

Indice

I	Principi della modulazione analogica	5
1	Introduzione	7
2	Modulazione di ampiezza	9
2.1	Modulazione double side band suppressed carrier (DSB-SC)	9
2.2	Modulazione double side band transmitted carrier (DSB-TC)	13
2.3	Realizzazione di un modulatore di ampiezza	16
2.4	Realizzazione di un demodulatore di ampiezza	20
2.5	Recupero della portante nei sistemi AM	26
2.6	Modulazioni di ampiezza lineari: schema generale	28
2.7	Single side band transmission carrier (SSB-TC)	37
2.8	Quadrature amplitude modulation (QAM)	37
2.9	Accesso multiplo a divisione di frequenza (FDM)	40
2.10	Prestazioni	41
3	Modulazione angolare	47
3.1	Definizioni e parametri associati	47
3.2	Narrowband FM	51
3.3	Wideband FM	53
3.4	Modulatori	55
3.5	Demodulatori	58
3.6	Prestazioni	59
3.7	Preenfasi e deenfasi nella FM	64
4	Confronto tra i vari metodi di modulazione ed esempi di sistemi	65
4.1	Esempi di sistemi di trasmissione analogici	65
4.1.1	Ricevitore supereterodina	66
4.1.2	Radio FM	69
4.1.3	Radio FM stereo	69
4.1.4	Segnale televisivo	70
II	Esercizi sui filtri numerici e FFT	73
5	Filtri numerici	75
5.1	Stabilità BIBO di un filtro numerico	75
5.2	Risposta in frequenza razionale	76
5.3	Equazioni alle differenze	79
5.4	L'uso di MATLAB per il progetto dei filtri	82

6	Esercizi sulla FFT	97
	Bibliografia	103

Note sulla modulazione analogica e il progetto di filtri numerici. La parte relativa alla modulazione è a cura del prof. N. Benvenuto. La parte relativa ai filtri numerici è a cura del prof. R. Rinaldo.

Parte I

**Principi della modulazione
analogica**

Capitolo 1

Introduzione

In questo capitolo riportiamo le tecniche di modulazione analogiche in cui il segnale di informazione è continuo sia nel tempo che in ampiezza. Innanzitutto chiariamo il significato del termine modulazione: essa è una trasformazione che elabora il segnale di informazione, ad esempio voce o video, per renderlo atto ad essere inviato sul canale di trasmissione. Di fatto il segnale di informazione prodotto dalla sorgente è tipicamente di tipo passa basso mentre il canale, ad esempio radio, è di tipo banda passante. Sorge allora la necessità di traslare in frequenza il segnale prima di inviarlo sul canale trasmissivo. Il modello di un sistema di comunicazione è riportato in Figura 1.1 in cui $a(t)$ è il segnale di informazione o segnale modulante mentre $s(t)$ è il segnale modulato ottenuto come trasformazione di $a(t)$.

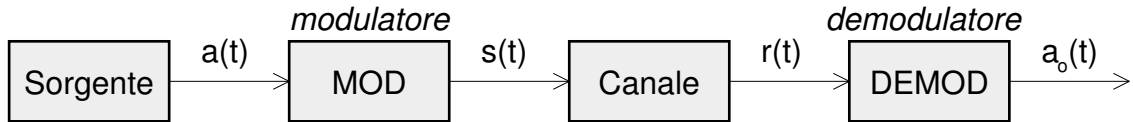


Figura 1.1: Modello di un sistema di trasmissione analogico.

Il canale, in generale, distorce $s(t)$ ed introduce del rumore. In queste note supporremo il canale non distortore per cui il segnale all'ingresso del ricevitore $r(t)$ sarà dato da

$$r(t) = \mathcal{G}_{Ch,0} s(t) + w(t) \quad , \quad (1.1)$$

in cui $\mathcal{G}_{Ch,0}$ è il guadagno del canale mentre $w(t)$ è un rumore AWGN con PSD $\mathcal{P}_w(f) = N_0/2$. Più comunemente si esprime $\mathcal{G}_{Ch,0}$ in termini dell'attenuazione di potenza \mathbf{a}_d del canale, come

$$\mathcal{G}_{Ch,0} = \frac{1}{\sqrt{\mathbf{a}_d}} \quad , \quad (1.2)$$

per cui in dB risulta

$$(\mathcal{G}_{Ch,0})_{dB} = -(\mathbf{a}_d)_{dB} \quad . \quad (1.3)$$

In base al segnale ricevuto $r(t)$, il compito del ricevitore è di ripristinare il segnale di informazione $a(t)$ fornendo una replica $a_o(t)$.

Le tecniche di modulazione analogiche fanno riferimento ad un segnale sinusoidale, detta portante, del tipo

$$p(t) = A \cos(2\pi f_0 t + \varphi_0) \quad , \quad (1.4)$$

e ne variano l'involuppo istantaneo A , fase istantanea φ_0 e frequenza istantanea f_0 [1] in modo proporzionale al segnale modulante $a(t)$. Ciò corrisponde rispettivamente alle modulazioni di ampiezza (*amplitude modulation*, AM), di fase (*phase modulation*, PM) e di frequenza (*frequency modulation*, FM). Vedremo che in effetti le PM e FM sono caratterizzazioni della modulazione angolare, per cui verranno analizzate congiuntamente.

Anticipiamo che la modulazione, oltre che per traslare in frequenza il segnale modulante, viene utilizzata anche per altri due motivi: 1) stabilire un *trade-off* tra la banda richiesta al canale e potenza trasmessa e 2) semplificare la realizzazione del demodulatore.

Nel seguito assumeremo che il segnale modulante $a(t)$, a valori reali, sia del tipo banda passante con una banda B come illustrato in Figura 1.2. In effetti, tipicamente la banda passante di $a(t)$ va da una frequenza $f_{a,1}$ ad $f_{a,2}$ con $f_{a,1} > 0$. In altre parole $a(t)$ non ha un contenuto spettrale attorno la componente continua (DC). Un esempio è fornito dal segnale voce che, per trasmissioni telefoniche ha una banda passante che va da 300 a 3400 Hz. D'altra parte, poichè $f_{a,1} \ll f_{a,2}$, assumeremo inoltre che la banda di $a(t)$ sia data da $B = f_{a,2} - f_{a,1} \simeq f_{a,2}$. In tutto questo capitolo assumeremo inoltre che $f_0 > B$ per le modulazioni di ampiezze e $f_0 \gg B$ per le modulazioni angolari.

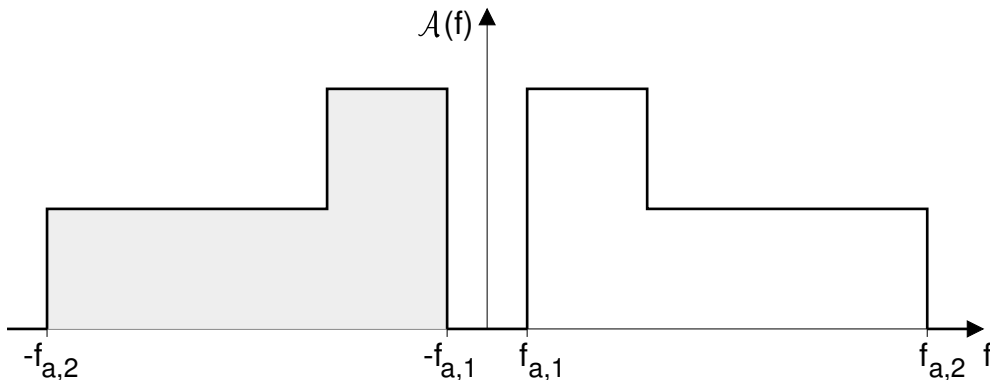


Figura 1.2: Andamento in frequenza di un tipico segnale modulante $a(t)$. Essendo $a(t)$ reale, la trasformata di Fourier $\mathcal{A}(f)$ risulta Hermitiana, con $\mathcal{A}(-f) = \mathcal{A}^*(f)$.

Iniziamo con l'illustrare il principio della tecnica di modulazione *double side band* (DSB) che sta alla base di tutte le modulazioni di ampiezze.

Capitolo 2

Modulazione di ampiezza

2.1 Modulazione double side band suppressed carrier (DSB-SC)

L'estensione *suppressed carrier* sarà chiarita in seguito. Il principio della modulazione DSB è riportato in Figura 2.1 mentre le trasformazioni dei vari segnali nel dominio della frequenza sono illustrate in Figura 2.2.

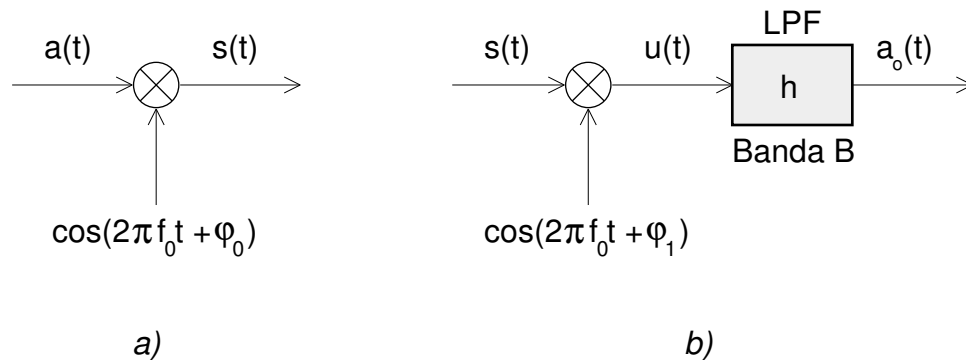


Figura 2.1: Modulazione DSB: a) modulatore e b) demodulatore.

Modulatore

Il segnale modulato è dato dal prodotto, effettuato tramite un *mixer*, tra il segnale modulante $a(t)$ e la portante,

$$s(t) = a(t) \cos(2\pi f_0 t + \varphi_0) \quad . \quad (2.1)$$

Nel dominio della frequenza avremo

$$\mathcal{S}(f) = \frac{e^{j\varphi_0}}{2} \mathcal{A}(f - f_0) + \frac{e^{-j\varphi_0}}{2} \mathcal{A}(f + f_0) \quad . \quad (2.2)$$

Questa trasformazione è riportata in Figura 2.2 utilizzando per $\mathcal{A}(f)$ l'andamento di Figura 1.2. Notiamo che la modulazione di $a(t)$ con la portante oltre ad attenuare il

segnale di un fattore $\frac{1}{2}$ ha traslato le frequenze di $s(t)$ in su e giù della frequenza della portante f_0 . Inoltre ha sfasato le componenti a frequenze positive di $s(t)$ di φ_0 e quelle a frequenze negative di $-\varphi_0$.

Una prima conseguenza di questa operazione è che il segnale modulato $s(t)$ ha una banda $2B$, doppia rispetto a quella del segnale modulante. Un primo beneficio è di avere un segnale ad alta frequenza che può essere trasmesso in modo efficiente. Ad esempio, nei sistemi radio, in cui l'antenna deve avere delle dimensioni legate alla lunghezza d'onda del segnale, si può utilizzare una antenna di dimensioni ragionevoli.

Un secondo beneficio della traslazione in frequenza è che più segnali possono condividere lo spettro radio senza interferenza reciproca, utilizzando per ciascun segnale una portante $f_{0,i}$ tale che i vari segnali modulati non si sovrappongano in frequenza. La Figura 2.3 illustra il segnale complessivo nel caso di due segnali.

Questo metodo di condividere il canale tra segnali di più utenti prende il nome di moltiplicazione a divisione di frequenza (*frequency division multiplexing*, FDM) e verrà approfondito in seguito.

In Figura 2.4 riportiamo un esempio della trasformazione $a \rightarrow s$ dato dalla (2.1) per un segnale $a(t)$ del tipo,

$$a(t) = \sin(2\pi f_1 t + \varphi_1) + 0.5 \sin(2\pi f_2 t + \varphi_2) + 0.25 \sin(2\pi f_3 t + \varphi_3) \quad , \quad (2.3)$$

con $f_1 = 400$ Hz, $f_2 = 500$ Hz, $f_3 = 600$ Hz, $\varphi_1 = \frac{\pi}{3}$, $\varphi_2 = \pi$, $\varphi_3 = \frac{\pi}{2}$. La portante è pari a $f_0 = 3500$ Hz. Notiamo che a meno del segno il segnale $a(t)$ è contenuto nell'involuppo del segnale $s(t)$.

Demodulatore

Come illustrato in Figura 2.1 il demodulatore, tramite un *mixer*, effettua il prodotto tra il segnale modulato $s(t)$ e la portante $\cos(2\pi f_0 t + \varphi_1)$ che deve avere la stessa frequenza e fase di quella trasmessa. Comunque per analizzare gli effetti di una possibile differenza di fase, assumeremo che la fase della portante in ricezione φ_1 possa essere diversa da quella in trasmissione. L'uscita del *mixer* viene filtrata da un filtro passa basso di banda B , pari a quella di $a(t)$, per fornire una stima di $a(t)$ che indicheremo con $a_o(t)$. Analiticamente, il segnale all'uscita del *mixer* è dato da

$$\begin{aligned} u(t) &= s(t) \cos(2\pi f_0 t + \varphi_1) \\ &= a(t) \cos(2\pi f_0 t + \varphi_0) \cos(2\pi f_0 t + \varphi_1) \\ &= \frac{a(t)}{2} \left[\cos(\varphi_0 - \varphi_1) + \cos(2\pi 2f_0 t + \varphi_0 + \varphi_1) \right] \quad , \quad (2.4) \end{aligned}$$

utilizzando note identità trigonometriche. Riconosciamo che in (2.4) $\frac{a(t)}{2} \cos(\varphi_0 - \varphi_1)$ è proporzionale al segnale modulante mentre $\frac{a(t)}{2} \cos(2\pi 2f_0 t + \varphi_0 + \varphi_1)$ rappresenta $a(t)$ traslato in frequenza attorno a $\pm 2f_0$. Utilizzando un filtro h del tipo passa basso con banda passante $(0, B)$, per non distorcere il primo termine in (2.4), e banda attenuata $(2f_0 - B, 2f_0 + B)$, per attenuare il secondo termine in (2.4), all'uscita del filtro, di guadagno unitario, avremo

$$a_o(t) = a(t) \frac{\cos(\varphi_0 - \varphi_1)}{2} \quad . \quad (2.5)$$

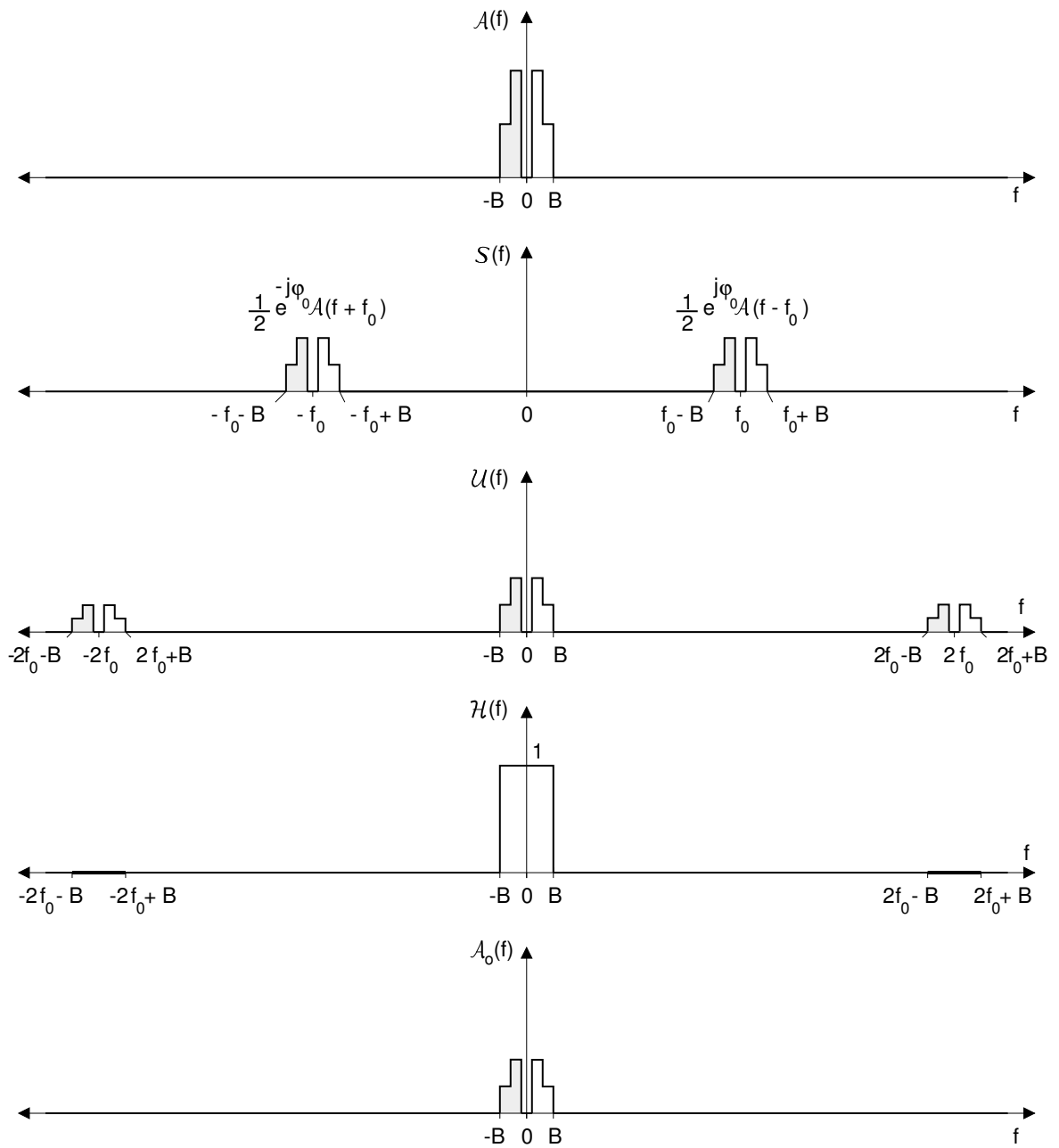


Figura 2.2: Illustrazione dei vari segnali nel dominio della frequenza in una modulazione DSB.

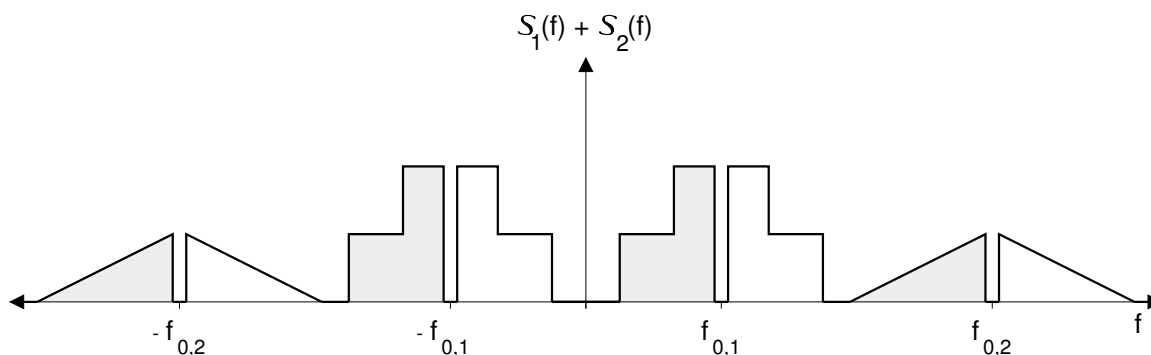


Figura 2.3: Principio della modulazione a divisione di frequenza.

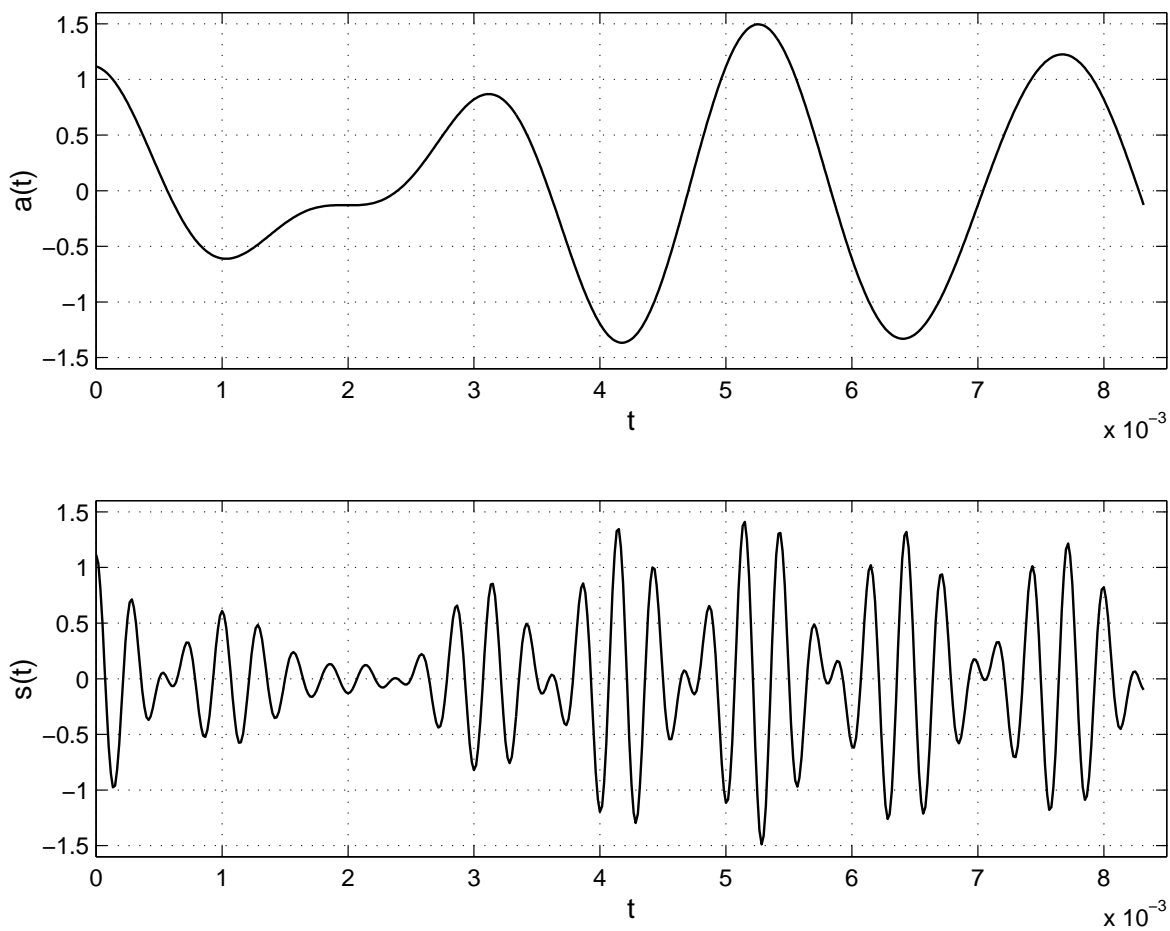


Figura 2.4: Esempio di segnale di informazione $a(t)$ e corrispondente segnale modulato DSB.

In effetti, notiamo che il segnale modulante viene ricostruito a meno di una costante pari a $\frac{\cos(\varphi_0 - \varphi_1)}{2}$. Di conseguenza è molto importante che $\varphi_1 = \varphi_0$, altrimenti il segnale ricostruito risulta ulteriormente attenuato.

2.2 Modulazione double side band transmitted carrier (DSB-TC)

Una formulazione generale della DSB porta ad avere un segnale modulato $s(t)$ del tipo

$$\begin{aligned} s(t) &= a(t) \cos(2\pi f_0 t + \varphi_0) + A \cos(2\pi f_0 t + \varphi_0) \\ &= (a(t) + A) \cos(2\pi f_0 t + \varphi_0) \quad , \end{aligned} \tag{2.6}$$

cioè oltre al segnale modulato per la portante viene trasmessa la portante stessa. Poichè la trasformata di Fourier della portante comporta un paio di delta di Dirach alle frequenze $\pm f_0$, l'andamento in frequenza di un tipico segnale modulato in ampiezza con portante trasmessa è riportato in Figura 2.5.

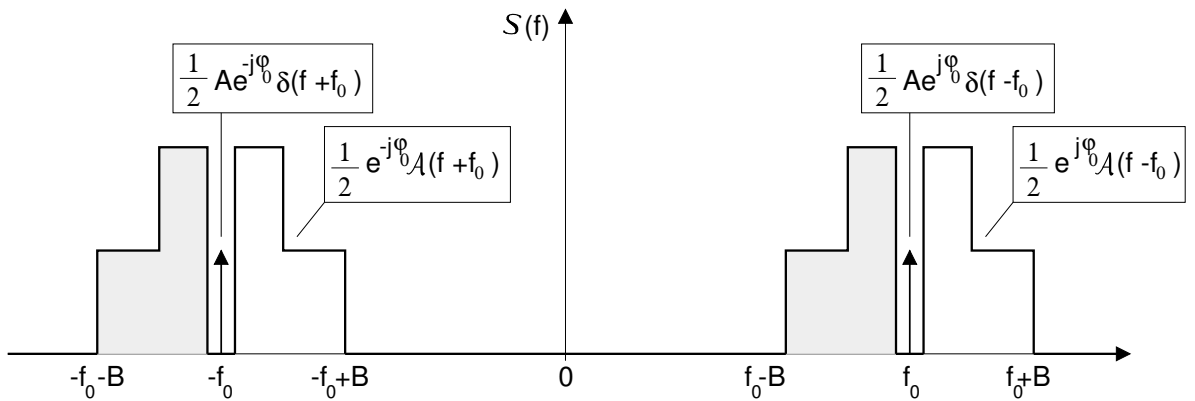


Figura 2.5: Andamento in frequenza di un segnale modulato DSB con trasmissione della portante.

Come nella DSB-SC, la banda della DSB-TC è pari a $2B$, doppia di quella del segnale modulante. Una diversità tra DSB-SC e DSB-TC è che in quest'ultima la trasformazione $a \rightarrow s$ non è più lineare.

Introduciamo il valore minimo di $a(t)$, in valore assoluto,

$$a_m = - \min_t a(t) \quad . \tag{2.7}$$

Se $a(t)$ è un processo aleatorio con ampiezza non limitata, a_m viene definito tramite la probabilità ε che il segnale assuma valori inferiori a $-a_m$,

$$P[a(t) < -a_m] = \varepsilon \quad , \tag{2.8}$$

con $\varepsilon \ll 1$. Tipicamente $\varepsilon \in \{10^{-3}, 10^{-4}\}$. Ad esempio se $a(t)$ è un processo aleatorio gaussiano con media nulla e varianza σ_a^2 risulta

$$a_m = \sigma_a Q^{-1}(\varepsilon) \quad , \tag{2.9}$$

dove Q^{-1} è la funzione inversa della funzione Q definita in [1]. Per $\varepsilon = 10^{-3}$ si ha che $Q^{-1}(10^{-3}) = 3.05$ e quindi $a_m \simeq 3.05 \sigma_a$. In altre parole le ampiezze di $a(t)$ sono inferiori a $-a_m = -3.05 \sigma_a$ con probabilità 10^{-3} .

Nel dominio del tempo, per un segnale $a(t)$ del tipo (2.3), riportiamo in Figura 2.6 il corrispondente segnale modulato DSB-TC nei due casi: *i*) per $A < a_m$ e *ii*) per $A > a_m$.

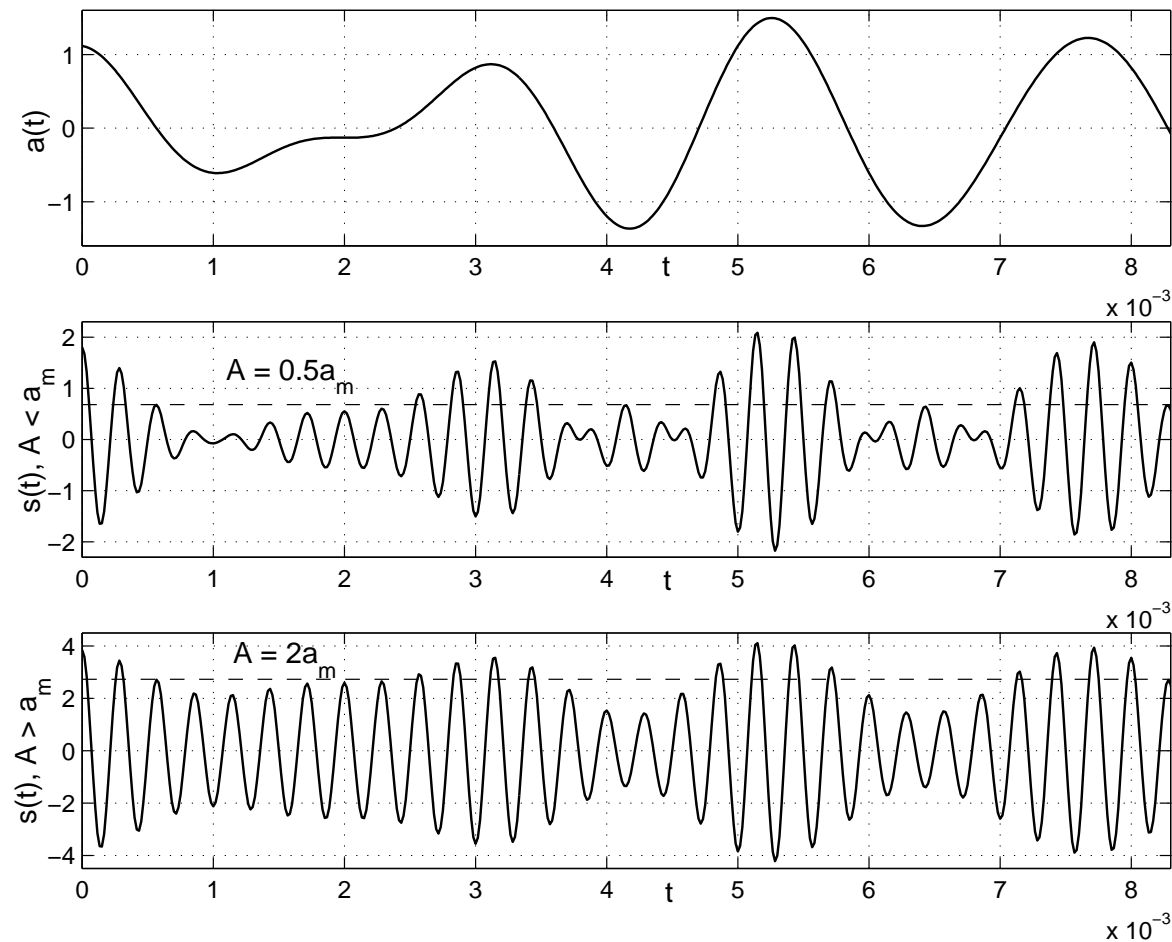


Figura 2.6: Andamento nel tempo del segnale modulante e corrispondente segnale modulato in ampiezza con trasmissione della portante (DSB-TC): *a*) segnale modulante $a(t)$, *b*) $s(t)$ per $A < a_m$, *c*) $s(t)$ per $A > a_m$.

Notiamo che nel caso $A > a_m$, denominata anche *condizione per l'assenza di distorsione dell'involuppo*, l'involuppo del segnale $s(t)$ coincide con $a(t) + A$. Nel caso $A < a_m$ invece l'involuppo di $s(t)$ coincide con $|a(t) + A|$.

In effetti è importante stabilire il rapporto tra a_m ed A che prende il nome di *indice di modulazione*

$$m = \frac{a_m}{A} \quad . \quad (2.10)$$

Notiamo che $m > 0$ e la scelta $A > a_m$ corrisponde ad un indice di modulazione inferiore ad 1.

Introduciamo inoltre il segnale di informazione normalizzato in ampiezza

$$\bar{a}(t) = \frac{a(t)}{a_m} \quad . \quad (2.11)$$

Assumendo, come si verifica di solito, che $\max_t a(t) = -\min_t a(t)$, $\bar{a}(t)$ risulta un segnale con ampiezza limitata -1 e 1; in ogni caso il valore minimo è -1.

Utilizzando la (2.10) e (2.11), una forma alternativa alla (2.6) è data da

$$s(t) = A (1 + m \bar{a}(t)) \cos (2\pi f_0 t + \varphi_0) \quad . \quad (2.12)$$

Questa formulazione è quella più generale di un segnale AM.

Introduciamo alcuni parametri associati al segnale (2.12).

POTENZA. Definiamo la potenza di un generico segnale $x(t)$

$$M_x = \lim_{T_w \rightarrow \infty} \frac{1}{T_w} \int_{-\frac{T_w}{2}}^{\frac{T_w}{2}} x^2(t) dt \quad . \quad (2.13)$$

Questa definizione coincide con la potenza statistica di un processo aleatorio definita come $E[x^2(t)]$, atteso che il processo aleatorio sia ergodico.

FATTORE DI FORMA. Associato al segnale normalizzato $\bar{a}(t)$, definiamo la sua potenza come

$$k_f^2 = M_{\bar{a}} = \frac{M_a}{a_m^2} \quad , \quad (2.14)$$

dove l'ultima uguaglianza discende dalla (2.11). Ad esempio se $a(t)$ è un segnale sinusoidale risulta $k_f = \frac{1}{\sqrt{2}}$, mentre se $a(t)$ ha una distribuzione di ampiezza uniforme tra $-a_m$ e a_m si ha che $k_f = \frac{1}{\sqrt{3}}$. Infine, se $a(t)$ ha una distribuzione di ampiezza gaussiana con media nulla allora $k_f = \frac{1}{Q^{-1}(\varepsilon)}$.

EFFICIENZA DELLA MODULAZIONE. In base alla struttura della (2.6) o (2.12) del segnale modulato, l'efficienza η della modulazione DSB-TC è definita dal rapporto tra la potenza del termine desiderato che porta l'informazione $a(t) \cos (2\pi f_0 t + \varphi_0)$ e la potenza del segnale complessivo $s(t)$. D'altra parte se $f_0 > B$ la potenza di $a(t) \cos (2\pi f_0 t + \varphi_0)$ è pari a $\frac{M_a}{2}$ per cui

$$\eta = \frac{\frac{M_a}{2}}{M_s} \quad . \quad (2.15)$$

Ora sempre sotto l'ipotesi che $f_0 > B$ dalla (2.12) risulta

$$M_s = \frac{A^2}{2} (1 + m^2 M_{\bar{a}}) \quad . \quad (2.16)$$

Utilizzando la (2.14) risulta

$$M_s = \frac{A^2}{2} (1 + m^2 k_f^2) \quad , \quad (2.17)$$

mentre dalla (2.11) si ha

$$M_a = a_m^2 k_f^2 \quad . \quad (2.18)$$

Sostituendo (2.17) e (2.18) in (2.15) ed utilizzando la definizione di indice di modulazione (2.10), risulta

$$\eta = \frac{m^2 k_f^2}{1 + m^2 k_f^2} . \quad (2.19)$$

Notiamo che più piccolo è l'indice di modulazione m minore è l'efficienza di modulazione η . In particolare, essendo $k_f < 1$, per $m < 1$ risulta $\eta < \frac{1}{2}$.

2.3 Realizzazione di un modulatore di ampiezza

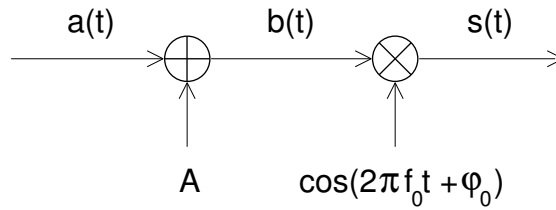


Figura 2.7: Modulatore di ampiezza con trasmissione della portante.

Uno schema che realizza la modulazione di ampiezza è riportato in Figura 2.7. Esso fa uso di un oscillatore a frequenza f_0 la cui uscita viene moltiplicata per il segnale $b(t) = a(t) + A$ tramite un *mixer*. Naturalmente per $A = 0$ abbiamo come caso particolare la DSB-SC.

Un'alternativa all'oscillatore che produce la portante è utilizzare un segnale periodico di periodo $T_0 = \frac{1}{f_0}$, ad esempio un treno di impulsi *rect* del tipo

$$p(t) = \sum_{n=-\infty}^{+\infty} \text{rect} \left(\frac{t - nT_0}{dT_0} \right) , \quad (2.20)$$

dove d , $0 < d < 1$, esprime la durata del generico impulso normalizzata al periodo T_0 di $p(t)$. Il segnale $p(t)$ è illustrato in Figura 2.8 per $d = 0.25$. Notiamo che se f_0 è molto elevata la durata del generico impulso *rect* può risultare molto corta. Utilizzando per $p(t)$ la espressione (2.20) il modulatore viene riportato in Figura 2.9. Il vantaggio di utilizzare come impulso fondamentale un *rect* è che il prodotto di $b(t)$ con $p(t)$ viene sostituito da un interruttore che “fa passare” il segnale $b(t)$ ogni T_0 secondi per una durata di dT_0 secondi.

Tramite lo sviluppo in serie di Fourier di $p(t)$ risulta

$$p(t) = d + 2d \sum_{k=1}^{\infty} \text{sinc}(kd) \cos(2\pi k f_0 t) . \quad (2.21)$$

Come illustrato in Figura 2.10, il prodotto di $p(t)$ con il segnale $b(t)$ produce in frequenza tante repliche di $\mathcal{B}(f)$ attorno alle frequenze kf_0 , $k = 0, 1, 2, \dots$, di ampiezza rispettivamente $d \text{sinc}(kd)$.

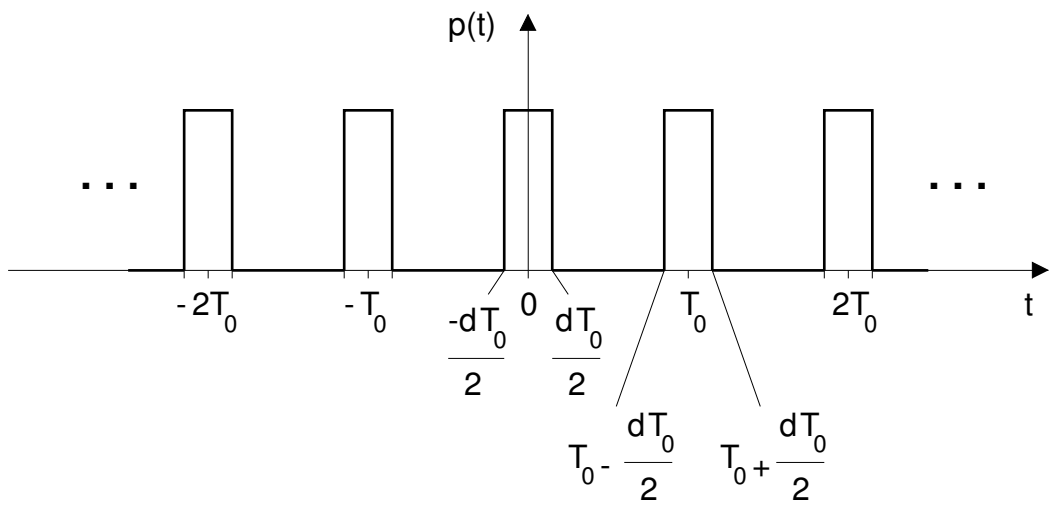


Figura 2.8: Treno di impulsi rect con $d = 0.25$.

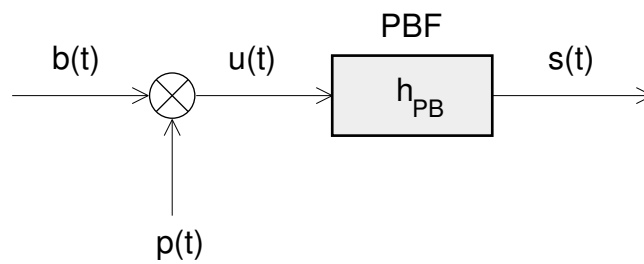


Figura 2.9: Modulatore di ampiezza a tenuta (*gated modulator*).

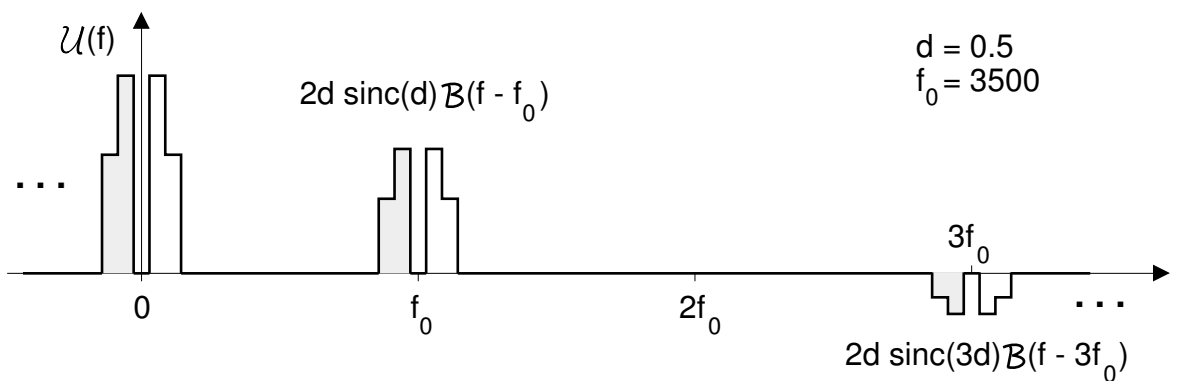


Figura 2.10: Andamento in frequenza del segnale $u(t)$, prodotto di $b(t)$ per un treno di impulsi rect di durata normalizzata $d = \frac{1}{2}$.

Per selezionare la replica attorno a f_0 è sufficiente utilizzare un filtro del tipo banda passante h_{PB} con centro banda f_0 e banda $2B$. Se il guadagno di h_{PB} è unitario alla sua uscita avremo

$$\begin{aligned} s(t) &= b(t) 2d \operatorname{sinc}(d) \cos(2\pi f_0 t) \\ &= b(t) \frac{2}{\pi} \sin(\pi d) \cos(2\pi f_0 t) \quad . \end{aligned} \quad (2.22)$$

Tipicamente $d = \frac{1}{2}$ e

$$s(t) = \frac{2}{\pi} b(t) \cos(2\pi f_0 t) \quad . \quad (2.23)$$

Le varie trasformazioni dello schema di Figura 2.9 sono illustrate in Figura 2.11.

Uno schema alternativo è rappresentato dal modulatore con legge quadratica (*square-law modulator*), riportato in Figura 2.12. A partire dal segnale modulante $b(t)$ e la portante, si effettua la loro somma,

$$v(t) = b(t) + \cos(2\pi f_0 t + \varphi_0) \quad . \quad (2.24)$$

Ora si trasforma $v(t)$ in modo non lineare, ad esempio tramite una legge quadratica

$$u(t) = v^2(t) \quad , \quad (2.25)$$

che tra i vari termini conterrà il prodotto desiderato tra $b(t)$ e $\cos(2\pi f_0 t + \varphi_0)$. Infatti la (2.24) in (2.25) fornisce

$$v^2(t) = b^2(t) + \cos^2(2\pi f_0 t + \varphi_0) + 2b(t) \cos(2\pi f_0 t + \varphi_0) \quad , \quad (2.26)$$

con un contenuto attorno DC di banda $2B$, una riga in $\pm 2f_0$ e il segnale modulato desiderato attorno a $\pm f_0$. Per selezionare il termine desiderato si utilizza un filtro del tipo banda passante centrato in f_0 .

Più in generale al posto della (2.25) è sufficiente utilizzare un elemento non lineare avente legame ingresso-uscita con uno sviluppo in serie del tipo

$$u(t) = c_1 v(t) + c_2 v^2(t) + c_3 v^3(t) + \dots \quad , \quad (2.27)$$

dove il termine desiderato è in $v^2(t)$. Ebbene, se c'è un modo per isolare questo termine l'elemento non lineare (2.27) può essere utilizzato al posto di quello quadratico (2.25). Ad esempio, un diodo a semiconduttore fornisce tra i terminali di uscita e di ingresso una legge non lineare data dai primi due termini della (2.27) e può quindi essere usato come approssimazione della legge quadratica. Notiamo che se in (2.27) con uno sviluppo fino al termine quadratico risulta $c_1 \simeq 0$, allora possiamo utilizzare un modulatore con legge quadratica per realizzare una DSB-SC. Nel caso $c_1 \neq 0$ avremo anche una riga attorno a $\pm f_0$ dovuta al termine $c_1 v(t)$. Se nell'elemento non lineare appaiono anche termini del terzo ordine che creano interferenza al termine desiderato, si può utilizzare il metodo di cancellazione di Figura 2.13 per rimuovere i termini con esponente dispari del tipo $b(t)$ e $b^3(t)$.

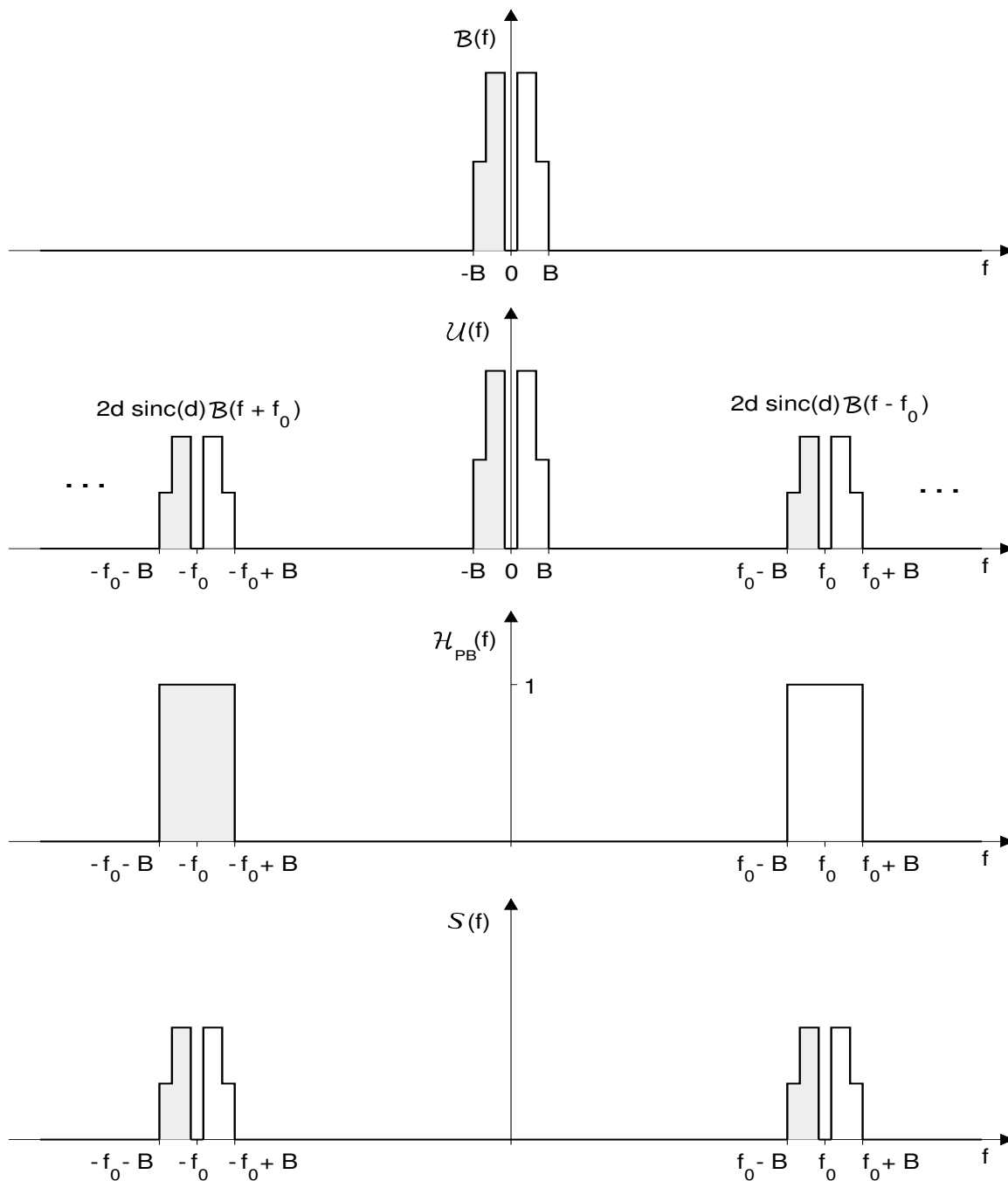


Figura 2.11: Illustrazione delle trasformazioni introdotte dallo schema di Figura 2.9.

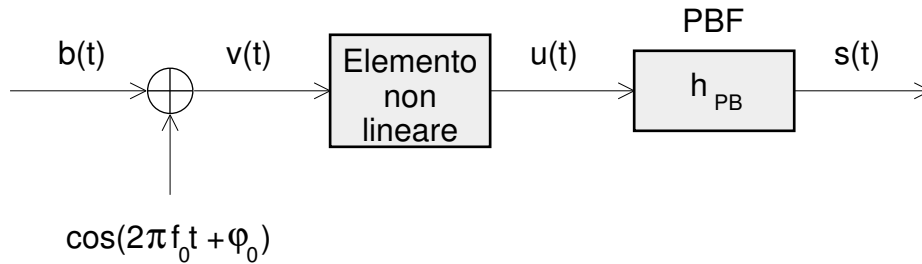


Figura 2.12: Modulatore di ampiezza con legge quadratica (*square-law modulator*).

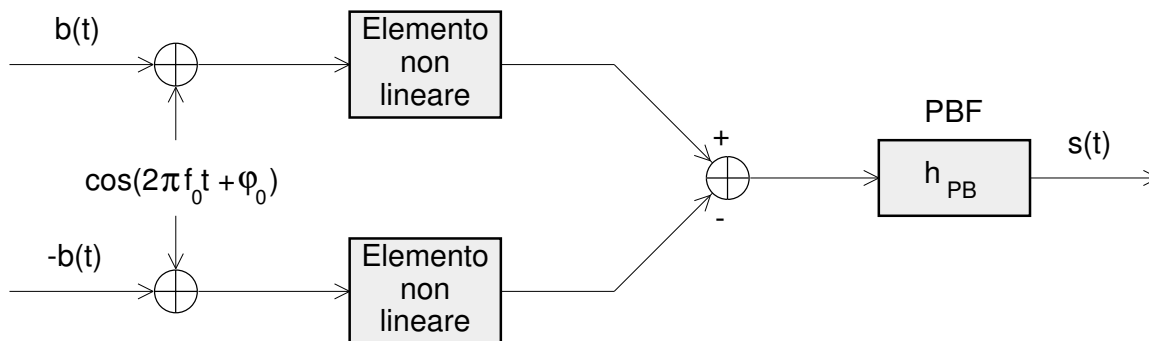


Figura 2.13: Modulatore di ampiezza bilanciato (*balanced modulator*).

2.4 Realizzazione di un demodulatore di ampiezza

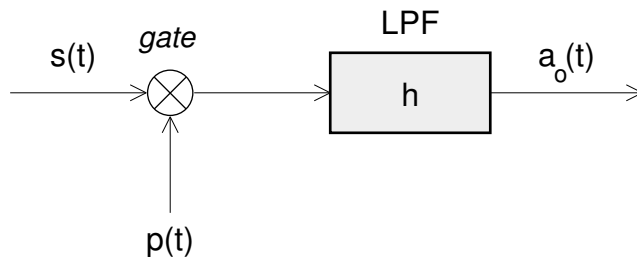
I demodulatori vengono suddivisi in due classi: *coerenti* e *non coerenti*. Nel caso coerente il ricevitore fa uso della conoscenza sia della frequenza che della fase della portante trasmessa¹. In altre parole, oltre alla forma d'onda della portante trasmessa (individuata dalla frequenza f_0) bisogna conoscere anche la temporizzazione (individuata dalla fase φ_0) con cui arriva la portante all'ingresso del ricevitore. Schemi non coerenti prescindono da queste conoscenze.

Demodulatori coerenti

Lo schema base di un demodulatore coerente è riportato in Figura 2.1b e consiste nel rimodulare $s(t)$ tramite la portante, o meglio una stima della portante, e filtrare il prodotto tramite un filtro passa basso.

Anche in questo caso il prodotto tra $s(t)$ e la portante locale può essere effettuato tramite uno degli schemi visti per il modulatore. Rivediamoli in dettaglio. Lo schema *gated demodulator* è riportato in Figura 2.14 dove $p(t)$ è un treno di impulsi sincronizzati con quelli di trasmissione.

¹In effetti dovremo dire della portante trasmessa valutata però all'uscita del canale. Nel nostro modello, essendo il canale ideale, le due locuzioni sono uguali.

Figura 2.14: *Gated demodulator*.

Per $p(t)$ del tipo (2.20) risulta

$$a_o(t) = \frac{\sin(\pi d)}{\pi} b(t) \quad , \quad (2.28)$$

dove $b(t) = a(t)$ nel caso DSB-SC mentre $b(t) = a(t) + A$ nel caso DSB-TC. In quest'ultimo caso per ottenere una replica di $a(t)$ bisogna rimuovere la componente continua di $a_o(t)$ tramite un filtro *notch* con guadagno nullo per $f = 0$ (DC). Naturalmente se $a(t)$ ha un contributo a DC il filtro notch introdurrà distorsione.

Lo schema *square-law demodulator*, riportato in Figura 2.15, è invece scarsamente utilizzato nei sistemi DSB-SC mentre si presta per sistemi DSB-TC con $A > a_m$ (in tal caso non serve la portante in ricezione e $K = 0$). Consideriamo la relazione generale

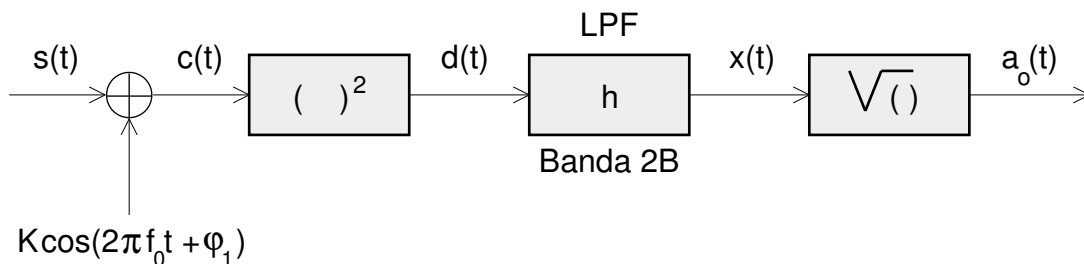
$$c(t) = s(t) + K \cos(2\pi f_0 t + \varphi_1) \quad , \quad (2.29)$$

per un elemento non lineare di tipo quadratico,

$$d(t) = c^2(t) \quad . \quad (2.30)$$

Un segnale DSB-SC $s(t) = a(t) \cos(2\pi f_0 t + \varphi_0)$ per $\varphi_1 = \varphi_0$ porge

$$\begin{aligned} d(t) &= (a(t) + K)^2 \cos^2(2\pi f_0 t + \varphi_0) \\ &= \frac{1}{2} (a(t) + K)^2 + \frac{1}{2} (a(t) + K)^2 \cos(2\pi 2f_0 t + 2\varphi_0) \quad . \end{aligned} \quad (2.31)$$

Figura 2.15: *Square-law demodulator* per DSB-SC.

Ora il secondo termine può essere rimosso da un filtro passa basso, alla cui uscita rimane

$$x(t) = \frac{1}{2} (a(t) + K)^2 = \frac{1}{2} (a^2(t) + 2K a(t) + K^2) \quad . \quad (2.32)$$

Purtroppo in (2.32) non è possibile isolare il termine $a(t)$ dal termine $a^2(t)$ poichè i due termini si sovrappongono in frequenza. Solo se $(a(t) + K)$ è non negativo², cioè per $K > a_m$, allora prendendo la radice quadrata di (2.32) avremo

$$a_o(t) = \sqrt{x(t)} = \frac{1}{\sqrt{2}} |a(t) + K| = \frac{1}{\sqrt{2}} (a(t) + K) \quad , \quad (2.33)$$

ottenendo un segnale legato ad $a(t)$ a meno della DC.

Notiamo che se in (2.29) utilizziamo una portante in ricezione con $\varphi_1 \neq \varphi_0$, allora potremo ricorrere al *quadrature demodulator* di Figura 2.16, per ottenere in generale, sia per DSB-SC che DSB-TC,

$$a_o(t) = \frac{1}{2} |b(t)| \quad . \quad (2.34)$$

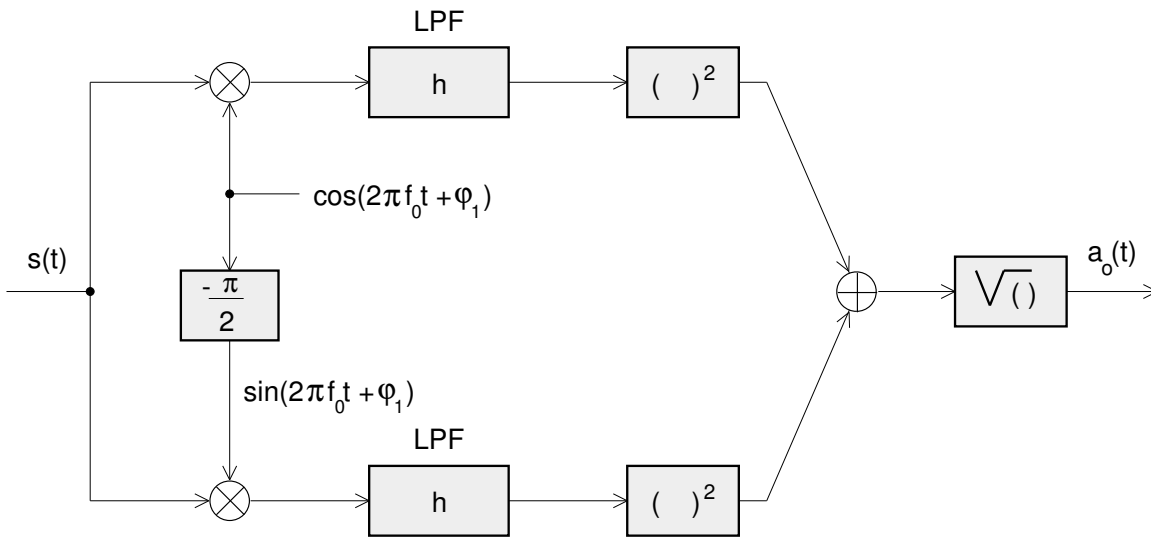


Figura 2.16: *Quadrature demodulator* per DSB-TC.

Anche questo schema però è di scarsa qualità se non per una modulazione DSB-TC con $A > a_m$.

Demodulatori non coerenti

Questi schemi prescindono dal recupero della portante in ricezione. L'inconveniente

²Sottolineiamo che in questa diseuguaglianza K è una ampiezza generata localmente al ricevitore mentre a_m è l'ampiezza minima del segnale ricevuto che spesso non è nota.

è che bisogna trasmettere la portante ad un livello sufficientemente elevato in modo che

$$a(t) + A \geq 0 \quad \forall t \quad , \quad (2.35)$$

cioè deve essere

$$A \geq a_m \quad , \quad (2.36)$$

dove a_m è definito in (2.7). Vediamo due schemi applicati alla DSB-TC. Lo *square-law demodulator* viene riportato in Figura 2.17 in cui $s(t) = (a(t) + A) \cos(2\pi f_0 t + \varphi_0)$.

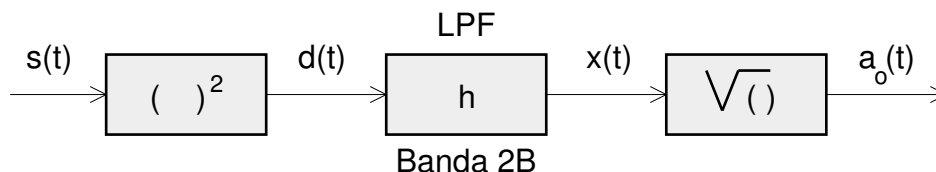


Figura 2.17: *Square-law demodulator* per segnali AM.

All'uscita del filtro h di tipo passa basso, di banda $2B$ per far passare le armoniche del segnale $(a(t) + A)^2$, avremo solo il termine a bassa frequenza $\frac{1}{2} (a(t) + A)^2$. Prendendo la radice quadrata avremo

$$a_o(t) = \frac{1}{\sqrt{2}} |a(t) + A| \quad . \quad (2.37)$$

Se $A > a_m$ il segnale all'uscita risulta proporzionale al segnale di informazione a meno di una componente continua, che al solito viene rimossa utilizzando un filtro *notch*,

$$a_o(t) = \frac{1}{\sqrt{2}} (a(t) + A) \quad . \quad (2.38)$$

Per un segnale $a(t)$ del tipo riportato in Figura 1.2, il filtro *notch* avrà le specifiche riportate in Figura 2.18.

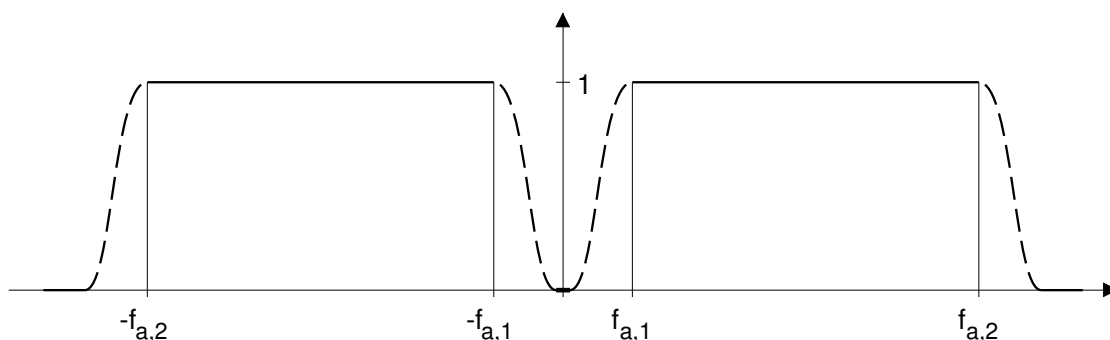


Figura 2.18: Specifiche del filtro *notch* per rimuovere la componente continua.

Al posto della funzione quadratica in Figura 2.17 si può utilizzare un raddrizzatore a onda piena o a semionda le cui relazioni ingresso-uscita sono le seguenti.

raddrizzatore a onda piena

$$d(t) = |s(t)| = \begin{cases} s(t) & , s(t) > 0 \\ -s(t) & , s(t) < 0 \end{cases} \quad (2.39)$$

raddrizzatore a semionda

$$d(t) = \begin{cases} s(t) & , s(t) > 0 \\ 0 & , s(t) < 0 \end{cases} \quad (2.40)$$

Consideriamo il primo caso per

$$d(t) = |s(t)| = |a(t) + A| |\cos(2\pi f_0 t + \varphi_0)| \quad (2.41)$$

Ora dalla (2.35) abbiamo che $|a(t) + A| = a(t) + A$. Inoltre, considerato lo sviluppo in serie di Fourier della funzione $|\cos(2\pi f_0 t + \varphi_0)|$, periodica di periodo $\frac{1}{2f_0}$, otteniamo

$$|\cos(2\pi f_0 t + \varphi_0)| = \frac{2}{\pi} + \sum_{k=1}^{\infty} c_k \cos(2\pi 2k f_0 t + \varphi_0) \quad (2.42)$$

Allora $d(t)$ è dato dalla somma di repliche di $b(t) = a(t) + A$ attorno DC, $\pm 2f_0$, $\pm 4f_0$, Se $2f_0 - B > B$, cioè per $f_0 > B$, all'uscita del LPF h rimarrà solo il primo termine e

$$a_o(t) = \frac{2}{\pi} (a(t) + A) \quad (2.43)$$

Questo schema è riportato in Figura 2.19 e prende il nome di *rectifier demodulator*.

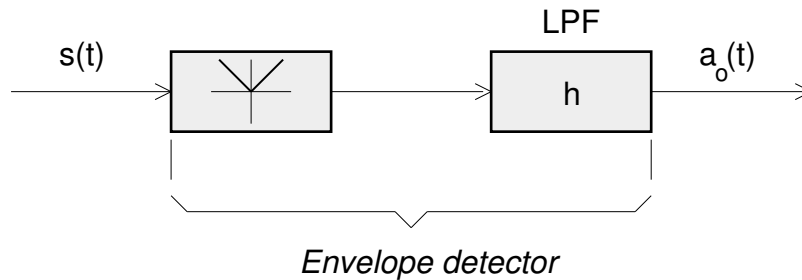


Figura 2.19: *Rectifier demodulator* per segnali AM.

Nel dominio del tempo le trasformazioni introdotte da un raddrizzatore a onda piena sono illustrate in Figura 2.20.

È interessante osservare che tale operazione coincide con moltiplicare $s(t)$ per un treno di impulsi avente la stessa frequenza, $\frac{1}{T_0} = f_0$, e fase della portante in ricezione, dato da

$$p_{Rc}(t) = \sum_{k=-\infty}^{+\infty} (-1)^k \text{rect}\left(\frac{t - k\frac{T_0}{2}}{dT_0}\right) \quad (2.44)$$

con $T_0 = \frac{1}{f_0}$ e $d = \frac{1}{2}$.

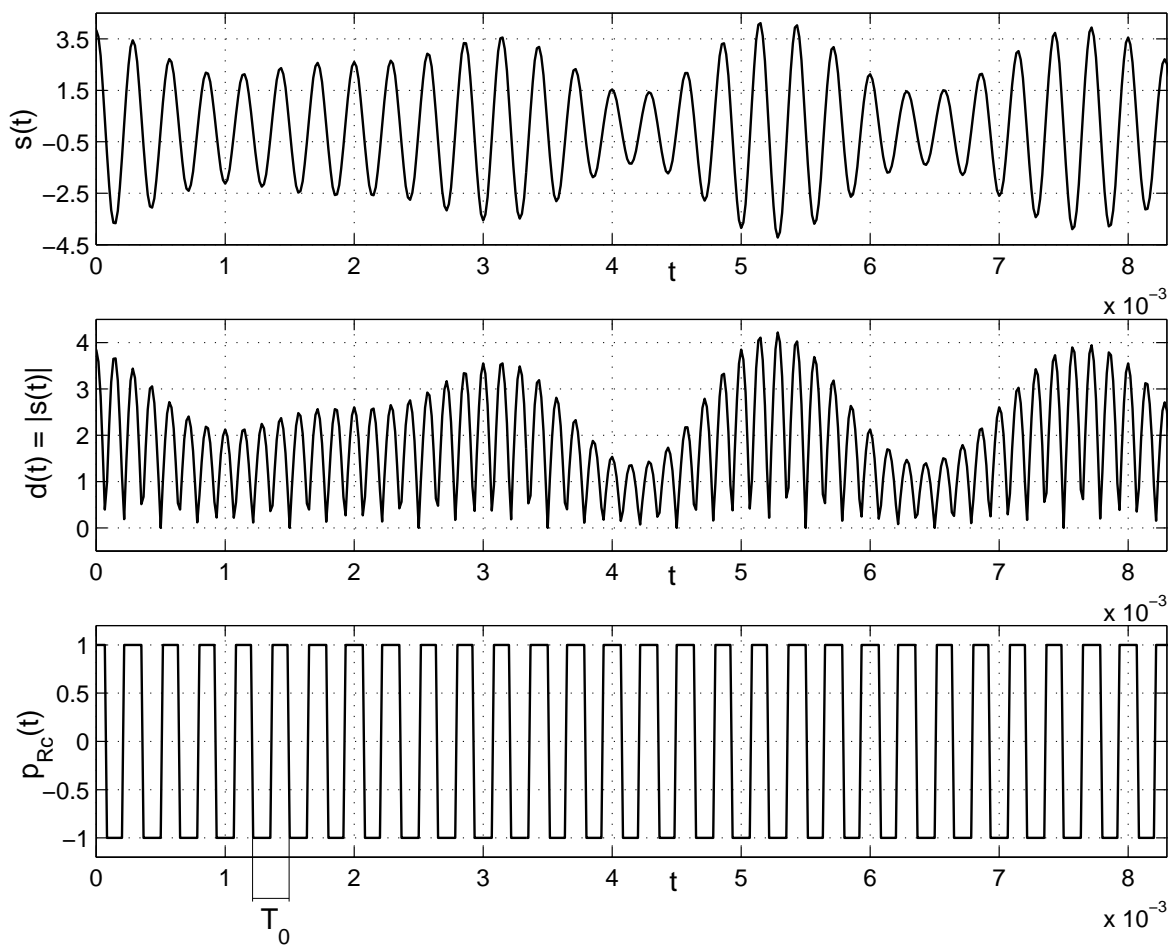


Figura 2.20: Andamento segnale AM e corrispondente segnale all'uscita del raddrizzatore a onda piena.

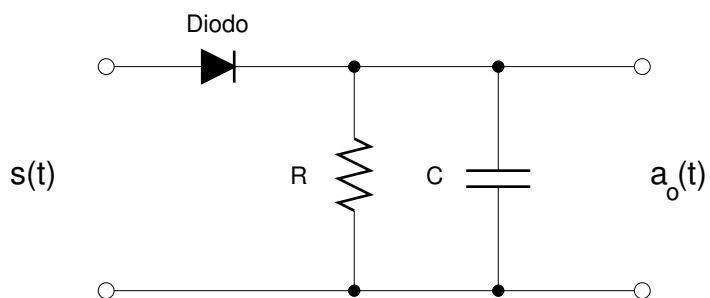


Figura 2.21: *Envelope detector*.

La cascata del raddrizzatore e del filtro h prende il nome di rivelatore di inviluppo o *envelope detector* poichè fornisce l'inviluppo del segnale $s(t)$. Se $f_0 \gg B$ l'*envelope detector* può essere realizzato tramite un semplice diodo seguito da un filtro RC, come illustrato in Figura 2.21.

In questo schema mentre il condensatore tende a caricare l'uscita $a_o(t)$ al valore massimo di ingresso, la resistenza permette un' scarica della tensione quando l'ingresso $s(t)$ è inferiore a $a_o(t)$. Le due operazioni sono illustrate in Figura 2.22. Naturalmente deve essere

$$B \ll \frac{1}{RC} \ll f_0 \quad . \quad (2.45)$$

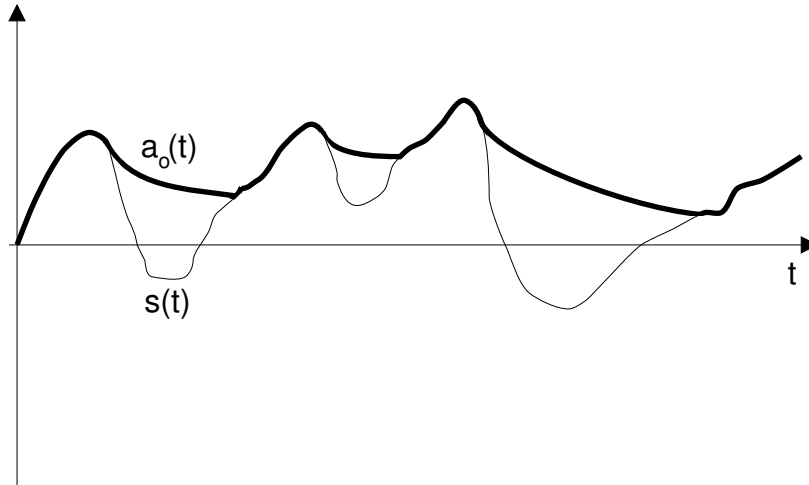


Figura 2.22: Illustrazione delle operazioni effettuate da un *envelope detector* nel dominio del tempo.

2.5 Recupero della portante nei sistemi AM

Abbiamo evidenziato precedentemente come sia importante nei sistemi AM con demodulazione coerente fare in modo che la frequenza e fase della portante in ricezione siano uguali a quelle della portante trasmessa. Per ottenere ciò è fondamentale derivare (stimare) questi parametri dal segnale ricevuto.

Per tutti i casi AM un primo schema è trasmettere assieme al segnale di informazione anche la portante, per cui otteniamo un segnale DSB-TC

$$s(t) = a(t) \cos(2\pi f_0 t + \varphi_0) + A \cos(2\pi f_0 t + \varphi_0) \quad . \quad (2.46)$$

Un metodo per estrarre la portante trasmessa è utilizzare un filtro a banda stretta, avente una banda B_{NBF} , sintonizzato sul valore nominale di f_0 . Come illustrato in Figura 2.23, all'uscita del filtro h_{NBF} , supposto ideale e con guadagno unitario, avremo

$$p_{Rc}(t) = A \cos(2\pi f_0 t + \varphi_0) + \int_{f_0 - \frac{B_{NBF}}{2}}^{f_0 + \frac{B_{NBF}}{2}} 2 \operatorname{Re} [\mathcal{A}(f - f_0) e^{j(2\pi f t + \varphi_0)}] df \quad . \quad (2.47)$$

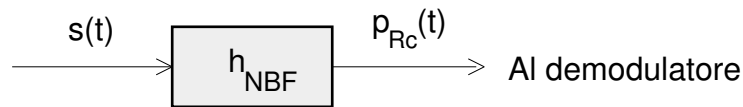


Figura 2.23: Filtro a banda stretta per estrarre la portante.

Se il filtro h_{NBF} ha una banda sufficientemente stretta e/o il contenuto di $a(t)$ è piccolo attorno DC, il secondo termine in (2.47) può essere reso sufficientemente piccolo rispetto al primo termine. Di conseguenza $p_{Rc}(t)$ risulta una replica della portante trasmessa.

Una alternativa al filtro a banda stretta è utilizzare un *phase locked loop* (PLL), come illustrato in Figura 2.24, il quale fa uso di un oscillatore controllato in tensione (*voltage controlled oscillator*, VCO) la cui fase istantanea insegue quella della componente

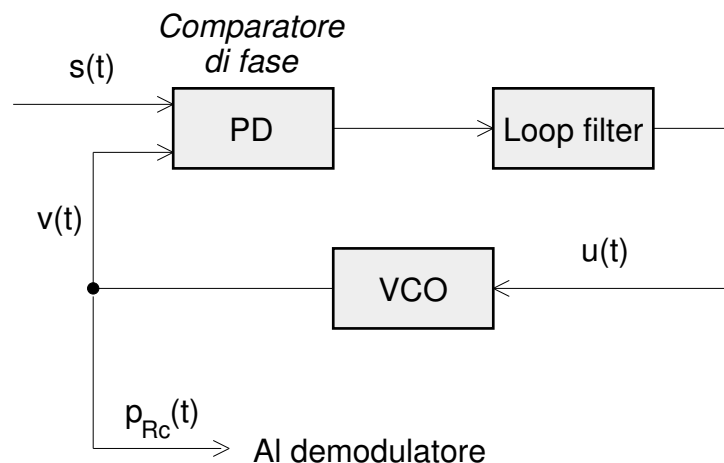


Figura 2.24: *Phase locked loop* per estrarre la portante.

periodica del segnale di ingresso $s(t)$. Il recupero della portante, anche in presenza del rumore introdotto dal canale, è analizzato in [1, cap. 14]. Qui richiamiamo solo la relazione ingresso-uscita

$$v(t) = A \sin \left(2\pi f_0 t + 2\pi \mathcal{K} \int_{-\infty}^t u(\tau) d\tau \right) . \quad (2.48)$$

A regime l'errore di fase tra il segnale d'ingresso e $v(t)$ sarà nullo e $u(t) = 0$. In tal caso $v(t)$ riprodurrà la portante a meno di uno sfasamento di $\frac{\pi}{2}$.

In assenza di trasmissione della portante, in ricezione vengono utilizzate trasformazioni non lineari sul segnale per creare delle righe spettrali che successivamente vengono estratte tramite un PLL.

2.6 Modulazioni di ampiezza lineari: schema generale

Oltre alla DSB esistono altre due tecniche di modulazione di ampiezza lineari che vengono spesso utilizzate.

Lo schema generale di un modulatore di ampiezza lineare, cioè senza trasmissione della portante, è illustrato in Figura 2.25 e consiste di un *mixer*, che effettua il prodotto tra il segnale modulante $a(t)$ e la portante $\cos(2\pi f_0 t + \varphi_0)$, seguito da un opportuno filtro del tipo banda passante h_{PB} . Prima di essere inviato sul canale il segnale viene amplificato ad un opportuno livello.

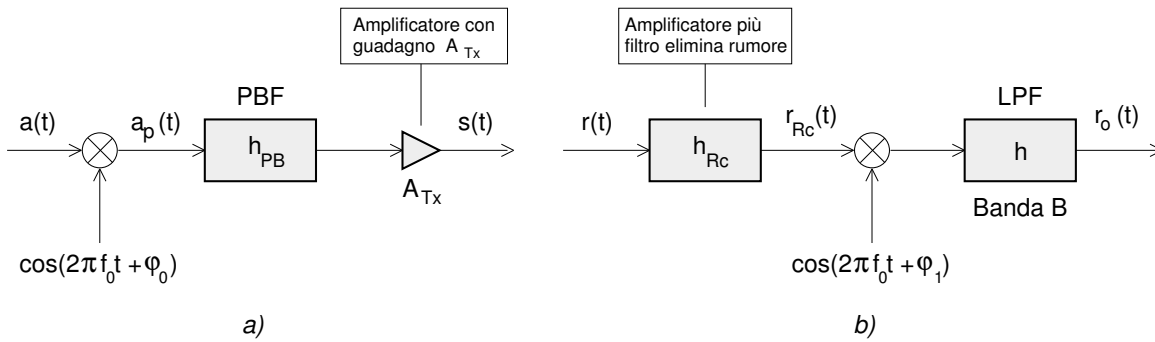


Figura 2.25: Schema base di un modulatore di ampiezza lineare: a) trasmettitore, b) ricevitore.

Definito il segnale

$$a_p(t) = a(t) \cos(2\pi f_0 t + \varphi_0) \quad , \quad (2.49)$$

l'espressione del segnale modulato è data da

$$s(t) = A_{Tx} (a_p * h_{PB})(t) \quad . \quad (2.50)$$

Nel dominio della frequenza queste relazioni diventano

$$\mathcal{A}_p(t) = \frac{e^{j\varphi_0}}{2} \mathcal{A}(f - f_0) + \frac{e^{-j\varphi_0}}{2} \mathcal{A}(f + f_0) \quad , \quad (2.51)$$

e

$$\mathcal{S}(t) = A_{Tx} \mathcal{A}_p(f) \mathcal{H}_{PB}(f) \quad . \quad (2.52)$$

In ricezione, il segnale $r(t)$ all'uscita del canale viene amplificato e filtrato mediante un filtro h_{Rc} avente una banda passante uguale a quella del segnale desiderato $s(t)$; questa specifica porta ad eliminare il rumore introdotto dal canale al di fuori della banda desiderata.

Il ricevitore, utilizzando un *mixer*, prende il segnale filtrato $r_{Rc}(t)$ e lo moltiplica per una portante che deve avere la stessa frequenza e fase della portante all'uscita del canale.

Inizialmente ci poniamo nelle condizioni di canale ideale e assenza di rumore per cui $r(t) = s(t)$. Inoltre, con riferimento allo schema di Figura 2.25 considereremo solo le trasformazioni introdotte dalla modulazione e demodulazione, come riportato in Figura 2.26.

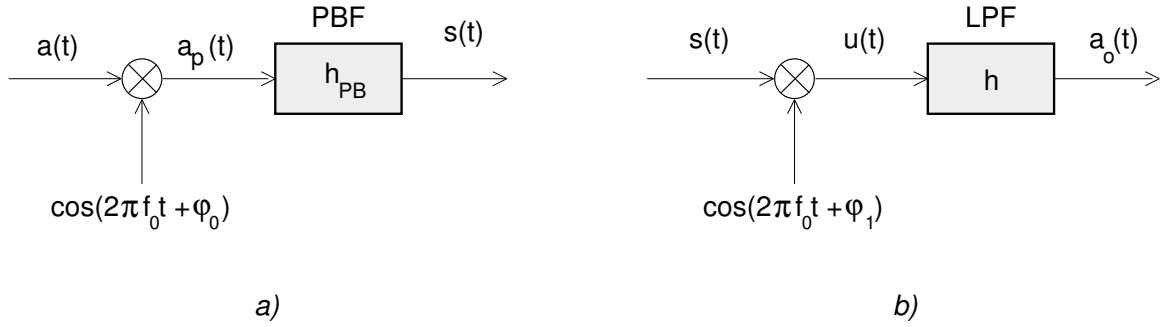


Figura 2.26: Schema base di un modulatore di ampiezza lineare: a) modulatore, b) demodulatore.

Notiamo che in ricezione il filtro h_{Rc} è stato omesso poichè irrilevante in assenza del rumore. È stato omesso anche l'amplificatore di trasmissione dal momento che il suo effetto si manifesta semplicemente tramite un fattore di scala che però non modifica la forma del segnale trasmesso $s(t)$.

La scelta del filtro h_{PB} determina il tipo di modulazione di ampiezza. Vediamo tre casi tipici.

Il primo schema è la modulazione DSB-SC trattata in Sezione 2.1. Essa è ottenuta omettendo il filtro h_{PB} in Figura 2.26, come si può osservare dallo schema di Figura 2.1.

Single side band (SSB)

Modulatore con filtro in banda passante

La tecnica di modulazione a banda laterale unica prevede di trasmettere metà spettro di $a(t)$, ad esempio quello a frequenze positive

$$\mathcal{A}^{(+)}(f) = \mathcal{A}(f) 1(f) = \begin{cases} \mathcal{A}(f) & , f > 0 \\ 0 & , f < 0 \end{cases} \quad (2.53)$$

o quello a frequenze negative

$$\mathcal{A}^{(-)}(f) = \mathcal{A}(f) 1(-f) = \begin{cases} 0 & , f > 0 \\ \mathcal{A}(f) & , f < 0 \end{cases} \quad (2.54)$$

Poichè $a(t)$ è reale risulta $\mathcal{A}^{(-)}(f) = [\mathcal{A}^{(+)}(-f)]^*$, per cui dato $\mathcal{A}^{(+)}(f)$ oppure $\mathcal{A}^{(-)}(f)$ è immediato risalire al segnale complessivo dato da

$$\mathcal{A}(f) = \mathcal{A}^{(+)}(f) + \mathcal{A}^{(-)}(f) \quad (2.55)$$

Illustriamo in dettaglio il caso di trasmissione della banda laterale superiore (*upper side band*), o SSB_+ . In questo caso la caratteristica del filtro h_{PB} deve essere tale da rimuovere le componenti di $a_p(t)$ nell'intervallo di frequenze $(f_0 - B, f_0)$. Le specifiche ideali di h_{PB} in frequenza sono le seguenti

$$\mathcal{H}_{PB}(f) = \begin{cases} 1 & , |f| \in (f_0, f_0 + B) \\ 0 & , |f| \in (f_0 - B, f_0) \\ \text{irrilevante} & , \text{altrove} \end{cases} \quad (2.56)$$

In pratica, sfruttando il fatto che $a(t)$ non ha componenti vicino la continua, le specifiche di $\mathcal{H}_{PB}(f)$ possono avere una banda di transizione tra $(-f_{a,1}+f_0)$ e $(f_{a,1}+f_0)$, dove ricordiamo $f_{a,1}$ è l'estremo inferiore della banda passante di $a(t)$. L'introduzione di questa banda di transizione rende il filtro realizzabile.

In trasmissione viene inviato il segnale $s(t)$ legato ad $a(t)$ dalla relazione in frequenza

$$\mathcal{S}(f) = \mathcal{A}_p(f) \mathcal{H}_{PB}(f) \quad , \quad (2.57)$$

dove $\mathcal{A}_p(f)$ è dato dalla (2.51). Utilizzando la (2.56) in (2.57) e le (2.53) e (2.54) risulta

$$\begin{aligned} \mathcal{S}(f) &= \frac{e^{j\varphi_0}}{2} \mathcal{A}(f-f_0) 1(f-f_0) + \frac{e^{-j\varphi_0}}{2} \mathcal{A}(f+f_0) 1(-f-f_0) \\ &= \frac{e^{j\varphi_0}}{2} \mathcal{A}^{(+)}(f-f_0) + \frac{e^{-j\varphi_0}}{2} \mathcal{A}^{(-)}(f+f_0) \quad . \end{aligned} \quad (2.58)$$

Le varie trasformazioni sono illustrate in Figura 2.27 .

Ora la banda di $s(t)$ è B , pari a quella di $a(t)$. Notiamo che la modulazione SSB richiede metà banda rispetto alla DSB. Naturalmente ciò è possibile al costo di una maggiore complessità realizzativa. Infatti è stato introdotto il filtro di trasmissione h_{PB} . Inoltre, per non introdurre distorsione, il recupero della frequenza e fase della portante tramite il PLL deve essere alquanto accurato. Infatti vedremo che un offset di fase tra la portante in trasmissione e quella in ricezione non si ripercuote solamente in una attenuazione del segnale di informazione ma anche in una componente di distorsione.

Modulatore con filtro in banda base

Un metodo alternativo per realizzare il modulatore SSB è tramite il filtro di Hilbert $(-\frac{\pi}{2})$ avente una risposta in frequenza data da

$$\mathcal{H}^{(h)}(f) = -j \operatorname{sgn}(f) = e^{-j\frac{\pi}{2} \operatorname{sgn}(f)} = \begin{cases} -j & , f > 0 \\ j & , f < 0 \end{cases} \quad . \quad (2.59)$$

In altre parole il filtro di Hilbert sfasa le armoniche a frequenze positive del segnale di ingresso di $-\frac{\pi}{2}$ e quelle a frequenze negative di $\frac{\pi}{2}$. Ad esempio, per un segnale di ingresso del tipo $\cos(2\pi f_0 t + \varphi_0)$ l'uscita del filtro di Hilbert è data da $\sin(2\pi f_0 t + \varphi_0)$.

Riportiamo ora una proprietà che verrà utilizzata in seguito: un segnale $a(t)$ e il corrispondente segnale trasformato secondo il filtro di Hilbert $a^{(h)}(t)$ sono ortogonali,

$$\langle a , a^{(h)} \rangle = 0 \quad . \quad (2.60)$$

Infatti poichè nel dominio della frequenza

$$\mathcal{A}^{(h)}(f) = -j \operatorname{sgn}(f) \mathcal{A}(f) \quad , \quad (2.61)$$

utilizzando il teorema di Parseval risulta

$$\langle a , a^{(h)} \rangle = \int_{-\infty}^{+\infty} a(t) a^{(h)}(t) dt = \int_{-\infty}^{+\infty} \mathcal{A}(f) (-j \operatorname{sgn}(f) \mathcal{A}(f))^* df \quad . \quad (2.62)$$

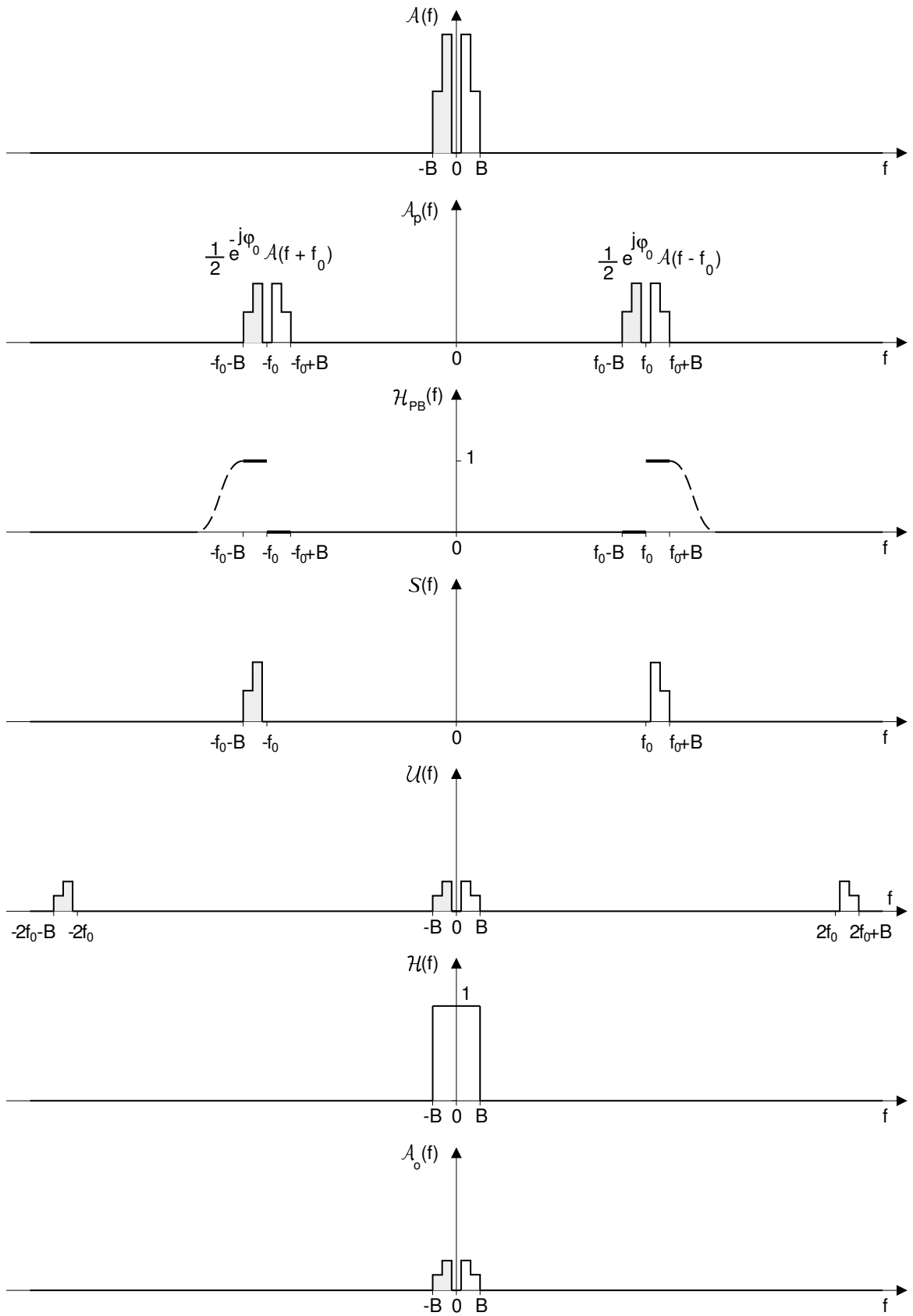


Figura 2.27: Illustrazione dei vari segnali nel dominio della frequenza in una modulazione SSB₊.

Allora

$$\langle a, a^{(h)} \rangle = j \int_{-\infty}^{+\infty} |\mathcal{A}(f)|^2 \operatorname{sgn}(f) df = 0 \quad , \quad (2.63)$$

dal momento che $|\mathcal{A}(f)|^2$ è una funzione pari mentre $\operatorname{sgn}(f)$ è dispari. Un'altra relazione semplice da provare è che

$$M_{a^{(h)}} = M_a \quad . \quad (2.64)$$

Utilizzando la relazione

$$1(f) = \frac{1 + \operatorname{sgn}(f)}{2} \quad , \quad (2.65)$$

in (2.58), risulta

$$\begin{aligned} \mathcal{S}(f) &= \frac{e^{j\varphi_0}}{2} \mathcal{A}(f - f_0) \frac{1 + \operatorname{sgn}(f - f_0)}{2} \\ &+ \frac{e^{-j\varphi_0}}{2} \mathcal{A}(f + f_0) \frac{1 + \operatorname{sgn}(-f - f_0)}{2} \\ &= \frac{1}{2} \left\{ \frac{e^{j\varphi_0} \mathcal{A}(f - f_0) + e^{-j\varphi_0} \mathcal{A}(f + f_0)}{2} \right. \\ &\quad \left. - \frac{e^{j\varphi_0} \mathcal{A}(f - f_0) (-j \operatorname{sgn}(f - f_0)) - e^{-j\varphi_0} \mathcal{A}(f + f_0) (-j \operatorname{sgn}(f + f_0))}{2j} \right\} , \end{aligned} \quad (2.66)$$

in cui abbiamo utilizzato la relazione $\operatorname{sgn}(-f - f_0) = -\operatorname{sgn}(f + f_0)$, essendo sgn una funzione dispari. Ebbene, è facile constatare che la (2.67), nel dominio del tempo, corrisponde alla relazione

$$s(t) = \frac{a(t)}{2} \cos(2\pi f_0 t + \varphi_0) - \frac{a^{(h)}(t)}{2} \sin(2\pi f_0 t + \varphi_0) \quad , \quad (2.67)$$

Lo schema del modulatore basato sulla precedente equazione è riportato in Figura 2.28.

Ricordando che in pratica si riesce realizzare un filtro a meno di un ritardo t_D , di conseguenza in (2.67) avremo $a^{(h)}(t - t_D)$ al posto di $a^{(h)}(t)$. Allora, il segnale $s(t)$ che viene moltiplicato per la portante, la portante stessa e così quella in quadratura devono essere ritardati di t_D .

È facile verificare che per realizzare un modulatore SSB a banda laterale inferiore, in cui vengono trasmesse le componenti a frequenze negative di $a(t)$, il filtro h_{PB} sarà del tipo

$$\mathcal{H}_{PB}(f) = \begin{cases} 1 & , \quad |f| \in (f_0 - B, f_0) \\ 0 & , \quad |f| \in (f_0, f_0 + B) \\ \text{irrelevante} & , \quad \text{altrove} \end{cases} \quad . \quad (2.68)$$

Utilizzando il filtro di Hilbert la (2.67) viene sostituita dalla relazione

$$s(t) = \frac{a(t)}{2} \cos(2\pi f_0 t + \varphi_0) + \frac{a^{(h)}(t)}{2} \sin(2\pi f_0 t + \varphi_0) \quad . \quad (2.69)$$

Nello schema di Figura 2.28 è sufficiente non invertire di segno il segnale che proviene dal ramo inferiore al sommatore.

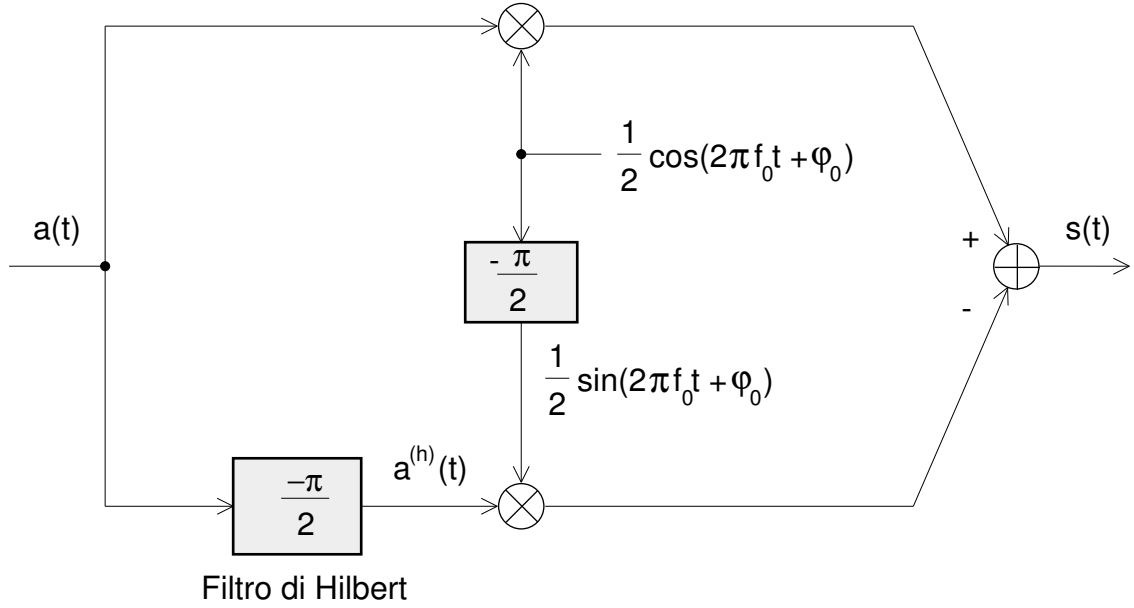


Figura 2.28: Realizzazione di un modulatore SSB a banda laterale superiore utilizzando il filtro di Hilbert.

Demodulatore

Con riferimento allo schema di Figura 2.26b, il segnale all'uscita del *mixer* nel dominio della frequenza è dato da

$$U(f) = \frac{e^{j\varphi_1}}{2} \mathcal{S}(f - f_0) + \frac{e^{-j\varphi_1}}{2} \mathcal{S}(f + f_0) \quad , \quad (2.70)$$

per cui utilizzando la (2.58) risulta

$$\begin{aligned} U(f) &= \frac{e^{j\varphi_0}}{2} \frac{e^{j\varphi_1}}{2} \mathcal{A}^{(+)}(f - 2f_0) \\ &+ \frac{e^{-j\varphi_0}}{2} \frac{e^{j\varphi_1}}{2} \mathcal{A}^{(-)}(f) \\ &+ \frac{e^{j\varphi_0}}{2} \frac{e^{-j\varphi_1}}{2} \mathcal{A}^{(+)}(f) \\ &+ \frac{e^{-j\varphi_0}}{2} \frac{e^{-j\varphi_1}}{2} \mathcal{A}^{(-)}(f + 2f_0) \\ &= \frac{e^{j(\varphi_0 - \varphi_1)}}{4} \mathcal{A}^{(+)}(f) + \frac{e^{-j(\varphi_0 - \varphi_1)}}{4} \mathcal{A}^{(-)}(f) \\ &+ \frac{e^{j(\varphi_0 + \varphi_1)}}{4} \mathcal{A}^{(+)}(f - 2f_0) + \frac{e^{-j(\varphi_0 + \varphi_1)}}{4} \mathcal{A}^{(-)}(f - 2f_0) \quad . \quad (2.71) \end{aligned}$$

Mentre il terzo e quarto termine dell'equazione precedente vengono eliminati dal filtro passa basso h , per $\varphi_0 = \varphi_1$, il primo e secondo termine porgono, rispettivamente, le componenti a frequenze positive e negative del segnale desiderato.

Di conseguenza per $\varphi_0 = \varphi_1$ risulta

$$a_o(t) = \frac{1}{4} a(t) \quad . \quad (2.72)$$

In effetti se $\varphi_1 \neq \varphi_0$, $a_o(t)$ conterrà, oltre ad una componente proporzionale al segnale desiderato $a(t)$, anche una componente di distorsione. Vediamo di ricavarla dalla (2.67). Dall'espressione

$$\begin{aligned}
u(t) &= s(t) \cos(2\pi f_0 t + \varphi_1) \\
&= \frac{a(t)}{2} \cos(2\pi f_0 t + \varphi_0) \cos(2\pi f_0 t + \varphi_1) \\
&\quad - \frac{a^{(h)}(t)}{2} \cos(2\pi f_0 t + \varphi_0) \cos(2\pi f_0 t + \varphi_1) \\
&= \frac{a(t)}{2} \cos(\varphi_0 - \varphi_1) + \frac{a(t)}{2} \cos(2\pi 2f_0 t + \varphi_0 + \varphi_1) \\
&\quad - \frac{a^{(h)}(t)}{2} \sin(\varphi_0 - \varphi_1) - \frac{a^{(h)}(t)}{2} \sin(2\pi 2f_0 t + \varphi_0 + \varphi_1) \quad , \quad (2.73)
\end{aligned}$$

risulta che il secondo e quarto termine hanno componenti spettrali attorno a $2f_0$, le quali vengono eliminate dal filtro passa basso, per cui all'uscita avremo

$$a_o(t) = \frac{a(t)}{2} \cos(\varphi_0 - \varphi_1) - \frac{a^{(h)}(t)}{2} \sin(\varphi_0 - \varphi_1) \quad . \quad (2.74)$$

Ora ricordando la proprietà (2.63), il segnale $a^{(h)}(t)$ non ha termini proporzionali ad $a(t)$ essendo ortogonale. Allora in (2.74) il termine $\frac{a^{(h)}(t)}{2} \sin(\varphi_0 - \varphi_1)$ si manifesta come un termine interferente nei confronti del termine utile $a(t) \cos(\varphi_0 - \varphi_1)$. È allora molto importante che in un sistema SSB il PLL sia sufficientemente accurato in modo che $\varphi_1 \simeq \varphi_0$ e rendere così trascurabile la distorsione.

Vestigial side band (VSB)

Modulatore

In talune applicazioni e tipicamente in presenza di segnali video, il filtro h_{PB} per modulazioni SSB risulta molto selettivo, nel senso che il rapporto tra la larghezza della banda di transizione e la larghezza della banda passante è molto piccola. In tali casi soddisfare le caratteristiche di ampiezza del filtro comporta necessariamente una notevole distorsione di fase. Si preferisce allora rilassare le specifiche del filtro h_{PB} e "richiedere" una banda maggiore al canale di trasmissione.

Come vedremo in seguito il filtro h_{PB} deve avere un andamento attorno alla portante f_0 tale che

$$\mathcal{H}_{PB}(f + f_0) + \mathcal{H}_{PB}(f - f_0) = K \quad , \quad \text{per } |f| \leq B \quad , \quad (2.75)$$

con K costante reale.

Consideriamo il caso di trasmettere tutte le armoniche a frequenze positive di $a(t)$ e solo parte delle armoniche a frequenze negative: ciò dà luogo alla modulazione VSB₊.

Nelle illustrazioni che seguono assumeremo che il filtro h_{PB} abbia una caratteristica per frequenze positive del tipo

$$\mathcal{H}_{PB}^{(+)}(f) = \begin{cases} 1 & , \quad f_0 + \rho B < f < f_0 + B \\ \frac{1}{2} \left[1 + \sin \left(\frac{\pi}{2} \frac{f-f_0}{\rho B} \right) \right] & , \quad f_0 - \rho B < f < f_0 + \rho B \\ 0 & , \quad f_0 - B < f < f_0 - \rho B \\ \text{irrelevante} & , \quad \text{altrove} \end{cases} \quad , \quad (2.76)$$

illustrata in Figura 2.29. Naturalmente essendo h_{PB} a valori reali, $\mathcal{H}_{PB}^{(-)}(f)$ si ottiene per simmetria, $\mathcal{H}_{PB}^{(-)}(f) = (\mathcal{H}_{PB}^{(+)}(-f))^*$. È facile verificare che tale caratteristica soddisfa la (2.75) con $K = 1$. In (2.76) ρ è un parametro tra 0 e 1 e determina la banda richiesta pari a $B(1 + \rho)$. Per $\rho = 0$ otteniamo la modulazione SSB₊.

Un approccio pratico alla progettazione di h_{PB} è partire da un filtro del tipo banda passante con banda di transizione che va da $(f_0 - \rho B)$ a $(f_0 + \rho B)$, banda passante con guadagno unitario che si estende alla destra di $(f_0 + \rho B)$ e banda attenuata con valore ideale nullo che si estende alla sinistra di $(f_0 - \rho B)$. Ora è sufficiente variare l'ordine del filtro fino a quando l'ampiezza del filtro vale 0.5 per $f = f_0$. Tipicamente se il *ripple* in banda passante e quello in banda attenuata sono scelti uguali, il filtro avrà un andamento che verifica la (2.75) con $K = 1$.

In ogni caso il segnale modulato, nel dominio della frequenza, è dato da

$$\mathcal{S}(f) = \left[\frac{e^{j\varphi_0}}{2} \mathcal{A}(f - f_0) + \frac{e^{-j\varphi_0}}{2} \mathcal{A}(f + f_0) \right] \mathcal{H}_{PB}(f) \quad . \quad (2.77)$$

Demodulatore

Dalla relazione generale (2.70) e (3.58), risulta

$$\begin{aligned} \mathcal{U}(f) &= \frac{e^{j(\varphi_0 - \varphi_1)}}{4} \mathcal{A}(f) \mathcal{H}_{PB}(f + f_0) \\ &+ \frac{e^{-j(\varphi_0 - \varphi_1)}}{4} \mathcal{A}(f) \mathcal{H}_{PB}(f - f_0) \\ &+ \frac{e^{j(\varphi_0 + \varphi_1)}}{4} \mathcal{A}(f - 2f_0) \mathcal{H}_{PB}(f - 2f_0) \\ &+ \frac{e^{-j(\varphi_0 + \varphi_1)}}{4} \mathcal{A}(f + 2f_0) \mathcal{H}_{PB}(f + 2f_0) \quad . \end{aligned} \quad (2.78)$$

Al solito gli ultimi due termini della precedente relazione vengono ottenuti dal filtro passa basso h , alla cui uscita per $\varphi_1 = \varphi_0$ risulta

$$\mathcal{A}_o(f) = \frac{\mathcal{A}(f)}{4} \left[\mathcal{H}_{PB}(f + f_0) + \mathcal{H}_{PB}(f - f_0) \right] \quad . \quad (2.79)$$

Di conseguenza, se il filtro di modulazione h_{PB} soddisfa la relazione (2.75) ed ha un guadagno unitario, avremo che $K = 1$ e

$$a_o(t) = \frac{a(t)}{4} \quad . \quad (2.80)$$

Complessivamente, in una modulazione VSB il filtro h_{PB} essendo meno selettivo rispetto al caso di modulazione SSB risulta più semplice da realizzare. La conseguenza è che il sistema richiede una banda più larga.

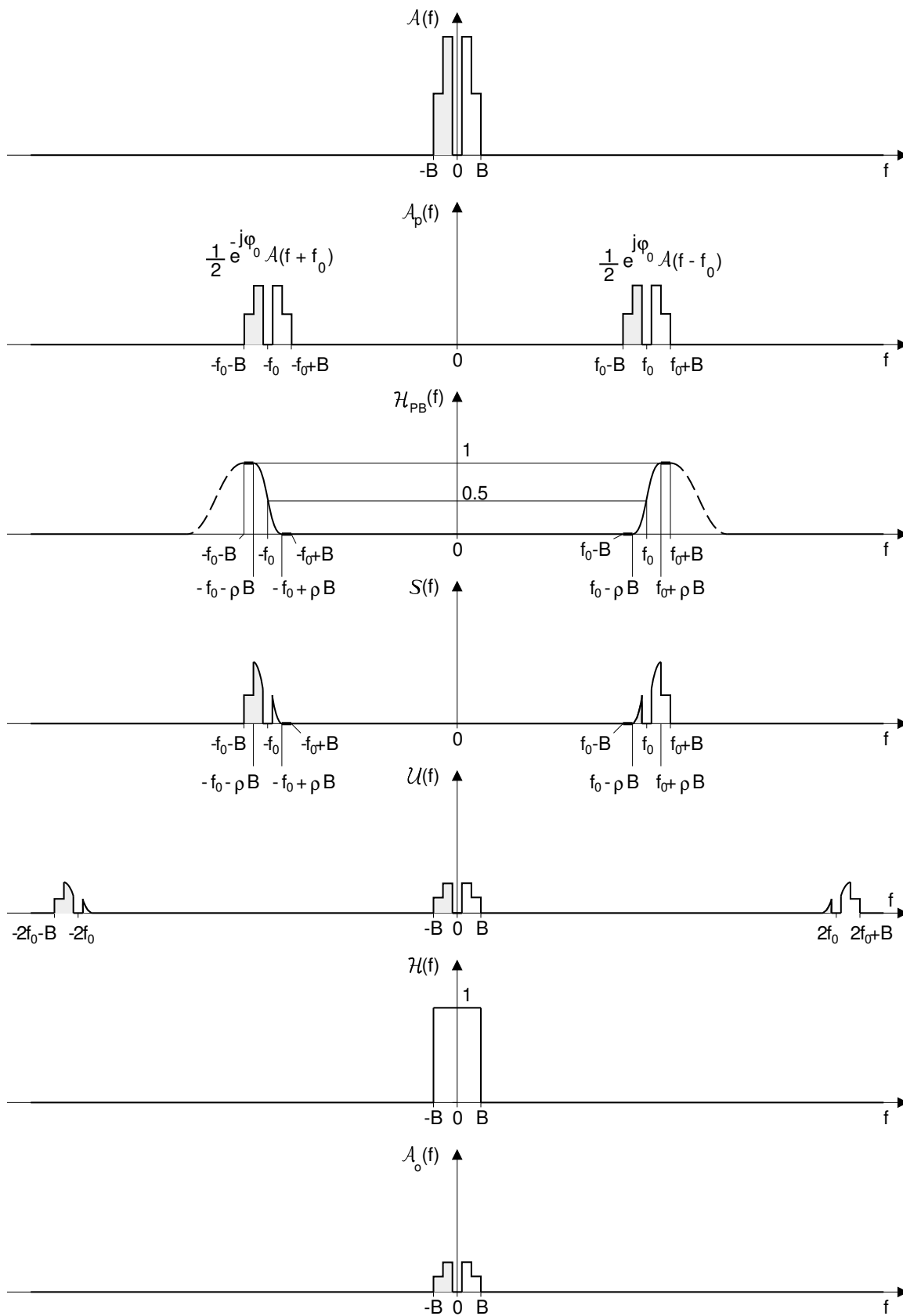


Figura 2.29: Illustrazione dei vari segnali nel dominio della frequenza in una modulazione VSB₊.

2.7 Single side band trasmission carrier (SSB-TC)

È possibile utilizzare la tecnica di inviare la portante anche in una modulazione SSB. Dalla (2.67) per una SSB₊-TC il segnale trasmesso è dato da

$$s(t) = \frac{a(t)}{2} \cos(2\pi f_0 t + \varphi_0) - \frac{a^{(h)}(t)}{2} \sin(2\pi f_0 t + \varphi_0) + A \cos(2\pi f_0 t + \varphi_0) \quad (2.81)$$

Anche in questo caso possiamo introdurre l'efficienza della modulazione che in analogia alla (2.15) è ora data da (vedi anche relazione (2.64))

$$\eta = \frac{\left(\frac{M_a}{4} + \frac{M_{a^{(h)}}}{4}\right) \frac{1}{2}}{M_s} = \frac{M_a}{M_s} \quad (2.82)$$

dove $M_s = \frac{M_a}{4} + \frac{A^2}{2}$.

2.8 Quadrature amplitude modulation (QAM)

Modulatore

Come illustrato in Figura 2.30 questo sistema è analogo alla modulazione di ampiezza in quadratura per segnali dati. Siano $a_1(t)$ e $a_2(t)$ due segnali di informazione a valori reali aventi una banda B . Moduliamo ora $a_1(t)$ per la portante $\cos(2\pi f_0 t + \varphi_0)$ mentre $a_2(t)$ per la portante in quadratura $\sin(2\pi f_0 t + \varphi_0)$, e formiamo il segnale modulato

$$s(t) = a_1(t) \cos(2\pi f_0 t + \varphi_0) - a_2(t) \sin(2\pi f_0 t + \varphi_0) \quad (2.83)$$

Un modo alternativo per ricavare $s(t)$ è introducendo il segnale complesso

$$a(t) = a_1(t) + j a_2(t) \quad (2.84)$$

Modulando $a(t)$ per la portante complessa $e^{j(2\pi f_0 t + \varphi_0)} = \cos(2\pi f_0 t + \varphi_0) + j \sin(2\pi f_0 t + \varphi_0)$ e trasmettendo la sola parte reale si ottiene $s(t)$. Analiticamente

$$s(t) = \text{Re}[a(t) e^{j(2\pi f_0 t + \varphi_0)}] \quad (2.85)$$

Notiamo che la banda di $s(t)$ è pari a $2B$, però il segnale porta due segnali di banda B per cui risulta la stessa efficienza spettrale della SSB. La illustrazione dello spettro dei vari segnali è riportata in Figura 2.31.

Demodulatore

Sia la fase della portante ricostruita uguale a quella ricevuta, cioè $\varphi_1 = \varphi_0$. Formiamo due segnali $u_1(t)$ e $u_2(t)$ moltiplicando $s(t)$ rispettivamente per $\cos(2\pi f_0 t + \varphi_0)$

e $\sin(2\pi f_0 t + \varphi_0)$. Sia $u_1(t)$ che $u_2(t)$ vengono filtrati da un filtro passa basso $h(t)$ di banda B . L'espressione di $u_1(t)$ sarà data da

$$\begin{aligned} u_1(t) &= s(t) \cos(2\pi f_0 t + \varphi_0) \\ &= a_1(t) \cos(2\pi f_0 t + \varphi_0) \cos(2\pi f_0 t + \varphi_0) \\ &\quad - a_2(t) \sin(2\pi f_0 t + \varphi_0) \cos(2\pi f_0 t + \varphi_0) \\ &= \frac{a_1(t)}{2} + \frac{a_1(t)}{2} \cos(2\pi 2f_0 t + 2\varphi_0) - \frac{a_2(t)}{2} \sin(2\pi 2f_0 t + 2\varphi_0) \quad .(2.86) \end{aligned}$$

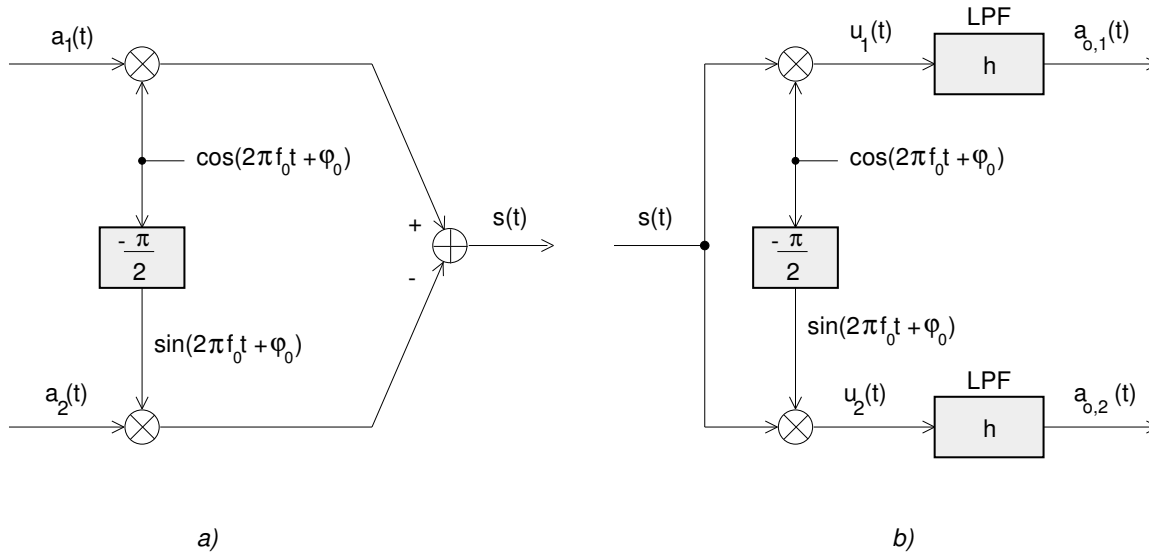


Figura 2.30: Schema base di un QAM: a) modulatore, b) demodulatore.

Ora sia il secondo che il terzo termine vengono attenuati dal filtro h alla cui uscita rimane

$$a_{o,1}(t) = \frac{a_1(t)}{2} \quad . \quad (2.87)$$

In modo analogo, il segnale all'uscita del secondo ramo è dato da

$$a_{o,2}(t) = \frac{a_2(t)}{2} \quad . \quad (2.88)$$

Un metodo alternativo per ricavare $a_{o,1}(t)$ e $a_{o,2}(t)$ è utilizzando la notazione complessa. Sia

$$u(t) = s(t) e^{-j(2\pi f_0 t + \varphi_0)} = u_1(t) + j u_2(t) \quad , \quad (2.89)$$

che in frequenza porge

$$\mathcal{U}(f) = \mathcal{S}(f - f_0) e^{-j\varphi_0} \quad , \quad (2.90)$$

e definendo $a_o(t) = (u * h)(t)$, è facile verificare che risulta

$$a_{o,1}(t) = \text{Re}[a_o(t)] \quad , \quad (2.91)$$

e

$$a_{o,2}(t) = \text{Im}[a_o(t)] \quad . \quad (2.92)$$

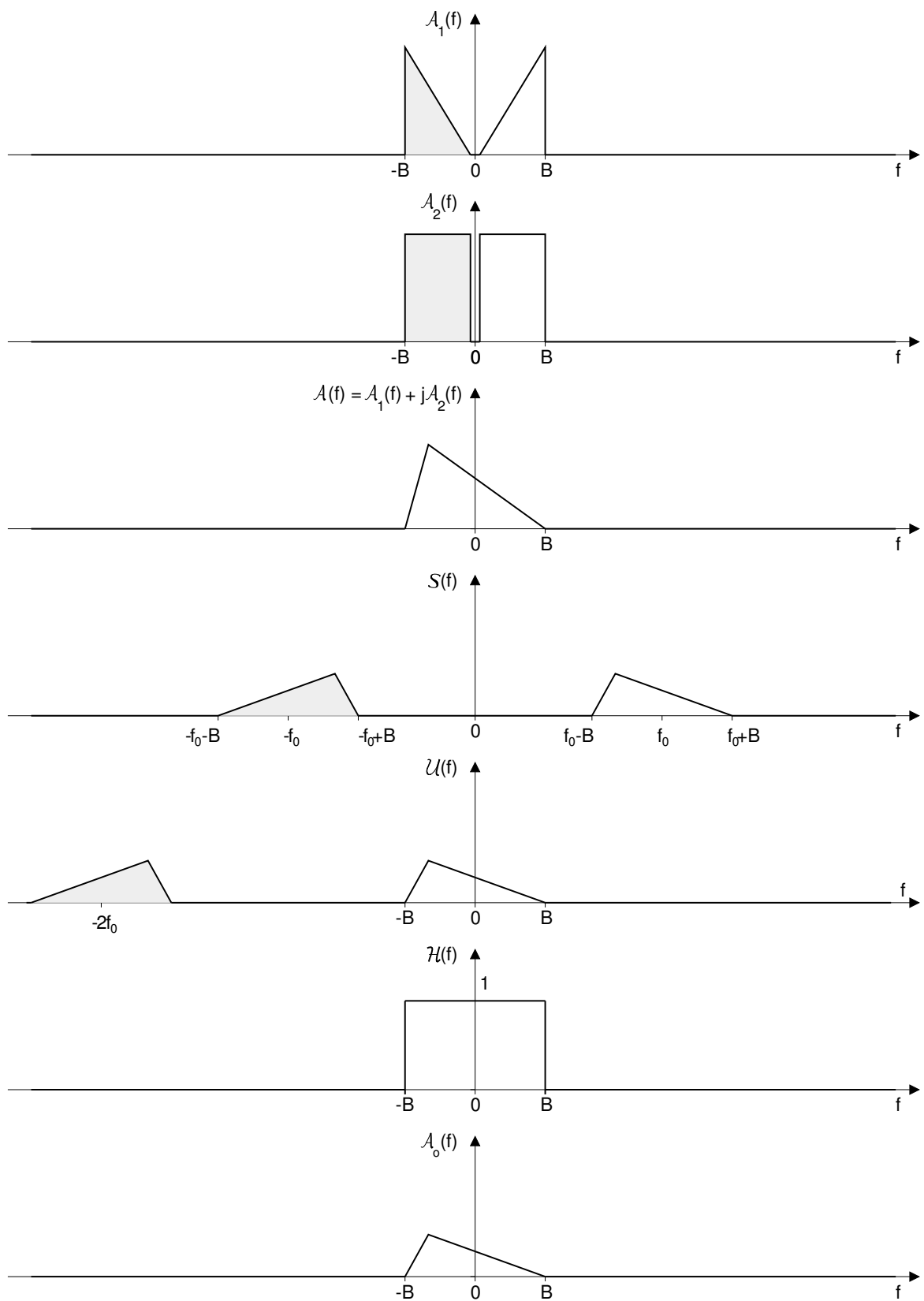


Figura 2.31: Illustrazione dei vari segnali nel dominio della frequenza in una QAM.

Notiamo che in questo schema i due segnali $a_1(t)$ e $a_2(t)$ vengono sovrapposti sia nel tempo che in frequenza. In effetti essi sono ancora separabili poichè la moltiplicazione per le funzioni $\cos(2\pi f_0 t + \varphi_0)$ e $\sin(2\pi f_0 t + \varphi_0)$ li rende ortogonali³, sotto l'ipotesi $f_0 > B$. Il prezzo di questa ortogonalità è aver raddoppiato la banda richiesta da B a $2B$.

2.9 Accesso multiplo a divisione di frequenza (FDM)

Riportiamo in Figura 2.32 il caso di N segnali modulati DSB-SC le cui portanti sono separate di $2B$ Hz. Notiamo che questa è la minima separazione in frequenza, altrimenti c'è interferenza tra due segnali. In pratica, per semplificare la realizzazione del filtro h al ricevitore è opportuno che venga introdotta una banda di guardia tra i segnali. Ciò si ottiene imponendo che la separazione tra le portanti sia maggiore di $2B$. In conclusione, tramite un certo spreco di banda si può ottenere un filtro h con una banda di transizione non nulla e di conseguenza realizzabile.

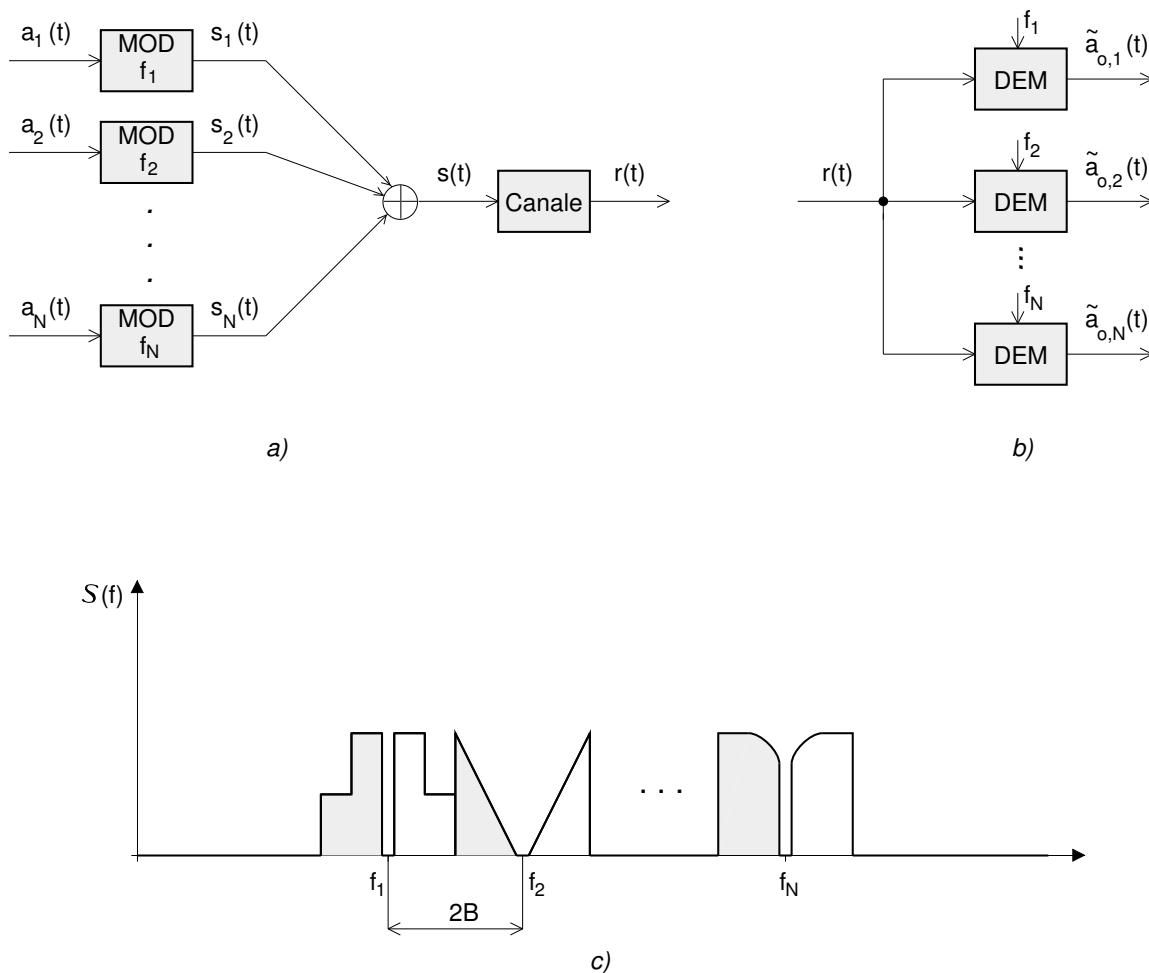


Figura 2.32: Schema base del principio FDM per N segnali: a) modulatore, b) demodulatore e c) illustrazione dello spettro occupato.

³In effetti la TDM e FDM sono casi particolari di ortogonalità tra segnali.

2.10 Prestazioni

Valutiamo ora le prestazioni dei vari schemi di modulazione di ampiezza per un canale rumoroso però non distorcente in cui il segnale all'ingresso del ricevitore è dato dalla (1.1).

Introduciamo innanzitutto il *il rapporto segnale-rumore di riferimento* Γ all'uscita del canale. Sia M_s la potenza del segnale trasmesso e B la banda del segnale modulante $a(t)$. Definiamo Γ come rapporto tra la potenza del segnale desiderato all'uscita del canale e la potenza del rumore valutata su una banda convenzionale pari a B :

$$\Gamma = \frac{M_s |\mathcal{G}_{Ch,0}|^2}{\frac{N_0}{2} 2B} = \frac{M_s}{N_0 B a_d} \quad , \quad (2.93)$$

dove a_d è l'attenuazione di potenza introdotta dal canale (vedi (1.2)). Per un dato segnale $a(t)$ con banda B e prefissato canale con attenuazione a_d , Γ dipende dal rapporto tra la potenza del segnale trasmesso e la densità spettrale del rumore. Sottolineiamo che la banda di $s(t)$ non coincide necessariamente con B , vedi ad esempio le modulazioni DSB e VSB.

D'ora in poi, per semplificare la notazione porremo $\mathcal{C} = \mathcal{G}_{Ch,0}$.

Analizzeremo distintamente i casi di ricevitore coerente e non coerente. In generale indicheremo con $r_o(t)$ il segnale all'uscita del ricevitore. In assenza di distorsione, $r_o(t)$ sarà composto da una componente utile $a_o(t)$ proporzionale al segnale di informazione $a(t)$, $a_o(t) = c a(t)$, e una componente addittiva rumorosa $w_o(t)$. In definitiva possiamo scrivere

$$r_o(t) = a_o(t) + w_o(t) = c a(t) + w_o(t) \quad . \quad (2.94)$$

Le prestazioni del sistema vengono misurate dal rapporto tra la potenza del termine desiderato $c^2 M_a$ e la potenza del rumore in uscita M_{w_o} ,

$$\Lambda_o = \frac{c^2 M_a}{M_{w_o}} \quad . \quad (2.95)$$

Il nostro obiettivo sarà esprimere Λ_o in funzione di Γ per le diverse modulazioni.

Demodulatori coerenti

DSB-SC

Dalle (2.1) e (1.1), il segnale all'ingresso del ricevitore è dato da

$$r(t) = \mathcal{C} a(t) \cos(2\pi f_0 t + \varphi_0) + w(t) \quad . \quad (2.96)$$

Dallo schema di Figura 2.25 all'uscita del filtro h_{Rc} , del tipo banda passante attorno f_0 e banda $2B$,

$$r_{Rc}(t) = \mathcal{C} a(t) \cos(2\pi f_0 t + \varphi_0) + w_{Rc}(t) \quad , \quad (2.97)$$

in cui, utilizzando la notazione [1]

$$w_{Rc}(t) = w_{Rc,I}(t) \cos(2\pi f_0 t + \varphi_1) - w_{Rc,Q}(t) \sin(2\pi f_0 t + \varphi_1) \quad , \quad (2.98)$$

dove $w_{Rc,I}(t)$ e $w_{Rc,Q}(t)$ sono rumori incorrelati, del tipo banda base di banda B , ciascuno con uno spettro

$$\mathcal{P}_{w_{Rc,I}}(f) = \mathcal{P}_{w_{Rc,Q}}(f) = N_0 |\mathcal{H}_{Rc}(f)|^2 = N_0 \operatorname{rect}\left(\frac{f}{2B}\right) . \quad (2.99)$$

Di conseguenza le potenze di $w_{Rc,I}$ e $w_{Rc,Q}$ sono uguali e pari a

$$\mathbf{M}_{w_{Rc,I}} = \mathbf{M}_{w_{Rc,Q}} = N_0 2B . \quad (2.100)$$

All'ingresso del filtro h il rumore è dato da

$$\begin{aligned} w_{Rc}(t) \cos(2\pi f_0 t + \varphi_1) &= \frac{1}{2} \left[w_{Rc,I}(t) + w_{Rc,I}(t) \cos(2\pi 2f_0 t + \varphi_0 + \varphi_1) \right. \\ &\quad \left. - w_{Rc,Q}(t) \sin(2\pi 2f_0 t + \varphi_0 + \varphi_1) \right] , \end{aligned} \quad (2.101)$$

di cui solo il primo termine passa attraverso il filtro e di conseguenza

$$w_o(t) = \frac{1}{2} w_{Rc,I}(t) . \quad (2.102)$$

In base alla (2.100) la potenza di w_o è pari a

$$\mathbf{M}_{w_o} = \frac{1}{4} N_0 2B = \frac{N_0 B}{2} . \quad (2.103)$$

Utilizzando infine l'espressione (2.5) del segnale utile all'uscita, la quale a causa dell'attenuazione \mathcal{C} del canale porge in (2.94)

$$c = \frac{\mathcal{C}}{2} \cos(\varphi_0 - \varphi_1) . \quad (2.104)$$

In definitiva per $\varphi_0 = \varphi_1$ il rapporto (2.95) è dato da

$$\Lambda_o = \frac{\frac{\mathcal{C}^2}{4} \mathbf{M}_a}{\frac{N_0 B}{2}} = \frac{\mathcal{C}^2 \mathbf{M}_a}{N_0 2B} . \quad (2.105)$$

Con riferimento al rapporto di riferimento Γ , dato dalla (2.93), poichè dalla (2.1) $\mathbf{M}_s = \frac{\mathbf{M}_a}{2}$, avremo che

$$\Lambda_o = \Gamma . \quad (2.106)$$

È utile introdurre anche il rapporto segnale-rumore Λ_i all'uscita del filtro h_{Rc} dove la potenza del rumore viene misurata con riferimento alla banda B_s del segnale $s(t)$

$$\Lambda_i = \frac{\mathcal{C}^2 \mathbf{M}_s}{\frac{N_0}{2} 2B_s} = \frac{\mathcal{C}^2 \mathbf{M}_s}{N_0 B_s} . \quad (2.107)$$

Poichè $B_s = 2B$ risulta

$$\Lambda_o = 2 \Lambda_i . \quad (2.108)$$

Osserviamo che il processo di demodulazione ha raddoppiato il rapporto segnale-rumore. Infatti le componenti del segnale desiderato a frequenza positiva e negativa vengono a sommarsi coerentemente per dar luogo ad un raddoppio del segnale e quindi

quadruplicando la potenza del segnale. Invece le componenti del rumore a frequenza positiva sono incorrelate con quelle a frequenza negativa per cui sommandosi la potenza aumenta solo di un fattore 2: complessivamente c'è un guadagno di un fattore 2 nel rapporto segnale-rumore.

SSB

Analizziamo il caso SSB_+ . L'analisi è simile per una modulazione SSB_- .

Dalla (1.1) e (2.67) risulta

$$r(t) = \frac{\mathcal{C}}{2} \left[a(t) \cos(2\pi f_0 t + \varphi_0) - a^{(h)}(t) \sin(2\pi f_0 t + \varphi_0) \right] + w(t) \quad . \quad (2.109)$$

Il filtro h_{Rc} , in base alla (2.56), avrà l'andamento ideale

$$\mathcal{H}_{Rc}(f) = \text{rect} \left(\frac{f - (f_0 + \frac{B}{2})}{B} \right) + \text{rect} \left(\frac{-f - (f_0 + \frac{B}{2})}{B} \right) \quad . \quad (2.110)$$

All'uscita di h_{Rc} , il segnale sarà dato da

$$r_{Rc}(t) = \frac{\mathcal{C}}{2} \left[a(t) \cos(2\pi f_0 t + \varphi_0) - a^{(h)}(t) \sin(2\pi f_0 t + \varphi_0) \right] + w_{Rc}(t) \quad . \quad (2.111)$$

in cui $w_{Rc}(t)$ ha uno spettro

$$\mathcal{P}_{w_{Rc}}(f) = \frac{N_0}{2} |\mathcal{H}_{Rc}(f)|^2 \quad , \quad (2.112)$$

e potenza

$$M_{w_{Rc}} = \frac{N_0 2B}{2} = N_0 B \quad . \quad (2.113)$$

Procedendo in modo analogo al caso DSB risulta

$$M_{w_o} = \frac{1}{4} M_{w_{Rc}} = \frac{N_0 B}{4} \quad , \quad (2.114)$$

mentre dalla (2.72), per un canale con ampiezza \mathcal{C} ,

$$M_{a_o} = \frac{\mathcal{C}^2}{16} M_a \quad . \quad (2.115)$$

In conclusione

$$\Lambda_o = \frac{\frac{\mathcal{C}^2}{16} M_a}{\frac{N_0 B}{4}} = \frac{\frac{\mathcal{C}^2}{4} M_a}{N_0 B} \quad . \quad (2.116)$$

Utilizzando la relazione (2.67) e il fatto che $a(t)$ e $a^{(h)}(t)$ sono segnali ortogonali, abbiamo che

$$M_s = \frac{1}{8} M_a + \frac{1}{8} M_{a^{(h)}} = \frac{M_a}{4} \quad , \quad (2.117)$$

utilizzando la (2.64). Allora

$$\Lambda_o = \Gamma \quad . \quad (2.118)$$

Inoltre poichè $B_s = B$, è facile verificare che

$$\Lambda_o = \Lambda_i \quad . \quad (2.119)$$

Il risultato importante è che le modulazioni DSB-SC e SSB hanno le stesse prestazioni in termini di rapporto segnale-rumore all'uscita del ricevitore, con $\Lambda_o = \Gamma$. La SSB però richiede metà banda, a costo di una maggiore complessità realizzativa.

DSB-TC e SSB-TC

È facile verificare che per entrambi i sistemi DSB e SSB, quando oltre al segnale di informazione viene trasmessa anche la portante, nel caso di demodulazione coerente l'analisi del sistema segue le stesse linee esposte sopra. In particolare il rapporto segnale-rumore all'uscita del ricevitore è dato dalle (2.106) e (2.118), rispettivamente per DSB e SSB. Considerando ora che solo una frazione η (definita dalla (2.15)) della potenza di $s(t)$ è utilizzata per la trasmissione del segnale di informazione, in entrambi i sistemi di modulazione avremo che

$$\Lambda_o = \Gamma \eta \quad , \quad (2.120)$$

con ovviamente una perdita di prestazioni, a parità di potenza trasmessa, rispetto al caso SC.

Demodulatori non coerenti

DSB-TC (AM)

Consideriamo la condizione per l'assenza di distorsione dell'involuppo, $A > a_m$. All'uscita del filtro h_{Rc} , di tipo banda passante di banda $2B$, avremo il segnale

$$\begin{aligned} r_{Rc}(t) &= C A (1 + m \bar{a}(t)) \cos(2\pi f_0 t + \varphi_0) \\ &+ w_{Rc,I}(t) \cos(2\pi f_0 t + \varphi_0) - w_{Rc,Q}(t) \sin(2\pi f_0 t + \varphi_0) \\ &= \left[C A (1 + m \bar{a}(t)) + w_{Rc,I}(t) \right] \cos(2\pi f_0 t + \varphi_0) \\ &- w_{Rc,Q}(t) \sin(2\pi f_0 t + \varphi_0) \quad , \end{aligned} \quad (2.121)$$

in cui $w_{Rc,I}(t)$ e $w_{Rc,Q}(t)$ definiti in 2.98.

Introdotte le funzioni modulo e deviazione di fase di $r_{Rc}(t)$,

$$M_{r_{Rc}}(t) = \left[\left(C A (1 + m \bar{a}(t)) + w_{Rc,I}(t) \right)^2 + w_{Rc,Q}^2(t) \right]^{\frac{1}{2}} \quad (2.122)$$

$$\Delta\varphi_{r_{Rc}}(t) = -\tan^{-1} \left(\frac{w_{Rc,Q}(t)}{C A (1 + m \bar{a}(t)) + w_{Rc,I}(t)} \right) \quad , \quad (2.123)$$

possiamo scrivere $r_{Rc}(t)$ nel seguente modo

$$r_{Rc}(t) = M_{r_{Rc}}(t) \cos(2\pi f_0 t + \varphi_0 + \Delta\varphi_{r_{Rc}}(t)) \quad . \quad (2.124)$$

L'uscita di un demodulatore di inviluppo con ingresso $r_{Rc}(t)$ porgerà proprio $M_{r_{Rc}}(t)$, utilizzando uno degli schemi illustrati in Sezione 2.4.

Ora, per separare in (2.122) la parte desiderata da quella rumorosa bisogna introdurre qualche approssimazione. Posto

$$s_{Rc}(t) = \mathcal{C} A (1 + m \bar{a}(t)) \quad , \quad (2.125)$$

assumeremo che

$$w_{Rc,Q}(t) \ll s_{Rc}(t) + w_{Rc,I}(t) \quad . \quad (2.126)$$

Tale condizione è verificata con elevata probabilità se il rapporto segnale-rumore Γ all'uscita del canale è sufficientemente elevato cioè se il segnale $s_R(t) = \mathcal{C} s(t) = s_{Rc}(t)$ all'ingresso del ricevitore ha una potenza sufficientemente più elevata della potenza del rumore $w_{Rc}(t)$ all'uscita di h_{Rc} . In base alla (2.126), la (2.122) diviene

$$M_{r_{Rc}}(t) = \left| \mathcal{C} A (1 + m \bar{a}(t)) + w_{Rc,I}(t) \right| \quad , \quad (2.127)$$

che, sempre in base alla (2.126), può essere scritta come

$$M_{r_{Rc}}(t) = \mathcal{C} A (1 + m \bar{a}(t)) + w_{Rc,I}(t) \quad . \quad (2.128)$$

Filtrando la componente DC della (2.128), rimane il segnale

$$\begin{aligned} a_o(t) &= \mathcal{C} A m \bar{a}(t) + w_{Rc,I}(t) \\ &= \mathcal{C} a(t) + w_{Rc,I}(t) \quad , \end{aligned} \quad (2.129)$$

dove abbiamo utilizzato le (2.10) e (2.11). Allora il rapporto segnale-rumore all'uscita del ricevitore è dato da

$$\Lambda_o = \frac{\mathcal{C} M_a}{N_0 2B} \quad . \quad (2.130)$$

Utilizzando le (2.15) e (2.93), risulta

$$\Lambda_o = \Gamma \eta = \Gamma \frac{m^2 k_f^2}{1 + m^2 k_f^2} \quad , \quad (2.131)$$

dove nell'ultimo passaggio abbiamo utilizzato la (2.19). Ricordando che per l'assenza di distorsione dell'inviluppo deve essere $m < 1$, allora risulta $\Lambda_o < \frac{\Gamma}{2}$; in altre parole, a parità di Γ , un demodulatore (non coerente) per DDS-TC ha almeno 3 dB di perdita in Λ_o rispetto ad un demodulatore (coerente) per DSB-SC.

Concludiamo ricordando che con rapporti Γ non sufficientemente elevati, la (2.131) non vale più e Λ_o dipende da Γ con legge quadratica [3].

SSB-TC

In linea di principio si potrebbe utilizzare un demodulatore non coerente anche per una modulazione SSB. In assenza del rumore, il segnale ricevuto è dato dalla (2.67) moltiplicato per \mathcal{C} , dovuto al canale, più il segnale dovuto alla portante $\mathcal{C} A \cos(2\pi f_0 t + \varphi_0)$:

$$\mathcal{C} \left[\left(\frac{a(t)}{2} + A \right) \cos(2\pi f_0 t + \varphi_0) - \frac{a^{(h)}(t)}{2} \sin(2\pi f_0 t + \varphi_0) \right] \quad . \quad (2.132)$$

L'involuppo della (2.132) è dato da

$$C \left[\left(\frac{a(t)}{2} + A \right)^2 + \left(\frac{a^{(h)}(t)}{2} \right)^2 \right]^{\frac{1}{2}} . \quad (2.133)$$

Notiamo che solo se A è molto elevato, allora il primo termine della somma nella (2.133) prevale e l'involuppo risulta proporzionale ad $a(t)$. Sfortunatamente ciò comporta un sistema con una efficienza η molto bassa, per cui un tale sistema è raramente utilizzato.

Capitolo 3

Modulazione angolare

3.1 Definizioni e parametri associati

L'espressione generale della portante modulata è data da (vedi 1.4)

$$s(t) = A \cos(2\pi f_0 t + \varphi_0 + \lambda(t)) \quad , \quad (3.1)$$

dove $\lambda(t) = \Delta\varphi_s(t)$ è la deviazione di fase istantanea [1]. Ricordiamo inoltre la relazione tra deviazione di fase istantanea $\Delta\varphi_s(t)$ e deviazione di frequenza istantanea $\Delta f_s(t)$,

$$\Delta f_s(t) = \frac{1}{2\pi} \frac{d}{dt} \lambda(t) \quad , \quad (3.2)$$

oppure, inversamente

$$\lambda(t) = 2\pi \int_{-\infty}^t \Delta f_s(\tau) d\tau \quad . \quad (3.3)$$

Ebbene, se il segnale modulante $a(t)$ varia $\Delta f_s(t)$ abbiamo la modulazione di frequenza (*frequency modulation*, FM) mentre se $a(t)$ varia $\lambda(t)$ abbiamo la modulazione di fase (*phase modulation*, PM). Entrambe le modulazioni vanno sotto il nome di modulazione angolare (*angle modulation*): analizziamole separatamente.

MODULAZIONE DI FREQUENZA

Sia K_F una costante di proporzionalità. Nella FM risulta

$$\Delta f_s(t) = K_F a(t) \quad . \quad (3.4)$$

Corrispondentemente la frequenza istantanea complessiva di $s(t)$ è pari a

$$f_s(t) = f_0 + K_F a(t) \quad , \quad (3.5)$$

mentre dalla (3.3) la deviazione di fase è data da

$$\lambda(t) = 2\pi K_F \int_{-\infty}^t a(\tau) d\tau \quad . \quad (3.6)$$

Il segnale modulato ha la seguente espressione

$$s(t) = A \cos(2\pi f_0 t + \varphi_0 + 2\pi K_F \int_{-\infty}^t a(\tau) d\tau) \quad . \quad (3.7)$$

E' facile verificare che se $f_0 \gg B$ la potenza media di $s(t)$ è

$$M_s = \frac{A^2}{2} \quad (3.8)$$

indipendente da $a(t)$.

D'altra parte dalla (3.6) la frequenza istantanea di $s(t)$ varia da un valore minimo

$$f_0 + K_F \min_t a(t) \quad (3.9)$$

ad un valore massimo

$$f_0 + K_F \max_t a(t) \quad (3.10)$$

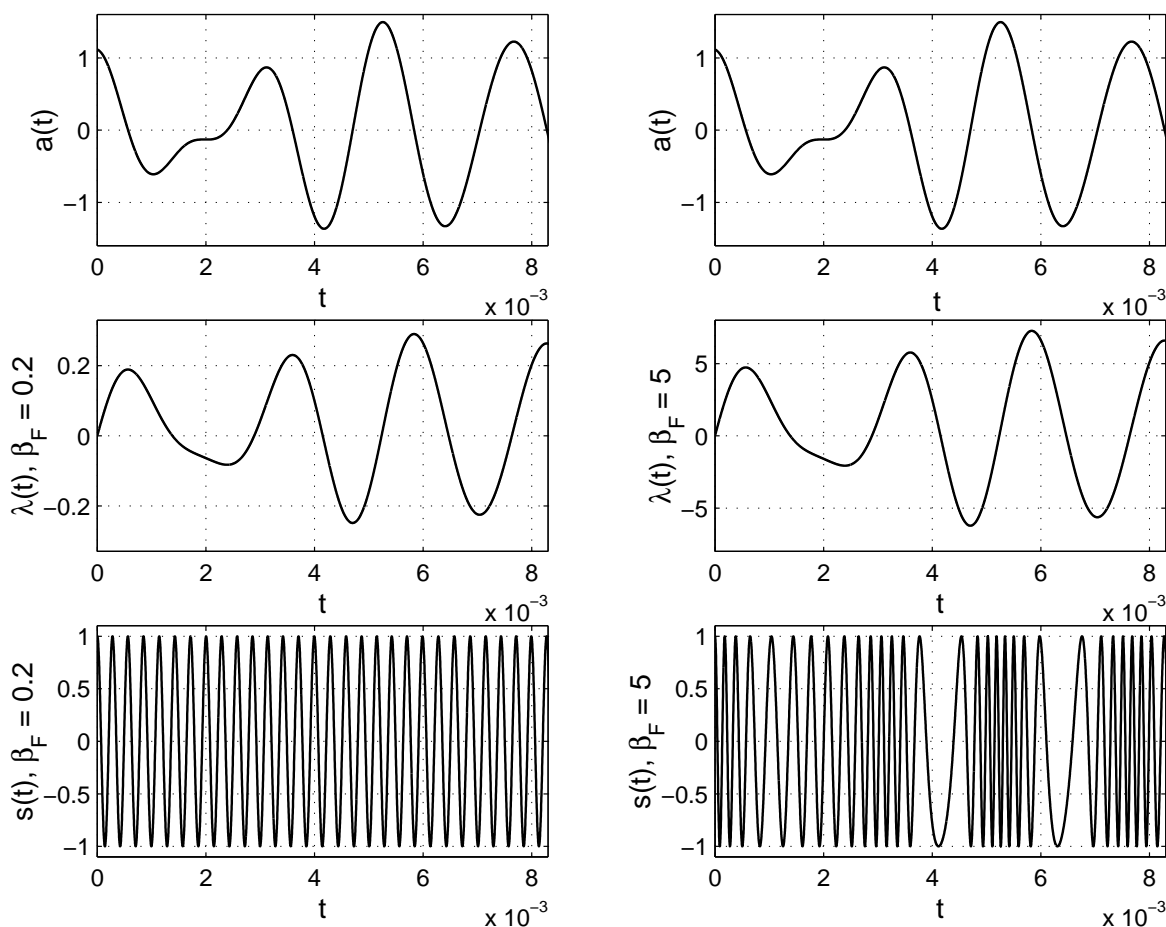


Figura 3.1: Esempio di segnale di informazione $a(t)$, corrispondente fase istantanea $\lambda(t)$ e segnale FM per $f_0 = 3500 Hz$, e due valori di $K_F = \frac{\beta_F B}{a_M}$: a) $\beta_F = 0.2$ e b) $\beta_F = 5$.

Questi due valori non devono però essere confusi con la banda passante di $s(t)$. Infatti, mentre la trasformata di Fourier di $s(t)$ è un integrale nel tempo, la frequenza istantanea non prende in considerazione la durata di ciascuna frequenza di $s(t)$.

Considerando come segnale modulante il segnale (2.3) costituito dalla somma di tre sinusoidi, riportiamo in Figura 3.1 l'andamento del segnale FM $s(t)$ e della fase istantanea $\lambda(t)$ per $f_0 = 3500 Hz$ e due valori di K_F .

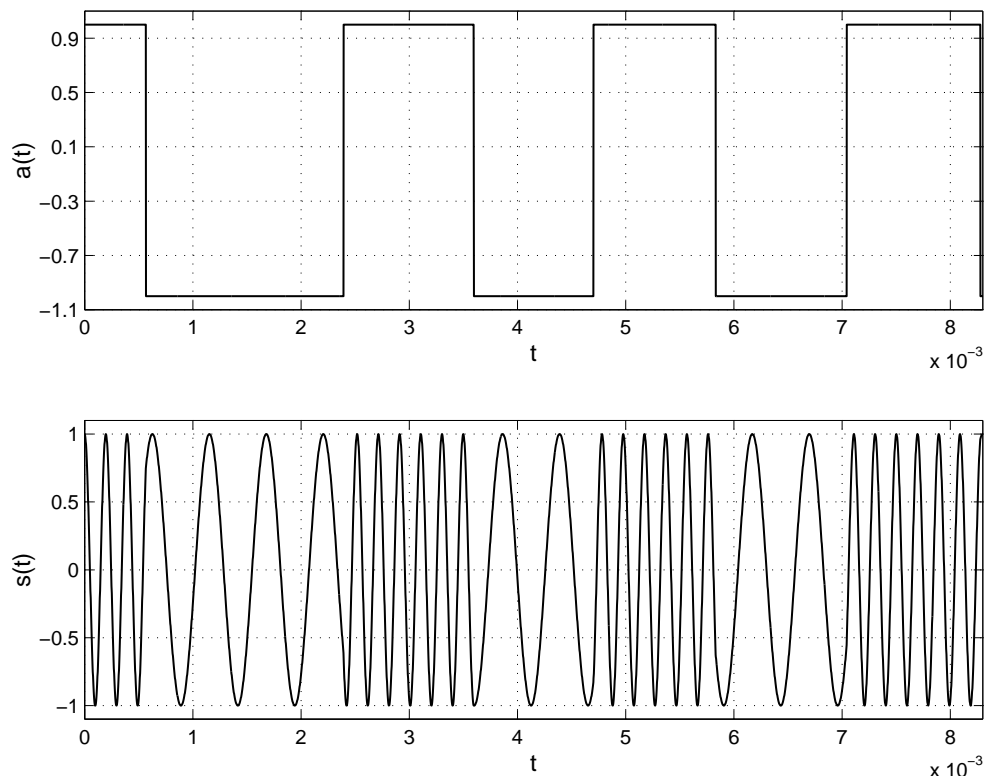


Figura 3.2: Segnale FM per un segnale modulante $a(t)$ del tipo costante a tratti.

Un secondo esempio illustrativo di segnale FM è riportato in Figura 3.2 in cui $a(t)$ è costante a tratti.

In effetti, nella FM l'informazione del segnale modulante è nella frequenza istantanea del segnale modulato. In termini approssimati possiamo dire che il numero di attraversamenti per lo zero di $s(t)$ nell'unità di tempo (cioè la densità degli zeri di $s(t)$) è proporzionale a $a(t)$.

MODULAZIONE DI FASE (PM)

Sia K_P una costante di proporzionalità. In questo caso la deviazione di fase è proporzionale al segnale modulante,

$$\lambda(t) = K_P a(t) . \quad (3.11)$$

Corrispondente

$$s(t) = A \cos(2\pi f_0 t + \varphi_0 + K_P a(t)) . \quad (3.12)$$

CORRISPONDENZA TRA SEGNALI FM E PM

Per un segnale di informazione $b(t)$, confrontando (3.7) e (3.12) abbiamo che un segnale FM può essere interpretato come PM in cui

$$a(t) = \int_{-\infty}^t b(\tau) d\tau . \quad (3.13)$$

Viceversa un segnale PM può essere interpretato come FM in cui

$$a(t) = \frac{d}{dt} b(t) \quad . \quad (3.14)$$

Allora, come illustrato in Figura 3.3, a partire da uno schema di modulazione, PM o FM, è possibile ottenere l'altro schema elaborando opportunamente il segnale di ingresso.

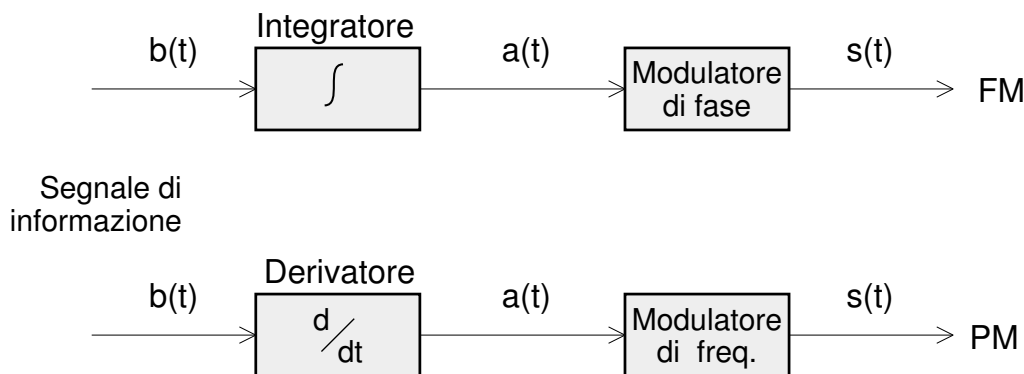


Figura 3.3: Generazione di un segnale FM utilizzando un PM e viceversa.

In ogni caso la modulazione di argomento non è lineare rispetto al segnale modulante.

INDICE DI MODULAZIONE

Per un segnale modulante $a(t)$ di banda B sia

$$a_M = \max_t |a(t)| \quad . \quad (3.15)$$

Anche in questo caso, se $a(t)$ ha ampiezza non limitata a_M viene definita tramite la probabilità

$$P[|a(t)| < a_M] = \varepsilon \quad , \quad (3.16)$$

con ε sufficientemente piccolo.

Per semplificare la notazione, supponiamo che $a_m = a_M$, dove a_m è definito in (2.7). Inoltre richiamiamo la definizione di fattore di forma k_f^2 dato della (2.14).

Definiamo l'indice di modulazione β per una modulazione angolare:

$$\begin{aligned}
 PM : \quad \beta_P &= \max_t |\Delta\varphi_s(t)| = K_P a_M = K_P \frac{\sqrt{M_a}}{k_f} \\
 FM : \quad \beta_F &= \frac{\max_t |\Delta f_s(t)|}{B} = \frac{K_F a_M}{B} = \frac{K_F \sqrt{M_a}}{B k_f}
 \end{aligned} \quad (3.17)$$

In altre parole, mentre nella PM β è dato dal valore massimo della deviazione di fase, nella FM β è dato dal valore massimo della deviazione di frequenza, diviso la banda di $a(t)$.

3.2 Narrowband FM

Per semplicità di scrittura poniamo

$$d(t) = \int_{-\infty}^t a(\tau) d\tau , \quad (3.18)$$

la quale, nel dominio della frequenza, porge

$$\mathcal{D}(f) = \frac{\mathcal{A}(f)}{j2\pi f} . \quad (3.19)$$

Il segnale modulato (3.7) può essere scritto nel seguente modo,

$$\begin{aligned} s(t) &= A \cos(2\pi f_0 t + \varphi_0 + 2\pi K_F d(t)) \\ &= A \cos(2\pi f_0 t + \varphi_0) \cos(2\pi K_F d(t)) \\ &\quad - A \sin(2\pi f_0 t + \varphi_0) \sin(2\pi K_F d(t)) . \end{aligned} \quad (3.20)$$

E' interessante esaminare l'espressione di $s(t)$ nel caso in cui K_F sia sufficientemente piccolo in modo che l'angolo $2\pi K_F d(t)$ sia sempre molto piccolo. In tal caso possiamo utilizzare le approssimazioni $\cos(2\pi K_F d(t)) \simeq 1$ e $\sin(2\pi K_F d(t)) \simeq 2\pi K_F d(t)$. Corrispondentemente abbiamo

$$s(t) = A \cos(2\pi f_0 t + \varphi_0) - A 2\pi K_F d(t) \sin(2\pi f_0 t + \varphi_0) . \quad (3.21)$$

A parte il primo termine, la portante, ora il legame tra $d(t)$ e $s(t)$ è lineare. Alla (3.21) corrisponde nel dominio della frequenza il segnale

$$\begin{aligned} \mathcal{S}(f) &= \frac{A}{2} [e^{j\varphi_0} \delta(f - f_0) + e^{-j\varphi_0} \delta(f + f_0)] \\ &\quad + \frac{A 2\pi K_F}{4\pi} \left[e^{j\varphi_0} \frac{\mathcal{A}(f - f_0)}{f - f_0} - e^{-j\varphi_0} \frac{\mathcal{A}(f + f_0)}{f + f_0} \right] . \end{aligned} \quad (3.22)$$

Per un segnale $a(t)$ a banda limitata B , il corrispondente segnale modulato ha uno spettro attorno f_0 di forma proporzionale ad $\mathcal{A}(f)/f$, come illustrato in Figura 3.4.

In tal caso, il segnale $s(t)$, denominato *FM a banda stretta*, ha una banda pari a $2B$, come nel caso DSB.

Possiamo inoltre affermare che nel caso K_F sia sufficientemente piccolo la modulazione FM è in grado di soddisfare gli obiettivi di un sistema di modulazione e cioè di fornire un segnale modulato a frequenze opportunamente elevate e di banda finita, in modo che più segnali possano condividere lo spettro. Il terzo obiettivo su come ricavare $a(t)$ dato $s(t)$ sarà esaminato in seguito.

Considerazioni analoghe valgono per un segnale PM in cui invece di $d(t)$ abbiamo direttamente $a(t)$.

Analizziamo ora la forma del segnale modulato $s(t)$ per il segnale $a(t)$ di tipo sinusoidale di frequenza B ,

$$a(t) = \cos(2\pi B t) . \quad (3.23)$$

Corrispondente

$$d(t) = \frac{\sin(2\pi B t)}{2\pi B} , \quad (3.24)$$

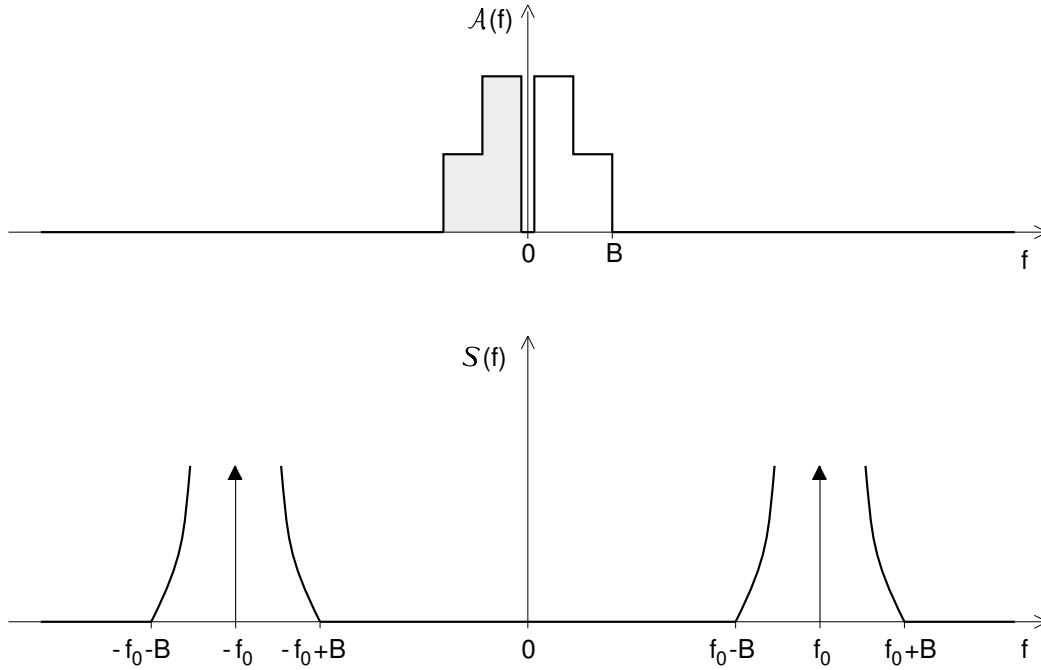


Figura 3.4: Segnale FM nel dominio della frequenza per K_F sufficientemente piccolo.

e sempre sotto l'ipotesi che $2\pi K_F d(t) \ll 1$, cioè per $\frac{K_F}{B} \ll 1$, dalla (3.21) risulta

$$s(t) = A \left[\cos(2\pi f_0 t + \varphi_0) - \frac{2\pi K_F}{2B} \sin(2\pi B t) \sin(2\pi f_0 t + \varphi_0) \right] \quad (3.25)$$

Utilizzando le relazioni trigonometriche

$$\cos(2\pi f_0 t) = \operatorname{Re} [e^{j(2\pi f_0 t + \varphi_0)}] \quad (3.26)$$

$$\sin(2\pi f_0 t) = \operatorname{Re} [-j e^{j(2\pi f_0 t + \varphi_0)}] \quad (3.27)$$

$$\sin(2\pi B t) = \frac{e^{j2\pi B t} - e^{-j2\pi B t}}{2j}, \quad (3.28)$$

abbiamo che

$$s(t) = A \operatorname{Re} \left\{ e^{j(2\pi f_0 t + \varphi_0)} \left[1 - \frac{K_F}{2B} e^{j2\pi B t} + \frac{K_F}{2B} e^{-j2\pi B t} \right] \right\} \quad (3.29)$$

A parte il termine di fase che cresce linearmente nel tempo (in particolare compie un angolo giro ogni $1/f_0$ secondi), dovuto alla frequenza della portante, il termine

$$\left[1 - \frac{K_F}{2B} e^{j2\pi B t} + \frac{K_F}{2B} e^{-j2\pi B t} \right], \quad (3.30)$$

comporta una fase aggiuntiva che dipende dal segnale di informazione e si compone come illustrato in Figura 3.5. Ricordiamo che in assenza di approssimazioni il modulo di questo fasore dovrebbe essere unitario.

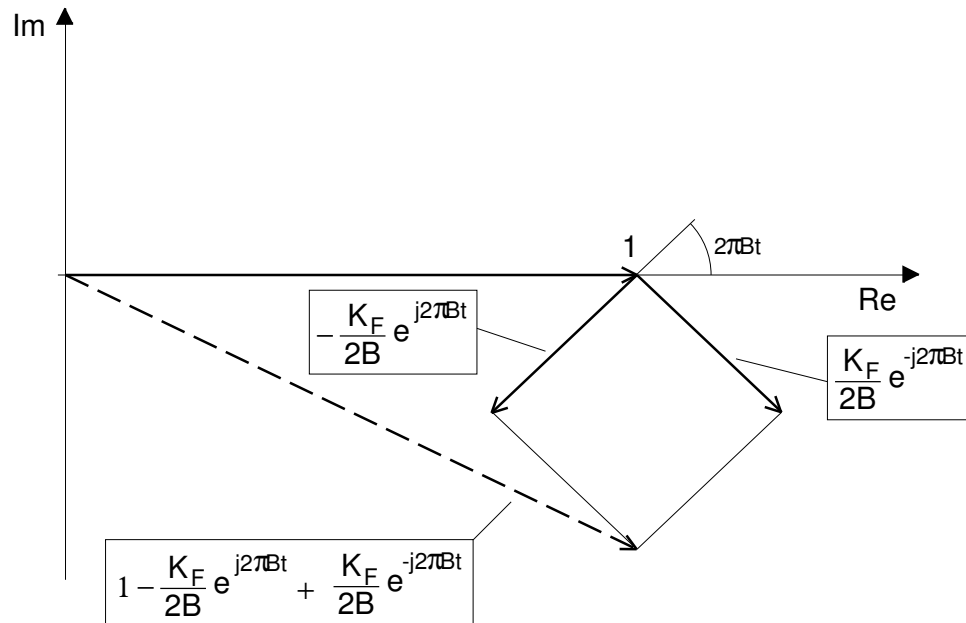


Figura 3.5: Composizione del termine di fase in un segnale FM per un segnale modulante sinusoidale di frequenza B .

3.3 Wideband FM

Nel caso generale in cui K_F non sia sufficientemente piccolo, si parla di *FM a banda larga*. In tal caso l'approssimazione (3.21) non è verificata e bisognerà riferirsi al segnale modulato

$$s(t) = A \cos(2\pi f_0 t + \varphi_0 + 2\pi K_F d(t)) \quad , \quad (3.31)$$

dove $d(t)$ è definito in (3.18).

Dal momento che il legame tra $a(t)$ e $s(t)$ è non lineare in generale non è possibile esprimere in forma analitica la trasformata di Fourier di $s(t)$ in funzione della trasformata di Fourier di $a(t)$. Comunque, nelle PM e FM è consuetudine utilizzare come banda di $s(t)$ la formula di Carson,

$$B_s = 2B(1 + \beta) \quad , \quad (3.32)$$

dove B è la banda del segnale modulante e β è l'indice di modulazione definito in (3.17). In realtà la (3.32) dà solamente una indicazione della banda effettiva di $s(t)$ ed è opportuno ricorrere a simulazioni per una più precisa valutazione dello spettro del segnale modulato.

E' interessante comunque ripercorrere l'analisi di Carson per la determinazione della (3.32). Innanzitutto l'analisi è stata fatta per un segnale modulante di tipo sinusoidale a frequenza B ,

$$a(t) = a_M \cos(2\pi B t) \quad , \quad (3.33)$$

per il quale nella FM

$$2\pi K_F d(t) = \beta_F \sin(2\pi B t) \quad . \quad (3.34)$$

Nella PM invece

$$K_P a(t) = \beta_P \cos(2\pi B t) \quad . \quad (3.35)$$

In entrambi i casi (a parte la sostituzione della funzione $\sin(2\pi Bt)$ con $\cos(2\pi Bt)$ passando da FM a PM),

$$\begin{aligned} s(t) &= A \cos(2\pi f_0 t + \varphi_0 + \beta \sin(2\pi Bt)) \\ &= A \operatorname{Re} [e^{j(2\pi f_0 t + \varphi_0)} e^{j\beta \sin(2\pi Bt)}] . \end{aligned} \quad (3.36)$$

Utilizzando lo sviluppo in serie di Fourier della funzione $e^{j\beta \sin(2\pi Bt)}$, periodica di periodo $1/B$, risulta

$$e^{j\beta \sin(2\pi Bt)} = \sum_{k=-\infty}^{+\infty} J_k(\beta) e^{jk2\pi Bt} , \quad (3.37)$$

dove

$$J_k(\beta) = \frac{1}{2\pi} \int_0^{2\pi} e^{j(\beta \sin u - ku)} du , \quad (3.38)$$

è la funzione di Bessel di prima specie di ordine k [3] Otteniamo così

$$s(t) = A \sum_{k=-\infty}^{+\infty} J_k(\beta) \cos(2\pi(f_0 + kB)t + \varphi_0) . \quad (3.39)$$

In altre parole, la trasformata di Fourier di un segnale FM modulato da un segnale sinusoidale di frequenza B è composta da una sequenza infinita di righe attorno alla portante, di ampiezze $\{J_k(\beta)\}$. In pratica, adottando il criterio di considerare solo i termini entro gli indici $[-k_{max}, k_{max}]$ per i quali si ottiene una potenza superiore al 90% di quella del segnale complessivo, il numero di termini significativi in (3.39) dipende da β però è finito ed è nell'ordine di $2k_{max} \simeq 2\beta + 3$. Poichè due righe spettrali distano di B Hz, si ottiene il risultato (3.32).

Benchè la formula di Carson sia stata derivata per segnali modulanti sinusoidali, sperimentalmente essa si è rivelata valida anche per segnali non sinusoidali di banda B . Per una spiegazione intuitiva di questo fatto rimandiamo a [3].

Vediamo alcune osservazioni:

1. Notiamo che in base alla (3.32) utilizzare un elevato valore di β comporta una banda elevata del sistema. D'altra parte, come vedremo in seguito, un elevato β migliora le prestazioni del ricevitore.
2. Definita la massima deviazione di frequenza del segnale modulato

$$\Delta F = \max_t |\Delta f_s(t)| = K_F a_M , \quad (3.40)$$

per un segnale FM la banda richiesta si può scrivere nel seguente modo

$$B_s = 2(B + \Delta F) . \quad (3.41)$$

3. Per $\beta \ll 1$, oppure in modo equivalente per $\Delta F \ll B$, la (3.32) fornisce il valore di $B_s \simeq 2B$ come avevamo visto per un segnale FM a banda stretta.

3.4 Modulatori

Metodo indiretto

Iniziamo con lo schema di Figura 3.6 in cui viene riportata la realizzazione di un modulatore FM a banda stretta, cioè con un basso indice di modulazione β_F per il quale vale la relazione (3.21). Analogamente, per un modulatore PM a banda stretta,

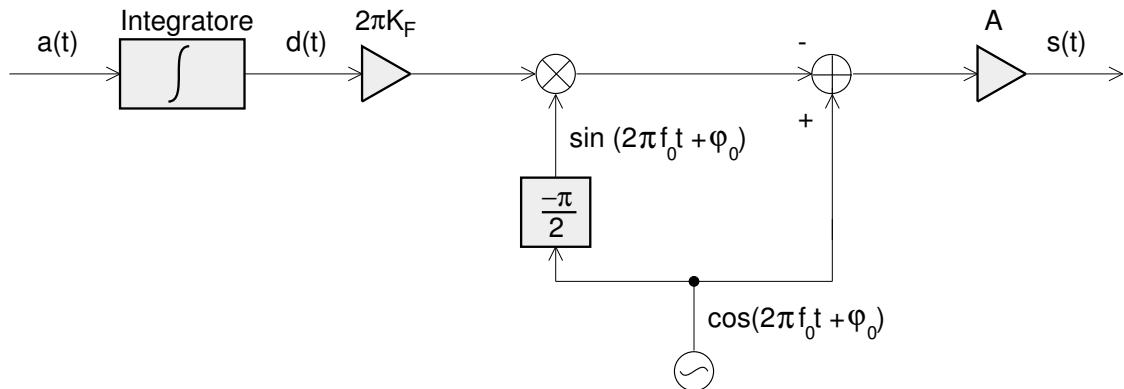


Figura 3.6: Modulatore FM a banda stretta.

con $\beta_P \ll 1$, riportiamo in Figura 3.7 uno schema realizzativo.

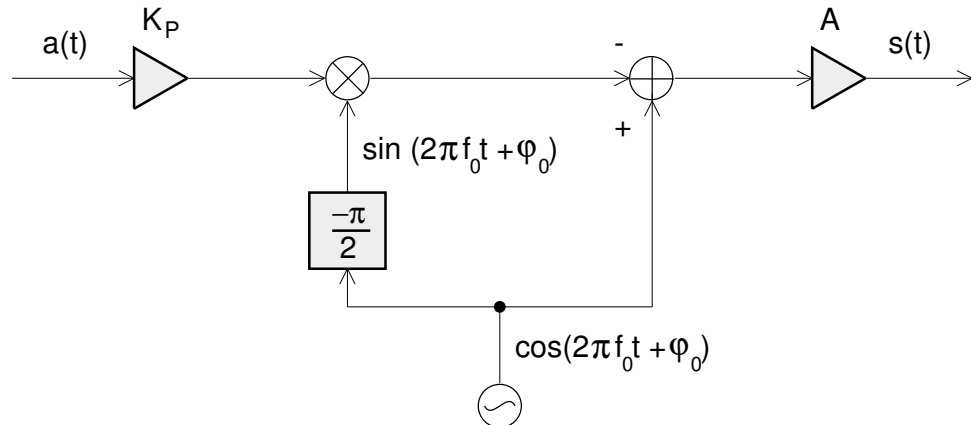


Figura 3.7: Modulatore PM a banda stretta.

Ora per ottenere un modulatore FM a larga banda si può utilizzare lo schema di Armstrong riportato in Figura 3.8 in cui viene generato un segnale FM a banda stretta $s'(t)$ dato da

$$A'[\cos(2\pi f'_o t + \varphi'_o) - 2\pi K_F d(t) \sin(2\pi f'_o t + \varphi'_o)] , \quad (3.42)$$

che in base alla (3.21) può essere scritto come

$$s'(t) = A' \cos(2\pi f'_o t + \varphi'_o + 2\pi K_F d(t)) . \quad (3.43)$$

Il *limiter* di Figura 3.8 normalizza l'involuppo di (3.42) ad una costante. Vediamone il funzionamento.

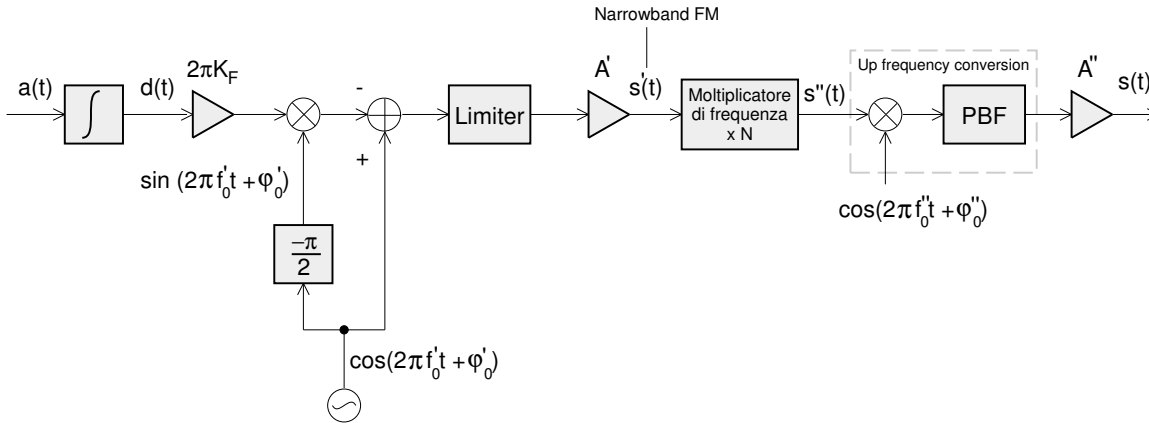


Figura 3.8: Metodo indiretto o di Armstrong per generare un segnale FM a larga banda.

Dato un segnale $x(t) = M_x(t) \cos(\varphi_x(t))$, di banda B_x e portante f_0 , la caratteristica di un limitatore (vedi Figura 3.9) è di normalizzare l'involucro di $x(t)$ squadrando il segnale di ingresso,

$$y(t) = \begin{cases} 1, & x(t) > 0 \\ -1, & x(t) < 0 \end{cases} \quad (3.44)$$

e rimuovendo tramite un PBF le armoniche di $y(t)$ oltre quelle desiderate attorno f_0 , per cui idealmente

$$\mathcal{H}_{PBF}(f) = \text{rect} \frac{f - f_0}{B_x} + \text{rect} \frac{f + f_0}{B_x} . \quad (3.45)$$

In uscita avremo

$$z(t) = \cos(\varphi_x(t)) . \quad (3.46)$$

Ritornando allo schema di Figura 3.8, tramite un *moltiplicatore di frequenza* la frequenza istantanea di $s'(t)$ viene moltiplicata per N , formando un segnale $s''(t)$ con frequenza istantanea

$$f_{s''}(t) = (f'_0 + K_F a(t))N = f'_0 N + K_F N a(t) . \quad (3.47)$$

In effetti mentre la banda del segnale modulante è sempre B , la massima deviazione di frequenza e corrispondentemente l'indice di modulazione sono aumentati di un fattore N . A partire dal segnale $s''(t)$, utilizzando un opportuno stadio di conversione di frequenza, denominato *up-frequency conversion*, si può "spostare" $s''(t)$ attorno la frequenza desiderata $f_0 = f'_0 N + f''_0$.

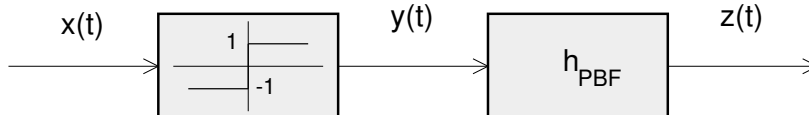


Figura 3.9: Schema di un limitatore (*limiter*).

Un metodo per elevare l'indice di modulazione è tramite lo schema di Figura 3.10 in cui $s'(t)$ del tipo (3.43) viene trasformato tramite un dispositivo non lineare con

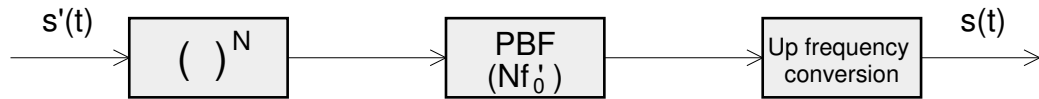


Figura 3.10: Metodo per elevare l'indice di modulazione di un segnale FM.

legge esponenziale di ordine N cui risultato è la creazione di repliche del segnale $s'(t)$ attorno alle frequenze nf_0' , $n = 1, \dots, N$. Di tutte queste repliche il filtro selettivo seleziona la replica desiderata attorno $f_0'N$ alla quale corrisponde un segnale con indice di modulazione pari a quello di $s'(t)$ moltiplicato per N .

Metodo diretto

Un metodo diretto per generare un segnale FM è tramite un VCO, come illustrato in Figura 3.11, all'interno di una configurazione PLL che insegue una frequenza di riferimento f_0/N generata da un oscillatore. Il divisore di frequenza è necessario per abbassare l'indice di modulazione, diciamo a circa 0.2, del segnale all'uscita del VCO e così avere disponibile un segnale FM a banda stretta del tipo (3.21) con una riga alla frequenza f_0/N sempre presente. La configurazione PLL serve per creare un'uscita $s(t)$ con una portante f_0 molto stabile, grazie al meccanismo di feedback. Esistono inoltre

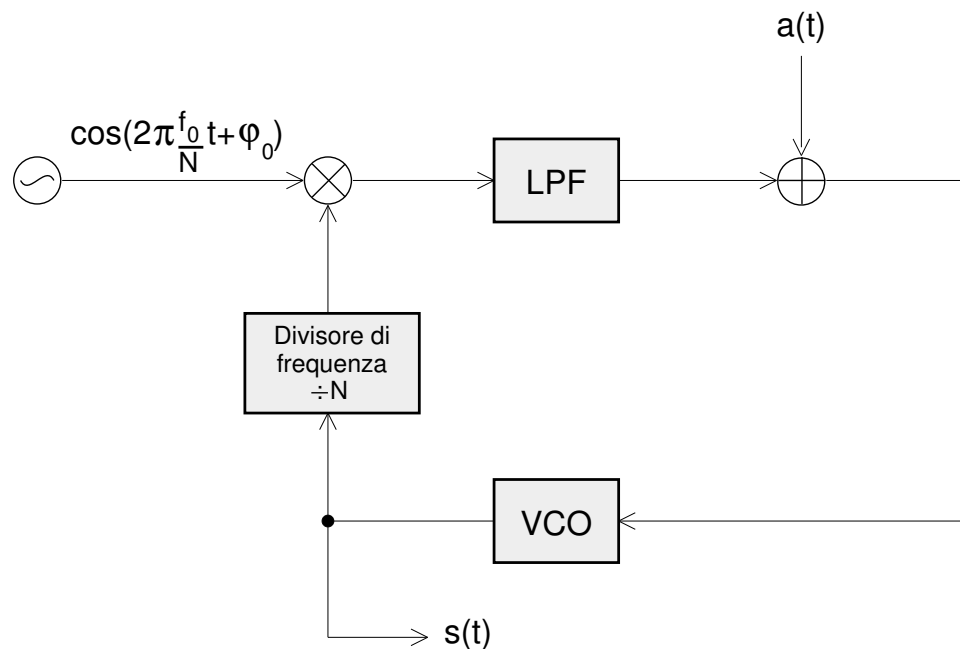


Figura 3.11: Metodo diretto per generare un segnale FM utilizzando un PLL.

schemi più semplici che utilizzano il solo principio del VCO controllato dal segnale di informazione [2, 3].

3.5 Demodulatori

Discriminatore

Lo schema è riportato in Figura 3.12 e fa uso di un filtro derivatore con specifiche ideali del tipo

$$\mathcal{H}_d(f) = jK_d f \quad (3.48)$$

In effetti è sufficiente che siano verificate le condizioni

$$\begin{cases} |\mathcal{H}_d(f)| = K_d f \\ \arg \mathcal{H}_d(f) = \frac{\pi}{2} \end{cases} \quad \text{per } |f - f_0| < B_s/2 \quad (3.49)$$

e $\mathcal{H}_d(f) = \mathcal{H}_d^*(-f)$ per $|f + f_0| < B_s/2$. Allora, entro la banda passante di $s(t)$ avremo

$$\begin{aligned} s_d(t) &= \bar{K}_d \frac{d}{dt} s(t) \\ &= -\bar{K}_d 2\pi A (f_0 + K_F a(t)) \sin(2\pi f_0 t + \varphi_0 + 2\pi K_F \int_{-\infty}^t a(\tau) d\tau) \quad (3.50) \end{aligned}$$

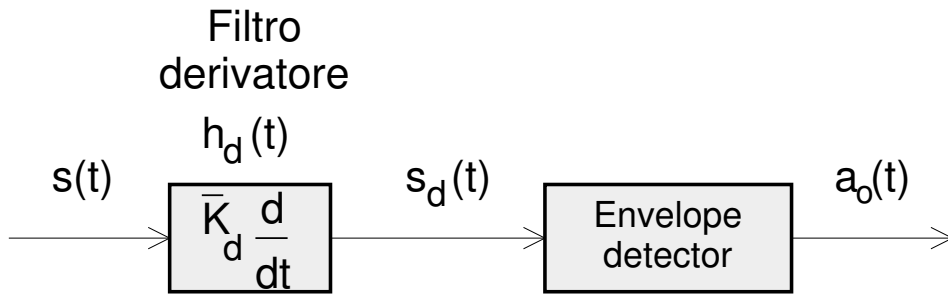


Figura 3.12: Demodulatore di frequenza del tipo discriminatore.

in cui $\bar{K}_d = \frac{K_d}{2\pi}$ è una costante non essenziale.

Notiamo che il filtro derivatore ha tramutato un segnale FM in AM. Ora in (3.50) il termine sinusoidale è un segnale FM con banda passante da $f_0 - B_s/2$ a $f_0 + B_s/2$ mentre il termine $2\pi(f_0 + K_F a(t))$ ha una banda B . Sotto l'ipotesi che $B < f_0 - B_s/2$, estraendo l'involuppo della (3.50) avremo, a parte fattori non essenziali

$$\begin{aligned} a_0(t) &= 2\pi |f_0 + K_F a(t)| \\ &= 2\pi (f_0 + K_F a(t)) \end{aligned} \quad (3.51)$$

essendo $f_0 \gg K_F a_m = \beta B$. Utilizzando un filtro *notch* per eliminare la DC dalla (3.51) riotteniamo il segnale desiderato. La cascata del filtro derivatore e del demodulatore di involuppo prende il nome di *discriminatore*.

Tra le varie strutture per realizzare un derivatore ricordiamo quella che fa uso dell'approssimazione

$$s(t) - s(t - \Delta T) = \Delta T \frac{d}{dt} s(t) \quad , \quad (3.52)$$

se ΔT è sufficientemente piccolo o meglio $\Delta T \ll 1/B_s$. Questa relazione porta al demodulatore *phase shift* illustrato in Figura 3.13.

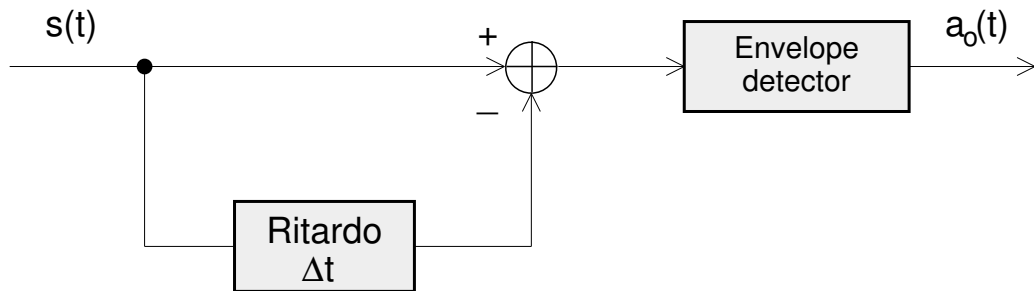


Figura 3.13: Demodulatore di frequenza del tipo *phase shift*.

Phase locked loop

Lo schema è riportato in Figura 3.14. Notiamo che il blocco *phase-difference* (PD) è

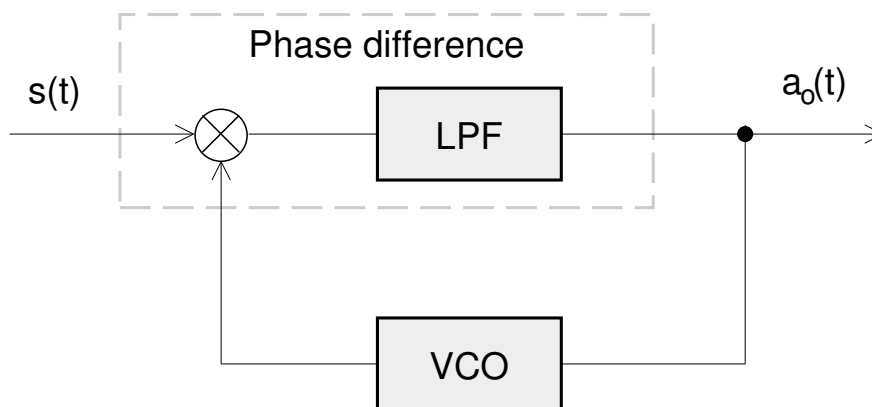


Figura 3.14: Demodulatore di frequenza tramite PLL.

stato realizzato tramite un prodotto a cui segue un filtro passa basso. Si può verificare che il VCO nell'inseguire la fase di $s(t)$ produce un errore all'uscita del PD proporzionale alla deviazione di frequenza istantanea di $s(t)$, $\Delta f_s(t)$. Di conseguenza l'uscita del blocco PD è proprio il segnale desiderato.

In tutti questi schemi, se il segnale di ingresso, a causa della distorsione introdotta dal canale non ideale, non ha un inviluppo costante, la derivata del segnale produce un termine di distorsione proporzionale alla derivata dell'inviluppo. E' allora importante precedere il derivatore da un *limiter*, del tipo illustrato in Figura 3.9.

3.6 Prestazioni

FM

Lo schema complessivo del ricevitore è riportato in Figura 3.15.

In presenza del rumore introdotto dal canale, il segnale al ricevitore $r(t)$ viene filtrato da un filtro del tipo banda passante di banda B_s attorno la portante f_0 . Il segnale assume la forma

$$r_{Rc}(t) = \zeta A \cos(2\pi f_0 t + \varphi_0 + 2\pi K_F d(t)) + \omega_{Rc}(t) \quad (3.53)$$

dove ζ è dovuta all'attenuazione del canale mentre $\omega_{Rc}(t)$ è il filtraggio di $\omega(t)$, rumore bianco con spettro $N_0/2$.

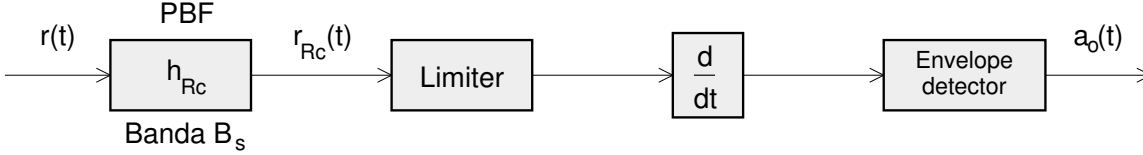


Figura 3.15: Schema complessivo semplificato di un ricevitore FM.

Utilizzando la usuale scomposizione di $\omega_{Rc}(t)$ nelle componenti in fase e quadratura attorno f_0 ,

$$\omega_{Rc}(t) = \omega_{Rc,I}(t) \cos(2\pi f_0 t + \varphi_0) - \omega_{Rc,Q}(t) \sin(2\pi f_0 t + \varphi_0) \quad , \quad (3.54)$$

possiamo scrivere

$$r_{Rc}(t) = Re \left\{ e^{j(2\pi f_0 t + \varphi_0)} \left[\mathcal{C} A e^{j2\pi K_F d(t)} + \omega_{Rc,I}(t) + j\omega_{Rc,Q}(t) \right] \right\} \quad (3.55)$$

E' proprio la fase del segnale

$$r_{Rc}^{(bb)}(t) = \mathcal{C} A e^{j2\pi K_F d(t)} + \omega_{Rc,I}(t) + j\omega_{Rc,Q}(t) \quad (3.56)$$

che porta il segnale di informazione. La variazione rispetto a $2\pi K_F d(t)$ è dovuta al rumore (vedi Figura 3.16).

L'ampiezza di $r_{Rc}^{(bb)}(t)$ è irrilevante poichè viene normalizzata dal limitatore. Indichiamo con $\vartheta(t)$ l'errore di fase introdotto dal rumore:

$$M e^{j\vartheta(t)} = \frac{r_{Rc}^{(bb)}(t)}{\mathcal{C} A e^{j2\pi K_F d(t)}} = 1 + \frac{\omega_{Rc,I}(t) + j\omega_{Rc,Q}(t)}{\mathcal{C} A e^{j2\pi K_F d(t)}} \quad (3.57)$$

Ebbene, la derivata di $\vartheta(t)$ è il segnale che va a sommarsi a $2\pi K_F \dot{a}(t)$ all'uscita della cascata derivatore-rivelatore di inviluppo. Qui omettiamo la derivazione e riportiamo solamente i punti salienti nel caso particolare il rapporto segnale-rumore Γ sia sufficientemente elevato.

- i) L'errore di fase $\vartheta(t)$ dovuto al rumore all'uscita di filtro di ricezione ha uno spettro bianco all'interno della banda $(-B, B)$ del segnale desiderato:

$$\mathcal{P}_\vartheta(f) = \frac{N_0}{\mathcal{C}^2 A^2} \quad , \quad |f| < B \quad (3.58)$$

- ii) All'uscita del filtro derivatore, $\dot{\vartheta}(t)$ avrà uno spettro di forma parabolica:

$$\mathcal{P}_{\dot{\vartheta}}(f) = (2\pi f)^2 \frac{N_0}{\mathcal{C}^2 A^2} \quad , \quad |f| < B \quad (3.59)$$

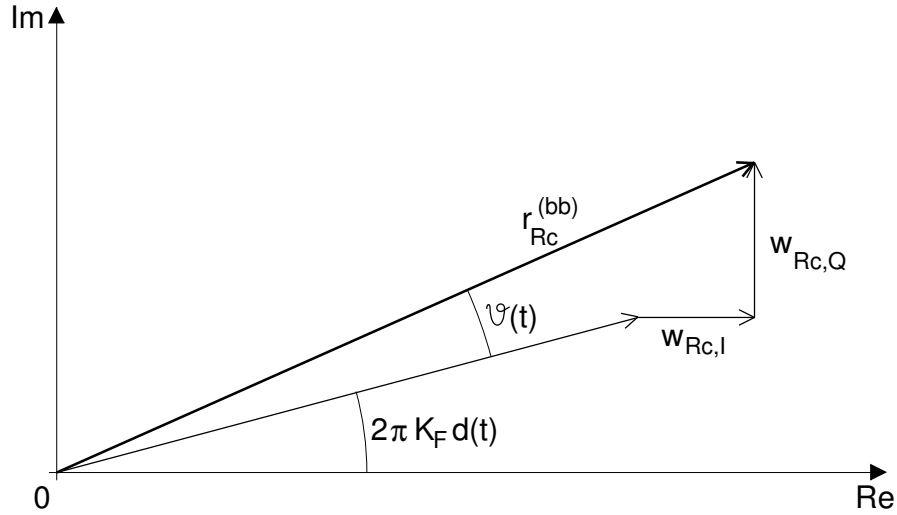


Figura 3.16: Composizione del fasore $r_{Rc}^{(bb)}(t)$.

iii) La potenza di $\dot{\vartheta}(t)$ è data da

$$M_{\dot{\vartheta}}(f) = \int_{-B}^B \mathcal{P}_{\dot{\vartheta}}(f) df = (2\pi)^2 \frac{N_0}{C^2 A^2} \frac{2}{3} B^3 \quad (3.60)$$

Richiamando la potenza del termine desiderato (3.51), rimosso della DC, data da

$$M_{a_o} = (2\pi K_F)^2 M_a \quad , \quad (3.61)$$

il rapporto segnale-rumore all'uscita del ricevitore ha la seguente espressione

$$\Lambda_o = \frac{(2\pi K_F)^2 M_a}{(2\pi)^2 \frac{N_0}{C^2 A^2} \frac{2}{3} B^3} = \frac{(CA)^2}{2} \frac{3 K_F^2 M_a}{N_0 B B^2} \quad (3.62)$$

Moltiplicando e dividendo la precedente espressione per k_f^2 e richiamando la definizione di β_F (3.17) e quella di Γ (2.93) dove

$$M_s = \frac{A^2}{2} \quad , \quad (3.63)$$

risulta

$$\Lambda_o = 3k_f^2 \beta_F^2 \Gamma \quad . \quad (3.64)$$

Dalla espressione (3.64) notiamo che se $k_f \beta_F > \frac{1}{\sqrt{3}}$, cioè per valori elevati dell'indice di modulazione, il ricevitore MF è il più efficiente di quello AM. L'inconveniente è che un β_F elevato comporta anche una elevata banda richiesta, come si osserva dalla formula di Carson (3.32).

Il termine $3k_f^2 \beta_F^2$ prende anche il nome di guadagno del ricevitore G_{Ric} . Ad esempio nella radiodiffusione del suono in MF si utilizza una banda $B = 15 \text{ kHz}$ e $\beta_F = 5$. Corrispondentemente $B_s = 2B(1 + \beta_F) = 180 \text{ kHz}$ e assumendo un fattore di forma di $a(t)$ pari a $k_f^2 = \frac{1}{2}$, valore tipico per un segnale voce, risulta un guadagno del ricevitore pari a

$$(G_{Ric})_{dB} = 15.7 \text{ dB} \quad . \quad (3.65)$$

Concludiamo questa sezione sulle prestazioni di un ricevitore FM ricordando che le precedenti considerazioni e il risultato (3.64) valgono nell'ipotesi di basso rumore rispetto al segnale desiderato e ciò comporta di aver un rapporto Γ superiore ad una certa soglia data da

$$\Gamma_{th} = 20(1 + \beta_F) . \quad (3.66)$$

In altre parole solo *operando sopra soglia*, ossia per $\Gamma \geq \Gamma_{th}$, allora vale la (3.64). Se $\Gamma < \Gamma_{th}$ si può constatare un decadimento delle prestazioni secondo la legge

$$\Lambda_o = 3k_f^2 \beta_F^2 \frac{1}{1 + \frac{2(1+\beta)}{\Gamma}} ; \quad (3.67)$$

il quale è riportato in Figura 3.17

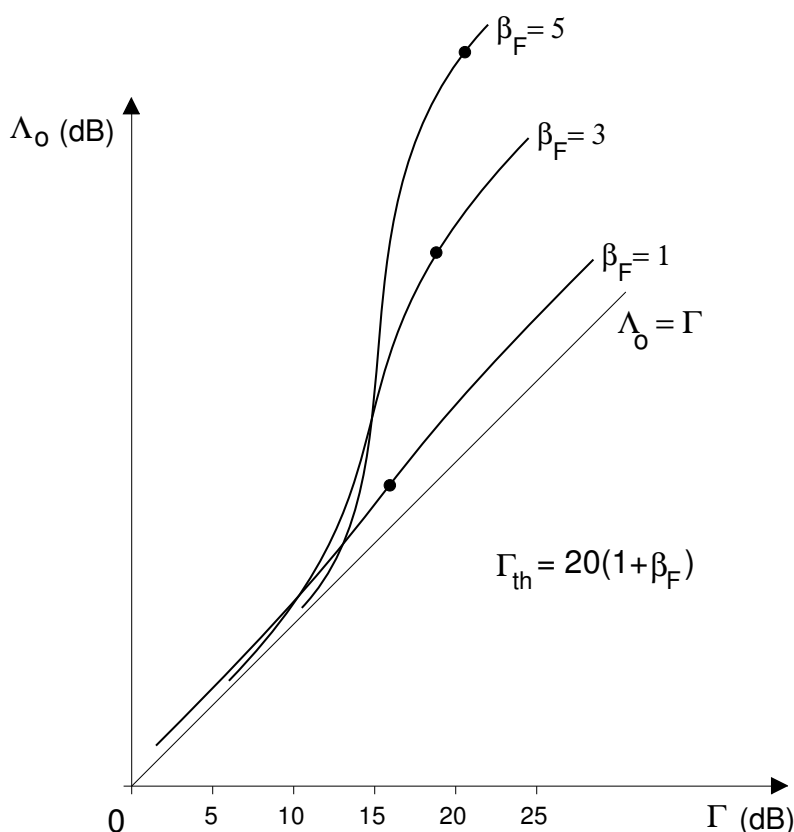


Figura 3.17: Andamento di Λ_o in funzione di Γ per diversi valori di β in un ricevitore FM.

Dal momento che β determina sia le prestazioni Λ_o che la banda richiesta B_s , bisognerà determinare quale delle due condizioni è più restrittiva. Vediamo il seguente esempio.

Esempio

Si desidera progettare un sistema FM in modo tale che la potenza trasmessa M_s sia minima. I requisiti e parametri del sistema siano i seguenti

$$\Lambda_o = 40 \text{ dB} , \quad B_s = 120 \text{ kHz}$$

$$B = 10 \text{ kHz} , \quad k_f^2 = \frac{1}{2} , \quad \frac{N_0}{2} = 0.5 \cdot 10^{-8} \text{ (V}^2/\text{Hz)} , \quad \mathcal{C} = 1 .$$

Ora il vincolo sulla banda impone che

$$2B(1 + \beta_F) \leq 120 \text{ kHz} , \quad (3.68)$$

per cui deve essere

$$\beta \leq 5 . \quad (3.69)$$

Il vincolo su Λ_o , supponendo di lavorare sopra soglia, impone che

$$3 k_f^2 \beta_F^2 20(1 + \beta_F) \geq 10^4 , \quad (3.70)$$

per cui deve essere

$$\beta_F \geq 6.\bar{6} . \quad (3.71)$$

Di conseguenza il sistema è limitato in banda e scegliamo

$$\beta_F = 5 . \quad (3.72)$$

Lavorando sopra soglia segue che deve essere

$$\Lambda_o = 3 k_f^2 \beta_F^2 \Gamma = 10^4 \quad (3.73)$$

da cui segue

$$(\Gamma)_{dB} = 24 \text{ dB} . \quad (3.74)$$

Dall'espressione $\Gamma = \frac{M_s}{N_0 B}$ si ricava infine

$$(M_s)_{dBm} = 14.3 \text{ dBm} . \quad (3.75)$$

Se non ci fosse stato il vincolo in banda allora per $\beta_F = 6.\bar{6}$ e $(\Gamma)_{dB} = 24 \text{ dB}$ sarebbe risultato

$$(M_s)_{dBm} = 11.8 \text{ dBm} , \quad (3.76)$$

con un risparmio di 2.5 dB nella potenza trasmessa, però con una occupazione di banda di 153 kHz .

PM

Secondo lo schema di Figura 3.3, per ottenere un demodulatore di fase è sufficiente integrare l'uscita di un ricevitore FM.

Dal punto di vista analitico si tratta di estrarre la fase del segnale all'uscita del *limiter*. Procedendo come nel caso MF, si ottiene un rapporto segnale-rumore all'uscita del ricevitore pari a

$$\Lambda_o = k_f^2 \beta_P^2 \Gamma , \quad (3.77)$$

dove β_P è definito in (3.17).

3.7 Preenfasi e deenfasi nella FM

Nella FM abbiamo osservato come la densità spettrale del rumore all'uscita del ricevitore abbia un andamento del tipo f^2 , cioè sia più concentrata verso le alte frequenze. L'idea è allora di filtrare il segnale ricostruito $a_0(t)$ con un ulteriore filtro $h_{de}(t)$ che attenui maggiormente le alte frequenze. Per non distorcere il segnale di informazione ricostruito bisognerà però filtrare il segnale di informazione $a(t)$ con un filtro h_{pe} il cui andamento in frequenza è il reciproco di quello h_{de} . In altre parole deve essere

$$\mathcal{H}_{pe}(f) \simeq \frac{1}{\mathcal{H}_{de}(f)} \quad , \quad |f| \leq B \quad , \quad (3.78)$$

secondo lo schema di Figura 3.18.

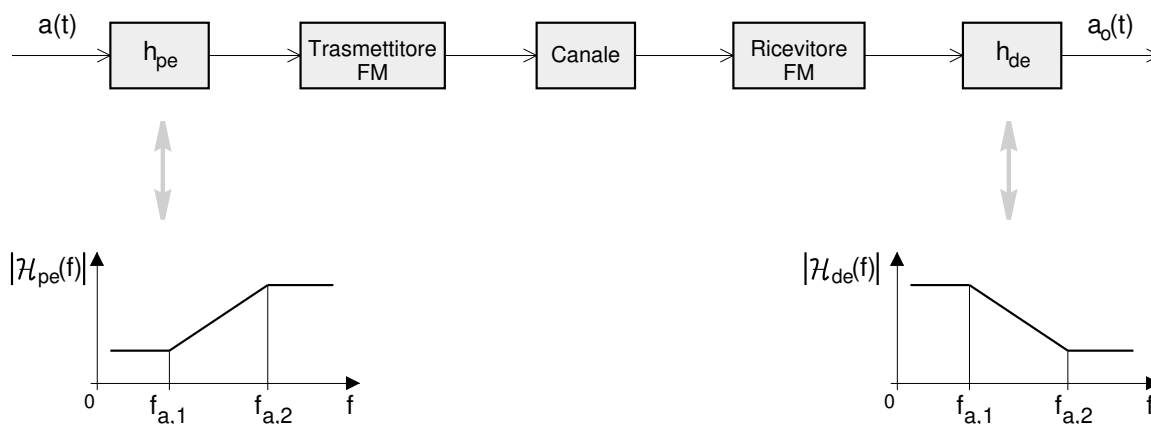


Figura 3.18: FM con preenfasi e deenfasi.

Naturalmente bisognerà scalare il segnale all'uscita del filtro h_{pe} in modo che rimanga invariata la potenza del segnale modulante. Diversamente potrebbe aumentare β_F e di conseguenza la banda del segnale modulato.

Nelle radio FM tipicamente

$$\mathcal{H}_{de}(f) = \frac{1}{1 + j \frac{f}{f_c}} \quad (3.79)$$

con $f_c = 2100 \text{ Hz}$ e $B = 15 \text{ kHz}$. A parità di potenza trasmessa, il rapporto Λ_o migliora di 6, 12 dB.

Capitolo 4

Confronto tra i vari metodi di modulazione ed esempi di sistemi

Il criterio di confronto si basa sulla banda occupata B_s , prestazioni (guadagno) del ricevitore e complessità realizzativa.

Banda B_s

La modulazione più efficiente in banda occupata è la SSB-SC con $B_s = B$. Essa è utilizzata nelle applicazioni con restrizioni in banda in cui $a(t)$ non ha DC, tipo nelle trasmissioni voce alle microonde e nelle trasmissioni satellitari. Per trasmissioni di segnale video si utilizza la VSB. Di certo non si utilizza la FM nei sistemi in cui la banda è scarsa.

Efficienza in potenza (o guadagno G_{Ric})

La FM si utilizza nelle applicazioni in cui la potenza trasmessa è critica, ad esempio nelle trasmissioni radio ad alta fedeltà. Di certo le DSB-TC e SSB-TC sono le meno efficienti.

Complessità realizzativa

AM con demodulatore non coerente e FM sono i sistemi con ricevitore più semplice da realizzare. Di fatto questi sistemi sono molto usati nelle applicazioni *broadcasting* tipo radio AM, video, radio FM ad alta fedeltà ed FM stereo.

La SSB-SC richiede un demodulatore sincronizzato che è alquanto complesso, per cui non viene utilizzata nelle applicazioni *broadcasting*. Tantomeno viene utilizzata la DSB-SC che richiede il doppio della banda.

4.1 Esempi di sistemi di trasmissione analogici

Prima di passare in rassegna alcuni sistemi di trasmissione, riportiamo in dettaglio la tipica configurazione del *front – end* di un ricevitore radio che va sotto il nome di supereterodina [1, Cap. 18]. Recentemente, per le radio digitali sono emerse architetture alternative con “conversione diretta”, le quali hanno il vantaggio di essere integrabili in un singolo chip [1].

4.1.1 Ricevitore supereterodina

L'obiettivo del *front – end* di un ricevitore radio è di estrarre il segnale modulato desiderato, corrotto dal rumore, dalla miriade di segnali presenti nell'aria. La struttura è riportata in Figura 4.1. Il principio è di traslare in giù lo spettro del segnale desiderato corrotto dal rumore, diciamo dalle radio frequenze (RF) alle frequenze intermedie (IF). A questo punto il segnale può essere rivelato tramite un demodulatore del tipo esaminato nei capitoli precedenti. Il filtro RF, del tipo passabanda, ha il compito di amplificare il segnale prima di entrare nel *mixer*, il quale, tra l'altro introduce esso stesso del rumore.

Il compito del filtro RF è anche quello di attenuare molti dei segnali che si trovano spettralmente vicino al segnale desiderato, secondo specifiche che vedremo tra poco. Ad ogni modo la *reiezione dei canali adiacenti* è demandata al filtro alle frequenze intermedie posto dopo il *mixer*.

Sia f_0 la frequenza della portante del segnale desiderato e B_s la sua banda. Selezioniamo inoltre una volta per tutte la frequenza portante f_{IF} del filtro IF. scegliamo allora come frequenza dell'oscillatore locale

$$f_{LO} = f_0 + f_{IF} . \quad (4.1)$$

In tal caso sarà $f_{LO} > f_0$. Il sintonizzatore seleziona come frequenza centrale del filtro RF proprio f_0 e come frequenza dell'oscillatore locale f_{LO} , data dalla (4.1).

Notiamo che a differenza di f_{IF} che è fissa, f_0 e di conseguenza f_{LO} sono variabili e dipendono dalla "stazione radio" su cui desideriamo sintonizzarci.

All'uscita del *mixer* avremo due repliche, o immagini, del segnale filtrato $r_{RF}(t)$, una attorno la frequenza $f_0 + f_{LO} = 2f_0 + f_{IF}$ e l'altra attorno $f_0 - f_{LO} = -f_{IF}$. Per simmetria avremo anche immagini complesso coniugate in ampiezza attorno $-(f_0 + f_{LO})$ e $-(f_0 - f_{LO}) = f_{IF}$.

Successivamente il filtro IF farà passare solo le immagini attorno $\pm f_{IF}$ ed eliminerà quelle attorno $\pm(2f_0 + f_{IF})$. Naturalmente la banda passante del filtro RF deve essere specificata altrimenti il segnale desiderato all'uscita del filtro IF sarà effetto da interferenza. Notiamo infatti che se in $r_{RF}(t)$ è presente un segnale con portante $f_0 + 2f_{IF}$, dopo il *mixer* esso creerà una immagine attorno f_{IF} che andrà ad interferire il segnale desiderato. Come illustrato in Figura 4.2, se l'amplificatore RF ha una banda B_{RF} compresa tra B_s e $2f_{IF} - B_s$, attorno f_0 , non vengono create immagini dei segnali interferenti entro la banda desiderata attorno f_{IF} . Tipicamente si seleziona per il filtro RF una frequenza di taglio (in corrispondenza dell'inizio della banda attenuata) circa pari a $f_0 \pm f_{IF}$. Se Δf è la spaziatura tra le portanti dei canali vicino a quello desiderato è facile vedere che la banda di transizione del filtro IF può estendersi da $f_{IF} \pm \frac{B_s}{2}$ ad $f_{IF} \pm (\Delta f - \frac{B_s}{2})$.

Ci sono schemi in cui $f_{LO} < f_0$, per $f_{LO} = f_0 - f_{IF}$. All'uscita del *mixer* avremo un'immagine desiderata attorno $f_0 - f_{LO} = f_{IF}$ e una attorno $f_0 + f_{LO} = 2f_0 - f_{IF}$ che viene attenuata dal successivo filtro IF. In questo caso il segnale desiderato non risulta "invertito in frequenza". Anche in questo caso la banda del filtro RF può estendersi da B_s fino ad un valore massimo $2f_{IF} - B_s$, altrimenti ci possono essere sovrapposizioni in banda IF con altre stazioni trasmettenti.

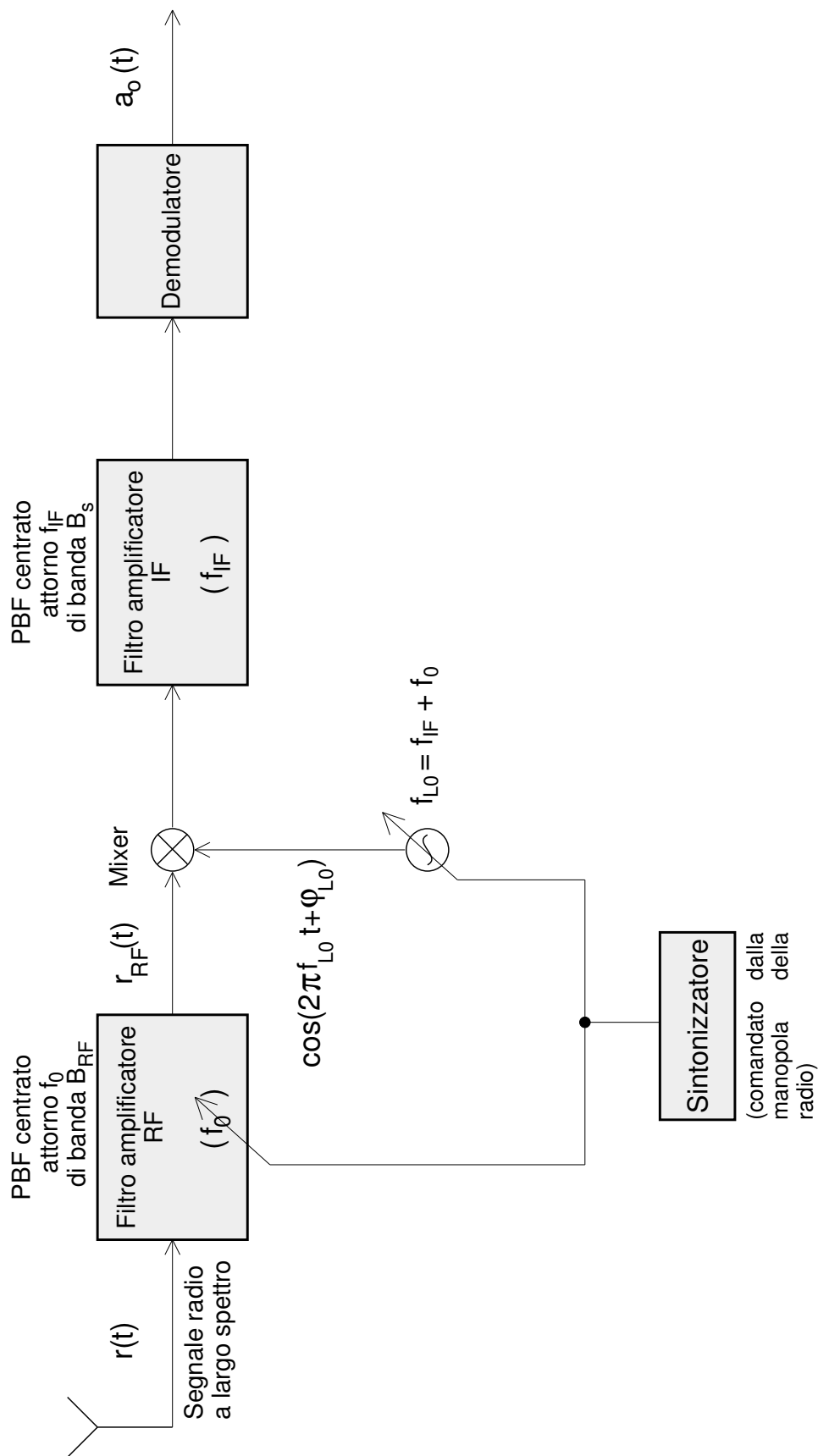


Figura 4.1: Ricevitore radio con *front – end* del tipo supereterodina.

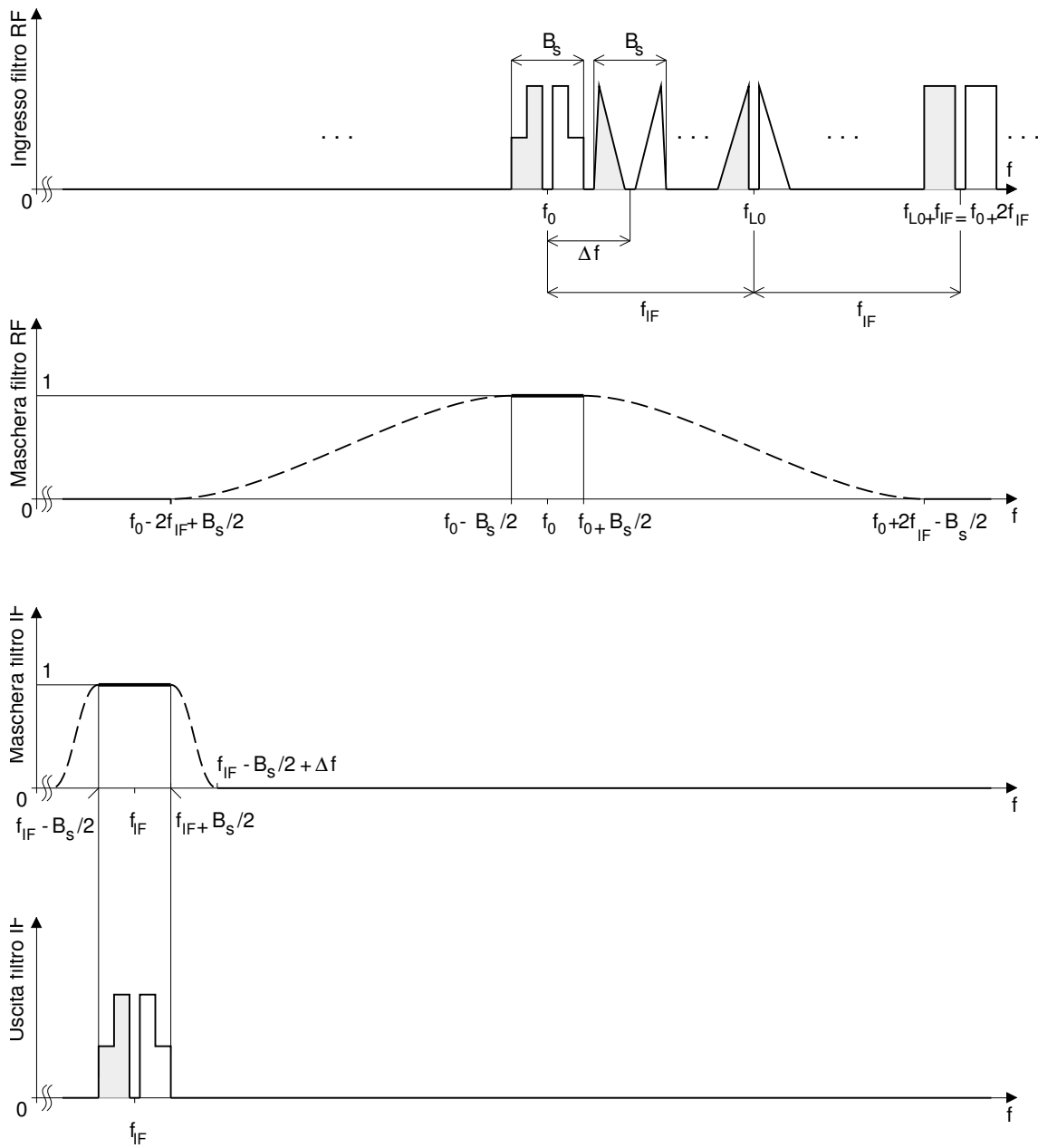


Figura 4.2: Illustrazione dei segnali nel dominio della frequenza in un ricevitore supereterodina.

4.1.2 Radio FM

Il sistema radio FM viene utilizzato per trasmettere voce e musica con una banda di 15 kHz . Il sistema utilizza la tecnica FDM e occupa lo spettro $88 \div 108\text{ MHz}$ con portanti separate di 200 kHz . La modulazione è FM con $\beta_F = 5$ ($\Delta F = 75\text{ kHz}$) e viene utilizzata la preenfasi. Il ricevitore, con *front - end* di tipo supereterodina è riportato in Figura 4.3 dove $f_{IF} = 10.7\text{ MHz}$

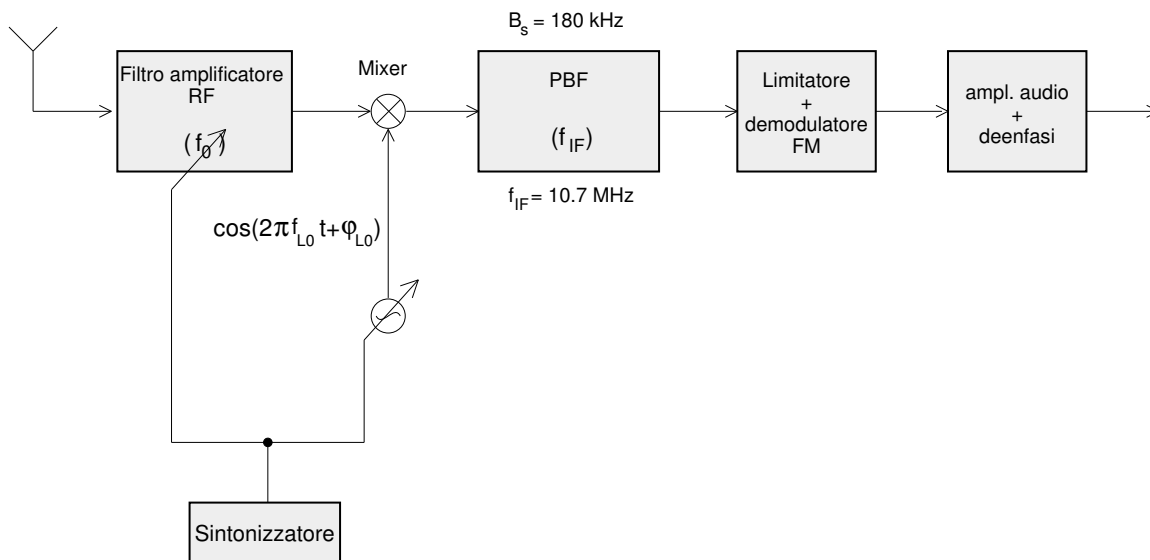


Figura 4.3: Ricevitore radio FM.

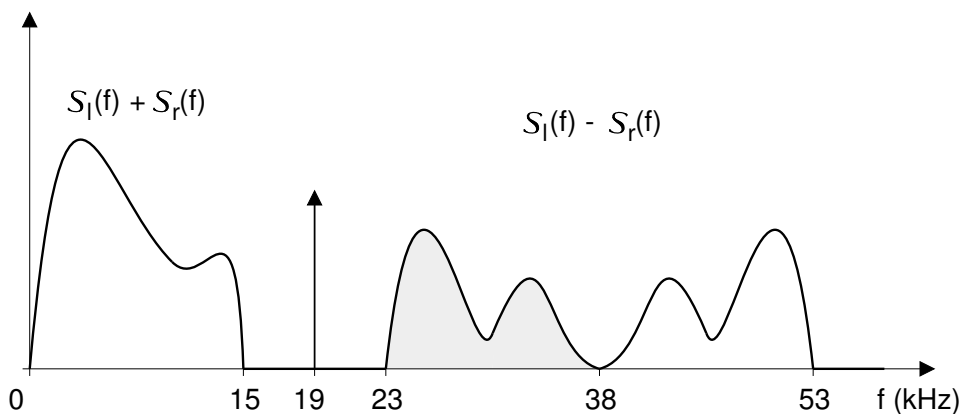


Figura 4.4: Spettro dei segnali all'ingresso del modulatore FM.

4.1.3 Radio FM stereo

A partire dai due segnali stereo $s_l(t)$ e $s_r(t)$, vengono formati i segnali somma $s_l(t) + s_r(t)$ e differenza $s_l(t) - s_r(t)$, i quali vengono moltiplicati in frequenza (FDM) utilizzando la tecnica DSB-SC per il segnale differenza (vedi Figura 4.4). In tal modo il sistema è

compatibile con il sistema tradizionale, non stereo, prelevando solo le componenti fino a 15 kHz . Come illustrato in Figura 4.5, il segnale composto dalla somma, segnale DSB-SC e la portante (utilizzata in ricezione per demodulare il segnale DSB) vengono sommati e modulati FM utilizzando una banda di 200 kHz ($\beta \simeq 0.9$).

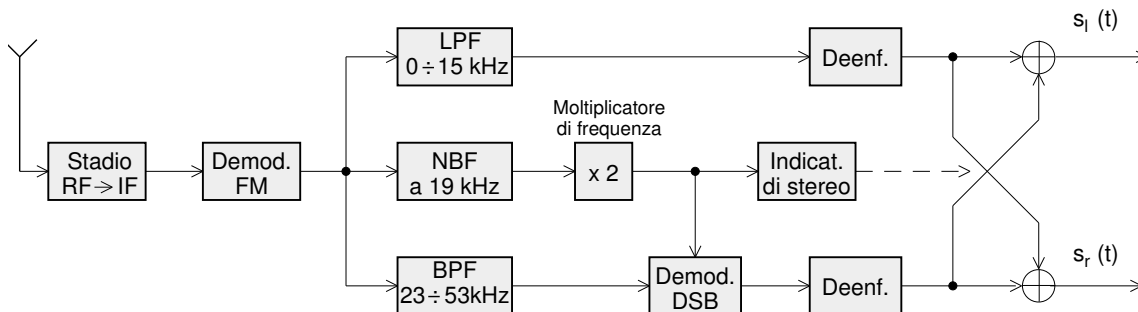
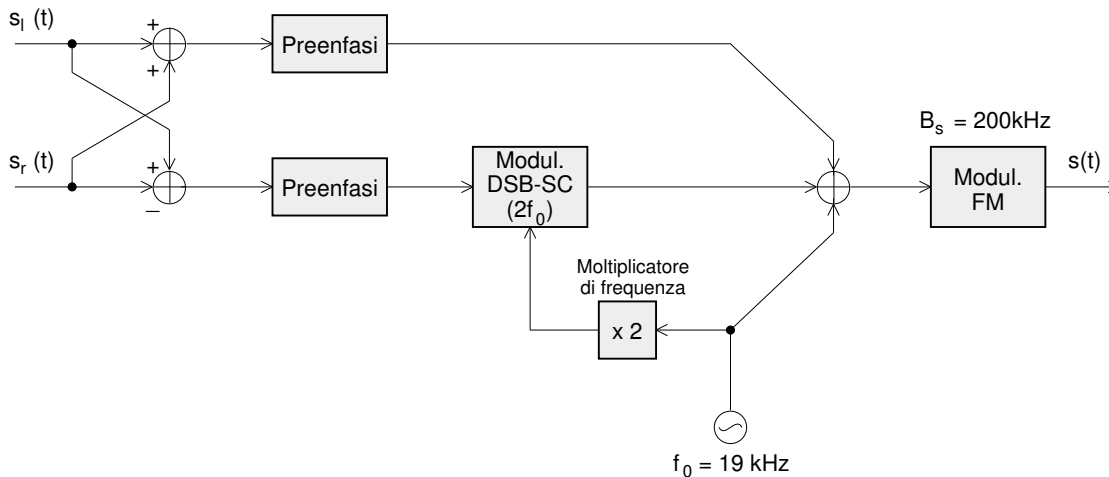


Figura 4.5: Trasmettitore e ricevitore per radio FM stereo.

In ricezione abbiamo le operazioni inverse.

4.1.4 Segnale televisivo

Un canale TV occupa una banda di 6 MHz mentre il segnale video commerciale ha una banda di 4.2 MHz . In effetti per motivi realizzativi lo spettro del segnale modulato è di tipo VSB-TC dopo il filtro IF in ricezione, mentre in trasmissione occupa una banda maggiore, secondo l'illustrazione di Figura 4.6. Il segnale trasmesso, con tecnica FDM, include in parte il segnale video con banda di 4.2 MHz e in parte l'audio, modulato

FM attorno 4.5 MHz . Per uno schema realizzativo del trasmettitore e ricevitore, rimandiamo a [4].

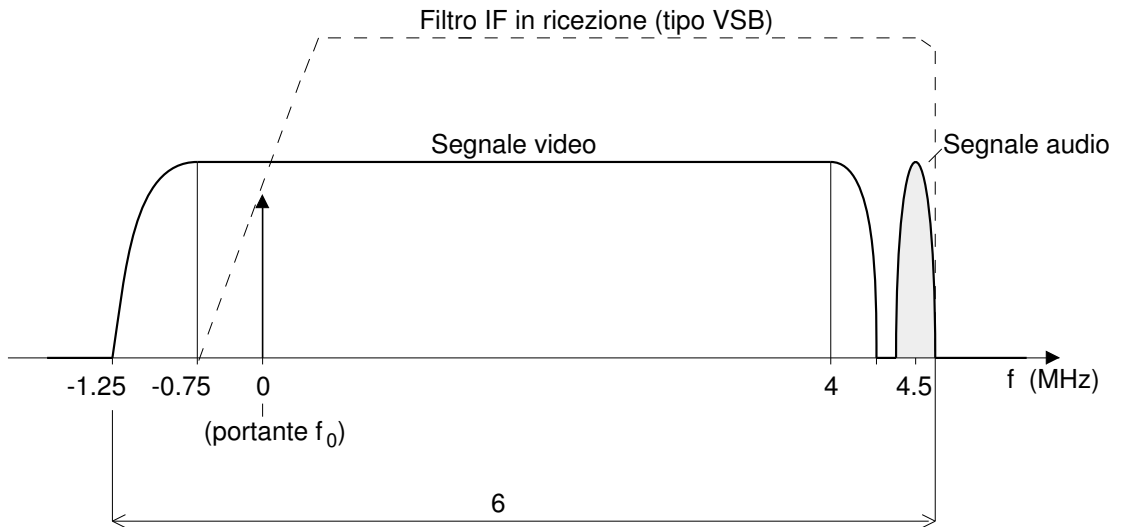


Figura 4.6: Occupazione i banda del segnale TV.

Parte II

Esercizi sui filtri numerici e FFT

Capitolo 5

Filtri numerici

In questa sezione, considereremo soprattutto alcuni esercizi relativi al filtraggio dei segnali. L'attenzione verrà posta al filtraggio dei segnali a tempo discreto, e si supporrà dunque di avere a che fare con segnali definiti su $Z(T)$.

La relazione ingresso-uscita di un filtro definito su $Z(T)$ con risposta impulsiva $h(t), t \in Z(T)$, risulta come noto data dall'operazione di convoluzione

$$y(nT) = T \sum_{k=-\infty}^{+\infty} h(kT)x(nT - kT) = T \sum_{k=-\infty}^{+\infty} x(kT)h(nT - kT). \quad (5.1)$$

Cominceremo con l'introdurre alcuni concetti fondamentali per la realizzazione e la comprensione dei filtri a tempo discreto. Il lettore potrà rivelarne l'analogia con concetti relativi ad altre discipline, in cui si considerano sistemi a tempo continuo.

5.1 Stabilità BIBO di un filtro numerico

Un filtro numerico si dice *stabile* in senso BIBO (Bounded Input-Bounded Output) se l'uscita ad un ingresso limitato è anch'essa limitata. Si ricorda che il segnale $x(nT)$ è limitato se esiste un numero finito M tale che

$$|x(t)| \leq M, \quad \text{per ogni } t \in Z(T). \quad (5.2)$$

Vengono ora analizzate le condizioni per cui la trasformazione lineare (5.1) risulta stabile in senso BIBO. Come è logico attendersi, le condizioni riguardano la risposta impulsiva $h(nT)$. Si ha infatti

$$\begin{aligned} |y(nT)| &\leq \sum_{k=-\infty}^{+\infty} T|h(kT)||x(nT - kT)| \\ &\leq M \sum_{k=-\infty}^{+\infty} T|h(kT)|, \end{aligned}$$

dove si è sfruttato il fatto che, per un ingresso limitato, vale la (5.2). Se dunque la somma dei moduli della risposta impulsiva del filtro è finita, così risulta l'uscita $y(nT)$. La relazione precedente stabilisce una condizione sufficiente per la stabilità BIBO. Tale condizione è anche necessaria: se infatti fosse

$$\sum_{k=-\infty}^{+\infty} T|h(kT)| = +\infty \quad (5.3)$$

e $h(kT)$ non fosse limitata, basterebbe prendere l'ingresso $x(nT) = \delta_{Z(T)}(nT)$ per ottenere un'uscita $y(nT) = h(nT)$ non limitata. Se infine valesse la (5.3) e $h(kT)$ fosse limitata, basterebbe prendere l'ingresso $x(nT - kT) = h^*(kT)/|h(kT)|$, per $|h(kT)| \neq 0$ e 0 altrove, per avere $y(nT) = +\infty$ all'istante nT . Possiamo dunque concludere con il seguente Teorema.

Teorema 1 *Un filtro sui tempi discreti caratterizzato dalla risposta impulsiva $h(t)$, $t \in Z(T)$ è stabile in senso BIBO se e solo se*

$$\sum_{k=-\infty}^{+\infty} T|h(kT)| = +\infty.$$

Dunque, un filtro risulta stabile in senso BIBO se la sua risposta impulsiva decade in maniera sufficientemente rapida, per nT che tende a più o meno infinito, in modo che valga la

$$\sum_{k=-\infty}^{+\infty} T|h(kT)| < +\infty$$

Si introduce la seguente nomenclatura: un filtro a tempo discreto si dice di tipo FIR (Finite Impulse Response) se la sua risposta impulsiva è di durata *finita*. Viceversa, esso si dice di tipo IIR (Infinite Impulse Response) se la sua risposta impulsiva ha durata non limitata.

5.2 Risposta in frequenza razionale

Una classe molto importante di filtri a tempo discreto è la classe dei filtri causali a risposta in frequenza razionale, tali cioè che la trasformata di Fourier $H(f)$, $f \in R/Z(T)$, della risposta impulsiva $h(t)$, $t \in Z(T)$, assume la forma

$$H(f) = \frac{T \sum_{k=0}^r b_k e^{-j2\pi f k T}}{\sum_{k=0}^q a_k e^{-j2\pi f k T}}, \quad (5.4)$$

dove si può porre, senza perdita di generalità, $a_0 = 1$. Nel caso infatti fosse $a_0 \neq 1$, basterebbe dividere numeratore e denominatore per a_0 . La risposta in frequenza $H(f)$ risulta, come è noto e come si vede immediatamente dalla (5.4), una funzione periodica in f di periodo $F_s = 1/T$.

Il termine *razionale* risulta più comprensibile se, anziché considerare la trasformata di Fourier (5.4), ci riferiamo alla trasformata zeta della risposta impulsiva, detta anche *funzione di trasferimento del filtro*,

$$H_z(z) \triangleq T \sum_{n=-\infty}^{+\infty} h(nT) z^{-n},$$

la quale risulta un rapporto fra polinomi in z^{-1} . Infatti, ricordando che la trasformata di Fourier si ottiene dalla trasformata zeta ponendo $z = \exp(j2\pi f T)$, nell'ipotesi di esistenza di entrambe, dalla (5.4) si ottiene

$$H_z(z) = \frac{T \sum_{k=0}^r b_k z^{-k}}{\sum_{k=0}^q a_k z^{-k}}. \quad (5.5)$$

Vedremo come la classe dei filtri a risposta in frequenza razionale possa essere realizzata mediante algoritmi numerici estremamente semplici.

Consideriamo dapprima alcuni semplici esempi.

Esempio 1. La sequenza esponenziale.

Consideriamo il filtro a tempo discreto con risposta impulsiva *causale* $h(nT) = p^n \mathbf{1}_0(nT)$. Il filtro è stabile in senso BIBO se risulta $|p| < 1$. In tali ipotesi, la risposta in frequenza è una funzione razionale e risulta

$$H(f) = \frac{T}{1 - pe^{-j2\pi fT}}, \quad f \in R/Z(1/T).$$

La trasformata zeta vale

$$H(z) = \frac{T}{1 - pz^{-1}} \quad (5.6)$$

e presenta un polo in $z = p$ all'interno del cerchio unitario del piano complesso e uno zero nell'origine. La regione di convergenza della trasformata zeta è la regione $|z| > |p|$.

Consideriamo invece un filtro con risposta in frequenza

$$H(f) = \frac{T}{1 - pe^{-j2\pi fT}}, \quad f \in R/Z(1/T),$$

dove però sia $|p| > 1$. In questo caso, l'antitrasformata risulta, come è immediato verificare, un segnale *anti-causale* ed uguale a

$$h(nT) = -p^n \mathbf{1}_0(-nT - T).$$

Si tratta anche in questo caso di un filtro stabile in senso BIBO. La trasformata zeta vale

$$H(z) = \frac{T}{1 - pz^{-1}}$$

ma questa volta la regione di convergenza risulta $|z| < |p|$.

Infine, un filtro con risposta impulsiva causale $h(nT) = p^n \mathbf{1}_0(nT)$, con $|p| \geq 1$, non è ovviamente stabile in senso BIBO. La sua trasformata zeta assume ancora la forma (5.6) e converge per $|z| > |p|$, mentre la sua trasformata di Fourier non esiste (almeno in senso ordinario).

Un facile esercizio, risolvibile applicando più volte la regola di derivazione della trasformata di Fourier, permetterebbe di stabilire che la sequenza causale

$$h(nT) = \frac{(n+k-1)!}{n!(k-1)!} p^n \mathbf{1}_0(nT), \quad nT \in Z(T), \quad (5.7)$$

stabile in senso BIBO per $|p| < 1$, ha trasformata di Fourier

$$H(f) = \frac{T}{(1 - pe^{-j2\pi fT})^k}, \quad f \in R/Z(T) \quad (5.8)$$

e trasformata zeta, convergente per $|z| > |p|$,

$$H_z(z) = \frac{T}{(1 - pz^{-1})^k}. \quad (5.9)$$

In conclusione, le (5.7)-(5.9), valide per $k \geq 1$, forniscono il legame segnale-trasformata relativo a sequenze *causali* e *stabili* con andamento esponenziale nel dominio del tempo e forma razionale nel dominio della trasformata. In particolare, dato che ci limitiamo a considerare

sequenze causali, la trasformata zeta (5.9) ha un polo, di molteplicità k , all'interno del cerchio unitario e uno zero di molteplicità k nell'origine. \square

Consideriamo ora il caso della generica trasformata zeta razionale (5.5). Il calcolo della antitrasformata è banale se $q = 0$. In tale caso, infatti, $H_z(z)$ si riduce ad un polinomio in z^{-1} e i coefficienti della risposta impulsiva risultano uguali a $h(nT) = b_n$ per $n = 0, \dots, r$, e 0 altrove. La risposta impulsiva è dunque, in questo caso, di tipo FIR.

Nel caso $q \neq 0$, il modo di procedere è quello di sviluppare la (5.5) in *frazioni parziali*: possiamo scrivere

$$\begin{aligned} H_z(z) &= T \frac{b_0 + b_1 z^{-1} + \dots + b_r z^{-r}}{1 + a_1 z^{-1} + \dots + a_m z^{-q}} \\ &= T \frac{b_0 + b_1 z^{-1} + \dots + b_r z^{-r}}{(1 - p_1 z^{-1})^{m_1} \dots (1 - p_s z^{-1})^{m_s}}, \end{aligned} \quad (5.10)$$

dove p_i è il polo i -esimo della trasformata zeta, e m_i la rispettiva molteplicità, con $m_1 + \dots + m_s = q$. Supponendo $r < q$ nella espressione della trasformata zeta¹, la (5.10) può essere scritta nella forma

$$H_z(z) = \sum_{j=1}^{m_1} \frac{A_{1j}}{(1 - p_1 z^{-1})^j} + \dots + \sum_{j=1}^{m_s} \frac{A_{sj}}{(1 - p_s z^{-1})^j}, \quad (5.11)$$

dove, come è noto,

$$A_{ij} = \frac{1}{(m_i - j)!} \frac{1}{(-p_i)^{m_i - j}} \frac{d^{m_i - j}}{d(z^{-1})^{m_i - j}} [H_z(z)(1 - p_i z^{-1})^{m_i}]_{z=p_i}. \quad (5.12)$$

Si noti che la derivata viene fatta rispetto alla variabile $w = z^{-1}$. Si noti inoltre che nel caso di un polo p_i con molteplicità unitaria, il coefficiente A_{i1} si calcola semplicemente dalla (5.12) valutando la funzione $H_z(z)(1 - p_i z^{-1})$ per $z = p_i$.

Possiamo calcolare l'antitrasformata di $H_z(z)$ usando i risultati dell'esempio precedente. Se ne conclude che, *se tutti i poli p_i sono in modulo minori di 1*, l'antitrasformata della $H_z(z)$ corrisponde alla combinazione lineare di sequenze con andamento esponenziale della forma (5.7). Nel caso fosse $r \geq q$, alla combinazione di esponenziali si aggiungerebbe nella risposta impulsiva un contributo di durata finita, corrispondente all'antitrasformata di un polinomio in z^{-1} .

Esempio 2.

Si consideri un filtro causale con funzione di trasferimento

$$H_z(z) = T \frac{1.5 + 0.5z^{-1} + 0.2z^{-2}}{1 - 0.7z^{-1} + 0.1z^{-2}}. \quad (5.13)$$

Si ricava facilmente che i poli di $H_z(z)$ valgono $z = 0.5$ e $z = 0.2$ e sono quindi interni al cerchio di raggio unitario. Se ne deduce che $H_z(z)$ è la funzione di trasferimento di un filtro

¹Se così non fosse, basterebbe prendere il resto della divisione del numeratore per il denominatore, scomponendo $H_z(z)$ nella somma di un polinomio in z^{-1} e di una funzione razionale della forma voluta.

causale e stabile. Applicando la scomposizione in frazioni parziali, dopo avere calcolato il resto della divisione fra numeratore e denominatore, si trova

$$H_z(z) = T \left[2 + \frac{5.5}{1 - 0.5z^{-1}} - \frac{6}{1 - 0.2z^{-1}} \right].$$

La risposta impulsiva del filtro risulta pertanto di tipo IIR

$$h(nT) = 2T\delta_{Z(T)}(nT) + 5.5 \cdot 0.5^n 1_0(nT) - 6 \cdot 0.2^n 1_0(nT).$$

□

5.3 Equazioni alle differenze

Si consideri un filtro con una risposta in frequenza razionale, come nell'equazione (5.4). La relazione ingresso-uscita del filtro, espressa nel dominio della frequenza, risulta

$$Y(f) = H(f)X(f), \quad f \in R/Z(T), \quad (5.14)$$

dove $Y(f)$ e $X(f)$ sono la trasformata di Fourier dell'uscita e dell'ingresso, rispettivamente. Moltiplicando entrambi i membri della (5.14) per il denominatore di $H(f)$ e ricordando la regola di traslazione per la trasformata di Fourier, si ottiene nel dominio del tempo la seguente relazione ricorsiva fra $y(nT)$ e $x(nT)$

$$y(nT) = - \sum_{k=1}^q a_k y((n-k)T) + T \sum_{k=0}^r b_k x((n-k)T). \quad (5.15)$$

L'equazione *lineare alle differenze finite* (5.15), corrispondente ad una funzione di trasferimento razionale, può facilmente essere realizzata mediante un algoritmo numerico, come esemplificato nello schema di Fig. 5.1.

La sua struttura suggerisce anche la forma normalizzata

$$y(nT) = - \sum_{k=1}^q a_k y((n-k)T) + \sum_{k=0}^r b'_k x((n-k)T),$$

in cui vengono definiti i coefficienti $b'_k = T b_k$.

Esempio 3. La funzione di trasferimento (5.13) corrisponde alla seguente equazione alle differenze

$$y(nT) = 0.7y((n-1)T) - 0.1y((n-2)T) + 1.5Tx(nT) + 0.5Tx((n-1)T) + 0.2Tx((n-2)T). \quad (5.16)$$

Come si vede, l'uscita al tempo nT dipende dai valori precedenti dell'ingresso e dell'uscita. Se però l'ingresso $x(nT)$ è nullo per $nT < n_0T$, anche l'uscita $y(nT)$ risulta nulla per $nT < n_0T$, dato che il filtro è causale e risulta $h(nT) = 0$ per $nT < 0$. In tali ipotesi, il sistema (5.16) evolve a partire da *condizioni iniziali* nulle a partire dall'istante nT_0 .

È possibile scrivere il seguente programma MATLAB per il calcolo dell'uscita da $n_0T = 0$ a $nT = 100T$ corrispondente all'ingresso $x(nT) = \sin(2\pi f_0 nT) 1_0(nT)$, con $T = 1$, $f_0 = 0.1T$.

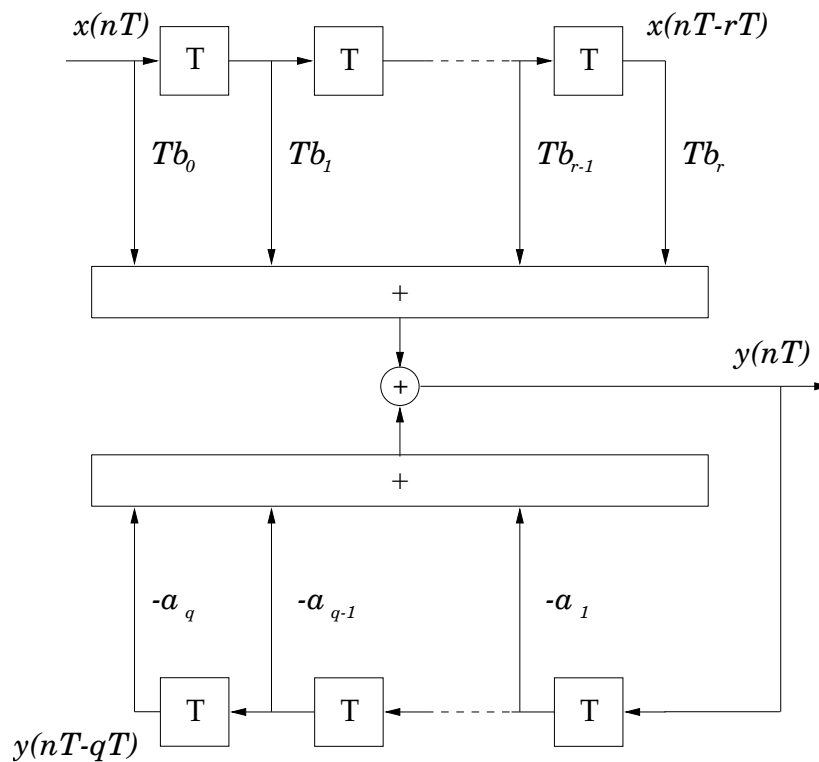


Figura 5.1: Schema per la realizzazione dell'equazione a differenze finite.

```

% condizioni iniziali -caricamento dello "stato" del filtro

xs=[0,0];
ys=[0,0];
r1=length(xs);
q1=length(ys);

% istante iniziale
n0=0;

n=n0;
T=1;

while (n>=n0 & n<=100),

% nuovo ingresso

    xnT=sin(2*pi*0.1*n);

% calcolo dell'uscita: l'indice n+1 e' relativo al campione y(nT)

    y(n+1)=[0.7 -0.1]*ys' + [1.5 0.5 0.2]*[xnT xs]';

% aggiornamento dello stato

    ys=[y(n+1),    ys(1:q1-1)];
    xs=[xnT,      xs(1:r1-1)];

```



```
n=n+1;
end;
```

□

Nell'esempio precedente abbiamo visto come possiamo utilizzare la (5.15) per il calcolo dell'uscita di un filtro causale con funzione di trasferimento razionale quando l'ingresso è nullo prima di un certo istante. Come possiamo calcolare l'uscita nel caso in cui l'ingresso sia diverso da zero, in generale, da meno infinito a più infinito? Ovviamente, saremo interessati dal punto di vista operativo a conoscere l'uscita a partire da un certo istante t di osservazione. Per poter utilizzare la (5.15) dovremmo però conoscere le condizioni iniziali, specificate anche dai valori precedenti dell'uscita, che purtroppo in generale non conosciamo dato che stiamo cercando di calcolarla.

La soluzione del problema si ottiene in maniera semplice se il segnale di ingresso $x(t), t \in Z(T)$, ha una durata convenzionale *praticamente* limitata. In questo caso, infatti, è sufficiente pensare che il segnale sia nullo al di fuori di un certo intervallo ed usare la (5.15) per calcolare l'uscita del filtro con la precisione voluta.

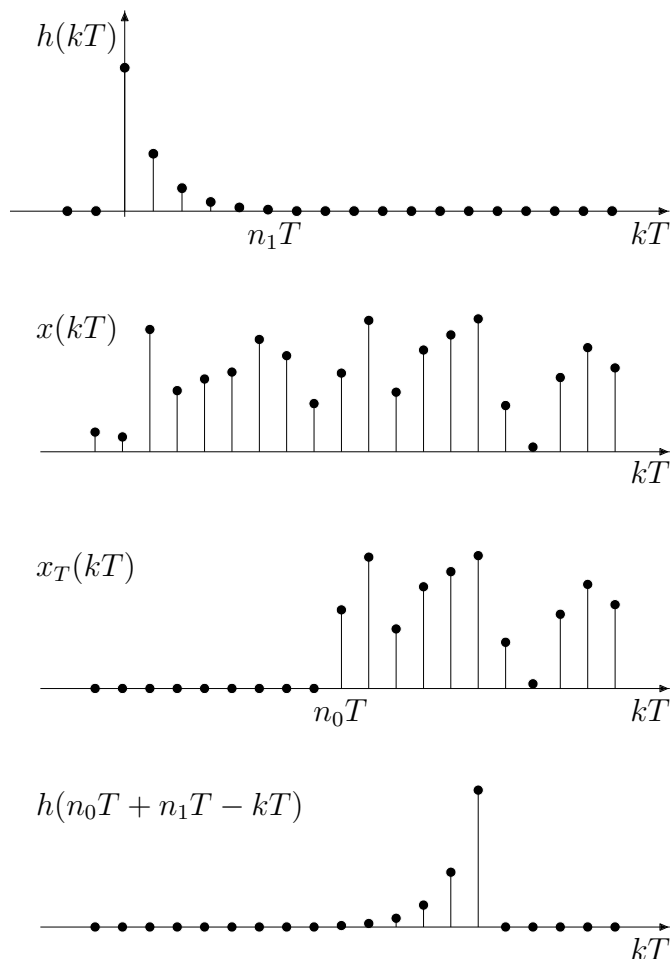


Figura 5.2: Nel calcolo della convoluzione, all'istante $n_0T + n_1T$ le uscite corrispondenti a $x(t)$ e $x_T(t)$ coincidono (praticamente).

Nel caso invece il segnale abbia un andamento persistente, come ad esempio avviene per i segnali sinusoidali, il segnale a gradino, o anche, in generale, per le realizzazioni dei

processi aleatori, è utile fare la seguente osservazione. Come abbiamo visto, la risposta impulsiva $h(nT)$ di un filtro causale stabile con risposta in frequenza razionale ha un andamento dato dalla somma di una sequenza di durata finita e di un certo numero di sequenze esponenziali. Nei limiti di una precisione prefissata, tale risposta impulsiva può considerarsi in ogni caso trascurabile a partire da un certo istante $t = n_1T$ che dipende dalla velocità di decadimento degli esponenziali, e quindi dal modulo dei poli della funzione di trasferimento. Più i poli sono piccoli in modulo, più velocemente si esaurisce la risposta impulsiva. Di conseguenza, se approssimiamo l'ingresso $x(nT)$ con una sua versione troncata $x_T(nT) = x(nT)1_0(nT - n_0T)$, le uscite $y(nT) = x * h(nT)$ e $y_T(nT) = x_T * h(nT)$ sono praticamente uguali per $t \geq n_0T + n_1T$ (vedi Fig. 5.2). Il calcolo di $y_T(nT)$ può dunque essere effettuato utilizzando l'equazione alle differenze (5.15) a partire da condizioni iniziali nulle: dopo un numero di campioni di uscita pari a n_1T , ovvero dopo l'esaurimento del *transitorio* del filtro, i campioni di $y_T(nT)$ coincidono praticamente con quelli che si sarebbero ottenuti filtrando l'ingresso originario. Si dice in questo caso che, dopo l'esaurimento del transitorio, l'uscita del filtro è in *regime permanente*. Si noti l'analogia con la nomenclatura ed i concetti usati, anche in altre discipline, per i segnali a tempo continuo.

5.4 L'uso di MATLAB per il progetto dei filtri

Abbiamo visto nella sezione precedente come sia possibile realizzare tramite un algoritmo numerico un filtro a tempo discreto con risposta in frequenza razionale. Esistono in realtà dei criteri di organizzazione dei calcoli che rendono la realizzazione del filtro più robusta, ad esempio qualora sia necessario utilizzare processori che utilizzano un'aritmetica a precisione finita. Il programma MATLAB fornisce con il *Signal Processing Toolbox* un insieme di procedure per il filtraggio a tempo discreto. In particolare, la funzione `filter(b,a,x)`, supponendo uguale ad 1 il primo elemento `a(1)` del vettore `a`, filtra i dati contenuti nel vettore `x` di ingresso realizzando l'equazione alle differenze

$$y(n) = b(1)*x(n) + b(2)*x(n-1) + \dots + b(nb+1)*x(n-nb) - a(2)*y(n-1) - \dots - a(na+1)*y(n-na).$$

Si noti che i coefficienti del vettore `b(i)` devono essere posti uguali a Tb_{i-1} , secondo quanto richiesto nella (5.15)².

La funzione `[H,f]=freqz(b,a,N,Fs)` calcola N campioni complessi equispaziati fra 0 e $F_s/2$, con $F_s = 1/T$, della trasformata di Fourier corrispondente alla funzione di trasferimento razionale

$$H_z(z) = \frac{b(1) + b(2)z^{-1} + \dots + b(nb+1)z^{-nb}}{1 + a(2)z^{-1} + \dots + a(na+1)z^{-na}}.$$

In particolare, gli N valori della risposta in frequenza vengono posti nel vettore `H`, mentre i rispettivi valori della frequenza, in Hz, vengono posti nel vettore `f`. Si noti che la trasformata di Fourier $H(f) = H_z(e^{j2\pi fT})$ ha simmetria hermitiana, e quindi ci si può limitare a specificarla tra 0 e $F_s/2$.

Per l'analisi di una funzione di trasferimento, sono utili anche le procedure `residue` e `roots` che calcolano, rispettivamente, la scomposizione in frazioni parziali di un rapporto di polinomi e le radici di un polinomio.

²Gli algoritmi di MATLAB assumono che $T=1$. Ci si può facilmente ricondurre al caso generale supponendo di riferirci ai coefficienti normalizzati Tb_k .

Un argomento di fondamentale importanza, cui qui accenneremo soltanto parzialmente, riguarda il *progetto* di filtri numerici. Il problema che qui ci si pone è il seguente: assegnata una determinata risposta in frequenza del filtro, come è possibile determinare i coefficienti b_i e a_i di una risposta in frequenza razionale che approssimi l'andamento voluto? Esistono a questo proposito molte tecniche di progetto, relative sia a filtri FIR ($q = 0$ nella (5.5)), che a filtri IIR.

Tipicamente, i filtri che si desidera realizzare hanno un andamento della risposta in frequenza che approssima uno degli andamenti ideali di Fig. 5.3, relativi a filtri *passa-basso*, *passa-alto*, *passa-banda* e *elimina-banda*, rispettivamente. Nella Fig. 5.3, le frequenze sono normalizzate rispetto alla frequenza di campionamento F_s . Trattandosi di filtri reali, è sufficiente rappresentare l'andamento della risposta in frequenza fra $f = 0$ e $f = F_s/2$. Come vedremo, i vari metodi di progetto si limitano molto spesso a specificare solamente il *modulo* della risposta in frequenza desiderata.

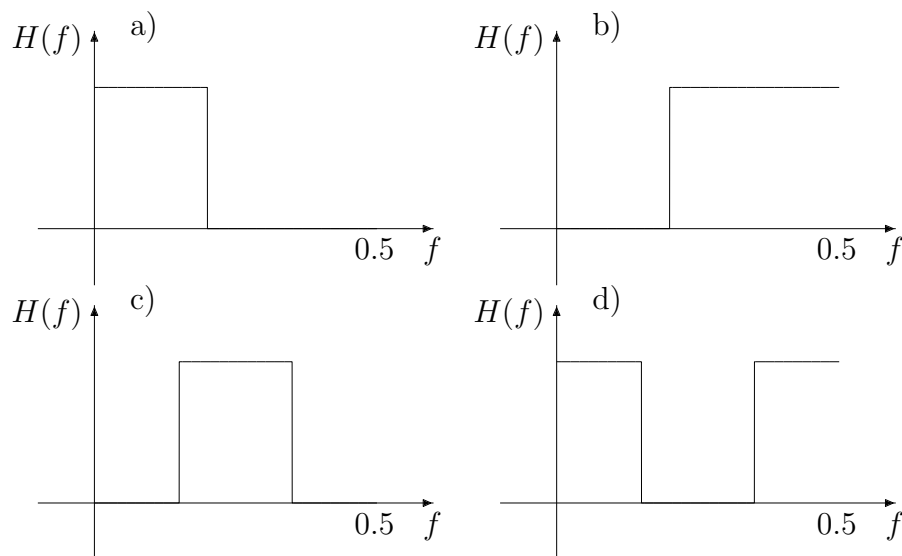


Figura 5.3: Andamento dei filtri ideali a) passa-basso, b) passa-alto, c) passa-banda e d) elimina-banda

Trattandosi di un problema di approssimazione, occorrerà specificare nel progetto le tolleranze ammesse nella realizzazione del filtro. A questo proposito, si consideri la Fig. 5.4. In essa, viene rappresentato il modulo della risposta in frequenza (l'asse delle ascisse è parametrizzato rispetto a f/F_s) di un tipico filtro numerico che approssima un filtro passa-basso ideale. I parametri da considerare sono:

1. la *banda passante*, specificata dalla frequenza W_p . Si richiede che il filtro abbia modulo della risposta in frequenza circa uguale ad 1 per $0 \leq f \leq W_p$. Per un filtro passabanda, andrà in genere specificato un intervallo di frequenze entro il quale si desidera che il filtro abbia modulo della trasformata di Fourier unitario;
2. la *banda oscura*, specificata in questo caso dalla frequenza W_s . Si richiede che il filtro abbia modulo della risposta in frequenza circa uguale a 0 per $W_p \leq f \leq F_s/2$. L'intervallo $[W_p, W_s]$ specifica la cosiddetta *banda di transizione* del filtro;
3. il *ripple* in banda passante, quantificato tramite il valore R_p (spesso espresso in dB) che permette di determinare l'errore rispetto all'andamento desiderato;

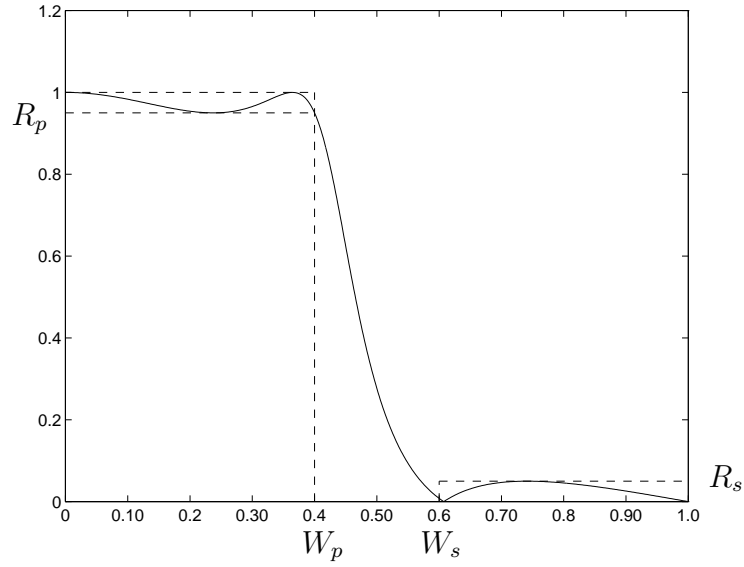


Figura 5.4: Parametri di progetto di un filtro passa-basso.

4. il *ripple* in banda attenuata, specificato tramite il valore massimo R_s del modulo della risposta in frequenza del filtro in banda attenuata.

Progetto di filtri IIR in MATLAB. Il programma MATLAB fornisce varie procedure per il progetto di filtri IIR ($q > 0$ nella (5.5)). Considereremo in particolare il caso dei filtri *ellittici*, i quali forniscono una approssimazione uniforme della risposta in frequenza sia in banda passante che in banda oscura. Si può dimostrare che i filtri ellittici sono ottimi, nel senso che, fissati i valori W_p , R_p e R_s , la banda di transizione $W_s - W_p$ è la minima possibile per un dato grado q del denominatore del filtro (q viene detto *l'ordine* del filtro).

Il progetto dei filtri ellittici passa-basso in MATLAB viene effettuato utilizzando le due procedure `ellipord` e `ellip` del Signal Processing Toolbox. La sintassi di `ellipord` è la seguente:

```
[N, Wn] = ellipord(Wp, Ws, RpdB, RsdB).
```

In ingresso alla funzione vanno specificate le frequenze W_p e W_s (vedi Fig. 5.4), normalizzate rispetto a $F_s/2$. Inoltre, occorre specificare i valori di R_p e R_s , in dB. L'uscita della funzione fornisce il valore N dell'ordine q del filtro che soddisfa le specifiche, e il valore della frequenza naturale Wn da usare per l'effettivo progetto del filtro, effettuato con la procedura

```
[b,a] = ellip(N, RpdB, RsdB, Wn).
```

Nei vettori \mathbf{b} ed \mathbf{a} , si trovano i coefficienti del numeratore e denominatore della risposta in frequenza razionale (5.5) che soddisfa le specifiche. Si noti che i valori nel vettore \mathbf{b} includono la moltiplicazione per il quanto temporale T . Il metodo di progetto con filtri ellittici non permette di specificare la caratteristica di *fase* del filtro.

Esempio 4. Supponiamo che la frequenza di campionamento del sistema numerico sia $F_s = 8$ kHz, e si voglia progettare un filtro passa-basso con banda passante da 0 a 1.6 kHz e banda oscura da 2.4 kHz a 4 kHz= $F_s/2$. Si supponga di imporre $R_p = 0.95$ e $R_s = 0.05$. Il progetto e la visualizzazione della risposta in frequenza del filtro, mostrata nelle figure 5.5 e 5.6, viene effettuata con il seguente programma MATLAB.

```
% Parametri di progetto

Fs=8000;
RpdB=-20*log10(0.95);
RsdB=-20*log10(0.05);
Wp=1600/Fs*2;
Ws=2400/Fs*2;

% Progetto del filtro

[N,Wn]=ellipord(Wp, Ws, RpdB, RsdB);
[b,a]=ellip(N, RpdB, RsdB, Wn);

% Visualizzazione del filtro

[H,f]=freqz(b,a,512,Fs);
plot(f,abs(H));
plot(f,angle(H));
```

Si noti dalla Fig. 5.5 che il filtro soddisfa effettivamente le specifiche. Nella Fig. 5.6, si noti la discontinuità di π attorno alla frequenza 2.5 kHz, dovuta al cambiamento di segno di

$H(f)$ (corrispondente al punto angoloso nell'andamento del modulo). La fase è rappresentata modulo 2π . L'ordine del filtro è risultato $q = 3$, e quindi i vettori \mathbf{b} e \mathbf{a} hanno dimensione 4. Il filtro può essere utilizzato per l'elaborazione dell'ingresso posto nel vettore \mathbf{x} mediante il comando `filter(b,a,x)`. \square

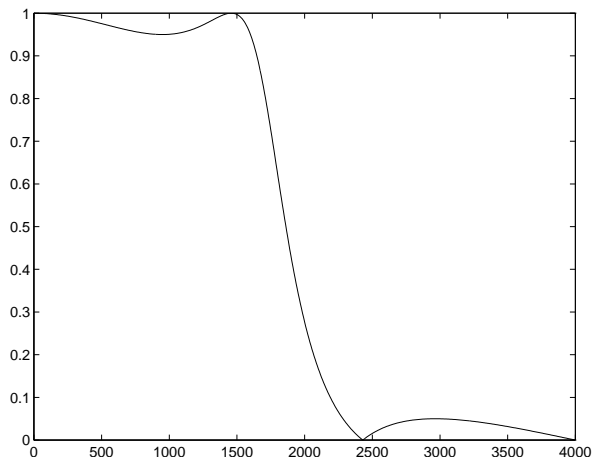


Figura 5.5: Modulo della risposta in frequenza del filtro ellittico (f in Hz).

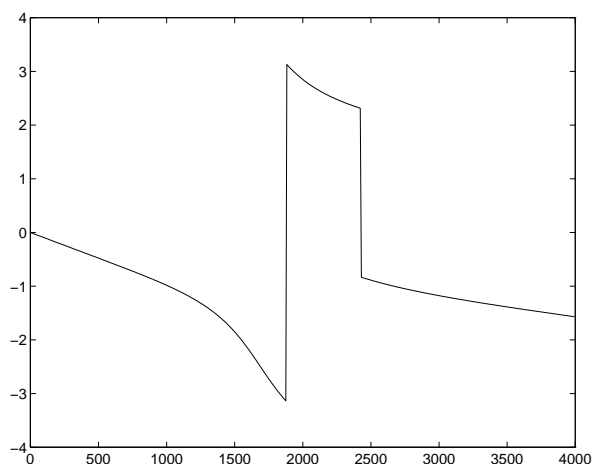


Figura 5.6: Fase (radianti) della risposta in frequenza del filtro ellittico (f in Hz)

Le funzioni `ellipord` e `ellip` possono essere utilizzate per il progetto di filtri passa-alto, passa-banda e elimina-banda. Si considerino in particolare i seguenti esempi esplicativi.

Esempio 5. Progetto di un filtro passa-alto ($F_s = 8$ kHz) con banda oscura $[0,1.5]$ kHz e banda passante $[2.5,4]$ kHz, $R_p = 0.99$ e $R_s = 0.01$.

`% Parametri di progetto`

```
Fs=8000;
RpdB=-20*log10(0.99);
RsdB=-20*log10(0.01);
Wp=2500/Fs*2;
Ws=1500/Fs*2;
```

```
% Progetto del filtro
```

```
[N,Wn]=ellipord(Wp, Ws, RpdB, RsdB);
[b,a]=ellip(N, RpdB, RsdB, Wn, 'high');
```

Si noti che $W_p > W_s$ e l'uso del parametro 'high' nella chiamata a `ellip`. L'ordine del filtro è risultato $q = 4$. □

Esempio 6. Progetto di un filtro passa-banda ($F_s = 8$ kHz) con banda oscura negli intervalli $[0,1]$ kHz e $[3,4]$ kHz e banda passante $[1.5,2.5]$ kHz, $R_p = 0.95$ e $R_s = 0.01$.

```
% Parametri di progetto
```

```
Fs=8000;
RpdB=-20*log10(0.95);
RsdB=-20*log10(0.01);
Wp=[1500/Fs*2, 2500/Fs*2];
Ws=[1000/Fs*2, 3000/Fs*2];
```

```
% Progetto del filtro
```

```
[N,Wn]=ellipord(Wp, Ws, RpdB, RsdB);
[b,a]=ellip(N, RpdB, RsdB, Wn);
```

Si noti che W_p e W_s sono dei vettori, con componenti le frequenze che delimitano la banda passante e le bande oscure. L'ordine del filtro risulta pari a due volte il valore N ritornato da `ellipord`. In questo caso, è risultato $N=4$ e $q = 8$. □

Esempio 7. Progetto di un filtro stop-banda ($F_s = 8$ kHz) con banda oscura $[1.5,2.5]$ kHz e banda passante negli intervalli $[0,1]$ kHz e $[3,4]$ kHz, $R_s = 0.95$ e $R_p = 0.01$.

```
% Parametri di progetto
```

```
Fs=8000;
RpdB=-20*log10(0.95);
RsdB=-20*log10(0.01);
Ws=[1000/Fs*2, 3000/Fs*2];
Wp=[1500/Fs*2, 2500/Fs*2];
```

```
% Progetto del filtro
```

```
[N,Wn]=ellipord(Wp, Ws, RpdB, RsdB);
[b,a]=ellip(N, RpdB, RsdB, Wn, 'stop');
```

Si noti che W_p e W_s sono dei vettori, con componenti le frequenze che delimitano le bande passanti e la banda oscura. Si noti inoltre l'uso del parametro 'stop' in ingresso alla procedura `ellip`. L'ordine del filtro risulta anche in questo caso pari a due volte il valore N ritornato da `ellipord`. È risultato in questo esempio $N=4$ e $q = 8$. □

Altre procedure per il progetto di filtri IIR disponibili in MATLAB sono `butter`, `cheby1`, `cheby2`, `besself`, `invfreqz`. In particolare, `invfreqz` permette il progetto di filtri di cui si specifica sia il modulo che la fase.

Progetto di filtri FIR in MATLAB. I filtri FIR hanno alcuni svantaggi rispetto ai filtri IIR ma anche alcuni importanti vantaggi. In generale, l'ordine di un filtro FIR (ovvero il parametro r nella (5.5)) risulta in generale molto più grande dell'ordine di un filtro IIR che soddisfa le stesse specifiche. D'altro canto, la durata del transitorio di un filtro FIR è finita e pari alla durata della risposta impulsiva del filtro. Inoltre, esistono metodi di progetto efficienti che garantiscono che il filtro FIR abbia esattamente *fase lineare*. Come è noto, questo garantisce che una sinusoide in ingresso al filtro subisce in uscita un ritardo indipendente dalla frequenza. Se dunque la risposta in frequenza del filtro ha l'espressione

$$H(f) = e^{-j2\pi ft_0} H_0(f),$$

dove $H_0(f)$ vale 1 nella banda passante, e 0 altrove³ e il segnale di ingresso ha un'estensione spettrale contenuta nella banda passante, l'uscita ha trasformata di Fourier

$$Y(f) = H(f)X(f) = e^{-j2\pi ft_0} X(f).$$

Nel dominio del tempo, risulta $y(t) = x(t - t_0)$ e l'uscita è una versione non distorta (secondo Heaviside) del segnale di ingresso.

Un filtro FIR causale di durata NT a fase lineare (a tratti) soddisfa la condizione

$$h(nT) = h((N - 1 - n)T), \quad n = 0, \dots, (N - 1)T.$$

La risposta impulsiva è dunque *simmetrica*, come evidenziato nelle Fig. 5.7a e 5.7b nel caso di N pari e N dispari, rispettivamente. In questo caso, la risposta in frequenza del filtro

$$H(f) = T \sum_{n=0}^{N-1} h(nT) e^{-j2\pi fnT}$$

risulta, come è immediato verificare

$$H(f) = \begin{cases} T e^{-j2\pi fT \frac{N-1}{2}} \left[h\left(\frac{NT-T}{2}\right) + 2 \sum_{n=0}^{\frac{N-3}{2}} h(nT) \cos\left(2\pi fT \left(n - \frac{N-1}{2}\right)\right) \right], & N \text{ dispari,} \\ T e^{-j2\pi fT \frac{N-1}{2}} \left[2 \sum_{n=0}^{\frac{N-2}{2}} h(nT) \cos\left(2\pi fT \left(n - \frac{N-1}{2}\right)\right) \right], & N \text{ pari.} \end{cases}$$

La risposta in frequenza è dunque uguale al prodotto di una funzione *reale* per un fattore a fase lineare. Eventuali discontinuità di π nella fase possono presentarsi in corrispondenza di cambi di segno della funzione reale (da cui la dizione *fase lineare a tratti*). Per N dispari, è possibile progettare i coefficienti $h(nT)$ in modo che la funzione reale sia circa uguale ad 1 in un certo intervallo di frequenze e 0 altrove, cosicché un ingresso con estensione spettrale contenuta nella banda passante si ritrova ritardato in uscita di $(N - 1)/2$ campioni e praticamente non distorto.

³Dovendo $H(f)$ essere periodica di periodo $1/T$, questo può avvenire solamente se t_0 è un multiplo di T .

Se N è pari, $(N-1)/2$ non è un numero intero e l'uscita non può essere in nessun caso una semplice versione ritardata dell'ingresso a tempo discreto. Tuttavia, se l'ingresso è ottenuto per campionamento su $Z(T)$ di una forma d'onda continua $x_a(t), t \in R$, nell'ipotesi in cui l'estensione spettrale del segnale campionato sia contenuta nella banda passante del filtro, l'uscita risulta pari a $y(nT) = x_a(nT - T(N-1)/2)$. Essa si ottiene dunque per campionamento della versione del segnale a tempo continuo ritardato di $t_0 = T(N-1)/2$.

Nota. Se N è pari risulta in ogni caso $H(F_s/2) = 0, F_s = 1/T$. Infatti, $H(F_s/2)$ risulta uguale all'area del segnale

$$h(nT)e^{-j2\pi\frac{F_s}{2}nT} = (-1)^n h(nT)$$

che risulta nulla, data la simmetria di $h(nT)$ (vedi Fig. 5.7a).

Non è dunque conveniente cercare di progettare un filtro FIR passa-alto a fase lineare con un numero pari di campioni.

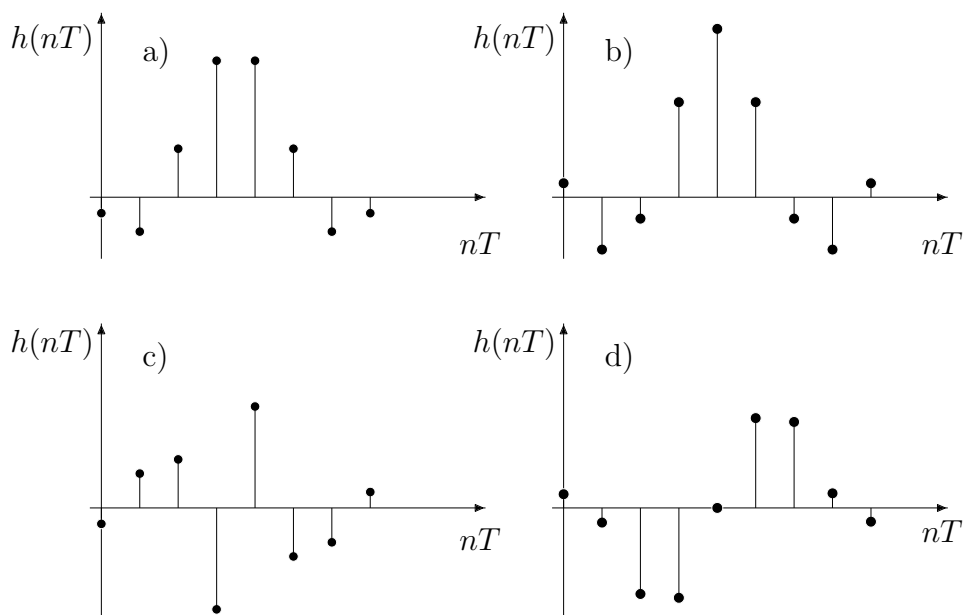


Figura 5.7: Andamenti della risposta impulsiva di filtri FIR simmetrici e antisimmetrici.

Un'altra classe di filtri FIR la cui risposta in frequenza presenta un fattore a fase lineare è quella dei filtri antisimmetrici, la cui risposta impulsiva verifica la condizione

$$h(nT) = -h((N-1-n)T).$$

Gli andamenti tipici sono mostrati nelle Fig. 5.7c e 5.7d per N pari e N dispari rispettivamente.

La risposta in frequenza risulta

$$H(f) = \begin{cases} -jT e^{-j2\pi f T \frac{N-1}{2}} \left[2 \sum_{n=0}^{\frac{N-3}{2}} h(nT) \sin \left(2\pi f T \left(n - \frac{N-1}{2} \right) \right) \right], & N \text{ dispari,} \\ -jT e^{-j2\pi f T \frac{N-1}{2}} \left[2 \sum_{n=0}^{\frac{N-2}{2}} h(nT) \sin \left(2\pi f T \left(n - \frac{N-1}{2} \right) \right) \right], & N \text{ pari.} \end{cases}$$

La risposta in frequenza risulta dunque il prodotto di un fattore a fase lineare e di una funzione puramente *immaginaria*. In particolare, il fattore $-j$ contribuisce ad una fase costante di $-\pi/2$.

Nota. Data l'antisimmetria, l'area della risposta impulsiva è nulla e risulta dunque $H(0) = 0$. Non è dunque conveniente progettare un filtro passa-basso utilizzando filtri FIR antisimmetrici.

Per N dispari, risulta anche $H(F_s/2) = 0$, per cui il progetto di filtri passa-alto non è consigliabile in questo caso.

Il progetto di filtri FIR simmetrici in MATLAB si effettua utilizzando la procedura

$$\mathbf{b} = \text{remez}(\mathbf{r}, \mathbf{f}, \mathbf{M}, \mathbf{w}).$$

Il filtro fornito dall'algoritmo di Remez, particolarizzato da Parks e McClellan al progetto di filtri FIR, è ottimo nel senso che minimizza il massimo errore di approssimazione nelle varie bande a parità di specifiche. La risposta in frequenza presenta una oscillazione uniforme nelle varie bande (ed è dunque, come si dice, di tipo *equiripple*). I parametri di ingresso alla procedura sono

1. l'ordine \mathbf{r} del filtro, la cui risposta impulsiva risulta pertanto di lunghezza $N = \mathbf{r} + 1$ campioni;
2. un vettore di frequenze \mathbf{f} , normalizzate rispetto a $F_s/2$, $F_s = 1/T$, che specifica le frequenze che delimitano le bande attenuate e oscure della risposta in frequenza;
3. un vettore \mathbf{M} che specifica le ampiezze desiderate alle frequenze specificate nel vettore \mathbf{f} ;
4. un vettore di pesi \mathbf{w} che specifica il rapporto fra le ampiezze degli errori di approssimazione nelle varie bande. Se il filtro non è soddisfacente in termini di errore *assoluto*, occorre aumentare l'ordine del filtro.

Il vettore \mathbf{b} di $\mathbf{r}+1$ elementi contiene i coefficienti del filtro, che includono la moltiplicazione per il quanto temporale T . Il filtro può essere utilizzato per l'elaborazione dell'ingresso posto nel vettore \mathbf{x} mediante il comando `filter(b,1,x)`.

Esempio 8. Supponiamo che la frequenza di campionamento del sistema numerico sia $F_s = 8$ kHz, e si voglia progettare un filtro FIR simmetrico passa-basso di ordine $r = 18$ con banda passante da 0 a 1.6 kHz e banda oscura da 2.4 kHz a 4 kHz = $F_s/2$. Detto δ_p il massimo errore nel modulo della risposta in frequenza in banda passante e δ_s il massimo errore in banda oscura, si desidera che $\delta_p/\delta_s = 2$, ovvero si tollera un errore doppio in banda passante rispetto alla banda oscura.

Il progetto e la visualizzazione del filtro possono essere ottenuti con i seguenti comandi MATLAB.

```

% Specifiche di progetto.

Fs=8000;
f=[0, 1600, 2400 4000]*2/Fs;
M=[1, 1, 0, 0];
w=[1, 2];
r=18;

% Progetto del filtro

b=remez(r,f,M,w);

% Visualizzazione del filtro

[H,f]=freqz(b,1,512,Fs);
figure;
plot(f,abs(H));
figure;
plot(f,angle(H));

```

Il modulo e la fase della risposta in frequenza del filtro ottenuto sono mostrati nelle Fig. 5.8 e 5.9. Si noti che la fase è rappresentata modulo 2π e che si hanno discontinuità di π in banda attenuata, corrispondenti a cambi di segno della risposta in frequenza. Una stima approssimata della lunghezza del filtro FIR con banda di transizione $\nu = W_s - W_p$ (espressa in termini di frequenze normalizzate rispetto a F_s) e con massimi errori δ_p e δ_s in banda passante e banda oscura, può essere ottenuta dalla relazione empirica [4],[5]

$$N \simeq \frac{-10 \log_{10}(\delta_p \delta_s) - 15}{14\nu} + 1.$$

Si può altrimenti usare la procedura MATLAB `remezord`.

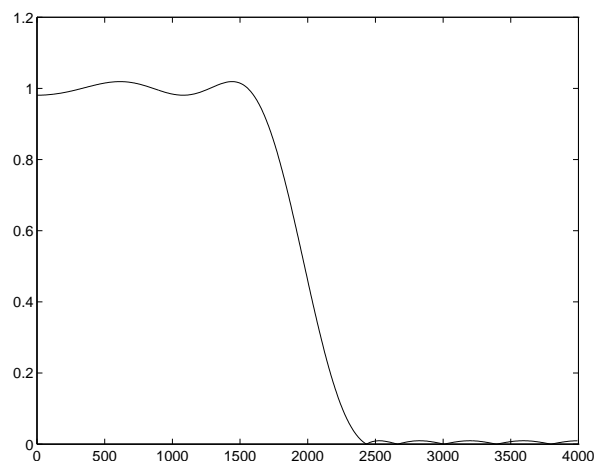


Figura 5.8: Modulo della risposta in frequenza del filtro FIR simmetrico (f in Hz).

□

Esempio 9. Progetto di un filtro passa-alto con $r + 1 = 19$ coefficienti ($F_s = 8$ kHz) con banda oscura $[0,1.5]$ kHz e banda passante $[2.5,4]$ kHz, con un errore massimo in banda passante uguale a quello in banda oscura.

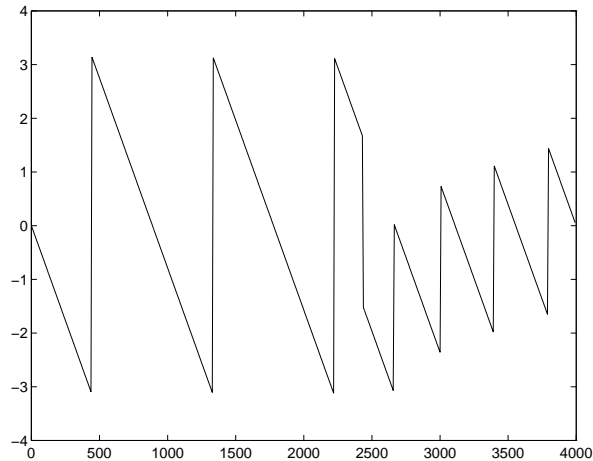


Figura 5.9: Fase (radianti) della risposta in frequenza del filtro FIR simmetrico (f in Hz).

% Specifiche di progetto.

```
Fs=8000;
f=[0, 1500, 2500, 4000]*2/Fs;
M=[0, 0, 1, 1];
w=[1, 1];
r=18;
```

% Progetto del filtro

```
b=remez(r,f,M,w);
```

□

Esempio 10. Progetto di un filtro multi banda con $r + 1 = 52$ ($F_s = 8$ kHz) con bande passanti $[0,1]$ kHz, $[2,3]$ kHz e bande oscure $[1.2,1.8]$ kHz, $[3.2,4]$ kHz. Nella banda $[0,1]$ kHz si tollera un errore pari a $1/2$ di quello nelle bande $[2,3]$ kHz e $[3.2,4]$ kHz. Nella banda $[1.2,1.8]$ kHz si tollera un errore pari a $1/3$ di quello nelle bande $[2,3]$ kHz e $[3.2,4]$ kHz.

% Specifiche di progetto.

```
Fs=8000;
f=[0, 1000, 1200, 1800, 2000, 3000, 3200, 4000]*2/Fs;
M=[1, 1, 0, 0, 1, 1, 0, 0];
w=[2, 3, 1, 1];
r=51;
```

% Progetto del filtro

```
b=remez(r,f,M,w);
```

□

Il progetto di filtri antisimmetrici in MATLAB si effettua ancora con la procedura `remez`, usando il parametro `'Hilbert'`. Si considerino i seguenti esempi.

Esempio 11. Progetto di un filtro passa-alto anti-simmetrico con $r + 1 = 18$ coefficienti ($F_s = 8$ kHz) con banda oscura $[0, 1.5]$ kHz e banda passante $[2.5, 4]$ kHz, con un errore massimo in banda passante uguale a quello in banda oscura. È opportuno scegliere una lunghezza pari del filtro, trattandosi di un filtro passa-alto antisimmetrico.

```
% Specifiche di progetto.

Fs=8000;
f=[0, 1500, 2500, 4000]*2/Fs;
M=[0, 0, 1, 1];
w=[1, 1];
r=17;

% Progetto del filtro

b=remez(r,f,M,w,'Hilbert');
```

□

Esempio 12. L'uso della parola 'Hilbert' deriva dal fatto che la fase costante $-\pi/2$ dovuta alla presenza del fattore $-j$ nella risposta in frequenza rende i filtri antisimmetrici adatti al progetto di filtri di Hilbert discreti, con risposta in frequenza di modulo unitario a tutte le frequenze e fase $-\pi/2$ per $0 < f < F_s/2$. Per il progetto di un filtro di Hilbert di ordine 30 si possono utilizzare i seguenti comandi

```
Fs=8000;
r=30;

b=remez(r,[.1 0.9],[1 1],'Hilbert')
```

Il filtro che si ottiene, a parte l'introduzione di un ritardo di 15 campioni, realizza approssimativamente la risposta in frequenza desiderata, come si può vedere con i seguenti comandi

```
N=r+1;
[H,f]=freqz(b,1,512,Fs);
figure;
plot(f,abs(H));

% Disegna la fase, compensando il fattore corrispondente al ritardo
% dell'uscita.

figure;
plot(f,angle(H.*exp(j*2*pi*f/Fs*(N-1)/2)));
```

□

Esercizi

1. Progettare un filtro passa-basso IIR ellittico che soddisfa le seguenti specifiche ($F_s = 2$ Hz): $R_p = 0.98$, $R_s = 0.05$, $W_p = 0.6$ Hz, $W_s = 0.8$ Hz. Si scriva una procedura per il calcolo dell'uscita e la si testi con gli ingressi $x(t) = \delta_{Z(T)}(t)$, $t \in Z(T)$ e $x(t) = \cos(2\pi 0.4t)1_0(t)$, $t \in Z(T)$. Si confrontino i propri risultati con quelli ottenuti usando la procedura `filter`.

2. Progettare un filtro passa-basso FIR a fase lineare che soddisfa le seguenti specifiche ($F_s=2$ Hz): $R_p = 0.98$, $R_s = 0.05$, $W_p = 0.6$ Hz, $W_s = 0.8$ Hz. Si provino vari ordini del filtro finché le specifiche non sono soddisfatte esattamente. Si scriva una procedura per il calcolo dell'uscita e la si testi con gli ingressi $x(t) = \delta_{Z(T)}(t)$, $t \in Z(T)$ e $x(t) = \cos(2\pi 0.4t)1_0(t)$, $t \in Z(T)$. Si confrontino i propri risultati con quelli ottenuti usando la procedura `filter`.
3. Progettare un filtro passa-basso IIR ellittico che soddisfa le seguenti specifiche ($F_s = 1/T = 8$ kHz): $R_p = 0.95$, $R_s = 0.02$, $W_p = 1.84$ kHz, $W_s = 2.48$ kHz. Dopo avere calcolato i poli della funzione di trasferimento con la procedura `roots`, si stimi la durata convenzionale della risposta impulsiva all'1%. Si consideri l'ingresso $x(nT) = 2 \sin(2\pi f_0 nT) + \cos(2\pi f_1 nT)$, con $f_0 = 1$ kHz e $f_1 = 3.2$ kHz e si valuti la risposta a regime permanente del filtro. Si confronti l'uscita con il segnale $x_1(nT) = 2 \sin(2\pi f_0 nT)$.
4. *Importanza della fase lineare.* Progettare un filtro passa-basso IIR ellittico che soddisfa le seguenti specifiche ($F_s = 1/T = 8$ kHz): $R_p = 0.95$, $R_s = 0.02$, $W_p = 2.4$ kHz, $W_s = 2.8$ kHz. Si consideri l'ingresso $x(nT) = 2 \sin(2\pi f_0 nT) + \cos(2\pi f_1 nT)$, con $f_0 = 2.35$ kHz e $f_1 = 1.5$ kHz e si valuti la risposta a regime permanente del filtro. Si progetti un filtro FIR simmetrico che soddisfa alle stesse specifiche. Dopo avere confrontato l'ordine dei due filtri, si consideri l'uscita in regime permanente all'ingresso $x(nT)$. Si confrontino fra loro le uscite ottenute dai due filtri e l'ingresso $x(nT)$.
5. *Notch filter.* Si consideri un filtro con funzione di trasferimento

$$H_z(z) = C \frac{1 - 2 \cos(2\pi f_0 T) z^{-1} + z^{-2}}{1 - 2\rho \cos(2\pi f_0 T) z^{-1} + \rho^2 z^{-2}}, \quad C = \frac{1 + 2\rho \cos(2\pi f_0 T) + \rho^2}{2 + 2 \cos(2\pi f_0 T)}.$$

Si valutino in forma chiusa i poli e gli zeri della funzione di trasferimento. e si disegni il modulo e la fase della risposta in frequenza al variare di $0 < \rho < 1$ e $0 < f_0 < F_s/2$ con $F_s = 1/T$. In particolare, si valuti la frequenza in cui $H(f)$ si annulla, e si considerino valori di ρ prossimi a 1. Sia $F_s = 8$ kHz. Si progetti un filtro che elimini un eventuale tono sinusoidale a 50 Hz sovrapposto al segnale utile. Tale disturbo potrebbe essere generato da interferenze con la tensione di rete.

6. Nel file

`frase_8192.mat`

sono contenuti i campioni di segnale vocale relativi alla frase “*E soprattutto, forse, ciò che gli occhi da soli non avrebbero potuto vedere.*” La frequenza di campionamento è $F_s = 8192$ Hz. Il file può essere caricato con il comando `load` ed i campioni vengono posti nel vettore `orig`.

Il segnale vocale è costituito da una successione di suoni elementari detti *fonemi*: ad esempio alla parola “fonemi” corrisponde la successione dei suoni elementari /f/ /o/ /n/ /e/ /m/ /i/. Inoltre le vocali (e in generale anche le consonanti *vocalizzate*, tipo “l”, “m”) hanno tipicamente un andamento quasi periodico. Questo contrasta con l'andamento irregolare e *rumoroso* delle *fricative*, tipo la

“s” o la “f”. Infine, le *plosive*, tipo la “p” o la “b”, sono precedute da una pausa prima dell'emissione del suono. Nella frase che stiamo esaminando, si possono osservare le caratteristiche dei vari segmenti del segnale con il comando `plot`. Se l'elaboratore è fornito di scheda audio, è possibile da MATLAB ascoltare la frase o una sua parte con il comando `sound`.

Nel file

`frase_r330.mat`

è contenuta una versione della frase corrotta “in fase di registrazione” aggiungendo al segnale un fastidioso tono a 330 Hz. Progettare un filtro notch che elimini tale disturbo.

7. Con riferimento all'esercizio precedente, nel file

`frase_r.mat`

è contenuta una versione della frase corrotta in fase di registrazione con un tono di frequenza non nota. Usando una tecnica di stima spettrale (ad esempio, il periodogramma), si stimi la frequenza della sinusoidale sovrapposta al segnale e la si elimini con un filtro notch. *Nota: il filtro notch è molto selettivo, ed è quindi importante determinare accuratamente la frequenza della sinusoidale per poterla cancellare efficacemente.*

8. Per ragioni tecniche, si vuole convertire la frequenza di campionamento del segnale vocale dell'esercizio 6 da 8192 Hz alla frequenza doppia 16384 Hz. Si realizzi il convertitore di frequenza mediante interpolazione seguita da opportuno filtraggio.
9. Il segnale telefonico viene usualmente campionato a $F_s = 8$ kHz, ma la sua estensione spettrale può essere limitata senza perdita di intelligibilità a $[-3.33, 3.33]$ kHz. Si progetti uno schema di interpolazione/campionamento per passare dalla frequenza $F_s = 8$ kHz alla frequenza $F'_s = 6.66$ kHz. Occorre dunque interpolare il segnale di un fattore 5 e sottocampionare di un fattore 6 (usando un prefiltro). Il file

`frase_8000.mat`

contiene la frase dell'esercizio 6 acquisita con $F_s = 8$ KHz.

Bibliografia

- [1] A. Oppenheim, R. Schaffer, “Digital Signal Processing”.
- [2] A. Peled, B. Liu, “Digital Signal Processing”.
- [3] G.C. Temes, S.K. Mitra, “Modern Filter Theory and Design”.
- [4] R. Crochiere, L. Rabiner, “Multirate Digital Signal Processing”.
- [5] L. Rabiner et al., “Some Comparisons between FIR and IIR Digital Filters”, Bell Systems Technical Journal, vol. 53, Feb. 1974.
- [6] M. Bellanger, “Digital Processing of Signals”.

Capitolo 6

Esercizi sulla FFT

In questa sezione vengono proposti alcuni semplici esercizi riguardanti l'uso della FFT per il calcolo della trasformata di Fourier di segnali a tempo discreto e a tempo continuo.

Si ricorda che se $x(t), t \in Z(T)/Z(NT)$ è un segnale a tempo discreto di periodo NT , la sua trasformata di Fourier ha l'espressione

$$X(kF) = T \sum_{n=0}^{N-1} x(nT) e^{-j2\pi \frac{nk}{N}}, \quad F = \frac{1}{NT}. \quad (6.1)$$

La trasformata di Fourier $X(kF)$ risulta essere definita su $Z(F)/Z(NF)$ ed è quindi completamente specificata dagli N valori assunti sulla cella elementare $f \in \{0, \dots, (N-1)F\}$. Il segnale $x(nT)$ viene recuperato da $X(kF)$ mediante la formula di inversione

$$x(nT) = F \sum_{n=0}^{N-1} X(nT) e^{j2\pi \frac{nk}{N}}. \quad (6.2)$$

L'algoritmo di FFT permette di calcolare, a partire dai valori $\{x(0), \dots, x((N-1)T)\}$, la sommatoria (6.1) in modo efficiente. In particolare, se N è una potenza di 2, la (6.1) viene calcolata in $\mathcal{O}(N \log_2 N)$ operazioni, il che consente una riduzione drastica rispetto al calcolo diretto che richiede $\mathcal{O}(N^2)$ operazioni.

La procedura MATLAB `X=fft(x,N)` calcola mediante un algoritmo di FFT la sommatoria

$$\mathbf{X}(k+1) = \sum_{n=0}^{N-1} \mathbf{x}(n+1) e^{-j2\pi \frac{nk}{N}}$$

a partire da un vettore di ingresso \mathbf{x} . Il risultato viene posto nel vettore di uscita \mathbf{X} , di dimensione N , che non è vincolato a essere una potenza di 2. Se il vettore di ingresso ha dimensione minore di N , esso viene esteso aggiungendo in coda valori nulli. Si noti che l'algoritmo di MATLAB suppone $T=1$, e che occorre in generale premoltiplicare l'ingresso per T , ponendo $\mathbf{x}(n+1) = Tx(nT)$.

Per l'antitrasformata, si può usare la procedura MATLAB `x=ifft(X)`, che calcola in modo efficiente la sommatoria

$$\mathbf{x}(k+1) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{X}(n+1) e^{j2\pi \frac{nk}{N}}$$

a partire dal vettore \mathbf{X} di dimensione N . Nel caso $T \neq 1$, occorre porre $\mathbf{X}(n+1) = X(nT)/T$.

L'algoritmo di FFT può dunque essere utilizzato per il calcolo della trasformata di Fourier di segnali a tempo discreto e periodici. Vedremo di seguito la possibilità di utilizzarla per il calcolo della trasformata di segnali a tempo discreto non periodici e di segnali continui.

Segnali a tempo discreto. Se $x(nT), t \in Z(T)$ è un segnale a tempo discreto, in generale non periodico, con estensione contenuta in $\{0, \dots, (N-1)T\}$, possiamo usare la FFT per calcolare i *campioni* della sua trasformata di Fourier. Infatti, dalla definizione di trasformata, si ottiene

$$X(f) = T \sum_{n=-\infty}^{+\infty} x(nT) e^{-j2\pi f nT} = T \sum_{n=0}^{N-1} x(nT) e^{-j2\pi f nT}, \quad f \in R/Z(1/T).$$

Se calcoliamo la relazione precedente per $f = kF$, dove $F = 1/NT$, otteniamo proprio

$$X(kF) = T \sum_{n=0}^{N-1} x(nT) e^{-j2\pi \frac{nk}{N}},$$

che ha la stessa struttura della (6.1). Dunque, N campioni in un periodo di $X(f)$, $X(kF), k = 0, \dots, (N-1)F$, possono essere calcolati usando la FFT. Si noti che possiamo infittire i campioni in cui viene calcolata la trasformata $X(f)$ semplicemente aumentando la dimensione da N a $M > N$: in tale caso, i campioni aggiuntivi $x(NT), \dots, x((M-1)T)$ da usare per il calcolo usando la (6.1) sono evidentemente uguali a zero.

Se il segnale $x(nT)$ ha durata finita $D \leq NT$, ma la sua estensione non è contenuta nella cella elementare $\{0, \dots, (N-1)T\}$, occorre osservare che ai fini del calcolo dei *campioni* $X(kF), k = 0, \dots, (N-1)F$, possiamo considerare la versione *periodicizzata* del segnale di ingresso

$$x_p(nT) = \text{per}_{NT} x(nT).$$

Infatti, se $x(nT)$ ha estensione contenuta in $\{n_0T, (n_0+1)T, \dots, (n_0+N-1)T\}$, si ha

$$X(kF) = T \sum_{n=n_0}^{n_0+N-1} x(nT) e^{-j2\pi nTkF} = T \sum_{n=0}^{N-1} x_p(nT) e^{-j2\pi \frac{nk}{N}}.$$

Nel calcolo dei campioni della trasformata tramite la (6.1) occorre dunque considerare i valori nella cella $\{0, \dots, (N-1)T\}$ del *segnale periodicizzato*.

Infine, se il segnale $x(nT)$ ha durata praticamente limitata, è possibile considerarlo nullo entro una precisione prefissata al di fuori di un certo intervallo I_0 . Questo equivale in effetti a moltiplicare il segnale di ingresso per una finestra che vale 1 per $t \in I_0$ e 0 altrove. I valori calcolati dei campioni della trasformata saranno pertanto relativi alla convoluzione tra la trasformata di Fourier del segnale di ingresso con quella della finestra.

Stima della trasformata per segnali a tempo continuo. Supponiamo di avere un segnale a tempo continuo $x(t), t \in R$ e di volere calcolare una stima della sua trasformata di Fourier $X(f), f \in R$. Per potere utilizzare a tale scopo un algoritmo numerico, e in particolare la FFT, occorre dapprima campionare il segnale di ingresso, ponendo $x_c(t) = x(t), t \in Z(T)$. La scelta di T deve essere fatta in modo

da limitare l'*aliasing*, e in generale deve essere $F_c = 1/T > B$, dove B è la larghezza di banda del segnale. Una volta campionato, il segnale ha una trasformata di Fourier periodica $X_c(f) = \text{rep}_{F_c} X(f)$, $f \in R/Z(F_c)$ i cui campioni in un periodo possono essere stimati mediante un algoritmo di FFT, secondo quanto esposto relativamente ai segnali a tempo discreto. Se non si conosce a priori la larghezza di banda B del segnale, occorrerà procedere per tentativi, valutando se le stime che si ottengono variano entro la precisione desiderata al variare di T .

Se il segnale $x(t)$ è reale, la sua trasformata di Fourier ha simmetria hermitiana ed estensione spettrale simmetrica $[-B/2, B/2]$. L'algoritmo di FFT fornisce i campioni della trasformata periodica $X_c(f) = \text{rep}_{F_c} X(f)$ relativamente al periodo $[0, F_c]$, mentre il confronto con $X(f)$ è più significativo se si considera la cella elementare $[-F_c/2, F_c/2]$. Posto $F = F_c/N$, una volta calcolati i valori $X_c(0), \dots, X_c((N-1)F)$ con un algoritmo di FFT, converrà dunque riferirsi al vettore

$$\left[X_c\left(\frac{N}{2}F\right), X_c\left(\left(\frac{N}{2} + 1\right)F\right), \dots, X_c((N-1)F), X_c(0), \dots, X_c\left(\left(\frac{N}{2} - 1\right)F\right) \right],$$

per N pari, e a

$$\left[X_c\left(\frac{N+1}{2}F\right), X_c\left(\left(\frac{N+1}{2} + 1\right)F\right), \dots, X_c((N-1)F), X_c(0), \dots, X_c\left(\left(\frac{N-1}{2}\right)F\right) \right],$$

per N dispari, che contengono la sequenza dei campioni relativi al periodo $[-F_c/2, F_c/2]$. Tale operazione di ordinamento degli elementi del vettore di uscita può ottenersi con la funzione MATLAB `fftshift`.

Esercizi

1. Si consideri il segnale $x(t)$, $t \in Z(T)$, $T = 2$, con $x(0) = 1$, $x(\pm T) = 2$, $x(\pm 2T) = 3$ e $x(nT) = 0$ altrove. Si calcoli in forma chiusa la trasformata di Fourier $X(f)$, $f \in R/Z(1/T)$. Si calcolino inoltre N campioni della trasformata in un periodo usando un algoritmo di FFT, per $N = 8, 16, 256$, e si verifichi in un grafico la coincidenza fra i campioni calcolati e l'espressione teorica.

Si valuti inoltre approssimativamente il numero di operazioni richieste per il calcolo degli N campioni della trasformata usando la FFT ed il calcolo diretto. Si tenga presente che il segnale di ingresso ha solamente 5 campioni diversi da zero.

Nota: $X(f)$ è in generale una funzione complessa, ed occorre considerarne di volta in volta la parte reale e immaginaria, oppure il modulo e la fase.

2. Si calcoli la DFT dei seguenti segnali, definiti su $Z(T)$, e si disegni il grafico del modulo, cercando di spiegarne le caratteristiche. Sia $F = 1/T = 8$ kHz.

- a. Trasformata su $N = 512$ punti. Segnale di ingresso

$$x(nT) = 1, \quad 0 \leq n \leq 63 \\ 0, \quad \text{altrove.}$$

- b. Trasformata su $N = 256$ punti. Segnale di ingresso

$$x(nT) = 0.5 \sin\left(\frac{2\pi n}{256} 9\right) + 0.7 \sin\left(\frac{2\pi n}{256} 100\right), \quad 0 \leq n \leq 255.$$

c. Trasformata su $N = 256$ punti. Segnale di ingresso

$$x(nT) = 0.5 \sin\left(\frac{2\pi n}{256} \frac{100}{3}\right) + 0.7 \sin\left(\frac{2\pi n}{256} \frac{200}{3}\right), \quad 0 \leq n \leq 255.$$

3. Una sinusoidale di frequenza 50 Hz viene campionata con 500 campioni al secondo e analizzata usando una FFT su 64 campioni.
 - a. Quale campione della trasformata sarà maggiore in modulo?
 - b. Quale campione della trasformata avrà un'ampiezza immediatamente inferiore a quella del campione di modulo maggiore, e qual'è l'ampiezza relativa dei due campioni (esprimere il risultato in dB)?

Si verifichino i risultati con MATLAB.

4. Siano dati i seguenti segnali, definiti su $Z(T)$ per $T = 1$.
 - a. $x(nT) = \sin\left(\frac{2\pi n}{256} 172\right)$,
 - b. $x(nT) = \sin(n)$.

Si analizzino i due segnali mediante FFT su 256 punti. Si preveda quale sarà il coefficiente della trasformata più grande in modulo nei due casi, e l'ampiezza relativa di tali campioni della trasformata. Si verifichi il risultato con MATLAB.

5. Questo esercizio mostra come occorra fare attenzione nell'uso della FFT per il calcolo della trasformata di Fourier di segnali sinusoidali. Si consideri il segnale $x(t) = \cos 2\pi f_0 t$, $t \in Z(T)$, $T = 1$, $f_0 = 0.25$ e si costruisca il vettore MATLAB

$$\mathbf{x}(n+1) = x(nT), \quad n = 0, \dots, N-1$$

per $N = 32$ e $N = 64$. Si calcoli poi la FFT di \mathbf{x} nei due casi su N punti e la si disegni usando la funzione `stem`. Si discuta il risultato ottenuto. Si ripeta lo stesso esperimento per $f_0 = 0.254$ e si discuta il risultato ottenuto. Si risolva inoltre l'esercizio per via analitica e si confronti con la soluzione trovata per via numerica.

Nota: per $f_0 = 0.25$ e i due valori di N considerati, $x(t)$ è periodico di periodo NT . Per $f_0 = 0.254$, $x(t)$ non è periodico di periodo NT .

6. Si scriva una procedura analoga alla procedura MATLAB `freqz`, usando la FFT.
7. Dati i segnali $x(nT)$ con estensione $0, \dots, (N-1)T$ e $h(nT)$ con estensione $0, \dots, (M-1)T$, si scriva una procedura MATLAB che calcola la convoluzione $y(nT) = h * x(nT)$ mediante l'uso della FFT. Si confronti il risultato con quello fornito dalla funzione MATLAB `y=T*conv(h,x)`, in cui si pone $\mathbf{x}(n+1) = x(nT)$, $nT = 0, \dots, (N-1)T$, e $\mathbf{h}(n+1) = h(nT)$, $nT = 0, \dots, (M-1)T$. Si scriva un programma che calcola la convoluzione di due segnali generici con estensioni $-5T, \dots, 4T$ e $3T, \dots, 5T$.

Soluzione proposta per la prima parte:

```

function y=convv(T,x,h);

N=length(x); M=length(h);

L=N+M-1;

X=T*fft(x,L); H=T*fft(h,L);

Y=X.*H;

y=ifft(Y)/T;

% Se x e h sono reali, y e' reale. Si pu\`o eliminare
% l'eventuale parte immaginaria, non identicamente nulla a causa
% degli errori di calcolo.

if (isreal(x) & isreal(h)) y=real(y);

end;

```

8. *Metodo Overlap-Add.* Con riferimento all'esercizio precedente, se il segnale $x(t)$, $t \in Z(T)$ ha una durata NT molto più grande di quella MT di $h(t)$, non è efficiente calcolare direttamente la convoluzione usando la FFT, che deve essere valutata per entrambi i segnali su almeno $M + N - 1$ punti. È conveniente allora suddividere il segnale di ingresso in blocchi non sovrapposti lunghi L ,

$$x_i(nT) = \begin{cases} x(nT), & n = iL, \dots, iL + L - 1, \\ 0, & \text{altrove} \end{cases}$$

ponendo

$$x(nT) = \sum_i x_i(nT).$$

Il risultato della convoluzione viene ottenuto *sommando* le uscite $x_i * h(nT)$, ciascuna delle quali può essere calcolata usando la FFT. Si noti che le uscite $x_i * h(nT)$ risultano sovrapposte di $M - 1$ campioni.

Si scriva un programma MATLAB che calcola la convoluzione di due segnali con il metodo *Overlap-Add*, ponendo $M = 4$ e $N = 64$, e si confronti il risultato con il metodo diretto e quello dell'esercizio 7.

9. Dati due segnali $x_1(t)$ e $x_2(t)$, $t \in Z(T)$ a durata limitata LT , è possibile considerarne le ripetizioni periodiche $x_{1,p}(t) = \text{rep}_{LT}x_1(t)$, $t \in Z(T)/Z(LT)$ e $x_{2,p}(t) = \text{rep}_{LT}x_2(t)$, $t \in Z(T)/Z(LT)$, senza perdita di informazione. Dimostrare e verificare su un esempio sviluppato con MATLAB che la convoluzione *ciclica* $y_p(nT) = x_{1,p} * x_{2,p}(nT)$ è uguale alla ripetizione periodica della convoluzione $y(nT) = x_1 * x_2(nT)$, ovvero $y_p(nT) = \text{rep}_{LT}y(nT)$.
10. *Metodo Overlap-Save.* Si considerino, come nell'Esercizio 8, due segnali $x(t)$, $t \in Z(T)$ e $h(t)$ con estensione $[0, \dots, (N - 1)T]$ e $[0, \dots, (M - 1)T]$, rispettivamente, e $N \gg M$. Nel metodo *Overlap-Save* per il calcolo della convoluzione, il segnale

$x(nT)$ viene suddiviso in blocchi $x_k(nT)$ lunghi L campioni. Il metodo consiste nel calcolare una convoluzione ciclica (vedi Esercizio 9) tra $h(nT)$ e $x_k(nT)$, identificando quella parte della convoluzione ciclica che corrisponde alla normale convoluzione. In particolare, supponendo $L \geq M$, nella convoluzione ciclica di $h(nT)$ e $x_k(nT)$ calcolata su L punti, risulta che i primi $M - 1$ non sono corretti, mentre i rimanenti punti sono gli stessi che otterremmo dalla normale convoluzione (perchè? Vedi l'Esercizio 9). Convien dunque sezionare $x(nT)$ in segmenti di lunghezza L in modo che ogni segmento si sovrapponga al precedente per $M - 1$ punti.

Dopo avere definito i segmenti $x_k(nT)$ nel modo seguente

$$x_k(nT) = x((n + k(L - M + 1))T), \quad 0 \leq n \leq L - 1,$$

si calcola pertanto la convoluzione ciclica $y_k(nT) = h * x_k(nT)$ su L punti e se ne scartano i primi $M - 1$. I rimanenti punti di ogni sottosequenza $y_k(nT)$ vengono giustapposti gli uni di seguito agli altri, fino ad ottenere l'uscita filtrata finale.

Si scriva un programma MATLAB che calcola la convoluzione di due segnali con il metodo *Overlap-Save*, ponendo $M = 4$ e $N = 64$, e si confronti il risultato con il metodo diretto e quello dell'Esercizio 7.

11. Si stimi la trasformata di Fourier del segnale a tempo continuo

$$x(t) = e^{-|6t|}, \quad t \in R.$$

12. Si stimi la trasformata di Fourier del segnale a tempo continuo

$$x(t) = e^{-|6(t-1)|}, \quad t \in R.$$

13. Si stimi la trasformata di Fourier del segnale a tempo continuo

$$x(t) = e^{-|6t|} \cos(2\pi 0.5t), \quad t \in R.$$

14. Si stimi la trasformata di Fourier del segnale a tempo continuo

$$x(t) = \frac{\sin(\pi(1-r)t/T) + 4r(t/T) \cos(\pi(1+r)t/T)}{\pi[1 - (4rt/T)^2]t/T} \text{rect} \frac{t}{8T}, \quad r = 0.125, \quad t \in R.$$

Bibliografia

- [1] N. Benvenuto and G. Cherubini, “Algorithms for communications systems and their application”, London: John Wiley & Sons, 2002.
- [2] L.W. Couch II, “Digital and analog communication systems ”, Upper Saddle River, NJ: Prentice-Hall, 1997.
- [3] M.S. Roden, “Analog and digital communication systems ”, Upper Saddle River, NJ: Prentice-Hall, 1996.
- [4] J.G. Proakis and M. Salehi, “Communication systems engineering”, Englehood Cliffs, NJ: Prentice-Hall, 1994.