

Fondamenti di elaborazione numerica dei segnali

Fabio Rocca

*Dipartimento di Ingegneria Elettronica
e Informazione
Politecnico di Milano*

Milano, 2004

Indice

Prefazione	xiii
1 Il Campionamento	1
1.1 Introduzione	1
1.2 Ricostruzione di un segnale	6
1.3 Campionamento passa banda	7
1.4 Il filtro antialias	8
2 La trasformata di Fourier discreta	11
2.1 Campionare in tempo e in frequenza	11
2.2 Proprietà della DFT	13
2.3 Filtraggi con la DFT	14
2.3.1 Convoluzioni aperiodiche: sequenze limitate . . .	15
2.3.2 Sequenze illimitate: overlap-save; overlap-add . .	15
2.4 Antitrasformazione di Fourier discreta	18
2.4.1 Il derivatore e la costante di tempo	20
2.5 Interpolazione con la DFT	22
2.5.1 Inserzione di zeri e repliche spettrali	22
2.6 Il calcolo della DFT: la FFT	25
2.6.1 Le farfalle	27
2.6.2 Occupazione di memoria e <i>bit reversal</i>	28
2.7 Filtraggio di sequenze illimitate	29
2.7.1 Costo computazionale della FFT	29
2.7.2 Segmentazione di sequenze illimitate	30
2.8 Trasformata di Fourier in base 4	31
3 Ricostruzione di segnali continui	33
3.1 Introduzione: sequenze limitate	33

INDICE

3.2	Ricostruzione di polinomi	37
3.3	Interpolatori derivati dal mantenitore	38
3.3.1	Interpolatore cubico troncato	39
3.3.2	Interpolatori con filtro ideale troncato	40
3.4	Rigidità e sovraelongazioni	41
3.5	La linea elastica o "spline"	43
3.5.1	Analisi dello "spline"	43
3.5.2	L'equazione dello spline	44
3.5.3	La risposta all'impulso dello spline	46
3.6	Il campionamento irregolare	48
3.6.1	Conversione di frequenze di campionamento	48
3.6.2	Interpolazione irregolare con la funzione <i>sinc</i>	50
4	La trasformata z	53
4.1	Introduzione: sequenze e polinomi	53
4.2	Sistemi lineari discreti	56
4.3	Sinusoidi campionate; zeri	56
4.4	Il piano z	59
4.4.1	Le sequenze nel piano z	59
4.4.2	Zeri a fase minima e fase massima	60
4.4.3	Sequenze a fase minima, massima o mista	63
4.4.4	Trasformata di Fourier e trasformata z	64
4.5	Poli	65
4.6	Equazioni alle differenze finite	68
4.7	Filtri passa banda	70
4.7.1	Sfasatori puri	71
4.7.2	Oscillatori, divani, splines	74
5	Filtri numerici	79
5.1	Introduzione	79
5.2	Finestre rettangolari e triangolari	80
5.3	Finestre cosinusoidali	82
5.4	Finestre diverse	84
5.5	Progetto con la tecnica MINIMAX	85
5.6	Lunghezza dei filtri FIR passa - basso	87
5.7	Filtri IIR	89
5.7.1	Derivazione numerica	89
5.7.2	La trasformazione bilineare	95

INDICE

5.8	L'invarianza della risposta impulsiva	98
5.9	Filtri da combinazione di sfasatori puri	99
5.10	Filtri di Butterworth, Chebyshev, ellittici	102
5.11	Sintesi di filtri tutti poli dall'autocorrelazione	103
5.11.1	Le equazioni di Yule Walker	103
5.11.2	La ricorsione di Levinson	108
5.11.3	La sintesi del filtro	110
6	Campionamento in frequenza	113
6.1	Introduzione	113
6.2	Progetto dei filtri a campionamento in frequenza	117
6.3	Quantizzazione dei coefficienti	123
6.4	Filtri a campionamento in frequenza irregolare	126
7	Decimazione ed interpolazione	129
7.1	Introduzione	129
7.2	Decimazione e interpolazione	131
7.3	I filtri polifase	139
7.4	Banchi di filtri	140
7.4.1	La matrice di modulazione	140
7.4.2	La ricostruzione perfetta	144
7.4.3	Banchi di due filtri	145
7.4.4	I filtri in quadratura a specchio (QMF)	146
7.4.5	Equivocazione e ricostruzione perfetta	148
7.5	Banchi di M filtri	151
8	Effetti di quantizzazione	153
8.1	Strutture dei filtri numerici	153
8.2	Approssimazioni nella realizzazione di filtri	154
8.2.1	La quantizzazione dei coefficienti	154
8.2.2	Sfasatori puri	158
8.3	La quantizzazione dei segnali	159
8.3.1	Rumore di quantizzazione dei segnali	160
8.3.2	Effetti di trabocco (overflow)	164
8.4	Quantizzazione e DFT	165
8.5	I cicli limite	165

INDICE

9	La trasformata di Hilbert	169
9.1	Trasformate di sequenze causali	169
9.2	Sfasatori puri di $\pi/2$	172
9.3	Sequenze periodiche e segnali continui	174
9.3.1	Singularità nella trasformazione di Hilbert	175
9.4	Funzioni a fase minima	176
9.5	Il segnale analitico	179
10	Le strutture riverberanti	181
10.1	Introduzione	181
10.2	Strutture riverberanti e sistemi numerici	183
10.3	Cascata di tronchi di linea	186
10.4	Il filtro a traliccio ricorsivo	187
10.5	La ricorsione di Levinson	190
10.5.1	Aggiunta di tronchi di linea	194
10.6	Calcolo dei coefficienti di riflessione	194

Prefazione

Il corso di "Fondamenti di Elaborazione numerica dei segnali" viene tenuto per gli allievi del corso di laurea in Ingegneria delle Telecomunicazioni al Politecnico di Milano dal 2003 - 2004. Questi allievi hanno seguito in precedenza un corso di *Segnali* dove vengono studiati i sistemi lineari, la trasformazione di Fourier, i principi di teoria delle probabilità ed i processi stocastici. Inoltre, nel terzo anno sono disponibili i corsi di "Algoritmi e Circuiti di Comunicazione", "Elaborazione e Trasmissione delle Immagini", dove vengono sviluppate con maggiore respiro le possibili applicazioni. La verifica viene fatta con un esame scritto; il compito a metà semestre copre il programma fino ai filtri FIR inclusi.

Nel testo, ho cercato per quanto possibile di partire da considerazioni elementari dando immagini intuitive (*spesso scritte in corsivo*) di molti tra gli argomenti trattati. Certamente la trattazione analitica è una guida più sicura e la via intuitiva non è sempre la più semplice, ma poi è più facile da memorizzare. Mi auguro poi che lo studente possa avere strumenti di tipo ingegneristico più che matematico per comprendere le relazioni intercorrenti tra le ipotesi presunte ed i risultati ottenuti e possa valutare cosa può accadere nei frequenti casi in cui queste ipotesi non sono più attendibili.

Il corso è di prevalente natura metodologica, anche se vengono sempre proposte delle applicazioni.

Questo testo è stato fatto per essere un ausilio alla didattica in aula, piuttosto che uno strumento per l'autodidattica. In questo caso, consiglio di cominciare con i libri di McClellan o di Mitra, inseriti tra i riferimenti bibliografici, ricchi di esempi *MATLABTM*. Il testo di Proakis

PREFAZIONE

e Manolakis è più completo, ma può risultare più difficile ad una prima lettura.

In questo testo non si troveranno esercizi perchè questi, insieme a complementi e a numerosi testi d'esame sono contenuti nel volume: *"Elaborazione Numerica dei Segnali: Temi d'esame '96/00"* a cura di U. Spagnolini, F. Rocca, L. Zanzi, Edizioni CUSL, Milano.

Aggiungo una breve bibliografia commentata perchè lo studente possa trovare riferimenti completi per ogni argomento del corso.

Il corso ha un suo sito web: <http://www.elet.polimi.it/dsp/courses/fens/> dove gli studenti troveranno il codice *MATLABTM*, reso disponibile agli iscritti al corso, indirizzi di posta elettronica per semplificare l'interazione docente/studente, esercizi vari con i relativi codici .m, data e luogo degli esami, esiti, soluzioni dei compiti, e materiale utile in generale.

Bibliografia.

- Mitra, S., 1998, *Digital Signal Processing, a computer - based approach*, McGraw-Hill, pp 864. Libro elementare, ricco di esempi basati su *MATLABTM*.
- Mc Clellan, J., Schafer, R., Yoder, M., 1998, *DSP first, A Multimedia Approach*, Prentice Hall, pp.514. Libro molto elementare, con esempi basati su *MATLABTM*.
- Proakis J., Manolakis D., 1995: *Digital Signal Processing*, Prentice Hall, 947 pp. e quattro appendici. Libro introduttivo didattico, semplice ma molto diffuso.
- Mitra S., Kaiser J., 1993: *Handbook for Digital Signal Processing*, Wiley Interscience, 1168 pp. Completo manuale che copre ampiamente il testo.
- Fliege, N., 1994, *Multirate Digital Signal Processing*, J. Wiley & Sons, pp. 340. Un buon libro sui filtri polifase e sui problemi di decimazione e interpolazione.

Capitolo 1

Il Campionamento

1.1 Introduzione

I segnali naturali sono definiti per valori continui della variabile tempo o spazio: è possibile misurare il loro valore in ogni istante o posizione, in un intervallo assegnato.

Volendo fare un'elaborazione è necessario effettuare un **campionamento** e cioè passare da una funzione definita su un asse (ad esempio spaziale) continuo, ad una sequenza di numeri corrispondenti alle misure del segnale in posizioni assegnate (segnale campionato). L'operazione di campionamento è corretta se è possibile ricostruire il segnale continuo a partire dai campioni misurati. E' ovvio che in generale ciò è impossibile.

Per fissare le idee, si consideri un sistema di allarme antincendio (le dimensioni possono essere qualsiasi: una foresta, un edificio, un sistema elettronico); è necessario campionare e controllare le temperature ovunque per lanciare l'allarme quando è necessario. Se avessimo a disposizione un solo ambiente, illimitato, sarebbe ragionevole spaziare i sensori di temperatura ad intervallo regolare, dipendente dalle dimensioni dei focolai di incendio che si vogliono poter rivelare (l'ordine di grandezza dell'intervallo di campionamento sarà il chilometro per le foreste, il decimetro per i sistemi elettronici).

Usualmente, si suppone di dover campionare un segnale definito su un dominio spaziale (o temporale) illimitato e di essere in condizioni di stazionarietà, e cioè in una situazione in cui le caratteristiche statistiche del segnale studiato non cambino con lo spazio (o con il tempo). Sia

CAPITOLO 1 IL CAMPIONAMENTO

$x(t)$ il segnale da campionare (nel caso precedente, la temperatura x , funzione della posizione t del sensore); il segnale sia reale e supponiamo di volerne fare misure ad intervallo T ; cioè si supponga nota la sequenza x_n ; $0 \leq n \leq N - 1$:

$$x_n = x(nT + T_o) \quad (1.1)$$

Il segnale continuo $\hat{x}(t)$ ricostruito a partire dalla sequenza dipenderà in generale da T_o e cioè dalla fase del reticolo di campionamento (nel caso del sistema antincendio, i sensori di temperatura siano disposti a distanza regolare T , e sia T_o l'ascissa del primo sensore). Il campionamento è corretto quando la ricostruzione non dipende da T_o (vogliamo valutare correttamente la consistenza di un incendio anche se il focolaio capita tra due sensori).

Gli errori della ricostruzione sono chiamati equivocazione (aliasing) e intuitivamente si possono dividere in due categorie: errori dovuti alla mancata registrazione di un evento, che è caduto "tra due misure" (*l'incendio non osservato*) o viceversa misure di eventi che, pur essendo di durata molto limitata, vengono interpretati come eventi lunghi circa T (vedi fig. 1.1) (si pensi al detto popolare "prendere lucciole per lanterne"). Se si vogliono evitare fenomeni di equivocazione sarà opportuno che anche gli eventi di durata minima siano campionati almeno due volte perchè siano delimitati (nel tempo o nello spazio) e ne venga evitata sia la perdita che l'esagerazione. Diventa quindi essenziale garantire che nel segnale misurato ad intervallo T siano assenti eventi brevi che creerebbero equivocazione. Questo si ottiene mediando (smussando) il segnale prima del campionamento e cioè **prefiltrando il segnale da campionare** (filtraggio antiequivocazione o antialias); in altre parole, lo strumento la cui misura è destinata ad essere campionata ad intervallo T deve essere "sensibile" non al valore locale del segnale, bensì al valore "medio" del segnale in quell'intervallo.

In questo modo l'errore di equivocazione viene eliminato sistematicamente. Non sempre ciò è facile o possibile. Ad esempio, sarebbe facile far equivocare il sistema d'allarme antiincendio considerato precedentemente, anche se questo fosse complesso; basterebbe accostare un fiammifero ad un sensore. Infatti in questo caso il campione di temperatura misurato dal sensore avrebbe valore elevato e molto differente da quello medio dell'ambiente; tuttavia, se vi è un solo sensore per intervallo, il

1.1 INTRODUZIONE

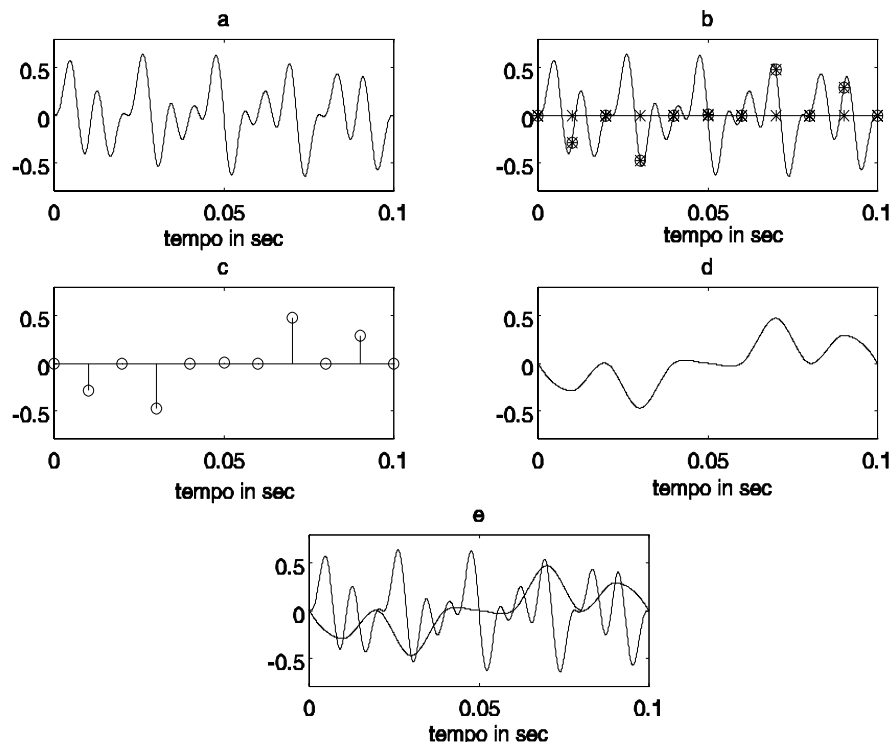


Figura 1.1: Errori di ricostruzione di un segnale dovuti ad un campionamento insufficiente

sistema ricostruisce da quella misura la temperatura di tutto l'intervallo, equivoca e lancia l'allarme.

Nel caso in cui si operi su segnali funzione del tempo, il filtraggio anti equivocazione (antialias) è facile. Infatti è implicita la disponibilità di tutti i campioni del segnale, per qualsivoglia istante di misura. Non è affatto così se si considerano variabili funzione dello spazio: in questo caso ogni campione del segnale comporta il costo di una postazione di misura. Il filtraggio anti equivocazione deve essere fatto usando strumenti opportuni che fanno a basso costo la media del segnale nell'intervallo; non deve essere fatto campionando fittamente (il che sarebbe costosissimo) facendo seguire eventualmente l'operazione di media e il successivo campionamento a frequenza inferiore (sottocampionamento).

Passando ad un'analisi quantitativa, se si campiona un segnale ad

CAPITOLO 1 IL CAMPIONAMENTO

intervallo T a partire da un istante qualsiasi T_0 , sarà necessario, per evitarne l'equivocazione, rimuovere tutte le componenti a frequenza superiore alla frequenza di Nyquist f_N , metà della frequenza di campionamento f_c ,

$$f_N = \frac{f_c}{2} \quad (1.2)$$

con un filtro opportuno. Infatti il segnale campionato é equivalente ad una sequenza di impulsi ideali con ampiezze proporzionali a quelle dei campioni $\{x_n\}$, equispaziati nel tempo (fig.1.2):

$$x_c(t) = x(t) \sum_{n=-\infty}^{+\infty} \delta(t - nT - T_0) = \sum_{n=-\infty}^{+\infty} x(nT) \delta(t - nT - T_0) \quad (1.3)$$

Moltiplicare nel dominio del tempo il segnale originario per il pettine di impulsi equivale a fare la convoluzione delle trasformate nel dominio delle frequenze; la trasformata del pettine di impulsi spazati di T é un pettine di impulsi nelle frequenze, spazati di $f_c = 1/T$. In formule si ha, indicando con la freccia il passaggio nel dominio coniugato delle frequenze:

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT - T_0) \leftrightarrow \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} \delta(\omega - 2\pi m f_c); \quad (1.4)$$

La trasformata di Fourier di $x_c(t)$ é quindi:

$$X_c(\omega) = \int_{-\infty}^{+\infty} x_c(t) e^{-j\omega t} dt = X(\omega) * \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} \delta(\omega - 2\pi m f_c) = \quad (1.5)$$

$$= \frac{1}{T} \sum_{m=-\infty}^{+\infty} e^{-j2\pi m f_c T_0} X(\omega - 2\pi m f_c) \quad (1.6)$$

Con $X(\omega)$, $X_c(\omega)$ sono rispettivamente indicate le trasformate di Fourier di $x(t)$, $x_c(t)$ e con $*$ il prodotto di convoluzione. Le repliche si sommano con sfasamenti proporzionali a T_0 .

Rimuovendo da $x(t)$ tutte le componenti a frequenza superiore a quella di Nyquist, si impedisce la sovrapposizione, nel dominio delle frequenze dello spettro originale con le sue repliche centrate a $m f_c$; $m \neq 0$ e quindi si permette la ricostruzione.

1.1 INTRODUZIONE

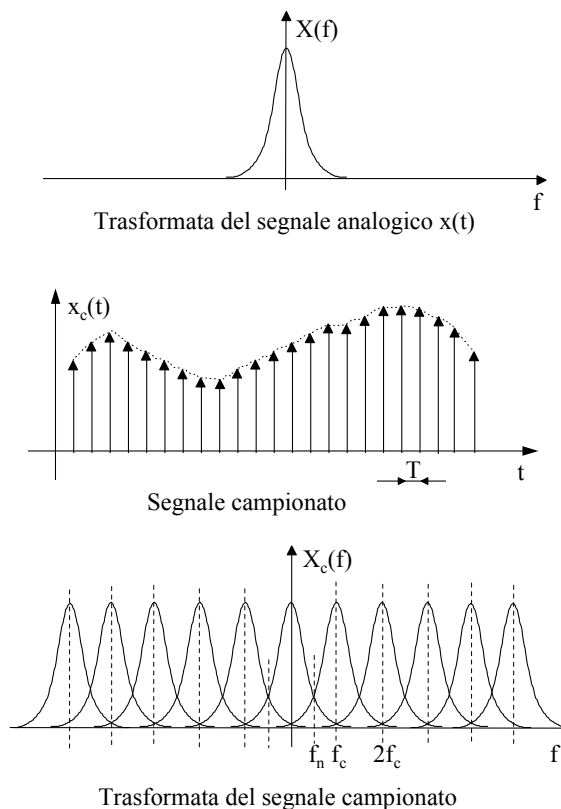


Figura 1.2: Relazione tra trasformata del segnale continuo e trasformata del segnale campionato

Un altro tipico effetto dell'equivocazione è il fenomeno per cui una sinusoide di frequenza f qualsiasi, se campionata ad intervallo T qualsiasi, può essere ricostruita (equivocando) come una sinusoide della stessa ampiezza ma a frequenza f_{eq} , inferiore alla frequenza di Nyquist.

$$f_{eq} = f - \frac{m}{T}; \quad 0 \leq f_{eq} \leq f_N \quad (1.7)$$

Per rendersi conto del fenomeno, basta pensare di illuminare con una luce regolarmente intermittente un disco metà bianco e metà nero che ruoti a velocità angolare costante. Se due lampi di luce successivi mostrano il disco nella stessa posizione, l'angolo di rotazione può essere stato nullo o un qualsiasi numero intero di giri m ; quindi si ha incertezza sulla

CAPITOLO 1 IL CAMPIONAMENTO

velocità angolare effettiva del disco, a meno che non si abbiano delle informazioni in più, ad esempio sull'intervallo in cui si prevede sia la velocità di rotazione; questo punto sarà discusso tra poco nel paragrafo sul campionamento passa banda.

1.2 Ricostruzione di un segnale

L'operazione di ricostruzione di un segnale continuo dalla sequenza dei suoi campioni (interpolazione) verrà analizzata nel prossimo capitolo; per ora è sufficiente osservare che dallo spettro $X_c(\omega)$ sembra semplice ottenere lo spettro del segnale ricostruito $\hat{X}(\omega)$ rimuovendo con un filtro ideale passa basso tutte le repliche di $X_c(\omega)$ centrate alle frequenze $m f_c$; $m \neq 0$ (fig. 1.3). Che la cosa non sia poi così facile lo si vedrà in

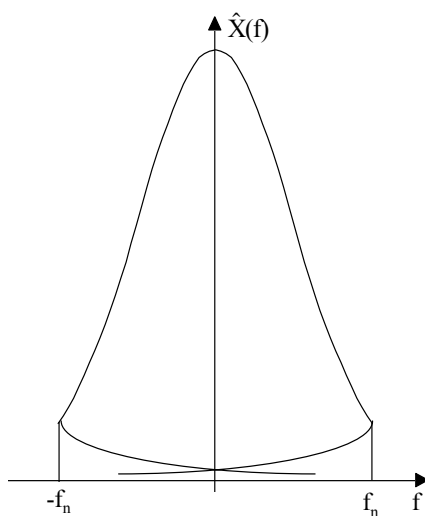


Figura 1.3: Spettro del segnale ricostruito, ottenuto rimuovendo le repliche dello spettro $X(f)$ con un filtro ideale

seguito; qui basti osservare che il filtro ideale ha una risposta all'impulso di durata illimitata; se la sequenza a disposizione è di lunghezza limitata, come sempre accade, dovremo risolvere il problema delle "condizioni ai limiti" e cioè cosa ipotizzare al di fuori dell'intervallo in cui la funzione è nota, per effettuare una ricostruzione "corretta".

1.3 Campionamento passa banda

Nel paragrafo precedente non é stata fatta alcuna ipotesi sullo spettro di $x(t)$ tranne che esso occupi frequenze inferiori a f_N , condizione necessaria per ottenere una corretta ricostruzione; inoltre, si é supposto $x(t)$ reale. Generalizzando invece il caso precedentemente introdotto (il disco rotante) del campionamento di una sinusoide a frequenza maggiore della

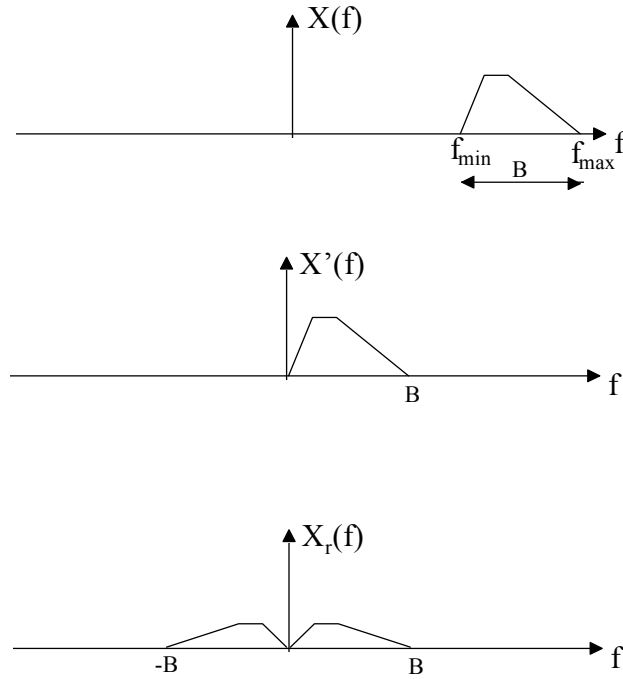


Figura 1.4: Spettri relativi al campionamento passa-banda

frequenza di Nyquist che ruota di un angolo $\varphi > \pi$ tra un campione ed il successivo, supponiamo che $x(t)$ sia un segnale complesso che occupi una banda B compresa tra le frequenze $B \ll f_{\min} \div f_{\max} = f_{\min} + B$; (lo spettro non occupi le frequenze negative).

$$\begin{aligned} X(f) &\neq 0; & f_{\min} &\leq f \leq f_{\max}; \\ f_{\max} - f_{\min} &= B; & f_0 &= \frac{f_{\max} + f_{\min}}{2} \end{aligned} \quad (1.8)$$

CAPITOLO 1 IL CAMPIONAMENTO

E' ovvio che usare come frequenza di campionamento $2f_{max}$ sarebbe uno spreco; possiamo operare in vario modo. Se conosciamo f_{min} e moltiplichiamo il segnale $x(t)$ per la sinusoida complessa $e^{-j2\pi f_{min}t}$ otteniamo un segnale complesso $x'(t)$ il cui spettro é contenuto tra 0 e B e non ha componenti a frequenze negative, come già osservato.

$$X'(f) \neq 0; \quad 0 \leq f \leq B \quad (1.9)$$

Lo spettro della parte reale di questo segnale occupa le frequenze da $-B \div B$; lo spettro della parte immaginaria occupa lo stesso intervallo di frequenza; sommato a quello della parte reale, risultano eliminate le componenti spettrali a frequenze negative.

Campionando ad intervallo $1/2B$ la parte reale e cioè usando $2B$ campioni per secondo, otterremo una sequenza da cui potremmo fare la ricostruzione del segnale originale (fig. 1.4).

Similmente, potremmo centrare (fig. 1.5) lo spettro del segnale convertito sull'origine e cioè moltiplicare il segnale $x(t)$ per la sinusoida complessa $e^{-j2\pi f_0 t}$ e ottenere un segnale complesso $x''(t)$ di banda $B/2$

$$X''(f) \neq 0; \quad -\frac{B}{2} \leq f \leq \frac{B}{2} \quad (1.10)$$

Questo segnale è da campionare con B campioni complessi al secondo; in un modo o nell'altro, si vede che un segnale di banda B si traduce in $2B$ numeri reali al secondo. E' tuttavia necessario, per potere fare correttamente il campionamento passabanda, conoscere a priori l'intervallo $f_{min} \div f_{max}$.

1.4 Il filtro antialias

In questa sezione, discuteremo brevemente del dimensionamento di un filtro antiequivocazione. Supponiamo di avere un segnale sinusoidale a pulsazione ω , da campionare ad intervallo T .

$$x(t) = e^{j\omega t}; \quad x_n = e^{j\omega nT}$$

Ovviamente, il filtro antiequivocazione ideale è:

$$h(t) = \frac{\sin(\pi t/T)}{\pi t/T}$$

1.4 IL FILTRO ANTIALIAS

$$H(\omega) = T; \quad |\omega| < \pi/T$$

di lunghezza illimitata ma di difficile adozione nella pratica; infatti, per determinare un campione del segnale, si dovrebbe conoscerne l'andamento per tutti i valori del tempo. Il più semplice filtro antialias è il mantenitore, che ha bisogno dei dati solo nell'intervallo di durata T , :

$$h_0(t) = 1; \quad |t| < T/2$$

$$H_0(\omega) = T \frac{\sin(\pi f T)}{\pi f T}$$

questo filtro tuttavia attenua bene solo le componenti del segnale a frequenza $1/T$ e multiple e non, come sarebbe necessario, tutte le componenti a frequenze al di sopra della frequenza di Nyquist, $1/2T$. Un filtro più accettabile è un filtro gaussiano

$$h_g(t) = e^{-t^2/2\sigma^2}$$

$$H_g(\omega) = \sqrt{2\pi}\sigma \times e^{-\omega^2\sigma^2/2}$$

di lunghezza limitata e di banda limitata. Per avere un dimensionamento, si può pensare di imporre un'attenuazione di 40dB alla frequenza di campionamento:

$$e^{-\omega^2\sigma^2/2} = e^{-4,6}; \quad \omega = \frac{2\pi}{T} \longrightarrow \sigma = .48T$$

Nei tempi si vede che:

$$h_g(T) = .11$$

e cioè che il filtro si estende fino ai campioni adiacenti. Volendo invece un'attenuazione di 40 dB alla frequenza di Nyquist, la banda si dimezza, la durata raddoppia e si ha:

$$h_g(T) = .58; \quad h_g(2T) = .11;$$

In altre parole, un filtro antialias che sopprima abbastanza bene le componenti alla frequenza di Nyquist, ha approssimativamente una forma triangolare e una durata di circa 4 campioni. Peraltro, un filtro triangolare di base 4 campioni ha funzione di trasferimento:

$$H(\omega) = \left[\frac{\sin(4\pi f T)}{4\pi f T} \right]^2 16T^2$$

CAPITOLO 1 IL CAMPIONAMENTO

e alla frequenza di Nyquist vale:

$$H\left(\frac{\pi}{T}\right) = .025H(0)$$

e cioè circa -32dB rispetto al valore a frequenza 0.

Capitolo 2

La trasformata di Fourier discreta

2.1 Campionare in tempo e in frequenza

Consideriamo una sequenza di N impulsi di ampiezze x_n , $0 \leq n \leq N-1$, e la loro trasformata di Fourier:

$$X(f) = \sum_{n=0}^{N-1} x_n e^{-j2\pi f n T} \quad (2.1)$$

Poichè la $X(f)$ è trasformata di una sequenza di impulsi e quindi di un segnale campionato, è funzione periodica della frequenza, con periodo $1/T$. Se campioniamo anche $X(f)$ **nelle frequenze** ad intervallo Δf la sequenza di impulsi si replicherà nei tempi ad intervallo $1/\Delta f$. Scegliendo $\Delta f = 1/NT$ le repliche della sequenza di impulsi non si sovrappongono, ed il campionamento anche nel dominio della frequenza non crea equivocazione. Tuttavia, l'infinita ampiezza degli impulsi, sia nel dominio delle frequenze che nel dominio dei tempi, correlativa dell'illimitata periodicità delle funzioni nell'altro dominio e pertanto di un'energia illimitata, complica inutilmente il problema. E' peraltro intuitivo che debba esistere una relazione tra le N "ampiezze" dei campioni nel dominio dei tempi che quelle dei campioni nel dominio trasformato delle frequenze.

Definiremo ora una trasformazione lineare, **la trasformazione discreta di Fourier** (in inglese: Discrete Fourier Transform DFT); essa lega una sequenza di N **campioni** complessi (non più impulsi ma numeri)

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

ad un'altra sequenza di N numeri complessi. Questa seconda sequenza corrisponde peraltro alle ampiezze degli impulsi di cui è costituita la trasformata di Fourier della sequenza di impulsi originaria; gli uni (nel dominio dei tempi) sono distanti tra loro dell'intervallo T e gli altri (nel dominio delle frequenze) dell'intervallo $1/NT$. Dimosteremo che la trasformazione così definita è invertibile; peraltro, la cosa è intuibile, visto che le due operazioni di campionamento non provocano equivocazione. Quindi, la trasformata di Fourier discreta di una sequenza di N numeri complessi è costituita da altrettanti numeri complessi che danno una rappresentazione alternativa della sequenza data nel dominio delle frequenze, come somma di altrettante sinusoidi.

La coppia di relazioni che legano una sequenza alla sua trasformata di Fourier discreta possono essere derivate sia direttamente, sia dalla trasformata di Fourier consueta sostituendo alle variabili continue t, f le variabili discretizzate $nT, k\Delta f$. La trasformata di Fourier consueta è:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (2.2)$$

L'integrale diventa una sommatoria, se la $x(t)$ è una somma di impulsi ai tempi nT : il fattore T servirebbe poi anche a ricordare la dimensionalità dell' integrale di Fourier.

$$X(f) = T \sum_{n=-\infty}^{+\infty} x_n e^{-j2\pi fnT} \quad (2.3)$$

Tuttavia, nel testo useremo sempre la definizione di formula (2.1). Inoltre in questo caso la $X(f)$ è periodica di periodo $1/T$; infatti:

$$X(f + m/T) = X(f) \quad (2.4)$$

Se poi la funzione $X(f)$ è campionata a intervallo $\Delta f = 1/NT$ la sequenza di impulsi di ampiezze $\{x_n\}$ è replicata con periodo NT . Definiamo allora come Trasformata di Fourier Discreta di una sequenza di N numeri complessi, la sequenza di N numeri complessi:

$$X(k\Delta f) = X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi k\Delta fnT} = \sum_{n=0}^{N-1} x_n W_N^{-nk}; \quad (2.5)$$

2.2 PROPRIETÀ DELLA DFT

$$W_N = \exp\left(j\frac{2\pi}{N}\right)$$

Si verifica che la Trasformazione discreta di Fourier è invertibile e la antitrasformata è:

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{nk} \quad (2.6)$$

Per convincersene, basta applicare il principio della sovrapposizione degli effetti e quindi limitarsi a supporre che la sequenza da trasformare $\{\tilde{x}_n\}$ sia costituita da un solo impulso ad un istante qualsiasi:

$$\tilde{x}_n = 1; n = n_0; \tilde{x}_n = 0; n \neq n_0$$

Allora:

$$\begin{aligned} \tilde{X}_k &= \sum_{n=0}^{N-1} \tilde{x}_n W_N^{-nk} = W_N^{-n_0 k} \\ \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}_k W_N^{nk} &= \frac{1}{N} \sum_{k=0}^{N-1} W_N^{-n_0 k} W_N^{nk} = \frac{1}{N} \sum_{k=0}^{N-1} W_N^{(n-n_0)k} = \tilde{x}_n \end{aligned}$$

Infatti, il versore W_N corrisponde ad una rotazione di un angolo pari a $2\pi/N$; se $n \neq n_0$ questo versore fa un numero intero di giri e la somma genera un poligono chiuso e quindi a somma zero. Se invece $n = n_0$ il versore non si sposta dal valore unitario ed il risultato è ovvio.

2.2 Proprietà della DFT

La discretizzazione dell'integrale di Fourier comporta un fattore di dimensione $[T]$ che però si omette nella definizione consueta della trasformata di Fourier discreta (DFT). Nel seguito cercheremo però di ricordare sempre l'origine fisica della trasformazione e quindi ci riferiremo alle dimensionalità originali degli assi di frequenze e tempi, per permettere un più efficace uso dell'intuizione. Nel dominio trasformato (che continuiamo pertanto ad indicare come dominio delle frequenze), vista la periodicità della trasformata, si osserva che i valori dell'indice $0 \leq k \leq N/2 - 1$ corrispondono alle frequenze positive; i valori $N/2 \leq k \leq N - 1$ corrispondono alle frequenze negative; in particolare, $k = N - 1$ corrisponde alla frequenza

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

$f = -1/NT$; la frequenza di Nyquist corrisponde al valore dell'indice $N/2$ se N è pari.

Similmente a quanto accade per la trasformata di Fourier continua si hanno diverse relazioni utili, tra cui la relazione di Parseval:

$$\frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 = \sum_{n=0}^{N-1} |x_n|^2 \quad (2.7)$$

il ritardo (circularizzato: $x(n+N) = x_n$):

$$x_{n+m} \leftrightarrow X(k)W_N^{-mk} \quad (2.8)$$

la convoluzione circolare:

$$\frac{1}{N} \sum_{n=0}^{N-1} x_m y_{n-m} \leftrightarrow X(k)Y(k) \quad (2.9)$$

il coniugio e l'inversione dell'asse temporale:

$$x_{-n}^* \leftrightarrow X^*(k) \quad (2.10)$$

2.3 Filtraggi con la DFT

Abbiamo visto che la trasformazione di Fourier discreta corrisponde alla decomposizione di una sequenza di N campioni complessi nella somma di N sinusoidi complesse di frequenze k/NT ; $0 \leq k \leq N-1$; è possibile convolvere due sequenze **entrambe di N campioni** moltiplicando le loro trasformate di Fourier.

Bisogna però ricordare (o verificare, se non si vuole sempre risalire alle origine fisiche della trasformazione) che così si realizza una **convoluzione periodica** delle due sequenze; infatti, essendo gli spettri campionati nelle frequenze, le sequenze di impulsi sono replicate nel tempo. Inoltre, il prodotto degli spettri campionati ha significato solo se i campioni nel dominio delle frequenze sono rilevati alle stesse frequenze, e pertanto le due sequenze (nei tempi e nelle frequenze) devono avere eguale numero di campioni, perchè l'operazione di prodotto abbia senso fisico.

Poi, il prodotto dei due spettri corrisponde alla convoluzione delle due sequenze originali, periodicizzate e cioè replicate dopo N campioni;

2.3 FILTRAGGI CON LA DFT

la sequenza risultante sarà ancora una sequenza periodica su N campioni, diversa dalla sequenza di $2N - 1$ campioni corrispondente alla convoluzione **aperiodica** delle due sequenze. Tutto passa quindi come se l'asse temporale fosse circolare e il risultato della convoluzione periodica è inquinato dal mescolamento dei primi e degli ultimi campioni della sequenza, come se fossero vicini tra di loro.

Anche se in alcuni casi (sistemi in coordinate polari) l'asse circolare può essere utile, in generale interessano solo convoluzioni aperiodiche; anch'esse sono possibili con la trasformata di Fourier discreta a patto di usare accortezze opportune o di accettare delle approssimazioni.

2.3.1 Convoluzioni aperiodiche: sequenze limitate

Si voglia effettuare la convoluzione di una sequenza $\{x_n\}$ lunga kN campioni, (si pensi a kN grande, dell'ordine di migliaia di campioni) con un filtro $\{h_m\}$ di lunghezza M campioni, con M dell'ordine delle centinaia. Come prima cosa, abbiamo osservato che **prima della trasformata di Fourier** si devono portare le due sequenze ad avere eguale lunghezza, perché le due trasformate siano campionate agli stessi valori della frequenza. Per fare ciò aggiungiamo $M - 1$ campioni nulli (che indicheremo con zeri, sperando che non vi sia confusione con gli zeri della trasformata z che non hanno nulla a che vedere) alla sequenza lunga N ed $N - 1$ zeri alla sequenza lunga M . Dopo questa operazione di allungamento con zeri (zero padding, in inglese) le due sequenze hanno eguale lunghezza $N + M - 1$ e si procede alle loro trasformate di Fourier (fig. 2.1).

Poi si moltiplicano le trasformate (che noi indicheremo spesso come gli spettri) e si antitrasforma; il risultato è la convoluzione periodica delle due sequenze allungate con zeri e quindi è esattamente la convoluzione aperiodica delle due sequenze originali. Infatti aggiungendo zeri abbiamo creato lo spazio, per fare sì che la convoluzione non si riavvolgesse sulla parte iniziale dei dati.

2.3.2 Sequenze illimitate: overlap-save; overlap-add

Qualora, come spesso accade, la sequenza originaria è di lunghezza illimitata, per procedere alla convoluzione con un filtro di lunghezza M

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

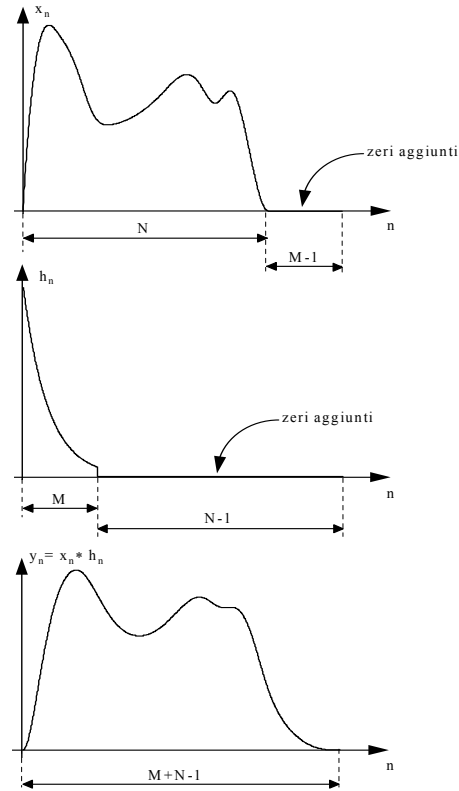


Figura 2.1: Convoluzione aperiodica con la DFT

dobbiamo scegliere tra due metodi simili ma alternativi per effettuare la convoluzione aperiodica dell'intero segnale. I metodi sono chiamati **Overlap and add** e **Overlap and save**. In sostanza in entrambi i casi si deve:

- segmentare il segnale di lunghezza illimitata in sottosequenze lunghe N campioni,
- tenere conto degli effetti della convoluzione periodica, nel moltiplicare le trasformate di segnale e filtro.

Supponendo che il filtro occupi i primi M campioni della sequenza, si può:

Overlap and add: azzerare gli ultimi $M - 1$ campioni della sequenza lunga N , moltiplicare gli spettri ed antitrasformare. L'antitrasforma-

2.3 FILTRAGGI CON LA DFT

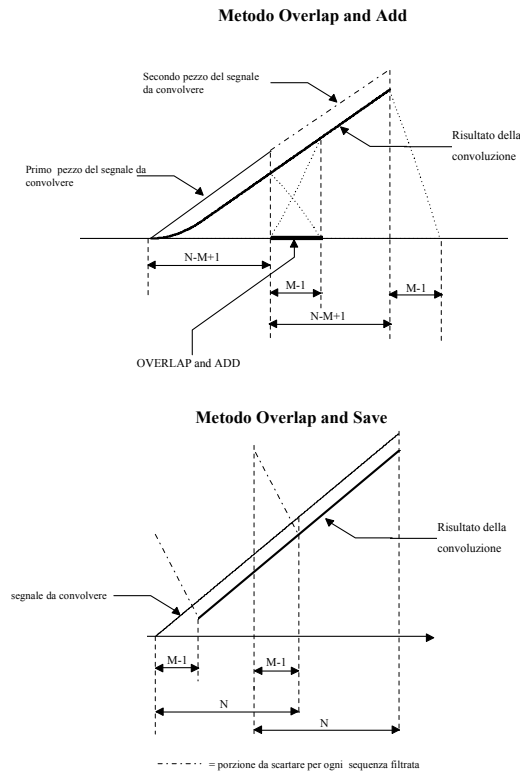


Figura 2.2: Filtraggio di sequenze illimitate con la DFT

ta contiene la convoluzione corretta del segnale parzializzato; nei primi $M - 1$ campioni però manca la coda dei campioni della sottosequenza precedente dove il segnale di ingresso era stato azzerato. Per arrivare al risultato corretto, si deve quindi riprendere il segnale originale dal campione di posizione $N - M + 1$, prenderne altri $N - M + 1$ campioni, allungarli con $M - 1$ zeri e così via. Il risultato finale si ottiene dai risultati parziali sommando le sequenze sovrapposte (overlap and add).

Overlap and save: in questo caso non si azzerano gli ultimi $M - 1$ campioni della sequenza lunga N campioni e ne risulta la creazione di $M - 1$ campioni errati **all'inizio** della sequenza convoluta che risulta dalla antitrasformazione di Fourier del prodotto degli spettri. Questi campioni del risultato finale, errati perchè inquinati dagli ultimi campioni

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

per effetto della circolarità della convoluzione, vengono annullati. Ogni sottosequenza deve quindi iniziare con $M - 1$ campioni di anticipo come nel caso overlap - add discusso prima. Le sottosequenze successive, che contribuiscono ognuna $N - M + 1$ campioni validi (gli ultimi), vanno giustapposte e non sommate.

2.4 Antitrasformazione di Fourier discreta

Nella sezione precedente si è discusso come convolvere tra loro sequenze di N campioni, create nel dominio dei tempi. Spesso accade però di avere dei segnali da antitrasformare, definiti nel dominio delle frequenze; nel decidere la lunghezza N bisogna stare attenti che sia rispettata, almeno approssimativamente, la condizione di non sovrapposizione delle repliche nei tempi (assenza di equivocazione temporale). Se si ha a che fare con risposte all'impulso illimitate (ad esempio come capita con filtri ideali), questo non è possibile e pertanto è necessario convivere con l'equivocazione ed accettare approssimazioni. Per fare un esempio semplicissimo, si voglia ritardare di τ la sequenza (fig. 2.3); in altre parole, si vuole la sequenza corrispondente al segnale continuo originario ricampionato con un pettine ritardato di τ rispetto al pettine di campionamento originale: per fare ciò basta moltiplicare la trasformata della sequenza per $e^{-j\omega\tau}$. Anzitutto, i campioni della trasformata vanno collocati nella posizione corretta (N sia un numero pari):

$$R(2\pi k\Delta f) = R_k = \exp(-j\frac{2\pi k}{NT}\tau); \quad (2.11)$$

$$R_{N-k} = R_k^*; \quad 0 \leq k < N/2; \quad R_{N/2} = R_{N-N/2}^* = \cos \frac{\pi\tau}{NT}$$

Inoltre, non si deve dimenticare l'effetto del campionamento dell'asse delle frequenze; l'antitrasformata della sequenza nel dominio delle frequenze è la sequenza nel dominio dei tempi che si ottiene come segue;

a) si considera l'impulso ideale, lo si ritarda di τ e lo si filtra passa basso con un filtro ideale di banda $B = 1/2T$; questo filtraggio è realizzato implicitamente dal limitare i valori della trasformata alla pulsazione di Nyquist $|\omega| \leq \pi/T$ (fig. 2.4).

b) si campiona questo segnale con intervallo T (la periodicità nel dominio delle frequenze è implicita nella trasformata di Fourier discreta).

2.4 ANTITRASFORMAZIONE DI FOURIER DISCRETA

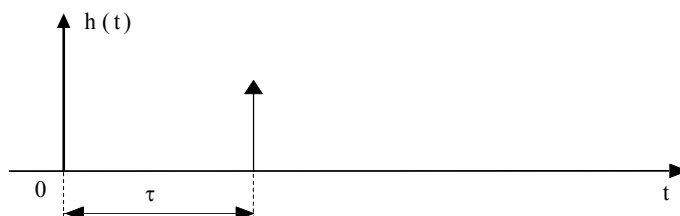


Figura 2.3: Risposta all'impulso ideale di un ritardatore

c) si periodicizza ad intervallo NT (si ritarda di mNT la sequenza di lunghezza infinita e si sommano tutte le repliche). Quest'effetto è dovuto al campionamento nelle frequenze.

Si osservi ora che la risposta nei tempi del ritardatore, non periodicizzata, ha durata infinita, e converge a zero con l'andamento del sinc e cioè come $1/n$; gli errori potrebbero non essere trascurabili.

Cosa si deve fare per evitare l'errore di equivocazione, temporale questa volta? Si deve smussare la trasformata nel dominio delle frequenze, prima di procedere al campionamento.

Ad esempio, nel caso del ritardatore indicato, prima di campionare, si può convolvere la trasformata, **prima del campionamento** con un triangolo di durata significativa, in modo che la risposta all'impulso, prima della replica, decada come $1/n^3$. La convoluzione va fatta in via analitica oppure numerica, partendo da un campionamento della funzione $\exp(-j2\pi f\tau)$ più fitto che $\Delta f = 1/NT$. Ciò corrisponde a considerare la sequenza corrispondente al ritardatore su un intervallo molto lungo, moltiplicarla per una funzione decrescente con n , e poi replicarla, per effetto del campionamento.

Sarebbe utile poter fare gli smussamenti nel dominio delle frequenze convolvendo con dei sinc, per fare veri e propri troncamenti nel dominio dei tempi, ma ovviamente la cosa è molto difficile numericamente; quindi il tutto equivale a trasferire il problema del ripiegamento dal dominio dei tempi dove avviene, a quello delle frequenze dove dovrebbe essere evitato con un filtraggio ideale, altrettanto difficile da effettuare, se non di più.

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

Vedremo ora questi effetti con due esempi semplici.

2.4.1 Il derivatore e la costante di tempo

Nella sezione precedente abbiamo visto come sia semplice realizzare un ritardatore di τ ma anche che è inevitabile accettare delle approssimazioni. Vedremo in questa sezione come realizzare un derivatore; basterebbe moltiplicare la trasformata per $j\omega$. Così facendo, si devono ripetere le osservazioni fatte nella sezione precedente. Pertanto: derivare operando nel dominio delle frequenze con la DFT equivale (operando nel dominio dei tempi) a convolvere il segnale con la versione **campionata e replicata** del derivatore ideale filtrato passa basso $d_f(t)$. Quest'ultimo è ottenuto campionando la derivata della risposta all'impulso del filtro ideale passa basso (fig. 2.5).

$$d_f(t) = \frac{d}{dt} \frac{\sin \pi t/T}{\pi t/T} = \frac{1}{\pi t^2} \left(\pi t \cos \frac{\pi}{T} t - T \sin \frac{\pi}{T} t \right);$$
$$d_f(nT) = \frac{1}{nT} (\cos n\pi) = \frac{(-1)^n}{nT}; \quad n \neq 0; \quad d_f(0) = 0$$

Per prefiltrare prima del campionamento in frequenza dobbiamo smusare nelle frequenze l'operatore di derivazione, che peraltro è molto liscio tranne che nell'origine e alla frequenza di Nyquist. In particolare il punto angoloso nell'origine è fortemente influenzato dagli effetti di troncamento o di ripiegamento nei tempi.

Consideriamo ad esempio l'effetto del troncamento sulla componente continua e cioè il valore medio dell'operatore nei tempi; essendo i campioni successivi dell'operatore a segni alterni nei tempi, qualsiasi troncamento comporta un valor medio non nullo e quindi la perdita dello zero alla frequenza zero. Per riportare a zero la funzione di trasferimento in continua si deve fare qualcosa che comunque deforma l'operatore di derivazione; il ripiegamento invece non perturba il valor medio dell'operatore, ma ha altri effetti dannosi. In conclusione, l'unica cosa fattibile semplicemente è scegliere N sufficientemente grande così che l'inevitabile effetto di ripiegamento nei tempi comporti conseguenze accettabili.

Come ultimo esempio consideriamo l'antitrasformata di Fourier dell'esponenziale, che ha funzione di trasferimento:

$$A(\omega) = \frac{1}{1 + j\omega\tau} \quad (2.12)$$

2.4 ANTITRASFORMAZIONE DI FOURIER DISCRETA

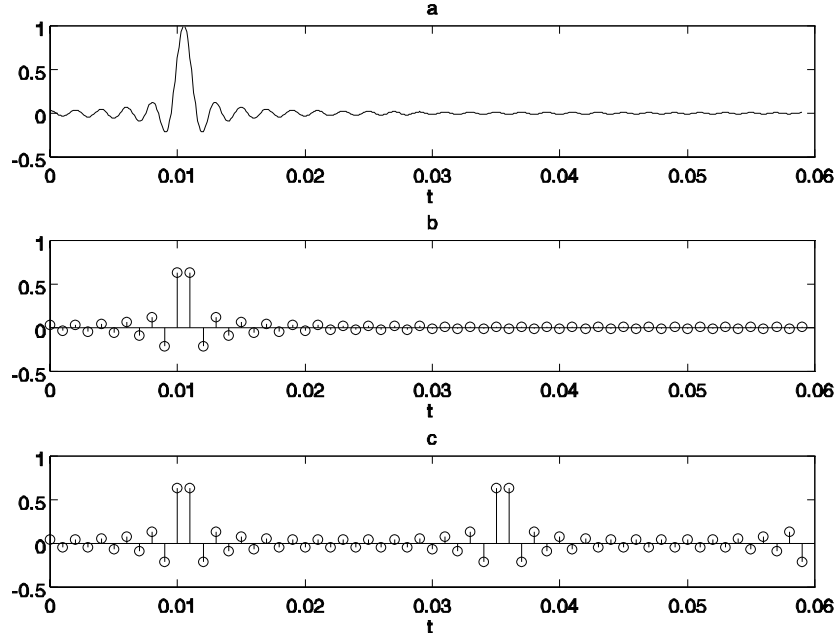


Figura 2.4: a) Risposta impulsiva di un ritardatore filtrata passabasso b) Campionamento nei tempi c) Campionamento nelle frequenze (Periodicizzazione nei tempi)

campionando nelle frequenze si replica l'esponenziale nei tempi; quindi la durata totale NT deve essere ad esempio almeno $NT > 3\tau$ (fig. 2.6). Ancora dobbiamo fissare il valore di T e questo verrà fatto in base alla necessità di evitare equivocazione nelle frequenze e cioè per campionare correttamente nei tempi: alla frequenza di Nyquist, dove sarà sovrapposta la replica proveniente dal campionamento, è necessario che lo spettro sia già molto ridotto; p. e. ad un $1/100$ del valore assunto a frequenza 0; si ricava allora

$$\frac{\pi\tau}{T} \simeq 100 \rightarrow T = .0314\tau; \quad \frac{N\pi}{100} > 3 \rightarrow N > 95 \quad (2.13)$$

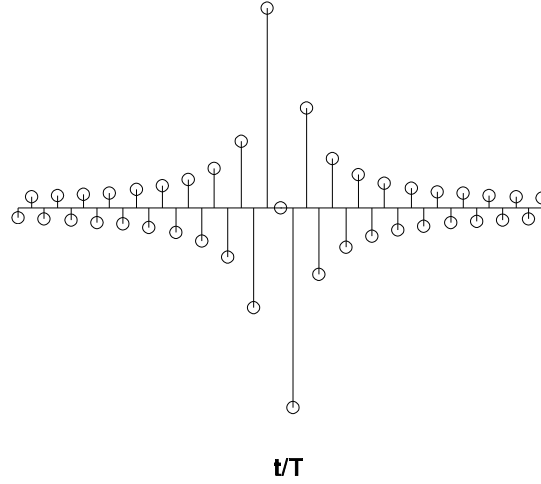


Figura 2.5: Derivatore ideale a banda limitata, non causale, campionato

2.5 Interpolazione con la DFT

2.5.1 Inserzione di zeri e repliche spettrali

L'interpolazione di sequenze per ottenere un segnale ad intervallo di campionamento più piccolo può essere fatta in un modo molto semplice mediante la DFT. Sia assegnata una sequenza $\{x_n; n = 0, \dots, N-1\}$ campionata ad intervallo T ; per fare un esempio di duplicazione della frequenza di campionamento e quindi di interpolazione su un intervallo di durata metà, consideriamo la sequenza di lunghezza doppia ma di eguale durata $\{x'_m; m = 0, \dots, 2N-1\}$ che si ottiene dalla sequenza precedente inserendo campioni di ampiezza nulla ad intervallo $T/2$. Si ha allora:

$$x'_{2n} = x_n; \quad x'_{2n+1} = 0;$$

La DFT della nuova sequenza $\{x'_m\}$ è:

$$X'_k = \sum_{m=0}^{2N-1} x'_m W_{2N}^{-mk} = \sum_{n=0}^{N-1} x_n e^{-j\frac{2\pi nk}{N}} = X_k; \quad k = 0, \dots, 2N-1 \quad (2.14)$$

Ma si ha anche:

$$X'_{k+N} = X_k; \quad k = 0, \dots, N-1$$

2.5 INTERPOLAZIONE CON LA DFT

Quindi, $\{X'_k\}$, trasformata della sequenza di $2N$ campioni $\{x'_m\}$ coincide con la sequenza X_k DFT della $\{x_n\}$, però replicata 2 volte nel dominio delle frequenze. In generale, se in una sequenza vengono inseriti $M - 1$ campioni nulli tra un campione ed il successivo (e quindi la lunghezza totale della sequenza viene moltiplicata per M), la DFT risulta replicata M volte.

Soppressione delle repliche spettrali con filtraggio ideale

Sopprimiamo ora tutte le repliche spettrali moltiplicando la sequenza X'_k per degli zeri disposti in modo opportuno ed anche per un opportuno fattore di scala M (supponiamo N pari: per conservare la simmetria della componente alla frequenza di Nyquist originale $\pm 1/2T$ si adotta il fattore $M/2$):

$$X''_k = MX'_k; \quad k = 0, \dots, N/2 - 1; \quad k = MN - N/2 + 1, \dots, MN - 1$$

$$X''_k = MX'_k/2; \quad k = N/2; \quad k = MN - N/2$$

$$X''_k = 0; \quad k \neq 0, \dots, N/2; \quad k \neq MN - N/2, \dots, MN - 1$$

allora la sequenza antitrasformata $\{x''_m\}$ è una sequenza della stessa durata ma di MN campioni, e sarà l'interpolata della sequenza originale su un intervallo di campionamento pari a T/M . L'interpolazione avviene mediante la convoluzione della sequenza inframmezzata con zeri per il filtro ideale campionato e periodicizzato la cui DFT è appunto:

$$F_k = M; \quad k = 0, \dots, N/2 - 1; \quad k = MN - N/2 + 1, \dots, MN - 1$$

$$F_k = M/2; \quad k = N/2; \quad k = MN - N/2$$

$$F_k = 0; \quad k \neq 0, \dots, N/2; \quad k \neq MN - N/2, \dots, MN - 1$$

I campioni della sequenza che costituisce la risposta all'impulso del filtro interpolatore, antitrasformata della $\{F_k\}$, sono facilmente calcolabili dalla solita espressione della serie geometrica troncata (N pari) con l'aggiunta dell'effetto del termine di ampiezza $M/2$:

$$f_n = \frac{M}{MN} \frac{\sin(\pi n N / MN)}{\tan(\pi n / MN)} = \frac{1}{N} \frac{\sin(\pi n / M)}{\tan(\pi n / MN)} \quad (2.15)$$

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

Il fattore di scala M nasce dal fatto che si desidera un'interpolazione e pertanto si vuole che:

$$f_0 = 1$$

Quindi, la componente continua della DFT non si conserva con l'interpolazione ma viene accresciuta di un fattore M . Si osservi che se la sequenza di partenza avesse avuto una durata illimitata, il filtro da utilizzare come interpolatore sarebbe stato il filtro ideale passa basso con frequenza di taglio $1/2T$ campionato con intervallo di campionamento $1/MT$ e cioè:

$$f'_n = \frac{\sin(\pi n/M)}{\pi n/M}$$

L'effetto della periodicizzazione è modesto se MN è grande; peraltro, se $M = N = 2$ si ha $\{f_i\} = \{1, .5, 0, .5\}$ e si trova un interpolatore a triangolo.

Nel caso in cui N è dispari si ha la consueta espressione del *sinc* discreto, antitrasformata di uno spettro "rettangolare" centrato intorno a frequenza 0:

$$F_k = M; \quad k = 0, \dots, (N-1)/2; \quad k = MN - (N-1)/2, \dots, MN-1$$

$$F_k = 0; \quad k \neq 0, \dots, (N-1)/2; \quad k \neq MN - (N-1)/2, \dots, MN-1$$

e risulta infine (N dispari):

$$f_n = \frac{1}{N} \frac{\sin(\pi n/M)}{\sin(\pi n/MN)}$$

L'interpolazione con aggiunta di zeri (zero padding)

In conclusione, la tecnica per interpolare con la DFT una sequenza di lunghezza N (pari) per farne una sequenza di lunghezza MN è la seguente:

- si calcola la DFT della sequenza di lunghezza N ;
- si allunga lo spettro aggiungendo $(M-1)N-1$ campioni nulli, inseriti tra i campioni di posto $N/2$ e $N/2+1$;
- si moltiplica per M e, per mantenere le simmetrie, si suddivide il valore del campione di posto $N/2$ tra i campioni di posto $N/2$ e $MN - N/2$, e infine si antitrasforma.

Il risultato sarà la sequenza di partenza, interpolata con la funzione *sinc* periodicizzata descritta nella formula (2.15).

2.6 Il calcolo della DFT: la FFT

Per calcolare la sequenza $\{X_k\}$ a partire dalla sequenza $\{x_n, 0 \leq n \leq N-1\}$ dobbiamo applicare la matrice $\underline{\mathbf{W}}_N^*$ di dimensioni $[N \times N]$:

$$X_k = \sum_{n=0}^{N-1} x_n W_N^{-nk}, \quad W_N = e^{j2\pi/N} \quad (2.16)$$

$$\underline{\mathbf{X}} = \underline{\mathbf{W}}^* \underline{\mathbf{x}} \quad (2.17)$$

dove con $\underline{\mathbf{x}}$, $\underline{\mathbf{X}}$ abbiamo indicato due vettori colonna contenenti rispettivamente la sequenza e la sua trasformata discreta di Fourier.

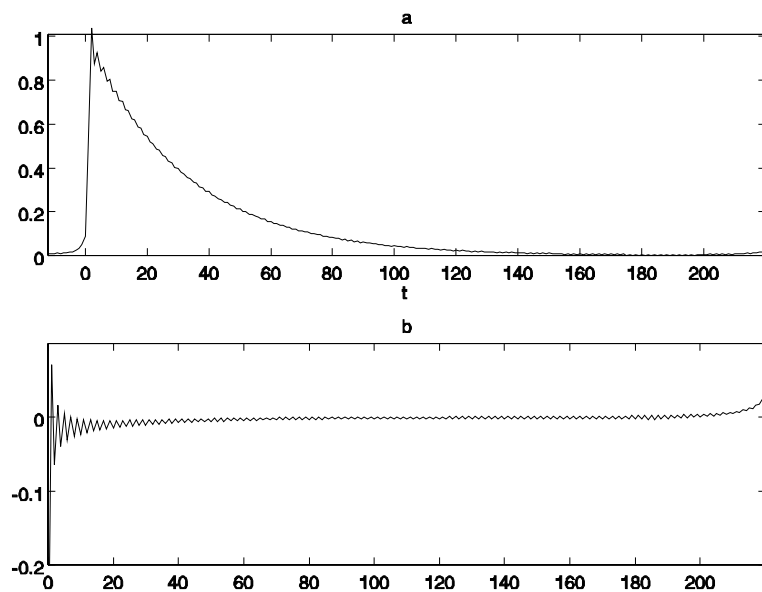


Figura 2.6: a) Antitrasformata discreta di una F.d.T con un polo reale; b) differenza tra antitrasformata discreta e antitrasformata continua

La matrice $\underline{\mathbf{W}}$ appare trasposta (o coniugata essendo Hermitiana),

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

per i segni negativi negli esponenti della trasformazione diretta.

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ \dots \\ X(N-2) \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & W_N^{-1} & W_N^{-2} & \dots & W_N^{-2} & W_N \\ 1 & W_N^{-2} & W_N^{-4} & \dots & W_N^{-4} & W_N^{-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & W_N^{-(N-2)} & W_N^{-4} & \dots & W_N^{-4} & W_N^{-2} \\ 1 & W_N & W_N^2 & \dots & W_N^{-2} & W_N^{-1} \end{bmatrix} \times \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_{N-2} \\ x_{N-1} \end{bmatrix} \quad (2.18)$$

I campioni della riga n -esima corrispondono alle successive posizioni di un vettore (una sinusoide complessa) che fa $n - 1$ giri in N campioni.

La seconda riga è una sinusoide complessa che fa un ciclo in senso orario in N campioni (frequenza $f = 1/NT$) e così via fino all'ultima riga che corrisponde ad una sinusoide complessa che fa $N - 1$ giri in N campioni; in altre parole, ruota in senso antiorario di un giro in N campioni (frequenza $f = -1/NT$). Se N è pari, la frequenza di Nyquist è alla riga $N/2 + 1$; è necessario ricordare come la frequenza 0 sia nella prima riga. La matrice $\underline{\mathbf{W}}$ non è una matrice ortogonale vera e propria perchè in quel caso dovrebbe essere:

$$\underline{\mathbf{A}}^{-1} = \underline{\mathbf{A}}^* \quad (2.19)$$

Invece si ha, combinando trasformazione ed antitrasformazione di Fourier:

$$\frac{1}{N} \underline{\mathbf{W}} \underline{\mathbf{W}}^* = \underline{\mathbf{I}}; \quad \rightarrow \quad \frac{1}{N} \underline{\mathbf{W}}^* = \underline{\mathbf{W}}^{-1} \quad (2.20)$$

Per calcolare $\underline{\mathbf{X}}$ è necessario applicare questa matrice e la cosa sarebbe numericamente assai costosa (circa N^2 moltiplicazioni) se non fosse stata scoperta la trasformata di Fourier veloce (FFT: Fast Fourier Transform) che permette di fare questo calcolo con un numero assai limitato di operazioni, crescente solo logicamente con N . Alla base della FFT c'è l'idea che una qualsiasi rotazione di k/N giri si può ottenere come combinazione di $m = \log_2 N$ rotazioni di $1/2$ giro, $1/4$ di giro ecc. fino ad $1/N$ giro, sempre che N sia una potenza di 2 e cioè $N = 2^m$.

2.6.1 Le farfalle

Il meccanismo che porta alla FFT può essere analizzato in breve: si consideri la trasformata di Fourier discreta (DFT) di una sequenza di N campioni e per semplicità si supponga N potenza di 2. Si decomponga la sommatoria in due sommatorie che si riferiscono la prima ai termini di posto pari della sequenza, e la seconda a quelli di posto dispari. Si osserva che le due sommatorie che compaiono sono ancora delle DFT, che si riferiscono questa volta a sequenze di $N/2$ campioni l'una. Indichiamo con $G(k)$, $H(k)$ queste due trasformate; risulta:

$$X(k) = \sum_{n=0}^{N/2-1} x_{2n} W_N^{-2nk} + \sum_{n=0}^{N/2-1} x_{2n+1} W_N^{-(2n+1)k} = \quad (2.21)$$

$$\sum_{n=0}^{N/2-1} x_{2n} W_{N/2}^{-nk} + W_N^{-k} \sum_{n=0}^{N/2-1} x_{2n+1} W_{N/2}^{-nk} = G(k) + W_N^{-k} \times H(k)$$

$$X(k + N/2) = G(k) - W_N^{-k} \times H(k)$$

In scrittura matriciale:

$$\begin{bmatrix} X(k) \\ X(k + N/2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} G(k) \\ W_N^{-k} H(k) \end{bmatrix} \quad (2.22)$$

dove si è fatto uso della proprietà delle trasformate di sequenze di $N/2$ campioni di essere periodiche di periodo $N/2$ e quindi:

$$H(k + N/2) = H(k); \quad G(k + N/2) = G(k)$$

inoltre:

$$W_N^{k+N/2} = -W_N^k$$

Nasce così il grafo di flusso che dà gli elementi di una trasformata di Fourier di N campioni, combinando quelli di due trasformate di Fourier di $N/2$ campioni l'una; in particolare viene indicato con il nome di *farfalla*, (*butterfly*) (fig. 2.7) il grafo che connette $X(k)$, $X(k+N/2)$ a $G(k)$, $H(k)$. La decomposizione si può ripetere fino ad avere sequenze di 2 campioni.

Allo stesso risultato si può arrivare ripetendo le considerazioni fatte in precedenza per l'interpolazione di una sequenza 2:1 eseguita inserendo

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

una sequenza di zeri in posizioni alterne e poi sopprimendo la replica spettrale. In questo caso, invece, la sequenza originale è la somma di **due sequenze** che hanno campioni nulli in posizioni alterne, e cioè:

- la sequenza dei campioni pari (con campioni nulli in posizioni dispari) che ha trasformata discreta di Fourier G_k , periodica di periodo $N/2$;
- la sequenza dei campioni dispari, ritardata di T rispetto alla sequenza dei campioni pari, con campioni nulli in corrispondenza delle posizioni pari, che ha trasformata discreta di Fourier $H_k W_N^{-k}$, periodica di periodo $N/2$. Infatti, nel dominio trasformato, il **ritardo di un campione** tra le due sequenze è ottenuto moltiplicando la H_k per il termine $\exp(-j2\pi k \Delta f T)$ e poichè $\Delta f = 1/NT$ questo ritardo corrisponde appunto al termine W_N^{-k} .

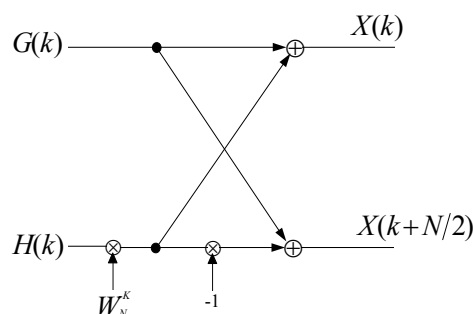


Figura 2.7: Grafo a "farfalla"

2.6.2 Occupazione di memoria e *bit reversal*

La trasformazione di Fourier veloce ha il vantaggio di potere essere fatta sul posto: in altre parole, non c'è bisogno di ulteriore memoria per effettuare la trasformazione. Una volta memorizzata la sequenza di N campioni, si possono realizzare i primi grafi a farfalla 2×2 e successivamente gli altri scambiando i campioni memorizzati con quelli via via calcolati. Alla fine, ci si troverà con la trasformata memorizzata là dove era prima memorizzata la sequenza. Supponiamo ora che $N = 2^m$, e che quindi l'indice del campione sia un numero binario. Si verifica che, per avere i dati della trasformata (antitrasformata) memorizzati in ordine

2.7 FILTRAGGIO DI SEQUENZE ILLIMITATE

progressivo, i dati originali vanno ordinati disponendo prima i campioni di posto pari (ultimo bit 0), e tra quelli, prima quelli di posto pari (penultimo bit 0) e poi ancora tra quelli, quelli di posto pari (terzultimo bit 0) ecc. e così per tutti gli stadi del calcolo. Quindi, per avere la trasformata nell'ordine naturale, i dati vanno scritti nell'ordine inverso dei bit del loro indice e cioè *bit reversed*. Se i dati sono trascritti nell'ordine *bit reversed*, partendo da dati in ordine naturale, l'ordinamento della trasformata risulterà progressivo. Similmente, se i dati sono scritti in ordine naturale, sarà la trasformata ad essere disposta in ordine *bit reversed*. Ad esempio, con $m = 3$, l'ordine dei campioni del segnale *bit reversed* è:

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

Come osservato, si può ottenere quest'ordine da quello naturale senza ulteriore occupazione di memoria, a parte due registri per ogni coppia di numeri complessi da scambiare di posizione.

2.7 Filtraggio di sequenze illimitate

2.7.1 Costo computazionale della FFT

Determiniamo anzitutto il costo computazionale di una trasformata di Fourier veloce (trascurando il costo computazionale delle somme). Ad ogni stadio della trasformazione devono essere calcolati $N/2$ grafi a farfalla che costano una moltiplicazione complessa l'uno (fig.2.8.) Sono da calcolare $\log_2 N$ stadi e quindi un totale di $(N \log_2 N)/2$ moltiplicazioni complesse ognuna corrispondente a 4 moltiplicazioni reali l'una. Se poi si considera che con una trasformata di Fourier veloce si possono fare 2 trasformate di Fourier di due sequenze reali di N campioni l'una, disponendo una delle due sequenze sull'asse reale e la seconda sull'asse immaginario, e poi combinando in modo opportuno le parti reale ed immaginaria della trasformata, si conclude che il costo computazionale di una trasformata di Fourier di una sequenza reale è di $\log_2 N$ moltiplicazioni per campione. Infatti, assegnate le funzioni reali $s_i(t)$; $i = 0, 1$, consideriamo il segnale complesso:

$$s(t) = s_0(t) + js_1(t) \leftrightarrow S(\omega) = S_0(\omega) + jS_1(\omega) \quad (2.23)$$

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

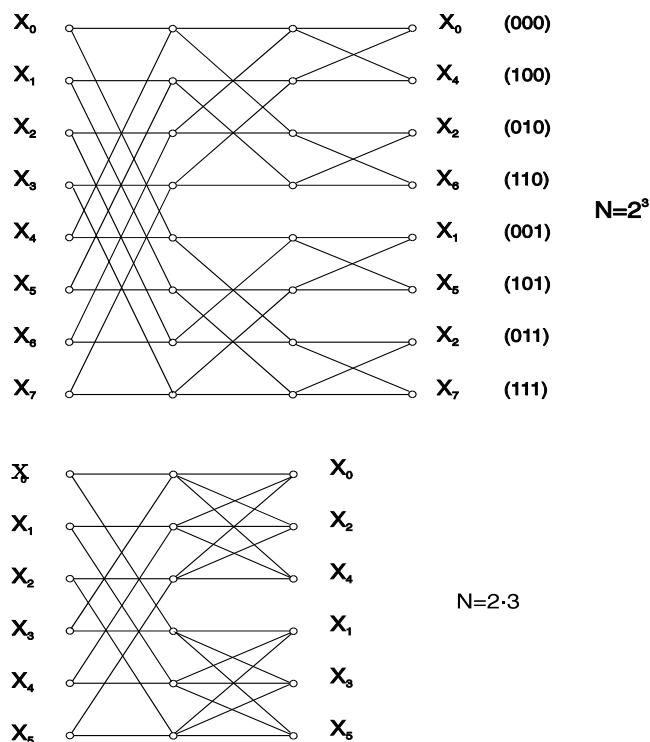


Figura 2.8: Grafi corrispondenti a trasformazioni di Fourier veloci

per le simmetrie delle trasformate delle funzioni reali:

$$S_i(-\omega) = S_i^*(\omega); \quad i = 0, 1; \quad (2.24)$$

Pertanto:

$$2j^i S_i(\omega) = S(\omega) + (-1)^i S^*(-\omega); \quad i = 0, 1 \quad (2.25)$$

2.7.2 Segmentazione di sequenze illimitate

Supponiamo di volere filtrare una sequenza di lunghezza illimitata con un filtro di M campioni; come scegliere N , lunghezza delle trasformate di Fourier? Se N è troppo piccolo, sarà eccessivo il numero dei campioni da eliminare o ricalcolare a causa del meccanismo overlap save o overlap add. Peraltro, se N è troppo grande, sprecheremo calcoli per fare trasformate di Fourier troppo lunghe.

2.8 TRASFORMATATA DI FOURIER IN BASE 4

L'ottimo si calcola subito: il numero di campioni utili per trasformata è pari a: $N' = N - M + 1$; il costo di ogni trasformata è $N \log_2 N$ moltiplicazioni reali; il costo effettivo per campione sarà quindi

$$c = \frac{N \log_2 N}{N - M + 1} \quad (2.26)$$

che ha un valore minimo per $N \simeq 10M$. Tenendo conto del costo computazionale di trasformata, antitrasformata e prodotto di spettri, l'uso della trasformata di Fourier risulta quindi preferibile non appena il filtro è più lungo di una ventina di campioni; tuttavia ciò comporta una latenza (ritardo della disponibilità) di N campioni (invece che M), ritardo che è inaccettabile in alcuni casi.

2.8 Trasformata di Fourier in base 4

Nella sezione precedente abbiamo visto come si effettuano trasformate di Fourier veloci con N potenza di 2. Lo stesso ragionamento si estende al caso in cui N è scomponibile in fattori primi piccoli (2, 3, 5); allora non avremo più farfalle ma strutture di calcolo un po' più complesse: un caso particolare si ha per la base 4, per cui ora faremo i conti per esteso. Suddividendo la sequenza in quattro sottosequenze, il calcolo della FFT procede così:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/4-1} x_{4n} W_N^{-4nk} + \sum_{n=0}^{N/4-1} x_{4n+1} W_N^{-(4n+1)k} + \\ &+ \sum_{n=0}^{N/4-1} x_{4n+2} W_N^{-(4n+2)k} + \sum_{n=0}^{N/4-1} x_{4n+3} W_N^{-(4n+3)k} = \\ &= \sum_{n=0}^{N/4-1} x_{4n} W_{N/4}^{-nk} + W_N^{-k} \sum_{n=0}^{N/4-1} x_{4n+1} W_{N/4}^{-nk} + \\ &+ W_N^{-2k} \sum_{n=0}^{N/4-1} x_{4n+2} W_{N/4}^{-nk} + W_N^{-3k} \sum_{n=0}^{N/4-1} x_{4n+3} W_{N/4}^{-nk} \end{aligned} \quad (2.27)$$

CAPITOLO 2 LA TRASFORMATTA DI FOURIER DISCRETA

Inoltre, poiché $W_N^{N/4} = j$

$$\begin{aligned} X(k + mN/4) = & \sum_{n=0}^{N/4-1} x_{4n} W_{N/4}^{-nk} + j^m W_N^{-k} \sum_{n=0}^{N/4-1} x_{4n+1} W_{N/4}^{-nk} + \\ & + (-1)^m W_N^{-2k} \sum_{n=0}^{N/4-1} x_{4n+2} W_{N/4}^{-nk} + (-j)^m W_N^{-3k} \sum_{n=0}^{N/4-1} x_{4n+3} W_{N/4}^{-nk}; \quad m = 0, 1, 2, 3 \end{aligned} \quad (2.28)$$

In forma matriciale:

$$\begin{bmatrix} X(k) \\ X(k + N/4) \\ X(k + N/2) \\ X(k + 3N/4) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix} \times \begin{bmatrix} G(k) \\ W_N^{-k} H(k) \\ W_N^{-2k} L(k) \\ W_N^{-3k} M(k) \end{bmatrix} \quad (2.29)$$

dove con G, H, L, M sono state indicate le trasformate delle quattro sottosequenze, ciascuna di lunghezza $N/4$. Moltiplicare per j non costa numericamente, e quindi si osserva che la FFT in base 4 ha 3 moltiplicazioni complesse per farfalla, $N/4$ farfalle per stadio, $\log_4 N$ stadi: in totale il numero di moltiplicazioni complesse é

$$n'' = 3 \times \frac{N \log_2 N}{4 \cdot 2} = 0.75 \times \frac{N}{2} \log_2 N \quad (2.30)$$

con un risparmio, rispetto alla trasformata di Fourier veloce in base 2, del 25%, che é significativo .

Capitolo 3

Ricostruzione di segnali continui

3.1 Introduzione: sequenze limitate

Nel trattamento del segnale numerico, vi sono numerosissime occasioni per effettuare delle interpolazioni: il caso ovvio è quello in cui abbiamo un segnale che è noto in certe posizioni e che deve essere valutato in altre, dove non è disponibile una misura diretta. È anche utile mescolare sequenze numeriche campionate con diverse frequenze di campionamento, oppure sequenze campionate con orologi imperfetti: si pensi ad esempio ad una registrazione numerica di un segnale musicale, a cui si voglia sommare un segnale audio di commento, raccolto con un altro registratore, magari dotato di una mediocre temporizzazione. Operando su segnali analogici, basta un nodo sommatore ed il problema è banale: non così nel caso di segnali campionati, in cui si deve trasformare un segnale campionato su un pettine di campionamento in un segnale campionato su un altro pettine. In più i due pettini di campionamento possono essere entrambi irregolari a causa di imperfezioni negli orologi. È chiaro che le regole del campionamento vanno rispettate, e pertanto i segnali vanno filtrati ognuno nella propria banda di Nyquist; uno dei due sarà quindi sovracampionato e cioè non occuperà tutta la banda disponibile. Si vedrà che la precisione con cui va fatta l'operazione di interpolazione deve essere molto elevata, se si desidera conservare un'elevata qualità del segnale campionato. In questo capitolo, analizzeremo l'operazione di interpo-

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

lazione intesa come ricostruzione di un segnale continuo, eventualmente ricampionabile con le regole esposte nel capitolo 1. Invece, nel capitolo 7, discuteremo di tecniche che trasformano direttamente una sequenza in un'altra campionata ad altra frequenza senza passare, nemmeno concettualmente, per il momento in cui il segnale è continuo.

La ricostruzione di un segnale continuo a partire da una sequenza di campioni posizionati regolarmente lungo un asse è possibile inserendo un filtro ideale passa basso che opera sulla sequenza di impulsi ideali corrispondenti ai campioni, che opera come interpolatore. Tuttavia, nascono problemi con il ritardo infinito apportato da questo filtro; in altre parole, per fare la ricostruzione ad un dato tempo servono infiniti campioni del segnale sia prima che dopo quell'istante. Il segnale continuo $x(t)$ è ottenuto dalla sequenza dei campioni come:

$$x(t) = \sum_{n=-\infty}^{+\infty} x_n \frac{\sin \frac{\pi(t-nT)}{T}}{\frac{\pi(t-nT)}{T}} = \sum_{n=-\infty}^{+\infty} x_n \operatorname{sinc} \frac{t-nT}{T} \quad (3.1)$$

Abbiamo adottato la definizione

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (3.2)$$

E' ora opportuno determinare tecniche di ricostruzione di un segnale continuo che siano più semplici ed in particolare diano un risultato anche con sequenze corte; quali sono i requisiti da imporre, visto che non sarà comunque possibile garantire di ottenere di nuovo la $x(t)$ iniziale? Infatti questa corrisponderebbe al risultato dell'interpolazione fatta con un filtro ideale, disponendo di un illimitato numero di campioni prima e dopo l'istante di interesse. Se i campioni disponibili sono limitati, e quindi "pochi", come accadrà in genere, cosa è opportuno fare?

Un interpolatore è quindi un filtro definito su un asse temporale o spaziale continuo che, alimentato da una sequenza limitata di impulsi ideali di ampiezza proporzionale a quella dei campioni della sequenza numerica, permette di ricostruire un segnale $\hat{x}(t)$ quanto possibile simile al segnale originario $x(t)$. In sostanza sarà un buon filtro passabasso, in grado di rimuovere al meglio le repliche dello spettro del segnale generate dal campionamento (fig. 3.1). Diremo che un interpolatore "rispetta" (to honor, in inglese) il dato, se il segnale continuo ricampionato restituisce

3.1 INTRODUZIONE: SEQUENZE LIMITATE

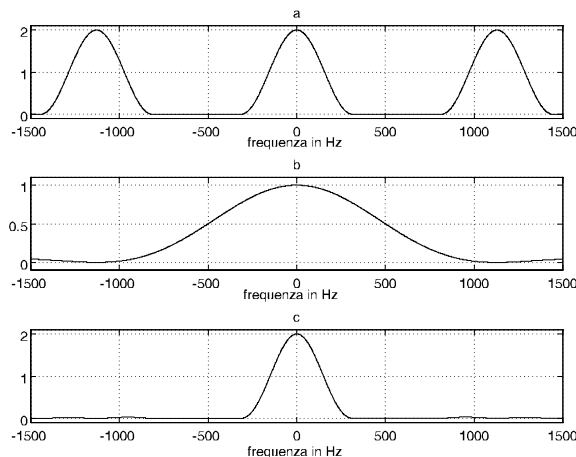


Figura 3.1: Rimozione delle repliche spettrali di un segnale campionato con un filtro ricostruttore triangolare

il campione assegnato. Non sempre gli interpolatori rispettano i dati; a volte, specialmente quando si può supporre che alla sequenza di ingresso sia sovrapposto del disturbo, può essere preferibile che il segnale continuo in uscita, ricampionato, restituisca una sequenza non identica a quella originale, ma smussata, proprio per ridurre l'effetto del disturbo. Si suppone cioè che sui campioni da interpolare sia sommata una sequenza indipendente che va rimossa.

Ancora, per descrivere un interpolatore, si può affermare che rispetta la costante, la retta, la parabola ecc, quando si ha che, se la sequenza di ingresso è ottenuta campionando una costante, una retta, una parabola ecc, il segnale continuo di uscita è appunto lo stesso segnale di partenza; si osservi che un interpolatore può rispettare le rette ma non i dati o viceversa (fig. 3.2).

Rispettare la costante significa che la funzione interpolata è costante, se tutti i campioni della sequenza sono eguali tra loro; più in generale, se la funzione campionata fosse un polinomio di grado k , la funzione ricostruita dall'interpolatore, di complessità crescente con k , dovrà essere identica al polinomio di partenza.

Indichiamo con ordine dell'interpolatore il grado dei polinomi utilizzati per effettuare l'interpolazione. L'interpolatore di ordine $k = 0$ è detto mantenitore: è quello che conserva le costanti (e solo quelle). Possi-

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

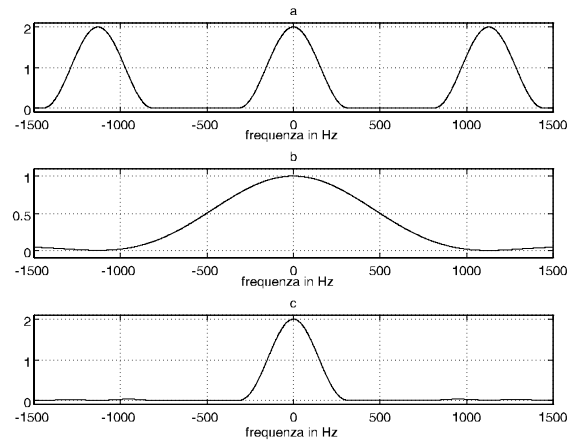


Figura 3.2:

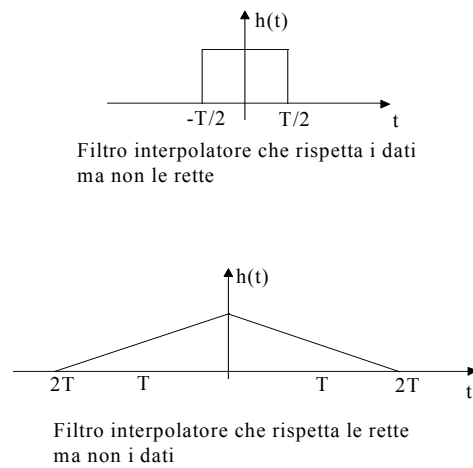


Figura 3.3:

3.2 RICOSTRUZIONE DI POLINOMI

amo poi considerare l'interpolatore lineare (ordine $k = 1$) che conserva le rette e tra queste quella con coefficiente angolare nullo e cioè la costante. L'interpolatore quadratico (ordine $k = 2$), conserverà le parabole, e cioè, quando i dati di ingresso sono disposti su una parabola campionata, il risultato dell'interpolazione sarà proprio la stessa parabola continua nel tempo. Si considera anche l'interpolatore cubico (ordine $k = 3$); gli interpolatori utilizzati in pratica poi sono il mantenitore, l'interpolatore lineare e un particolare interpolatore cubico detto anche bacchetta elastica o "spline".

Per caratterizzare un interpolatore, utilizzeremo la risposta all'impulso, e cioè la funzione tempo continua $h_k(t)$, uscita dell'interpolatore, quando all'ingresso vi è la sequenza corrispondente all'impulso unitario $x_n = \delta_n$:

$$\delta_n = 0; \forall n \neq 0; \delta_0 = 1 \quad (3.3)$$

La risposta all'impulso $h_k(t)$ di una classe di interpolatori di ordine k , che ora considereremo, ha la durata $(k + 1)T$ ed è costituita da $k + 1$ archi di polinomi di grado k -esimo. Indicheremo con $H_k(\omega)$ la trasformata di Fourier di $h_k(t)$. Indicata con $x_k(t)$ la funzione interpolata con l'interpolatore di ordine k si ha, se la funzione ricostruita rispetta il dato:

$$x_k(t) = \sum_{n=-\infty}^{+\infty} x_n h_k(t - nT) \quad (3.4)$$

$$x_k(nT) = x_n; \rightarrow h_k(0) = 1; \quad h_k(nT) = 0; \quad n \neq 0$$

Si noti che in genere $x_k(t) \neq x(t)$ e cioè la funzione interpolata non è eguale alla funzione originale, a meno che questa non fosse polinomiale del grado k .

3.2 Ricostruzione di polinomi

E' facile trovare le condizioni che devono essere soddisfatte dalla risposta all'impulso di un interpolatore perché un polinomio di grado k sia ricostruito correttamente. Partiamo da $k = 0$; basta applicare la sovrapposizione degli effetti: se tutti i campioni della sequenza sono eguali tra loro ($x_n = x_0$), la funzione interpolata sarà la sovrapposizione di tante repliche eguali della risposta all'impulso $h_k(t)$, sfalsate di T l'una rispetto all'altra. Ma replicare con passo T nel dominio del tempo implica

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

campionare a passo $f_c = 1/T$ nel dominio delle frequenze e quindi la trasformata della funzione interpolata sarà eguale alla trasformata della risposta all'impulso campionata alle frequenze m/T ; perché poi questa funzione campionata si riduca ad un impulso a $f_c = 0$ e quindi sia la trasformata di una costante, bisogna che:

$$H_k(m\omega_c) = 0; \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0 \quad (3.5)$$

Quindi per rispettare le costanti basta che la trasformata della risposta all'impulso dell'interpolatore si annulli in corrispondenza della frequenza di campionamento e delle sue armoniche. Tra tutti gli interpolatori che rispettano la costante, ce ne è uno, più semplice di tutti, chiamato *mantenitore*.

Se il segnale $x(t)$ ha uno spettro che non occupa solo la frequenza zero come una costante, ma è concentrato alle basse frequenze come una rampa, non basterà che la trasformata della risposta all'impulso dell'interpolatore sia nulla solo alle frequenze mf_c per ottenerne una interpolazione corretta. Essa deve essere nulla in tutto un intorno di queste frequenze, in modo da *azzerare* le repliche dello spettro. Si verifica che non solo $H_k(\omega)$ deve essere nulla alle frequenze mf_c ma devono anche essere nulle almeno le prime k derivate rispetto alla pulsazione ω . Parimenti, in un intorno dell'origine, k derivate rispetto a ω della $H_k(\omega)$ devono essere nulle, e cioè la $H_k(\omega)$ deve essere quanto più costante possibile, intorno a frequenza zero. Se k tende ad infinito, si ottiene il filtro ideale.

Nel caso dell'interpolatore a bacchetta elastica, accade appunto che le derivate della funzione di trasferimento (fino alla derivata terza compresa) sono nulle alle armoniche della frequenza di campionamento; a frequenza zero, la funzione è costante fino a termini di quarto ordine. Quindi la bacchetta elastica onora i polinomi di terzo grado.

3.3 Interpolatori derivati dal mantenitore

L'interpolatore più semplice che onora la costante è il mantenitore cui corrisponde la risposta all'impulso (fig. 3.3.a):

$$h_0(t) = 1; \quad |t| \leq \frac{T}{2} \quad \leftrightarrow \quad H_0(\omega) = T \frac{\sin \pi f T}{\pi f T} = T \operatorname{sinc}(fT) \quad (3.6)$$

3.3 INTERPOLATORI DERIVATI DAL MANTENITORE

Infatti,

$$H_0(m\omega_c) = T \frac{\sin \pi m}{\pi m} = 0; \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0$$

Immediatamente piú complesso, ma ancora molto semplice é l'interpolatore lineare che ha risposta all'impulso e funzione di trasferimento (fig. 3.3.b):

$$h_1(t) = 1 - \frac{|t|}{T}; \quad |t| \leq T; \quad h_1(t) = \frac{1}{T} (h_0(t) * h_0(t)) \quad (3.7)$$

$$H_1(\omega) = T \operatorname{sinc}(fT)^2 \quad (3.8)$$

Si vede subito che

$$H_1(m\omega_c) = T \left(\frac{\sin \pi m}{\pi m} \right)^2 = 0; \quad \frac{dH_1}{d\omega}(m\omega_c) = 0 \quad \omega_c = \frac{2\pi}{T}; \quad m \neq 0$$

In entrambi i casi la risposta all'impulso é molto corta (T o 2T); poi possono essere interessanti gli interpolatori

$$H_2(\omega) = T \operatorname{sinc}(fT)^3; \quad \leftrightarrow \quad h_2(t) = \frac{1}{T} (h_1(t) * h_0(t)) \quad (3.9)$$

$$H_3(\omega) = T \operatorname{sinc}(fT)^4; \quad \leftrightarrow \quad h_3(t) = \frac{1}{T} (h_2(t) * h_0(t)) \quad (3.10)$$

(il valore della costante si trova imponendo pari a T il guadagno a frequenza zero). Il problema di questi ultimi due interpolatori (fig. 3.3.c, 3.3.d) é che rispettano polinomi di grado k ma non sequenze qualsiasi di dati; infatti le risposte all'impulso $h_2(t)$ e $h_3(t)$ non si annullano per $t = T$ (vedi fig 3.3); infatti la funzione di trasferimento decade per valori bassi della frequenza e quindi la sequenza risultato del ricampionamento del segnale interpolato é una versione **smussata** della sequenza assegnata.

3.3.1 Interpolatore cubico troncato

Un interpolatore interessante per la sua semplicità ed accuratezza è il seguente interpolatore cubico, che ha durata $4T$. Le equazioni sono:

$$y(-x) = y(x) = 1 - (1+a)x^2 + ax^3; \quad 0 \leq x = \frac{t}{T} \leq 1 \quad (3.11)$$

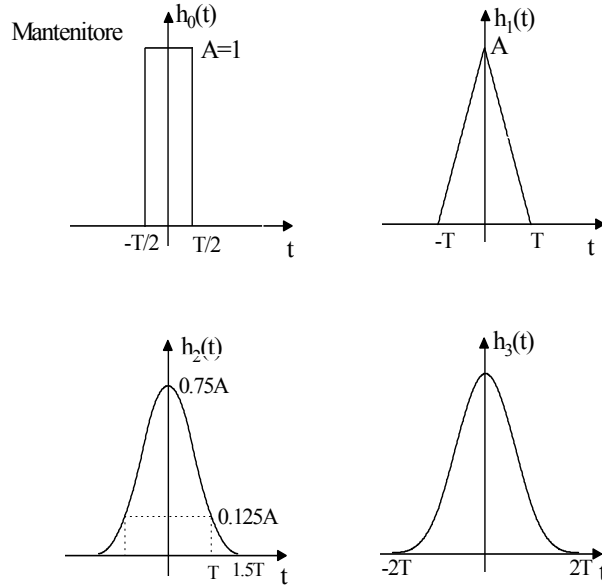
Interpolatori di ordine k derivati dal mantenitore

Figura 3.4: Interpolatori di ordine k derivati dalla convoluzione con se stesso del mantenitore: a) $k = 0$ mantenitore; b) $k = 1$ lineare; c) $k = 2$ parabolico; d) $k = 3$ cubico

$$y(-x) = y(x) = (a - 2)(x^3 - 5x^2 + 8x - 4); \quad 1 \leq x \leq 2 \quad (3.12)$$

Il passaggio per lo 0 a $\pm T$ garantisce che i dati sono onorati. inoltre, per tutti i valori del parametro a la derivata prima è continua; la derivata seconda è discontinua agli estremi e pertanto è inutile imporre la continuità all'interno. Per $a = 3/2$ la funzione di trasferimento è massimamente piatta (fino al 4° ordine) nell'origine e gli zeri alle frequenze multiple di $1/T$ sono di secondo ordine; questo quindi è il valore consigliato per a .

3.3.2 Interpolatori con filtro ideale troncato

Altri interpolatori si possono trovare ricorrendo al filtro ideale, e limitandone la durata troncandolo con una finestra p.e. a coseno rialzato

come si vedrà nel capitolo 5.

3.4 Rigidità e sovraelongazioni

Consideriamo ora vari interpolatori, dal mantenitore o dall'interpolatore lineare, che non hanno memoria da un campione al successivo, al filtro ideale, dove l'influenza di un campione si estende all'infinito. Nel caso più semplice (il mantenitore) non è garantita nemmeno la continuità della funzione; poi, via via, si accresce la *rigidità* dell'interpolatore e cioè si impone la continuità della funzione (interpolatore lineare) e poi di un crescente numero di derivate. Nel caso estremo del filtro ideale, vi è continuità di tutte le derivate in corrispondenza dei campioni; quindi non vi è più possibilità di ritrovare la posizione del reticolo di campionamento dal segnale ricostruito, e di fatto il segnale ricostruito è quello di partenza, indipendente dalla posizione del reticolo di campionamento.

Le condizioni agli estremi sono tanto più influenti quanto più l'interpolatore è rigido. Nel caso del mantenitore e dell'interpolatore lineare, non sono rilevanti; nel caso del filtro ideale invece, se la sequenza da interpolare è limitata, si possono arbitrariamente assegnare tutte le derivate del segnale da interpolare in corrispondenza del primo e dell'ultimo campione. Ciò influenza l'andamento del segnale all'interno dell'intervallo, pur senza alterare i valori dei campioni, all'interno dell'intervallo. **Pertanto, è arbitrario usare interpolatori rigidi (come il filtro ideale) con sequenze corte.** È ragionevole utilizzare interpolatori con lunghezza della risposta all'impulso molto più breve della lunghezza della sequenza da interpolare. In questo modo, gli effetti di bordo sono limitati.

Nel caso dell'interpolatore lineare, i valori minimi e massimi del segnale interpolato sono quelli dei campioni; interpolando invece i dati come con una bacchetta flessibile l'escursione dei valori è maggiore. Utilizzando il filtro ideale, imponendo la continuità di tutte le derivate e quindi una risposta all'impulso dell'interpolatore di lunghezza infinita, può addirittura accadere che l'uscita sia illimitata per alcune sequenze di dati.

Per fare un caso evidente, consideriamo una sequenza alla frequenza di Nyquist, che moltiplichiamo per la funzione $\text{sign}(n) = n/|n|$; $\text{sign}(0) = 0$ (fig. 3.4). Il segnale interpolato con il filtro ideale tende all'infinito in corrispondenza del cambio di segno. Poiché la singolarità nasce dalla somma dei *sinc* (una serie armonica) che si combinano con segni eguali,

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

l'ampiezza cresce con il logaritmo del numero di campioni della sequenza interpolata. In proposito, si ricordi che

$$\sum_{n=1}^N |\text{sinc}(\frac{t}{T} - n)| = |\text{sinc}(\frac{t}{T})| \sum_{n=1}^N \frac{1}{1 + n\frac{T}{t}}$$

e poi che la serie logaritmica è:

$$\sum_{n=1}^{+\infty} \frac{x^n}{n} = \log \frac{1}{1-x}; \quad |x| < 1 \quad (3.13)$$

Al limite, per $x = 1$, si trova la serie armonica. Si vedrà poi al capitolo 6 quando si discuterà della trasformazione di Hilbert, che l'involuppo di un segnale interpolato dipende dalla trasformata di Hilbert del segnale stesso; la trasformata di Hilbert della funzione *sign* ed in generale di ogni gradino ha una singolarità di tipo logaritmico che si ritrova sia nell'involuppo che nel segnale interpolato.

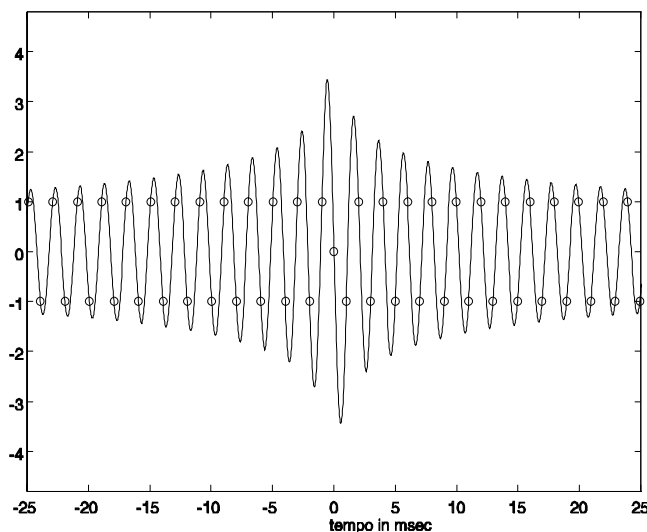


Figura 3.5: Esempio di sovraelongazioni dovute all'eccessiva rigidità del filtro ideale usato come interpolatore

3.5 La linea elastica o "spline"

Esiste un altro interpolatore cubico, diverso da $H_3(\omega)$ e dall'interpolatore cubico troncato visto in precedenza che non solo rispetta rette, parabole e cubiche ma rispetta anche i dati, cosa che non accade per $H_3(\omega)$. Questo interpolatore, noto come linea elastica (inglese: *spline* = bacchetta elastica), è molto utilizzato sia in sistemi monodimensionali che multidimensionali.

Intuitivamente, si può pensare di tracciare su un foglio di carta le posizioni dei campioni del segnale e puntarvi degli spilli. Per costruire una curva continua che li congiunga, si può usare un filo teso tra uno spillo e l'altro (ed allora risulta l'interpolazione lineare) oppure usare come curvilineo una bacchetta flessibile (un filo di acciaio, p.e.) la cui posizione di equilibrio rappresenta il segnale interpolato (fig. 3.5).

L'interpolazione effettuata con la bacchetta elastica (la linea elastica) rispetta costanti e rette (ma anche parabole e cubiche) ed anche il dato. Peraltro, il risultato dell'interpolazione dipende anche da come la bacchetta è fissata agli estremi dell'intervallo. lasciando liberi gli estremi, la posizione di equilibrio della bacchetta corrisponde al minimo dell'energia elastica immagazzinata e pertanto la curvatura media della linea è minima; il segnale interpolato è allora, pur sempre onorando i dati, il più *liscio* possibile. La bacchetta proseguirà fuori dell'intervallo in cui è nota la sequenza fornendo una particolare estrapolazione dei dati. Altrimenti, si può incastrare agli estremi la bacchetta fissando così la derivata prima del segnale interpolato agli estremi dell'intervallo (non necessariamente a zero).

3.5.1 Analisi dello "spline"

Passando ad un'analisi quantitativa della posizione di equilibrio della bacchetta, osserviamo che la sua curvatura è proporzionale al momento flettente in quel punto (fig. 3.9); le reazioni vincolari dei nodi sono forze che la inflettono: il momento flettente dovuto all'insieme delle reazioni vincolari, ha un andamento lineare a tratti, cambiando pendenza in corrispondenza di ogni nodo e di ogni nuova reazione vincolare. In conclusione, la posizione di equilibrio della bacchetta e pertanto la curva interpolante ha una curvatura che varia linearmente tra un nodo e l'altro. Se si fa l'ipotesi addizionale che gli scostamenti $y(x)$ della bacchetta

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

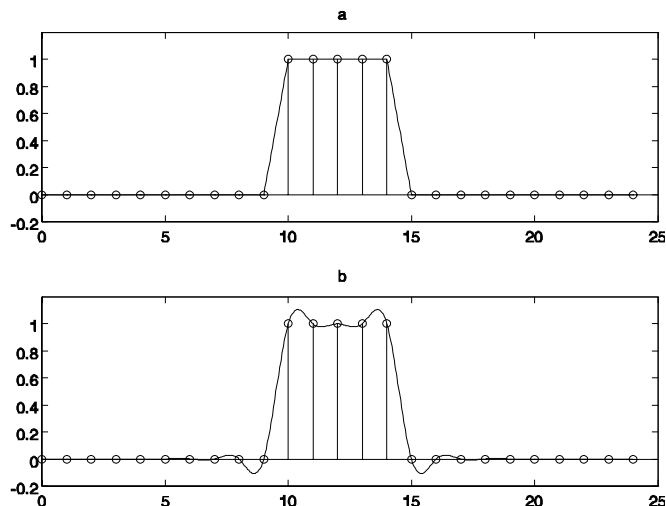


Figura 3.6: Interpolazione di un impulso rettangolare: lineare (a); con bacchetta elastica (b).

dall'asse delle ascisse siano molto piccoli, si può dire che la sua curvatura è proporzionale alla derivata seconda della curva, che è quindi lineare a tratti. Infatti si ha che la curvatura di una linea, inverso del raggio di curvatura $r(x)$ è:

$$\frac{1}{r(x)} = \frac{y''(x)}{[1 + y'(x)^2]^{3/2}} \simeq y''(x); \quad |y'(x)| < 1 \quad (3.14)$$

In ogni tratto tra due campioni, la curva ricostruita sarà pertanto una cubica, definita cioè con quattro parametri, ad esempio i valori della funzione e delle derivate in corrispondenza dei campioni. Le derivate sono incognite e vanno determinate utilizzando le condizioni di continuità delle derivate prime e seconde da un tratto di curva all'altro.

3.5.2 L'equazione dello spline

Indichiamo con y_i, y'_i i valori della funzione da interpolare e della sua derivata in corrispondenza del campione i -esimo. La funzione interpolante cercata è un polinomio di terzo grado (e cioè con una derivata seconda che varia linearmente con la variabile x). Nell'intervallo $i, i + 1$

3.5 LA LINEA ELASTICA O "SPLINE"

la funzione interpolante $f(x)$ è descritta dalla seguente equazione:

$$f(x) = y_i(1 - 3x^2 + 2x^3) + y_{i+1}(3x^2 - 2x^3) + \quad (3.15)$$

$$+ y'_i(x - 2x^2 + x^3) + y'_{i+1}(-x^2 + x^3)$$

Per verificare l'equazione precedente, si supponga diverso da zero solo uno dei quattro parametri y_i , y_{i+1} , y'_i , y'_{i+1} , e si suppongano nulli gli altri tre. I polinomi che risultano, hanno appunto valore nullo in loro corrispondenza: ad esempio, nel caso di y_i , il polinomio $1 - 3x^2 + 2x^3$ ha valore 1 in $x = 0$ è nullo per $x = 1$ e le sue derivate sono nulle in $x = 0, 1$. Nel caso di y'_{i+1} , il polinomio $-x^2 + x^3$ ha valore 0 in $x = 0; 1$ ha derivate prime nulle in 0 ed unitaria in 1. Pertanto la $f(x)$ descritta dall'equazione precedente è il più generale polinomio di terzo grado che abbia assegnati i valori della funzione e delle derivate agli estremi.

E' possibile calcolare la derivata seconda del polinomio ed imporre che al nodo i -esimo le derivate seconde calcolate per l'intervallo $i, i + 1$ e $i - 1, i$ coincidano. Ne risulta la seguente equazione alle differenze finite, che ci permette di calcolare i valori delle derivate della funzione da interpolare in corrispondenza dei nodi :

$$y'_{i-1} + 4y'_i + y'_{i+1} = 3(y_{i+1} - y_{i-1}) \quad (3.16)$$

Questa equazione alle differenze finite sarà studiata con le tecniche della trasformata z nel capitolo 3. Qui osserviamo anche che imponendo che la derivata seconda sia nulla ai nodi terminali per imporre la condizione di immagazzinamento della minima energia elastica, si ottiene:

$$y'_0 = \frac{-y'_1 + 3(y_1 - y_0)}{2} \quad (3.17)$$

$$y'_N = \frac{-y'_{N-1} + 3(y_{N-1} - y_N)}{2} \quad (3.18)$$

Abbiamo quindi un sistema di N equazioni nelle N incognite (le y'_i) che è risolvibile facilmente essendo tridiagonale (e cioè avendo nella matrice dei coefficienti solo la diagonale e le due subdiagonali adiacenti diverse da zero). La soluzione viene fatta per sostituzione, a partire dalla prima

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

equazione.

$$\begin{aligned}
 & \begin{bmatrix} 2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 4 & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} y'_0 \\ y'_1 \\ y'_2 \\ \dots \\ y'_{N-2} \\ y'_{N-1} \\ y'_N \end{bmatrix} = \quad (3.19) \\
 & = \begin{bmatrix} -3 & 3 & 0 & \dots & 0 & 0 & 0 \\ -3 & 0 & 3 & \dots & 0 & 0 & 0 \\ 0 & -3 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 3 & 0 \\ 0 & 0 & 0 & \dots & -3 & 0 & 3 \\ 0 & 0 & 0 & \dots & 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ y_{N-2} \\ y_{N-1} \\ y_N \end{bmatrix}
 \end{aligned}$$

Una volta trovate le N incognite y'_i esse vanno sostituite nell'equazione (5.48) e quindi si determinano i vari archi di polinomi di terzo grado.

Si osservi anche che il concetto di bacchetta elastica si estende con facilit  al caso di campioni non equidistanti; in questo caso, i coefficienti lungo la diagonale e le subdiagonali non sono costanti.

3.5.3 La risposta all'impulso dello spline

Vogliamo ora determinare la risposta all'impulso dello spline $s_0(t)$ cio  la funzione che si ottiene quando solo $y_0 = 1$ e tutti gli altri $y_i = 0$; $-\infty < i < \infty$.

Lo spettro di $s_0(t)$ si pu  calcolare con il seguente metodo. Anzitutto, dall'equazione alle differenze finite (5.48) si trova la trasformata di Fourier della sequenza delle derivate della $s_0(t)$ in corrispondenza dei campi-

3.5 LA LINEA ELASTICA O "SPLINE"

oni; infatti indicando rispettivamente con $Y'_c(\omega)$, $Y_c(\omega)$ le trasformate di Fourier delle sequenze delle derivate e dei campioni.

$$y'_{i-1} + 4y'_i + y'_{i+1} = 3(y_{i+1} - y_{i-1}) \rightarrow \quad (3.20)$$

$$Y'_c(\omega)[e^{j\omega T} + 4 + e^{-j\omega T}] = 3Y_c(\omega)[e^{j\omega T} - e^{-j\omega T}]$$

$$Y'_c(\omega) = Y_c(\omega) \frac{3j \sin \varphi}{2 + \cos \varphi}; \quad \varphi = \omega T \quad (3.21)$$

Nel caso particolare in cui $Y_c(\omega) = 1$, e cioè per la risposta all'impulso si ha:

$$Y'_c(\omega) = \frac{3j \sin \varphi}{2 + \cos \varphi} \simeq j\varphi; \quad (3.22)$$

Si osservi che questo derivatore numerico ha un'ottima funzione di trasferimento, nel senso che essa somiglia molto alla funzione di trasferimento desiderata per un derivatore. Infatti

$$Y'_c(\omega) \simeq j\omega Y_c(\omega) = j\omega T Y_c(\omega) \left(1 - \frac{1}{180} \omega^4 T^4 \dots \right); \quad |\omega T| < 1 \quad (3.23)$$

Vedremo poi nel seguito altre tecniche per trovare le derivate di una

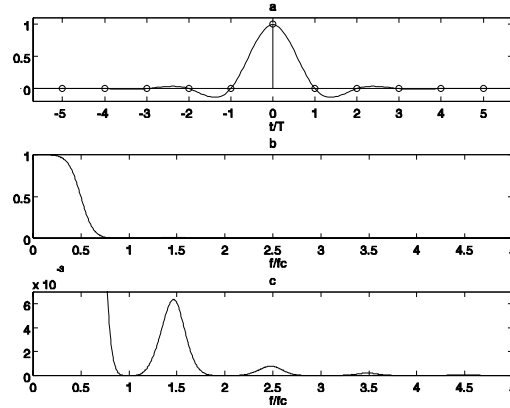


Figura 3.7: Risposta all'impulso (a) e trasformata (b) dello spline; in (c) è ingrandita la trasformata per vederne i valori in corrispondenza delle frequenze $n f_c$.

sequenza numericamente e cioè trovare, agli istanti di campionamento, la sequenza delle derivate del segnale continuo ottenuto filtrando la sequenza iniziale con un filtro ideale.

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

Utilizzando poi le equazioni che danno la funzione a partire dai valori ai nodi dei campioni e delle derivate, si trova la trasformata della risposta all'impulso cercata. Risulta che la risposta all'impulso della bacchetta elastica ha trasformata $S_0(\omega)$ (fig. 3.6) con equazione:

$$S_0(\omega) = \frac{3 \operatorname{sinc}^4(fT)}{1 + 2 \cos^2 \pi fT} \quad (3.24)$$

In particolare $S_0(\omega)$ ha zeri di quarto ordine alle frequenze multiple di $1/T$ ed è costante fino al quarto ordine nell'origine.

3.6 Il campionamento irregolare

3.6.1 Conversione di frequenze di campionamento

In molte occasioni può essere utile convertire un segnale da una frequenza di campionamento ad un'altra: p.e. inserire un segnale audio da un CD in una linea telefonica numerica e quindi passare da circa 44KHz a 8KHz di frequenza di campionamento; a parte il necessario filtraggio per sopprimere l'alias, sarà anche necessario interpolare il segnale su un altro pettine di campionamento, totalmente indipendente; ovviamente, sarebbe possibile ritornare ad un segnale analogico e ricampionare, ma ciò comporta delle perdite di rapporto segnale rumore. Ancora, è utile mescolare sorgenti numeriche di origini diverse, anche se nominalmente alla stessa frequenza di campionamento; per esempio, si vogliono mescolare segnali audio registrati da varie posizioni in un teatro, durante lo stesso concerto, da diversi registratori numerici. I vari clock usati per le registrazioni possono essere imperfetti e le loro frequenze possono pendolare rispetto alla frequenza nominale: i segnali audio devono essere allineati temporalmente e poi i segnali numerici corrispondenti devono essere ricampionati in modo simultaneo. I vari campioni devono essere interpolati. Un ulteriore caso è la tecnica analogica varispeed in cui l'asse temporale viene accelerato o ritardato per ottenere degli effetti audio; è opportuno realizzare la stessa funzione con tecniche numeriche.

Inoltre, vi sono variazioni casuale della frequenza di orologio dovute a imperfezioni del sistema di trasmissione numerico indicate con il nome di **jitter**. È approssimativamente possibile rimuovere il jitter di trasmissione da una sequenza; non quello di registrazione, tuttavia, a meno di

3.6 IL CAMPIONAMENTO IRREGOLARE

non disporre di altre fonti di sincronizzazione come ad esempio i segnali stessi. La sequenza ricevuta viene memorizzata su una memoria FIFO (First In First Out), e quindi questa memoria viene riletta al ritmo opportuno. Questo ritmo deve essere un pò più veloce se la memoria va svuotata, più lento se la memoria tende a riempirsi. I tempi di restituzione sono confrontati con i tempi dell'orologio di campionamento, registrati, ed i campioni ai tempi necessari sono interpolati.

In ogni modo, la conversione della frequenza di campionamento va fatta in due passi: per prima cosa si ottiene una sequenza, approssimativamente campionata al ritmo desiderato, ma che conserva le eventuali imprecisioni del ritmo di campionamento iniziale. Ad esempio la si sovra-campiona di un fattore M per poi sottocampionarla di un fattore N . Per fare ciò, si adottano le tecniche di interpolazione e decimazione discusse nel capitolo 8. Quindi, si fa un'operazione di precisione, calcolando i campioni di una sequenza esattamente temporizzata.

In sostanza, si dispone di una sequenza registrata a tempi "errati" ed irregolari ma noti; vogliamo ricostruire la sequenza in modo corretto. È sorprendente la precisione dell'interpolazione necessaria per garantire un buon rapporto segnale - rumore p.e. 100 dB, corrispondente ai 16 bit con cui i campioni sono registrati su un CD. Il segnale desiderato è:

$$s(nT); \quad -\infty < n < \infty$$

Il segnale disponibile sia invece :

$$s(t_n); \quad t_n - nT = \tau_n \ll T$$

$$s(nT) = s(t_n - \tau_n) \approx s(t_n) - \tau_n \left. \frac{ds(t)}{dt} \right|_{t=t_n}$$

L'errore quadratico medio è pertanto:

$$\overline{\varepsilon^2} = \overline{\tau^2} \times \overline{\left| \frac{ds}{dt} \right|^2}$$

Nel caso di sequenze a spettro uniforme, utilizzando il teorema di Parseval si ottiene, indicando con σ^2 il valore quadratico medio del segnale:

$$\left| \frac{ds}{dt} \right|^2 = \sigma^2 \frac{\pi^2}{3T^2} \longrightarrow SNR^{-1} = 10^{-10} = \frac{\overline{\varepsilon^2}}{\sigma^2} = \frac{\pi^2}{3} \times \frac{\overline{\tau^2}}{T^2} \longrightarrow \sqrt{\frac{\overline{\tau^2}}{T^2}} = .55 \times 10^{-5}$$

CAPITOLO 3 RICOSTRUZIONE DI SEGNALI CONTINUI

L'errore di temporizzazione nell'interpolazione deve essere dell'ordine del **milionesimo di campione** se i successivi errori temporali sono perfettamente casuali e quindi creano un disturbo casuale. Se invece supponiamo che i successivi valori τ_n siano tutti eguali tra loro (o lentamente variabili) allora l'errore diventa praticamente inosservabile perchè equivale ad avere un errore nella funzione di trasferimento pari a:

$$\Delta(\omega) = e^{j\omega\tau}$$

Tuttavia questo ritardo (o anticipo) indesiderato va ridotto a valori accettabili (ben maggiori del milionesimo di campione) se si vogliono temporizzare più segnali assieme.

3.6.2 Interpolazione irregolare con la funzione *sinc*

Rimane da discutere il problema dell'interpolazione di funzioni campionate irregolarmente e cioè ad intervalli ineguali di tempo o spazio. In linea teorica, per N tanto grande da poter applicare in pratica un filtro ideale, se si hanno a disposizione N campioni nell'intervallo NT , anche se non equispaziati, è tuttora possibile determinare gli N campioni equispaziati originali e di qui l'interpolazione con un filtro ideale; basta scrivere un sistema di N equazioni in N incognite dove i termini noti sono gli N valori assunti dalla funzione x_n nei tempi assegnati (t_n irregolari) e le incognite i valori dei campioni a_n ai tempi nT . Il sistema:

$$x(t) = \sum_{n=-N/2}^{+N/2-1} a_n \text{sinc}\left(\frac{t - nT}{T}\right); \quad x(t_n) = x_n; \quad n = -N/2, \dots, N/2 - 1 \quad (3.25)$$

ha tante equazioni quante incognite. In termini matriciali si ha:

$$\underline{\mathbf{G}} \underline{\mathbf{a}} = \underline{\mathbf{x}}$$

dove $\underline{\mathbf{x}}$, $\underline{\mathbf{a}}$ sono i vettori di componenti $\{x_n\}$, $\{a_n\}$ rispettivamente e la matrice $\underline{\mathbf{G}}$ ha componenti

$$g_{i,j} = \text{sinc}\left(\frac{t_i - t_j}{T}\right)$$

Il sistema diventa mal condizionato non appena il campionamento diventa fortemente irregolare; se vi sono due campioni molto vicini, due righe

3.6 IL CAMPIONAMENTO IRREGOLARE

della matrice $\underline{\mathbf{G}}$ saranno molto simili tra loro ed un autovalore almeno della matrice $\underline{\mathbf{G}}$ sarà molto basso, ecc. Quindi, questa tecnica di interpolazione, apparentemente infallibile, è destinata a dare risultati poco utili.

Capitolo 4

La trasformata z

4.1 Introduzione: sequenze e polinomi

Nell'elaborazione numerica dei segnali viene fatto continuo uso della Trasformata z , e cioè di una corrispondenza che viene stabilita tra sequenze di campioni (complessi o reali) e polinomi in una variabile complessa che verrà indicata con z . In questo capitolo cercheremo di riassumere le caratteristiche della trasformata z usando solo concetti elementari. Anzi-tutto ricordiamo che la trasformata z di una sequenza di lunghezza finita $\{a_{1,n}\}$ contenuta tra le posizioni indicizzate con $-M, \dots, N$ è il "polinomio" (con esponenti sia positivi che negativi) nella variabile complessa z :

$$A_1(z) = \sum_{n=-M}^N a_{1,n} z^{-n} \quad (4.1)$$

Ad esempio, la sequenza costituita da un campione di ampiezza 2 al posto caratterizzato dall'indice $n = 18$ e di un altro campione di ampiezza -3 al posto caratterizzato dall'indice $n = -1$ è il "polinomio":

$$A_1(z) = -3 \times z + 2 \times z^{-18}$$

Quindi la trasformata z di una sequenza $\{a_{1,n}\}$ di lunghezza finita a meno del fattore z^M è il "polinomio" nella variabile complessa z^{-1} :

$$A_1(z) = z^M P(z^{-1})$$

dove $P(z)$ è un polinomio vero e proprio (contenente solo potenze positive cioè) di ordine $M+N$. Ritardare di K campioni una sequenza corrisponde

CAPITOLO 4 LA TRASFORMATTA Z

a moltiplicarne la trasformata z per z^{-K} e viceversa anticiparla di H campioni corrisponde a moltiplicarne la trasformata per z^H .

Vediamo ora perchè la trasformata z serve: consideriamo la sequenza $\{c_n\}$ convoluzione di due sequenze $\{a_n\}$, $\{b_n\}$. Si abbia cioè:

$$\{c_n\} = \{a_n\} * \{b_n\}$$

$$c_n = \sum_k a_{n-k} b_k$$

La trasformata z della sequenza $\{c_n\}$ è il prodotto delle trasformate z delle sequenze $\{a_n\}$, $\{b_n\}$. Ciò è immediatamente verificabile mediante la sovrapposizione degli effetti, p.e. rappresentando la sequenza $\{b_n\}$ come somma di impulsi ritardati di k campioni, con k variabile. La sequenza $\{c_n\}$ è la somma delle repliche della sequenza $\{a_n\}$ progressivamente sfalsate di k campioni e moltiplicate per b_k . Dette infatti:

$$A(z) = \sum_l a_l z^{-l}$$

$$B(z) = \sum_k b_k z^{-k}$$

$$C(z) = \sum_n c_n z^{-n}$$

si ha:

$$C(z) = \left(\sum_k b_k z^{-k} \right) \left(\sum_l a_l z^{-l} \right) = B(z) \times A(z)$$

Poi, è anche vero che moltiplicando due "polinomi" se ne ottiene un terzo che ha per sequenza dei coefficienti la convoluzione delle sequenze dei coefficienti dei primi due. A prodotti di convoluzione di sequenze (indicati con $*$), estremamente utili da calcolare, corrispondono quindi i prodotti delle rispettive trasformate z .

$$C(z) = \left(\sum_k b_k z^{-k} \right) \left(\sum_l a_l z^{-l} \right) = \quad (4.2)$$

$$= B(z) \times A(z) = \sum_k \sum_l a_l b_k z^{-(k+l)} \quad (4.3)$$

4.1 INTRODUZIONE: SEQUENZE E POLINOMI

e posto:

$$n = k + l$$

si verifica che:

$$C(z) = \sum_n \left(\sum_k a_{n-k} b_k \right) z^{-n} = \sum_n c_n z^{-n}$$

Per semplicità supporremo che le sequenze studiate abbiano sempre il primo campione nell'origine ($M = 0$) e pertanto le loro trasformate z siano polinomi che contengono il termine di grado zero.

Per il teorema fondamentale dell'algebra ogni polinomio di ordine N a coefficienti complessi ha N radici nel campo complesso z_1, z_2, \dots, z_N ; quindi **ogni sequenza di $N + 1$ campioni complessi che inizia a partire dall'origine dei tempi (il primo campione è posto nella posizione di indice 0) si può decomporre nella convoluzione di N sequenze elementari costituite da 2 campioni complessi di cui il primo è di ampiezza unitaria ed è posizionato nell'origine ed il secondo, posizionato al posto di indice 1, ha valore $-z_n, n = 1, \dots, N$. Infatti una qualsiasi sequenza di N campioni che inizia nell'origine ha trasformata z pari a:**

$$A(z) = \sum_{n=0}^N a_n z^{-n}$$

Ma, per il teorema fondamentale dell'algebra possiamo scrivere:

$$A(z) = \sum_{n=0}^N a_n z^{-n} = a_0 \prod_{n=1}^N (1 - z_n z^{-1}); \quad (4.4)$$

Abbiamo mostrato che la sequenza di $N+1$ campioni $\{a_n; n = 0, 1, 2, \dots, N\}$ è perciò decomponibile nel prodotto di convoluzione che segue:

$$\{a_n\} = a_0 * \{1, -z_1\} * \{1, -z_2\} * \{1, \dots\} * \{1, -z_N\} \quad (4.5)$$

Questo fatto, che **non** si estende a due o più dimensioni, ci porta a grandi semplificazioni; infatti, qualsiasi sequenza di lunghezza finita potrà essere vista come la convoluzione di sequenze elementari di due campioni, che chiameremo zeri per il motivo indicato nella sezione 3.

4.2 Sistemi lineari discreti

Un sistema lineare, a cui si applica la sovrapposizione degli effetti, è caratterizzato da un ingresso ed un'uscita; la linearità permette di calcolare l'uscita supponendo che l'ingresso sia costituito da una successione di segnali impulsivi, che possono essere considerati uno per volta; nel caso discreto, che è quello di interesse nel seguito, il sistema ha per ingresso la sequenza illimitata di campioni $\{x_n\}$ e per uscita la sequenza $\{y_n\}$; **l'intervallo di campionamento in ingresso ed uscita è identico.**

La risposta del sistema ad un impulso nella posizione i -esima sarà una sequenza $h_{i,k}$ con l'indice k compreso tra i valori $-M_i$ e N_i ; la sequenza $h_{i,k}$ può cambiare al variare dell'indice i . Sempre per la sovrapposizione degli effetti, l'uscita sarà

$$y_n = \sum_{k=-M_{n-k}}^{N_{n-k}} x_{n-k} h_{n-k,k} = \sum_{l=n+M_l}^{n-N_l} x_l h_{l,n-l}$$

Ora, supponiamo che la risposta all'impulso del sistema non dipenda dall'indice ma sia tempo (spazio) invariante. L'uscita è

$$y_n = \sum_{k=-M}^N x_{n-k} h_k = \sum_{l=n+M}^{n-N} x_l h_{n-l}; \quad \{y_n\} = \{x_n\} * \{h_n\}$$

e cioè la convoluzione tra le due sequenze $\{x_n\}$ e $\{h_n\}$.

Nel caso che la sequenza di ingresso ad un sistema tempo (spazio) invariante sia una sinusoide campionata è noto (ma lo dimostreremo di nuovo nella prossima sezione) che la risposta è ancora una sinusoide campionata della stessa frequenza di quella di ingresso ma di diversa ampiezza e fase, dovute al transito per il sistema; le caratteristiche di ampiezza e di fase del sistema lineare (al variare della frequenza della sinusoide di ingresso) ne corrispondono alla **funzione di trasferimento**.

4.3 Sinusoidi campionate; zeri

Abbiamo visto come ogni sequenza si può decomporre in una convoluzione di tante sequenze elementari di due campioni; per trovare la funzione di trasferimento di una sequenza di $N + 1$ campioni e quindi cosa accade

4.3 SINUSOIDI CAMPIONATE; ZERI

quando convolviamo una sinusoide (campionata) con una sequenza di $N + 1$ campioni, basta convolvere questa sinusoide successivamente con tutte le N sequenze elementari componenti (la funzione di trasferimento della sequenza è quindi il prodotto delle funzioni di trasferimento delle sequenze elementari). Ci riferiremo a sinusoidi generalizzate, e cioè esponenziali complessi il cui modulo può crescere o decrescere geometricamente con l'indice n . Queste sinusoidi generalizzate sono delle sequenze $\{x_n\}$ illimitate descritte dalla formula:

$$x_n = B\rho^n \times \exp(j2\pi fnT) ; \quad -\infty \leq n \leq \infty \quad (4.6)$$

dove B è un numero complesso e la frequenza f può avere un valore qualsiasi; peraltro è inutile fare uscire f dall'intervallo $-f_N < f < f_N$ dove $f_N = 1/2T$ è la frequenza di Nyquist, e ciò per evitare equivocazione. Nel caso $\rho = 1$ si ritrovano le sinusoidi complesse; le consuete sinusoidi e cosinusoidi ne costituiscono le parti reali ed immaginaria: si ha cioè:

$$B = B_R + jB_I = |B|e^{j\varphi} ; \quad |B| = \sqrt{B_R^2 + B_I^2}; \quad \tan \varphi = \frac{B_I}{B_R}$$

$$Be^{j2\pi fnT} = |B| [\cos(2\pi fnT + \varphi) + j \sin(2\pi fnT + \varphi)] ;$$

Indichiamo ora con z_1 la variabile complessa

$$z_1 = \rho e^{j2\pi fT}$$

Una sinusoide generalizzata è quindi la sequenza:

$$\{x_n\} = \{Bz_1^n\}; \quad -\infty \leq n \leq \infty \quad (4.7)$$

Se si convolve una sinusoide generalizzata con la sequenza di due campioni $1, -z_0$ posizionati nell'origine e nel campione successivo, si osserva che il segnale in uscita è:

$$\{Bz_1^n\} * \{1, -z_0\} = \{Bz_1^n - Bz_0z_1^{n-1}\} = \{B(1 - z_0z_1^{-1})z_1^n\} \quad (4.8)$$

e cioè ancora la stessa sinusoide generalizzata, moltiplicata per il fattore complesso $1 - z_0z_1^{-1}$, nullo quando

$$z_1 = z_0$$

CAPITOLO 4 LA TRASFORMATTA Z

Piú in generale, utilizzando la formula (4.4) si verifica che la convoluzione di una sinusoide generalizzata con una sequenza qualsiasi può essere ottenuta come la cascata delle convoluzioni con le sequenze elementari di due campioni in cui la sequenza è stata decomposta.

Pertanto, il risultato della convoluzione di una sinusoide generalizzata campionata ed una sequenza qualsiasi $\{a_n\}$ è ancora la stessa sinusoide generalizzata moltiplicata per una fattore complesso che è definito come la **funzione di trasferimento della sequenza**; questo fattore è $A(z_1)$ e cioè il valore che assume la trasformata z della sequenza, assegnando all'argomento il valore z_1 della frequenza generalizzata della sinusoide $z_1 = \rho e^{j2\pi fT}$.

Abbiamo quindi trovato un significato fisico alla variabile z , a parte quello di corrispondere ad un anticipo nella trasformata z . Un qualsiasi

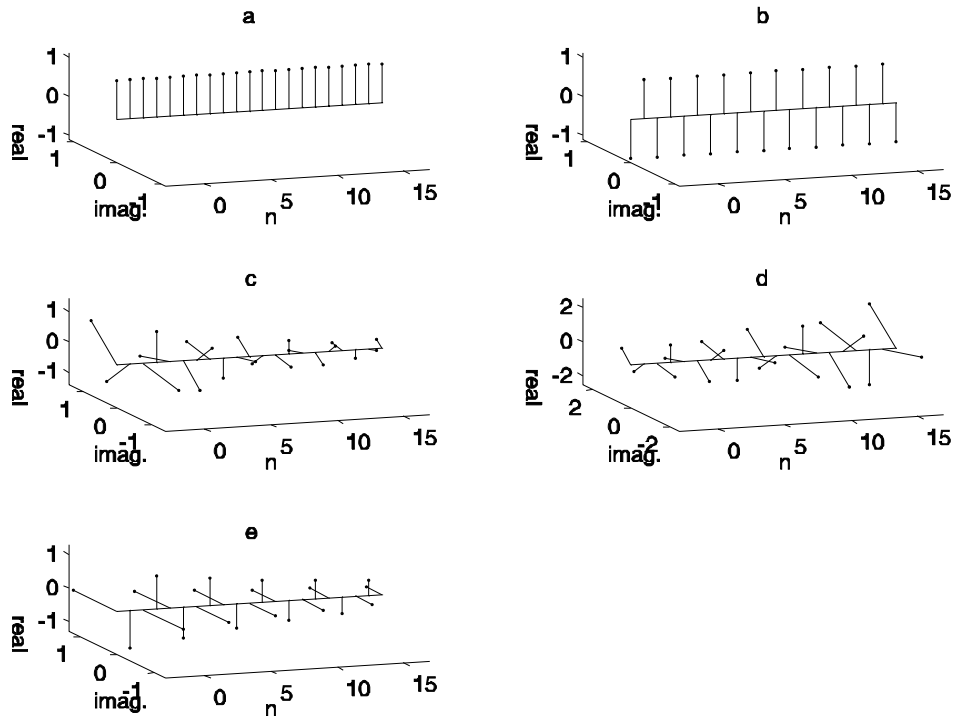


Figura 4.1: Esempi di sinusoidi generalizzate campionate $x_n = A\rho^n e^{jn\varphi}$; $\varphi = 2\pi f/f_c$: a) $\rho = 1$, $f/f_c = 0$; b) $\rho = 1$, $f/f_c = 0.5$; c) $\rho = 0.95$, $f/f_c = 2/5$; d) $\rho = 1.05$, $f/f_c = 2/5$; e) $\rho = 0.95$, $f/f_c = 1/4$.

valore di z può rappresentare l'argomento di una sinusoidale generalizzata; il valore $A(z_1)$ della trasformata z di una sequenza $\{a_n\}$ calcolata per il valore z_1 è il fattore per cui deve essere moltiplicata la sinusoidale generalizzata di argomento z_1 quando viene convoluta con la sequenza $\{a_n\}$.

Si comprende perché una sequenza di due campioni viene indicata come uno **zero**: infatti la radice (lo zero) del polinomio di primo grado corrispondente alla sua trasformata z è l'argomento della sinusoidale generalizzata che, convoluta con la sequenza elementare, viene annullata.

4.4 Il piano z

L'argomento delle sinusoidi generalizzate è rappresentabile come un punto di un piano complesso (figura 4.2) detto piano z ; la distanza dall'origine corrisponde a ρ ; l'angolo formato con l'asse delle ascisse è $\omega T = 2\pi fT$. In questo piano, viene tracciato un cerchio di raggio unitario, che è il luogo degli argomenti corrispondenti a sinusoidi campionate vere e proprie. La sinusoidale a frequenza 0, (il segnale costante) corrisponde al punto $z_1 = 1 + j \times 0$; la sinusoidale alla frequenza di Nyquist è invece corrispondente al punto $z_2 = -1 + j \times 0$.

4.4.1 Le sequenze nel piano z

Una qualsiasi sequenza di $N + 1$ campioni è rappresentabile nel piano z con la disposizione degli N zeri del polinomio corrispondente alla sua trasformata z ; compare anche una singolarità (un polo di ordine N nell'origine del piano z) che però non ha un vero e proprio significato fisico e nasce dalla scelta dell'esponente negativo per rappresentare i ritardi nella trasformata z .

Come già osservato, è immediato trovare la funzione di trasferimento corrispondente ad una sequenza qualsiasi; infatti, basta calcolare il valore della trasformata z della sequenza sul punto del cerchio unitario corrispondente alla frequenza desiderata. Procedendo per via geometrica, si vede che nel caso di una sequenza di due campioni, la funzione di trasferimento è (fig. 4.2):

$$A(z) = 1 - z_0 z^{-1} = \frac{(z - z_0)}{z} \quad (4.9)$$

CAPITOLO 4 LA TRASFORMATTA Z

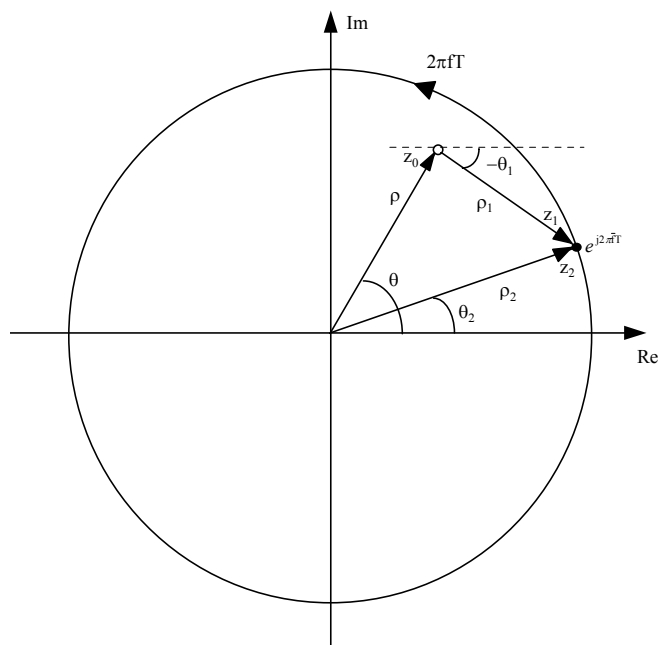


Figura 4.2: Calcolo della F.d.T. di una sequenza relativa ad uno zero per via geometrica

e quindi corrisponde al rapporto tra i due vettori complessi che congiungono il punto rappresentativo della sinusoide sul cerchio unitario e lo zero e l'origine del piano z rispettivamente. E' quindi possibile tracciare semplicemente la caratteristica di ampiezza corrispondente a zeri variamente posizionati nel piano z ed a sequenze più complesse; ovviamente, la funzione di trasferimento di una sequenza è il prodotto delle funzioni di trasferimento delle sequenze elementari così come la sequenza è convoluzione delle sequenze elementari. Sequenze reali hanno zeri reali o complessi coniugati, come ovvio visto che corrispondono a radici di polinomi a coefficienti reali.

4.4.2 Zeri a fase minima e fase massima

Si possono avere zeri sia sul cerchio unitario (ed allora annullano sinusoidi vere e proprie) che fuori (annullano sinusoidi generalizzate crescenti con

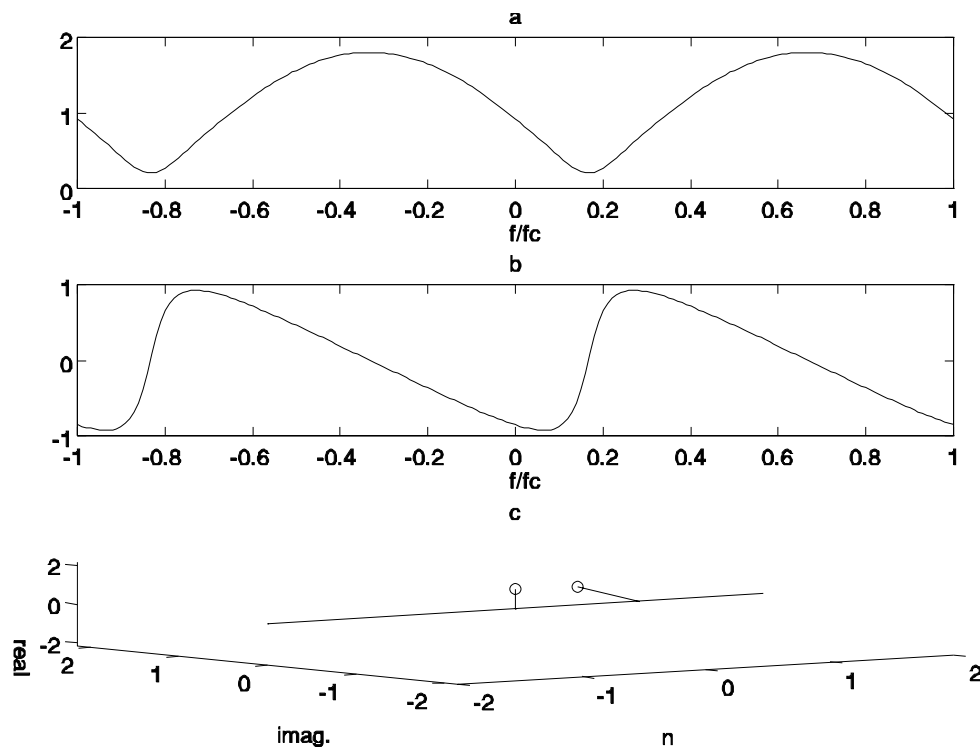


Figura 4.3: Caratteristiche di ampiezza e fase dello zero complesso della figura 4.2

l'indice) che infine dentro (annullano sinusoidi generalizzate decrescenti con l'indice).

Gli zeri posizionati dentro il cerchio unitario sono chiamati zeri a fase minima e zeri a fase massima quelli all'esterno (fig. 4.4). Il motivo si evince dalle caratteristiche di fase delle rispettive funzioni di trasferimento; per gli zeri a fase minima, quando il punto rappresentativo della frequenza della sinusoide fa un intero giro lungo il cerchio unitario, dal punto corrispondente alla frequenza zero fino a tornarvi, la fase delle funzione di trasferimento ritorna al valore iniziale; non così nel caso di zeri a fase massima, in cui la fase ritorna al valore iniziale a meno di 2π . Infatti in questo secondo caso, il vettore z fa un intero giro intorno all'origine mentre il vettore $z - z_0$ no.

Nel caso di zeri a fase massima pertanto vi sarà una discontinuità di

CAPITOLO 4 LA TRASFORMATTA Z

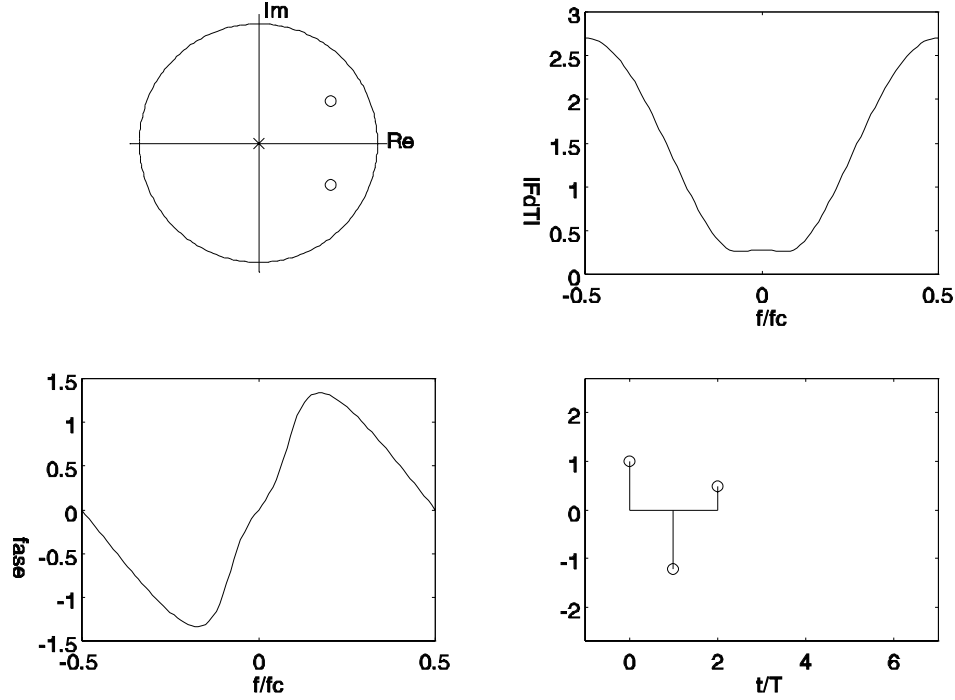


Figura 4.4: Caratteristiche di fase ed ampiezza di una F.d.T a fase minima

2π della fase della funzione di trasferimento (fig. 4.5).

E' interessante osservare che due zeri reciproci e coniugati (ossia due sequenze di 2 campioni le cui trasformate z sono caratterizzate da zeri $z_0, 1/z_0^*$ in posizione reciproca e coniugata) hanno eguale andamento della caratteristica di ampiezza (**ma solo sul cerchio unitario**) e diversa caratteristica di fase . Ponendo $z = e^{j\varphi}$

$$|A(z)|^2 = |1 - z_0 e^{-j\varphi}|^2 = 1 + |z_0|^2 - 2\text{Re}(z_0 e^{-j\varphi}); \quad (4.10)$$

$$|A_r(z)|^2 = \frac{|z_0^* - e^{-j\varphi}|^2}{|z_0|^2} = \frac{1 + |z_0|^2 - 2\text{Re}(z_0 e^{-j\varphi})}{|z_0|^2} = \frac{|A(z)|^2}{|z_0|^2} \quad (4.11)$$

Consideriamo una sequenza causale $\{a_n\}$; $a_n = 0$; $\forall n < 0$ la cui trasformata z è un polinomio $A(z)$ a coefficienti complessi . Indicheremo con $\tilde{A}(z)$ lo stesso polinomio $A(z)$ ma con i coefficienti coniugati e nell'ordine inverso; esso è trasformata z della sequenza $\{\tilde{a}_n\}$ che si ottiene

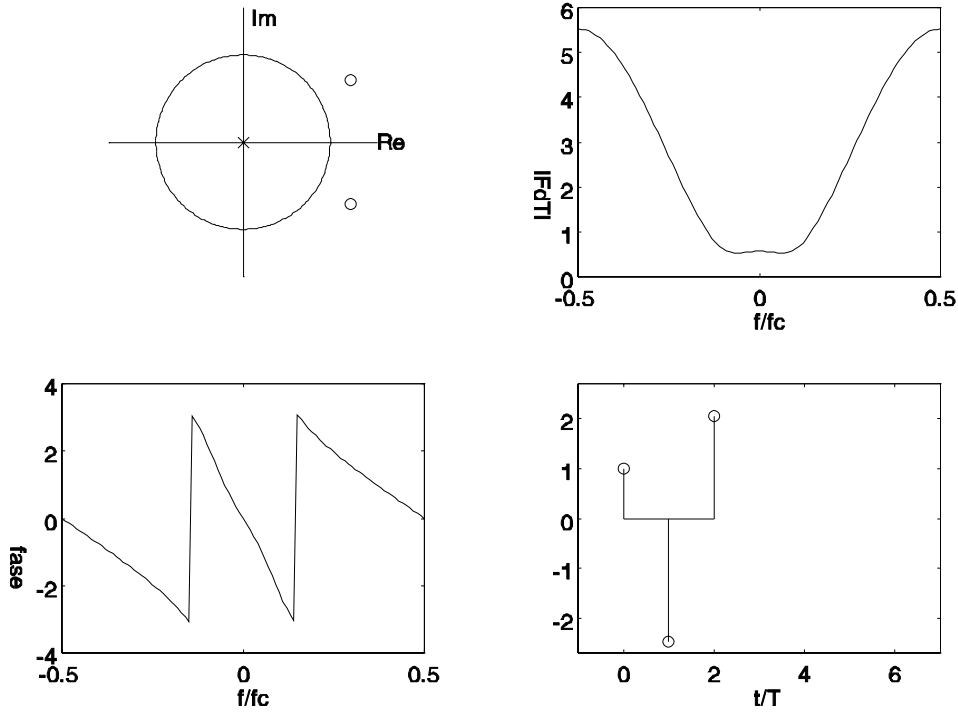


Figura 4.5: Caratteristiche di fase ed ampiezza di una F.d.T a fase massima

dalla sequenza $\{a_n\}$, coniugandola e ribaltandola rispetto all'istante 0 e ritardandola in modo da riavere una sequenza causale. Risulta che:

$$\tilde{A}(z) = z^{-N} A^*(1/z); \quad |A(z)|_{z=e^{j\omega T}} = |\tilde{A}(z)|_{z=e^{j\omega T}} \quad (4.12)$$

Bisogna rammentare che il campione di posto 0 della sequenza $\{\tilde{a}_n\}$ è $\tilde{a}_0 = a_N \neq 1$.

4.4.3 Sequenze a fase minima, massima o mista

Una sequenza $A(z)$ la cui trasformata z ha tutti gli zeri nel cerchio unitario è detta una sequenza a fase minima. Se tutti gli zeri sono fuori del cerchio unitario, la sequenza è detta a fase massima; ad esempio, $\tilde{A}(z)$ è a fase massima, se $A(z)$ è a fase minima. La sequenza è detta a fase mista nei casi in cui parte degli zeri sono dentro e parte sono fuori del cerchio.

CAPITOLO 4 LA TRASFORMATTA Z

Convolvere una sequenza $\{a_n\}$ di N campioni con la sequenza $\{\tilde{a}_n\}$ (quella che si ottiene ribaltando e ritardando la sequenza $\{a_n\}$ come visto al paragrafo precedente) si ottiene una sequenza lunga $2N - 1$ campioni e simmetrica rispetto al campione di posto $N - 1$; ciò si può verificare con una sequenza corrispondente ad uno zero. Nel dominio delle trasformate si passa da $A(z)$ ad $\tilde{A}(z)$ ribaltando tutti gli $N - 1$ zeri della trasformata nella posizione reciproca rispetto al cerchio unitario (e coniugata se si tratta di zero complesso). Pertanto la caratteristica di fase di $A(z)\tilde{A}(z)$ sarà quella corrispondente ad un ritardo pari a $(N - 1)T$ e cioè:

$$\angle[A(z)\tilde{A}(z)]|_{z=e^{j\omega T}} = -(N - 1)\omega T \quad (4.13)$$

Quindi la caratteristica di fase della trasformata z di una sequenza a fase massima si deduce facilmente da quella della sequenza a fase minima corrispondente. Ribaltando gli zeri rispetto al cerchio unitario, non cambia la caratteristica di ampiezza, come si è visto, ma si ribalta (e si ritarda) la sequenza nel tempo.

4.4.4 Trasformata di Fourier e trasformata z

E' evidente che la trasformata z di una sequenza di N campioni e cioè:

$$X(z) = \sum_{n=0}^{N-1} x_n z^{-n}$$

e la trasformata di Fourier della stessa sequenza:

$$X_F(f) = \sum_{n=0}^{N-1} x_n e^{-j2\pi f n T}$$

sono legate tra loro ed in particolare:

$$X_F(f) = X(z = e^{j2\pi f n T})$$

e cioè i valori assunti dalla trasformata z lungo il cerchio unitario sono appunto i valori della trasformata di Fourier della sequenza stessa. I valori della trasformata di Fourier discreta si otterranno poi campionando lungo il cerchio unitario, ad angoli crescenti multipli di $2\pi/N$, i valori della trasformata z .

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi k \Delta f n T} = X(z = e^{j2\pi n k \Delta f T}) = X(z = e^{j\frac{2\pi n k}{N}})$$

4.5 Poli

E' interessante studiare sequenze aventi trasformata z inversa di quella corrispondenti agli zeri, e cioè del tipo

$$C(z) = \frac{1}{B(z)} = \frac{1}{1 - z_0 z^{-1}}; \quad B(z) = 1 - z_0 z^{-1}; \quad (4.14)$$

Convolvere una sequenza qualsiasi prima con la sequenza di trasformata $B(z)$ (lo zero) e poi con quella di trasformata $C(z)$ è come non fare nulla o fare una convoluzione con un impulso nell'origine, o moltiplicarne la trasformata per una trasformata z unitaria. Abbiamo quindi introdotto una sequenza $\{c_n\}$ [l'antitrasformata di $C(z)$] "inversa" alla sequenza $\{b_n\} = \{1, -z_0\}$ corrispondente allo zero $B(z)$; infatti convolvere per $\{c_n\}$ *compensa* le modifiche apportate alla funzione di trasferimento dalla convoluzione con la sequenza $\{b_n\}$ [corrispondente a $B(z)$]. Le sequenze "inverse" delle sequenze elementari di due campioni (zeri) sono indicate con il nome di **poli**.

Poiché però la funzione $C(z)$ non è un polinomio di grado finito, la sequenza $\{c_n\}$ non è univocamente determinata.

Infatti vedremo ora che l'operazione di antitrasformazione di $C(z)$ non è immediata; questa operazione è univoca solo se la funzione è un polinomio di ordine finito nella variabile z^{-1} o magari z se vi sono anticipi invece che ritardi; in questo caso, basta la diretta identificazione del coefficiente dei termini di grado i -esimo per trovare il termine i -esimo della sequenza.

Il problema è sostanziale e non solo formale: infatti esistono infinite sequenze, di durata illimitata, che convolute con una sequenza di due campioni danno come risultato l'impulso unitario nell'origine e cioè hanno una trasformata z inversa di quella di uno zero. La verifica è semplice: si consideri la sequenza di due campioni $b_0 = 1$; $b_1 = -z_0$ e si cerchi la sequenza $\{c_n\}$ tale che:

$$\{b_n\} * \{c_n\} = \delta(n); \quad (4.15)$$

$$\begin{aligned} \delta(n) &= 1; & n &= 0; & \delta(n) &= 0; & n &\neq 0 \\ c_n - z_0 c_{n-1} &= \delta(n) \end{aligned} \quad (4.16)$$

Questa è un'equazione alle differenze finite che si può risolvere per sostituzione. Supponiamo ora che z_0 sia uno zero a minima fase e cioè $|z_0| < 1$.

CAPITOLO 4 LA TRASFORMATTA Z

Poniamo $c_{-1} = 0$ e risolviamo per indici diversi da -1 ; risulta:

$$c_n = 0; \quad n < 0; \quad c_n = z_0^n; \quad n \geq 0 \quad (4.17)$$

Peraltro, se avessimo posto $c_1 = 0$ avremmo trovato:

$$c_n = 0; \quad n \geq 0; \quad c_n = -z_0^n; \quad n < 0 \quad (4.18)$$

Quindi abbiamo trovato per la sequenza $\{c_n\}$ almeno due diverse soluzioni, ed anzi infinite perchè infinite combinazioni lineari di queste due soluzioni sono ancora soluzioni del problema. Nel caso quindi di uno zero a minima fase, la sequenza causale (quella nulla per $n < 0$) è convergente e quella anticausale è divergente. Si osserva invece che con $|z_0| > 1$ e cioè con uno zero a massima fase, è vero il contrario e cioè la sequenza anticausale è convergente e quella causale è divergente. Allo stesso risultato si arriva dal dominio delle trasformate utilizzando la serie geometrica; è possibile scrivere che:

$$\frac{1}{1 - z_0 z^{-1}} = 1 + (z_0 z^{-1}) + (z_0 z^{-1})^2 + (z_0 z^{-1})^3 + \dots \rightarrow 1, z_0, z_0^2, z_0^3, \dots \quad (4.19)$$

sequenza causale e convergente se lo zero è a minima fase; ma anche

$$\begin{aligned} \frac{1}{1 - z_0 z^{-1}} &= -\frac{1}{z_0} \frac{z}{1 - z_0^{-1} z} = \\ &= -\frac{z}{z_0} \left(1 + \frac{z}{z_0} + \left(\frac{z}{z_0}\right)^2 + \left(\frac{z}{z_0}\right)^3 + \dots \right) \end{aligned} \quad (4.20)$$

sequenza anticausale e convergente se lo zero è a massima fase.

Da un altro punto di vista, si può osservare che lo zero a fase minima o a fase massima annulla una sinusoide generalizzata che decresce o cresce con l'indice. Si divida ora la sinusoide generalizzata in due sottosequenze in corrispondenza dell'origine, e cioè $\{z_0^n; n \geq 0\}$ e $\{z_0^n; n < 0\}$; il risultato della convoluzione di ognuno dei due tronconi con lo zero sarà sempre nullo tranne che nell'origine. Ancora, poiché sommando i due tronconi si ritrova la sinusoide generalizzata, si conclude che i due impulsi che si trovano convolvendo i due tronconi destro o sinistro sono di segno opposto (fig. 4.6).

La sequenza corrispondente ad un polo può essere sia l'uno che l'altro dei due tronconi (a meno del segno); più in generale una qualsiasi combinazione lineare dei due tronconi; in un modo o nell'altro si trova una

4.5 POLI

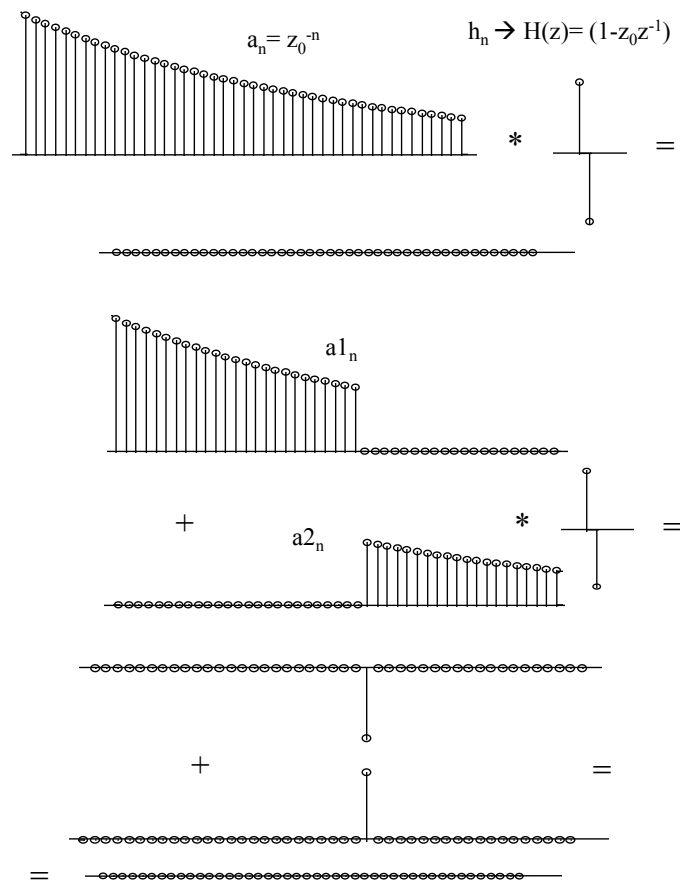


Figura 4.6: Rappresentazione delle sequenze "inverse" degli zeri

CAPITOLO 4 LA TRASFORMATTA Z

sequenza che convoluta con uno zero dà un impulso; è importante osservare che sempre solo uno dei due tronconi è convergente; l'altro tende ad infinito.

Da quanto osservato, si conclude che vi sono infiniti modi per *compensare* l'effetto di uno zero. Quindi, ad un polo corrispondono infinite sequenze; il grado di libertà che ne consegue può essere utilizzato in vario modo: ad esempio imponendo che la sequenza inversa assuma un valore assegnato ad un certo istante. **Pertanto, tra tutte le sequenze inverse possibili, una sola sarà convergente**, ed in particolare sarà causale nel caso di poli a fase minima (inversi di zeri a fase minima) e anticausale nel caso di poli a fase massima.

La sequenza inversa è unica qualora il problema esige la convergenza (valore assegnato pari a zero a tempo infinito), o qualora esige la causalità (valore assegnato pari a zero per $n = -1$), o infine se si impone che la risposta all'impulso abbia ad un tempo qualsiasi un valore qualsiasi.

4.6 Equazioni alle differenze finite

Si può fare la convoluzione di una sequenza causale corrispondente ad un polo (di lunghezza infinita), con un segnale numerico qualsiasi al costo molto basso di una sola moltiplicazione ed una addizione per campione, come se si trattasse di uno zero. Infatti sia

$$Y(z) = \frac{X(z)}{1 - z_0 z^{-1}} \quad (4.21)$$

Antitrasformando, se ne ricava l'equazione alle differenze finite

$$y_n = x_n + z_0 y_{n-1} \quad (4.22)$$

da cui si calcola la sequenza in uscita e pertanto il risultato della convoluzione cercato.

La condizione di causalità è applicata implicitamente se si fa il calcolo per indici crescenti iniziando p.e. da y_0 assegnando ad y_{-1} un valore qualsiasi; il risultato sarà convergente se il polo è a minima fase.

Se invece il polo fosse a massima fase, per avere un risultato ancora convergente, si dovrebbe risolvere all'incontrario l'equazione di cui sopra andando verso indici decrescenti, assegnando cioè ad y_1 un valore

4.6 EQUAZIONI ALLE DIFFERENZE FINITE

qualsiasi e quindi risolvendo a ritroso:

$$y_{n-1} = \frac{1}{z_0}(y_n - x_n) \quad (4.23)$$

Ovviamente in questo secondo caso, si ha convergenza per $z_0 > 1$.

Sia ora in generale:

$$Y(z) = X(z) \frac{A(z)}{B(z)} \quad (4.24)$$

dove $X(z)$ è la trasformata della sequenza $\{x_n\}$ e $A(z)$, $B(z)$ sono polinomi qualsiasi.

Un sistema di questo tipo ha poli posizionati sia fuori sia dentro il cerchio unitario e anche zeri. In questo caso è necessario prima capire se è la soluzione convergente quella cercata: in questo caso, si deve:

trovare tutte le radici del denominatore (poli) e decomporre il denominatore $B(z)$ nel prodotto di due polinomi, l'uno $B_{max}(z)$ con le radici tutte fuori dal cerchio unitario (poli a fase massima), l'altro $B_{min}(z)$ con le radici tutte dentro il cerchio unitario (poli a fase minima);

$$Y(z) = X(z) \frac{A(z)}{B_{min}(z)B_{max}(z)} = X(z) \frac{A(z)}{B_{min}(z)} \times \frac{1}{B_{max}(z)} \quad (4.25)$$

convolvere con la parte causale e convergente della risposta all'impulso del sistema, e cioè quella corrispondente ai poli a fase minima e agli zeri; ciò si attua utilizzando l'ingresso $\{x_n\}$ come termine forzante dell'equazione alle differenze finite corrispondente a questa parte del sistema. In sostanza si considera il sistema:

$$Y_1(z) = X(z) \frac{A(z)}{B_{min}(z)} \quad (4.26)$$

Sia ora la sequenza $\{y_{1,n}\}$ l'uscita della parte causale del sistema.

infine, si deve convolvere la sequenza $\{y_{1,n}\}$ con la risposta all'impulso corrispondente ai poli a fase massima, che è convergente e quindi anticausale.

$$Y(z) = \frac{Y_1(z)}{B_{max}(z)} \quad (4.27)$$

CAPITOLO 4 LA TRASFORMATTA Z

Per fare ciò in modo semplice ed a basso costo computazionale si deve invertire l'asse temporale. Formalmente, si può scrivere:

$$Y(1/z) = \frac{Y_1(1/z)}{B_{max}(1/z)} \quad Y_1(1/z) \leftrightarrow \{y_{1,-n}\} \quad (4.28)$$

Infatti la sequenza $Y_1(1/z)$ è la trasformata della sequenza $\{y_{1,n}\}$, dopo che sia stata ribaltata nei tempi.

A sua volta, il polinomio $B_{max}(1/z)$ ha le radici reciproche di quelle del polinomio $B_{max}(z)$ e che quindi sono all'interno del cerchio unitario; l'equazione alle differenze finite descritta dalla precedente equazione è quindi corrispondente ad un sistema causale e convergente che può essere risolto come nel caso precedente. Infine, la sequenza $Y(1/z)$ deve venire ancora ribaltata nel tempo per avere il risultato finale e cioè la sequenza $Y(z)$.

Si osservi che per ribaltare l'asse temporale di una sequenza, ne devono essere disponibili tutti i campioni, presenti e futuri. La cosa non può essere fatta in tempo reale con sistemi fisici che si evolvono nel tempo (sistemi causali).

Tuttavia, anche nel caso in cui non si hanno a disposizione tutti i dati, ma si deve operare con un ritardo finito, l'equazione con i poli fuori del cerchio unitario può essere risolta approssimativamente con la tecnica discussa nel paragrafo sulle sequenze geometriche troncate.

4.7 Filtri passa banda

I poli si prestano per realizzare a basso costo computazionale funzioni di trasferimento a banda molto stretta; infatti le sequenze rappresentate da poli sono lunghe, qualora il polo sia prossimo al cerchio unitario. Ad esempio, supponiamo che il polo sia in $z = (1 - \varepsilon)e^{j\omega_0 T}$ con $\varepsilon \ll 1$ (figg. 4.7; 4.8). Usando la tecnica vettoriale per il calcolo della funzione di trasferimento, si osserva che essa ha un valore massimo di $1/\varepsilon$ per $\omega = \omega_0$ e cala di 3 dB (un fattore pari a $\sqrt{2}$) fuori dell'intervallo $\Delta\omega T \simeq 2\varepsilon$.

Basta scegliere ε piccolo, e risultano filtri a banda molto stretta. Come vedremo poi, discutendo della quantizzazione dei coefficienti, questi filtri sono molto sensibili alle variazioni dei coefficienti causate dalla loro rappresentazione con un limitato numero di cifre binarie.

4.7 FILTRI PASSA BANDA

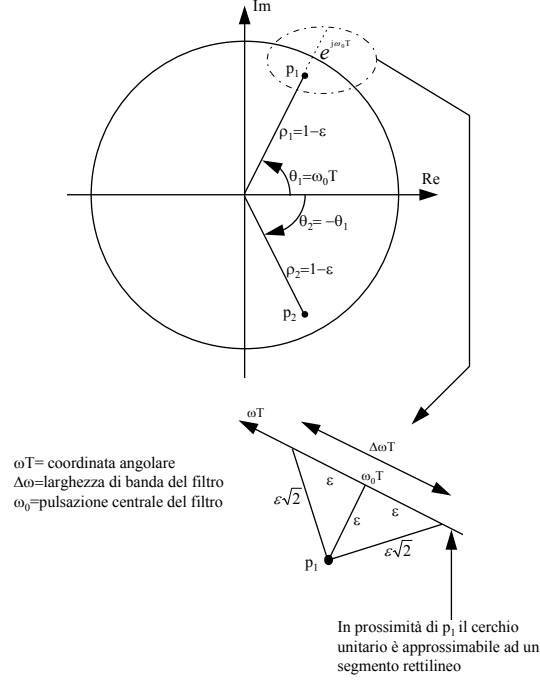


Figura 4.7: Tecnica vettoriale per il progetto di un filtro passa-banda realizzato con due poli complessi coniugati p_1 e p_2 molto vicini al cerchio di raggio unitario

Di notevole interesse è la realizzazione di filtri arresta banda fatta con una coppia di un polo posto presso il cerchio unitario ed uno zero sul cerchio unitario, isofrequenziale. La funzione di trasferimento si annulla alla frequenza ω_0 ma ritorna 3 dB al di sotto del valore unitario appena fuori dell'intervallo $\Delta\omega T \simeq 2\varepsilon$ (fig. 4.9).

4.7.1 Sfasatori puri

Altro caso interessante è quello dello sfasatore puro; qui abbiamo un polo ed uno zero messi in posizione reciproca coniugata rispetto al cerchio unitario; pertanto il modulo della funzione di trasferimento è costante, ma non così la fase che avrà una rapida variazione all'interno del solito intervallo di frequenza $\Delta\omega T \simeq 2\varepsilon$ (fig. 4.10). Si osservi che trasformate

CAPITOLO 4 LA TRASFORMATTA Z

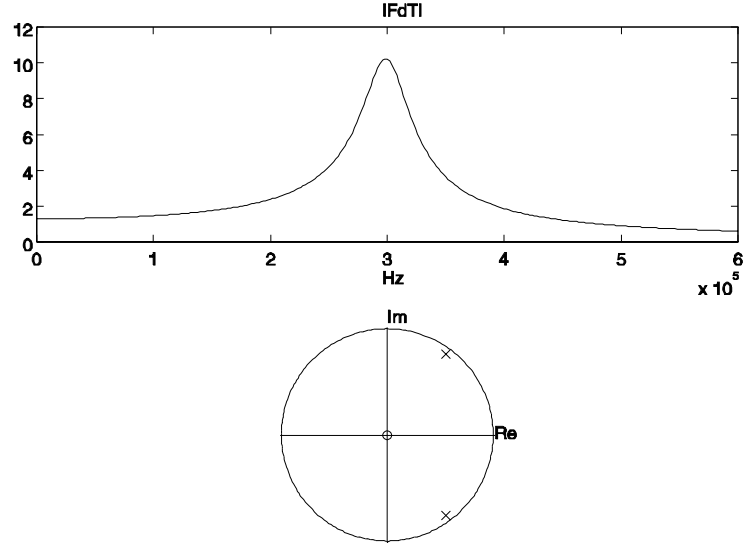


Figura 4.8: Caratteristiche di un filtro passa-banda avente una banda passante di $40kHz$ attorno alla frequenza di $300kHz$; $f_c = 2MHz$

z del tipo

$$\frac{\tilde{A}(z)}{A(z)} = Q(z); \quad |Q(z)|_{z=e^{j\omega T}} = 1 \quad (4.29)$$

hanno modulo sempre unitario (sul cerchio unitario, ovviamente) e rappresentano sfasatori puri, causali e convergenti se il polinomio $A(z)$ è a minima fase.

La caratteristica di fase di uno sfasatore puro causale e convergente è sempre monotona decrescente; per dimostrarlo basta considerare il caso di un sistema con un solo zero ed un solo polo reali; casi di maggiore complessità si riporteranno a cascata di sistemi come questo, centrati a differenti frequenze e quindi ruotati intorno al cerchio unitario in vario modo:

$$H(z) = \frac{-a + z^{-1}}{1 - az^{-1}} = \frac{\widetilde{A(z)}}{A(z)};$$

La caratteristica di fase è:

$$\angle H(z) = \phi(\omega) = -\angle A(z) + \angle \widetilde{A(z)} = -\omega - 2\angle A(z)$$

$$\phi(\omega) = -\omega T - 2 \tan^{-1} \frac{a \sin \omega T}{1 - a \cos \omega T}$$

4.7 FILTRI PASSA BANDA

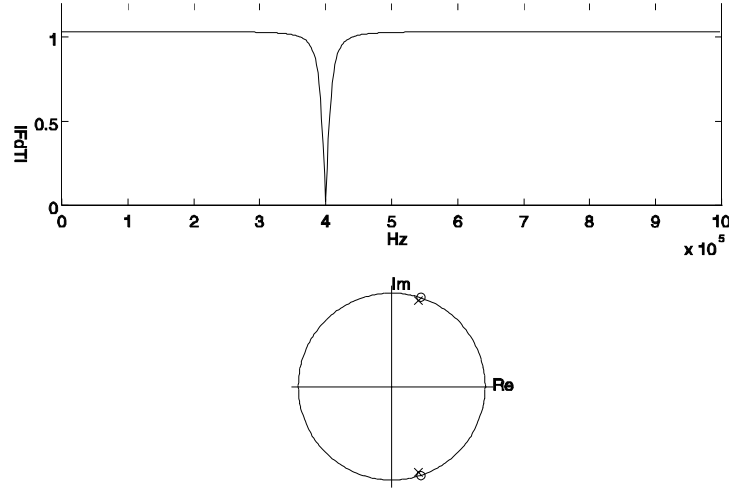


Figura 4.9: Caratteristiche di un filtro arresta banda avente una banda soppressa di $20kHz$ attorno alla frequenza di $400kHz$; $f_c = 2MHz$

Il ritardo di gruppo $\tau(\omega)$, ritardo di un segnale passabanda centrato a pulsazione ω , ottenuto differenziando la fase rispetto a ω e cambiando segno, è necessariamente positivo per i sistemi causali. Infatti:

$$-\frac{d\phi}{d\omega} = \tau(\omega) = T \frac{1 - a^2}{1 + a^2 - 2a \cos \omega T} > 0 \text{ se } a = 1 - \varepsilon < 1$$

Pertanto nel caso in cui $a = 1 - \varepsilon < 1$, (sistema causale convergente) la fase è sempre decrescente. Si osservi che il grafico che dà il ritardo di gruppo in funzione della frequenza è molto semplice. Inoltre:

$$\tau_{\max} = \frac{(1 + a)T}{(1 - a)} \sim \frac{2T}{\varepsilon};$$

Pertanto, il ritardo massimo può crescere quanto si desidera: p.e., teoricamente, uno sfasatore puro potrebbe essere utilizzato al posto di una memoria passabanda di durata $2T/\varepsilon$ e banda $\Delta f = \varepsilon/\pi T$. Invece, con $a > 1$, la fase è monotonamente crescente e lo sfasatore puro, per essere stabile, deve essere anticausale.

CAPITOLO 4 LA TRASFORMATA Z

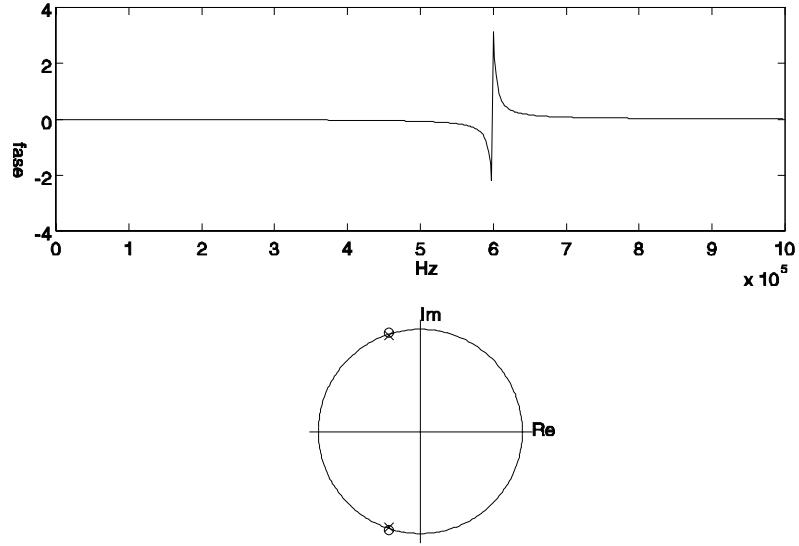


Figura 4.10: Caratteristiche di uno sfasatore puro con sfasamento intorno ai $600kHz$ in un intervallo di $50kHz$; $f_c = 2MHz$

4.7.2 Oscillatori, divani, splines

Anzitutto consideriamo un esempio di equazione alle differenze finite che si riferisce ad un sistema che si evolve nel tempo e pertanto soggetto alla relazione di causalità. Questo è il caso del sistema con due poli in equazione (4.30); il sistema viene qui indicato come oscillatore numerico poichè, per opportuni valori dei parametri, la sua risposta all'impulso può essere una sinusoide campionata; al sistema corrisponde la seguente equazione alle differenze finite:

$$Y(z) = \frac{X(z)}{1 - 2z^{-1}\rho \cos \phi + \rho^2 z^{-2}}; \quad \phi = \omega_0 T \quad (4.30)$$

corrispondente ad un sistema causale perché riferita ad un oggetto fisico che opera nel tempo e non risponde prima della sua eccitazione. Qualora $X(z) = 1$ la risposta del sistema si riduce alla combinazione delle risposte ai due poli complessi coniugati (posti vicino o sul cerchio unitario se $\rho \leq 1$) e quindi a due sequenze causali complesse coniugate la cui somma è:

$$y_n = \rho^n \frac{\sin[(n+1)\phi]}{\sin \phi}; \quad n \geq 0; \quad (4.31)$$

4.7 FILTRI PASSA BANDA

Si osservi come il primo campione è sempre 1, e al decrescere di ϕ l'oscillazione ha ampiezze crescenti. Se poi $\rho \geq 1$ l'oscillazione diverge ed il sistema andrà incontro a fenomeni di saturazione. Se vogliamo che l'uscita dell'oscillatore sia una cosinusoide esponenzialmente decrescente, a fase γ preassegnata, si dovrà aggiungere uno zero per ottenere lo sfasamento desiderato alla pulsazione della sinusoide; p.e. ponendo:

$$Y_1(z) = \frac{1}{2} \left(\frac{e^{j\gamma}}{1 - \rho e^{j\phi} z^{-1}} + \frac{e^{-j\gamma}}{1 - \rho e^{-j\phi} z^{-1}} \right) = \frac{\cos \gamma - z^{-1} \rho \cos(\gamma - \phi)}{1 - 2z^{-1} \rho \cos \phi + \rho^2 z^{-2}}$$

La risposta nei tempi è ottenuta sviluppando le due frazioni in serie geometriche e poi sommando i valori coniugati:

$$y_{1n} = \rho^n \cos(n\phi + \gamma)$$

e si ritrova il caso precedente (a meno di un fattore) ponendo $\gamma = \phi - \pi/2$.

Vediamo ora un'applicazione ad un sistema non causale: un divano a molle. Si supponga di avere un sistema meccanico formato da nodi collegati da molle; il nodo i -esimo è collegato al suolo da una molla verticale di rigidità (rapporto tra forza e deformazione) k_1 ed è collegato ai nodi adiacenti $i - 1$, $i + 1$ da molle di rigidità k_2 ; ogni nodo è poi soggetto alla forza verticale f_i . L'equazione di equilibrio del nodo i -esimo è, indicando con x_i lo spostamento elastico verticale del nodo i :

$$k_2(x_i - x_{i-1}) + k_2(x_i - x_{i+1}) + k_1 x_i = f_i \quad (4.32)$$

Passando alle trasformate z delle sequenze degli spostamenti $X(z)$ e delle forze $F(z)$ si ha:

$$X(z) [(k_1 + 2k_2) - (z + z^{-1})k_2] = F(z) \quad (4.33)$$

e quindi la risposta all'impulso ($F(z) = f_0$ e cioè una sola forza posizionata sul campione di posto 0) è:

$$X(z) = \frac{f_0}{(k_1 + 2k_2) - (z + z^{-1})k_2} \quad (4.34)$$

La trasformata z della risposta è di nuovo una funzione di trasferimento simile a quella del caso precedente: inoltre, i poli del sistema sono reali, positivi e posti in posizioni reciproche rispetto al cerchio unitario, in

CAPITOLO 4 LA TRASFORMATA Z

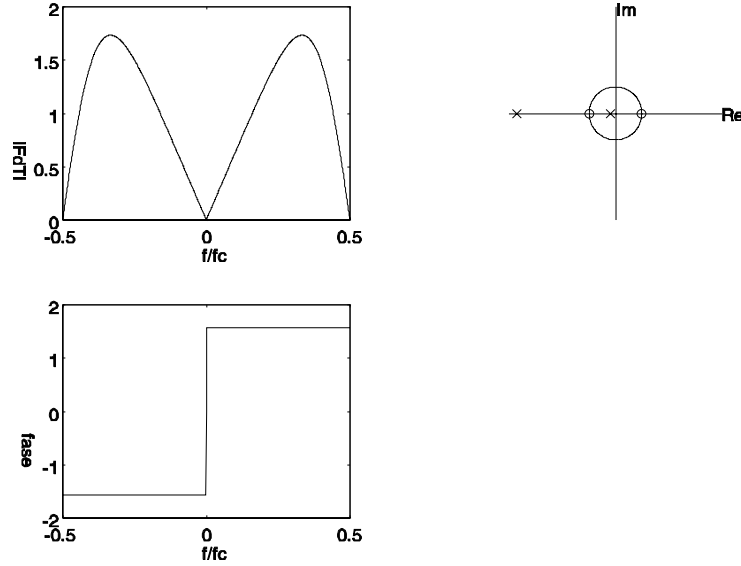


Figura 4.11: Caratteristiche relative alla sequenza delle derivate della risposta impulsiva dello spline in corrispondenza dei campioni

quanto il prodotto delle radici del denominatore è 1 ed il discriminante è positivo.

La risposta da attendersi non è però quella causale, come nel caso precedente, e cioè non ci attendiamo spostamento nullo per indici inferiori a 0, e vogliamo una risposta convergente; (*sappiamo che il divano non si deforma solo a destra di chi si siede e non esplode*). Quindi sceglieremo la soluzione causale per il polo interno al cerchio e quella anticausale per il polo esterno; sostanzialmente cerchiamo una soluzione che sia simmetrica negli spazi.

Posto:

$$q = \frac{2k_2 + k_1 - \sqrt{k_1^2 + 4k_1k_2}}{2k_2}; \quad 0 \leq q \leq 1 \quad (4.35)$$

$$X(z) = -f_0 \frac{1 + q^2}{(k_1 + 2k_2)(1 - q^2)} \left(\frac{1}{1 - qz} + \frac{1}{1 - qz^{-1}} - 1 \right) \quad (4.36)$$

avremo:

$$x_n = -f_0 \frac{1 + q^2}{(k_1 + 2k_2)(1 - q^2)} q^{|n|} \quad (4.37)$$

4.7 FILTRI PASSA BANDA

In sostanza, se le molle orizzontali sono molto flessibili, il divano si infossa solo localmente ($q \simeq 0$); se sono rigide, il divano si abbassa quasi senza deformarsi ($q \leq 1$).

Come ultimo caso consideriamo l'equazione alle differenze finite che, per una bacchetta elastica, lega le pendenze ai nodi y'_i e gli spostamenti dei nodi y_i ; con la solita convenzione per le trasformate z rispettive, si ha, ricordando quanto osservato nel capitolo II:

$$(z) = 3Y(z) \frac{z - z^{-1}}{z + 4 + z^{-1}} \quad (4.38)$$

I poli del sistema sono reali e negativi in $-2 + \sqrt{3}$, $-2 - \sqrt{3}$ dentro e fuori il cerchio unitario in posizioni reciproche (infatti il prodotto delle radici del denominatore é unitario) (fig. 4.11).

Dobbiamo ancora separare il polo dentro da quello fuori del cerchio e scrivere:

$$Y'(z) = Y(z) \left(\frac{1}{1 + .268z} - \frac{1}{1 + .268z^{-1}} \right) \quad (4.39)$$

nel caso di $Y(z) = 1$ (ingresso impulsivo, solo il campione di posto 0 è non nullo) allora $Y'(z)$ é differenza di due sequenze, l'una causale e l'altra anticausale, simmetriche.

$$y'_n = -(-.268)^{|n|} \text{sign}(n); \quad y'_0 = 0 \quad (4.40)$$

Infatti la soluzione da scegliere in questo caso é ancora quella convergente; una bacchetta elastica la cui deflessione è nulla ad intervalli regolari, se abbassata o sollevata in corrispondenza del campione centrale, convergerà a zero con spostamenti a segni alterni; le derivate nei nodi sono anch'esse decrescenti esponenzialmente a segni alterni.

CAPITOLO 4 LA TRASFORMATTA Z

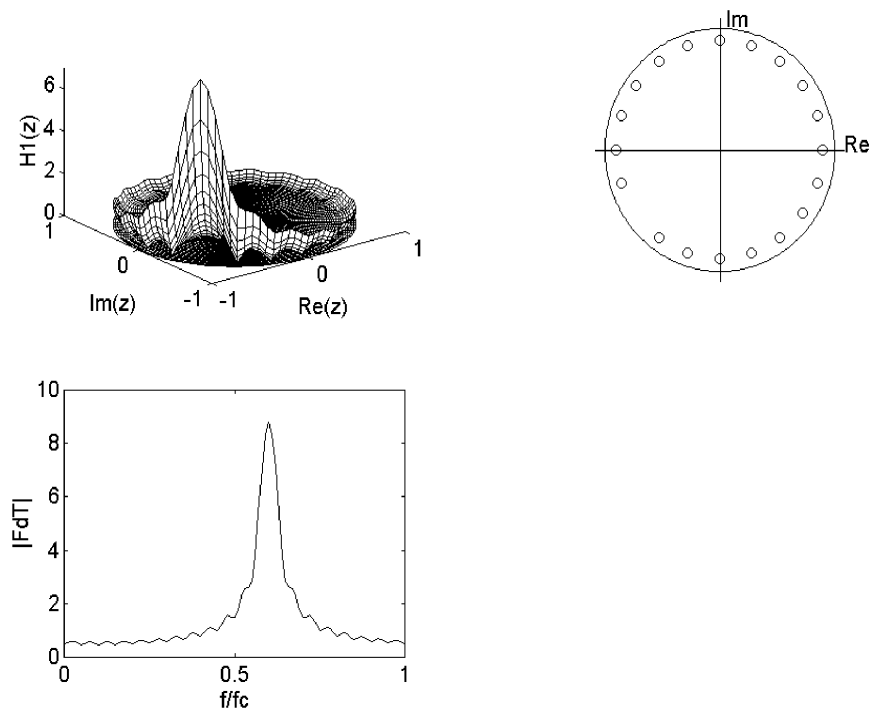


Figura 4.12: Caratteristiche della FdT $H_1(z)$ caratterizzata da $N = 19$ zeri

Capitolo 5

Filtri numerici

5.1 Introduzione

In questo capitolo discuteremo delle tecniche di progetto di filtri numerici: essi possono essere di due categorie e cioè FIR (Finite Impulse Response) o IIR (Infinite Impulse response) a seconda che la loro trasformata z consista di soli zeri o di poli e zeri. Una buona progettazione di un filtro numerico deve tendere ad ottenere il massimo risultato e cioè la massima aderenza alle specifiche progettuali; queste sono assegnate nel dominio delle frequenze mediante una maschera all'interno della quale deve essere contenuta la funzione di trasferimento del filtro sintetizzato (fig. 5.1). Inoltre, si deve minimizzare il costo e cioè minimizzare il numero di operazioni per campione; come vedremo ciò non corrisponde necessariamente al minor numero di moltiplicazioni; strutture di filtri più complesse possono essere convenienti se usano moltiplicazioni per numeri con pochi bit diversi da zero (potenze di due, ad esempio) o in genere se sono meno sensibili alla quantizzazione dei coefficienti e quindi operano correttamente con moltiplicazioni meno costose.

Tuttavia, per comprendere le basi delle tecniche di sintesi dei filtri FIR é opportuno discutere anzitutto delle tecniche di finestramento (windowing) e cioè gli effetti del troncamento nel tempo delle risposte all'impulso desiderate con finestre (rastremazioni a zero) di vario tipo. **Per semplicità l'analisi degli effetti del finestramento verrà fatta per segnali tempo continui, invece che campionati.** Questo semplifica molto le formule e le rende più leggibili.

CAPITOLO 5 FILTRI NUMERICI

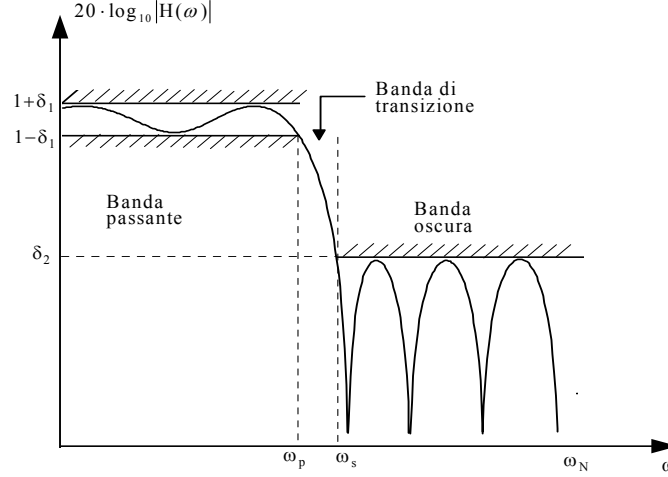


Figura 5.1: Maschera di un filtro passa-basso

5.2 Finestre rettangolari e triangolari

Indichiamo con $D(\omega)$ la risposta all'impulso desiderata per il filtro che si deve sintetizzare. La antitrasformata di Fourier verrà indicata con $d(t)$ ed in generale sarà illimitata nei tempi. Vogliamo vedere cosa accade troncando il segnale $d(t)$ con una finestra rettangolare di lunghezza NT . Nel caso discreto, considereremo la sequenza troncata dopo N campioni (fig. 5.2).

$$d'(t) = d(t) \times w_R(t) \quad (5.1)$$

$$w_R(t) = 1; |t| \leq \frac{NT}{2}; \quad w_R(t) = 0; |t| \geq \frac{NT}{2}$$

Moltiplicare nei tempi equivale a convolvere nelle frequenze e cioè a considerare la

$$D'(\omega) = D(\omega) * W_R(\omega); \quad W_R(\omega) = NT \operatorname{sinc}(fNT) \quad (5.2)$$

La $W_R(\omega)$ tende ad un impulso ideale se N diventa molto grande e questo corrisponde all'ovvia considerazione che finestrare con finestre lunghe

5.2 FINESTRE RETTANGOLARI E TRIANGOLARI

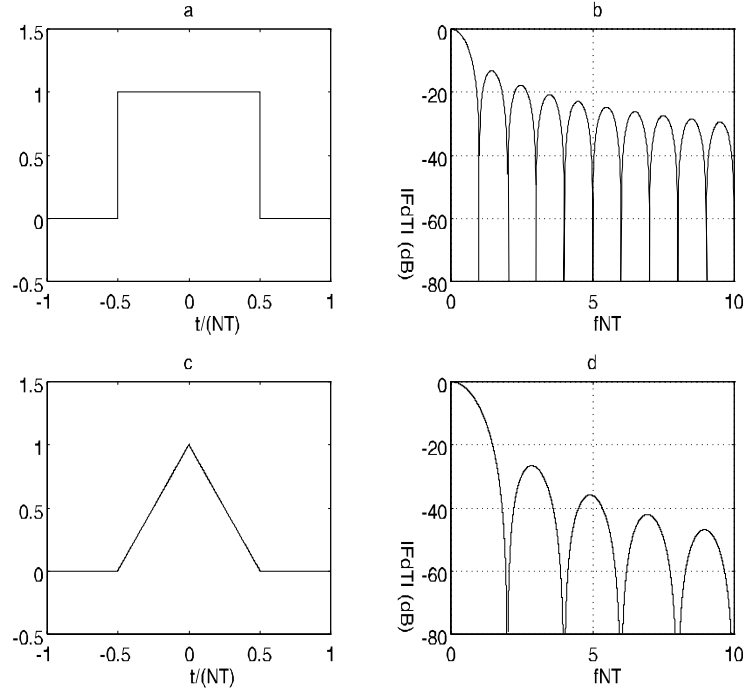


Figura 5.2: a) Finestra rettangolare b) Trasformata di Fourier della finestra rettangolare in dB c) Finestra triangolare d) Trasformata di Fourier della finestra triangolare in dB

porta a piccole distorsioni. Quello che interessa però sono appunto le distorsioni della risposta all'impulso dovute al finestramento: esse sono di due tipi diversi: quelle dovute alla convoluzione per il **lobo principale** del sinc che ha i primi zeri in $f = \pm 1/NT$ e quelle dovute alla convoluzione per i **lobi laterali**, che come è noto decrescono come $1/\Delta f NT$ dove con Δf abbiamo indicato l'intervallo in frequenza tra la posizione del lobo principale e quella del lobo secondario considerato. La convoluzione per il lobo principale smussa l'andamento in frequenza del filtro, e lo rende meno selettivo. I fronti di salita e di discesa, nel dominio della frequenza, non potranno essere più brevi di $\simeq 2/NT$; questa è una diretta conseguenza delle limitazione della lunghezza del filtro.

Il brusco troncamento della risposta $d(t)$ porta alla creazione di componenti ad alta frequenza che sarebbero altrimenti state assenti dalla risposta desiderata; questo fatto viene osservato, nel dominio delle fre-

CAPITOLO 5 FILTRI NUMERICI

quenze, con la formazione di componenti ad alta frequenza dovute alla convoluzione con i lobi laterali della finestra. Siamo quindi di fronte a due fenomeni diversi che però interagiscono: consideriamo infatti ora una finestra triangolare della durata ancora di N campioni;

$$w_T(t) = 1 - \frac{2|t|}{NT}; \quad |t| \leq \frac{NT}{2} \quad W_T(\omega) = \frac{NT}{2} \sin^2 \frac{fNT}{2}; \quad (5.3)$$

Come si vede, per la finestra triangolare, i lobi laterali decrescono come $1/(\Delta f NT)^2$ ma il lobo principale ha larghezza doppia. In generale se $W(\omega)$ e $\{w_n\}$ sono legate da una trasformata di Fourier, allora la larghezza di banda equivalente B_{eq} la durata equivalente T_{eq} sono così legate:

$$\begin{aligned} \frac{1}{2\pi} \int W(\omega) e^{j\omega t} d\omega &= W(0) \times B_{eq} = w(0) = 1; \\ \int w(t) dt &= 1 \times T_{eq} = W(0); \rightarrow B_{eq} \times T_{eq} = 1 \end{aligned} \quad (5.4)$$

Pertanto, se per ridurre l'effetto dei lobi laterali si addolciscono le transizioni temporali (si apre e chiude la finestra in modo graduale e quindi si decresce T_{eq} a pari durata NT) nello stesso tempo si allarga l'area sottesa dalla $W(\omega)$ perchè ne allarga il lobo centrale di banda B_{eq} .

5.3 Finestre cosinusoidali

Una forma di finestre particolarmente efficiente è quella delle finestre cosinusoidali (coseno rialzato, Hann, Hamming) che danno un miglior compromesso lobi laterali - lobo principale che non la finestra triangolare. La finestra di Hann, o cosinusoidale è così costituita:

$$w_C(t) = \cos \left(\frac{\pi t}{NT} \right); \quad |t| < \frac{NT}{2} \quad (5.5)$$

La sua trasformata è deducibile in modo sintetico con il seguente ragionamento: poichè la $w_C(t)$ è il prodotto, nei tempi, di una cosinusoide di periodo $2NT$ e di una finestra rettangolare di durata NT , la sua trasformata sarà la convoluzione con il sinc della trasformata della cosinusoide e cioè due impulsi di eguale segno in $f = \pm 1/2NT$. Ma allora, i lobi laterali dei due sinc si sommano dopo una traslazione di frequenza

5.3 FINESTRE COSINUSOIDALI

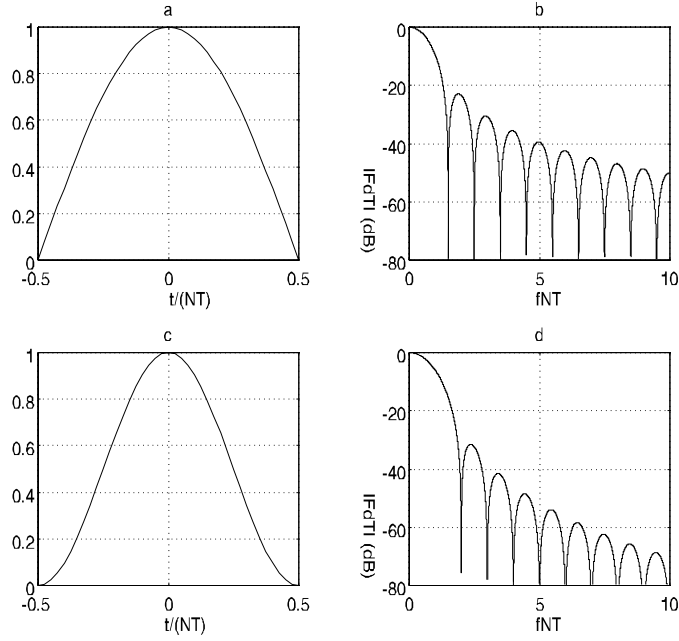


Figura 5.3: a) Finestra cosinusoidale b) Trasformata di Fourier della finestra cosinusoidale in dB c) Finestra a coseno rialzato d) Trasformata di Fourier della finestra a coseno rialzato in dB

pari a $1/NT$ e quindi le loro ampiezze si sottraggono, avendo i due lobi successivi segni opposti.

Pertanto, l' n -esimo lobo laterale decrescerà con una legge del tipo (fig. 5.3):

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)} \quad (5.6)$$

e quindi quadraticamente con n . Peraltro, la larghezza del lobo principale è $3/NT$ e cioè il 25% meno del lobo principale della finestra triangolare. Quindi, la finestra cosinusoidale è meglio della finestra triangolare, per queste applicazioni: il vero vantaggio della finestra triangolare sta nel fatto che lo spettro è sempre positivo, cosa non vera per la finestra cosinusoidale; ma questo vantaggio sarà utile per la stima spettrale. Si possono studiare finestre più complesse, ma fondate sullo stesso principio:

CAPITOLO 5 FILTRI NUMERICI

ad esempio la finestra a coseno rialzato (fig. 5.3.c,d):

$$w_R(t) = \frac{1}{2} \left[1 + \cos \left(\frac{2\pi t}{NT} \right) \right]; \quad |t| < \frac{NT}{2} \quad (5.7)$$

Qui, la trasformata si trova come convoluzione del sinc e della somma di un impulso di ampiezza .5 nell'origine e due impulsi di ampiezza .25 alle frequenze $f \pm 1/NT$. In questo caso, i lobi laterali decrescono con legge cubica come risulta da

$$-\frac{1}{4(n-1)} + \frac{1}{2n} - \frac{1}{4(n+1)} = \frac{1}{2n(n^2-1)} \quad (5.8)$$

Il lobo centrale ha larghezza $4/NT$ come la finestra triangolare: anche questa finestra è più efficiente della finestra triangolare, perchè a pari lobo centrale, i lobi laterali decrescono con legge cubica invece che quadratica.

5.4 Finestre diverse

In generale, si può pensare di costruire tanti tipi diversi di finestra, che ottimizzano diverse quantità; p.e. il raggio di girazione della risposta all'impulso, e cioè l'area incrementale moltiplicata per il quadrato della distanza dall'origine. Esistono anche famiglie di finestre (Finestre di Kaiser) che assomigliano a delle gaussiane e quindi tendono ad avere eguale aspetto nelle frequenze e nei tempi. La finestra di Kaiser è:

$$w_K(t) = \frac{I_0 \left[k \sqrt{1 - \left(\frac{2t}{NT} \right)^2} \right]}{I_0(k)}; \quad |t| \leq \frac{NT}{2} \quad (5.9)$$

La funzione I_0 è la funzione di Bessel modificata del primo tipo, di ordine 0; Il parametro k serve ad ottenere il compromesso desiderato tra ampiezza dei lobi laterali e larghezza del lobo principale. Una finestra interessante è la finestra di Hamming (fig. 5.4),

$$w_H(t) = .54 + .46 \cos \left(\frac{2\pi t}{NT} \right); \quad |t| < \frac{NT}{2} \quad (5.10)$$

Questa strana finestra ha un piccolo gradino agli estremi: essa quindi ha lobi che discendono come $1/N$ ma con piccola ampiezza perchè lo sbilanciamento dei tre lobi che vengono mediati nella convoluzione è piccolo. Il merito sta nel fatto che il primo lobo laterale ha ampiezza minima.

5.5 PROGETTO CON LA TECNICA MINIMAX

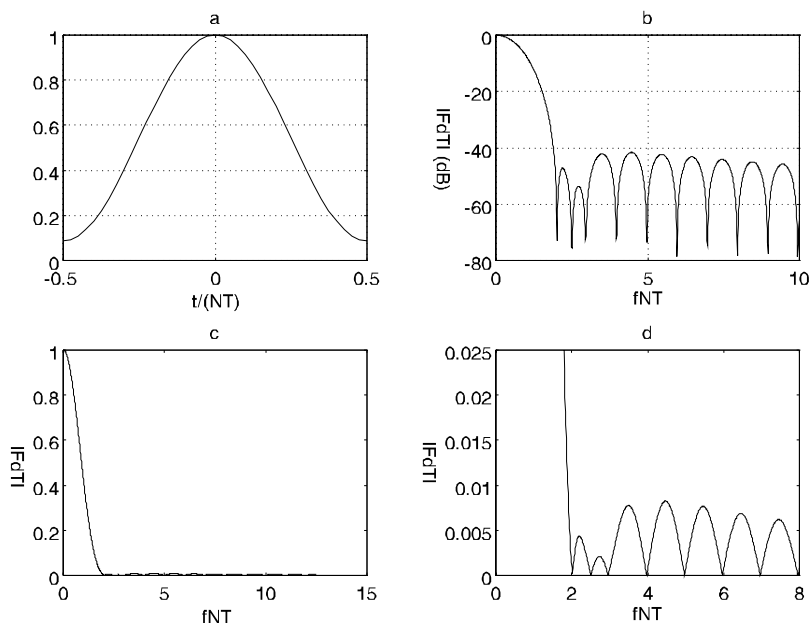


Figura 5.4: a) Finestra di Hamming b) Trasformata della finestra di Hamming in dB c) Trasformata della finestra di Hamming d) Particolare della trasformata della finestra di Hamming

5.5 Progetto con la tecnica MINIMAX

Nei casi studiati finora, per un dato lobo centrale, e quindi per un dato smussamento della risposta in frequenza, i lobi laterali decrescono con la differenza di frequenza; quindi le zone della funzione di trasferimento più lontane dalle transizioni ripide sono meno disturbate dalla finestatura.

Può essere utile progettare un filtro in modo che tutti i lobi laterali abbiano la stessa ampiezza e cioè in modo da minimizzare l'errore massimo dovunque capiti. Questo tipo di progetto, detto MINIMAX, è effettuato in modo iterativo. In questo caso, è essenziale che il numero dei parametri che caratterizza il filtro sia limitato e ci si riferirà ad un segnale a fase zero, cioè simmetrico nei tempi: si dimostra infatti che, a parità di lunghezza del filtro, i filtri a fase zero hanno le transizioni di ampiezza più ripide, nel dominio delle frequenze. Sia $\{h_n\}$; $|n| \leq N$ la sequenza corrispondente al filtro, lungo quindi $2N + 1$ campioni, di cui

CAPITOLO 5 FILTRI NUMERICI

però solo $N + 1$ sono diversi tra loro. La sua trasformata è:

$$H(\omega) = h_0 + \sum_{n=1}^N 2h_n \cos(n\omega T) \quad (5.11)$$

Vogliamo che

$$W(\omega)|H(\omega) - D(\omega)| \leq \delta \quad (5.12)$$

dove $W(\omega)$ è una funzione peso, positiva e la $D(\omega)$ è la funzione di trasferimento desiderata.

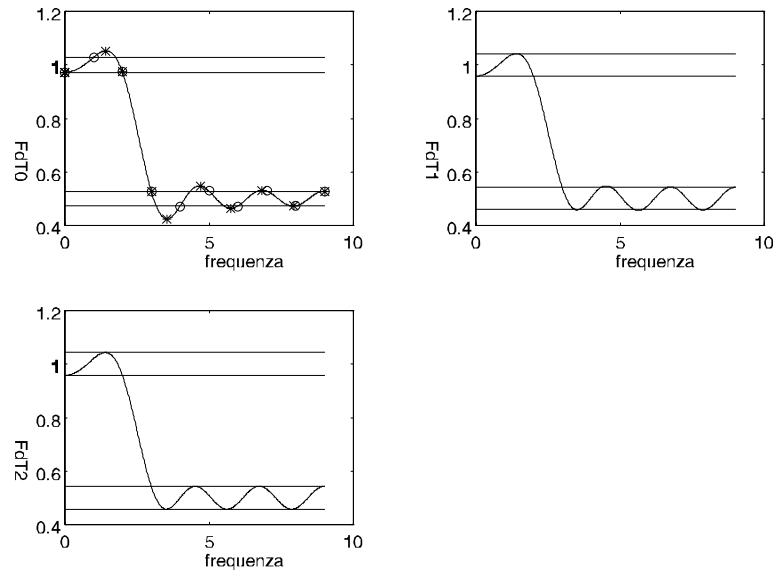


Figura 5.5: Progetto di un filtro FIR passa basso con la tecnica MINIMAX; progressive iterazioni

Esiste un teorema, dovuto a Chebyshev, che afferma che quando un polinomio trigonometrico di ordine N deve approssimare una funzione qualsiasi minimizzando l'errore massimo allora la differenza tra la funzione approssimata e quella approssimante ha $N + 2$ massimi e minimi successivi, di modulo eguale e segno alterno. Tuttavia, non sappiamo a quali frequenze la differenza sia massima; nasce allora una procedura iterativa.

5.6 LUNGHEZZA DEI FILTRI FIR PASSA - BASSO

Supponiamo di partire da un insieme di $N + 2$ frequenze $f_i^{(0)}; i = 1, \dots, N + 2$ e **imponiamo** che a queste frequenze le differenze tra funzioni approssimante ed approssimata siano tutte eguali ad un valore δ incognito. Inoltre, in corrispondenza delle discontinuità della maschera la differenza tra la funzione di trasferimento sintetizzata e quella desiderata raggiungerà un minimo od un massimo; pertanto, due frequenze estremali vanno anche posizionate in corrispondenza delle frequenze limite della banda passante e della banda attenuata dove appunto la caratteristica desiderata risulta essere discontinua. Scriveremo poi:

$$W(\omega_i^{(0)}) \left[D(\omega_i^{(0)}) - [h_0^{(0)} + \sum_{n=1}^N 2h_n^{(0)} \cos(n\omega_i^{(0)}T)] \right] = \quad (5.13)$$

$$= (-1)^i \delta^{(0)}; \quad i = 1, \dots, N + 2; \quad \omega_i^{(0)} = 2\pi f_i^{(0)}$$

Questo è un sistema lineare di $N + 2$ equazioni nelle $N + 2$ incognite $h_n^{(0)}, n = 0, \dots, N; \delta^{(0)}$, che dà sempre una soluzione. Tuttavia, in generale non sarà vero che le frequenze $f_i^{(0)}$ scelte all'inizio risultino essere le frequenze estremali, e cioè le frequenze alle quali la differenza (approssimante - approssimata) raggiunge i massimi e minimi: le frequenze estremali che risultano in questa prima approssimazione saranno invece altre, che indicheremo con $f_i^{(1)}$; pertanto alle frequenze estremali la differenza non risulta essere costante, come sarebbe richiesto. E' necessario iterare con il secondo insieme di frequenze; si risolve di nuovo il sistema di equazioni (5.13) e così via fino a che $f_i^{(k)} \simeq f_i^{(k+1)}$. A questo punto siamo alla convergenza ed accade appunto che a queste frequenze le differenze tra funzione approssimante ed approssimata hanno i valori (estremi) tutti eguali a $\delta^{(k+1)}$. A questo punto si verifica se $\delta^{(k+1)} \leq \delta$. Se sì, il problema è risolto. Altrimenti si deve ricominciare con N più grande. Naturalmente, se δ è troppo piccolo rispetto a quello richiesto, si può ricominciare con un valore di N minore (fig. 5.5).

5.6 Lunghezza dei filtri FIR passa - basso

È opportuno disporre di una stima del numero N di campioni da cui è costituito un filtro FIR, passa - basso; a priori, è chiaro che N non dipende tanto dalla frequenza di taglio del filtro quanto dalla larghezza Δf della banda di transizione del comportamento del filtro tra passaggio

CAPITOLO 5 FILTRI NUMERICI

e attenuazione; con f_c indichiamo la frequenza di campionamento. La lunghezza illimitata del filtro corrispondente alla risposta all'impulso del tipo *sinc* va delimitata con una finestra $\{w_n\}$ e quindi la risposta in frequenza convoluta con la trasformata della finestra $W(z)$. La larghezza del lobo principale della finestra determina quindi la larghezza della banda di transizione Δf . Utilizzando la tecnica MINIMAX per il calcolo del filtro, è disponibile una formula approssimata che dà la lunghezza del filtro in funzione dell'ondulazione (ripple) (δ_1) della funzione di trasferimento in banda passante e dell'attenuazione (δ_2) nella banda oscura.

$$N \approx .66 \times \frac{f_c}{\Delta f} \times \log_{10} \left(\frac{1}{10 \times \delta_1 \times \delta_2} \right)$$

Ad esempio, per $\delta_1 = \delta_2 = 10^{-4}$, $N \approx 4.6 \times f_c / \Delta f$. Poichè poi la larghezza della banda di transizione va commisurata con la larghezza di banda del filtro (ad esempio, il 20%) ne deriva una relazione tra la frequenza di taglio del filtro f_{\max} e la sua durata:

$$\text{Se: } \Delta f = .2 f_{\max} \quad \text{allora} \longrightarrow N \approx 20 f_c / f_{\max}$$

È chiaro che filtri passa basso molto selettivi sono molto costosi in termini di numero di moltiplicazioni; diventa interessante allora la tecnica di procedere per stadi e cioè:

- a) filtrare passa basso la sequenza iniziale riducendone la frequenza massima di un fattore ragionevole;
- b) sottocampionare p.e. dello stesso fattore la sequenza così ottenuta;
- c) iterare filtraggio passa basso e sottocampionamento.

In questo modo il risparmio può essere notevole, visto che la frequenza di taglio relativa dei filtri di ognuno degli stadi della sequenza è ben maggiore di quella finale, desiderata. In altre parole, per costruire un filtro a frequenza di taglio $f_c/1000$, si possono utilizzare 3 stadi di riduzione di un fattore 10, in cascata. La lunghezza del filtro necessario per ottenere il valore di Δf desiderato è ottenuta con un segnale fortemente sottocampionato e perciò a basso costo computazionale. Infatti, i filtri degli stadi successivi al primo operano su sequenze sottocampionate e pertanto hanno un costo ridotto di un fattore pari al fattore di sottocampionamento (10, nel caso considerato). Ulteriori analisi di questi filtri verranno fatte al capitolo 9, dedicato a decimazione e interpolazione.

5.7 Filtri IIR

Abbiamo visto che per realizzare filtri molto selettivi, è necessario utilizzare sequenze lunghe: ciò corrisponde a costi computazionali elevati a meno che non si ricorra all'uso di poli, e cioè costruire filtri con equazioni alle differenze finite recursive in cui il termine forzante è il segnale di ingresso e l'uscita è la soluzione. Come già visto, se i filtri sono causali, per assicurare la convergenza, sarà necessario garantire che i poli siano contenuti nel cerchio unitario. Tuttavia, ciò non è necessario se si opera in tempo differito, quando cioè tutti i dati di ingresso sono disponibili prima del trattamento e pertanto si può invertire l'asse temporale per ribaltare all'interno del cerchio unitario i poli esterni. Nel seguito, per semplicità ci riferiremo a filtri causali e pertanto avremo necessità di verificare la convergenza.

In generale, anche a scopo didattico, la progettazione di filtri IIR viene fatta trasponendo nel discreto i filtri analogici realizzati con sistemi di equazioni differenziali a coefficienti costanti. Un sistema lineare elettrico, meccanico, ecc. è rappresentato da un sistema di equazioni differenziali lineari, tempo continue; si desidera trasformare questi sistemi in sistemi alle differenze finite. Per fare ciò è necessario anzitutto definire l'operatore di derivazione per segnali numerici, e poi farne combinazioni lineari per ottenere l'equivalente di operatori di derivazione di ordine superiore.

5.7.1 Derivazione numerica

Se i dati fossero noti tutti, dal primo all'ultimo, e se potessimo operare in tempo differito, ed inoltre se le sequenze a disposizione fossero molto lunghe, non vi sarebbero problemi nel calcolo del derivatore. Infatti, come già visto nel capitolo sulla trasformazione di Fourier discreta, basta trasformare la sequenza, per poi moltiplicarla per $j2\pi k/NT$. Si ha così la trasformata della sequenza delle derivate, **periodicizzata**. Per comodità, ricalcoliamo la risposta all'impulso del derivatore campionato. La risposta del filtro ideale è:

$$x(t) = x = \frac{\sin \pi t/T}{\pi t/T};$$

CAPITOLO 5 FILTRI NUMERICI

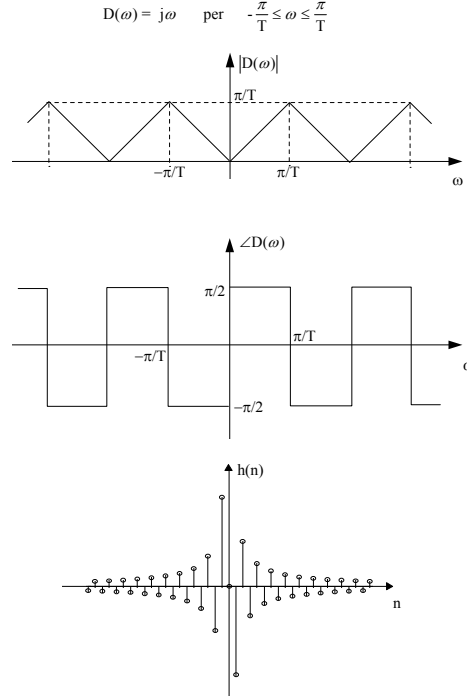


Figura 5.6: Derivatore ideale campionato

La derivata, campionata agli istanti $t = nT$, è:

$$\frac{dx}{dt} \Big|_{t=nT} = \frac{\cos n\pi}{nT} = \frac{(-1)^n}{nT}; n \neq 0; = 0; n = 0$$

Se la sequenza è lunga, la periodicizzazione non dà problemi, tranne che agli estremi (fig. 5.6). Però non è possibile operare in tempo reale, perchè abbiamo bisogno di sequenze lunghe e quindi verranno introdotti notevoli ritardi per attuare l'elaborazione nel dominio delle frequenze. Abbiamo pertanto necessità di tecniche che operino in tempo reale, e quindi dovremo determinare un operatore di derivazione causale; inoltre, vorremo che l'operatore di derivazione sia a fase minima perchè anche il suo inverso sia causale e stabile. Si consideri il sistema caratterizzato da

un'equazione differenziale di primo ordine:

$$\frac{dy(t)}{dt} = ay(t) + bx(t) \quad (5.14)$$

e trasformiamolo in un sistema discreto. La discretizzazione desider-

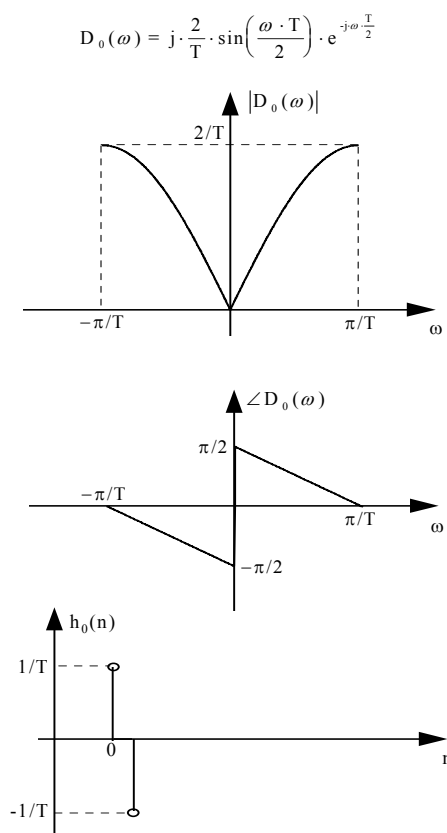


Figura 5.7: Derivatore non ricorsivo

abile corrisponderebbe al derivatore ideale discreto, che è come visto in precedenza:

$$d_f(nT) = \frac{(-1)^n}{nT}; \quad d_f(0) = 0$$

Questa soluzione però è antisimmetrica, e quindi non causale e quindi inaccettabile, perché non la si potrebbe realizzare in pratica con filtri

CAPITOLO 5 FILTRI NUMERICI

senza ritardo. La soluzione causale più semplice è:

$$\frac{y_n - y_{n-1}}{T} = ay_n + bx_n \quad (5.15)$$

in cui si è rappresentata la derivata con la differenza tra il campione attuale e quello precedente, divisa per T ; trasformiamo nelle frequenze per vedere la validità della scelta fatta:

$$D(\omega)Y(\omega) = Y(\omega)\frac{(1 - e^{-j\omega T})}{T} = aY(\omega) + bX(\omega) \quad (5.16)$$

$$D(\omega) = \frac{(1 - e^{-j\omega T})}{T} = \frac{2j}{T} \sin\left(\frac{\omega T}{2}\right) e^{-j\omega T/2} \simeq j\omega; \quad \omega T \ll 2 \quad (5.17)$$

L'operatore di derivazione $D(\omega)$ considerato va abbastanza bene, purché la frequenza sia piccola rispetto alla frequenza di Nyquist; sia la caratteristica di ampiezza che quella di fase differiscono però da quella di un derivatore esatto fino alla frequenza di Nyquist.

La caratteristica di ampiezza di $D(\omega)$ (fig. 5.7) ha alla frequenza di Nyquist valore $2/T$ invece che π/T ; inoltre nella caratteristica di fase, si osserva un ritardo pari a $T/2$, che fa sì che l'operatore non sia a fase minima, se non al limite; infatti, ha uno zero sul cerchio unitario. Il ritardo di mezzo campione è comprensibile perché la derivata è riferita all'istante $(n-1/2)T$. E' ragionevole quindi fare come proposto da Crank e Nicholson e cioè riferire anche il secondo membro dell'equazione (5.15) allo stesso istante.

$$\frac{y_n - y_{n-1}}{T} = a\frac{(y_n + y_{n-1})}{2} + b\frac{(x_n + x_{n-1})}{2} \quad (5.18)$$

Nel dominio delle frequenze questo diventa:

$$Y(\omega)\frac{(1 - e^{-j\omega T})}{T} = aY(\omega)\frac{1 + e^{-j\omega T}}{2} + bX(\omega)\frac{1 + e^{-j\omega T}}{2}$$

$$Y(\omega)\frac{2(1 - e^{-j\omega T})}{T(1 + e^{-j\omega T})} = aY(\omega) + bX(\omega) \quad (5.19)$$

L'operatore di derivazione è perciò rappresentato dall'operatore di derivazione ricorsiva:

$$D_r(\omega) = \frac{2(1 - e^{-j\omega T})}{T(1 + e^{-j\omega T})} = j\frac{2}{T} \tan \frac{\omega T}{2} \quad (5.20)$$

5.7 FILTRI IIR

La caratteristica di fase è quella giusta e cioè un anticipo costante di $\pi/2$; la caratteristica di ampiezza, invece, tocca valore infinito alla frequenza di Nyquist. Peraltro, dalla figura 5.8 è evidente come il derivatore $D_r(\omega)$ sia più preciso dell'operatore $D(\omega)$ fino a frequenze più elevate. Abbiamo quindi una tecnica per trasformare un sistema differenziale continuo in un sistema discreto: basterà sostituire alla variabile ω_a (la frequenza analogica) la variabile ω_n (la frequenza numerica):

$$\omega_n = \frac{2}{T} \tan^{-1} \left(\frac{\omega_a T}{2} \right) \quad (5.21)$$

Rappresentando ora la sequenza corrispondente alla $D_r(\omega)$ con la trasformata z , scriveremo la seguente corrispondenza:

$$j\omega \Rightarrow D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.22)$$

Di nuovo abbiamo un sistema che ha uno zero sul cerchio unitario e quindi rimane da dimostrare che un sistema differenziale stabile viene trasformato in un sistema discreto stabile e cioè poli nel semipiano sinistro nel caso di sistemi tempo continui vengono trasformati in poli all'interno del cerchio unitario con la corrispondenza dell'equazione (5.21). Piuttosto che fare una dimostrazione, faremo una verifica con il sistema RC della fig. 5.9:

La tensione $V_r(\omega)$ è:

$$V_r(\omega) = \frac{E(\omega)}{1 + j\omega\tau} \quad \tau = RC \quad (5.23)$$

Nel caso discreto, il segnale $e(t)$ si trasforma in una sequenza $\{e_n\}$, e l'equazione si trasforma in un'equazione alle differenze finite di cui presentiamo la trasformata

$$\begin{aligned} \frac{V(z)}{E(z)} &= 1 / \left(1 + \tau \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \right) \\ &= \frac{1 + z^{-1}}{1 + a + (1 - a)z^{-1}}; \quad a = \frac{2\tau}{T} \end{aligned} \quad (5.24)$$

Antitrasformando, si ha:

$$v_n = \frac{a - 1}{a + 1} v_{n-1} + \frac{1}{1 + a} (e_n + e_{n-1}) \quad (5.25)$$

CAPITOLO 5 FILTRI NUMERICI

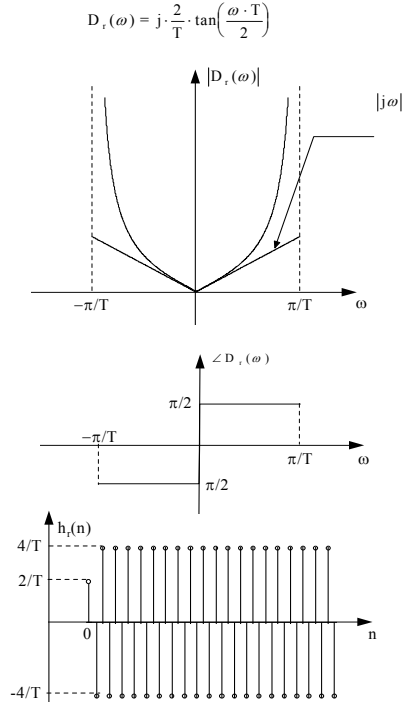


Figura 5.8: Derivatore ricorsivo

Il polo è nel cerchio unitario e cioè:

$$-1 < \frac{a-1}{a+1} < 1 \quad \text{se } \tau > 0 \quad (5.26)$$

e cioè se e solo se la resistenza R è positiva e quindi se il sistema che opera nel tempo continuo è stabile. In conclusione, la derivazione recursiva determinata con l'operatore

$$D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.27)$$

trasforma sistemi continui stabili in sistemi discreti stabili: inoltre la trasformazione non altera sostanzialmente l'asse delle frequenze fino a circa $1/3$ della frequenza di Nyquist, come vedremo al prossimo paragrafo.

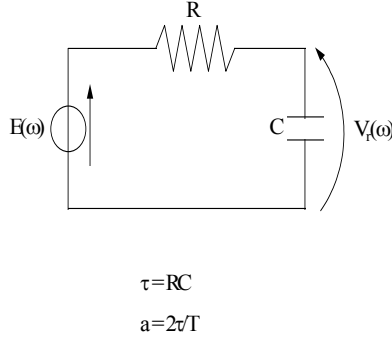


Figura 5.9: Filtro passa - basso RC

5.7.2 La trasformazione bilineare

La corrispondenza tra sistemi tempo continui causali e tempo discreti causali indotta dall'equazione (5.21)

$$j\omega \rightarrow D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} = j \frac{2}{T} \tan \frac{\omega T}{2} \quad (5.28)$$

può essere vista come una corrispondenza tra l'asse delle frequenze di un sistema continuo e quello di un sistema discreto (fig. 5.10). Infatti se alla variabile $j\omega$ si sostituisce la variabile $D_r(w)$ si ottiene un sistema discreto al posto di un sistema continuo, come si è visto. Se $H_c(\omega_a)$ è il valore della funzione di trasferimento del sistema tempo continuo ad una pulsazione analogica ω_a assegnata, la funzione di trasferimento del sistema tempo discreto corrispondente $H_d(\omega)$ avrà lo stesso valore alla pulsazione ω_d dove

$$\omega_d = \frac{2}{T} \tan^{-1} \left(\frac{\omega_a T}{2} \right) \quad (5.29)$$

$$\omega_a = \frac{2}{T} \tan \left(\frac{\omega_d T}{2} \right) \quad (5.30)$$

Quindi avviene che i valori assunti dalla funzione di trasferimento analogica nell'intervallo $0 \div \infty$ sono assunti dalla funzione di trasferimento numerica corrispondente nell'intervallo $0 \div f_N$. I filtri passa basso perciò vanno a zero alla frequenza di Nyquist. Ogni tecnica per il progetto di filtri analogici, p.e. i filtri di Butterworth o di Chebyshev può essere

CAPITOLO 5 FILTRI NUMERICI

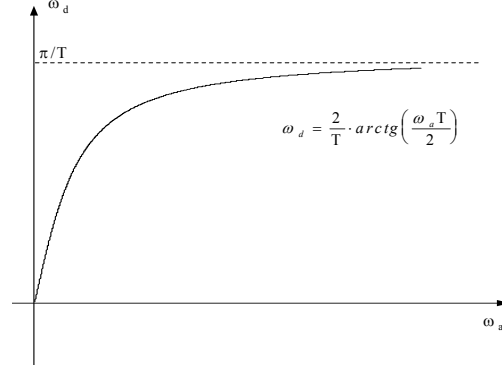


Figura 5.10: Trasformazione bilineare: relazione tra frequenza analogica e frequenza numerica

trasposta in campo numerico facilmente con la trasformazione bilineare, appunto. Si devono però poi applicare delle predistorsioni in maniera da compensare la distorsione non uniforme dell'asse delle frequenze indotta dalla trasformazione bilineare. Ad esempio, si verifica che:

$$\omega_a \simeq \omega_d \left[1 + \frac{\omega_d^2 T^2}{12} \right] + \dots \quad (5.31)$$

L'errore di frequenza commesso è pertanto inferiore al 10% per frequenze $f < \simeq .35 f_N$. Se il filtro da progettare è fortemente selettivo, la caratteristica di frequenza va predistorta. Per esempio, si voglia progettare un filtro passa banda tra $.8 f_N \div .998 f_N$; basterà cercare sulle tabelle il filtro analogico passabanda tra le frequenze:

$$\frac{1}{\pi T} \tan .4\pi \div \frac{1}{\pi T} \tan .499\pi$$

Quindi si applica la trasformazione bilineare per passare dal filtro analogico al filtro numerico, e il risultato è ottenuto. Ritornando all'esempio del circuito RC del paragrafo precedente, si può osservare che il sistema analogico ha come risposta all'impulso un esponenziale con costante di tempo τ ; il sistema numerico corrispondente ha un polo posizionato in $a - 1/a + 1$ ed uno zero alla frequenza di Nyquist. La risposta all'impulso

5.7 FILTRI IIR

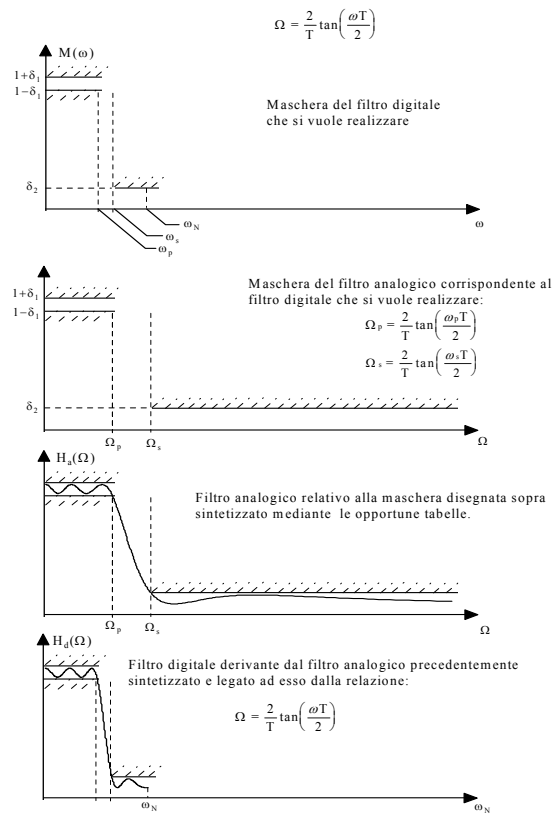


Figura 5.11: Esempio di progettazione di un filtro digitale mediante trasformazione bilineare

CAPITOLO 5 FILTRI NUMERICI

del sistema numerico è:

$$h_n = 0; \quad n < 0; \quad h_0 = \frac{1}{1+a}; \quad h_n = \frac{2a}{a^2-1} \left(\frac{a-1}{a+1} \right)^n \quad (5.32)$$

Il termine esponenziale è approssimabile poi come:

$$\frac{2a}{a^2-1} \left(\frac{a-1}{a+1} \right)^n \simeq \frac{T}{\tau} e^{n \log(a-1/a+1)} \simeq \frac{T}{\tau} e^{-nT/\tau}; \quad \log(1+x) \simeq x \quad (5.33)$$

e quindi approssima (a parte il solito fattore T) la risposta all'impulso del sistema analogico, campionata, ma ne è diversa soprattutto in corrispondenza del primo campione, come vedremo meglio al prossimo paragrafo

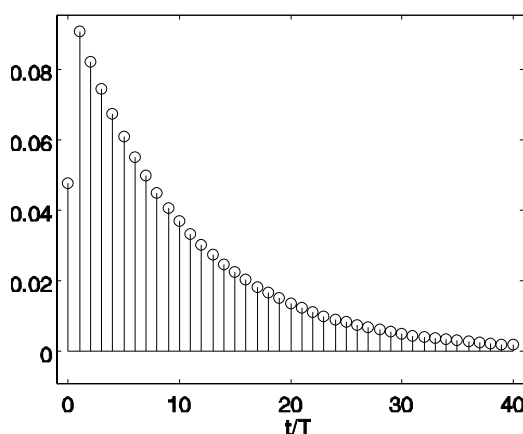


Figura 5.12: Risposta all'impulso del sistema numerico (RC); $a = \frac{2\tau}{T} = 20$

5.8 L'invarianza della risposta impulsiva

Una diversa corrispondenza tra sistemi analogici e sistemi numerici può essere stabilita imponendo che la risposta all'impulso dei secondi sia ottenuta dal campionamento della risposta all'impulso dei primi. Ad esempio alla funzione di trasferimento

$$H(\omega) = \frac{1}{1+j\omega\tau} \rightarrow h(t) = \frac{1}{\tau} e^{-t/\tau}$$

5.9 FILTRI DA COMBINAZIONE DI SFASATORI PURI

si farebbe corrispondere la sequenza:

$$h_n = \frac{T}{\tau} e^{-nT/\tau}; \quad n = 0, 1, \dots$$

e quindi la trasformata z

$$H(z) = \frac{T}{\tau} \frac{1}{1 - e^{-T/\tau} z^{-1}}$$

Di nuovo, a sistemi stabili corrisponderebbero sistemi stabili; le risposte all'impulso si somiglierebbero, ma le risposte in frequenza sarebbero piuttosto diverse tranne ovviamente che per frequenze molto basse e per $T \ll \tau$. Infatti, il campionamento, qualsiasi sia l'intervallo T , è **effettuato senza prefiltraggio** e quindi comporta equivocazione. In particolare, alla frequenza di Nyquist, dove vi è il ripiegamento, la funzione di trasferimento desiderata sarebbe sempre riprodotta con un errore del 100%. Nel caso della funzione di trasferimento discussa prima, risultati migliori si avrebbero imponendo (*per continuità*) che il valore assunto al tempo 0 sia $T/2\tau$ invece che T/τ ; così si aggiunge uno zero (non proprio alla frequenza di Nyquist come nel caso della trasformazione bilineare, ma molto vicino se $T \ll \tau$), riducendo il fenomeno dell'equivocazione. Infatti in questo caso risulta:

$$H'(z) = \frac{T}{2\tau} \frac{1 + e^{-T/\tau} z^{-1}}{1 - e^{-T/\tau} z^{-1}}$$

Pertanto, con la tecnica dell'invarianza impulsiva, sarebbe più difficile fare delle trasposizioni da un sistema analogico ad un sistema numerico che non con la trasformazione bilineare (fig. 5.10).

5.9 Filtri da combinazione di sfasatori puri

Da una semplice costruzione vettoriale ricaveremo ora un interessante classe di filtri, poi molto utilizzati nella pratica. Si considerino due sfasatori puri, e cioè filtri caratterizzati da poli all'interno del cerchio unitario (i denominatori sono funzioni a fase minima) e zeri nelle posizioni reciproche coniugate (i numeratori sono delle funzioni a fase massima) così da avere caratteristica di ampiezza unitaria. Si ricordi che uno

CAPITOLO 5 FILTRI NUMERICI

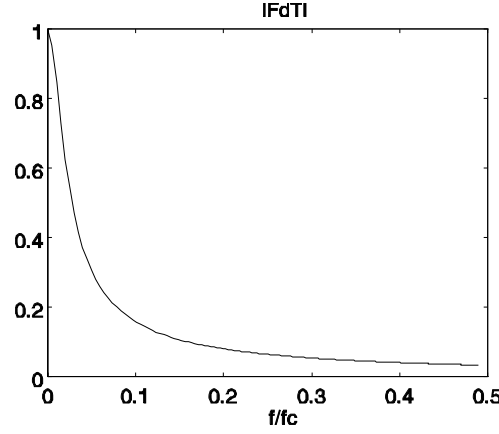


Figura 5.13: Modulo della f.d.t. del sistema (RC) numerico con $a = \frac{2\tau}{T} = 20$

sfasatore puro di ordine M comporta solo M moltiplicazioni per campione; inoltre, rimane sempre tale, anche al variare della quantizzazione dei coefficienti. Ad esempio, siano gli sfasatori puri $A_0(z)$, $A_1(z)$ di ordine N_0 , N_1 anche diverso:

$$A_0(z) = \frac{\sum_{n=0}^{N_0} a_{0,N_0-n}^* z^{-n}}{\sum_{n=0}^{N_0} a_{0,n} z^{-n}}; \quad |A_0(z)| = 1, \quad z = e^{j\omega T} \quad (5.34)$$

$$A_1(z) = \frac{\sum_{n=0}^{N_1} a_{1,N_1-n}^* z^{-n}}{\sum_{n=0}^{N_1} a_{1,n} z^{-n}}; \quad |A_1(z)| = 1, \quad z = e^{j\omega T} \quad (5.35)$$

$$H_0(z) = \frac{A_0(z) + A_1(z)}{2} \quad (5.36)$$

$$H_1(z) = \frac{A_0(z) - A_1(z)}{2} \quad (5.37)$$

Con la costruzione vettoriale di fig. 5.14 si vede che $H_0(\omega)$, $H_1(\omega)$ sono semidiagonali di un rombo che ha per lati $|A_0(\omega)| = |A_1(\omega)| = 1$; pertanto:

$$|H_0(z)|^2 + |H_1(z)|^2 = 1; \quad z = e^{j\omega T} \quad (5.38)$$

e cioè i due filtri $H_0(\omega)$, $H_1(\omega)$ sono complementari in potenza; alle frequenze a cui è passa basso l'uno sarà passa alto l'altro e viceversa.

5.9 FILTRI DA COMBINAZIONE DI SFASATORI PURI

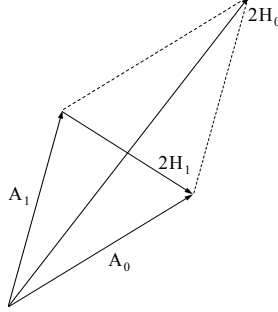


Figura 5.14: Filtri IIR ottenuti come somma e differenza di due sfasatori puri

Inoltre i due filtri sono anche complementari in ampiezza in quanto la loro somma è di ampiezza costante:

$$|H_0(z) + H_1(z)| = |A_0(z)| = 1; \quad z = e^{j\omega T} \quad (5.39)$$

I due filtri $H_0(\omega)$, $H_1(\omega)$ sono pertanto detti *doppiamente complementari*, in potenza ed ampiezza cioè. Ancora vi è un'altra utile caratteristica dei due filtri $H_0(\omega)$, $H_1(\omega)$; le sequenze causali corrispondenti ai numeratori delle due funzioni di trasferimento sono l'una simmetrica e l'altra antisimmetrica. Infatti:

$$A_0(\omega) = \frac{\tilde{D}_0(\omega)}{D_0(\omega)}, \quad A_1(\omega) = \frac{\tilde{D}_1(\omega)}{D_1(\omega)}; \quad (5.40)$$

$$H_{0,1}(\omega) = \frac{\tilde{D}_0(\omega)D_1(\omega) \pm \tilde{D}_1(\omega)D_0(\omega)}{2D_0(\omega)D_1(\omega)} = \frac{N(\omega) \pm \tilde{N}(\omega)}{2D_0(\omega)D_1(\omega)} \quad (5.41)$$

ma anche:

$$N(\omega) = \tilde{D}_0(\omega)D_1(\omega); \quad \tilde{N}(\omega) = \tilde{D}_1(\omega)D_0(\omega) \quad (5.42)$$

Si verifica che le famiglie di filtri di Butterworth, Chebyshev, ellittici sono appunto scomponibili nella semisomma o semidifferenza di due sfasatori puri.

Come trovare i due sfasatori puri date le funzioni di trasferimento? Nel caso interessante dei filtri ellittici (nonchè Butterworth e Chebyshev,

CAPITOLO 5 FILTRI NUMERICI

che ne sono casi limite), la cosa è semplice perchè si verifica che i poli di $A_0(\omega)$, $A_1(\omega)$ che sono anche poli di $H_0(\omega)$, $H_1(\omega)$ sono alternati nelle frequenze; in altre parole i poli di uno sfasatore puro si alternano con quelli dell'altro. Pertanto per fare la decomposizione, basta assegnare i poli a frequenze crescenti alternativamente ai due sfasatori. Gli zeri degli sfasatori sono dedotti dai poli. Nel caso di funzioni di trasferimento di ordine pari, l'osservazione è tuttora valida, ma i coefficienti degli sfasatori puri sono complessi [Vedi: Vaidyanathan, P. P., Multirate Systems and Filter Banks, Prentice Hall, 1993].

Elemento chiave per la progettazione dei filtri è il rapporto tra $|H_0(\omega)|$, $|H_1(\omega)|$ e cioè la funzione

$$K(\omega) = \frac{|N(\omega) + \tilde{N}(\omega)|}{|N(\omega) - \tilde{N}(\omega)|} \quad (5.43)$$

che ripartisce la potenza tra i due filtri, al variare della frequenza.

5.10 Filtri di Butterworth, Chebyshev, ellittici

Queste famiglie di filtri numerici, derivano dalle famiglie omologhe di filtri utilizzati per segnali analogici attraverso la trasformazione bilineare. Pertanto ne faremo solo un brevissimo cenno, limitato ai filtri di Butterworth, rinviando gli interessati alla letteratura specializzata ed inoltre al software *MATLABTM*.

Ad esempio, il filtro passabasso di Butterworth di ordine n e frequenza di taglio ω_0 ha la funzione di trasferimento

$$|B_n(\omega)|^2 = \frac{1}{1 + (\frac{\omega}{\omega_0})^{2n}} \quad (5.44)$$

I poli della funzione di trasferimento analogica sono su di un semicerchio del piano s di raggio ω_0 e di fasi:

$$\frac{\pi}{2n} \times (n + 1 + 2k); \quad 0 \leq k \leq n - 1 \quad (5.45)$$

Dalla posizione dei poli del filtro analogico, si ricava la funzione di trasferimento del filtro analogico e poi di quello numerico, usando sempre la

5.11 SINTESI DI FILTRI TUTTI POLI DALL'AUTOCORRELAZIONE

relazione:

$$j\omega \leftrightarrow \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (5.46)$$

La funzione di ripartizione della potenza in questo caso è

$$K(\omega) = |B_n(\omega_d)|^2 = \left[\frac{2}{\omega_0 T} \frac{|1 - z^{-1}|}{|1 + z^{-1}|} \right]^{2n} \quad (5.47)$$

Se $K(\omega) \rightarrow \infty$, la potenza del segnale è deviata attraverso H_0 ; attraverso H_1 , se invece $K(\omega) = 0$.

Nella fig.5.15 sono mostrate funzioni di trasferimento, risposte all'impulso e disposizione di poli e zeri di quattro filtri passa basso IIR di ordine 10. Si osservi come il filtro ellittico ha la massima ripidità di transizione e quindi transitorio di lunghezza maggiore mentre il filtro di Butterworth ha il transitorio più breve. Si osservi come l'ottimizzazione della distribuzione delle singolarità (*calcolando la funzione di trasferimento con il metodo grafico si comprende che le singolarità devono raggrupparsi vicino al cerchio unitario in corrispondenza delle transizioni ripide*) conduca a funzioni di trasferimento più vicine alla maschera di progetto. Dalla figura è poi facile comprendere la posizione delle singolarità dei due sfasatori puri in cui si potrebbe decomporre il filtro.

5.11 Sintesi di filtri tutti poli dall'autocorrelazione

5.11.1 Le equazioni di Yule Walker

Si desideri determinare un filtro tutti poli di ordine N , $1/A_N(z)$, la cui caratteristica di ampiezza approssimi la caratteristica di ampiezza desiderata, $|H(z)|$:

$$\left| \frac{1}{A_N(z)} \right| \sim |H(z)|$$

Il filtro ottenuto dovrà essere stabile e quindi a minima fase. Per trovare $A_N(z)$, utilizzeremo la tecnica di Yule Walker. Si osservi anzitutto che assegnare la funzione $|H(z)|$ implica anche il rendere disponibile l'autocorrelazione della risposta all'impulso $\{h_n\}$ del filtro $H(z)$. Infatti,

CAPITOLO 5 FILTRI NUMERICI

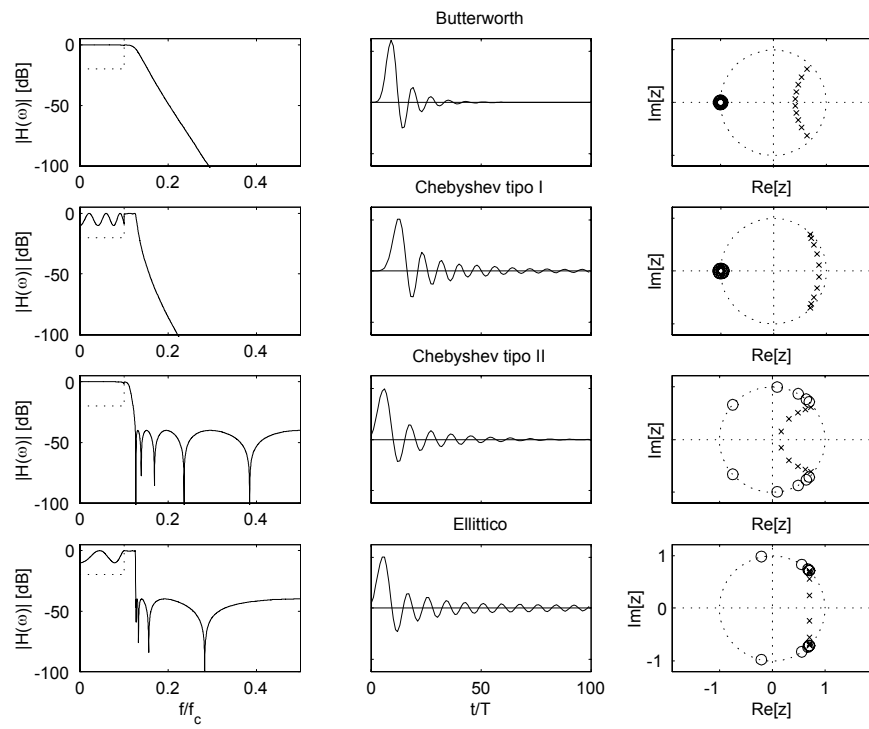


Figura 5.15: Filtri passa-basso IIR di ordine 10 [specifiche progetto: $f_p = f_c/4$, $\delta_1 = 0.5dB$, $\delta_2 = -40dB$]. La caratteristica della risposta in frequenza nella banda passante è amplificata (zona tratteggiata).

5.11 SINTESI DI FILTRI TUTTI POLI DALL'AUTOCORRELAZIONE

l'antitrasformata di $|H(z)|^2$ non é altro che l'autocorrelazione $\{r_m\}$ di $\{h_n\}$.

$$|H(z)||H(z)|^* \Leftrightarrow r_m = \sum_{n=0}^{\infty} h_n h_{n-m}^*$$

Utilizzeremo ora i primi N tra i valori assegnati di $\{r_m\}$ per costruire il filtro $A_N(z)$.

Sia $\{x_n\}$ la sequenza in uscita del filtro IIR di ordine N , $1/A_N(z)$ quando questo fosse alimentato da una sequenza bianca $\{w_n\}$, e cioè da una sequenza di campioni indipendenti. Senza ledere la generalit , $A_N(z)$   rappresentato da un polinomio monico, e cio  con il primo coefficiente pari ad 1. Quindi,

$$A_N(z) = 1 + \sum_{h=1}^N a_{h,N} z^{-h}$$

  descritto da N coefficienti $a_{h,N}$, $h = 1, ..N$. La sequenza $\{x_n\}$, uscita del filtro, si ottiene dall'equazione alle differenze finite:

$$x_n = - \sum_{h=1}^N a_{h,N} x_{n-h} + w_n \quad (5.48)$$

e passando alle trasformate z

$$X(z) = \frac{W(z)}{A_N(z)}$$

L'autocorrelazione r_m della sequenza $\{x_n\}$ che, se appunto il filtro A_N   alimentato da una sequenza bianca,   ancora l'autocorrelazione della risposta all'impulso del filtro, pu  essere dedotta moltiplicando la equazione (5.48) per il campione della sequenza x_{n-m}^* (anticipato di m campioni e coniugato) e poi considerando il valore atteso del prodotto:

$$r_x(m) = r_m = E[x_n x_{n-m}^*] = -E\left[\sum_{h=1}^N a_{h,N} x_{n-h} x_{n-m}^*\right] + E[w_n x_{n-m}^*] = \quad (5.49)$$

$$r_m = - \sum_{h=1}^N a_{h,N} r_{m-h} + \sigma_w^2 \delta(m); \quad \forall m \quad (5.50)$$

CAPITOLO 5 FILTRI NUMERICI

Infatti la sequenza $\{w_n\}$ è incorrelata con i campioni precedenti x_{n-m} ; $m > 0$. Per $m = 0$, si ha poi:

$$E[w_n^* x_{n-m}] = E[w_n^* w_n] = \sigma_w^2$$

Le equazioni (5.50) sono note come le equazioni di Yule - Walker e forniscono:

- gli N coefficienti del filtro $A_N(z)$ a partire dai primi N valori della funzione di autocorrelazione $\{r_m\}$, a parte r_0 , ottenuti con l' antitrasformazione del filtro desiderato $|H(z)|^2$.

- tutti i campioni dell'autocorrelazione dell'uscita del filtro, dati i coefficienti del filtro $A_N(z)$ ed il valore quadratico medio del segnale d'ingresso σ_w^2 . I primi N campioni saranno eguali a quelli desiderati, per costruzione. Non così i successivi, a meno che il filtro $H(z)$ non sia stato ricostruito perfettamente e cioè se $H(z)$ è un filtro a N_1 poli, e $N_1 < N$.

Ancora, con la tecnica che sarà vista nel capitolo sulle strutture riverberanti, da $A_N(z)$ si possono determinare gli N coefficienti di riflessione c_1, c_2, \dots, c_N che caratterizzano la struttura riverberante che consta di N celle e che approssima la funzione di trasferimento $H(z)$. Si osservi anche che nel caso di segnale e o coefficienti $a_{h,N}$ complessi, le equazione (5.50) vanno scritte tenendo conto che $r_{-m} = r_m^*$.

Per motivi mnemonici, può essere utile osservare che la formula (5.50) assomiglia ad una convoluzione tra le sequenze r_m ed $a_{h,N}$; non è una convoluzione, per la limitazione degli indici. L'equazione corrispondente a $m = 0$ contiene anche la varianza del segnale di ingresso σ_w^2 ; peraltro, il filtro non dipende da questo fattore di scala e si può sempre riscalare la funzione di autocorrelazione ponendo $r_0 = 1$ risolvendo poi le equazioni (Y-W) per $m = 1, \dots, N$.

Il fatto che l'equazione (5.50) sia definita per tutti i valori di m , implica che, una volta assegnata qualsiasi N -pla consecutiva di valori di r_m , poi possiamo determinare $A_N(z)$ e quindi i valori di r_m per qualsiasi altro valore di m .

Le equazioni di Yule Walker si possono scrivere in forma matriciale

5.11 SINTESI DI FILTRI TUTTI POLI DALL'AUTOCORRELAZIONE

indicando i seguenti vettori con i corrispondenti simboli:

$$\underline{\mathbf{a}}_N = \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix}; \quad \underline{\boldsymbol{\xi}}_N = \begin{bmatrix} x_{n-1}^* \\ x_{n-2}^* \\ x_{n-3}^* \\ \dots \\ x_{n-N}^* \end{bmatrix}; \quad \underline{\mathbf{r}}_N = \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \dots \\ r_x(N) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix}$$

Allora si ha:

$$x_n = -\underline{\boldsymbol{\xi}}_N^* \underline{\mathbf{a}}_N + w_n \quad (5.51)$$

E moltiplicando a sinistra per $\underline{\boldsymbol{\xi}}_N$ e considerando il valore atteso, visto che $E[\underline{\mathbf{x}}_N w_n] = \underline{\mathbf{0}}$:

$$E[\underline{\boldsymbol{\xi}}_N x_n] = \underline{\mathbf{r}}_N = E[\underline{\boldsymbol{\xi}}_N (-\underline{\boldsymbol{\xi}}_N^* \underline{\mathbf{a}}_N + w_n)] = -E[\underline{\boldsymbol{\xi}}_N \underline{\boldsymbol{\xi}}_N^*] \underline{\mathbf{a}}_N = -\underline{\mathbf{R}} \underline{\mathbf{a}}_N$$

La matrice dei coefficienti delle equazioni (Y-W) è una matrice di Toeplitz; cioè è simmetrica (o Hermitiana, per sequenze complesse) e tutti gli elementi appartenenti alla stessa diagonale o subdiagonale sono uguali tra loro. La matrice $\underline{\mathbf{R}}$ $[N \times N]$ è pertanto caratterizzata da N numeri. Il vettore dei termini noti $-\underline{\mathbf{r}}_N$ è anch'esso formato dai campioni della funzione di autocorrelazione, r_m ; $m = 1, \dots, N$, col segno cambiato. Riscrivendo in forma matriciale le equazioni di Yule Walker si ottiene:

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & r_{-1} \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 \end{bmatrix} \begin{bmatrix} a_{1,N} \\ a_{2,N} \\ a_{3,N} \\ \dots \\ a_{N,N} \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \dots \\ r_N \end{bmatrix} \quad (5.52)$$

ed in equazione vettoriale:

$$\underline{\mathbf{R}}_N \underline{\mathbf{a}}_N = -\underline{\mathbf{r}}_N \quad (5.53)$$

CAPITOLO 5 FILTRI NUMERICI

Per completezza, si aggiunge una formulazione completa, facilmente verificabile, in cui compare anche la prima equazione che contiene la varianza della sequenza bianca di ingresso $\sigma_{w,N}^2$.

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} & \dots & r_{-N} \\ r_1 & r_0 & r_{-1} & \dots & r_{-N+1} \\ r_2 & r_1 & r_0 & \dots & r_{-N+2} \\ \dots & \dots & \dots & \dots & r_{-1} \\ r_N & r_{N-1} & r_{N-2} & \dots & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} \sigma_{w,N}^2 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (5.54)$$

Si ricorda infine che in queste equazioni non sempre le incognite sono contenute nel vettore $\underline{\mathbf{a}}_N$; a volte è la funzione di autocorrelazione a non essere nota, oppure il valore di $\sigma_{w,N}^2$, ecc.

5.11.2 La ricorsione di Levinson

Nel calcolo di $A_N(z)$, si può utilizzare la stessa recursione di Levinson, incontrata nello studio delle strutture riverberanti. Infatti, si può risolvere l'equazione (5.53) a partire da $N = 1$, per poi crescere progressivamente. Si osserva che:

$$\underline{\mathbf{R}}_{N+1} = \begin{bmatrix} \underline{\mathbf{R}}_N & \underline{\tilde{\mathbf{r}}}_N \\ \underline{\tilde{\mathbf{r}}}_N^* & r_0 \end{bmatrix}; \quad \underline{\tilde{\mathbf{r}}}_N = \begin{bmatrix} r_{-N} \\ r_{-N+1} \\ \dots \\ r_{-1} \end{bmatrix}$$

il vettore colonna $\underline{\tilde{\mathbf{r}}}_N$ non è altro che il vettore colonna $\underline{\mathbf{r}}_N$, coniugato, ribaltato negli indici e ritardato per avere indici tutti positivi, con l'usuale convenzione che abbiamo visto legare le due sequenze $\tilde{A}(z)$ ed $A(z)$:

$$\tilde{A}_N(z) = z^{-N} A_N^*(1/z)$$

Il vettore riga $\underline{\tilde{\mathbf{r}}}_N^*$ è il trasposto Hermitiano (e quindi coniugato) del vettore colonna $\underline{\tilde{\mathbf{r}}}_N$.

$$\begin{bmatrix} \underline{\mathbf{R}}_N & \underline{\tilde{\mathbf{r}}}_N \\ \underline{\tilde{\mathbf{r}}}_N^* & r_0 \end{bmatrix} \underline{\mathbf{a}}_{N+1} = - \begin{bmatrix} \underline{\mathbf{r}}_N \\ r_{N+1} \end{bmatrix}$$

5.11 SINTESI DI FILTRI TUTTI POLI DALL'AUTOCORRELAZIONE

Si scriva ora la soluzione come:

$$\underline{\mathbf{a}}_{N+1} = \begin{bmatrix} \underline{\mathbf{a}}_N \\ 0 \end{bmatrix} + \begin{bmatrix} \underline{\mathbf{d}}_N \\ c_{N+1} \end{bmatrix}$$

$$\begin{bmatrix} \underline{\mathbf{R}}_N & \widetilde{\underline{\mathbf{r}}}_N \\ \widetilde{\underline{\mathbf{r}}}_N^* & r_0 \end{bmatrix} \left\{ \begin{bmatrix} \underline{\mathbf{a}}_N \\ 0 \end{bmatrix} + \begin{bmatrix} \underline{\mathbf{d}}_N \\ c_{N+1} \end{bmatrix} \right\} = - \begin{bmatrix} \underline{\mathbf{r}}_N \\ r_{N+1} \end{bmatrix}$$

Si trovano due equazioni : la prima, matriciale, riguardante le prime N righe, e la seconda, scalare, corrispondente alla $N + 1$ esima riga.

$$\underline{\mathbf{R}}_N (\underline{\mathbf{a}}_N + \underline{\mathbf{d}}_N) + \widetilde{\underline{\mathbf{r}}}_N c_{N+1} = - \underline{\mathbf{r}}_N$$

$$\widetilde{\underline{\mathbf{r}}}_N^* (\underline{\mathbf{a}}_N + \underline{\mathbf{d}}_N) + c_{N+1} r_0 = - r_{N+1}$$

profittando della (5.53) la prima equazione diventa:

$$\underline{\mathbf{d}}_N = - c_{N+1} \underline{\mathbf{R}}_N^{-1} \widetilde{\underline{\mathbf{r}}}_N$$

Qui si fa l'osservazione ovvia, ma fondamentale, che:

$$\underline{\mathbf{d}}_N = c_{N+1} \widetilde{\underline{\mathbf{a}}}_N$$

dove al solito:

$$\widetilde{\underline{\mathbf{a}}}_N = \begin{bmatrix} a_{N,N}^* \\ \vdots \\ \vdots \\ a_{2,N}^* \\ a_{1,N}^* \end{bmatrix}$$

L'ultima equazione serve infine per determinare c_{N+1}

$$\widetilde{\underline{\mathbf{r}}}_N^* (\underline{\mathbf{a}}_N + c_{N+1} \widetilde{\underline{\mathbf{a}}}_N) + c_{N+1} r_0 = - r_{N+1}$$

da cui:

$$c_{N+1} = - \frac{r_{N+1} + \widetilde{\underline{\mathbf{r}}}_N^* \underline{\mathbf{a}}_N}{r_0 + \widetilde{\underline{\mathbf{r}}}_N^* \widetilde{\underline{\mathbf{a}}}_N} \quad (5.55)$$

CAPITOLO 5 FILTRI NUMERICI

Ne consegue la tecnica detta di Levinson per la soluzione progressiva delle equazioni di Yule e Walker. Si ha cioè:

$$A_{N+1}(z) = A_N(z) + c_{N+1}z^{-1}\tilde{A}_N(z); \quad \tilde{A}_N(z) = z^{-N}A_N^*(1/z)$$

dove c_N è appunto il coefficiente di riflessione N -esimo della struttura riverberante che realizza la funzione di trasferimento $1/A_N(z)$ ed è ottenuto dai valori della funzione di autocorrelazione e dal polinomio $A_{N-1}(z)$ nel modo osservato in (5.55) e come verrà di nuovo dimostrato nel capitolo della predizione lineare. Si dimostra anche che:

$$\sigma_{w,N}^2 = r_0 (1 - |c_1|^2) (1 - |c_2|^2) \dots (1 - |c_N|^2)$$

e quindi:

$$\sigma_{w,N-1}^2 \leq \sigma_{w,N}^2$$

Questo modo di risolvere le equazioni di Yule e Walker è computazionalmente efficiente. Infatti, per risolvere un sistema di N equazioni in N incognite servono un ordine di grandezza di N^3 moltiplicazioni; usando la ricorsione di Levinson, invece, il costo computazionale di un incremento di indice corrisponde, a parte le somme, al costo di calcolo della (5.55) e quindi $2N$ moltiplicazioni ed una divisione. Il costo totale diventa quindi circa:

$$\sum_{n=1}^N 2n \sim 2 \frac{N^2}{2} = N^2$$

5.11.3 La sintesi del filtro

Riassumendo, se viene assegnata la caratteristica di ampiezza di un filtro da approssimare con una funzione trasferimento tutti poli di ordine N , $A_N(z)$, per effettuarne la sintesi e quindi per determinarne il denominatore:

$$|H(z)| \sim \frac{1}{|A_N(z)|}$$

si opererà così:

1) Si calcola $|H(\omega)|^2$ e, facendone l'antitrasformata z , si trova la funzione di autocorrelazione:

$$r(n); \quad n = 0, 1, \dots, N, \dots, \infty$$

5.11 SINTESI DI FILTRI TUTTI POLI DALL'AUTOCORRELAZIONE

2) Utilizzando le equazioni di Yule Walker, si trova, per ogni valore desiderato di N , il vettore $\underline{\mathbf{a}}_N$;

3) Si determina la funzione di trasferimento del filtro:

$$\frac{1}{A_N(z)} \sim H(z)$$

È necessario garantire che il filtro $1/A_N(z)$ sia stabile e pertanto che il polinomio $A_N(z)$ sia a minima fase; si dimostrerà che ciò accade **se e solo se** i campioni $r(n)$ che si utilizzano nelle equazioni di Yule Walker sono campioni di una funzione di autocorrelazione (a spettro positivo). Ciò è automaticamente ottenuto se $r(n)$ è antitrasformata di una funzione tutta positiva come $|H(\omega)|^2$. Poichè, per ottenere il filtro di ordine N , non serve **tutta** la funzione di autocorrelazione, ma ne bastano i primi N campioni, lo spettro $|H(\omega)|^2$ può essere smussato, purchè sia convoluto con una funzione tutta positiva come p.e. un sinc^2 .

Capitolo 6

Campionamento in frequenza

6.1 Introduzione

Una classe di filtri molto utilizzata e di facile progettazione è quella dei filtri a campionamento in frequenza. Si tratta di filtri che sono teoricamente dei filtri FIR, anche se con dei poli, in quanto questi poli sono cancellati da opportuni zeri. Nei fatti, visto che la cancellazione è solo parziale, risultano essere dei filtri IIR.

Si consideri un oscillatore numerico, come quello studiato al capitolo 3. Se questo oscillatore è eccitato da un impulso, la sua risposta è una sinusoide, di ampiezza proporzionale a quella dell'impulso. Se si applicano più impulsi nel tempo, la risposta sarà la somma delle risposte. Supponiamo ora che la sinusoide generata compia un numero intero k di cicli, nell'intervallo di tempo NT . Se allora inviamo all'oscillatore due impulsi di eguale ampiezza e segno opposto a distanza di NT la risposta globale sarà una sinusoide di frequenza k/NT troncata dopo k periodi (fig. 6.1). Quindi la risposta all'impulso del filtro è finita ed il filtro è FIR, anche se realizzato con oscillatori numerici e cioè con strutture IIR. La realizzazione è presentata in fig. 6.4.

Essa è ottenuta combinando un oscillatore numerico ed un blocco di ritardo di N campioni realizzato con una memoria di N elementi, per creare sistematicamente l'impulso di spegnimento N campioni dopo l'impulso di attivazione dell'oscillatore. Il costo computazionale è basso, visto che si richiede una moltiplicazione per campione nell'oscillatore ed un'altra per stabilire il guadagno.

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

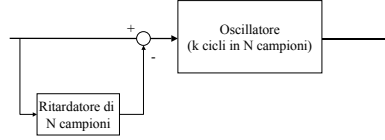


Figura 6.1: Funzionamento di un oscillatore a campionamento di frequenza

La risposta in frequenza é quella di una sinusoide a frequenza f_0 finestrata con una finestra rettangolare di durata NT ; quindi del tipo

$$H_c(\omega) = \text{sinc}[(f - f_0)NT] \quad (6.1)$$

A parte le repliche spettrali dovute al campionamento nei tempi, lo spettro corrisponde ad un sinc con il primo zero ad un intervallo di frequenza pari a $1/NT$ centrato sulla frequenza dell'oscillatore. Quindi, la selettività non é legata al costo computazionale, ma solo alla dimensione della memoria, che determina la durata della risposta all'impulso.

Poli e zeri

Il sistema di fig. 6.4 é caratterizzato da N zeri dovuti al blocco di ritardo e i poli dovuti all'oscillatore (uno se corrispondono a esponenziali complessi o due se a frequenze reali); la trasformata z della risposta é (fig. 6.2):

$$H(z) = \frac{1 - z^{-N}}{1 - 2 \cos(2\pi k/N)z^{-1} + z^{-2}}; \quad k \neq 0, N/2 \quad (6.2)$$

$$H(z) = \frac{1 - z^{-N}}{1 \mp z^{-1}} \quad k = 0, N/2. \quad (6.3)$$

Gli zeri del primo blocco corrispondono alle N radici N -esime dell'unit  e quindi sono disposti regolarmente lungo il cerchio unitario a distanza angolare pari a $2\pi/N$ con il primo zero nel punto $z = 1$ (la componente continua é soppressa dal primo blocco). I poli sono posizionati in

6.1 INTRODUZIONE

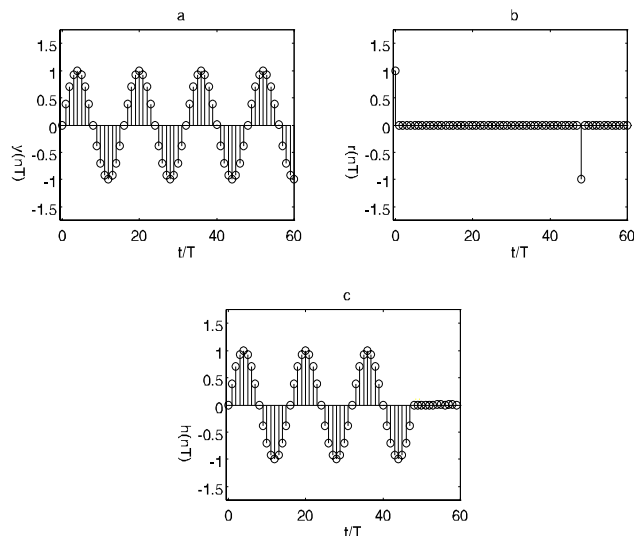


Figura 6.2: a) Risposta impulsiva dell'oscillatore b) Risposta impulsiva del ritardatore c) Risposta all'impulso del filtro a campionamento di frequenza

corrispondenza di qualche zero e ne sono cancellati, **se la frequenza dell'oscillatore é esatta**. Quindi il sistema é tutti zeri, FIR.

In generale, come rappresentato in figura 6.6 il blocco di ritardo può alimentare fino a $N/2 + 1$ oscillatori, alle frequenze $f_k = k/NT$; $k = 0..N/2$; in questo modo si potrebbero costruire complesse caratteristiche di ampiezza senza però avere vantaggi computazionali, perché con $N/2 + 1$ oscillatori il numero di moltiplicazioni per campione diventa pari a N come con un qualsiasi filtro FIR di pari lunghezza. Quindi i filtri a campionamento in frequenza hanno interesse dal punto di vista del costo computazionale solo se il numero di oscillatori alimentati é piccolo rispetto ad N e quindi se hanno delle caratteristiche di tipo passa banda piuttosto pronunciate.

Un altro motivo per cui i filtri a campionamento in frequenza sono interessanti, a parte quello computazionale visto in precedenza, sta nella modalità di progetto, assai semplice.

Infatti si consideri la caratteristica del filtro desiderato $D(\omega)$; (**attenzione: visto che il filtro approssimante sarà causale, anche il filtro approssimando deve essere causale!** Quindi $D(\omega)$ deve tenere conto

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

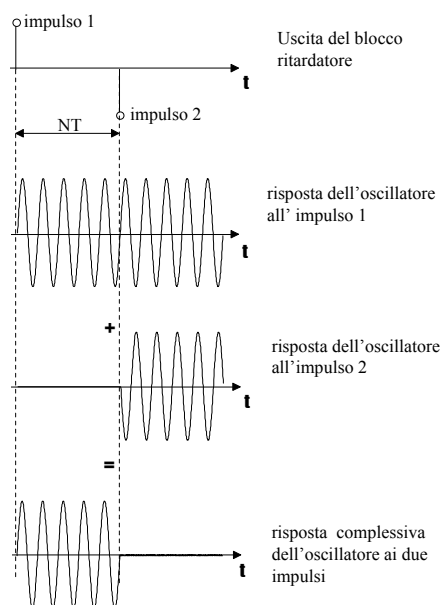


Figura 6.3: Modo alternativo per vedere il funzionamento di un filtro a campionamento di frequenza

di un ritardo pari a $N/2$ campioni) si campioni questa caratteristica alle $N/2 + 1$ pulsazioni $\omega_k = 2\pi k/NT$; $k = 0, \dots, N/2$ e siano D_k i valori (in generale complessi) di questi campioni. Ora, si interpolino questi campioni lungo l'asse delle frequenze con una funzione sinc che ha appunto il primo zero a intervallo $1/NT$; la funzione risultante $H(\omega)$ è appunto quella del sistema a campionamento in frequenza presentato nella figura 6.5, purché i guadagni H_k siano pari a D_k . Per ritrovare valori reali dei guadagni, si ricordi di accoppiare i rami corrispondenti alle frequenze positive con quelli delle frequenze negative.

La dimostrazione di quanto osservato segue dal fatto che stiamo facendo un'interpolazione (con un filtro ideale nel dominio delle frequenze) di un segnale costituito da una sequenza di impulsi (le sinusoidi alle frequenze $f_k = k/NT$); nel capitolo 2 un'interpolazione simile era stata fatta nel dominio dei tempi. La convoluzione con il *sinc* nel dominio delle frequenze corrisponde alla finestratura rettangolare fatta nel dominio dei tempi. Peraltro si ricordi che sia ha a che fare con segnali campionati e pertanto

6.2 PROGETTO DEI FILTRI A CAMPIONAMENTO IN FREQUENZA

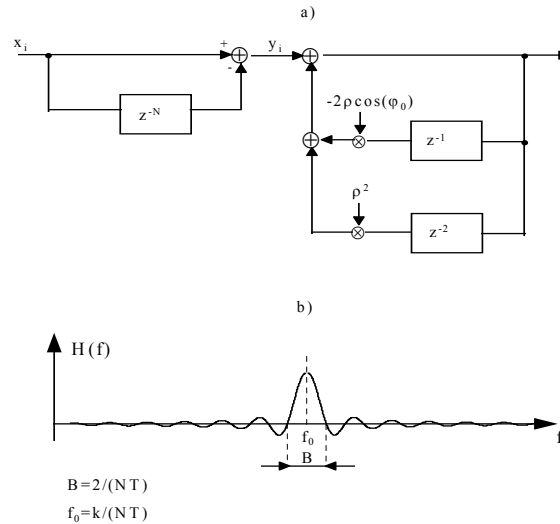


Figura 6.4: Schema a blocchi (a) e funzione di trasferimento (b) di un filtro a campionamento in frequenza

periodici nel dominio delle frequenze; la trasformata del *sinc* discreto è presentata nel Capitolo 4.

Al crescere di N , cresce la qualità dell'interpolazione e quindi decresce la differenza tra approssimante ed approssimata. Per filtri passa banda, come si è detto, c'è poi anche un vantaggio computazionale.

Si osservi che la somma in parallelo dei vari rami nel sistema di figura 6.6 comporta la formazione di nuovi zeri accanto alla cancellazione dei poli. Altrimenti si arriverebbe al paradosso che la risposta finale ha solo un campione.

6.2 Progetto dei filtri a campionamento in frequenza

Il progetto di filtri FIR utilizzando i campioni della trasformata di Fourier desiderata presi ad intervallo regolare $1/NT$ può essere ricondotto alla progettazione di filtri assegnando i valori dei campioni della DFT

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

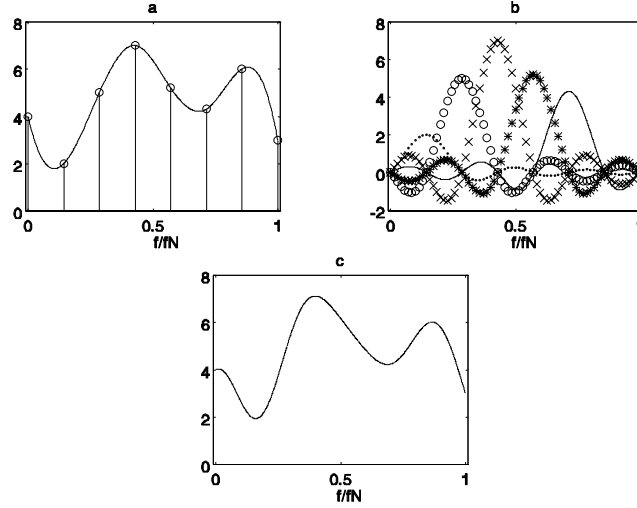


Figura 6.5: a) Campionamento della F.d.T. da sintetizzare b) Interpolazione con funzioni *sinc* c) Risultato dell' interpolazione

$\{H_k\}_{k=0}^{N-1}$. L'unica accortezza è che il filtro ricavato per questa via ha una risposta temporale periodica che deve essere troncata su un periodo per ottenere un filtro a risposta temporale limitata (FIR). Orbene, **il troncamento non può essere arbitrario**. Vengono presentati di seguito i passi principali:

a) Antitrasformazione discreta di Fourier: $\{H_k\}_{k=0}^{N-1} \Leftrightarrow \{\tilde{h}_n\}_{n=0}^{N-1}$. Si ricordi che la sequenza è periodica (ovvero $\tilde{h}_n = \tilde{h}_{n+N}$).

b) Troncamento su 1 periodo della sequenza periodica (sequenza a-periodica): $h_n = \tilde{h}_n$ per $n = 0, \dots, N-1$ e zero altrove.

c) Trasformata z della sequenza troncata:

$$\begin{aligned}
 H(z) &= \sum_{n=0}^{N-1} h_n z^{-n} = \sum_{n=0}^{N-1} \underbrace{\left[\frac{1}{N} \sum_{k=0}^{N-1} H_k W_N^{kn} \right]}_{\text{IDFT}} z^{-n} = \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \left[\sum_{n=0}^{N-1} W_N^{kn} z^{-n} \right] =
 \end{aligned}$$

6.2 PROGETTO DEI FILTRI A CAMPIONAMENTO IN FREQUENZA

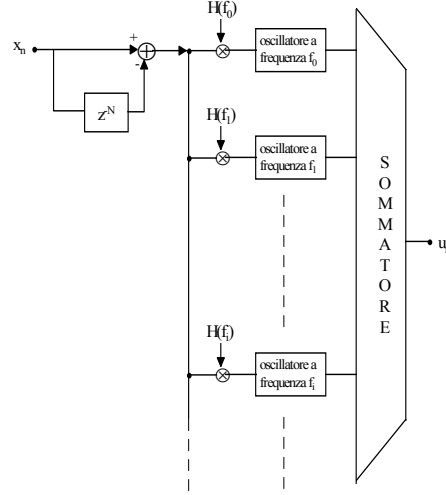


Figura 6.6: Schema a blocchi di un filtro a campionamento in frequenza i cui campioni assegnati siano $H(f_i)$

$$\begin{aligned}
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \left[\sum_{n=0}^{N-1} (W_N^k z^{-1})^n \right] \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} H_k \frac{1 - W_N^{kN} z^{-N}}{1 - W_N^k z^{-1}} = \\
 &= \frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - W_N^k z^{-1}} = \\
 &\quad \frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - \exp(j \frac{2\pi}{N} k) z^{-1}}
 \end{aligned}$$

ovvero la risposta all'impulso del filtro può essere realizzata da un insieme di oscillatori che generano sinusoidi complesse di ampiezza $|H_k|$ e fase $\angle H_k$. Gli oscillatori sono “spenti” dopo N campioni e avere quindi compiuto un numero intero k di cicli. Se la sequenza è reale allora $H_k = H_{N-k}^*$ e quindi la trasformata z diventa (separando il campione

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

alla continua, $k = 0$, e al Nyquist, $k = N/2$, quando quest'ultimo è presente e cioè se N è pari):

$$\begin{aligned}
 H(z) &= \frac{1 - z^{-N}}{N} \left[\frac{H_0}{1 - z^{-1}} + \frac{H_{N/2}}{1 + z^{-1}} + A \right] \\
 A &= \sum_{k=1}^{N/2-1} \left(\frac{H_k}{1 - \exp(j\frac{2\pi}{N}k)z^{-1}} + \frac{H_k^*}{1 - \exp(-j\frac{2\pi}{N}k)z^{-1}} \right) \\
 &= \frac{1 - z^{-N}}{N} \left[\frac{H_0}{1 - z^{-1}} + \frac{H_{N/2}}{1 + z^{-1}} + B \right] \\
 B &= 2 \sum_{k=1}^{N/2-1} \left(\frac{\operatorname{Re}(H_k) - \operatorname{Re}(H_k \exp(-j\frac{2\pi}{N}k)) z^{-1}}{1 - 2 \cos(\frac{2\pi}{N}k)z^{-1} + z^{-2}} \right)
 \end{aligned}$$

L'ultimo termine è una somma di oscillazioni sinusoidali **reali** opportunamente sfasate secondo il termine di fase indicato dallo zero ($\operatorname{Re}(H_k) - \operatorname{Re}(H_k \exp(-j\frac{2\pi}{N}k)) z^{-1}$). È opportuno osservare che per H_k reali la risposta all'impulso di ogni oscillatore è una cosinusoide.

d) Trasformata di Fourier della sequenza di durata limitata:

$$H(z = \exp(j\omega T)) = \frac{1 - e^{-j\omega NT}}{N} \sum_{k=0}^{N-1} \frac{H_k}{1 - e^{j(\frac{2\pi}{N}k - \omega T)}}$$

sviluppando ciascuno dei termini nella sommatoria la trasformata diventa:

$$\begin{aligned}
 H(z = \exp(j\omega T)) &= \sum_{k=0}^{N-1} H_k \cdot W(\omega T - \frac{2\pi}{N}k) \\
 \text{con } W(\omega) &= \frac{\exp(-j\omega T(N-1)/2)}{N} \cdot \frac{\sin(\omega NT/2)}{\sin(\omega T/2)}
 \end{aligned}$$

rileggendo questa relazione si osserva che la trasformata di Fourier della sequenza h_n può essere ricavata da H_k per (convoluzione) interpolazione con la funzione $W(\omega)$ (nota che $H(2\pi k/N) = H_k$), oppure la sequenza periodica è troncata da una finestra rettangolare (causale) di N campioni:

$$\begin{aligned}
 w_n &= \begin{cases} 1 & n = 0, 1, \dots, N-1 \\ 0 & \text{altrove} \end{cases} \longleftrightarrow \\
 W(\omega) &= \frac{\exp(-j\omega T(N-1)/2)}{N} \cdot \frac{\sin(\omega TN/2)}{\sin(\omega T/2)}.
 \end{aligned}$$

6.2 PROGETTO DEI FILTRI A CAMPIONAMENTO IN FREQUENZA

Si osserva quindi che esiste un limitato controllo delle caratteristiche del filtro in quanto la fdt è assegnata (in fase di progetto) solo attraverso i campioni in frequenza H_k . L'andamento effettivo della fdt tra i campioni è controllato dalla funzione $W(\omega)$. Il termine di ritardo $\exp(-j\omega T(N-1)/2)$ introduce una oscillazione nella risposta in frequenza del filtro che può essere eliminata solo se si introduce una traslazione antitetica opportuna (o ritardo) della risposta all'impulso periodica. Riscrivendo la $H(\omega)$:

$$H(\omega) = \exp(-j\omega T(N-1)/2) \sum_{k=0}^{N-1} \underbrace{H_k \exp(j \frac{2\pi}{N} k \times \frac{(N-1)}{2})}_{\text{anticipo seq. periodica}} \cdot C$$

$$C = \frac{\sin((\omega T - \frac{2\pi}{N} k)N/2)}{N \sin((\omega T - \frac{2\pi}{N} k)/2)}$$

si evidenziano tre contributi (nell'ordine): il ritardo di $(N-1)/2$ campioni necessario per avere un filtro causale, i coefficienti del filtro con uno sfasamento corrispondente ad un anticipo (il termine $\exp(j \frac{2\pi}{N} k \times \frac{(N-1)}{2})$) e il seno cardinale periodicizzato che effettua l'interpolazione. Per ovviare ai termini indesiderati indotti dall'interpolazione è necessario scegliere i coefficienti del filtro H_k in modo che questi contengano un ritardo di $(N-1)/2$ campioni tale da compensare l'anticipo visto prima. In altri termini, a causa della periodicità di \tilde{h}_n il troncamento da applicare per ottenere h_n non è arbitrario. In pratica è sempre **necessario** ritardare di $(N-1)/2$ campioni la sequenza per ovviare ad un troncamento troppo brusco all'inizio e alla fine della risposta h_n . Questo troncamento introdurrebbe termini ad alta frequenza indesiderati e zeri in posizioni difficilmente prevedibili. In fase di progetto il ritardo di $(N-1)/2$ campioni si realizza sfasando opportunamente i campioni in frequenza ovvero scegliendo opportunamente i campioni in frequenza \bar{H}_k , che sono:

$$H_k = \bar{H}_k \exp(-j\pi k \frac{N-1}{N}) = [\text{per } N \text{ grande}] \simeq (-1)^k \bar{H}_k,$$

per segnali reali : $H_k = H_{N-k}^*$

Questo giustifica l'inversione di segno necessaria quando si vuole realizzare un filtro passa banda combinando due risposte all'impulso a frequenze adiacenti. A titolo di esempio in figura si mostra la risposta all'impulso

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

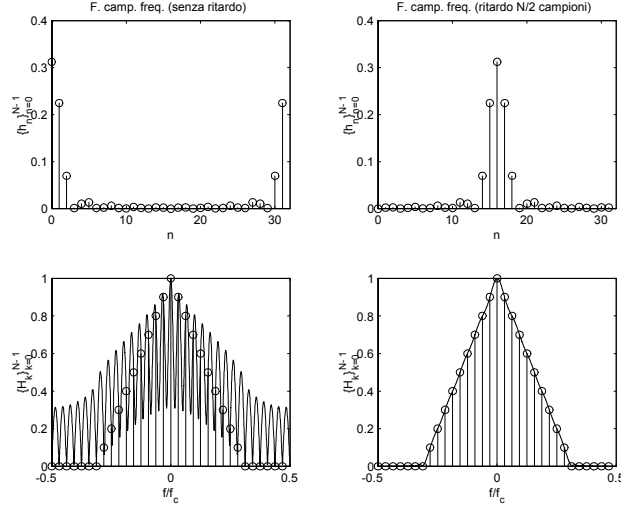


Figura 6.7: Risposta all'impulso (parte superiore) e trasformata di Fourier (parte inferiore) di un filtro passa basso a campionamento in frequenza in cui sono assegnati i campioni della fdt $\{H_k\}_{k=0}^{N-1}$ (pallini vuoti).

e la fdt di un filtro passa basso a campionamento in frequenza senza applicazione di ritardo e con un ritardo di $N/2$ campioni (per semplicità qui si è scelto di ritardare di $N/2$ campioni in quanto $N = 32 \gg 1$). In entrambi i casi la fdt coincide, per $\omega T = 2\pi k/N$, con i campioni della risposta in frequenza del filtro definiti in fase di progetto. Nel secondo caso la fdt appare più “ragionevole” per l'assenza di termini indesiderati ad alta frequenza (introdotti dal termine $\exp[-j\omega T(N-1)/2]$ nella $W(\omega)$). La posizione degli zeri per i due casi è indicata nella figura.

Osservazioni:

1) Il ritardo di $(N-1)/2$ campioni rappresenta un numero intero di campioni nel caso di N dispari mentre corrisponde ad un'interpolazione nei tempi nel caso di N pari (vedi esempio in figura). In quest'ultimo caso si potrebbe adottare un ritardo di $N/2$ a meno di un “piccolo” errore (l'errore è trascurabile solo per risposta all'impulso di durata limitata).

2) Ogni oscillatore ha (nel caso di risposta all'impulso reale) un costo di 3 moltiplicazioni/campione. Complessivamente ci sono $N/2-1$ oscillatori con un costo di (circa) $3N/2$ moltiplicazioni/campione. Se confrontato con un'implementazione diretta di un filtro FIR con risposta h_n (N

6.3 QUANTIZZAZIONE DEI COEFFICIENTI

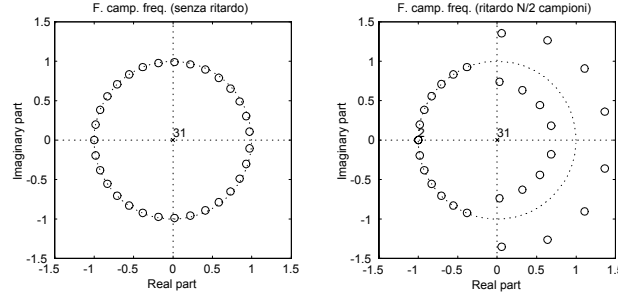


Figura 6.8: Poli/zeri filtro campionamento in frequenza senza e con ritardo di $N/2$ campioni. In entrambi i casi, i valori dei campioni in frequenza sono eguali!

moltiplicazioni/campione) il filtro a campionamento in frequenza ha un vantaggio computazionale solo se il numero di coefficienti H_k è piccolo (se il numero di oscillatori è limitato).

6.3 Quantizzazione dei coefficienti

Un importante problema realizzativo affligge i filtri a campionamento in frequenza: mentre le posizioni degli zeri della funzione di trasferimento sono corrispondenti proprio alle radici dell'unità, le posizioni dei poli dipendono dalla precisione con cui vengono realizzati i fattori $-2\cos(2\pi k/N)$; se uno di questi fattori è rappresentato con un numero binario finito, e quindi imprecisamente, la perfetta cancellazione polo - zero diventa impossibile. La sinusoide lanciata da un impulso non fa un numero intero di cicli in N campioni, e l'impulso di segno opposto che arriva dopo N campioni lancia quindi una sinusoide che non sarà in perfetta opposizione di fase con la precedente. La somma delle due sinusoidi non si annulla e quindi rimane una sinusoide, residuo della cancellazione imperfetta. La risposta all'impulso del filtro diventa IIR, *non convergente a zero*; il sistema diventa quindi instabile, come vedremo. Un oscillatore con funzione di trasferimento

$$H(z) = \frac{1}{1 + az^{-1} + z^{-2}}; \quad |a| \leq 2 \quad (6.4)$$

CAPITOLO 6 CAMPIONAMENTO IN FREQUENZA

genera una sinusoide campionata $\{y_n\}$ con pulsazione

$$\omega T = \varphi = \cos^{-1}\left(\frac{a}{2}\right) \quad (6.5)$$

Noi desidereremmo avere esattamente:

$$\omega = \frac{2\pi k}{N} \rightarrow a = -2 \cos(2\pi k/N)$$

Ma le piccole variazioni di a dovute alla quantizzazione dei coefficienti portano ad una variazione $\Delta\omega$ della pulsazione ω :

$$\Delta\omega = \frac{\Delta a}{2T \sin \varphi} \quad (6.6)$$

La differenza tra la sinusoide desiderata e quella generata dopo N campioni é:

$$\sin \alpha_1 - \sin \alpha_2 = 2 \cos \frac{\alpha_1 + \alpha_2}{2} \sin \frac{\alpha_1 - \alpha_2}{2} \quad (6.7)$$

$$\alpha_1 = (2\pi k/NT + \Delta\omega)(n + N)T; \quad \alpha_2 = (2\pi k/NT + \Delta\omega)nT \quad (6.8)$$

L'ampiezza della sinusoide residua è (supponendo unitaria l'ampiezza della sinusoide generata):

$$y_{r,n} = 2 \sin \frac{\Delta\omega NT}{2} \cos\left[(\omega + \Delta\omega)nT + \frac{\Delta\omega NT}{2}\right] \simeq \frac{N\Delta a}{4 \sin \varphi} \cos\left(n\varphi + \frac{N\Delta a}{4 \sin \varphi}\right) \quad (6.9)$$

Come si vede, rimane una sinusoide residua di ampiezza anche piccola se Δa è piccolo, ma di durata illimitata. La situazione é inaccettabile, perché ogni campione in ingresso poi lascerebbe un residuo sinusoidale di durata illimitata; dopo n campioni di ingresso casuali, di eguale varianza σ ed incorrelati tra loro, la varianza del segnale residuo, somma di n sinusoidi di ampiezza e fase casuale, cresce proporzionalmente a $\sigma\sqrt{n}$. Quindi, dopo qualche tempo, il filtro sarebbe comunque totalmente inutilizzabile, per buona che fosse la quantizzazione. Il sistema è diventato instabile.

Per stabilizzarlo, non rimane che riportare i poli degli oscillatori all'interno del cerchio unitario ($\rho < 1$), in modo da limitare in pratica la durata delle sinusoidi residue, che decadranno esponenzialmente a zero

6.3 QUANTIZZAZIONE DEI COEFFICIENTI

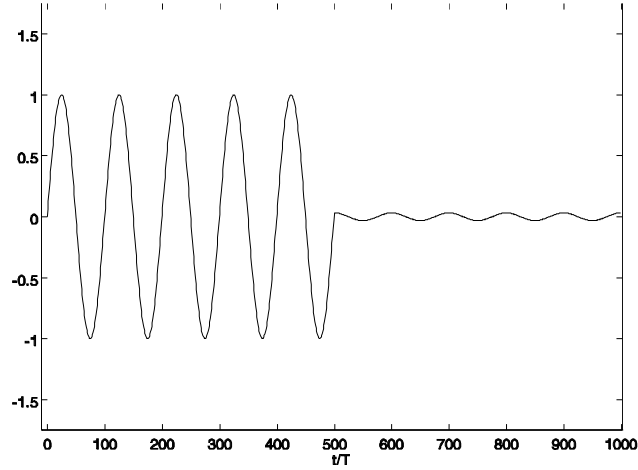


Figura 6.9: Risposta impulsiva di un filtro a campionamento di frequenza reale

($\rho^N \ll 1$). Se la quantizzazione dei coefficienti è rozza (fatta con pochi bit), sarà opportuno allontanare viepiù i poli dal cerchio unitario e quindi scegliere un minor valore per ρ . Questo comporta un maggiore degrado della caratteristica di ampiezza del filtro rispetto al caso ideale. Inoltre, per mantenere la cancellazione approssimata, devono rientrare all'interno del cerchio unitario anche gli zeri; questo si ottiene in modo naturale sottraendo dopo N campioni l'impulso di ingresso moltiplicato per ρ^N . In sostanza, la funzione di trasferimento di un filtro passa banda, robusto rispetto al problema della quantizzazione dei coefficienti, deve essere:

$$H(z) = \frac{1 - z^{-N}\rho^N}{1 - 2\rho \cos(2\pi k/N)z^{-1} + \rho^2 z^{-2}}; \quad k \neq 0, N/2; \quad \rho < 1$$

dove ρ è tanto più piccolo quanto più la quantizzazione dei coefficienti è rozza. Si osservi anche che il costo di questo filtro cresce da 1 a 3 moltiplicazioni per campione.

6.4 Filtri a campionamento in frequenza irregolare

I filtri che abbiamo visto nei paragrafi precedenti sono piuttosto semplici e poco costosi; può essere interessante renderli più complicati se i benefici superano i costi: ad esempio, per avere una struttura di lobi laterali più rapidamente decrescente a zero di quanto accade con il *sinc*, è possibile disporre in cascata due filtri a campionamento in frequenza identici o magari con due diversi valori per N . La caratteristica di ampiezza sarà il prodotto delle due e la decrescita dei lobi avverrà più rapidamente, al variare della frequenza; in sostanza, in questo caso, l'involuppo della sinusoide generata da ogni oscillatore sarà un trapezio (convoluzione di due rettangoli differenti) invece di un rettangolo.

È interessante notare che nel caso di sinusoidi reali, i due *sinc* centrati alle frequenze $\pm k/NT$ hanno una struttura di lobi laterali tale che i lobi positivi di un *sinc* si sovrappongono ai lobi negativi dell'altro comportando quindi una decrescita della funzione di trasferimento più veloce con la frequenza; infatti, il troncamento della sinusoide avviene in occasione di un passaggio per lo zero e quindi viene conservata "la continuità" della funzione.

Ancora, si può generalizzare il comportamento dei filtri a campionamento in frequenza al caso in cui la frequenza generata dall'oscillatore sia qualsiasi invece che k/NT . Allora, si possono spillare dalla memoria i campioni ritardati sia di N che di $N - 1$ campioni e si può alimentare in modo opportuno l'oscillatore in modo da troncare anche sinusoidi (eventualmente smorzate) aventi queste frequenze lanciando, dopo N campioni, delle sinusoidi sfasate dell'angolo opportuno, che può essere qualsiasi. Il polinomio a numeratore del filtro sarà quindi:

$$B(z) = 1 + az^{-(N-1)} + bz^{-N} \text{ invece che } 1 - z^{-N}$$

Indicando la pulsazione normalizzata della sinusoide con ϕ , risulta:

$$a = -\frac{\sin N\phi}{\sin \phi}$$

$$b = \frac{\sin (N-1)\phi}{\sin \phi}$$

6.4 FILTRI A CAMPIONAMENTO IN FREQUENZA IRREGOLARE

La frequenza centrale del filtro passabanda è quindi arbitraria e, per costruzione, due degli N zeri di $B(z)$ cancelleranno i due poli dell'oscillatore; gli altri zeri di $B(z)$ non saranno più disposti regolarmente attorno al cerchio unitario, ma se ne distanzieranno in modo abbastanza casuale. Peraltro, il troncamento della sinusoide **non** in corrispondenza di un passaggio per il valore 0, comporta necessariamente una decrescenza più lenta dei lobi laterali.

Capitolo 7

Decimazione ed interpolazione

7.1 Introduzione

E' spesso utile decomporre una sequenza di N campioni in un numero di sottosequenze $M \leq N$ mediante un banco di M filtri passa banda le cui frequenze centrali siano opportunamente sfalsate e le cui bande passanti (dette sottobande) non siano necessariamente eguali tra loro. Le applicazioni di questa decomposizione sono numerose; anzitutto è possibile **trattare in parallelo** le varie sottobande, utilizzando eventualmente tecniche di filtraggio adattativo, come si vedrà nel prossimo capitolo.

La decomposizione in sottobande (subband decomposition) permette di avvicinarsi alla situazione in cui lo spettro del segnale è uniforme in ogni sottobanda. Dal banco di filtri si ottengono infatti M sottosequenze che occupano bande diverse; se queste sottosequenze sono **convertite a frequenza zero** si ottengono M sottosequenze filtrate passabasso ed occupanti $1 : M$ -esimo della banda iniziale. Se poi queste sottosequenze sono massimamente decimate e cioè sono sottocampionate con rapporto $1 : M$ il loro spettro si estenderà fino alla nuova frequenza di Nyquist, $1 : M$ -esimo della frequenza di Nyquist iniziale. Le sottosequenze massimamente decimate tendono quindi ad essere bianche, **anche se la sequenza di partenza è colorata**, poiché nel segnale originale ricoprivano un intervallo di frequenza ridotto, in cui lo spettro può essere considerato costante.

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

Vi è allora il vantaggio, che sarà particolarmente apprezzabile dopo lo studio dei filtri adattativi, che i campioni successivi delle sequenze filtrate passa banda, convertite a frequenza zero, e infine massimamente decimate, sono poco correlati.

Tra le varie applicazioni della decomposizione in sottobande annoveriamo:

a) la realizzazione di filtri passabanda per l'estrazione di canali dallo spettro a radio frequenza;

b) i cosiddetti **transmultiplexer**, e cioè sistemi che trasformano un segnale multiplo a diversità di frequenza in un segnale multiplo a diversità di tempo, e viceversa;

c) la divisione in sottobande del segnale musicale, utile perchè le componenti di un segnale musicale che occupano frequenze superiori a quella di una componente forte, sono percepite con minore accuratezza. Questo effetto, detto **mascheramento**, può essere sfruttato per opportune tecniche di codifica;

d) la realizzazione numerica di circuiti utili per esempio in tecniche audio numeriche, per separare i segnali da avviare agli altoparlanti di bassa e di alta frequenza (woofer - tweeter), ecc.

Un requisito molto importante nella decomposizione in sottobande è quello della perfetta ricomposizione e cioè che un sistema di decomposizione messo in cascata con un sistema di ricomposizione dia il segnale di partenza:

i) senza buchi o picchi nella funzione di trasferimento;

ii) con la completa cancellazione del disturbo di equivocazione, generato dal sottocampionamento delle sottobande.

Infatti, la massima decimazione (sottocampionamento $1 : M$ dei segnali delle M sottobande) è necessaria se si vuole che il segnale complessivo delle M sottosequenze e cioè i segnali delle M sottobande abbia ancora lo stesso numero di campioni della sequenza iniziale.

La trasformazione di Fourier discreta è un caso estremo di questa decomposizione, in cui abbiamo il massimo numero di sottosequenze ($M = N$, tante quante i campioni) e la minima durata delle sequenze decimate (1 campione complesso per sottosequenza decimata, che è poi l'ampiezza di una sinusoide complessa). La tecnica della trasformazione di Fourier discreta permette infatti la decomposizione di una sequenza di N campioni complessi in altrettante sinusoidi complesse campionate e cioè segnali il cui spettro discreto è impulsivo e che, quando siano deci-

7.2 DECIMAZIONE E INTERPOLAZIONE

mate, si riducono ad un solo campione. All'estremo opposto, c'è il caso in cui $M = 2$; la sequenza viene quindi divisa in due sottosequenze di banda metà (l'una passa basso e l'altra passa alto) da sottocampionare $1 : 2$ e che quindi hanno $N/2$ campioni ognuna e così via. Una tecnica per operare è quella denominata *filtri a specchio in quadratura* (Quadrature Mirror Filters).

E' importante determinare **strutture efficienti** per questa decomposizione (i **filtri polifase**) ed inoltre studiare il problema della **equivocazione** che si formerà nella decimazione; infatti le bande passanti dei filtri devono sovrapporsi parzialmente, per evitare buchi spettrali dovuti alla non idealità della funzione di trasferimento; quindi, sembrerebbe inevitabile avere dell'equivocazione, a meno di sovracampionare e quindi non avere massima decimazione. Si **riesce invece a cancellare totalmente tale equivocazione** e pertanto a decomporre e ricomporre perfettamente il segnale. Purtroppo, tale cancellazione avviene solo se i vari canali non sono modificati prima di ricombinare il segnale in uscita. Altrimenti, perché la cancellazione poi avvenga effettivamente, il banco deve essere riprogettato in modo opportuno.

Contribuirà all'efficienza computazionale delle strutture filtranti il tenere conto del fatto che le uscite dei filtri passabanda vanno decimate; infatti, **non devono essere calcolati tutti i campioni delle uscite** ma solo quelli che non sarebbero eliminati. Ancora, costruendo il segnale multiplo corrispondente alla ricombinazione delle varie bande e quindi **interpolando** ogni sottosequenza ad un intervallo di campionamento inferiore, bisognerà evitare di fare calcoli inutili nella costruzione del segnale multiplo. E' inutile interpolare indipendentemente ogni sequenza e poi sommarle; si deve invece costruire il segnale combinato usando solo i campioni decimati.

7.2 Decimazione e interpolazione

Supponiamo di avere una sequenza reale campionata ad intervallo T , ma che occupa solo una frazione $1/M$ della banda $0 \div f_c$ e cioè una delle r sottobande $r = 0, \dots, M - 1$. Per semplicità considereremo $r = 0$ e quindi un segnale che occupa la banda base $\pm f_c/2M$. Per sottocampionare questa sequenza, è sufficiente prenderne 1 campione ogni M (fig. 7.1). Volendo invece partire da una sequenza qualsiasi, essa deve prima essere

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE



Figura 7.1: Decimazione di sequenze

filtrata con un filtro passa basso di banda $\pm f_c/2M$ (consideriamo la r -esima sottobanda con $r = 0$) e quindi sottocampionata. Se il filtro è lungo LT è inefficiente calcolare anche i campioni del segnale che saranno soppressi; in sostanza non è utile fare più di $\simeq L/M$ moltiplicazioni per campione di ingresso, calcolando così solo i contributi ai campioni di uscita (fig. 7.2 e fig.7.3). Nello schema a blocchi di figura, si traspongono

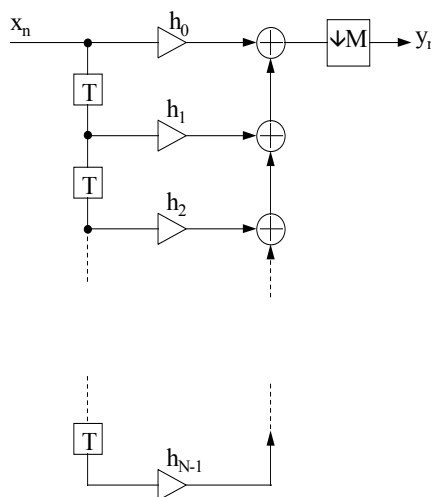


Figura 7.2: Decimazione di sequenze

le operazioni di moltiplicazione per i coefficienti $\{h_i\}$ e di decimazione per il fattore M , risparmiando così molte operazioni. Lo schema equivalente per l'interpolazione è riportato in fig. 7.4. Si osservi anche che per avere un buon filtro di banda $1/MT$ sono necessari ben più di M campioni e pertanto il numero di moltiplicazioni per campione di ingresso può essere di diverse unità. Vediamo come questo risultato modifica le figure 7.2, 7.3, 7.4, 7.5, 7.6. Indichiamo con $\{h_i\}$ la risposta all'impulso del filtro

7.2 DECIMAZIONE E INTERPOLAZIONE

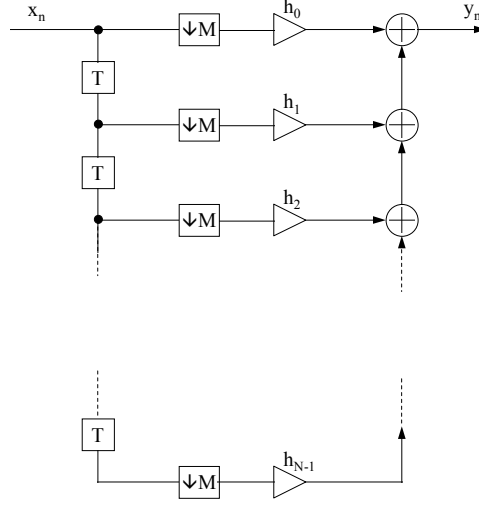


Figura 7.3: Schema efficiente per decimazione

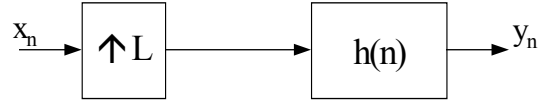


Figura 7.4: Schema per l'interpolazione

campionata ad intervallo T ; supponiamo sia quella di un filtro FIR di $L = KM$ coefficienti, con K intero, per fare in modo che ogni campione dell'uscita sia combinazione lineare di un egual numero di campioni di ingresso. Volendo, qualcuno dei coefficienti potrà essere nullo. Il segnale filtrato (campionamento pieno) è:

$$y_i = \sum_{l=0}^{L-1} h_l x_{i-l} \quad \leftrightarrow \quad H(z)X(z) \quad (7.1)$$

Per fissare le idee, poniamo $M = 2$; supponiamo che il filtro $\{h_l\}$ sia lungo $L = 4$ campioni e rimuoviamo dall'uscita i campioni di indice dispari. Al

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

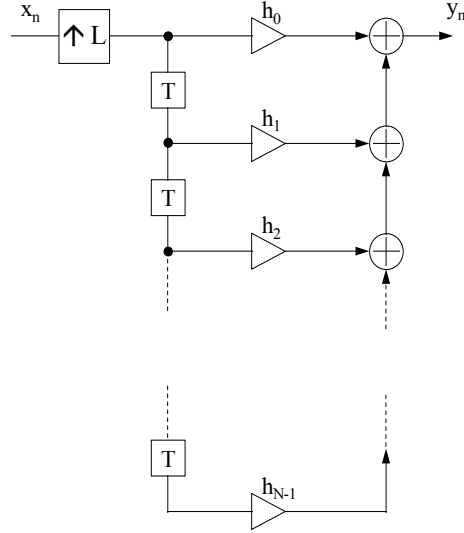


Figura 7.5: Schema di interpolazione

campione uscente di indice $2i$ contribuiscono i campioni di ingresso:

$$y_{2i} = x_{2i}h_0 + x_{2i-1}h_1 + x_{2i-2}h_2 + x_{2i-3}h_3 \quad (7.2)$$

Si vede subito che i campioni di ingresso pari sono convoluti con un "sottofiltro" realizzato con i campioni di indice pari del filtro passa basso iniziale, mentre i campioni di ingresso di indice dispari sono convoluti con il "sottofiltro" realizzato con i campioni di indice dispari. In generale, sottocampionando $1 : M$, suddivideremo l'ingresso in M sottosequenze ognuna delle quali è convoluta con un sottofiltro diverso lungo $K = L/M$ campioni realizzato sottocampionando $1 : M$ il filtro passa basso iniziale, con M pettini sfalsati tra loro di un campione. La situazione è descritta dalla formula che segue:

$$H(z) = \sum_{l=0}^{L-1} h_l z^{-l} = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} h_{kM+m} z^{-(kM+m)} = \quad (7.3)$$

$$\sum_{m=0}^{M-1} z^{-m} \sum_{k=0}^{K-1} h_{kM+m} (z^M)^{-k} = \sum_{m=0}^{M-1} z^{-m} E_m(z^M)$$

7.2 DECIMAZIONE E INTERPOLAZIONE

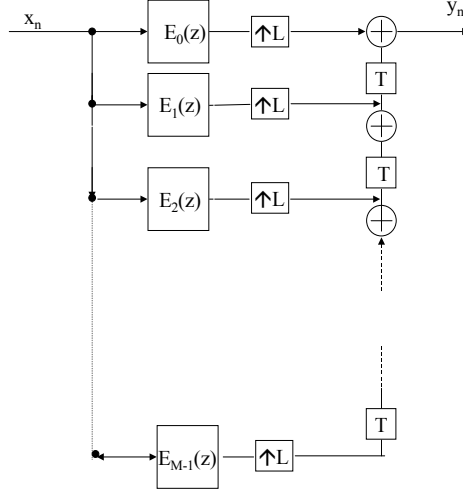


Figura 7.6: Schema efficiente per l'interpolazione

Indicando con:

$$E_m(z) = \sum_{k=0}^{K-1} e_{k,m} z^{-k}; \quad e_{k,m} = h_{kM+m} \quad (7.4)$$

In altre parole, con la risposta del filtro $\{h_l\}$, che come si ricorderà è lunga KM campioni, si formano M sottosequenze di K campioni, indicati come $\{e_{k,m}\}$, chiamati filtri polifase. Queste M sottosequenze sono ottenute dai campioni del filtro passa basso, decimati $1 : M$, con le M diverse posizioni del pettine di sottocampionamento. La struttura del filtro è indicata in fig. 7.7 (decimazione $1 : 3$, filtro di lunghezza 9 campioni) e può essere vista come un commutatore a spazzola con M posizioni diverse, che alimenta con ritardo progressivo gli M sottofiltri (a intervallo di campionamento MT) indicati come $E_m(z)$. L'uscita filtrata e decimata è la somma delle uscite degli M sottofiltri. Una struttura equivalente che minimizza il numero dei blocchi di memoria è presentata in fig. 7.8 (decimazione $1 : 2$, filtro di lunghezza 6 campioni). In particolare si osserva che servono $M - 1$ blocchi di ritardo unitario (intervallo di campionamento del segnale da decimare) per realizzare il "commutatore a spazzola" di ingresso e poi ancora $K - 1$ ritardi di M campioni (se ci riferisce all'intervallo di tempo da decimare, ma in realtà unitari se

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

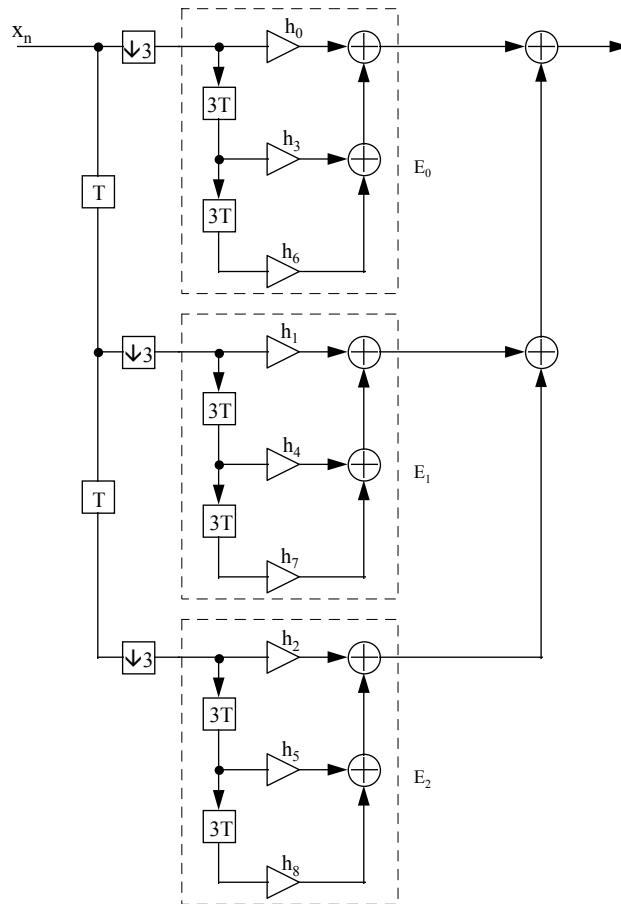


Figura 7.7: Filtro polifase per decimazione; il numero dei blocchi di memoria è pari a L

7.2 DECIMAZIONE E INTERPOLAZIONE

intesi come l'intervallo di campionamento del segnale decimato). Similmente, nella fig. 7.9 viene riportata la struttura per il filtro per effettuare un'interpolazione 1 : 3.

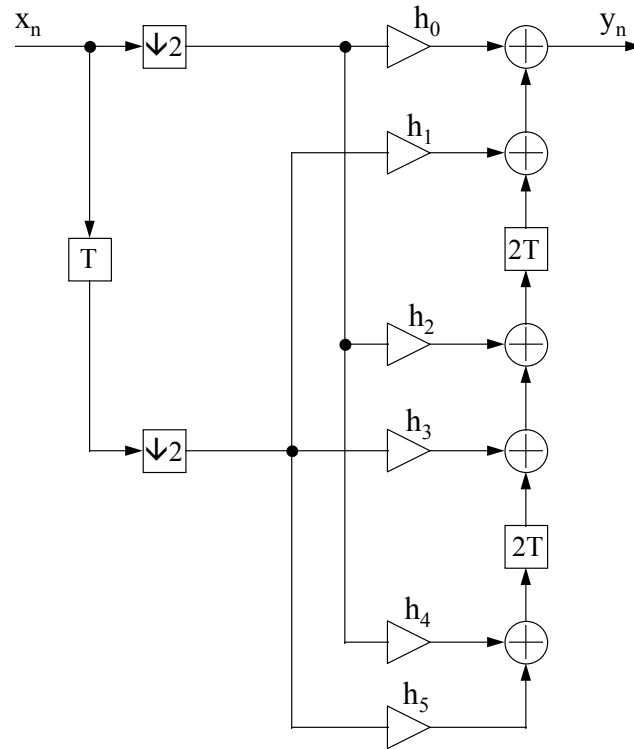


Figura 7.8: Filtro polifase per decimazione 1:2 (filtro lungo 6 campioni); ritardando i campioni dopo la moltiplicazione, il numero dei blocchi di memoria si riduce da $M - 1 + M(K - 1) = MK - 1$ come nel caso della figura precedente, a $M - 1 + K - 1$.

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

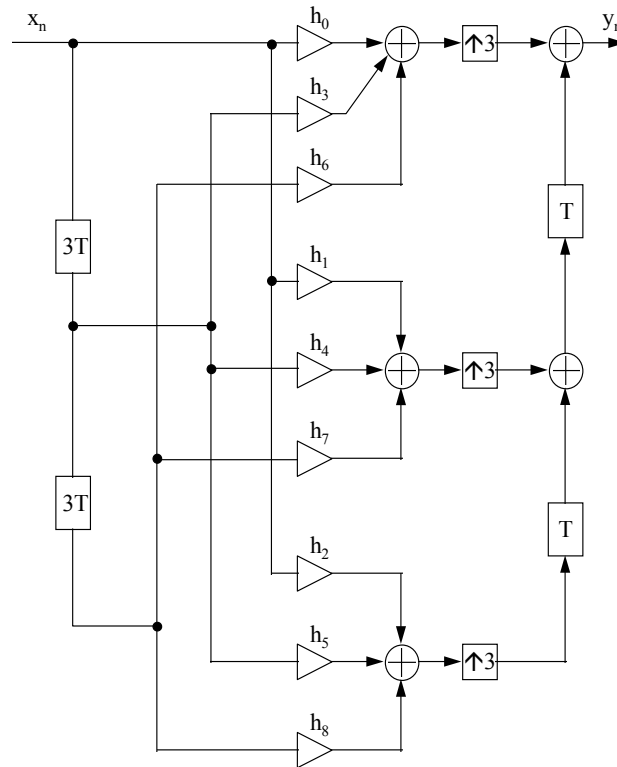


Figura 7.9: Interpolatore polifase a memoria minima; il segnale interpolando ha ritardi fino a $K - 1 = 2$ campioni; il segnale interpolato ha ritardi fino a $M - 1 = 2$ campioni (il commutatore a spazzola).

7.3 I filtri polifase

Vediamo ora il perchè del nome polifase, assegnato a questi filtri dall'inventore, M. Bellanger. Se il filtro originale $H(z)$ è un filtro passabasso di banda f_N/M , il sottocampionamento $1 : M$ provoca una replicazione della funzione di trasferimento fino ad ottenere in pratica, con le sovrapposizioni delle repliche, un ricoprimento con ampiezza pressochè uniforme di tutto l'intervallo di frequenze fino alla frequenza di Nyquist f_N (fig. 7.10); quindi, i filtri $E_m(z^M)$ sono filtri passatutto, e pertanto sfasatori

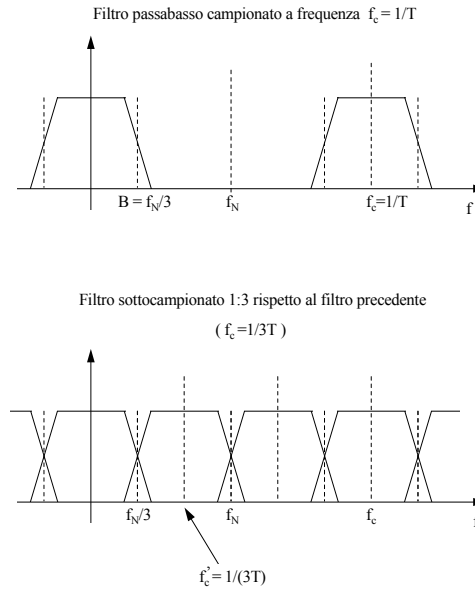


Figura 7.10: Sottocampionamento di un segnale

puri. Ma anche i sottofiltri $E_m(z)$ hanno spettro circa costante e sono quindi filtri passatutto, nella banda $0 \div 2\pi/MT$. Si osservi che in questa ultima espressione di trasformata z compare z e non z^M ; le sequenze $E_m(z)$ corrispondono ai sottofiltri **dopo** la decimazione.

Se il filtro $H(z)$ è approssimativamente un filtro passa basso ideale, l'effetto di uno dei filtri sottocampionati $z^{-m}E_m(z^M)$ è, a parte il filtraggio passabasso seguito dalla decimazione, un ritardo della sequenza da decimare di m campioni al ritmo originale sulla sequenza decimanda;

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

peraltro, a causa della decimazione, il filtro lascia anche passare le componenti spettrali al di fuori della banda $0 \div 2\pi/MT$. Lo stesso sottofiltro, dopo l'eliminazione dei campioni nulli, e quindi dopo il sottocampionamento, ha $E_m(z)$ come trasformata z e fa ancora la stessa operazione, e quindi si comporta ancora come un ritardo temporale che è ora, al ritmo ridotto $1 : M$, di una frazione di campioni $m/M \leq 1$; in particolare, si deduce che se l'ingresso è costante, il risultato del filtraggio con uno qualsiasi dei filtri $E_m(z)$, è ancora costante (per una costante, un ritardo è irrilevante).

Allo stesso risultato si può arrivare in modo intuitivo considerando il caso dell'interpolazione $M : 1$ di una sequenza. Infatti, la sequenza interpolata può essere ottenuta generando, a partire dalla sequenza di partenza a bassa frequenza di campionamento $f_c = 1/T$, $M - 1$ sequenze ad intervallo T , ritardate progressivamente di $kT/(M - 1)$, con $0 < k < M$. La sequenza interpolata è poi ottenuta dall'interallacciamento delle M sequenze con i vari ritardi (l'azione del commutatore a spazzola). I filtri polifase sono pertanto dei ritardatori puri di $kT/M - 1$, come già osservato.

7.4 Banchi di filtri

7.4.1 La matrice di modulazione

Si può capire ora come avviene la formazione di un intero banco di M filtri; utilizzeremo la sovrapposizione degli effetti supponendo di avere un segnale il cui spettro è diverso da zero solo nella sottobanda r -esima; se esso viene filtrato con i vari filtri $z^{-m}E_m(z^M)$; $m = 0, \dots, M - 1$, l'uscita totale deve essere nulla, visto che il filtro complessivo $H(z)$ è passa basso. Infatti accade proprio questo, e ce ne rendiamo conto nel seguente modo: il segnale alla frequenza centrale dell' r -esima sottobanda è la sinusoide $\exp(j\frac{2\pi}{M}nr)$; le m sottosequenze (gli m segnali provenienti dal sottocampionamento $1 : M$) sono delle costanti pari a $\exp(j\frac{2\pi}{M}mr) = W_M^{mr}$, utilizzando la notazione usata per lo studio della Trasformazione Discreta di Fourier; nel dominio sottocampionato, l'uscita totale sarà la somma delle uscite dei sottofiltri $E_m(z)$ ai cui ingressi sono applicate le

7.4 BANCHI DI FILTRI

costanti W_M^{mr} e cioè:

$$U = \sum_{m=0}^{M-1} W_M^{mr} E_m(z=1) = 0 \text{ se } r \neq 0$$

Se però moltiplicassimo i coefficienti dei vari sottofiltri $E_m(z)$ per il fattore W_M^{-mr} compenseremmo le rotazioni W_M^{mr} e la somma degli M sottofiltri ci darebbe una costante; in sostanza otterremmo la r -esima sottobanda, convertita a frequenza zero e decimata. Nasce allora la struttura di fig. (7.11) combinazione di un filtro polifase e di una DFT di M

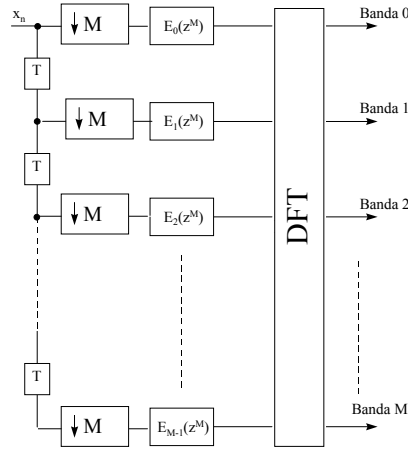


Figura 7.11: Banco di M filtri come combinazione di un filtro polifase e FFT

campioni, che viene realizzata ad ogni campione decimato. Le M uscite della FFT sono le M sottobande, già convertite a frequenza 0 e decimate massimamente.

Un altro modo di vedere la stessa cosa è il seguente: possiamo facilmente ottenere la conversione a frequenza zero della r -esima sottobanda; basta moltiplicare il segnale x_n per una sinusoide campionata posta alla frequenza centrale della sottobanda e cioè per il fattore W_M^{-nr} ma anche, tenendo conto delle periodicità, per il fattore W_M^{-mr} dove l'indice

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

m varia tra 0 a $M - 1$ e poi ancora tra 0 a $M - 1$ ecc. ($W_M^{Mr} = 1$); per ottenere tutte le sottobande, pertanto, si devono moltiplicare i campioni della M -pla x_{n-m} ; $m = 0, \dots, M - 1$, per la matrice di Fourier che funge da **Matrice di Modulazione**. C'è da osservare che con questo modo di vedere, la conversione a frequenza zero dovrebbe essere fatta **prima** del filtraggio; in realtà, le due operazioni di conversione e filtraggio possono essere trasposte poichè la sottosequenza m -esima che affluisce allo stesso filtro polifase $z^{-m}E_m(z^M)$ per contribuire alla sottobanda r -esima deve venire moltiplicata sempre per lo stesso numero complesso W_M^{-mr} ; in altre parole, la posizione del commutatore a spazzola che alimenta il filtro polifase m -esimo corrisponde sempre allo stesso fattore complesso W_M^{-mr} . La moltiplicazione può quindi essere eseguita **dopo** il filtraggio e pertanto si arriva di nuovo allo schema di fig. (7.11).

Procedendo analiticamente si può partire dalla trasformata z delle M sequenze periodiche di periodo M , $C_{M,\lambda}(n)$ costituite da un campione unitario posizionato al campione λ , e da $M - 1$ zeri. Esse possono essere viste come antitrasformate di Fourier di un ritardo di λ campioni, per sequenze periodiche su M campioni, e cioè:

$$C_{M,\lambda}(n) = \frac{1}{M} \sum_{m=0}^{M-1} W_M^{-\lambda m} W_M^{nm}$$

La moltiplicazione di una sequenza $x(n)$ per la sequenza $C_{M,\lambda}(n)$ (e quindi la decimazione di una sequenza ritardata di λ campioni) porta alla seguente trasformata z :

$$\begin{aligned} x(n)C_{M,\lambda}(n) &\rightarrow \sum_n x(n)C_{M,\lambda}(n)z^{-n} = \frac{1}{M} \sum_n \sum_{r=0}^{M-1} W_M^{-\lambda r} W_M^{nr} x(n)z^{-n} = \\ &= \frac{1}{M} \sum_{r=0}^{M-1} W_M^{-\lambda r} \sum_n x(n)(zW_M^{-r})^{-n} = \frac{1}{M} \sum_{r=0}^{M-1} W_M^{-\lambda r} X(zW_M^{-r}) \end{aligned}$$

e quindi:

- la trasformata z di una sequenza decimata $1 : M$ è la somma di M repliche della trasformata z della sequenza originale;
- la r -esima replica della trasformata z è ruotata nel piano z di un angolo W_M^{-r} e sfasata di un angolo $-2\pi\lambda r/M$, dove λ è il ritardo della

7.4 BANCHI DI FILTRI

sequenza decimata. Un risultato corrispondente era stato ottenuto nel capitolo 1, equazione 1.6. Adesso possiamo arrivare alla decomposizione polifase di un banco di filtri, utilizzando l'equazione dei filtri polifase che riscriviamo per comodità:

$$H(z) = \sum_{n=0}^{L-1} h_n z^{-n} = \sum_{k=0}^{K-1} \sum_{\ell=0}^{M-1} h_{kM+\ell} z^{-(kM+\ell)} = \quad (7.5)$$

$$\sum_{\ell=0}^{M-1} z^{-\ell} \sum_{k=0}^{K-1} h_{kM+\ell} (z^M)^{-k} = \sum_{\ell=0}^{M-1} z^{-\ell} E_{\ell}(z^M)$$

$$E_{\ell}(z) = \sum_{k=0}^{K-1} e_{k,\ell} z^{-k}; \quad e_{k,\ell} = h_{kM+\ell} \quad (7.6)$$

Trasformare il filtro $H(z)$ da passa basso in passa r -esima sottobanda coincide con il ruotare la trasformata z del filtro di un angolo $2\pi r/M$; ruotiamo quindi $H(z)$ di un angolo W_M^{-r} nel piano z ; questo equivale a sostituire z con zW_M^{-r} , oppure a moltiplicare la risposta all'impulso del filtro per la sequenza W_M^{-nr} (una sinusoide campionata che attua la conversione di frequenza). Il risultato è un filtro passa banda che lascia passare la sottobanda r -esima

$$\begin{aligned} H_r(z) &= H(zW_M^{-r}) = \sum_{\ell=0}^{M-1} z^{-\ell} W_M^{-\ell r} E_{\ell}(z^M W_M^{-rM}) = \\ &= \sum_{\ell=0}^{M-1} z^{-\ell} W_M^{-\ell r} E_{\ell}(z^M) \end{aligned}$$

dove abbiamo utilizzato il fatto che $W_M^{-rM} = 1$ e cioè che la sinusoide modulante sottocampionata $1 : M$ diventa una costante complessa. In conclusione, il filtro che passa la banda r -esima si ottiene combinando le uscite dei filtri polifase $\ell = 1, \dots, M$ con i coefficienti $W_M^{-\ell r}$. Per ottenere tutte le bande assieme si utilizza la moltiplicazione per la matrice della trasformazione di Fourier discreta in base M .

La realizzazione efficiente della moltiplicazione per la matrice di Fourier (e pertanto la trasformazione di Fourier) attraverso la FFT, permette poi di realizzare il banco di filtri in modo efficiente con K moltiplicazioni

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

per campione di ingresso a cui vanno sommate $\frac{1}{2} \log_2 M$ moltiplicazioni per campione di uscita; questi, contando tutte le sottobande decimate, sono tanti quanti i campioni di ingresso. In totale pertanto vanno effettuate $L/M + \frac{1}{2} \log_2 M$ moltiplicazioni per campione della sequenza (decomposta o no in sottobande), dove M è il numero delle sottobande. Per la ricomposizione, vi è una struttura perfettamente simmetrica.

7.4.2 La ricostruzione perfetta

Fino a questo momento, non abbiamo discusso della realizzazione del filtro passabasso $H(z)$, se non per dire che deve essere lungo KM campioni. Vogliamo ora definirne le regole di progetto. I requisiti sono di due tipi diversi: da un lato si desidera un ottimo comportamento filtrante e fase lineare, cosicchè la decomposizione in sottobande sia accurata e non vi siano sovrapposizioni o buchi spettrali; ancora però si desidera che le componenti degli spettri di ingresso, equivocate da una sottobanda all'altra per la massima decimazione, siano poi ricostruite correttamente. In altre parole, pur con una buona separazione delle bande, si desidera una ricombinazione corretta della sequenza di uscita e cioè:

- nessuna equivocazione pur con decimazione $1 : M$;
- guadagno globale costante;
- fase lineare.

E' ovvio che questi requisiti (la cosiddetta ricostruzione perfetta) sono assai pesanti per il filtro ma è possibile dimostrare che sono ottenibili, sia pure a costi crescenti via via che si approssima la situazione ideale. Si vedrà che la ricostruzione perfetta corrisponde alla compensazione del rumore di equivocazione dovuto alla decimazione $1 : M$ e lasciato passare dal filtraggio, necessariamente non ideale, che viene effettuato prima della decimazione. Si osservi però che, non appena il contenuto p.e. della banda r -esima viene modificato o magari semplicemente amplificato in modo differente da quello delle bande adiacenti, proprio per motivare in qualche modo la decisione di effettuare la decomposizione, tale compensazione scompare e quindi la ricostruzione non è più perfetta. In altre parole, non appena la decomposizione viene utilizzata per fare qualcosa di utile, la proprietà di ricostruzione perfetta scompare. Bisognerà addirittura riprogettare il sistema multibanda in queste nuove condizioni (*rifare tutto daccapo*). Se ne evince che è piuttosto difficile (anche se non impossibile) costruire filtri adattativi multibanda che garantiscono

la proprietà della ricostruzione perfetta, sia pure in modo solo asintotico. Su questo punto si ritornerà nelle sezioni finali del capitolo.

7.4.3 Banchi di due filtri

Vediamo anzitutto come sia possibile realizzare un banco di due filtri, $H_L(\omega)$ e $H_H(\omega)$ rispettivamente passa basso e passa alto che permetta la ricostruzione corretta anche con la massima decimazione. Questi due filtri sono chiamati filtri a specchio in quadratura (Quadrature Mirror Filter QMF) per i motivi che vedremo. Lo schema a blocchi del sistema è presentato in figura 7.12. Indichiamo rispettivamente con $X(\omega)$, $X_L(\omega)$,

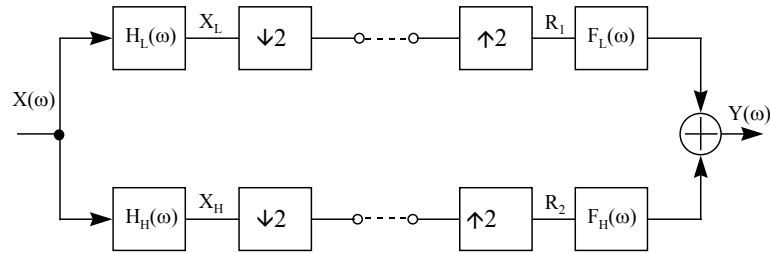


Figura 7.12: Schema a blocchi di un filtro a specchio in quadratura

$X_H(\omega)$ le trasformate di Fourier del segnale e delle sue due componenti passa basso e passa alto.

$$X_L(\omega) = X(\omega)H_L(\omega); \quad X_H(\omega) = X(\omega)H_H(\omega) \quad (7.7)$$

I segnali dopo decimazione ed interpolazione corrispondono alla replica in frequenza di $X_L(\omega)$, $X_H(\omega)$ a distanza $\omega_0 = \pi/T$:

$$R_1(\omega) = X_L(\omega) + X_L(\omega - \omega_0) \quad (7.8)$$

$$R_2(\omega) = X_H(\omega) + X_H(\omega - \omega_0) \quad (7.9)$$

La replica in frequenza è infatti dovuta al sottocampionamento 1 : 2 ed all'inserzione successiva di zeri. Il risultato della ricombinazione con due filtri $F_L(\omega)$, $F_H(\omega)$ è:

$$\hat{X}(\omega) = R_1(\omega)F_L(\omega) + R_2(\omega)F_H(\omega) = \quad (7.10)$$

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

$$\begin{aligned}
 &= X(\omega)[H_L(\omega)F_L(\omega) + H_H(\omega)F_H(\omega)] + \\
 &+ X(\omega - \omega_0)[F_L(\omega)H_L(\omega - \omega_0) + F_H(\omega)H_H(\omega - \omega_0)] = \\
 &X(\omega)T(\omega) + X(\omega - \omega_0)A(\omega)
 \end{aligned}$$

Vorremmo pertanto che:

$$T(\omega) = H_L(\omega)F_L(\omega) + H_H(\omega)F_H(\omega) = e^{-j\omega kT} \quad (7.11)$$

per avere guadagno costante e fase lineare; inoltre vorremmo:

$$A(\omega) = F_L(\omega)H_L(\omega - \omega_0) + F_H(\omega)H_H(\omega - \omega_0) = 0 \quad (7.12)$$

per non avere equivocazione. In teoria, ad ogni frequenza, abbiamo solo due equazioni per determinare quattro incognite (i valori delle funzioni di trasferimento dei quattro filtri) e pertanto abbiamo ampie possibilità di soluzione; tuttavia, si cerca una soluzione pratica.

7.4.4 I filtri in quadratura a specchio (QMF)

Imponiamo ora dei requisiti sui filtri $H_L(\omega)$, $H_H(\omega)$; anzitutto imponiamo che siano FIR, reali, lunghi N campioni e simmetrici e cioè a fase zero a parte un ritardo di $(N - 1)/2$ campioni. Per ogni zero della funzione di trasferimento, vi sarà sia lo zero coniugato perchè la risposta all'impulso è reale, sia lo zero reciproco, perchè essa è simmetrica. Imponendo quindi la simmetria coniugata rispetto all'asse reale del piano z si ha:

$$H_L(z) = H_L^*(z^*); \quad H_H(z) = H_H^*(z^*) \quad (7.13)$$

Imponiamo poi che i due filtri passa basso e passa alto siano ottenibili l'uno dall'altro con una semplice traslazione lungo l'asse delle frequenze e cioè:

$$H_H(\omega) = H_L(\omega - \omega_0); \quad \rightarrow \quad h_{H,n} = (-1)^n h_{L,n} \quad (7.14)$$

In termini di trasformata z si può scrivere che:

$$H_H(z) = H_L(ze^{-j\omega_0 T}) = H_L(-z) \quad (7.15)$$

e quindi la trasformata z di $H_H(z)$ è la riflessione speculare della trasformata z di $H_L(z)$ rispetto all'asse immaginario del piano z (in fig. 7.9 , 7.14 sono riportate la caratteristica di ampiezza e gli zeri di un filtro

7.4 BANCHI DI FILTRI

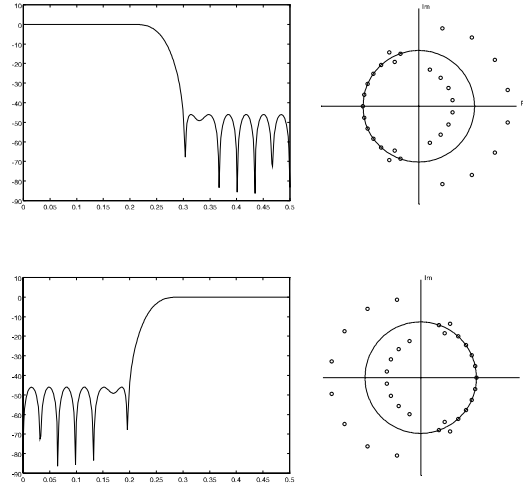


Figura 7.13: Filtri in quadratura a specchio

simmetrico di 32 campioni ottimizzato come indicato in [Chen, C. K., *Signal Processing*, Dicembre 1995, pp. 275]). Il nome di filtri a specchio in quadratura (QMF) nasce appunto dalla doppia simmetria rispetto alle rette del piano z : $Re(z) = 0$; $Im(z) = 0$. Per risparmiare sui coefficienti del filtro potremmo ancora imporre che ogni filtro sia **specularmente simmetrico** rispetto alla frequenza $f_c/4$;

$$H_L(\omega) + H_L(\omega_0 - \omega) = H_L(0); \quad H_H(\omega) + H_H(\omega_0 - \omega) = H_H(0); \quad (7.16)$$

Infatti, questo implica che la funzione di trasferimento sommata con la sua replica spettrale è una costante. Avevamo già visto questa condizione discutendo della ricostruibilità di costanti sovrapponendo le risposte all'impulso degli interpolatori di vario ordine; imporre che la somma della trasformata e della sue n repliche spettrali sia costante implica che nel dominio coniugato (nei tempi in questo caso) l'antitrasformata si annulli in corrispondenza del pettine di campionamento relativo alla replicazione, e quindi un campione ogni n . Se $n = 2$, circa metà dei campioni della risposta all'impulso (i campioni di posto pari tranne quello nell'origine) sono nulli con vantaggi di costo notevoli. I filtri con questa simmetria sono anche chiamati filtri di Nyquist e sono spesso utilizzati per la trasmissione numerica. Essi garantiscono infatti limitata o nulla inter-

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

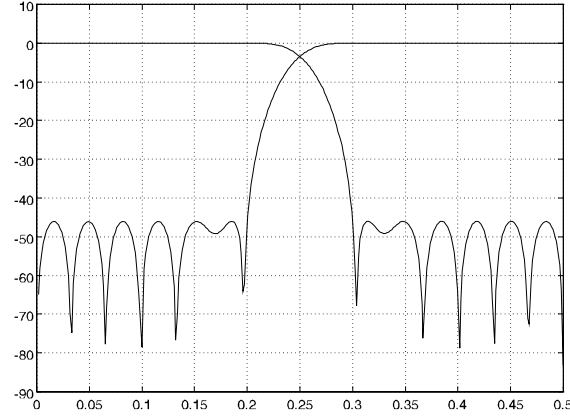


Figura 7.14: Simmetria delle funzioni di trasferimento (lo specchio).

ferenza intersimbolica, quando i simboli sono trasmessi in corrispondenza dei regolari passaggi per lo zero della risposta all'impulso.

7.4.5 Equivocazione e ricostruzione perfetta

Vediamo ora le conseguenze delle scelte fatte nel paragrafo precedente per quanto riguarda la perfetta ricostruzione; si decida inoltre che anche per i filtri $F_L(\omega)$, $F_H(\omega)$ valgano le stesse condizioni che per i filtri $H_L(\omega)$; $H_H(\omega)$, con la differenza che ora

$$F_H(\omega) = -F_L(\omega - \omega_0) \quad (7.17)$$

Scegliendo infine $F_L(\omega) = H_L(\omega)$ si ha:

$$T(\omega) = H_L^2(\omega) - H_L^2(\omega - \omega_0) = e^{-j\omega kT} \quad (7.18)$$

$$\begin{aligned} A(\omega) &= H_L(\omega - \omega_0)F_L(\omega) + H_H(\omega - \omega_0)F_H(\omega) = \\ &= H_L(\omega)H_L(\omega - \omega_0) - H_L(\omega - \omega_0)H_L(\omega) = 0 \end{aligned}$$

Si vede cioè che, sistematicamente, la ricostruzione non contiene rumore di equivocazione. Bisognerebbe ancora imporre che la funzione di trasferimento $T(\omega)$ sia unitaria, perchè la ricostruzione sia perfetta; si dimostra però che una soluzione esatta che assicuri valore perfettamente unitario

7.4 BANCHI DI FILTRI

per $|T(\omega)|$, porta a filtri $H_L(\omega)$ poco selettivi; il lettore, per esempio, consideri il caso:

$$H_L(z) = 0.5(1 + z^{-1}) \rightarrow H_H(z) = .5(1 - z^{-1}); T(z) = z^{-1}; A(z) = 0$$

Nelle zone in cui $H_L(\omega) \simeq 1 \rightarrow H_H(\omega) \simeq 0$; $|T(z)|$ è quindi circa unitario come desiderato; si tratta solo di ottenere un buon comportamento nella banda di transizione e questo è possibile con ottima approssimazione. Si osserva poi che, poichè $H_L(\omega)$ è FIR, lungo N campioni, a fase lineare $\angle H_L(\omega) = -\omega T(N-1)/2$ è necessario che sia N **pari** per evitare che $T(\omega)$ sia sistematicamente nulla alla frequenza $f_N/2$. Infatti, usando la ([?])e la ([?]):

$$T(\omega) = [|H_L(\omega)|^2 - (-1)^{N-1}|H_L(\omega - \omega_0)|^2]e^{-j\omega T(N-1)} \quad (7.19)$$

Si deve imporre quindi che N sia pari e poi si deve anche imporre che:

$$[|H_L(\omega)|^2 + |H_L(\omega - \omega_0)|^2] = |T(\omega)| \simeq 1 \quad (7.20)$$

Quest'ultima condizione, insieme al requisito di una buona attenuazione in banda oscura, che implica un buon comportamento in banda passante per le relazioni tra comportamento in banda passante e in banda oscura stabilite in precedenza, si realizza appunto con un'opportuna ottimizzazione dei coefficienti dei filtri.

Si può ora capire come si ottiene l'eliminazione del rumore di equivocazione (alias) (fig. 7.15) nella ricostruzione di $\hat{X}(\omega)$, utilizzando i due filtri $F_L(\omega)$; $F_H(\omega)$; gli spettri dei due segnali decimati e riportati al ritmo di campionamento iniziale con l'aggiunta di zeri, $R_1(\omega)$, $R_2(\omega)$ sono sempre visibili in fig. 7.15. Nella ricombinazione, ad esempio l'effetto del filtro $F_H(\omega)$ su $R_2(\omega)$ sarà quello di lasciare passare ulteriore equivocazione che **cancella** l'equivocazione lasciata in precedenza passare dal filtro $H_L(\omega)$. Infatti, a causa della rotazione di fase dovuta al ritardo, se si **somma** l'uscita del canale passa alto con quella del canale passa basso, si **sottrae** allo stesso tempo il segnale equivocato che ha viaggiato attraverso il filtro ad un'altra frequenza e quindi con un diverso sfasamento dovuto al ritardo. Quindi, i rumori di equivocazione lasciati passare dal filtro di ricostruzione, per le simmetrie, **sono cancellati** dai rumori di equivocazione dell'altro canale lasciati passare dai filtri di decimazione, ed alla fine risulta $A(\omega) = 0$. Si comprende meglio quanto osservato in

CAPITOLO 7 DECIMAZIONE ED INTERPOLAZIONE

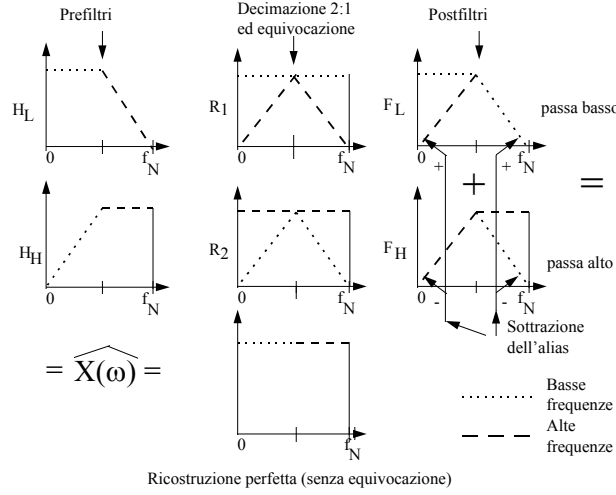


Figura 7.15: Ricostruzione perfetta (sottrazione delle componenti equivocate)

una sezione precedente, e cioè che se poi i due canali fossero amplificati in modo diseguale, tanto per motivare in qualche modo la decomposizione, tale compensazione non avverrebbe più.

La realizzazione polifase del filtro $H_L(\omega)$

Il filtro $H_L(\omega)$ ed anche il corrispondente filtro passa alto $H_H(\omega)$ possono venire realizzati con la tecnica polifase. Avremo allora:

$$H_L(z) = H_0(z^2) + z^{-1}H_1(z^2) \quad (7.21)$$

dove $H_i(z)$; $i = 0, 1$ sono i sottofiltri ottenuti dal sottocampionamento di $H_L(z)$. Si osservi che, nel caso si faccia la scelta di un progetto di filtro simmetrico rispetto a $f_c/4$, a causa dei numerosi passaggi per lo zero della forma d'onda dovuti alla simmetria appunto rispetto alla frequenza $f_c/4$, il filtro $H_1(z^2)$ viene ad avere solo un campione diverso da zero. Di nuovo, si trovano i campioni del filtro $H_0(z^2)$ con una procedura di ottimizzazione, imponendo un buon comportamento filtrante. Ancora, l'equivocazione viene cancellata sistematicamente per l'adozione della ([?]).

7.5 Banchi di M filtri

Abbiamo visto, nel caso di un banco di due filtri, che non è solo possibile ottenere sistematicamente la rimozione del rumore di equivocazione ma anche si può avere una funzione di trasferimento unitaria. Per estendere queste proprietà a banchi di più filtri si può operare in vario modo:

- formazione di alberi binari di filtri, ripetendo l'operazione di divisione in due bande di ogni banda;
- tecniche approssimate (filtri a quasi perfetta ricostruzione);
- tecniche esatte di perfetta ricostruzione.

la prima soluzione è di facile ottenimento; la sequenza iniziale è prima divisa in due sottosequenze ma poi viene ulteriormente frazionata in quattro e così via fino a raggiungere il numero desiderato di filtri, eventualmente anche con larghezze di banda diverse tra filtro e filtro. Questa tecnica ha il difetto di mettere in cascata numerosi filtri FIR (quelli di tutte le decomposizioni e ricomposizioni) e quindi può portare a ritardi globali inaccettabili; ovviamente il ritardo temporale maggiore è quello dovuto ai filtri di banda più stretta, relativi alla massima decimazione in assoluto; ma poi si deve tenere conto anche degli altri filtri in cascata con risultati spesso inaccettabili.

Piuttosto, si utilizzano tecniche per sintetizzare banchi di filtri in cui la rimozione del rumore di equivocazione è solo approssimata; si suppone cioè che solo il rumore prodotto dalla decimazione delle bande adiacenti sia significativo, mentre i segnali che transitano per filtri più lontani in frequenza non danno contributo di rumore. In questo modo il progetto dei filtri si semplifica.

Infine, la tecnica di perfetta ricostruzione si può estendere a banchi di M filtri. Tuttavia, i filtri così sintetizzati sono più complessi, a parità di selettività. Se poi si tiene conto del fatto che a causa della quantizzazione di coefficienti e segnali si ha sempre del rumore dovuto al filtraggio, si comprende come finora siano utilizzati preferenzialmente i filtri a quasi perfetta ricostruzione.

Capitolo 8

Effetti di quantizzazione

8.1 Strutture dei filtri numerici

La realizzazione pratica dei filtri con moltiplicatori e sommatore binari con un numero limitato di cifre ha notevoli effetti sul loro comportamento, come già osservato alla fine del precedente capitolo nel caso di filtri a campionamento in frequenza. In questo capitolo illustreremo anzitutto le forme con cui si possono realizzare efficientemente dei filtri numerici. Poi vedremo che la realizzazione efficiente che minimizza il numero delle moltiplicazioni non è sempre la soluzione più valida. Un qualsiasi filtro FIR o IIR è rappresentato dall'equazione alle differenze finite:

$$y_i = \sum_{n=0}^N b_n x_{i-n} + \sum_{m=1}^M a_m y_{i-m} \quad (8.1)$$

Una realizzazione pratica è presentata nella fig. 8.1 dove il blocco che realizza gli zeri viene prima del blocco che realizza i poli; se invece si fa precedere il blocco dei poli, (fig. 8.2) si vede che il contenuto delle memorie è identico nei due blocchi; evitando la duplicazione, si arriva allo schema di figura 8.3 che realizza l'equazione (8.1) con il minimo di blocchi di memoria, di moltiplicazioni e di interconnessioni. Non è necessario fare precedere i poli; è possibile fare precedere gli zeri ritardando i segnali moltiplicati invece che moltiplicare i segnali ritardati. In sostanza, è possibile scambiare il senso delle frecce nella figura 8.3 e si ritrova la stessa funzione di trasferimento purché si scambino i sensi di propagazione nelle memorie e i coefficienti a_i con i b_i . Il sistema poi si semplifica quando il

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

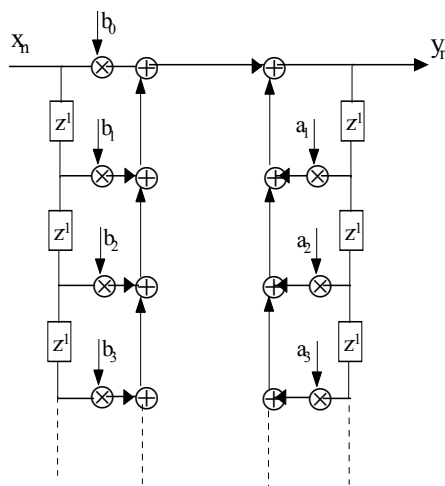


Figura 8.1: Realizzazione di un filtro FIR o IIR con zeri attuati prima dei poli

filtro FIR da realizzare ha fase zero (simmetria dei coefficienti rispetto al tempo 0) ma anche un ritardo per permetterne la realizzazione causale; esso ha quindi fase lineare e pertanto c'è simmetria dei coefficienti (fig. 8.4) rispetto ad un istante centrale. Se il numero dei campioni del filtro è dispari, vi è anche il campione centrale (fig. 8.4b). E' possibile ridurre il numero delle moltiplicazioni alla metà (o quasi, per N dispari) con la struttura di figura 8.4 in cui si calcola prima la somma in parentesi e poi il prodotto.

$$y'_i = \sum_{n=0}^{N/2-1} b_n (x_{i-n} + x_{i-N+1+n}) \quad (8.2)$$

8.2 Approssimazioni nella realizzazione di filtri

8.2.1 La quantizzazione dei coefficienti

Se ora le moltiplicazioni che compaiono nello schema di figura 8.3 vengono realizzate con moltiplicatori binari con un numero finito di cifre la

8.2 APPROSSIMAZIONI NELLA REALIZZAZIONE DI FILTRI

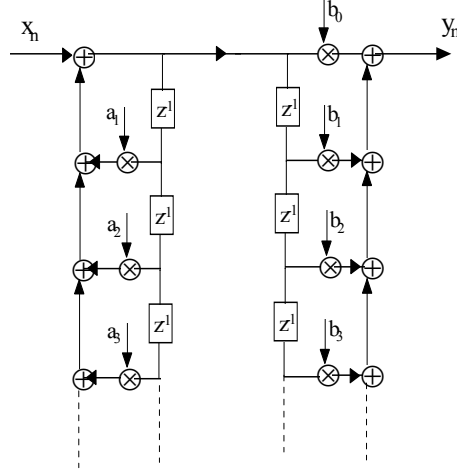


Figura 8.2: Realizzazione di un filtro FIR e IIR con poli attuati prima degli zeri

posizione dei poli e degli zeri delle funzioni di trasferimento cambia. La dipendenza delle radici di un polinomio qualsiasi dalla variazione dei coefficienti (sensitività) è calcolata come segue. Supponiamo per semplicità di avere un polinomio nella variabile z e non z^{-1} :

$$A(z) = a_0 + \sum_{m=1}^N a_{N-m} z^m = a_N \prod_{n=1}^N (z - z_n); \quad a_0 = a_N \prod_{n=1}^N (-z_n) \quad (8.3)$$

e di alterarne i coefficienti a causa della quantizzazione; vogliamo calcolare le sensitività come derivate parziali $\partial z_m / \partial a_k$ e cioè la variazione della posizione dello zero m -esimo in funzione della variazione del coefficiente k -esimo. Per fare ciò, possiamo osservare che:

$$\frac{\partial A(z)}{\partial z_h} = -a_N \prod_{n \neq h}^N (z - z_n); \quad \frac{\partial A(z)}{\partial a_k} = z^{N-k}; \quad (8.4)$$

Inoltre,

$$\frac{\partial A(z)}{\partial a_k} = z^{N-k} = \sum_{m=1}^N \frac{\partial A(z)}{\partial z_m} \frac{\partial z_m}{\partial a_k} = -a_N \sum_{m=1}^N \prod_{n \neq m}^N (z - z_n) \frac{\partial z_m}{\partial a_k} \quad (8.5)$$

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

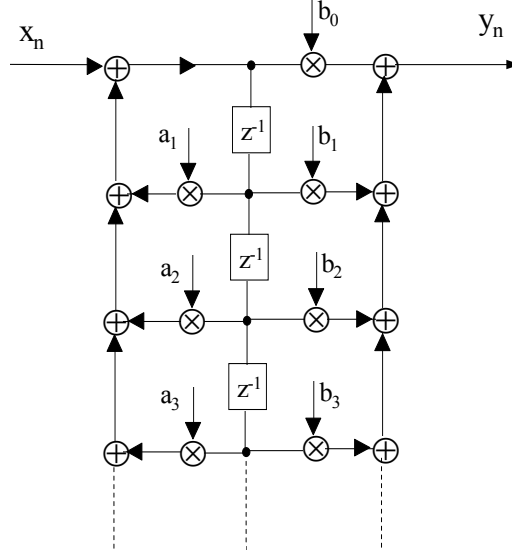


Figura 8.3: Realizzazione usuale dei filtri di fig 8.1 e fig 8.2

Ponendo $z = z_h$, tutti i termini della sommatoria si annullano tranne quello che non ha tra i fattori $(z - z_h)$,

$$z_h^{N-k} = -a_N \frac{\partial z_h}{\partial a_k} \prod_{n \neq h}^N (z_h - z_n) \quad (8.6)$$

e quindi:

$$\left| \frac{\partial z_h}{\partial a_k} \right| = \left| \frac{z_h^{N-k}}{a_N \prod_{n \neq h}^N (z_h - z_n)} \right| \quad (8.7)$$

Si vede pertanto che la sensitività e cioè il fattore di proporzionalità che lega lo spostamento di una radice alla variazione di un coefficiente è tanto maggiore quanto maggiore è il grado del polinomio e quanto più vicine sono le altre radici alla radice considerata. Quindi, per avere la minima sensitività sarà opportuno realizzare blocchi che contengano solo polinomi di grado minimo e cioè due per evitare di utilizzare aritmetica complessa. Quindi sistemi complicati verranno realizzati come cascate di forme biquadratiche e cioè funzioni di trasferimento che realizzano una coppia di poli ed una coppia di zeri.

8.2 APPROSSIMAZIONI NELLA REALIZZAZIONE DI FILTRI

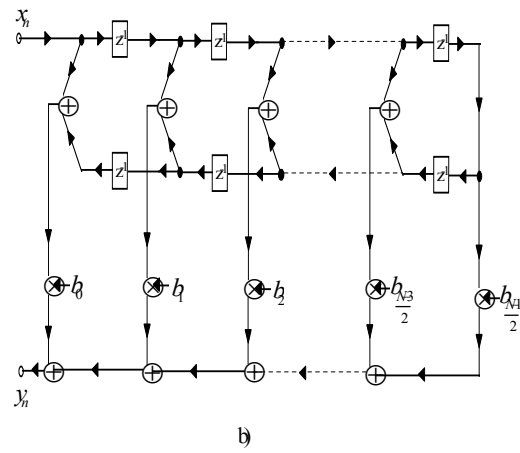
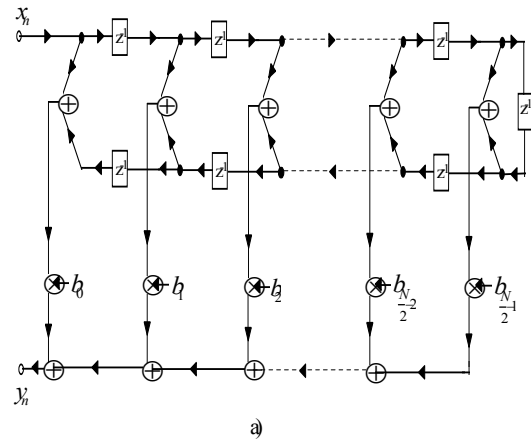


Figura 8.4: a) Realizzazione di un filtro FIR di ordine pari con fase lineare; b) Realizzazione di un filtro FIR di ordine dispari con fase lineare

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

Peraltro, come vedremo in un prossimo capitolo, sarà opportuno trovare delle strutture filtranti dove sia possibile garantire che i poli siano confinati all'interno del cerchio unitario, per rozza che sia la quantizzazione dei coefficienti. Ciò è possibile, con filtri fondati su analogie fisiche con le strutture riverberanti, per le quali si può garantire una stabilità intrinseca.

8.2.2 Sfasatori puri

E' importante osservare che se si ha uno sfasatore puro, caratterizzato cioè da una funzione di trasferimento del tipo:

$$A(z) = \frac{\tilde{D}(z)}{D(z)} \quad (8.8)$$

numeratore e denominatore sono eguali a parte l'ordine dei coefficienti. Anzitutto, è facile identificare delle strutture con cui, per realizzare uno sfasatore puro con N zeri ed N poli, bastano N moltiplicazioni per campione, invece di $2N$. Infatti, indicando con $X(z), Y(z)$ le trasformate z delle sequenze di ingresso e di uscita dello sfasatore puro, basta scriverne l'equazione alle differenze finite (supponendo $d_0 = 1$) :

$$Y(z) = \frac{\tilde{D}(z)}{D(z)} X(z)$$

$$D(z) Y(z) = \tilde{D}(z) X(z)$$

$$\begin{aligned} \sum_{m=0}^N d_m y_{n-m} &= \sum_{m=0}^N d_{N-m} x_{n-m} \rightarrow \sum_{m=0}^N d_m (y_{n-m} - x_{n-N+m}) = 0 \rightarrow \\ &\rightarrow y_n = x_{n-N} - \sum_{m=1}^N d_m (y_{n-m} - x_{n-N+m}) \end{aligned}$$

Inoltre, per qualsiasi quantizzazione dei coefficienti, lo sfasatore puro rimarrà sempre tale, in quanto la sequenza a numeratore $\tilde{D}(z)$ è comunque la ribaltata della sequenza a denominatore $D(z)$; bisogna però verificare che i poli non escano dal cerchio unitario. Se si ricorda che abbiamo a suo tempo osservato come si possano ottenere filtri passa basso e passa alto complementari in potenza da somme e differenze di due

sfasatori puri, si comprende che si possano così ottenere strutture fil-tranti robuste rispetto alla quantizzazione dei coefficienti. Nella sezione sulle strutture riverberanti, vedremo sistemi per realizzare sfasatori puri certamente stabili.

8.3 La quantizzazione dei segnali

Il risultato di una moltiplicazione tra due numeri rispettivamente di B_c (il coefficiente del filtro) e B_s (il segnale) cifre binarie, ha un numero di cifre binarie pari a $B_c + B_s$; è ragionevole ricondurre a B_s il numero di cifre binarie che rappresenta un segnale, o dopo il filtraggio, o, come accade ben più spesso, dopo ogni moltiplicazione di un segnale con un co-efficiente. Ciò implica un arrotondamento od un troncamento della cifra binaria, in funzione dell'aritmetica usata; con arrotondamento, indichi-amo l'operazione di approssimazione con il numero quantizzato (maggiore o minore) più vicino al numero dato. Invece, si ha troncamento quando il modulo del numero approssimante, quantizzato, è sempre inferiore al modulo del numero approssimato. Ancora, può a volte essere necessario saturare l'uscita per evitare trabocchi, se si supera il numero massimo che può essere rappresentato con l'aritmetica binaria utilizzata.

In questo paragrafo vogliamo identificare le implicazioni di questa operazione di quantizzazione: sono da considerare due casi di natura molto diversa tra loro:

a) il rumore dovuto alla quantizzazione è piccolo rispetto al segnale e può essere modellato come un disturbo indipendente dal segnale stesso, anche se questa indipendenza è di per sé falsa, visto che ad eguale segnale corrisponderà sempre eguale rumore di quantizzazione.

b) il rumore non può essere considerato indipendente dal segnale. Questo accade se il segnale è periodico o costante, e quindi il rumore di quantizzazione ha anch'esso delle periodicità. Ancora, può accadere che il rumore di quantizzazione, circolando attraverso il filtro numerico, viene amplificato, crea altro rumore ecc. fino alla formazione di un disturbo di ampiezza anche grande che può essere assai dannoso per il funzionamento del filtro stesso (ciclo limite).

8.3.1 Rumore di quantizzazione dei segnali

Se il numero di cifre binarie con cui si rappresenta un segnale viene riportato a B_s dopo ogni moltiplicazione, si addiziona al segnale un disturbo che, se il segnale è casuale e con una dinamica grande rispetto al rumore, è sostanzialmente indipendente dal segnale stesso.

Se il segnale ha una variazione di ampiezza tra picco positivo e picco negativo pari a V_{pp} , ed il rumore di quantizzazione nasce da un arrotondamento del segnale, esso ha una distribuzione di probabilità pressoché uniforme; l'escursione di ampiezza del rumore di quantizzazione, tra picco e picco, è $a = V_{pp}/2^{B_s}$; il valore medio è nullo (fig. 8.5). La varianza

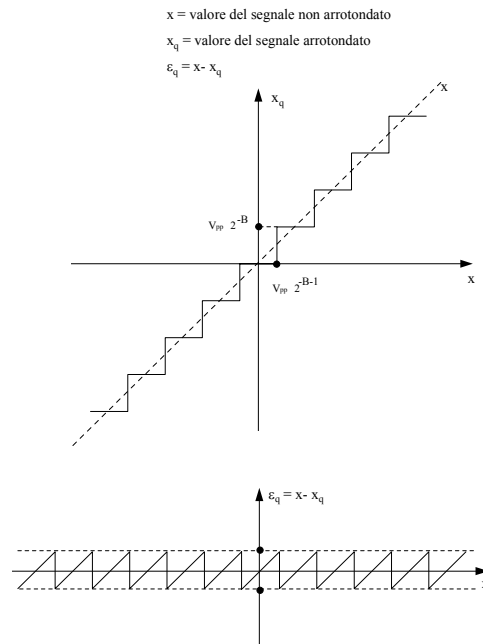


Figura 8.5: Rumore di quantizzazione: arrotondamento del segnale ed escursione di ampiezza del rumore

è:

$$\sigma_n^2 = \frac{1}{12} \frac{V_{pp}^2}{2^{2B_s}} = \frac{a^2}{12} \quad (8.9)$$

8.3 LA QUANTIZZAZIONE DEI SEGNALE

La situazione é diversa se il disturbo di quantizzazione nasce da un troncamento; allora il valore medio é nullo solo se il segnale ha valor medio nullo, e la varianza quadruplica (vedi figura 8.6).E' comune l'ipotesi che

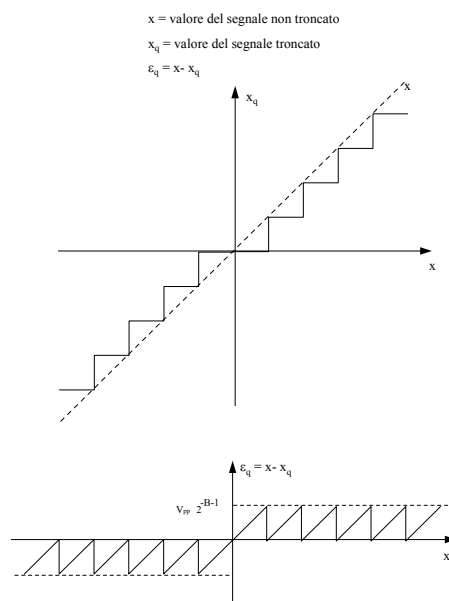


Figura 8.6: Rumore di troncamento: arrotondamento del segnale e ampiezza dell'errore

i campioni di questo rumore siano indipendenti tra loro; ancora, ciò non é vero in ogni caso come vedremo tra poco, ma é un'ipotesi valida nelle normali condizioni di lavoro di un filtro numerico.

Quindi, si può pensare che accanto ad ogni moltiplicatore (e sommatore) vi sia un iniettore di disturbo; il disturbo poi contribuisce all'uscita, dopo essere transitato per il filtro, attraverso una sua propria funzione di trasferimento, che é in generale diversa da quella del segnale. La potenza totale di disturbo in uscita si otterrà sommando le potenze di disturbo dovute alle varie sorgenti. Bisogna stare assai attenti all'amplificazione del segnale dovuta ai poli, che amplifica poi anche il disturbo.

Lo spettro continuo di potenza bilatero del rumore di quantizzazione iniettato, $|S_q(\omega)|^2$, é costante per l'ipotesi che i campioni successivi del

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

rumore siano indipendenti e quindi incorrelati e quindi la loro funzione di autocorrelazione sia impulsiva (per una discussione di questi concetti, si rinvia al capitolo sull'analisi spettrale). Il suo valore è determinabile con il teorema di Parseval; integrato fino alla frequenza di Nyquist, deve dare il valore quadratico medio del rumore.

$$\begin{aligned} E[n_q^2] &= \sigma_n^2 = \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 \frac{d\omega}{2\pi} = \frac{a^2}{12}; \\ \leftrightarrow |S_q(\omega)|^2 &= \frac{a^2 T}{12}; \end{aligned} \quad (8.10)$$

Se avessimo utilizzato la definizione di trasformazione discreta di Fourier del Cap. 4 avremmo trovato un diverso fattore di scala, come risulta nel capitolo dedicato al periodogramma. La sequenza dei campioni del rumore in uscita n_u e cioè $n_{u,i}$ corrisponde al rumore generato, amplificato dalla funzione di trasferimento $H(\omega)$.

$$E[n_u^2] = \sigma_u^2 = \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 |H(\omega)|^2 \frac{d\omega}{2\pi} \quad (8.11)$$

A ulteriore titolo di esempio, per un filtro del secondo ordine, ricorsivo, riporteremo la potenza di disturbo in uscita σ_o^2 dovuta al rumore di quantizzazione (fig. 8.9); le sorgenti di rumore sono due e possono essere trasferite al nodo sommatore di ingresso. La funzione di trasferimento tra ingresso ed uscita è:

$$H(z) = \frac{1}{1 - 2\rho \cos \theta z^{-1} + \rho^2 z^{-2}} \quad (8.12)$$

Per il teorema di Parseval:

$$\begin{aligned} E[n_o^2] = \sigma_o^2 &= \int_{-\pi/T}^{\pi/T} |S_q(\omega)|^2 |H(\omega)|^2 \frac{d\omega}{2\pi} = 2 \times \frac{a^2 T}{12} \int_{-\pi/T}^{\pi/T} |H(\omega)|^2 \frac{d\omega}{2\pi} = \\ &= \frac{a^2}{6} \frac{1 + \rho^2}{1 - \rho^2} \frac{1}{1 - 2\rho^2 \cos 2\theta + \rho^4} \end{aligned} \quad (8.13)$$

$$\text{Se } \rho = 1 - \varepsilon; \lim_{\varepsilon \rightarrow 0} \sigma_o^2 = \frac{a^2}{6} \times \frac{1}{4\varepsilon \sin^2 \theta}$$

8.3 LA QUANTIZZAZIONE DEI SEGNALE

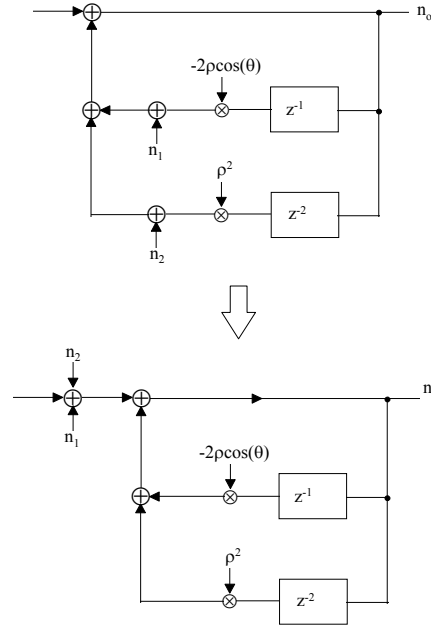


Figura 8.7: Filtro ricorsivo del secondo ordine: calcolo della potenza in uscita del rumore dovuto all'arrotondamento.

Quest'ultima formula è facilmente ottenibile considerando i due poli complessi coniugati distanti di ε dal cerchio unitario nel piano z ; in corrispondenza delle pulsazioni di guadagno massimo

$$\pm\omega_0 = \pm\theta/T$$

la funzione di trasferimento è deducibile dai diagrammi vettoriali:

$$|H(\omega)| = \frac{1}{2\sin\theta} \times \frac{1}{\sqrt{\varepsilon^2 + (\omega - \omega_0)^2 T^2}}$$

Allora, tenendo conto dei due contributi alle pulsazioni $\pm\omega_0$:

$$\sigma_o^2 = 2 \times \frac{a^2 T}{12} \int_{-\pi/T}^{\pi/T} |H(\omega)|^2 \frac{d\omega}{2\pi} \simeq \frac{a^2}{24\pi\varepsilon\sin^2\theta} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx =$$

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

$$= \frac{a^2}{24\varepsilon \sin^2 \theta}; \quad x = \frac{\omega T}{\varepsilon}$$

In sostanza, la potenza di rumore dovuta all'arrotondamento viene amplificata di un fattore proporzionale a $1/\varepsilon$.

Si osserva infine che il rumore dovuto alla quantizzazione dei segnali in una cella biquadratica cambia con il modo di realizzazione e cioè è diverso se si usi la struttura di figura (8.7) o la struttura con i sensi di propagazione e i coefficienti scambiati.

8.3.2 Effetti di trabocco (overflow)

Si prenderanno ora in considerazione gli effetti di trabocco, dovuti al fatto che i segnali descritti con cifre binarie (in virgola fissa) corrispondono a segnali a dinamica limitata (i valore massimi negativi e positivi sono finiti). Il valore massimo del modulo del segnale non deve superare troppo spesso la condizione di trabocco. Se il segnale è a densità di probabilità gaussiana, la probabilità che il suo valore di picco sia p.e. 3 volte superiore al valore efficace è pari a 2.7×10^{-3} . Per rendere trascurabile l'ulteriore contributo di rumore dovuto al trabocco, bisognerà fare in modo che in tutti i punti del filtro il valore efficace del segnale sia almeno 12 dB al di sotto del valore di V_p , condizione di trabocco del filtro; da questa osservazione e dal numero di bit che rappresentano i segnali si può determinare il rapporto segnale - rumore di quantizzazione risultante. Se poi si ricorda che i poli in generale tendono ad amplificare il segnale e gli zeri ad attenuarlo, sarà opportuno costruire i filtri numerici alternando poli e zeri e realizzando assieme con coppie biquadratiche i poli e gli zeri posizionati nelle stesse zone del piano z , così da mantenere abbastanza uniforme il livello del segnale.

Ancora, è opportuno che il trabocco non generi segnali di segno opposto e ampiezza vicina alla massima, come potrebbe accadere con aritmetica binaria senza controlli; infatti, ne potrebbero nascere ulteriori pericolose oscillazioni (cicli limite).

8.4 Quantizzazione e DFT

Una possibile fonte di disturbo nel trattamento dei segnali con la trasformazione di Fourier discreta deriva dalla quantizzazione dei coefficienti della trasformazione e dei segnali. Indicando con:

$$X(k) = \sum_{n=0}^{N-1} x_n W^{-nk}, \quad W = e^{j2\pi/N} \quad (8.14)$$

e supponendo che sia segnali che coefficienti siano quantizzati con errori di quantizzazione pari a δ_n e w_n rispettivamente:

$$X'(k) = X(k) + \Delta_k = \sum_{n=0}^{N-1} (x_n + \delta_n)(W^{-nk} + w_n);$$

$$E[w_n] = E[\delta_n] = E[\Delta_k] = 0;$$

$$E[|\Delta_k|^2] = N(E[|w_n|^2] \times \sigma_x^2 + E[\delta_n^2] \times 1 + E[\delta_n^2] \times E[w_n^2]) \simeq 2kN\sigma_x^2 2^{-2B} \quad (8.15)$$

se:

$$k = \frac{V_{pp}^2}{12\sigma^2}; \quad E[w_n^2] = k2^{-2B} \times \sigma_W^2 = 2^{-2B}; \quad E[\delta_n^2] = k2^{-2B}\sigma_x^2 \quad (8.16)$$

e quindi si determina il rumore che inevitabilmente è sovrapposto alla trasformata.

8.5 I cicli limite

Un altro tipo di disturbo è dovuto ai cicli limite: per comprendere il fenomeno, consideriamo un filtro ad un polo, situato assai vicino al cerchio unitario, ad esempio in $z = 1 - \varepsilon$ (fig. 8.10). Supponiamo di alimentare il filtro con un impulso di ampiezza A_0 e di lasciare che il sistema evolva; tuttavia l'uscita del moltiplicatore per $z = 1 - \varepsilon$ è quantizzata con B_s cifre binarie. La moltiplicazione per $z = 1 - \varepsilon$ seguita da una quantizzazione lascia il campione immutato non appena:

$$y_n = [(1 - \varepsilon)y_{n-1}]_{\text{quantizzato}} \quad (8.17)$$

CAPITOLO 8 EFFETTI DI QUANTIZZAZIONE

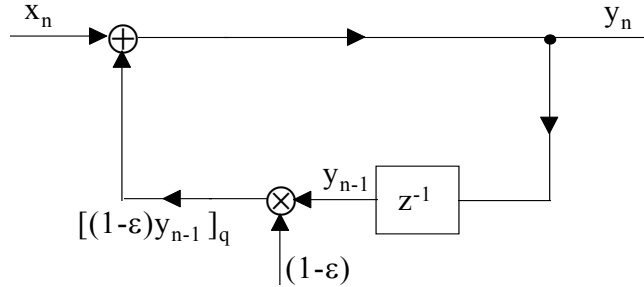


Figura 8.8: Effetti di arrotondamento: il ciclo limite

coincide con y_{n-1} e cioè quando:

$$\varepsilon y_{n-1} \leq \frac{\Delta}{2} \rightarrow y_{n-1} \leq \frac{\Delta}{2\varepsilon} = A_{lim} \quad \Delta = \frac{V_{pp}}{2^{B_s}} \quad (8.18)$$

da questo momento in poi, è come se il polo si fosse portato sul cerchio unitario, e ricircola il valore A_{lim} senza decrescere più. In altre parole, il sistema è diventato "appiccicoso" (sticky) e l'uscita non tende più a zero, ma rimane al valore A_{lim} . Si osservi che se ε è piccolo, A_{lim} può essere anche notevole.

Una cura per questo semplice caso di ciclo limite può essere un sistema simile a quello di scuotere una ciotola per fare scivolare sul fondo una pallina rimasta attaccata alle pareti appiccicose. In questo caso, si aggiunge all'ingresso del sistema (per poi sottrarlo all'uscita) una sequenza di dither (setaccio). Il dither è una sequenza pseudo casuale e cioè una sequenza periodica, nota, e sufficientemente ampia da spezzare il ciclo limite; la presenza del dither riporta il sistema nelle condizioni per cui il segnale è casuale ed ha ampia dinamica. All'uscita del filtro, si sottrae l'effetto del disturbo pseudo casuale, facilmente calcolabile nel caso di sistema privo di rumore di quantizzazione.

Nel caso del semplice filtro di figura 8.10 è stato assai facile calcolare il ciclo limite; ciò è assai più difficile nel caso di filtri numerici complessi. Non esistono regole sicure allora, se non quella della simulazione, oppure il ricorrere a strutture intrinsecamente senza cicli limite come le strutture riverberanti che vedremo in un prossimo capitolo.

8.5 I CICLI LIMITE

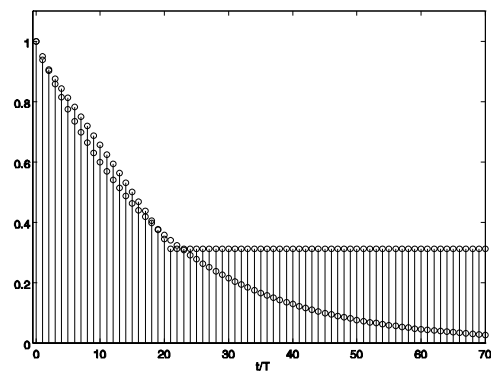


Figura 8.9: Risposta impulsiva con e senza l'effetto del ciclo limite

Capitolo 9

La trasformata di Hilbert

9.1 Trasformate di sequenze causali

Esiste una relazione tra la parte reale e la parte immaginaria della trasformata di Fourier di una sequenza causale; la stessa relazione intercorre tra modulo e fase della trasformata di una sequenza a fase minima; ancora, è la stessa relazione che vi è tra una funzione e quella che si ottiene ruotando di $\pi/2$ tutte le componenti della sua trasformata di Fourier: questa relazione è di tipo convolutivo e viene indicata come trasformazione di Hilbert, $\mathcal{H}(\cdot)$. In sostanza si tratta della convoluzione con una funzione del tipo $-1/(\pi t)$. A causa di campionamenti sia nei tempi e/o nelle frequenze, la funzione con cui si fa la convoluzione assume forme diverse, che vedremo. Tuttavia, la ripetuta applicazione della trasformazione di Hilbert porterà sempre alla funzione di partenza con il segno cambiato.

Consideriamo una sequenza **reale e causale** e cioè sia:

$$x_i = 0; \quad i < 0 \quad (9.1)$$

e si decomponga la sequenza nelle sue parti pari e dispari (fig. 6.8). Sia cioè:

$$x_{p,i} = x_i/2; \quad x_{p,-i} = x_i/2; \quad (9.2)$$

$$x_{d,i} = x_i/2; \quad x_{d,-i} = -x_i/2 \quad (9.3)$$

$$x_i = x_{p,i} + x_{d,i} \quad (9.4)$$

Si ha cioè che, definita la sequenza $sign(i)$ in modo che:

$$sign(i) = \frac{i}{|i|}; \quad i \neq 0; \quad sign(0) = 0; \quad (9.5)$$

CAPITOLO 9 LA TRASFORMATTA DI HILBERT

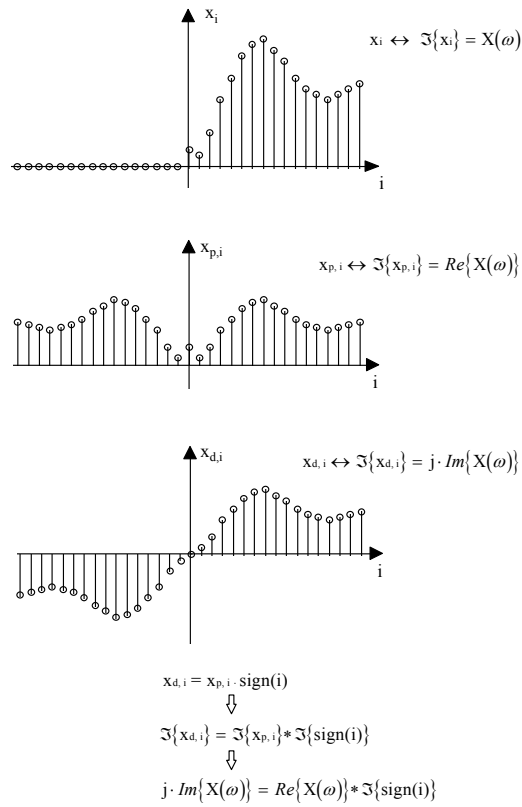


Figura 9.1: Decomposizione di una sequenza nelle componenti pari $x_{p,i}$ e dispari $x_{d,i}$

9.1 TRASFORMATE DI SEQUENZE CAUSALI

allora risulta :

$$x_{d,i} = x_{p,i} \text{sign}(i) \quad (9.6)$$

e quindi, dette $R(\omega)$, $jIm(\omega)$ le parti reali ed immaginarie della trasformata di Fourier della sequenza x_i , risulta che se:

$$X(\omega) = R(\omega) + jIm(\omega) \quad (9.7)$$

la parte immaginaria è la **Trasformata di Hilbert** della parte reale:

$$jIm(\omega) = \mathcal{H}[Re(\omega)] = Re(\omega) * F(\text{sign}(i)) \quad (9.8)$$

$$F[\text{sign}(i)] = \sum_{i=1}^{\infty} (z^{-i} - z^i) = \frac{1}{1 - z^{-1}} - \frac{1}{1 - z} = \frac{z + 1}{z - 1} = -j \cot(\omega T/2) \quad (9.9)$$

(sul cerchio unitario). Pertanto:

$$Im(\omega) = \mathcal{H}\{Re(\omega)\} = Re(\omega) * -\cot(\omega T/2) \quad (9.10)$$

Il risultato ottenuto è facilmente ricordato se si pensa al fatto che la "derivata numerica" della sequenza sign è un impulso. Allora le trasformate di Fourier della sequenza sign e dell'operatore derivata devono essere l'una proporzionale all'inverso dell'altra.

Infatti richiamiamo l'espressione dell'operatore di derivazione causale, recursiva:

$$D_r(z) = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} = \frac{2}{T} \frac{z - 1}{z + 1} \quad (9.11)$$

Ora, vediamo che le due espressioni $(D_r(z), F[\text{sign}(i)])$ corrispondono l'una all'inversa dell'altra. Tuttavia, non ci si stupisca se, convolvendo le due risposte in tempi NON si ottiene un impulso: in un caso (derivatore ricorsivo) il polo è visto come causale; nel caso invece della funzione sign il polo è sia causale che anticausale.

Abbiamo dimostrato che le parti reale ed immaginaria della trasformata di una sequenza causale sono legate dalla convoluzione con la funzione periodica $-\cot(\omega T/2)$, che è la trasformata di Fourier della sequenza $\text{sign}(i)$; $-\infty < i < \infty$; essa rappresenta una versione discreta del cosiddetto nocciolo di Hilbert (fig. 6.9).

La trasformazione descritta dalle precedenti equazioni è perciò chiamata **trasformazione di Hilbert discreta**, e corrisponde al caso in cui il segnale è campionato ed illimitato in un dominio (in questo caso quello dei tempi) e periodico e continuo nell'altro (in questo caso quello delle frequenze).

CAPITOLO 9 LA TRASFORMATA DI HILBERT

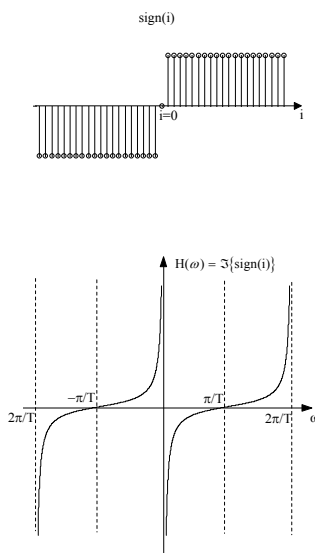


Figura 9.2: Nocciolo di Hilbert illimitato nel dominio coniugato

9.2 Sfasatori puri di $\pi/2$

Come osservato, esistono varie versioni del nocciolo di Hilbert, a seconda che il segnale sia continuo nei due domini, campionato e illimitato nell'uno e periodico e continuo nell'altro, o infine campionato e periodico in entrambi.

Consideriamo il caso che il campionamento sia nel dominio coniugato a quello dove si calcola la funzione $\text{sign}(i)$ che quindi viene generalizzata diventando una funzione continua e periodica, invece che discreta ed illimitata.

Il problema di riferimento in questo caso è il seguente: data una sequenza reale x_i se ne consideri la trasformata di Fourier $X(\omega)$ (si ricordi che questa è una funzione periodica!) e si ruotino di $\pi/2$ tutte le sue componenti di Fourier; in altre parole, si moltiplichino la trasformata per la funzione $j\text{sign}(\omega)$ periodicizzata sull'intervallo $-\pi/T \div \pi/T$. L'anti-trasformata è una nuova sequenza $\xi_i = \mathcal{H}\{x_i\}$, che anche viene identificata come la trasformata di Hilbert della sequenza x_i . La situazione però è differente da quella vista prima perchè la sequenza equivalente al nocciolo di Hilbert ha trasformata periodica e continua (fig. 6.10).

9.2 SFASATORI PURI DI $\pi/2$

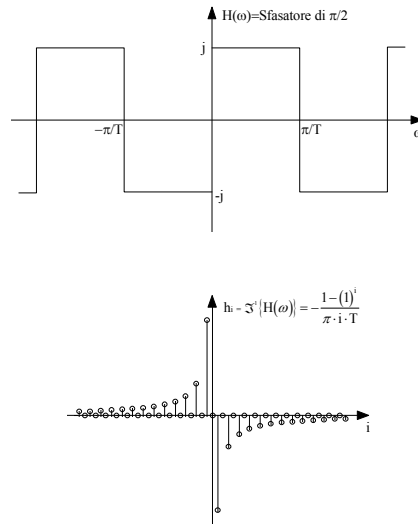


Figura 9.3: Nocciolo di Hilbert periodicizzato nel dominio coniugato

Si tratta quindi di calcolare l'antitrasformata della funzione periodizzata $j\text{sign}(\omega)$; $|\omega| \leq \pi/T$. Per fare ciò senza fatica, basta calcolare prima il segnale continuo antitrasformata della funzione $j\text{sign}(\omega)$ annullata fuori del periodo principale e poi campionare la funzione del tempo così ottenuta. Il risultato è:

$$h(t) = -\frac{1 - \cos \pi t/T}{\pi t/T}; \quad h(iT) = h_i = -\frac{1 - (-1)^i}{\pi i} \quad -\infty < i < \infty \quad (9.12)$$

Esso assomiglia all'andamento della cotangente con il segno cambiato, ma con differenze notevoli se si considera che la sequenza h_i non è un segnale continuo, è nulla a campioni alterni e non è periodica. La sequenza ξ_i , convoluzione della x_i e della h_i , è detta ancora trasformata di Hilbert della sequenza x_i .

$$\xi_i = \mathcal{H}\{x_i\} = x_i * h_i$$

Si noti che se una sequenza è una sinusoide (campionata), la sua trasformata di Hilbert è una cosinusoide etc.

9.3 Sequenze periodiche e segnali continui

Una terza versione del nocciolo di Hilbert si ha quando la trasformata di Fourier sia periodica e campionata e cioè quando la sequenza di partenza è campionata e periodica. E' evidente che in questo caso il concetto di causalità va riformulato, accettando di definire causale una sequenza nulla per la seconda metà dei campioni. Allora se ne deduce una sequenza *sign* periodicizzata di cui si cerca la trasformata.

$$X(k) = \sum_{n=-N/2+1}^{N/2} x_n W^{-nk};$$

$$\begin{aligned} W &= \exp(2\pi j/N); \\ x_n &= 1; \quad 0 < n < N/2 \\ x_n &= -1; \quad -N/2 < n < 0; \quad x_0 = x_{N/2} = 0 \end{aligned} \quad (9.13)$$

Risulta che:

$$X(k) = -j[1 - (-1)^k] \cot k\pi/N \quad (9.14)$$

e perciò si trova tanto l'effetto della discontinuità alla frequenza di Nyquist (la nullità a elementi alterni della trasformata) quanto l'effetto della periodicizzazione (la cotangente) (fig. 6.11).

L'ultimo caso è quello della continuità nei due domini: il nocciolo di Hilbert è allora:

$$h(t) = -\frac{1}{\pi t} \quad (9.15)$$

Possiamo ora tabellare le varie forme assunte dalla funzione che appare nella trasformazione di Hilbert; essa è sempre la trasformata di Fourier della funzione *sign*, che però può essere continua/discreta e illimitata/periodica, nel dominio dei tempi o delle frequenze ed inoltre, si faccia attenzione, a volte compare con un segno positivo ed a volte con quello negativo.

Funzione	<i>continua</i>	<i>continua</i>	<i>discreta</i>	<i>discreta</i>
Segnale	<i>illimitato</i>	<i>periodico</i>	<i>illimitato</i>	<i>periodico</i>
Espressione	$-\frac{1}{\pi t}$	$-\frac{1-(-1)^i}{\pi i}$	$-\cot \frac{\omega T}{2}$	$-\left[1 - (-1)^k\right] \cot \frac{k\pi}{N}$

9.3 SEQUENZE PERIODICHE E SEGNALI CONTINUI

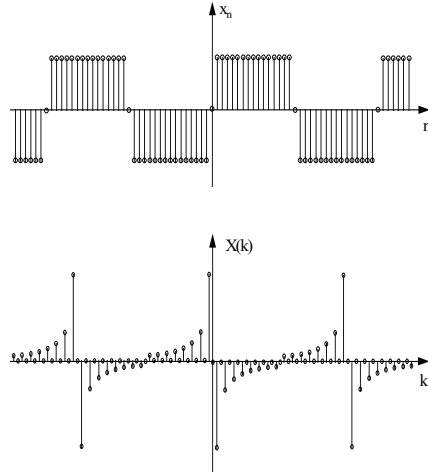


Figura 9.4: Sequenza $sign(n)$ periodicizzata e sua trasformata

9.3.1 Singularità nella trasformazione di Hilbert

Apparentemente, la trasformazione di Hilbert non dovrebbe creare singolarità perchè corrisponde a pure rotazioni di fase, senza amplificazioni dello spettro. Quindi il teorema di Parseval ci dovrebbe permettere di concludere che non sono da attendersi singolarità. Tuttavia, la trasformata di Hilbert della sequenza $sign(i)$ ha una singolarità di tipo logaritmico, che nasce dalla somma di una serie armonica (fig. 6.12).

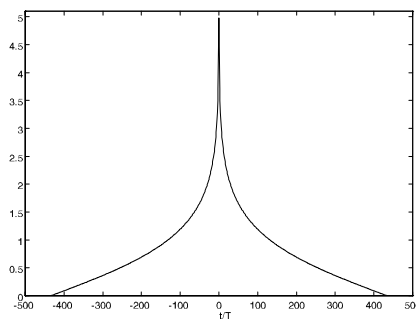


Figura 9.5: Trasformata di Hilbert della sequenza $sign(n)$

Questo fatto è connesso a quanto osservato nel capitolo 2, sull'interpolazione di sequenze illimitate con il filtro ideale. I segnali continui che interpolano sequenze con transitori bruschi e permanenti hanno picchi che crescono con il logaritmo del numero dei campioni della sequenza e quindi infiniti nel caso di sequenze per le quali il transitorio è illimitato.

9.4 Funzioni a fase minima

Supponiamo di avere una funzione di trasferimento di cui è nota solo la caratteristica di ampiezza. E' chiaro che possiamo assegnarne la caratteristica di fase con vari criteri: ad esempio possiamo imporre che la funzione di trasferimento sia a fase zero, ed allora la risposta all'impulso sarà non causale e simmetrica nel tempo; ancora possiamo imporre che la risposta all'impulso sia causale, aggiungendo un ritardo e quindi una fase che cresce linearmente con la frequenza.

E' importante il caso in cui vogliamo che la risposta sia sì causale, ma anche desideriamo che la sua durata sia minima e che cioè il ritardo della risposta sia il più breve possibile. Similmente, si può desiderare che la funzione di trasferimento abbia una funzione di trasferimento inversa causale e quindi sia a fase minima, come visto nel capitolo sulla Trasformata z . E' perciò un problema interessante quello di determinare la caratteristica di fase minima che si accompagna ad una caratteristica di ampiezza assegnata. Dimosteremo che vi è un facile legame tra caratteristica di ampiezza e caratteristica di fase minima, che corrisponde ad una trasformazione di Hilbert.

Si abbia infatti una sequenza a fase minima e indichiamo la sua trasformata z con $F(z)$; dimostriamo anzitutto che $G(z)$, logaritmo complesso di $F(z)$, è la trasformata z di una sequenza causale. Infatti, se $F(z)$ è a fase minima:

$$G(z) = \log F(z) = \log(|F(z)|e^{j\angle F(z)}) = \log |F(z)| + j\angle F(z) \quad (9.16)$$

La $F(z)$, trasformata z di una sequenza a fase minima, ha poli e zeri all'interno del cerchio unitario. Il suo logaritmo complesso $G(z)$ ha tutti i poli all'interno del cerchio unitario (i valori infiniti della funzione logaritmo sono raggiunti dove l'argomento del logaritmo ha valore zero o infinito e quindi dove la $F(z)$ ha zeri o poli). La funzione $G(z)$ si può annullare (in corrispondenza degli zeri della $G(z)$) in qualsiasi posto del

9.4 FUNZIONI A FASE MINIMA

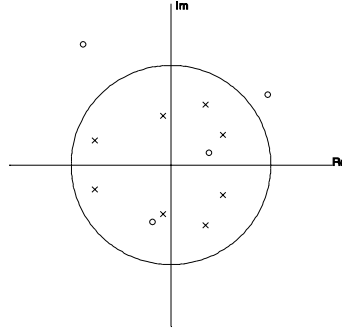


Figura 9.6: Poli e zeri del logaritmo della trasformata di una sequenza a fase minima

piano complesso, là dove $F(z) = 1$ (fig. 6.13)). Quindi la sequenza antitrasformata della $G(z)$ è una sequenza causale. Perciò le parti reale e parte immaginaria di $G(z)$ lungo il cerchio unitario sono legate dalla trasformazione di Hilbert e si può ottenere l'una dall'altra mediante la convoluzione con la funzione $-\cot(\omega T/2)$.

Ricordando poi che la parte immaginaria di $G(z)$ è la fase di $F(z)$ risulta:

$$\angle F(z) = \text{Im}[G(z)] = \text{Im}(\log F(z)) \quad (9.17)$$

$$\begin{aligned} \angle F(z)|_{z=e^{j\omega T}} &= \text{Im}(\log F(z)|_{z=e^{j\omega T}}) = \\ &= \mathcal{H}\{\text{Re}(\log F(z)|_{z=e^{j\omega T}})\} = \mathcal{H}\{\log |F(z)|_{z=e^{j\omega T}}\}. \end{aligned} \quad (9.18)$$

In conclusione, per ottenere la caratteristica di fase di una funzione di trasferimento a fase minima basta fare la trasformata di Hilbert del logaritmo naturale del modulo della sua caratteristica di ampiezza. Nel capitolo sulla predizione lineare troveremo un'importante applicazione di questo risultato.

Come esempio, si può considerare un filtro ideale di banda inferiore alla frequenza di Nyquist. La sequenza corrispondente è un *sinc* che poi deve essere opportunamente finestrato (p.e. con un coseno rialzato) per limitarne la durata. Il *sinc* finestrato è presentato in fig. 6.14; il logaritmo della caratteristica di ampiezza è presentato in fig. 6.15; la trasformata di Hilbert (la caratteristica di fase da associare alla caratteristica di ampiezza per avere fase minima) è in fig. 6.16. Infine la sequenza che

CAPITOLO 9 LA TRASFORMATTA DI HILBERT

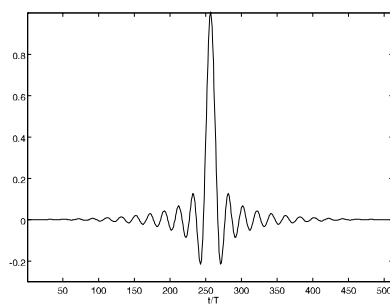


Figura 9.7: Sinc finestrato

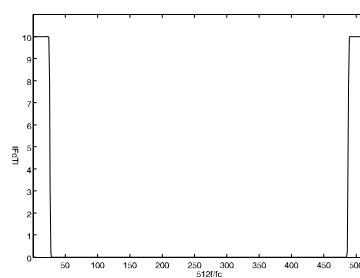


Figura 9.8: Caratteristica d'ampiezza

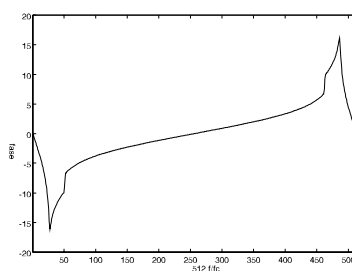


Figura 9.9: Caratteristica di fase

9.5 IL SEGNALE ANALITICO

ha le caratteristiche di ampiezza e di fase mostrate è presentata in fig. 6.17. E' evidente che la durata della sequenza in fig. 6.17 è ben inferiore

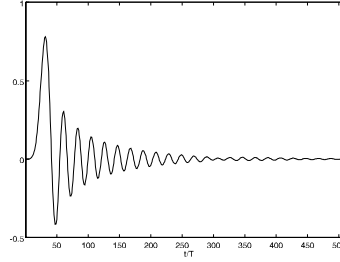


Figura 9.10: Sequenza a fase minima corrispondente ad un filtro passa basso ideale

a quella della sequenza in fig. 6.14.

9.5 Il segnale analitico

Consideriamo una sequenza reale x_i e aggiungiamogli come parte immaginaria una sequenza $-j\xi_i$, ottenuta dalla sequenza x_i mediante trasformazione di Hilbert e moltiplicazione per $-j$. La trasformata di Fourier di questa nuova sequenza complessa:

$$y_i = x_i - j\xi_i \quad (9.19)$$

che viene definita "il segnale analitico" corrispondente alla sequenza x_i ha la parte negativa dello spettro nulla. Infatti, la trasformata della sequenza $j\xi_i$ è identica a quella della sequenza x_i a meno della moltiplicazione per la funzione *sign*. Ancora, la trasformata della sequenza ξ_i è eguale a quella della sequenza x_i a meno di una rotazione di $-\pi/2$ di tutte le componenti. Infine, la trasformata della sequenza y_i è "causale" nelle frequenze e pertanto ha spettro nullo per $\omega < 0$. E' utile osservare che:

$$|y_i| = \sqrt{x_i^2 + \xi_i^2}$$

$$\angle y_i = \arctan(\xi_i/x_i) \quad (9.20)$$

CAPITOLO 9 LA TRASFORMATTA DI HILBERT

Si ricordi che se la sequenza x_i corrisponde ad una sinusoide, la sequenza ξ_i corrisponde alla stessa sinusoide sfasata di $-\pi/2$ e pertanto $|y_i|$ e $\angle y_i$ sono delle costanti e niente altro che l'ampiezza e la fase della sinusoide stessa. Pertanto con il segnale analitico è possibile ottenere un'estensione dei concetti di fase e di ampiezza di una sinusoide, e costruire inviluppo e fase istantanea di un segnale qualsiasi (fig. 6.18); naturalmente, è anche possibile definire la frequenza istantanea di un segnale qualsiasi, come derivata della fase istantanea, fatta rispetto al tempo.

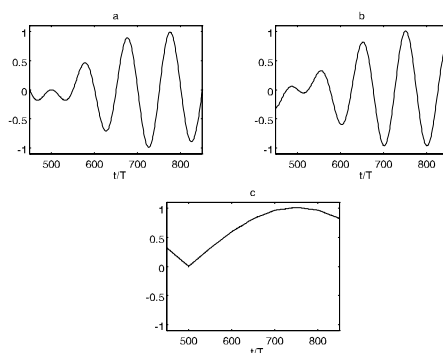


Figura 9.11: a) Segnale modulato in ampiezza x_i b) Trasformata di Hilbert del segnale modulato in ampiezza c) Inviluppo del segnale modulato

Capitolo 10

Le strutture riverberanti

10.1 Introduzione

Abbiamo visto nel capitolo sulla quantizzazione di coefficienti e segnali come piccole variazioni dei coefficienti di filtri stabili IIR possano portare ad instabilità. Abbiamo anche visto la possibilità di formazione di cicli limite in filtri IIR con poli vicini al cerchio unitario.

Entrambi questi fenomeni sono poco fisici; non ci attendiamo che un sistema fisico possa, con lievi alterazioni dei parametri, diventare da stabile instabile e generare energia invece di assorbirne. La questione è quindi che i filtri numerici, quali studiati finora, non corrispondono direttamente a delle strutture fisiche da cui possano mutuare pregevoli doti di stabilità.

Questa è appunto la corrispondenza che vogliamo studiare in questo capitolo: in particolare verranno introdotti i sistemi numerici chiamati **filtri a traliccio ricorsivi**, derivanti da sistemi fisici indicati come "strutture riverberanti"; mediante questi sistemi è possibile realizzare qualsiasi filtro causale IIR: i loro parametri però avranno un diretto significato fisico ed in particolare saranno i coefficienti di riflessione per onde trasmesse su cascate di tronchi di linee di trasmissione; per motivi energetici, i coefficienti di riflessione saranno sempre inferiori all'unità.

Consideriamo una linea di trasmissione illimitata, uniforme; questa linea può essere elettrica (p.e. un cavo coassiale), acustica (p.e. un trombone), o meccanica (p.e. una corda tesa). In questa linea si propagano onde (di tensione e corrente nel caso della linea elettrica, di velocità e

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

pressione nel caso della linea acustica, di velocità e tensione nel caso della corda tesa) nelle due direzioni che indicheremo con S (sinistra) e D (destra) (fig. 8 .1). Supponiamo ora di avere due tronchi di linea illimi-

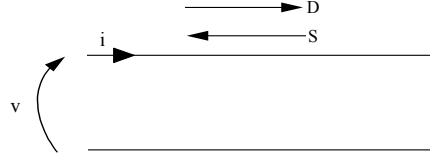


Figura 10.1: Convenzione sulla propagazione delle onde in una linea

tati, diversi, ma entrambi uniformi e collegati insieme (fig. 8 .2): l'onda A_i che arriva alla interfaccia da uno dei due tronchi viene parzialmente riflessa ed in parte viene trasmessa nel secondo tronco; supponendo unitaria l'ampiezza dell'onda incidente A_i indicheremo con c e t le ampiezze delle onde rispettivamente riflesse e trasmesse. Per la continuità delle grandezze fisiche all'interfaccia deve essere:

$$A_i + cA_i = tA_i ; \quad \rightarrow \quad 1 + c = t \quad (10.1)$$

Se ora indichiamo con Y_i ; $i = 1, 2$ le ammettenze dei due tronchi di linea e cioè dei parametri che legano l'energia trasportata dall'onda al modulo quadrato dell'ampiezza dell'onda stessa (ammettenze vere e proprie se si tratta di tensioni, impedenze se si tratta di correnti, altri parametri meccanici in altri tipi di linee ecc.) risulta che l'energia trasportata da un'onda di ampiezza A_i è E_i :

$$E_i = Y_i \times A_i^2 \quad (10.2)$$

Il bilancio energetico impone che la somma delle energie trasportate dalle onde riflessa e trasmessa sia eguale all'energia dell'onda incidente:

$$Y_2 \times 1^2 = Y_2 \times c^2 + Y_1 \times t^2 \quad (10.3)$$

Sostituendo nella relazione (8 .1) risulta:

$$c = \frac{Y_2 - Y_1}{Y_2 + Y_1} \quad (10.4)$$

10.2 STRUTTURE RIVERBERANTI E SISTEMI NUMERICI

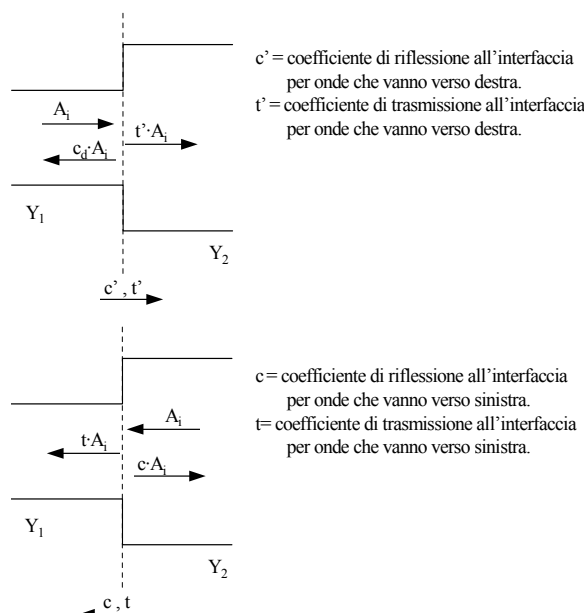


Figura 10.2: Connessione tra tronchi di linee illimitati

Poiché, per motivi energetici, $Y_i > 0$, ne risulta che i coefficienti di riflessione sono inferiori ad 1; $|c_i| < 1$; inoltre il coefficiente di trasmissione $0 < t < 2$.

Ancora, scambiando i due mezzi, il coefficiente di riflessione cambia segno $c' = -c$; $t' = 1 + c' = 1 - c$. Non deve stupire che il coefficiente di trasmissione possa essere maggiore di 1; basta pensare alle onde del mare che, arrivando vicino alla spiaggia, diventano più alte, senza che la loro energia sia amplificata.

10.2 Strutture riverberanti e sistemi numerici

Fino ad ora, niente appare collegare i sistemi numerici e le strutture riverberanti; pensiamo però ad una cascata di tronchi di linea, diversi tra loro, e che terminano in tronchi di linea illimitati, dalle due parti.

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

Supponiamo anche che tutti i tronchi, sia pure diversi tra loro, abbiano un eguale ritardo di propagazione pari a $T/2$.

Lanciamo ora su questa linea composita un'onda impulsiva da uno dei due tronchi illimitati; l'impulso si propagherà nei vari tronchi e lancerà altre onde impulsive nelle due direzioni. Tuttavia, sia l'onda che finalmente esce dalla struttura, sia l'onda riflessa saranno sequenze di infiniti impulsi intervallati di T ; infatti ad ogni interfaccia sono generate nuove e nuove onde che però si combinano in qualsiasi altra interfaccia sempre dopo intervalli di tempo multipli di T , e cioè dopo un numero intero di percorsi avanti e indietro nelle varie celle.

Per rendersi conto della situazione consideriamo una cascata di tre tronchi di linea (fig. 8 .3); indichiamo con Y_i ; $i = 1, 2, 3$ le tre ammettenze dei tre tronchi e c_i , $i = 1, 2$ i coefficienti di riflessione alle due interfacce, per onde che vadano verso destra (D); per le onde S, che vanno verso sinistra, i coefficienti di riflessione avranno segno opposto. L'onda impulsiva di ampiezza unitaria arriva al tempo 0 alla prima interfaccia e si dipartono due onde impulsive nelle due direzioni di ampiezze c_1 e t_1 ; l'onda riflessa nel primo mezzo si propaga all'indietro indisturbata, dopo un tempo $T/2$ (ritardo che rappresenteremo con $z^{-1/2}$); l'onda trasmessa arriva alla seconda interfaccia e qui partono due nuove onde, di ampiezza e ritardo $c_2 t_1 z^{-1/2}$ e $t_2 t_1 z^{-1/2}$; poi, la prima si riflette ancora sulla prima interfaccia, parte rifluendo nel primo mezzo e parte nel secondo, e così via.

L'onda trasmessa nel terzo mezzo é :

$$T(z) = t_1 t_2 z^{-1/2} (1 - c_2 c_1 z^{-1} + (c_2 c_1 z^{-1})^2 - (c_2 c_1 z^{-1})^3 \dots = \frac{t_1 t_2 z^{-1/2}}{1 + c_1 c_2 z^{-1}} \quad (10.5)$$

L'onda riflessa nel primo mezzo é:

$$R(z) = c_1 + (t_1 c_2 t'_1 z^{-1}) (1 - c_2 c_1 z^{-1} + (c_2 c_1 z^{-1})^2 - (c_2 c_1 z^{-1})^3 + \dots = (10.6)$$

$$c_1 + \frac{t_1 c_2 t'_1 z^{-1}}{1 + c_1 c_2 z^{-1}} = \frac{c_1 + c_2 z^{-1}}{1 + c_1 c_2 z^{-1}} \quad (10.7)$$

La sequenza corrispondente all'onda trasmessa è un polo, a parte il ritardo inessenziale di $1/2$ campione. L'onda riflessa invece ha un polo ed uno zero, e corrisponde ad uno sfasatore puro nel caso che $|c_2| = 1$: infatti, in questo caso l'energia è tutta riflessa e quindi è ragionevole che,

10.2 STRUTTURE RIVERBERANTI E SISTEMI NUMERICI

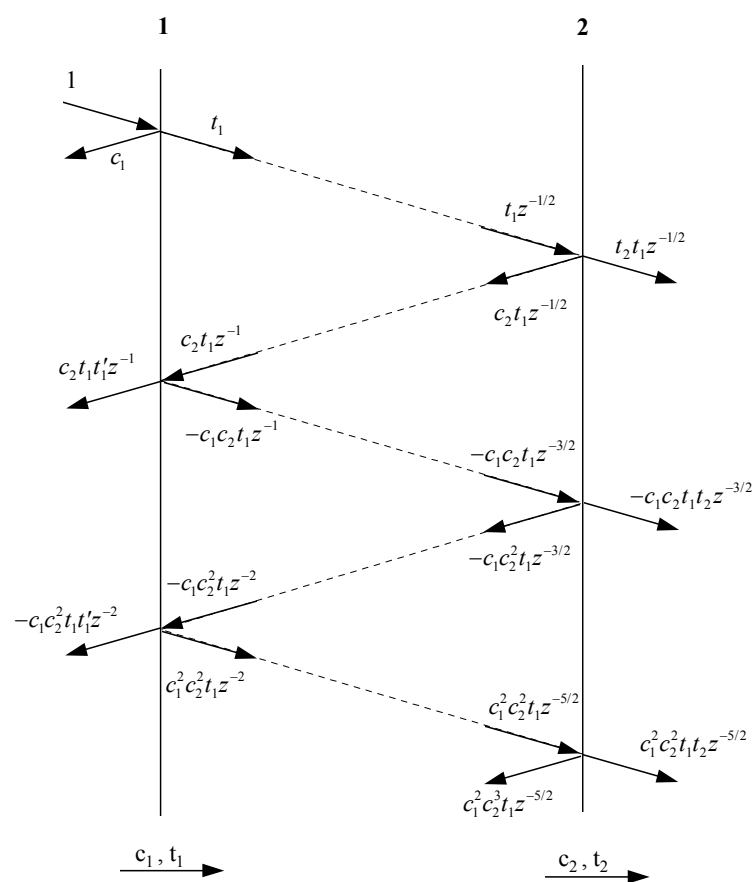


Figura 10.3: Riverberazioni in una cascata di tre tronchi di linea

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

sfasamenti a parte, la funzione di trasferimento abbia modulo unitario in tutta la banda.

Quando un coefficiente di riflessione è unitario, diremo che la struttura è totalmente disadattata ad un estremo.

10.3 Cascata di tronchi di linea

Consideriamo ora una cascata di tronchi di linea e determiniamo le relazioni tra i segnali ai due ingressi ed alle due uscite: consideriamo le due interfacce tra i tronchi $n-1; n$ ed $n, n+1$; indichiamo con c_n, t_n e $c'_n = -c_n, t'_n = 1 - c_n$ i coefficienti di riflessione e trasmissione per i segnali che si propagano verso sinistra (S) e destra (D) rispettivamente. Indichiamo poi con D_n, S_n i segnali nel tronco n , che si propagano verso destra e sinistra rispettivamente in corrispondenza dell'interfaccia verso il blocco $n-1$ e D'_n, S'_n i segnali (sempre nel tronco n) in corrispondenza dell'interfaccia verso il blocco $n+1$ (vedi figura 8.4). La relazione tra le

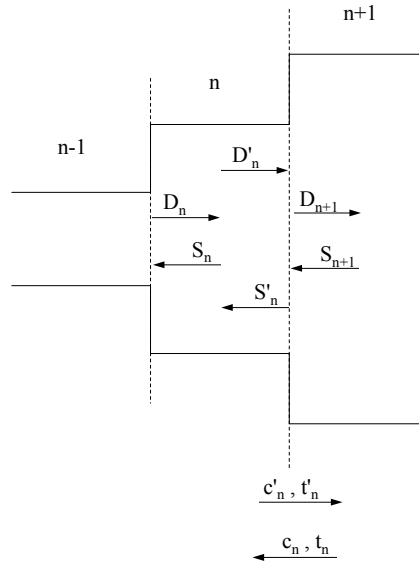


Figura 10.4: Coefficiente di riflessione e trasmissione all'($n+1$)-esima interfaccia

10.4 IL FILTRO A TRALICCIO RICORSIVO

ampiezze delle onde ai due lati dell'interfaccia è pertanto:

$$\begin{bmatrix} S'_n(z) \\ D_{n+1}(z) \end{bmatrix} = \begin{bmatrix} t_n & c'_n \\ c_n & t'_n \end{bmatrix} \times \begin{bmatrix} S_{n+1}(z) \\ D'_n(z) \end{bmatrix} \quad (10.8)$$

Ancora possiamo scrivere che:

$$S_n(z) = z^{-1/2} S'_n(z)$$

$$D_n(z) = z^{+1/2} D'_n(z)$$

e combinando le precedenti equazioni troviamo che:

$$\begin{bmatrix} S_{n+1}(z) \\ D_{n+1}(z) \end{bmatrix} = \frac{z^{1/2}}{t_n} \begin{bmatrix} 1 & c_n z^{-1} \\ c_n & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_n(z) \\ D_n(z) \end{bmatrix} \quad (10.9)$$

10.4 Il filtro a traliccio ricorsivo

Le strutture riverberanti si trasformano in sistemi numerici in modo semplice: infatti basta costruire lo schema riportato in figura 8.5 corrispondente all'equazione (10.9). In questo schema compaiono i blocchi:

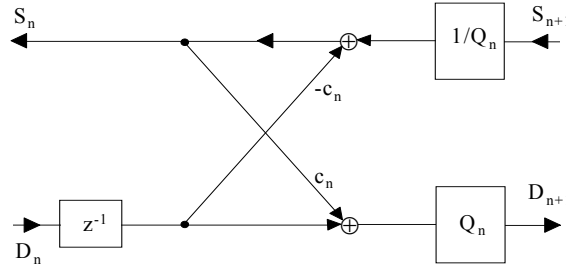


Figura 10.5: Schema della struttura numerica equivalente ad un tronco di linea

$$Q_n = \frac{z^{1/2}}{t_n}; \quad \frac{1}{Q_n} = \frac{t_n}{z^{1/2}} \quad (10.10)$$

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

che hanno funzioni di trasferimento reciproche e si possono elidere contemporaneamente, senza cambiare le funzioni di trasferimento esterne al blocco. Aggiungendo altre celle ed anche una cella iniziale disadattata corrispondente al coefficiente $|c_0| = 1$, vedremo che è possibile costruire:

- un filtro tutti poli di ordine N ;
- un'uscita ritardata di N campioni; (figg. 8 .6, 8 .7)
- uno sfasatore puro caratterizzato dagli stessi poli. Si osservi infatti

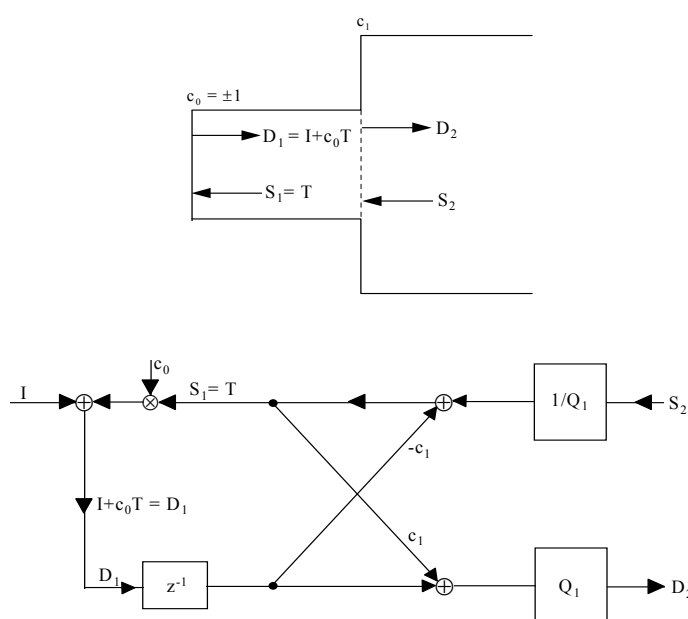


Figura 10.6: Cella iniziale corrispondente al disadattamento e suo corrispondente schema numerico

che il sistema numerico ha due ingressi possibili, $I(z)$ dalla parte della cella a coefficiente di riflessione unitario (funzione di trasferimento verso $C(z)$ caratterizzata da N poli e un ritardo di N campioni), ed $E(z)$ dalla parte adattata (non c'è ritardo tra ingresso ed uscita in $S_1(z)$); l'uscita può essere prelevata in C ed in S_1 , ma potremmo prelevare il segnale anche altrove. Quindi, è disponibile tra le altre anche una funzione di trasferimento tra l'ingresso E e l'uscita C , che ha sempre gli stessi poli,

10.4 IL FILTRO A TRALICCIO RICORSIVO

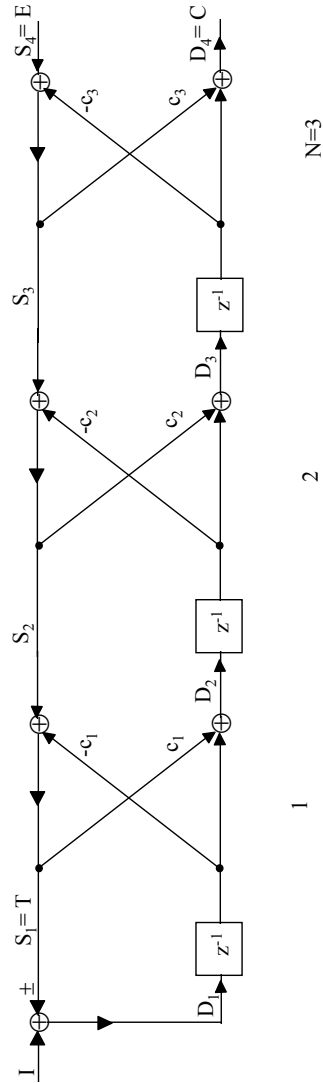


Figura 10.7: Realizzazione di filtri tutti poli con e senza ritardo e di sfasatori puri come cascata di celle elementari

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

ma è uno sfasatore puro, per motivi fisici già discussi in precedenza e che saranno rivisti ancora.

Le equazioni che governano il sistema numerico senza i blocchi Q_n , $1/Q_n$ sono facilmente deducibili dallo schema a blocchi di figura 8.5:

$$\begin{bmatrix} S_{n+1}(z) \\ D_{n+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & c_n z^{-1} \\ c_n & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_n(z) \\ D_n(z) \end{bmatrix} \quad (10.11)$$

Questi filtri vengono chiamati **filtri a traliccio ricorsivi**; se sono di ordine N , (pari a 3 in figura 8.7) sono necessarie $2N$ moltiplicazioni e $2N$ addizioni: essi permettono di realizzare sfasatori puri e funzioni di trasferimento tutti poli (con e senza ritardo) ma, a pari ordine, costano di più in termini di operazioni da compiere. Tuttavia, le moltiplicazioni sono tutte per fattori inferiori ad 1 (i coefficienti di riflessione c_n); **la natura fisica della struttura evita così sia cicli limite che instabilità** purchè, anche per valori dei c_n molto prossimi all'unità, il segnale riflesso sia sempre inferiore (in modulo) al segnale incidente.

In questo modo è possibile garantire l'incondizionata stabilità del sistema numerico, per rozza che sia la quantizzazione di coefficienti e segnali. Ancora, potremo rendere questi sistemi tempo varianti e potremo garantirne la stabilità in ogni istante, sempre che naturalmente $|c_n| \leq 1$. Ancora sarà possibile, imponendo $c_N = \pm 1$, ottenere filtri IIR (riverberazioni senza perdite di energia) che si comportano da oscillatori puri, non divergenti se $|c_n| \leq 1, n < N$; si è poi visto nel capitolo 5 che si possono ottenere filtri IIR di tipo Butterworth, Chebyshev o ellittici come somme e differenze di sfasatori puri.

10.5 La ricorsione di Levinson

Nel paragrafo precedente, abbiamo visto un filtro numerico derivante dalla cascata di più celle corrispondenti a tronchi di linea, chiuso da un lato con una cella corrispondente ad un coefficiente di riflessione unitario (disadattamento da un lato). Le funzioni di trasferimento di questo sistema numerico, a seconda dell'ingresso scelto, sono tutti poli (con o senza un ritardo di N campioni) o sfasatore puro; esse sono determinabili in modo semplice a partire dai coefficienti di riflessione delle varie celle, attraver-

10.5 LA RICORSIONE DI LEVINSON

so una formula ricorsiva nota come **ricorsione di Levinson**. Questa formula verrà poi rincontrata nello studio della predizione lineare.

Consideriamo quindi il caso in cui la struttura riverberante sia chiusa da un lato da un coefficiente di riflessione unitario: numereremo i tronchi di linea a partire dal coefficiente unitario, che diventa quindi il coefficiente $c_0 = \pm 1$; in questo modo possiamo allungare la struttura riverberante senza dovere rinumerare i coefficienti. I sistemi meccanici corrispondenti a questo sistema numerico sono molti: nel caso acustico, questo sistema corrisponde ad uno strumento musicale a fiato. La riflessione totale su c_0 corrisponde a quella delle vibrazioni acustiche sull'ancia o sulle labbra del suonatore che chiudono lo strumento (trombone, clarino, oboe, sassofono ecc.). La bocca dello strumento invece corrisponde alla cella adattata ed è svasata appunto per creare adattamento con la propagazione nell'atmosfera. La variabilità delle riverberazioni e quindi dei suoni generati è ottenuta alterando i coefficienti di riflessione (otturando i fori del flauto, allungando il trombone, ecc.). Nel caso del tratto vocale che vedremo al capitolo 15, la struttura riverberante è aperta sullo spazio libero dal lato della bocca. E' questa l'interfaccia caratterizzata da un coefficiente di riflessione pressochè unitario, e quindi, nel caso del tratto vocale numereremo i coefficienti di riflessione a partire dalle labbra, andando verso la gola. Un terzo esempio è quello delle onde sismiche che si propagano nel terreno; la superficie libera del terreno costituisce praticamente un riflettore ideale per le onde che, provenienti dal basso, sono riflesse di nuovo verso il basso.

Il segno del coefficiente di riflessione della cella disadattata iniziale, di modulo unitario, dipende dal tipo di onda che si riflette: nel caso del tratto vocale, avremo un minimo di pressione e quindi le onde di pressione vedono un coefficiente di riflessione $c_0 = -1$; avremo invece un massimo di velocità e per queste onde il coefficiente di riflessione sarà $+1$.

Passando ora al sistema numerico corrispondente senza i blocchi Q_n , $1/Q_n$, indichiamo con $I(z)$ l'ingresso dalla parte disadattata (fig. 8.7) e con $E(z)$ e $C(z)$ rispettivamente l'ingresso e l'uscita dalla parte adattata. La funzione di trasferimento tra ingresso dalla parte disadattata $I(z)$ ed uscita dalla parte adattata $C(z)$ ha solo poli, a parte un ritardo di N campioni ed un fattore di scala; la funzione di trasferimento tra $E(z)$ (ingresso dalla parte adattata) e $C(z)$ è invece uno sfasatore puro, perchè tutta l'energia viene riflessa all'indietro dal coefficiente di riflessione a modulo unitario della cella iniziale; infine, la funzione di trasferimento

CAPITOLO 10 LE STRUTTURE RIVERBERANTI

tra l'ingresso $E(z)$ e l'uscita prelevata dalla parte disadattata in $S_1(z)$ è sempre con gli stessi N poli, ma senza ritardo.

Sempre riferendosi alla figura 8.7 ed alle equazioni (10.11) supponiamo che c_2 e c_3 siano nulli. Le funzione di trasferimento $S_1(z)/S_2(z)$, $D_2(z)/S_2(z)$, $D_2(z)/I(z)$ sono rispettivamente:

$$D_1(z) = I(z) + c_0 S_1(z)$$

Se $S_2(z) = 0$:

$$D_2(z) = I(z) \times \frac{(1 - c_1^2)z^{-1}}{A_1(z)}$$

e se $I(z) = 0$:

$$S_1(z) = S_2(z) \times \frac{1}{A_1(z)}$$

$$D_2(z) = S_2(z) \times \frac{\tilde{A}_1(z)}{A_1(z)}$$

dove

$$A_1(z) = 1 + c_0 c_1 z^{-1}$$

e la sequenza

$$\tilde{A}_N(z) = z^{-N} A_N(z^{-1}) \quad (10.12)$$

è la sequenza $A_N(z)$, ribaltata nei tempi e ritardata di N campioni; in altre parole è la sequenza che si ottiene scambiando l'ordine dei coefficienti. Inoltre, $c_0 = 1$ è il coefficiente di riflessione corrispondente al disadattamento all'ingresso. Supponiamo ora che c_2 sia diverso da 0; si verifica (utilizzando la (10.9)) che:

$$S_3(z) = S_2(z) \times \frac{A_2(z)}{A_1(z)} = S_1(z) A_1(z) \times \frac{A_2(z)}{A_1(z)} = S_1(z) A_2(z)$$

$$D_3(z) = S_3(z) \times \frac{\tilde{A}_2(z)}{A_2(z)}$$

$$D_3(z) = I(z) \times \frac{(1 - c_1^2)(1 - c_2^2)z^{-2}}{A_2(z)}$$

a seconda che siano diversi da 0, $I(z)$ oppure $S_3(z)$. Inoltre,

$$A_2(z) = A_1(z) + c_0 c_2 z^{-1} \tilde{A}_1(z)$$

10.5 LA RICORSIONE DI LEVINSON

Proseguendo ricorsivamente si trova che se:

$$D_N(z) = S_N(z) \times c_0 \times \frac{\tilde{A}_{N-1}(z)}{A_{N-1}(z)} \quad (10.13)$$

$$\begin{bmatrix} S_{N+1}(z) \\ D_{N+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & c_N z^{-1} \\ c_N & z^{-1} \end{bmatrix} \times \begin{bmatrix} S_N(z) \\ c_0 S_N(z) \tilde{A}_{N-1}(z) / A_{N-1}(z) \end{bmatrix}$$

ed eliminando $S_N(z)$, si trova:

$$D_{N+1}(z) = S_{N+1}(z) \times c_0 \frac{\tilde{A}_N(z)}{A_N(z)}$$

dove:

$$A_N(z) = A_{N-1}(z) + z^{-1} c_0 c_N \tilde{A}_{N-1}(z) = 1 + \dots + z^{-N} c_0 c_N \quad (10.14)$$

Inoltre:

$$S_1(z) = S_{N+1}(z) \times \frac{1}{A_N(z)}$$

$$D_{N+1}(z) = I(z) \times \frac{z^{-N} \prod_{n=1}^N (1 - c_n^2)}{A_N(z)} \quad (10.15)$$

L'equazione (10.14) è nota come la ricorsione di Levinson e collega la funzione di trasferimento di un sistema riverberante a N celle disadattato da un lato con quello che si ottiene aggiungendo una cella con coefficiente di riflessione c_{N+1} (per le onde che vanno verso il disadattamento). Se il segno di c_0 è negativo, nella formula compare il segno negativo. I **polinomi** $A_N(z)$ **sono a minima fase**, per motivi fisici (*le trombe non esplodono*); procedendo matematicamente si ha:

$$A_{N+1}(z) = A_N(z) [1 + c_0 c_{N+1} z^{-1} \frac{\tilde{A}_N(z)}{A_N(z)}] = A_N(z) \times [1 + \mu_{N+1}(z)] \quad (10.16)$$

Se $A_N(z)$ è a minima fase dimostriamo che anche $A_{N+1}(z)$ lo sarà; infatti è prodotto di una funzione a minima fase per una funzione $[1 + \mu_{N+1}(z)]$ anche essa a fase minima; infatti purchè sia:

$$|\mu_{N+1}(z)| = |c_{N+1}| \leq 1 \quad (10.17)$$

e la fase del vettore $[1 + \mu_{N+1}(z)]$, somma del vettore unitario e di un vettore più piccolo, è compresa tra $\pm \pi/2$ e quindi non supera π ed è minima.

10.5.1 Aggiunta di tronchi di linea

L'equazione (10.14) è importante perchè ci permette di capire che il comportamento passa basso o passa alto del sistema è condizionato dal coefficiente di riflessione più vicino al disadattamento (*per generare un suono più basso (uuuu) o più acuto (iiiiii), si forma una cavità più grande o più piccola vicino ai denti*):

$$|A_{N+1}(z)|^2 = |A_N(z)|^2(1 + c_{N+1}^2) + 2\operatorname{Re}[c_{N+1}z^{N+1}A_N(z)^2] \quad (10.18)$$

Aggiungendo il coefficiente di riflessione c_{N+1} , a meno del fattore di scala $1 + c_{N+1}^2$, lo spettro rimane immutato, con l'aggiunta del termine:

$$2\operatorname{Re}[c_{N+1}z^{N+1}A_N(z)^2] \quad (10.19)$$

che, a causa del fattore $\operatorname{Re}(z^{N+1} \dots)$ ha circa $(N+1)/2$ ondulazioni tra la frequenza 0 e la frequenza di Nyquist f_N . Quindi, il primo coefficiente di riflessione caratterizza l'andamento globale con la frequenza; i successivi coefficienti provocano una sempre maggiore *ondulazione* dell'andamento dello spettro.

10.6 Calcolo dei coefficienti di riflessione

Siamo ora in grado di determinare i coefficienti di riflessione c_n ; $n = 1, \dots, N$ di una struttura riverberante disadattata da un lato in modo che i poli del sistema siano le radici di un polinomio $A_N(z)$: basta applicare la ricorsione di Levinson all'incontrario, partendo da N per arrivare ad 1, p.e. supponendo $c_0 = 1$. Si può quindi verificare (come proposto da Schur e Cohn) se un polinomio qualsiasi, il cui primo coefficiente è 1 come accade per il polinomio $A_N(z)$, sia a minima fase. Infatti, basta verificare che tutti i coefficienti di riflessione c_i siano in modulo minori di 1. Anzitutto si osservi (10.14) che:

$$c_N = a_{N,N} \quad (10.20)$$

e cioè che l'ultimo coefficiente del polinomio è un coefficiente di riflessione e pertanto deve essere in modulo inferiore ad 1.

Per calcolare c_{N-1} dovremo calcolare $A_{N-1}(z)$. Per fare ciò è conveniente scrivere i coefficienti dei polinomi $A_N(z)$ ed $A_{N-1}(z)$ in termini di

10.6 CALCOLO DEI COEFFICIENTI DI RIFLESSIONE

vettori colonna. La ricorsione di Levinson (10.14) diventa:

$$\begin{bmatrix} 1 \\ a_{1,N} \\ a_{2,N} \\ \dots \\ a_{N-1,N} \\ a_{N,N} \end{bmatrix} = \begin{bmatrix} 1 \\ a_{1,N-1} \\ a_{2,N-1} \\ \dots \\ a_{N-1,N-1} \\ 0 \end{bmatrix} + c_N \begin{bmatrix} 0 \\ a_{N-1,N-1} \\ a_{N-2,N-1} \\ \dots \\ a_{1,N-1} \\ 1 \end{bmatrix} \quad (10.21)$$

Per calcolare i coefficienti di $A_{N-1}(z)$, ad esempio $a_{h,N-1}$ dovremo risolvere il sistema:

$$a_{h,N} = a_{h,N-1} + c_N a_{N-h,N-1} \quad (10.22)$$

$$a_{N-h,N} = a_{N-h,N-1} + c_N a_{h,N-1} \quad (10.23)$$

da cui risulta :

$$a_{h,N-1} = \frac{a_{h,N} - c_N a_{N-h,N}}{1 - c_N^2} \quad (10.24)$$

Troveremo così $c_{N-1} = a_{N-1,N-1}$ e poi progressivamente tutti i coefficienti di riflessione; se qualcuno di essi ha modulo maggiore di 1, il polinomio di partenza $A_N(z)$ non è a minima fase (la verifica di Schur - Cohn).