \approx 34 \quad (3.17)$$

Per semplicità di notazione si definisce la sequenza complessa di finestra:

$$\omega_k(n) = w_k(n) \cdot e^{j \frac{2\pi Q n}{N_k}} \quad (3.18)$$

La (3.15) diventa la sommatoria del prodotto di due segnali nel tempo che per il teorema di Parseval può essere valutata nel dominio della frequenza come:

$$X_k = \frac{1}{N_k} \sum_{n=0}^{N_k-1} x(n) \cdot \omega'_k(n) = \frac{1}{N_k} \sum_{m=0}^{N_k-1} X_{m,k} \cdot \Omega'_{m,k} \quad (3.19)$$

dove $X_{m,k}$ e $\Omega_{m,k}$ sono i coefficienti calcolati con FFT delle sequenze $x(n)$ e $\omega(n)$ per $n=0,1,\dots,N_k$. Per analizzare un tono diverso è necessario ricalcolare le FFT su sequenze di un'altra lunghezza:

$$N_k = \frac{f_{SR}}{\delta_k} = \frac{f_{SR}}{f_k} \cdot Q \quad (3.20)$$

Questo metodo è computazionalmente inadeguato ad implementare un analizzatore di spettro che faccia un'analisi *real-time* della funzione di trasferimento del locale d'ascolto, e quindi non permette di percepire immediatamente l'effetto delle modifiche sui parametri dei filtri correttivi. Sarebbe più efficiente calcolare una sola operazione di FFT su un numero di campioni indipendente da f_k . La stessa J.C. Brown suggerisce [10] che è possibile calcolare l'operazione di FFT rispetto al caso peggiore $N_{max} = f_{SR} Q / f_{min}$, ma per un'analisi a 24-esimi d'ottava senza *overlapping*, la FFT verrebbe calcolata ogni 2,13 secondi.

Si è sperimentato il calcolo degli $X_{m,k}$, $\Omega'_{m,k}$ con un numero N arbitrario di campioni (comunque una potenza di due).

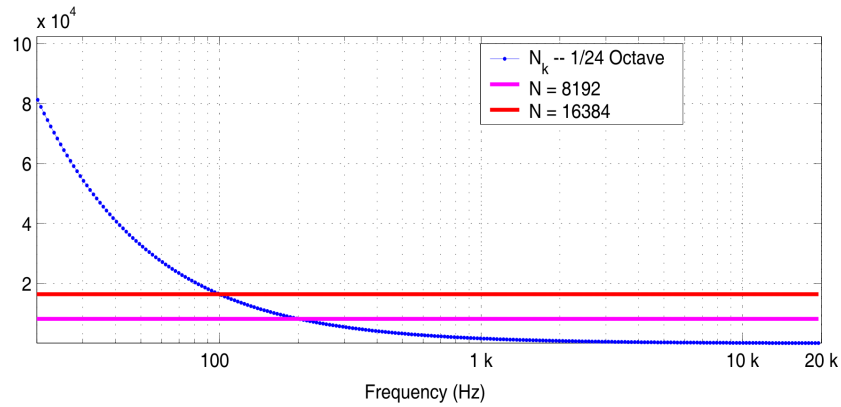


Figura 3.9.: Lunghezza delle sequenze nell'analisi a quarti di tono

La figura 3.9 illustra, al variare della frequenza, le lunghezze N_k definite dalla (3.20) per $f_{SR}=44100$. Nel paragrafo precedente è stata documentata la sperimentazione della FFT di una sequenza di 8192 campioni. Se si fissa nel grafico questo valore, si osserva che le sequenze relative alle $f_k > 200$ Hz sarebbero più corte, mentre 8192 campioni non sono sufficienti a contenere le sequenze relative alle $f_k < 200$ Hz. Come mostra la figura 3.10, questo significa che la risoluzione della FFT sarebbe migliore dell'intervallo di quarto di tono sopra i 200 Hz, ma peggiore al di sotto.

3.3. Constant-Q Transform

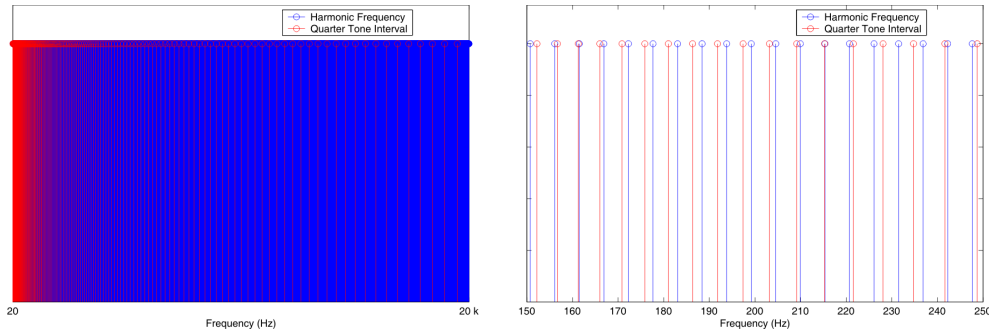


Figura 3.10.:

Tuttavia N va fissato scegliendo un compromesso tra la risoluzione e la periodicità con cui si intende aggiornare l'analizzatore di spettro. Si osserva che il valore $N = 16384$ permette una risoluzione spettrale che è maggiore dell'intervallo di quarto di tono già sopra a 100 Hz e un *refresh* del grafico ogni 370 ms. La figura 3.11 mostra i valori di N_k per la sequenza dei semi-toni ($Q \approx 17$) ed i toni ($Q \approx 8.5$) della scala temperata.

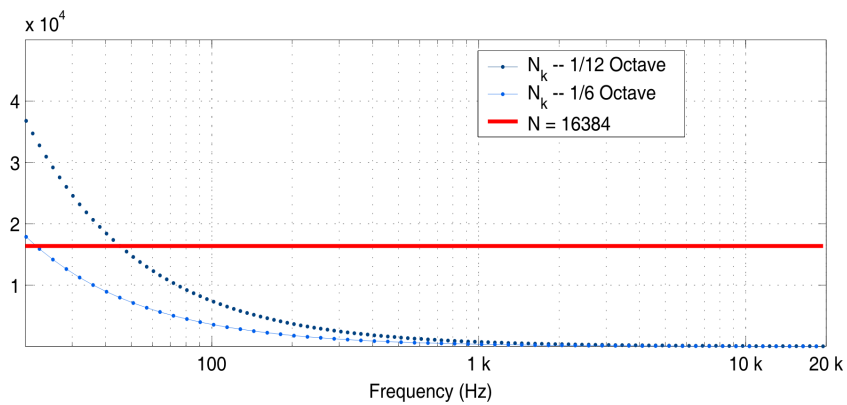


Figura 3.11.: Lunghezza delle sequenze per gli intervalli di tono e semi-tono.

Per sequenze finite di 16384 campioni, si osserva che la trasformata discreta

ha una risoluzione migliore di un quarto di tono a partire dalla terza ottava, migliore di un semi-tono a partire dalla seconda ottava e migliore di un tono su tutto la banda audio.

È stata scelta la lunghezza $N=16384$, indipendente da f_k , che permette di calcolare tutti gli X_k con una sola operazione di FFT sul segnale d'ingresso:

$$X_m = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi mn}{N}} \quad \Omega_{m,k} = \sum_{n=0}^{N-1} \omega_k(n) e^{-j \frac{2\pi mn}{N}} \quad (3.21)$$

La dipendenza da N_k è rimasta solo nella lunghezza della sequenza di finestrazione $\omega_k(n)$ ma non nel calcolo della sua FFT. Tuttavia si osservi che gli $\Omega_{m,k}$ non dipendono dalla sequenza $x(n)$ da analizzare e si possono calcolare una sola volta, *off-line*, quando l'applicazione viene inizializzata. La sequenza viene troncata ad N campioni nell'analisi di quelle frequenze tali che $N_k > N$; la sequenza viene estesa con l'operazione di *zero-padding* per quelle in cui $N_k < N$. Per inciso si fa notare che è stato necessario modificare la *routine* di FFT perché l'operazione di *zero-padding* cambia il valore medio del segnale e quindi anche la componente continua del suo spettro. La condizione per cui i coefficienti $\Omega_{m,k}$ siano reali è che la sequenza nel tempo sia coniugata e simmetrica: $\omega_k(n) = \omega'_k(-n)$. Questo si verifica se la sequenza di Hamming $w_k(n)$ è centrata rispetto all'intervallo di osservazione e nulla al di fuori di $[N/2 - N_k/2, N/2 + N_k/2]$. Alla luce di queste considerazioni la (3.19) può essere riscritta nella forma:

$$X_k = \sum_{m=0}^{N-1} X_m \cdot \Omega_{m,k} \quad (3.22)$$

È stato rimosso anche il termine $1/N$: è stata calcolata la sommatoria anziché

la media ma si ottenuto lo stesso risultato perché se si è usato come segnale di test il rumore rosa anziché il rumore bianco, preferito per quanto già citato nel paragrafo precedente. La trasformata Q costante è stata calcolata dunque come la somma pesata del risultato della FFT: per dimostrare l'effetto dei pesi si esplicitano i coefficienti $\Omega_{m,k}$ con l'aiuto della (3.18):

$$\begin{aligned}\Omega_{m,k} &= \sum_{n=0}^{N-1} w_k(n) \cdot e^{j\frac{2\pi Qn}{N_k}} e^{-j\frac{2\pi mn}{N}} = \\ &= \sum_{n=0}^{N-1} w_k(n) \cdot e^{-j2\pi(m-\frac{Q \cdot N}{N_k})\frac{n}{N}}\end{aligned}\quad (3.23)$$

e si definisce:

$$\frac{Q \cdot N}{N_k} = \frac{f_k}{f_{SR}} \cdot N = \frac{f_k}{\delta f} \quad \rightarrow \quad \hat{k} = \text{round}\left\{\frac{f_k}{\delta f}\right\}\quad (3.24)$$

Segue che:

$$\Omega_{m,k} = \sum_{n=0}^{N-1} w(n, k) \cdot e^{-j2\pi(m-\hat{k})\frac{n}{N}} = W_{m-\hat{k}}\quad (3.25)$$

dove W_m sono i coefficienti discreti di Fourier della sequenza di Hamming di durata N_k , mentre l'effetto dell'esponenziale è quello di sintonizzarne lo spettro attorno alla frequenza f_k . Si è dimostrato che la trasformata a Q costante è equivalente ad un banco di filtri con banda passante pari a 24-esimi d'ottava, ma più efficiente ed elegante computazionalmente. Nelle figure 3.12 e 3.13 sono state riportate solo le frequenze definite dalla progressione geometrica di ragione $2^{1/3}$ per rendere la rappresentazione intellegibile, e le sequenze di $N=8192$ campioni: si può verificare l'equivalenza con un banco di filtri passa-banda a terzi d'ottava.

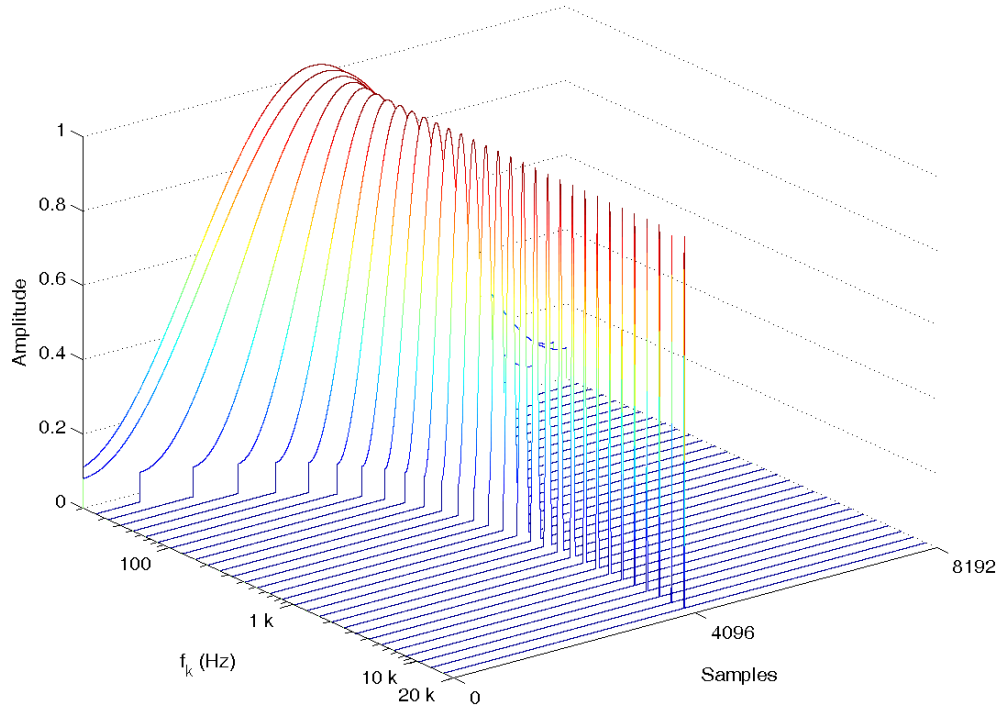


Figura 3.12.: Sequenze di Hamming nel dominio del tempo: $w_k(n)$.

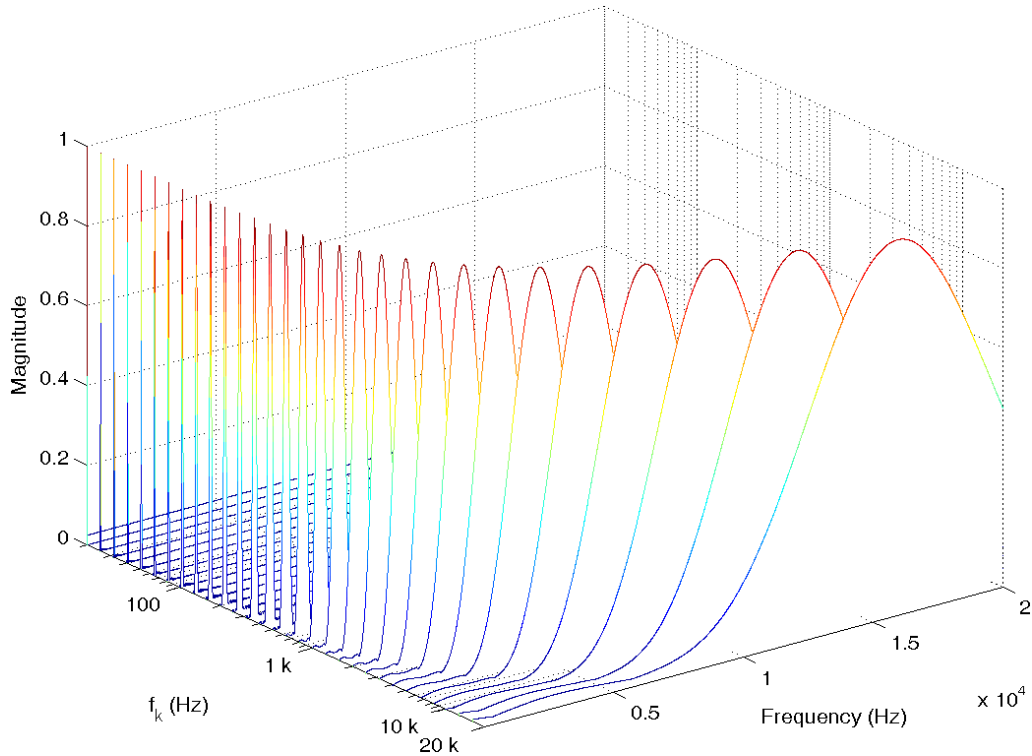


Figura 3.13.: Sequenze di Hamming nel dominio della frequenza: $W_{m-\hat{k}}$.

Sostituendo la (3.25) nella (3.22) si è giunti ad una formula che trae vantaggio dall'efficienza computazionale della FFT ma analizza lo spettro per le frequenze definite da una progressione geometrica:

$$X_k = \sum_{m=0}^{N-1} X_m \cdot W_{m-\hat{k}} \quad (3.26)$$

È stato possibile tuttavia ridurre ancora il numero di moltiplicazioni ed addizioni osservando che gli $W_{m-\hat{k}}$ assumono un'ampiezza trascurabile all'aumentare di $|(m-\hat{k})|$ e non danno contributo alla sommatoria. Sono stati trascurati i valori inferiori alla soglia 0.054 [11]; questo ha permesso di limitare

3.3. Constant-Q Transform

gli estremi della sommatoria per $|(m - \hat{k})| < C_k$, dove C_k rappresenta la metà del numero di operazioni sufficienti al calcolo di X_k . Come si può osservare dalla figura 3.14, la trasformazione (3.26) aggiunge un onere computazionale ridotto al numero di operazioni della FFT.

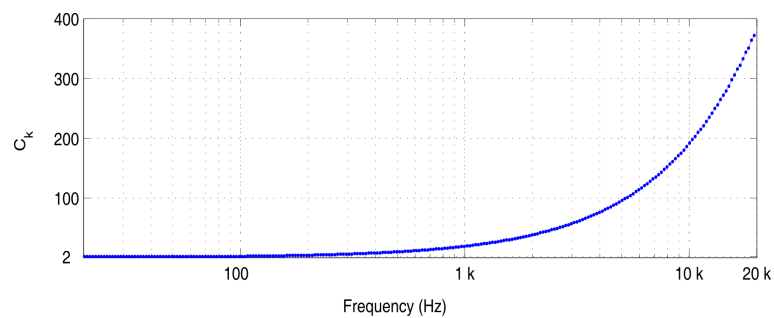


Figura 3.14.: Contributi non nulli nella 3.26.

Per sperimentare l'efficienza della trasformata Q costante si è posto l'analizzatore di spettro all'uscita dell'equalizzatore il cui ingresso è stato sollecitato da rumore rosa. Agendo sui parametri dei filtri è stato possibile simulare qualunque funzione di trasferimento. Come dimostra la figura 3.15

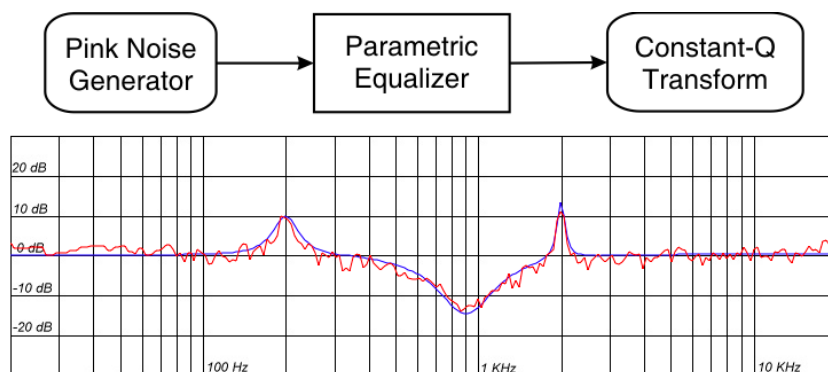


Figura 3.15.: Test d'efficienza della CQT mediante rumore rosa filtrato dall'equalizzatore parametrico.

l'analizzatore di spettro è in grado di misurare qualunque non linearità d'ampiezza nella risposta.

Si osservi che la figura mostra un'istantanea dell'analizzatore di spettro e le irregolarità della risposta sono fluttuazioni, dell'ordine di pochi dB, attorno al valore reale e sono da attribuirsi alla natura aleatoria del processo di rumore: lo *smooth* della risposta si realizza mediando nel tempo i valori che assume ciascun X_k .

Riguardo alla figura 3.7 si era fatto riferimento ad un segnale composto da cinque componenti armoniche tutte della stessa ampiezza. La figura 3.16 mostra come quello stesso segnale viene rilevato con la trasformata Q costante: si può osservare che fornisce un risultato più attendibile.

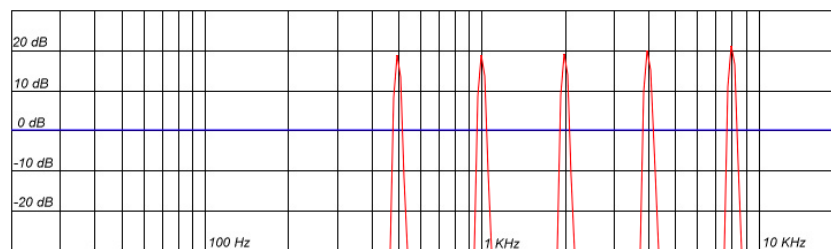


Figura 3.16.: Analisi di un segnale con cinque armoniche.

Terminata la descrizione della trasformata Q costante, può essere utile riepilogare il procedimento svolto per implementare l'analizzatore di spettro:

- per ciascun quarto di tono f_k è stata creata una sequenza di Hamming di lunghezza N_k , simmetrica rispetto ad un *buffer* di N campioni;
- è stata calcolata la FFT di ciascuna sequenza e lo spettro è stato sintonizzato attorno alla frequenza f_k ;
- in regime di analisi si calcola la FFT del segnale d'ingresso troncando la sequenza ogni N campioni;

- si calcolano i coefficienti X_k con la formula (3.26) trascurando i contributi non significativi al di sotto di una soglia;
- si traccia il grafico dello spettro sull'asse delle frequenze logaritmico interpolando gli X_k .

Si osservi che la trasformata Q costante non è invertibile: non esiste un'espressione per l'antitrasformata.