

This notes have been written by an ex-Studente, Ing. Matteo Beardo, that I kindly thanks!

Though they are in Italian, they can be useful for some of you, particularly for the first part related to:

- signal theory
- analytic signal
- noise

Please note that, from 2013 in the class "Digital Transmission" the parts on analog modulations (AM and FM) and on PCM are NOT ANY MORE part of the program.

I unfortunately didn't have time to review this material personally, but I decided to make it available because it appears as very well written.

Roberto Gaudino

Dispense di Comunicazioni Elettriche

Ing. Matteo Beardo

Versione del 4 marzo 2013

Indice

Introduzione	v
1 Richiami di teoria dei segnali determinati	1
1.1 Classificazione dei segnali	1
1.2 Segnali determinati	2
1.2.1 Medie temporali su segnali determinati	3
1.2.2 Segnali a potenza finita	3
1.2.3 Segnali a energia finita	5
1.2.4 Distribuzione δ di Dirac	5
1.3 Rappresentazione segnali su basi ortogonali	6
1.4 La Serie di Fourier	10
1.5 La Trasformata di Fourier	11
1.5.1 Proprietà della trasformata di Fourier	12
1.5.2 Trasformata di Fourier di segnali periodici	14
1.6 Sistemi lineari e tempo invarianti (LTI)	14
1.6.1 Sistemi causali	17
1.6.2 Risposta in frequenza	17
1.6.3 Sistemi LTI non distorcenti	18
1.7 Densità spettrale di energia e di potenza	19
1.7.1 Densità spettrale di energia	19
1.7.2 Densità spettrale di potenza	20
1.7.3 Funzione di autocorrelazione	21
1.8 Teorema del campionamento	22
1.8.1 Campionamento ideale	22
1.8.2 Ricostruzione del segnale campionato	23
2 Richiami sui processi casuali	27
2.1 Teoria della probabilità	27
2.1.1 Variabili casuali	28
2.1.2 Coppie di variabili casuali	33
2.1.3 Momenti	33
2.1.4 Indipendenza statistica e correlazione	36
2.1.5 Medie condizionate	36
2.2 Processi casuali	37
2.2.1 Caratterizzazione statistica di un processo casuale	39
2.2.2 Stazionarietà in senso stretto	41
2.2.3 Stazionarietà in senso lato (WSS)	41

2.2.4	Ergodicità	44
2.2.5	Densità spettrale di potenza	47
2.2.6	Trasformazioni LTI di processi casuali	49
2.3	Conclusione	51
3	Introduzione ai sistemi di trasmissione	53
3.1	Tipologie di segnali e di sistemi di trasmissione	53
3.1.1	Sistemi analogici	53
3.1.2	Sistemi digitali	55
3.1.3	Sistemi misti analogico-digitali	55
3.1.4	Segnali audio	57
3.1.5	Segnali video	59
3.2	Grandezze fondamentali di un sistema di trasmissione	59
3.2.1	Banda	60
3.2.2	Potenza	60
3.2.3	Complessità	60
3.3	Richiami sui decibel	60
3.4	Definizioni di banda	61
3.5	Definizione di canale	63
3.6	Propagazione guidata e in spazio libero	63
3.6.1	Sistemi via cavo	63
3.6.2	Sistemi via etere	66
3.7	Riassunto formule principali sulla propagazione	70
4	Trasmissioni analogiche in banda base	71
4.1	Introduzione	71
4.2	Trasmissioni analogiche in banda base	72
4.2.1	Canale ideale	72
4.2.2	Canale non ideale	74
4.2.3	Cascata di due trasformazioni	75
4.2.4	Esempio 1: canale ideale + rumore	75
4.2.5	Esempio 2: segnale + rumore gaussiano e filtro	76
4.3	Riassunto formule principali sulla trasmissione in banda base	76
5	Rumore termico e doppi bipoli	79
5.1	Rumore termico	79
5.1.1	Descrizione probabilistica del rumore termico	79
5.1.2	Temperatura equivalente di rumore di una sorgente	84
5.2	Caratterizzazione di doppi bipoli rumorosi	85
5.2.1	Cifra di rumore di un doppio bipolo	86
5.2.2	Temperatura equivalente di rumore di un doppio bipolo	87
5.2.3	Formule di conversione tra cifra di rumore e temperatura equivalente di rumore	87
5.2.4	Interpretazione fisica della cifra di rumore	88
5.2.5	Calcolo della cifra di rumore di un attenuatore	89
5.3	Cascate di doppi bipoli	90
5.3.1	Calcolo della cifra di rumore e della temperatura equivalente di rumore di una cascata di doppi bipoli	91

5.3.2	Sistemi multitratta	93
5.4	Conclusione	95
5.5	Riassunto formule principali sul rumore termico	95
6	Segnale analitico e inviluppo complesso	97
6.1	Classificazione spettrale dei segnali	97
6.1.1	Accenno alla modulazione	98
6.2	Segnale analitico	98
6.3	Inviluppo complesso	101
6.3.1	Rappresentazione geometrica di un segnale in banda traslata . .	102
6.4	Spettri	103
6.5	Equivalente passa-basso di un sistema passa-banda	106
6.6	Rumore in banda traslata	108
6.7	Conclusione	109
7	Modulazioni di ampiezza: AM	111
7.1	Introduzione	111
7.2	Modulazione AM standard	112
7.2.1	Calcolo densità spettrale di potenza per AM standard	112
7.2.2	Ricevitore a rivelazione di inviluppo	116
7.3	Modulazione AM-DSB-SC	120
7.4	Modulazione SSB	122
7.4.1	Schema alternativo per il modulatore SSB	124
7.4.2	Demodulazione di segnali SSB	126
7.5	Prestazioni modulazioni AM in presenza di rumore	127
7.5.1	Calcolo SNR per sistemi AM tradizionali con ricezione coerente .	129
7.5.2	SNR sistemi AM-DSB-SC con ricezione coerente	133
7.5.3	SNR sistemi AM-SSB con ricezione coerente	133
7.5.4	Calcolo prestazioni con ricevitore di inviluppo	134
7.6	Confronto tra le varie modulazioni di ampiezza	137
7.7	Riassunto formule principali su AM	140
8	Modulazioni angolari: FM e PM	141
8.1	Introduzione	141
8.2	Rappresentazione di segnali FM e PM	141
8.3	Spettri di modulazioni angolari	146
8.3.1	Spettro di segnale angolare con modulante sinusoidale	146
8.3.2	Occupazione spettrale di segnali FM/PM: regola di Carson . . .	153
8.3.3	Altro risultato	156
8.4	Implementazione di modulatori e demodulatori per sistemi FM	156
8.4.1	Modulatori FM	157
8.4.2	Demodulatori FM	158
8.5	Effetto del rumore sulle modulazioni angolari	159
8.5.1	Calcolo rapporto segnale-rumore per sistemi PM	161
8.5.2	Calcolo rapporto segnale-rumore per sistemi FM	168
8.5.3	Effetto soglia nelle modulazioni angolari	172
8.5.4	Tecniche di pre-enfasi e de-enfasi per la modulazione FM	173
8.6	Confronto tra sistemi di modulazione	175

8.7	Conclusione	176
8.8	Riassunto formule principali sulle modulazioni angolari	177
9	Sistemi PCM	179
9.1	Introduzione	179
9.2	Quantizzazione uniforme	181
9.2.1	Calcolo del rapporto segnale-rumore di quantizzazione	182
9.2.2	Prestazioni del PCM in presenza di errori sul canale	186
9.2.3	Confronto con i sistemi analogici	193
9.3	Quantizzazione non uniforme	194
9.3.1	Calcolo dell'errore di quantizzazione	197
9.3.2	Tecniche di companding	200
9.4	Conclusione	202
9.5	Riassunto formule principali sul PCM	203
10	Teoria dell'Informazione per sorgenti	205
10.1	Introduzione	205
10.2	Codifica di sorgente	205
10.3	Teoria dell'Informazione per sorgenti: premesse matematiche	207
10.3.1	Entropia di sorgenti senza memoria	209
10.3.2	Codifica di sorgenti senza memoria	211
10.4	La codifica di Huffman	213
10.5	Conclusione	216
10.6	Riassunto formule principali sulla codifica di sorgente	216
	Bibliografia	219

Introduzione

Queste dispense sono state inizialmente sviluppate per il corso di *Comunicazioni Elettriche per TLC* che teneva il professor Roberto Gaudino al Politecnico di Torino fino all'anno scorso. Si trattava di un corso che trattava le modulazioni analogiche e alcuni concetti sui sistemi digitali. Questo corso adesso non esiste più ed è stato sostituito dal corso più moderno *Digital Transmission*, che tratta esclusivamente di trasmissioni numeriche.

Nonostante il cambiamento, parecchi argomenti che qui affronteremo sono ancora trattati nel nuovo corso, quindi io e il professor Gaudino riteniamo che queste dispense possano essere ancora utili agli studenti.

Sottolineo che questo lavoro è ancora in una fase provvisoria, quindi sicuramente conterrà diversi errori e imprecisioni. Tutti i capitoli sono stati scritti a partire dagli appunti delle lezioni del professor Gaudino, che sono stati rivisti e ampliati. Ovviamente gli errori che troverete sono solo colpa del sottoscritto. Siete pregati di inviare segnalazioni di errori e/o commenti positivi e negativi a:

`mattebeardo@gmail.com`

così ci aiuterete a migliorare la qualità del materiale. Mi interesserebbe anche sapere se qualche argomento non è spiegato bene, in modo da poterlo sistemare.

Queste dispense, come il corso *Digital Transmission* del resto, presuppongono che lo studente abbia una conoscenza dei concetti fondamentali di Analisi Matematica, Teoria dei Segnali, Calcolo delle Probabilità e Processi Casuali. Non è richiesta tuttavia una conoscenza particolarmente approfondita di queste materie.

Gli argomenti trattati nei vari capitoli sono:

1. Richiami di Teoria dei Segnali Determinati.
2. Richiami sui Processi Casuali.
3. Introduzione ai sistemi di trasmissione analogici e digitali.
4. Trasmissioni analogiche in banda base.
5. Rumore e doppi bipoli rumorosi.
6. Segnale analitico e inviluppo complesso.
7. Modulazioni di ampiezza: AM.

8. Modulazioni angolari: FM e PM.
9. PCM.
10. Introduzione alla Teoria dell'Informazione per sorgenti.

Potete trovare queste dispense sul sito Internet

http://www.tlc.polito.it/~gaudino/com_elt/

assieme a parecchio materiale didattico che il professor Gaudino aveva preparato per il corso di Comunicazioni Elettriche. Si tratta di materiale valido che può essere ancora utile per il nuovo corso *Digital Transmission*.

Queste dispense sono state interamente composte con il programma $\text{\LaTeX} 2_{\epsilon}$. Si tratta di un programma di tipocomposizione gratuito e altamente professionale (lo usano anche per scrivere libri). Consiglio a tutti gli studenti di scaricare questo programma, imparare a usarlo e poi utilizzarlo per scrivere la tesi di laurea: il risultato sarà molto professionale. Per acquisire una conoscenza di base di $\text{\LaTeX} 2_{\epsilon}$ basta un pomeriggio (per imparare a usarlo bene serve molto più tempo ma per scrivere la vostra tesi vi basterà una conoscenza di base).

Oltre al professor Gaudino, desidero ringraziare la mia famiglia e dedicare questo lavoro a mia sorella Alessandra, che lo ha reso possibile.

Capitolo 1

Richiami di teoria dei segnali determinati

Questo capitolo costituisce un ripasso di teoria dei segnali determinati. Vengono rivisti i concetti principali su segnali e sistemi lineari che sono stati affrontati nei corsi di Teoria dei Segnali. Per una trattazione approfondita di questi argomenti si consiglia di consultare il testo [1].

1.1 Classificazione dei segnali

Un *segnale* è una funzione del tempo t . Anche se in teoria qualsiasi funzione matematica potrebbe rappresentare un segnale, al concetto di segnale è sempre associato un fenomeno fisico. Quindi, in pratica, soltanto le funzioni "fisicamente realizzabili" possono rappresentare un segnale. Possibili esempi di segnali sono: l'andamento della temperatura nell'arco di un certo periodo di tempo, la voce umana, le vibrazioni del suolo, il campo elettromagnetico ricevuto dal nostro telefonino.

In queste dispense siamo interessati alle applicazioni dei segnali nel settore delle telecomunicazioni. In questo campo, i segnali di interesse vengono prima di tutto trasformati da opportuni trasduttori in segnali elettrici, poi vengono elaborati oppure trasmessi sotto questa forma. I segnali che considereremo nel seguito rappresenteranno quindi tensioni, correnti o campi elettrici.

I segnali si possono classificare in diversi modi. Innanzitutto si può distinguere tra segnali a tempo continuo e segnali a tempo discreto.

Segnali a tempo continuo. Sono segnali in cui la variabile indipendente t è un numero reale: li indicheremo come $x(t)$.

Segnali a tempo discreto. Sono segnali in cui la variabile indipendente n è un numero intero: li indicheremo come $x[n]$.

Spesso i segnali a tempo discreto sono ottenuti campionando segnali a tempo continuo con un periodo di campionamento T_0 :

$$x[n] = x(nT_0)$$

Ci sono altri modi di classificare segnali. I segnali possono assumere valori numerici reali o complessi.

Segnali reali. Un segnale reale assume valori tra i numeri reali, cioè $x(t) \in \mathbb{R}$. Tutti i segnali realizzabili in pratica sono segnali reali.

Segnali complessi. Un segnale complesso assume valore tra i numeri complessi, cioè $x(t) \in \mathbb{C}$. I segnali complessi non sono fisicamente realizzabili, tuttavia si utilizzano per rappresentare alcuni segnali di interesse nel campo delle telecomunicazioni. Come i numeri complessi, un segnale complesso può essere rappresentato da due segnali reali. Questi due segnali possono essere parte reale e parte immaginaria oppure modulo e fase. Di solito in pratica si usano modulo e fase.

Un'altra importante distinzione è quella tra segnali determinati e processi casuali.

Segnali determinati. Segnali in cui a ogni istante di tempo t il valore $x(t)$ è un numero reale o complesso.

Processi casuali. Segnali in cui a ogni istante di tempo t il valore di $x(t)$ è una variabile casuale.

Nei prossimi paragrafi rivedremo i concetti principali sui segnali determinati. Parleremo dei processi casuali nel prossimo capitolo.

Distinguiamo infine tra segnali periodici e non periodici.

Segnali periodici. Un segnale periodico soddisfa la seguente proprietà:

$$x(t + T_0) = x(t) \quad (1.1)$$

per qualsiasi t . I segnali periodici hanno durata infinita, quindi non sono fisicamente realizzabili. Tuttavia sono utili come modello matematico per rappresentare i segnali ciclici.

Segnali non periodici. Sono tutti i segnali che non soddisfano la condizione (1.1).

1.2 Segnali determinati

In questa sezione consideriamo i segnali determinati a tempo continuo, che indicheremo come $x(t)$. Come abbiamo già detto, $x(t)$ rappresenta l'andamento nel tempo di una tensione, corrente o campo elettrico.

Un segnale $x(t)$ a tempo continuo è fisicamente realizzabile se ha le seguenti caratteristiche:

- ▷ $x(t)$ è reale;
- ▷ $x(t)$ ha ampiezza finita in qualsiasi istante di tempo;
- ▷ il supporto temporale è finito: il segnale ha un inizio e una fine;
- ▷ il supporto spettrale è finito: passa attraverso sistemi fisici a banda finita.

I supporti (sia nel tempo che in frequenza) saranno sempre intesi in maniera "ingegneristica": si intende cioè che il supporto del segnale sia quell'intervallo (di tempo o di frequenza) al di fuori del quale il segnale vada "pressoché a zero".

Molte di queste condizioni vengono spesso rimosse per comodità matematica. Ad esempio si considereranno i segnali periodici che sono illimitati nel tempo. Si useranno anche segnali complessi per comodità di rappresentazione. In conclusione, si tratterà spesso di segnali a durata illimitata nel tempo, eventualmente complessi.

1.2.1 Medie temporali su segnali determinati

Per caratterizzare un segnale si utilizzano spesso delle misure legate a medie nel dominio del tempo. Dato un segnale $x(t)$ si definisce la sua media temporale in questo modo:

$$\langle x(t) \rangle \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) dt$$

La media temporale è un operatore lineare:

$$\langle a_1 x_1(t) + a_2 x_2(t) \rangle = a_1 \langle x_1(t) \rangle + a_2 \langle x_2(t) \rangle$$

Se $x(t)$ è un segnale periodico, la sua media vale:

$$\langle x(t) \rangle = \frac{1}{T_0} \int_{a-\frac{T_0}{2}}^{a+\frac{T_0}{2}} x(t) dt$$

Notare che essendo $x(t)$ periodico, a può assumere un valore qualunque, cioè la media si può calcolare semplicemente su un periodo, preso in qualunque posizione sull'asse dei tempi.

Il valor medio di un segnale viene detto anche componente continua o DC:

$$\begin{aligned} x_{DC} &= \langle x(t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) dt \end{aligned}$$

Su segnali reali spesso la componente continua è misurata su un intervallo di tempo $[t_1, t_2]$, dunque:

$$x_{DC} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t) dt$$

1.2.2 Segnali a potenza finita

È noto dai corsi di elettrotecnica che la potenza istantanea dissipata ai capi di una resistenza R è data da:

$$P(t) = \frac{v^2(t)}{R} = R \cdot i^2(t) \quad (1.2)$$

Dall'espressione (1.2) si vede che la potenza di un segnale elettrico è proporzionale al suo valore al quadrato (sia che il segnale sia una tensione, sia che sia una corrente). Per comodità, nell'ambito della teoria dei segnali si considerano sempre potenze normalizzate su una resistenza con valore $R = 1 \Omega$.

Alla luce delle considerazioni appena svolte, dato un generico segnale $x(t)$ adimensionato, si definisce la sua *potenza istantanea* in questo modo:

$$P_x(t) \triangleq |x(t)|^2 \quad (1.3)$$

Per i segnali reali basta usare $x^2(t)$ al posto di $|x(t)|^2$. Per ottenere la potenza in Watt basterà poi moltiplicare per opportuni coefficienti: R se $x(t)$ è una corrente e $1/R$ se $x(t)$ è una tensione.

La *potenza media* di un segnale non sarà altro che il valor medio della potenza istantanea $P_x(t)$, calcolato su tutto l'asse dei tempi. Utilizzando il teorema della media integrale, si può quindi definire la *potenza media* di un segnale in questo modo:

$$P_x \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} |x(t)|^2 dt \quad (1.4)$$

Un segnale si dice *a potenza finita* quando la sua potenza media data dalla (1.4) è finita e non nulla:

$$0 < P_x < \infty$$

In realtà qualunque segnale fisicamente realizzabile ha durata limitata e dunque, con la definizione per T che tende a infinito, risulta avere $P_x = 0$. Tuttavia, nell'ambito della teoria delle comunicazioni, risulta spesso utile estendere il dominio dei segnali in gioco tra $-\infty$ e $+\infty$.

Sui segnali fisici spesso si definiscono delle potenze medie su intervalli *finiti*, e sono queste le quantità che gli strumenti misurano.

Esempio 1.1. Calcolare la potenza media del seguente segnale sinusoidale:

$$x(t) = A_p \cos(2\pi f_0 t + \varphi)$$

con $T = 1/f_0$.

Soluzione. La potenza istantanea vale:

$$P_{istantanea}(t) = A_p^2 \cos^2(2\pi f_0 t + \varphi)$$

La potenza di picco vale:

$$P_{picco} = A_p^2 \quad (1.5)$$

La potenza media vale:

$$P_{media} = \langle P_{istantanea}(t) \rangle \quad (1.6a)$$

$$= \frac{A_p^2}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \cos^2(2\pi f_0 t + \varphi) dt \quad (1.6b)$$

$$= \frac{A_p^2}{2T} \left[\int_{-\frac{T}{2}}^{+\frac{T}{2}} 1 dt + \int_{-\frac{T}{2}}^{+\frac{T}{2}} \cos(4\pi f_0 t + 2\varphi) dt \right] \quad (1.6c)$$

$$= \frac{A_p^2}{2T} T \quad (1.6d)$$

$$= \frac{A_p^2}{2} \quad (1.6e)$$

Per passare dalla (1.6b) alla (1.6c) si è utilizzata la formula di bisezione del coseno:

$$\cos^2(x) = \frac{1 + \cos(2x)}{2}$$

Nella (1.6c) il secondo integrale è nullo perché è un coseno integrato in un numero intero di periodi.

A questo punto, confrontando la (1.6e) con la (1.5) si ha che:

$$P_{media} = \frac{P_{picco}}{2} \quad (1.7)$$

Abbiamo quindi ottenuto che in un segnale sinusoidale la potenza media è pari a metà della potenza di picco. ■

1.2.3 Segnali a energia finita

È noto dalla fisica che la potenza è pari all'energia spesa in un certo intervallo di tempo:

$$\text{Energia [J]} = \text{Potenza [x]} \cdot \text{tempo [s]}$$

In questo modo, partendo dall'espressione (1.4) della potenza media di un segnale, si può definire l'*energia di un segnale* in questo modo:

$$E_x \triangleq \lim_{T \rightarrow \infty} \int_{-\frac{T}{2}}^{+\frac{T}{2}} |x(t)|^2 dt \quad (1.8)$$

Per i segnali reali basta usare $x^2(t)$ al posto di $|x(t)|^2$.

Un segnale si dice *a energia finita* quando la sua energia (1.8) è finita:

$$E_x < \infty$$

Si noti che:

- ▷ un segnale a energia finita ha sicuramente potenza media nulla;
- ▷ un segnale a potenza media finita deve avere energia infinita.

In pratica, matematicamente, un segnale è ad energia finita oppure a potenza media finita. Mai entrambi.

I segnali fisicamente realizzabili hanno sempre energia finita. Se limitati in un intervallo di tempo finito $[t_1, t_2]$, abbiamo:

$$E_x = \int_{t_1}^{t_2} |x(t)|^2 dt$$

1.2.4 Distribuzione δ di Dirac

È molto usata nell'ambito delle telecomunicazioni. Solitamente si pensa come limite per T che tende a 0 della funzione $p_T(t)/T$ (vedi figura 1.1):

$$\delta(t) = \lim_{T \rightarrow 0} \frac{1}{T} p_T(t)$$

Si tratta di un segnale non realizzabile praticamente. Tuttavia è una buona approssimazione per segnali di durata molto breve e con elevata intensità.

Proprietà principali della δ di Dirac:

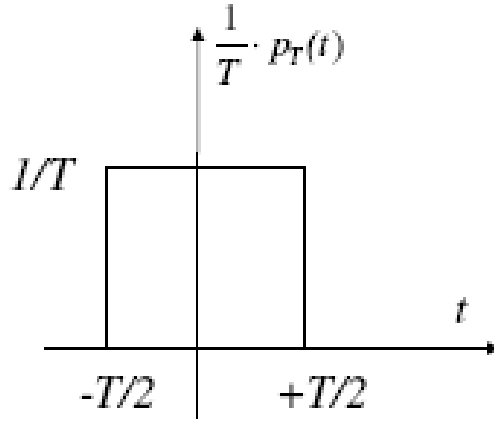


Figura 1.1: Funzione porta rettangolare: facendo tendere T a zero tende alla distribuzione di Dirac.

1. $\delta(t) = 0$ per qualsiasi t diverso da 0.
2. La $\delta(t)$ "campiona" le funzioni per cui viene moltiplicata:

$$f(t) \delta(t - a) = f(a) \delta(t - a) \quad (1.9)$$

3.

$$\int_{-\infty}^{+\infty} \delta(t) dt = 1 \quad (1.10)$$

4.

$$\int_{-\infty}^{+\infty} f(t) \delta(t) dt = f(0) \quad (1.11)$$

5.

$$\int_{-\infty}^{+\infty} e^{\pm j2\pi ft} df = \delta(t) \quad (1.12)$$

1.3 Rappresentazione segnali su basi ortogonali

In questo paragrafo viene introdotta la rappresentazione di un segnale su una base ortonormale. Date due funzioni $\varphi_1(t)$ e $\varphi_2(t)$ definite in un intervallo $[a, b]$, a volte è comodo considerarle come se fossero dei vettori. Si definisce allora il loro prodotto scalare in questo modo:

$$(\varphi_1(t), \varphi_2(t)) \triangleq \int_a^b \varphi_1(t) \varphi_2^*(t) dt$$

Usiamo la generalizzazione in ambito complesso, poiché a volte è comodo lavorare con funzioni complesse. In analogia a quello che succede nella geometria euclidea, effettuare il prodotto scalare tra $\varphi_1(t)$ e $\varphi_2(t)$ equivale a trovare la componente di $\varphi_1(t)$ rispetto a $\varphi_2(t)$.

Diremo allora che due funzioni $\varphi_1(t)$ e $\varphi_2(t)$, definite su un intervallo $[a, b]$, sono *ortogonali* se il loro prodotto scalare è nullo:

$$(\varphi_1(t), \varphi_2(t)) = \int_a^b \varphi_1(t) \varphi_2^*(t) dt = 0$$

Consideriamo ora un insieme di m funzioni $\varphi_i(t)$, con $i = 1, \dots, m$. Un insieme di funzioni sono *ortogonali* su un intervallo $[a, b]$ se:

$$(\varphi_i(t), \varphi_j(t)) = \int_a^b \varphi_i(t) \varphi_j^*(t) dt = 0 \quad \forall i \neq j$$

Se poi le funzioni sono a energia unitaria:

$$\int_a^b |\varphi_i(t)|^2 dt = 1$$

allora vengono dette *ortonormali* e solitamente si indicano come:

$$\hat{\varphi}_i(t)$$

Dato un insieme di funzioni $x_1(t), \dots, x_n(t)$, in alcune applicazioni si è interessati a trovare una base ortonormale completa per questi segnali:

$$x_i(t) = \sum_{j=1}^m a_j \hat{\varphi}_j(t)$$

con $m \leq n$. I coefficienti a_j sono detti coefficienti dell'espansione in serie e si trovano effettuando il seguente prodotto scalare:

$$a_j = (x_i(t), \hat{\varphi}_j^*(t)) = \int_a^b x_i(t) \hat{\varphi}_j^*(t) dt$$

Il metodo per trovare la base a partire dalle funzioni $x_i(t)$ è detto *algoritmo di Gram-Schmidt*.

Una volta nota la base di segnali ortonormali $\hat{\varphi}_1(t), \dots, \hat{\varphi}_m(t)$, ogni segnale $x_i(t)$ può essere rappresentato tramite i suoi coefficienti $[a_1, \dots, a_m]$ come un vettore in uno spazio a m dimensioni. Tipicamente si useranno insiemi di segnali rappresentabili con basi ortogonali di 1 oppure 2 elementi. Si possono quindi rappresentare i punti su una retta o su un piano.

Il tipo di rappresentazione appena introdotto è largamente utilizzato per descrivere gli schemi di modulazioni digitali. Questi schemi di modulazione non verranno trattati in queste dispense. Consigliamo al lettore interessato di fare riferimento a [3] oppure a [4].

Vediamo una proprietà importante. Una volta espresso $x_i(t)$ in termini di componenti rispetto alla base ortonormale:

$$x_i(t) = \sum_{j=1}^m a_j \hat{\varphi}_j(t)$$

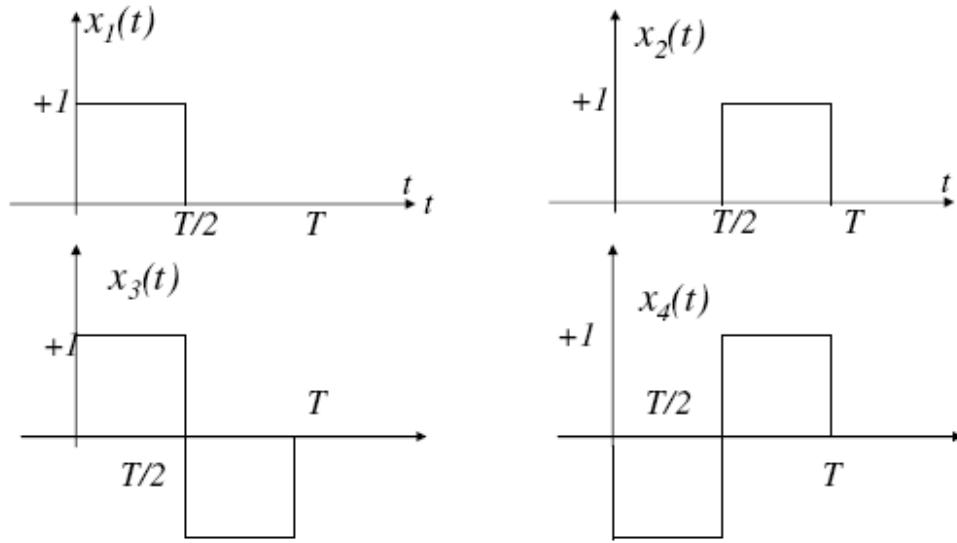


Figura 1.2: Segnali per l'esempio 1.2.

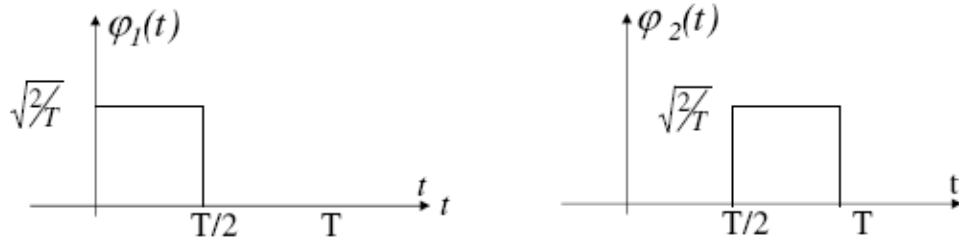


Figura 1.3: Base ortonormale per rappresentare i segnali mostrati in figura 1.2.

si ha che l'energia è pari al quadrato della distanza dall'origine:

$$E_x = \sum_{j=1}^m |a_j|^2$$

Esempio 1.2. Sia dato l'insieme di segnali $x_i(t)$ mostrati in figura 1.2. Calcolare i coefficienti della rappresentazione in termini di basi ortogonali di questi quattro segnali.

Soluzione. Per ispezione diretta dei segnali della figura 1.2, si vede che questi possono essere rappresentati usando la base ortonormale mostrata in figura 1.3. Questi segnali sono sicuramente ortogonali perché definiti in intervalli temporali disgiunti. Si ottiene facilmente che:

$$\begin{cases} x_1(t) = \sqrt{\frac{T}{2}} \varphi_1(t) \\ x_2(t) = \sqrt{\frac{T}{2}} \varphi_2(t) \\ x_3(t) = +\sqrt{\frac{T}{2}} \varphi_1(t) - \sqrt{\frac{T}{2}} \varphi_2(t) \\ x_4(t) = -\sqrt{\frac{T}{2}} \varphi_1(t) + \sqrt{\frac{T}{2}} \varphi_2(t) \end{cases}$$

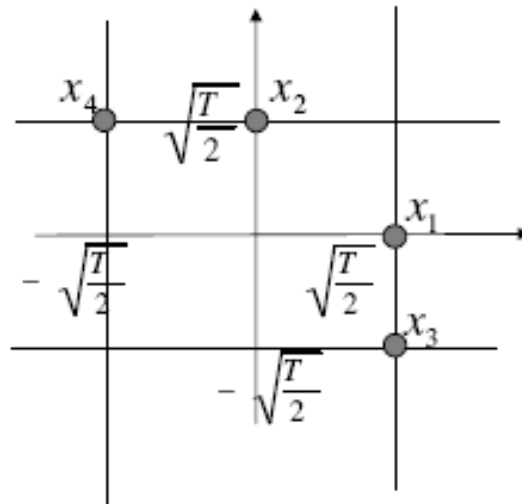


Figura 1.4: Rappresentazione vettoriale dei segnali dell'esempio 1.2.

Quindi i vettori corrispondenti ai segnali $x_1(t)$, $x_2(t)$, $x_3(t)$ e $x_4(t)$ sono i seguenti:

$$\begin{cases} x_1 = \left(\sqrt{\frac{T}{2}}; 0 \right) \\ x_2 = \left(0; \sqrt{\frac{T}{2}} \right) \\ x_3 = \left(+\sqrt{\frac{T}{2}}; -\sqrt{\frac{T}{2}} \right) \\ x_4 = \left(-\sqrt{\frac{T}{2}}; +\sqrt{\frac{T}{2}} \right) \end{cases}$$

Questi quattro vettori sono mostrati in figura 1.4.

Calcoliamo a titolo di esempio l'energia di x_3 . Con la definizione si trova:

$$\begin{aligned} E_{x_3} &= \int_{-\infty}^{+\infty} x_3^2(t) dt \\ &= \int_0^T 1 dt \\ &= T \end{aligned}$$

Il calcolo si può svolgere molto più semplicemente in termini vettoriali:

$$\begin{aligned} E_{x_3} &= \left(\sqrt{\frac{T}{2}} \right)^2 + \left(\sqrt{\frac{T}{2}} \right)^2 \\ &= \frac{T}{2} + \frac{T}{2} \\ &= T \end{aligned}$$

■

1.4 La Serie di Fourier

Consideriamo un segnale periodico $x(t)$ avente periodo T_0 e frequenza $f_0 = 1/T_0$. Se $x(t)$ soddisfa certe condizioni di regolarità (chiamate condizioni di Dirichlet), allora può essere espressa utilizzando la seguente base di segnali ortogonali:

$$\left\{ e^{j2\pi n f_0 t} \right\}_{n=-\infty}^{+\infty} \quad (1.13)$$

Il lettore è invitato a verificare che i segnali espressi dalla (1.13) sono ortogonali.

Lo sviluppo in serie di $x(t)$ rispetto alla base (1.13) viene detto serie di Fourier:

$$x(t) = \sum_{n=-\infty}^{+\infty} \mu_n e^{j2\pi n f_0 t} \quad (1.14)$$

In base a quello che abbiamo detto nella sezione 1.3, i coefficienti μ_n dello sviluppo in serie possono essere calcolati eseguendo il prodotto scalare $(x(t), e^{j2\pi n f_0 t})$ (il termine $1/T_0$ è dovuto al fatto che la base (1.13) non è ortonormale):

$$\mu_n = \frac{1}{T_0} \int_a^{a+T_0} x(t) e^{-j2\pi n f_0 t} dt \quad (1.15)$$

Si noti che:

- ▷ In generale i coefficienti μ_n sono numeri complessi, anche per $x(t)$ reale.
- ▷ Il parametro a presente nei limiti dell'integrale della (1.15) è arbitrario: può essere scelto per semplificare il calcolo dell'integrale. Di solito conviene scegliere $a = 0$ oppure $a = -T_0/2$.
- ▷ Le condizioni di Dirichlet sono solo condizioni sufficienti. Esistono segnali che non soddisfano queste condizioni e che ammettono lo stesso la serie di Fourier.

La serie di Fourier è un risultato molto importante. L'espressione (1.14) permette di esprimere una funzione periodica come somma di infinite funzioni esponenziali complesse $e^{j2\pi n f_0 t}$. La minima frequenza che compare nello sviluppo $f_0 = 1/T_0$ è detta *frequenza fondamentale* del segnale $x(t)$. Le altre frequenze $n f_0$, dette *armoniche* del segnale, sono tutte multiple della fondamentale.

Scrivendo μ_n in questo modo:

$$\mu_n = |\mu_n| e^{j\angle\mu_n}$$

si vede che $|\mu_n|$ fornisce l'ampiezza della componente relativa all' n -esima armonica e $\angle\mu_n$ ne fornisce la fase. È possibile rappresentare l'andamento dei coefficienti complessi μ_n per mezzo di due grafici in cui vengono tracciati modulo e fase di μ_n in funzione del parametro n o della frequenza n/T . Grafici di questo tipo vengono detti *spettro discreto* del segnale periodico $x(t)$ e ne descrivono l'andamento nel dominio della frequenza.

L'operazione di sviluppo in serie di Fourier è chiamata analisi armonica del segnale e riveste un ruolo fondamentale nella teoria delle comunicazioni.

Se $x(t)$ è un segnale reale, si ha che $\mu_{-n} = \mu_n^*$, cioè i coefficienti positivi e negativi sono complessi coniugati: $|\mu_n|$ ha simmetria pari mentre $\angle\mu_n$ ha simmetria dispari.

Questo mi dice anche che per un segnale reale tutta l'informazione necessaria per ricostruire il segnale è contenuta nei coefficienti relativi alle frequenze positive.

Si dimostra inoltre che la potenza del segnale può essere calcolata in questo modo:

$$P_x = \sum_{n=-\infty}^{+\infty} |c_n|^2$$

1.5 La Trasformata di Fourier

La serie di Fourier è un modo di espandere un segnale periodico in termini di funzioni esponenziali complesse. L'idea delle serie di Fourier può essere estesa al caso dei segnali non periodici: il risultato è la trasformata di Fourier, cioè uno spettro in cui le frequenze assumono valori in un insieme continuo.

La trasformata di Fourier di un segnale $x(t)$ è definita in questo modo:

$$X(f) \triangleq \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \quad (1.16)$$

Di solito per indicare che $X(f)$ è la trasformata di Fourier di $x(t)$ si usa questa notazione:

$$X(f) = F[x(t)]$$

La trasformata di Fourier può essere invertita: a partire da $X(f)$ è possibile ricostruire il segnale $x(t)$ effettuando una antitrasformata. L'antitrasformata di Fourier di $X(f)$ è definita in questo modo:

$$x(t) \triangleq \int_{-\infty}^{+\infty} X(f) e^{+j2\pi ft} df \quad (1.17)$$

Per indicare che $x(t)$ è l'antitrasformata di Fourier di $X(f)$ si usa questa notazione:

$$x(t) = F^{-1}[X(f)]$$

Da un punto di vista matematico, la questione dell'esistenza della trasformata di Fourier è una questione abbastanza spinosa. Noi non ci preoccupiamo di questo problema perché tutti i segnali di nostro interesse sono trasformabili e con trasformata invertibile.

Si noti che $X(f)$ è in generale una funzione complessa, anche per $x(t)$ reale. Il significato di $X(f)$ è chiaro dall'espressione dell'antitrasformata (1.17). Da qui si vede infatti che $x(t)$ può essere espressa come somma (integrale) di infinite funzioni esponenziali complesse $e^{+j2\pi ft}$, ciascuna con ampiezza infinitesima pari a

$$X(f) df$$

Il modulo $|X(f)|$ e la fase $\angle X(f)$ rappresentano ampiezza e fase delle componenti in frequenza di $x(t)$. Solitamente in pratica si è più interessati al modulo piuttosto che alla fase.

La funzione $X(f)$ è anche chiamata *spettro* del segnale $x(t)$: si tratta di un tipo di rappresentazione molto utilizzata nell'ambito delle comunicazioni elettriche. Si presti attenzione al fatto che la definizione di trasformata di Fourier $X(f)$ è una definizione matematica, fatta con una trasformazione integrale su un intervallo temporale infinito. Dei segnali utilizzati in pratica raramente si riesce a calcolare $X(f)$ con la definizione e si deve spesso ricorrere a tecniche numeriche (tipicamente FFT).

Esempio 1.3. Calcolare la trasformata di Fourier della seguente funzione, detta esponenziale unilatera decrescente:

$$x(t) = \begin{cases} e^{-at} & \text{per } t \geq 0 \\ 0 & \text{per } t < 0 \end{cases}$$

Soluzione. Utilizzando la definizione (1.16) si ottiene:

$$\begin{aligned} X(f) &= \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{+\infty} e^{-(a+j2\pi f)t} u(t) dt \\ &= \int_0^{+\infty} e^{-(a+j2\pi f)t} dt \\ &= \frac{1}{a + j2\pi f} \end{aligned}$$

In pratica non si fanno quasi mai i calcoli ma si applicano sempre le proprietà delle trasformate e le tavole delle trasformate. ■

Esempio 1.4. Calcolare la trasformata di Fourier della Delta di Dirac $\delta(t)$.

Soluzione. Utilizzando la definizione di trasformata di Fourier (1.16) e la proprietà (1.11) della delta di Dirac (si veda pagina 6) si ottiene:

$$\begin{aligned} X(f) &= \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{+\infty} \delta(t) e^{-j2\pi ft} dt \\ &= e^{-j2\pi ft} \Big|_{t=0} \\ &= 1 \end{aligned}$$

La trasformata di Fourier della Delta di Dirac è una funzione costante in frequenza. Questo è un risultato che viene utilizzato moltissimo nell'ambito delle telecomunicazioni. ■

1.5.1 Proprietà della trasformata di Fourier

Riportiamo ora le principali proprietà della trasformata di Fourier. Tali proprietà si utilizzano di solito per calcolare le trasformate senza bisogno di utilizzare la definizione. Nelle formule che seguono si ha $X(f) = F[x(t)]$ e $Y(f) = F[y(t)]$.

Linearità. La trasformata di Fourier è un operatore lineare. Questo significa che la trasformata di una combinazione lineare di funzioni è uguale alla combinazione lineare delle trasformate delle singole funzioni:

$$F \{a x(t) + b y(t)\} = a X(f) + b Y(f) \quad (1.18)$$

Shift temporale. Uno shift di $\pm\theta$ nel dominio del tempo provoca uno shift di fase di $e^{\pm j2\pi f\theta}$ nel dominio della frequenza:

$$F \{x(t \pm \theta)\} = X(f) e^{\pm j2\pi f\theta} \quad (1.19)$$

Questo risultato mi dice che ritardando un segnale nel dominio del tempo, il modulo della sua trasformata non cambia.

Scalamento asse temporale. Una compressione (allargamento) di $x(t)$ corrisponde a un allargamento (compressione) di $X(f)$

$$F \{x(kt)\} = \frac{1}{|k|} X\left(\frac{f}{k}\right) \quad (1.20)$$

Condizioni di parità. Se $x(t)$ è reale:

- $Re \{X(f)\}$ è pari, $Im \{X(f)\}$ è dispari;
- $|X(f)|$ è pari, $\angle X(f)$ è dispari.

Questa proprietà è importantissima perché stabilisce che se $x(t)$ è reale, tutta l'informazione utile per ricostruire il segnale sta nelle frequenze positive.

Modulazione.

$$F \{x(t) e^{\pm j2\pi f_0 t}\} = X(f \mp f_0) \quad (1.21)$$

$$F \{x(t) \cos(2\pi f_0 t)\} = \frac{1}{2} [X(f - f_0) + X(f + f_0)] \quad (1.22)$$

Nei prossimi capitoli vedremo che queste due proprietà sono fondamentali per lo studio dei sistemi di telecomunicazioni

Convoluzione. La trasformata del prodotto di convoluzione di due funzioni è uguale al prodotto delle trasformate delle due funzioni:

$$F \{x(t) * y(t)\} = X(f) Y(f) \quad (1.23)$$

dove si ricorda che il simbolo $*$ denota il prodotto di convoluzione, definito in questo modo:

$$\begin{aligned} x(t) * y(t) &\triangleq \int_{-\infty}^{+\infty} x(\tau) y(t - \tau) d\tau \\ &\triangleq \int_{-\infty}^{+\infty} y(\tau) x(t - \tau) d\tau \end{aligned}$$

Prodotto. La trasformata del prodotto di due funzioni è uguale al prodotto di convoluzione delle trasformate delle due funzioni:

$$F\{x(t)y(t)\} = X(f) * Y(f) \quad (1.24)$$

Derivazione.

$$F\left\{\frac{d}{dt}x(t)\right\} = j2\pi fX(f) \quad (1.25)$$

Dualità.

$$F\{X(t)\} = x(-f) \quad (1.26)$$

Questa proprietà è utile per usare le tavole delle trasformate e delle anti-trasformate in modo "simmetrico".

1.5.2 Trasformata di Fourier di segnali periodici

Vediamo ora l'espressione della trasformata di Fourier di un segnale periodico. Per ricavarla conviene partire dal segnale $x(t)$ sviluppato in serie di Fourier:

$$x(t) = \sum_{n=-\infty}^{+\infty} \mu_n e^{j2\pi n f_0 t}$$

Utilizzando la proprietà di linearità (1.18), la proprietà (1.21) relativa alla modulazione e il fatto che $F[\delta(t)] = 1$, si ottiene che:

$$X(f) = \sum_{n=-\infty}^{+\infty} \mu_n \delta(f - nf_0) \quad (1.27)$$

L'espressione (1.27) mi dice che un segnale periodico ha uno spettro "a righe": si tratta di un treno di Delta le cui ampiezze sono date dai coefficienti della serie di Fourier del segnale. Da questo si capisce bene come per un segnale periodico la trasformata di Fourier fornisca esattamente le stesse informazioni che fornisce la serie di Fourier.

1.6 Sistemi lineari e tempo invarianti (LTI)

Un gran numero di blocchi che costituiscono un sistema di comunicazione possono essere modellati da sistemi lineari tempo-invarianti (Linear Time Invariant, LTI). In particolare, i sistemi LTI sono modelli accurati per una vasta classe di canali di comunicazione. Alcuni componenti base di trasmettitori e ricevitori, come filtri, amplificatori e equalizzatori, sono sistemi LTI. Per questo motivo i sistemi LTI rivestono un ruolo così importante nelle applicazioni.

La figura 1.5 mostra schematicamente un sistema LTI. Il sistema ha un segnale in ingresso che indichiamo come $x(t)$. Il sistema effettua una trasformazione sul segnale $x(t)$ producendo in uscita un altro segnale $y(t)$. Descriveremo questa trasformazione con la lettera τ :

$$y(t) = \tau[x(t)]$$

Come suggerisce il nome stesso, un sistema LTI è:

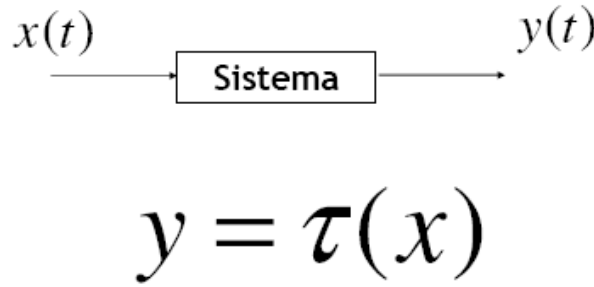


Figura 1.5: Rappresentazione schematica di un sistema lineare e tempo invariante (LTI).

1. **Lineare.** Significa che vale il principio di sovrapposizione degli effetti:

$$\begin{aligned} y(t) &= \tau[a_1 x_1(t) + a_2 x_2(t)] \\ &= a_1 \tau[x_1(t)] + a_2 \tau[x_2(t)] \end{aligned}$$

2. **Tempo invariante.** Significa che se

$$y(t) = \tau[x(t)]$$

allora

$$y(t - t_0) = \tau[x(t - t_0)]$$

È noto che i sistemi LTI possono essere caratterizzati tramite la risposta all'impulso $h(t)$: si tratta dell'uscita del sistema quando all'ingresso è posto $x(t) = \delta(t)$. È anche noto che per un generico sistema LTI l'uscita è data dal prodotto di convoluzione tra l'ingresso e la risposta all'impulso del sistema:

$$\begin{aligned} y(t) &= x(t) * h(t) \\ &= \int_{-\infty}^{+\infty} x(\tau) h(t - \tau) d\tau \\ &= \int_{-\infty}^{+\infty} h(\tau) x(t - \tau) d\tau \end{aligned}$$

Calcolo grafico del prodotto di convoluzione

Per calcolare il prodotto di convoluzione

$$y(t) = \int_{-\infty}^{+\infty} x(\tau) h(t - \tau) d\tau \quad (1.28)$$

con carta e penna, è possibile utilizzare un metodo grafico. Immaginiamo ad esempio di dover effettuare il prodotto di convoluzione tra i due segnali $x(t)$ e $h(t)$ mostrati in figura 1.6. L'espressione (1.28) mi dice che, per ogni valore di t , il valore di $y(t)$ deve essere calcolato tramite un integrale diverso. Per ogni valore di t :

- ▷ si considera la funzione $x(\tau)$ (scrivere $x(\tau)$ o $x(t)$ non cambia niente, la funzione è sempre la stessa: vedi figura 1.7 in basso);

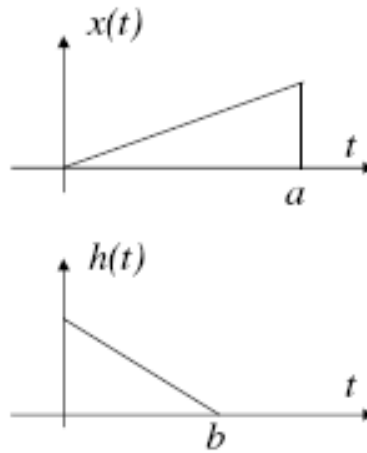


Figura 1.6: Esempio di segnali di cui effettuare prodotto di convoluzione. Si ha $a > b$.

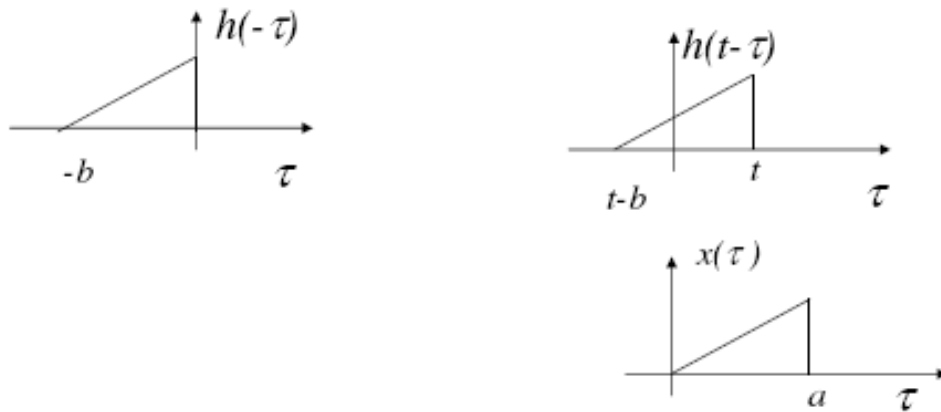


Figura 1.7: Esecuzione del prodotto di convoluzione tra le due funzioni della figura 1.6.

- ▷ si considera la funzione $h(t - \tau)$: si tratta di $h(\tau)$ vista allo specchio e traslata a destra o a sinistra di una quantità t (vedi figura 1.7 in alto);
- ▷ si effettua il prodotto tra $x(\tau)$ e $h(t - \tau)$ e si integra il risultato.

Utilizzando il procedimento appena descritto con i segnali di figura 1.6, si ottiene:

1. Per $t < 0$:

$$y(t) = 0$$

2. Per $0 \leq t \leq b$:

$$y(t) = \int_0^t x(\tau) h(t - \tau) d\tau$$

3. Per $b \leq t \leq a$:

$$y(t) = \int_{t-b}^t x(\tau) h(t - \tau) d\tau$$

4. Per $a \leq t \leq a + b$:

$$y(t) = \int_{t-b}^a x(\tau) h(t - \tau) d\tau$$

5. Per $t > a + b$:

$$y(t) = 0$$

Il calcolo analitico del prodotto di convoluzione è possibile solo per funzioni particolarmente semplici.

1.6.1 Sistemi causali

Dall'espressione del prodotto di convoluzione

$$y(t) = \int_{-\infty}^{+\infty} x(\tau) h(t - \tau) d\tau \quad (1.29)$$

è chiaro che, per segnali $x(t)$ e $h(t)$ generici, l'uscita del sistema a un istante t dipende sia dai valori passati che dai valori futuri dell'ingresso. Tuttavia, i sistemi che si possono implementare in pratica sono tutti *causali*: significa che, dato $x(t)$ nullo per $t < t_0$, allora l'uscita del sistema $y(t)$ è anch'essa nulla per $t < t_0$. In pratica si ha che l'uscita non precede l'ingresso.

Si può dimostrare che un sistema è causale se

$$h(t) = 0 \text{ per } t < 0$$

In effetti se si considera l'espressione (1.29), si vede che in questo caso l'uscita non dipende dai valori futuri dell'ingresso.

Osserviamo infine ci sono casi particolari in cui è comodo assumere che il sistema non sia causale: questo viene ad esempio fatto nelle modulazioni numeriche riguardo all'*interferenza intersimbolica*.

1.6.2 Risposta in frequenza

Consideriamo la relazione ingresso-uscita di un sistema LTI:

$$y(t) = x(t) * h(t)$$

Lavorando con le trasformate di Fourier, si può utilizzare la proprietà (1.23) (vedi pagina 13) e ottenere il seguente legame tra le trasformate:

$$Y(f) = X(f) \cdot H(f) \quad (1.30)$$

L'espressione (1.30) è importantissima e giustifica l'utilizzo della trasformata di Fourier per l'analisi dei sistemi lineari. Infatti, lavorando nel dominio trasformato, l'operazione di convoluzione (che in generale è complicata da svolgere) diventa un semplice prodotto algebrico.

La funzione

$$H(f) = F[h(t)]$$

viene detta **funzione di trasferimento** del sistema.

Risposta a segnali sinusoidali

Vediamo ora qual'è l'uscita di un generico sistema LTI quando in ingresso viene posto un segnale sinusoidale:

$$x(t) = A_p \cos(2\pi f_0 t)$$

Utilizzando le proprietà di linearità (1.18) e di modulazione (1.22) della trasformata di Fourier (vedi pagina 13), si ottiene che la trasformata di $x(t)$ vale:

$$X(f) = \frac{A_p}{2} [\delta(f - f_0) + \delta(f + f_0)]$$

La trasformata dell'uscita può essere calcolata mediante l'espressione (1.30):

$$\begin{aligned} Y(f) &= H(f) \cdot X(f) \\ &\stackrel{1}{=} \frac{A_p}{2} H(f) [\delta(f - f_0) + \delta(f + f_0)] \\ &\stackrel{2}{=} \frac{A_p}{2} [H(f_0) \delta(f - f_0) + H(-f_0) \delta(f + f_0)] \\ &\stackrel{3}{=} \frac{A_p}{2} [H(f_0) \delta(f - f_0) + H^*(f_0) \delta(f + f_0)] \end{aligned} \quad (1.31)$$

Nell'uguaglianza 2 si è sfruttata la proprietà (1.9) della Delta (si veda pagina 6). L'uguaglianza 3 deriva dal fatto che se $h(t)$ è reale, allora $H(-f) = H^*(f)$.

Per trovare l'espressione $y(t)$ dell'uscita, bisogna effettuare l'antitrasformata di Fourier della (1.31):

$$\begin{aligned} y(t) &= \frac{A_p}{2} [H(f_0) e^{+j2\pi f_0 t} + H^*(f_0) e^{-j2\pi f_0 t}] \\ &= \frac{A_p}{2} [|H(f_0)| e^{+j(2\pi f_0 t + \arg(H(f_0)))} + |H(f_0)| e^{-j(2\pi f_0 t + \arg(H(f_0)))}] \\ &= A_p |H(f_0)| \cos(2\pi f_0 t + \arg(H(f_0))) \end{aligned} \quad (1.32)$$

L'espressione (1.32) è importantissima per i sistemi lineari. Mi dice che, se in ingresso a un sistema lineare viene posto un ingresso sinusoidale con frequenza f_0 , l'uscita è sempre sinusoidale e con la stessa frequenza. Ciò che il sistema cambia sono l'ampiezza e la fase del segnale: l'ampiezza viene moltiplicata per il valore di $H(f)$ calcolato in f_0 , mentre la fase viene sommata alla fase di $H(f)$ calcolata in f_0 .

Per questo motivo la funzione di trasferimento $H(f)$ viene anche detta **risposta in frequenza**: i suoi valori mi dicono di quanto vengono amplificate/attenuate e sfasate le componenti dei segnali sinusoidali in ingresso.

1.6.3 Sistemi LTI non distortenti

Un sistema di trasmissione LTI è detto **non distortente** se, dato un ingresso $x(t)$, l'uscita $y(t)$ è una *replica fedele* dell'ingresso. Replica fedele significa che $y(t)$ è uguale a $x(t)$, eventualmente amplificato o ritardato di una quantità costante:

$$y(t) = A_p x(t - T_d) \quad \text{con } A_p, T_d > 0 \quad (1.33)$$

Si parla di replica fedele perché $x(t)$ e $y(t)$ hanno la stessa forma.

Vediamo quali caratteristiche deve avere la funzione di trasferimento di un sistema LTI non distortore. Conviene scrivere l'espressione (1.33) nel dominio della frequenza. Utilizzando la proprietà (1.19) dello shift temporale della trasformata di Fourier (si veda pagina 13) si trova:

$$Y(f) = A_p X(f) e^{-j2\pi f T_d}$$

quindi la risposta in frequenza del sistema deve valere:

$$\begin{aligned} H(f) &= \frac{Y(f)}{X(f)} \\ &= A_p e^{-j2\pi f T_d} \end{aligned} \quad (1.34)$$

Le condizioni su $H(f)$ affinché il sistema sia non distortore sono dunque:

1. Il modulo di $H(f)$ è costante e non dipende dalla frequenza;
2. La fase di $H(f)$ è uguale a kf (k è una costante arbitraria), cioè è una funzione lineare con la frequenza.

È chiaro che nessun sistema reale ha risposta in frequenza data dalla (1.34): questo presupporrebbe avere una banda infinita. Invece tutti i sistemi reali sono a banda finita. In pratica basta che le due condizioni siano soddisfatte sulla banda del segnale d'ingresso $X(f)$.

1.7 Densità spettrale di energia e di potenza

1.7.1 Densità spettrale di energia

Il teorema di Parseval mi dice che per un segnale $x(t)$ ad energia finita si ha:

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 dt = \int_{-\infty}^{+\infty} |X(f)|^2 df \quad (1.35)$$

Il modulo della trasformata di Fourier al quadrato $|X(f)|^2$ mi dice cioè quanto vale il contributo delle varie componenti in frequenza all'energia complessiva del segnale $x(t)$. La quantità $|X(f)|^2$ viene allora chiamata *spettro di energia* del segnale e indicata in questo modo:

$$\varepsilon_x(f) \triangleq |X(f)|^2 \quad (1.36)$$

Lo spettro di energia (1.36) di un segnale reale ha tutta una serie di proprietà utili:

1. È una funzione sempre positiva e pari;
2. L'energia del segnale si calcola come:

$$E_x = \int_{-\infty}^{+\infty} \varepsilon_x(f) df$$

3. Per sistemi LTI con funzione di trasferimento $H(f)$ si ha (vedi figura 10.3):

$$\varepsilon_y(f) = |H(f)|^2 \varepsilon_x(f)$$

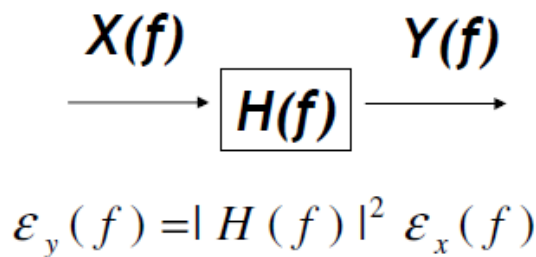


Figura 1.8: Relazione tra le densità spettrali di energia in un sistema LTI.

1.7.2 Densità spettrale di potenza

Nel campo delle telecomunicazioni è più frequente lavorare con segnali a potenza media finita piuttosto che su segnali a energia finita. Per questo tipo di segnali è necessario lavorare sulle potenze invece che sulle energie.

La *densità spettrale di potenza* viene definita in questo modo:

1. Sia $x_T(t)$ una versione di $x(t)$ troncata a T , cioè: $x_T(t) = p_T(t) x(t)$.
2. Sia: $X_T(f) = F[x_T(t)]$
3. La densità spettrale di potenza è definita così:

$$G_x(f) \triangleq \lim_{T \rightarrow \infty} \frac{|X_T(f)|^2}{T} \quad (1.37)$$

È utile grazie alle sue proprietà, analoghe a quelle dello spettro di energia:

1. $G_x(f)$ è sempre positiva e pari;
2. La potenza del segnale si calcola come:

$$P_x = \int_{-\infty}^{+\infty} G_x(f) df$$

3. Per sistemi LTI con funzione di trasferimento $H(f)$ si ha:

$$G_y(f) = |H(f)|^2 G_x(f)$$

Segnali periodici. Per i segnali periodici lo spettro di potenza assume la seguente forma:

$$G_x(f) = \sum_{n=-\infty}^{+\infty} |\mu_n|^2 \delta(f - nf_0) \quad (1.38)$$

dove μ_n sono i coefficienti della serie di Fourier.

1.7.3 Funzione di autocorrelazione

Per un segnale a potenza media finita la *funzione di autocorrelazione* $R_x(\tau)$ è definita come l'antitrasformata di Fourier dello spettro di potenza:

$$R_x(\tau) \triangleq F^{-1} \{G_x(f)\}$$

Si può dimostrare che la funzione di autocorrelazione può essere calcolata mediante la seguente media temporale:

$$\begin{aligned} R_x(\tau) &\triangleq \langle x(t) x(t + \tau) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t) x(t + \tau) dt \end{aligned} \quad (1.39)$$

Il risultato (1.39) prende il nome di *teorema di Wiener-Khinchin*. Può essere utilizzato come metodo per calcolare $G_x(f)$.

Esempio 1.5. Vogliamo calcolare lo spettro di potenza del segnale sinusoidale

$$x(t) = A_p \sin(2\pi f_0 t)$$

Convieni calcolare $G_x(f)$ passando da $R_x(\tau)$:

$$\begin{aligned} R_x(\tau) &= \langle A_p \sin(2\pi f_0 t) A_p \sin[2\pi f_0(t + \tau)] \rangle \\ &= A_p^2 \langle \sin(2\pi f_0 t) \sin[2\pi f_0(t + \tau)] \rangle \\ &= \frac{A_p^2}{2} \langle \cos(2\pi f_0 \tau) - \cos[2\pi f_0(2t + \tau)] \rangle \\ &= \frac{A_p^2}{2} \{ \langle \cos(2\pi f_0 \tau) \rangle - \langle \cos[2\pi f_0(2t + \tau)] \rangle \} \end{aligned} \quad (1.40)$$

Calcoliamo ora i due termini presenti nell'espressione (1.40):

$$\begin{aligned} \langle \cos(2\pi f_0 \tau) \rangle &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \cos(2\pi f_0 \tau) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \cos(2\pi f_0 \tau) \int_{-\frac{T}{2}}^{+\frac{T}{2}} dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \cos(2\pi f_0 \tau) T \\ &= \cos(2\pi f_0 \tau) \end{aligned} \quad (1.41)$$

$$\begin{aligned} \langle \cos[2\pi f_0(2t + \tau)] \rangle &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \cos[2\pi f_0(2t + \tau)] dt \\ &= 0 \end{aligned} \quad (1.42)$$

Sostituendo le espressioni (1.41) e (1.42) nella (1.40) otteniamo finalmente l'espressione della funzione di autocorrelazione:

$$R_x(\tau) = \frac{A_p^2}{2} \cos(2\pi f_0 \tau)$$

Quindi lo spettro di potenza vale:

$$\begin{aligned} G_x(f) &= F[R_x(\tau)] \\ &= \frac{A_p^2}{4} [\delta(f - f_0) + \delta(f + f_0)] \end{aligned}$$

La potenza ora può essere calcolata semplicemente come:

$$\begin{aligned} P_x &= \int_{-\infty}^{+\infty} G_x(f) df \\ &= \frac{A_p^2}{2} \end{aligned}$$

■

È possibile definire una funzione di autocorrelazione anche per segnali a energia finita, tuttavia si tratta di una definizione poco utile per le applicazioni. Il lettore interessato può fare riferimento al testo [1].

1.8 Teorema del campionamento

Al giorno d'oggi, con la diffusione delle tecnologie digitali, gran parte dell'elaborazione e trasmissione dei segnali viene effettuata per via numerica. I segnali analogici vengono convertiti in segnali digitali, che vengono poi elaborati e infine riconvertiti in segnali analogici di cui l'utente finale può usufruire.

In questo contesto vogliamo studiare la possibilità di *campionare* un segnale, cioè di trasformarlo in una sequenza di numeri da cui poi si possa ricavare nuovamente il segnale originario. Questa teoria va sotto il nome di *teorema del campionamento* e costituisce la base delle conversioni analogico-digitali.

1.8.1 Campionamento ideale

Consideriamo un segnale $x(t)$ limitato in banda, cioè tale per cui:

$$X(f) = 0 \quad \text{per } |f| > B$$

dove B è detta *banda assoluta* del segnale $x(t)$ (vedi figura 1.9). Vogliamo trasformare $x(t)$ in una sequenza di numeri $x[i]$ e fare in modo di poter poi ricostruire $x(t)$ a partire dai campioni di $x[i]$. Per fare questo occorre *campionare* il segnale $x(t)$ con un periodo di campionamento T_c , cioè prelevare un campione del segnale ogni T_c secondi.

Il metodo *matematico* più semplice per campionare un segnale e di effettuare la moltiplicazione per un treno di Delta di Dirac:

$$\Delta(t) = T_c \cdot \sum_{i=-\infty}^{i=+\infty} \delta(t - iT_c)$$

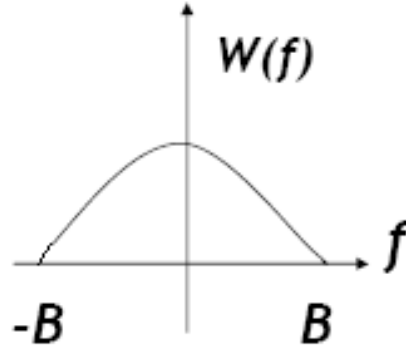


Figura 1.9: Esempio di segnale a banda limitata.

Moltiplicando il segnale $x(t)$ per $\Delta(t)$ si ottiene il *segnale campionato* $x_\delta(t)$:

$$x_\delta(t) = T_c \cdot x(t) \cdot \sum_{i=-\infty}^{i=+\infty} \delta(t - iT_c) \quad (1.43)$$

$$= T_c \cdot \sum_{i=-\infty}^{+\infty} x(iT_c) \cdot \delta(t - iT_c) \quad (1.44)$$

Il segnale (1.44) contiene solo i campioni di $x(t)$ agli istanti iT_c . Il tempo T_c viene detto *periodo di campionamento*, mentre $f_c = 1/T_c$ viene detta *frequenza di campionamento*. A questo punto il segnale campionato $x(iT_c)$ è fatto da una sequenza di numeri, che di solito si indica come $x[i]$.

1.8.2 Ricostruzione del segnale campionato

Per capire sotto quali condizioni $x(t)$ può essere ricostruito a partire dal segnale campionato $x_\delta(t)$, conviene lavorare nel dominio della frequenza. Trasformando secondo Fourier la (1.43) si ottiene:

$$X_\delta(f) = T_c \cdot X(f) * F \left[\sum_{i=-\infty}^{+\infty} \delta(t - iT_c) \right] \quad (1.45)$$

La trasformata di un treno di Delta spaziate di T_c nel tempo è pari a un treno di Delta spaziate di $1/T_c$ in frequenza:

$$F \left[T_c \cdot \sum_{i=-\infty}^{+\infty} \delta(t - iT_c) \right] = \sum_{i=-\infty}^{+\infty} \delta \left(f - \frac{i}{T_c} \right)$$

quindi la (1.45) diventa:

$$\begin{aligned} X_\delta(f) &= X(f) * \sum_{i=-\infty}^{+\infty} \delta \left(f - \frac{i}{T_c} \right) \\ &= \sum_{i=-\infty}^{+\infty} X \left(f - \frac{i}{T_c} \right) \end{aligned} \quad (1.46)$$

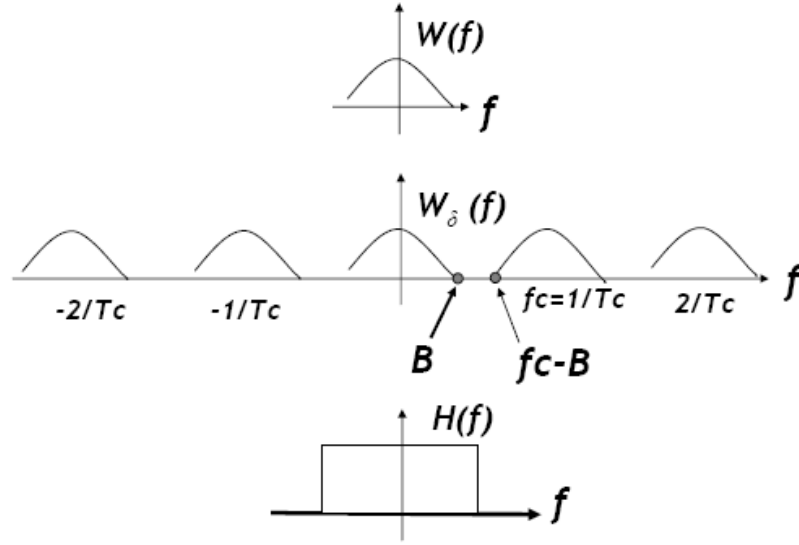


Figura 1.10: Esempio di spettro di segnale analogico (in alto), spettro del relativo segnale campionato (in mezzo) e maschera del filtro ricostruttore (in basso).



Figura 1.11: Filtro ricostruttore.

La (1.46) è l'espressione dello spettro del segnale campionato. Si tratta di uno spettro periodico ottenuto come somma di tutte le repliche di $X(f)$ centrate attorno ai multipli di $f_c = 1/T_c$. La figura 1.10 ne mostra un esempio.

Dalla figura 1.10 si vede che, per poter ricostruire esattamente il segnale $x(t)$, è necessario che le repliche adiacenti dello spettro non si sovrappongano. Bisogna quindi avere:

$$f_c - B \geq B$$

cioè:

$$f_c \geq 2B \quad (1.47)$$

La frequenza di campionamento minima è pari al doppio della banda del segnale $x(t)$. Questo è il cosiddetto teorema di Nyquist.

La figura 1.10 suggerisce anche un metodo per ricostruire il segnale $x(t)$ (almeno dal punto di vista teorico): basta filtrare $X_\delta(f)$ con un filtro passa-basso ideale con banda B (vedi figura 1.11):

$$X(f) = X_\delta(f) \cdot H(f)$$

Il filtraggio col filtro ricostruttore corrisponde a questa operazione nel dominio del

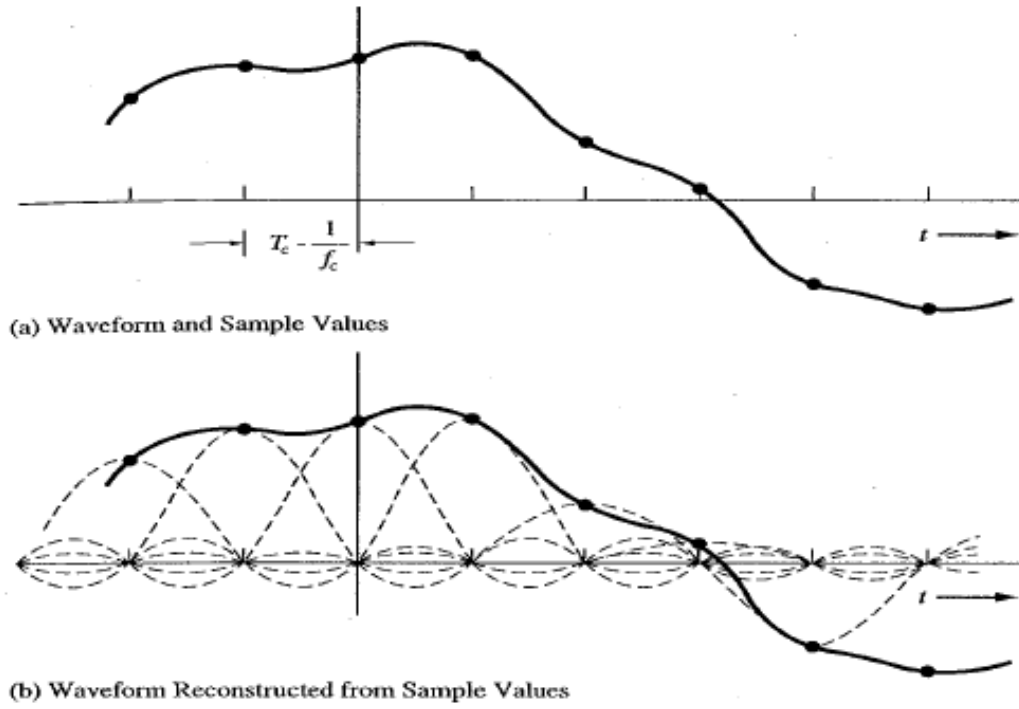


Figura 1.12: Illustrazione ricostruzione segnale campionato.

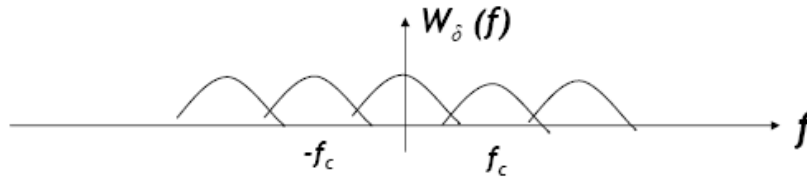


Figura 1.13: Esempio di spettro di segnale affetto da aliasing.

tempo (si veda la figura 1.12):

$$x(t) = \sum_{i=-\infty}^{+\infty} x(iT_c) \frac{\sin\left(\frac{\pi t}{T_c} - i\pi\right)}{\left(\frac{\pi t}{T_c} - i\pi\right)}$$

Se si sceglie una frequenza di campionamento $f_c < 2B$, si dice che il segnale è *sottocampionato*: succede quanto mostrato in figura 1.13. Si viene a formare una zona di sovrapposizione che va sotto il nome di *aliasing* ed il segnale $x(t)$ non è più ricostruibile esattamente. In pratica il fenomeno dell'aliasing è piuttosto fastidioso e bisogna sempre cercare di evitarlo.

In questo paragrafo abbiamo richiamato i principi della teoria del campionamento. Abbiamo trattato il campionamento ideale con il treno di Delta di Dirac, che non è fisicamente realizzabile. In pratica per campionare si utilizzano funzioni più regolari, però il principio di funzionamento di base non cambia molto. Per una trattazione più approfondita di questi argomenti si rimanda il lettore al testo [5].

Capitolo 2

Richiami sui processi casuali

In questo capitolo vengono richiamati i concetti principali sulla teoria della probabilità e sui processi casuali. Si tratta di strumenti fondamentali per affrontare lo studio di sistemi di telecomunicazioni.

2.1 Teoria della probabilità

La motivazione di base che ha portato alla definizione della teoria della probabilità è riconducibile all'esistenza di esperimenti casuali (detti anche esperimenti aleatori), cioè di esperimenti di cui non è possibile conoscere l'esito a priori. Esempi di esperimenti casuali sono:

- ▷ Il lancio di una moneta o di un dado.
- ▷ Il numero di lanci di una moneta necessari per ottenere testa per la prima volta.
- ▷ La lunghezza di una barra d'acciaio tagliata da un macchinario soggetto a imperfezioni.
- ▷ La trasmissione di un bit su un sistema di trasmissione digitale.

Rivediamo ora i risultati fondamentali della teoria della probabilità.

- ▷ La probabilità di un evento A è indicata come $P(A)$ ed è un numero sempre compreso tra 0 e 1:

$$0 \leq P(A) \leq 1 \quad (2.1)$$

- ▷ Dati due eventi A e B , la loro *probabilità congiunta* è la probabilità che si verifichi sia A che B . La probabilità congiunta viene indicata come

$$P(A \cap B)$$

oppure come

$$P(A, B)$$

- ▷ Due eventi A e B si dicono *disgiunti* se la loro probabilità congiunta è nulla:

$$P(A \cap B) = 0$$

- ▷ La probabilità dell'*evento unione* $A \cup B$ (avviene A oppure B) vale:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (2.2)$$

Se A e B sono disgiunti la probabilità dell'evento unione è uguale alla somma delle probabilità di A e B :

$$P(A \cup B) = P(A) + P(B)$$

- ▷ Dati due eventi A e B , la probabilità che si verifichi A sapendo che è avvenuto l'evento B viene detta *probabilità condizionata* e indicata come $P(A|B)$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.3)$$

Se A e B sono disgiunti si ha:

$$P(A|B) = 0$$

- ▷ Due eventi A e B si dicono *statisticamente indipendenti* se il realizzarsi o meno di uno dei due non modifica la probabilità di realizzazione dell'altro:

$$P(A|B) = P(A)$$

che equivale a dire che la probabilità congiunta è uguale al prodotto delle singole probabilità:

$$P(A \cap B) = P(A) P(B)$$

- ▷ *Teorema della probabilità totale.* Siano dati n eventi B_i disgiunti ed esaustivi. Per qualsiasi evento A si può scrivere che:

$$P(A) = \sum_{i=1}^n P(A|B_i) P(B_i) \quad (2.4)$$

- ▷ *Teorema di Bayes.* Serve a calcolare la probabilità di $A|B$ a partire dalla probabilità di $B|A$:

$$P(A|B) = \frac{P(A) P(B|A)}{P(B)} \quad (2.5)$$

2.1.1 Variabili casuali

Solitamente, per studiare un fenomeno aleatorio, si lavora con le variabili casuali. Una *variabile casuale* è una funzione che assume, a seconda del risultato dell'esperimento in esame, valori numerici a cui è possibile attribuire una certa probabilità di realizzazione.

Ad esempio, se consideriamo il lancio di un dado non truccato, possiamo descrivere l'esperimento con una variabile casuale che può assumere valori da 1 a 6, ognuno con probabilità $1/6$. Se si verifica l'evento "è uscito 1" la variabile casuale assume valore 1, se si verifica l'evento "è uscito 2" la variabile casuale assume valore 2 e così via con tutti gli altri numeri.

Le variabili casuali possono essere discrete o continue e si indicano solitamente con lettere greche o con lettere minuscole. Nel seguito ci concentreremo esclusivamente su variabili casuali continue. Una variabile casuale continua può essere descritta mediante la sua distribuzione cumulativa oppure, equivalentemente, con la sua densità di probabilità.

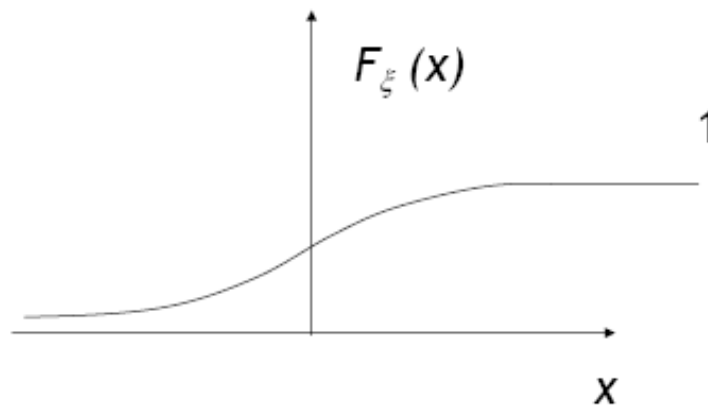


Figura 2.1: Esempio di distribuzione cumulativa.

Funzione di distribuzione cumulativa

Data una variabile casuale ξ , la sua *distribuzione cumulativa* $F_\xi(x)$ mi dice quanto vale la probabilità che ξ assuma un valore minore o uguale a x :

$$F_\xi(x) = P(\xi \leq x) \quad (2.6)$$

La distribuzione cumulativa ha le seguenti proprietà:

- ▷ $F_\xi(x)$ è non decrescente;
- ▷ $\lim_{x \rightarrow -\infty} F_\xi(x) = 0$;
- ▷ $\lim_{x \rightarrow +\infty} F_\xi(x) = 1$.

La figura 2.1 mostra un esempio di distribuzione cumulativa.

Densità di probabilità

La *densità di probabilità* $f_\xi(x)$ di una variabile casuale è data dalla derivata della distribuzione cumulativa:

$$f_\xi(x) = \frac{dF_\xi(x)}{dx}$$

La densità di probabilità ha le seguenti proprietà:

- ▷ È sempre non nulla:

$$f_\xi(x) \geq 0$$

- ▷ Il suo integrale vale 1 (probabilità di un evento certo):

$$\int_{-\infty}^{+\infty} f_\xi(x) dx = 1$$

- ▷ La probabilità che la variabile casuale appartenga a un intervallo $[a, b]$ è pari all'integrale tra a e b della densità di probabilità:

$$P\{\xi \in [a, b]\} = \int_a^b f_\xi(x) dx$$

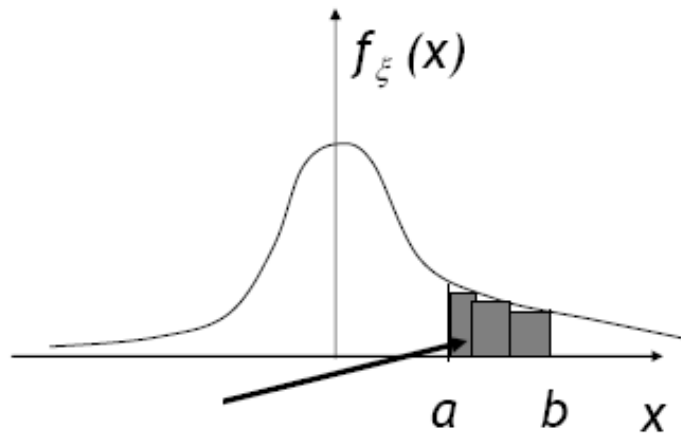


Figura 2.2: Esempio di densità di probabilità.

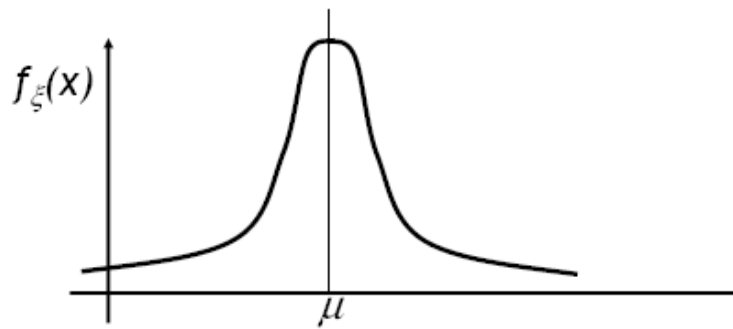


Figura 2.3: Densità di probabilità gaussiana.

La figura 9.14 mostra un esempio di densità di probabilità.

Vediamo ora due tipologie di variabili casuali molto utilizzate nelle applicazioni: la gaussiana e la variabile casuale uniforme.

Variabile casuale gaussiana

La variabile casuale gaussiana ha la seguente densità di probabilità (vedi figura 2.3):

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

dove:

- ▷ μ è il valore medio;
- ▷ σ^2 è la varianza della gaussiana, indice della sua "larghezza".

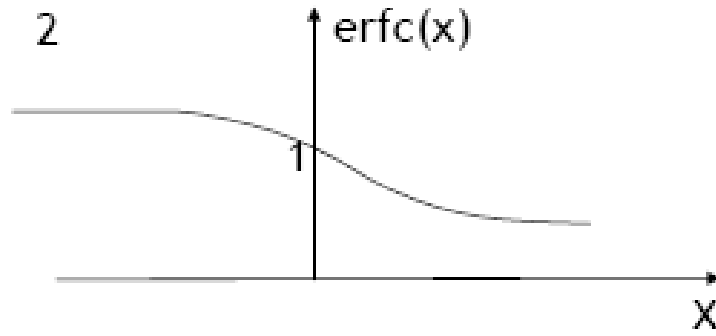


Figura 2.4: Andamento della funzione $\text{erfc}()$.

La distribuzione cumulativa della gaussiana è data da:

$$\begin{aligned} F_{\xi}(x) &= \int_{-\infty}^x f_{\xi}(t) dt \\ &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \end{aligned} \quad (2.7)$$

L'integrale (2.7) non può essere svolto per via analitica ma è disponibile tabulato o tramite algoritmi di calcolo approssimati.

Spesso al posto di $F_{\xi}(x)$ si utilizza la *funzione complementare di errore* $\text{erfc}()$, definita in questo modo:

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt \quad (2.8)$$

Anche questa funzione non è calcolabile analiticamente ma è disponibile tabulata o tramite algoritmi di calcolo approssimati. L'andamento della funzione $\text{erfc}()$ è mostrato in figura 2.4. Commenti sulla funzione $\text{erfc}()$:

- ▷ Tale funzione interessa soprattutto per argomenti positivi;
- ▷ Una approssimazione della $\text{erfc}(x)$:

$$\text{erfc}(x) \approx \frac{1}{\sqrt{\pi}} \frac{e^{-x^2}}{x}$$

Utilizzando la funzione $\text{erfc}()$ possiamo ad esempio calcolare (vedi figura 2.5):

$$\begin{aligned} P(\xi < a) &= \frac{1}{2} \text{erfc}\left(\frac{\mu - a}{\sqrt{2}\sigma}\right) \\ &= \frac{1}{2} \text{erfc}\left(\frac{d}{\sqrt{2}\sigma}\right) \end{aligned}$$

oppure:

$$P(\xi > a) = 1 - \frac{1}{2} \text{erfc}\left(\frac{\mu - a}{\sqrt{2}\sigma}\right)$$

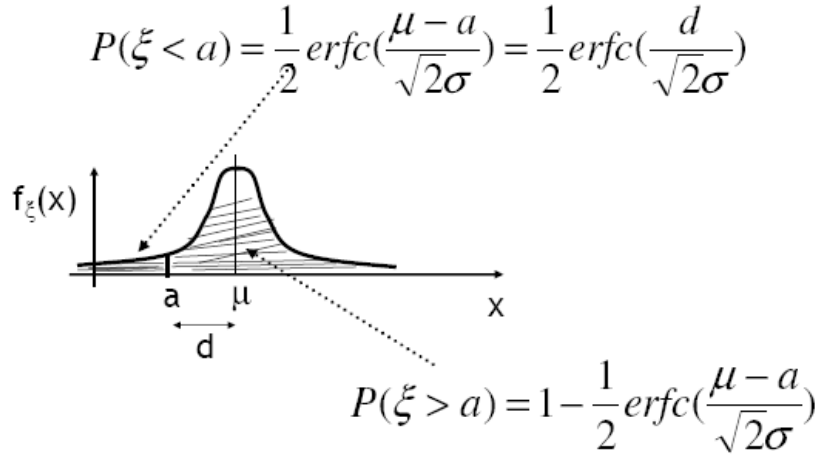
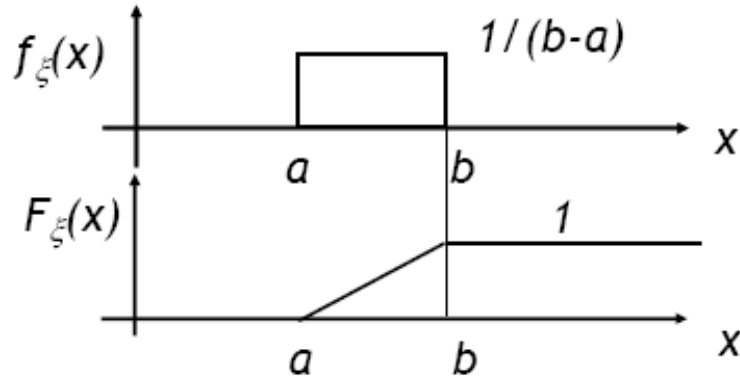
Figura 2.5: Funzione $\operatorname{erfc}()$.

Figura 2.6: Densità di probabilità di variabile casuale uniforme (in alto) e relativa distribuzione cumulativa (in basso).

Variabile casuale uniforme

La figura 9.7 mostra la distribuzione cumulativa e la densità di probabilità di una variabile casuale uniformemente distribuita.

Calcoliamo $P(\xi > c)$:

▷ Per $c < a$:

$$P(\xi > c) = 1$$

▷ Per $a < c < b$:

$$\begin{aligned} P(\xi > c) &= \int_c^b f_\xi(x) dx \\ &= \frac{b - c}{b - a} \end{aligned}$$

▷ Per $c > b$:

$$P(\xi > c) = 0$$

2.1.2 Coppie di variabili casuali

Quando si tratta più di una variabile casuale, è utile introdurre il concetto di probabilità congiunta. La *distribuzione cumulativa congiunta* di due variabili casuali è definita in questo modo:

$$F_{\xi\eta}(x, y) = P(\xi \leq x; \eta \leq y)$$

Proprietà di $F_{\xi\eta}(x, y)$:

- ▷ È non decrescente in x e y ;
- ▷ $F_{\xi\eta}(-\infty, -\infty) = 0$, $F_{\xi\eta}(+\infty, +\infty) = 1$;
- ▷ Le cumulative marginali sono:

$$F_{\xi\eta}(x, +\infty) = F_{\xi}(x)$$

$$F_{\xi\eta}(+\infty, y) = F_{\eta}(y)$$

La densità di probabilità congiunta è pari alla derivata della distribuzione cumulativa congiunta:

$$f_{\xi\eta}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{\xi\eta}(x, y)$$

Due variabili casuali si dicono statisticamente indipendenti se la loro densità di probabilità congiunta è uguale al prodotto delle densità marginali:

$$f_{\xi\eta}(x, y) = f_{\xi}(x) f_{\eta}(y)$$

2.1.3 Momenti

Le variabili casuali si caratterizzano mediante i momenti. Il momento di ordine k di una variabile casuale è definito in questo modo:

$$\begin{aligned} \mu_k &\triangleq \int_{-\infty}^{+\infty} x^k f_{\xi}(x) dx \\ &\triangleq E[\xi^k] \end{aligned}$$

Due momenti molto utilizzati sono:

- ▷ Media:

$$\mu_1 = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx$$

- ▷ Valore quadratico medio:

$$\mu_2 = \int_{-\infty}^{+\infty} x^2 f_{\xi}(x) dx$$

Inoltre si usano:

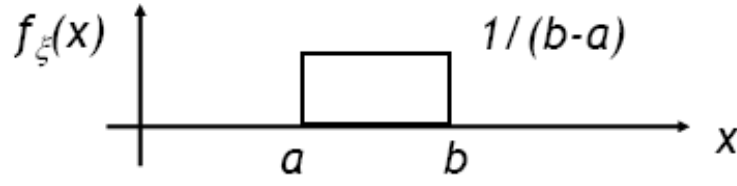


Figura 2.7: Densità di probabilità uniforme per l'esempio 2.1.

▷ Varianza: è il momento centrato di ordine 2.

$$\begin{aligned}
 \sigma_\xi^2 &\triangleq E[(\xi - \mu_1)^2] \\
 &= \int_{-\infty}^{+\infty} (x - \mu_1)^2 f_\xi(x) dx \\
 &= E[\xi^2] - (E[\xi])^2
 \end{aligned}$$

▷ Deviazione standard: è la radice quadrata della varianza.

$$\begin{aligned}
 \sigma_\xi &= \sqrt{\sigma_\xi^2} \\
 &= \sqrt{E[(\xi - \mu_1)^2]}
 \end{aligned}$$

Esempio 2.1 (Calcolo momenti di variabile casuale uniforme). Sia ξ una variabile casuale uniforme in $[a, b]$ (vedi figura 2.7). Calcoliamone i momenti. Il momento di ordine k vale:

$$\begin{aligned}
 E[\xi^k] &= \int_{-\infty}^{+\infty} x^k f_\xi(x) dx \\
 &= \int_a^b x^k \frac{1}{b-a} dx \\
 &= \frac{1}{b-a} \left[\frac{x^{k+1}}{k+1} \right]_a^b \\
 &= \frac{1}{b-a} \left(\frac{b^{k+1} - a^{k+1}}{k+1} \right)
 \end{aligned} \tag{2.9}$$

Dalla (2.9) si possono ricavare tutti i momenti che servono. La media (momento di ordine 1) vale:

$$\begin{aligned}
 E[\xi^k]_{k=1} &= \frac{1}{b-a} \left(\frac{b^2 - a^2}{2} \right) \\
 &= \frac{b+a}{2}
 \end{aligned}$$

Il valor quadratico medio (momento di ordine 2) vale:

$$E[\xi^k]_{k=2} = \frac{1}{b-a} \left(\frac{b^3 - a^3}{3} \right)$$

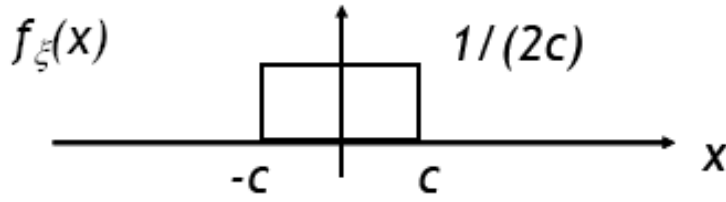


Figura 2.8: Densità di probabilità uniforme simmetrica per l'esempio 2.1.

La varianza (momento centrato di ordine 2) vale:

$$\begin{aligned}\sigma_{\xi}^2 &= E[\xi^2] - (E[\xi])^2 \\ &= \frac{b^3 - a^3}{3(b-a)} - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{(b-a)^2}{12}\end{aligned}$$

Consideriamo ora il caso in cui la densità di probabilità uniforme sia simmetrica (vedi figura 2.8). Il valor medio vale:

$$\begin{aligned}E[\xi^k]_{k=1} &= \left(\frac{b+a}{2}\right)_{a=-c, b=c} \\ &= 0\end{aligned}$$

Il valor quadratico medio vale:

$$\begin{aligned}E[\xi^k]_{k=2} &= \sigma_{\xi}^2 \\ &= \frac{1}{b-a} \left(\frac{b^3 - a^3}{3}\right)_{a=-c, b=c} \\ &= \frac{2c^3}{3(2c)} \\ &= \frac{c^2}{3}\end{aligned}$$

■

Teorema generale del valore atteso

Il valore atteso di una generica funzione di variabile casuale $g(\xi)$ è dato da:

$$E[g(\xi)] = \int_{-\infty}^{+\infty} g(x) f_{\xi}(x) dx$$

Permette di calcolare la media di una certa funzione di variabile casuale.

Proprietà: l'operatore di valore atteso è lineare:

$$E[a\xi + b] = aE[\xi] + b$$

2.1.4 Indipendenza statistica e correlazione

In generale il calcolo di medie di due variabili casuali:

$$E[\xi_1 \xi_2]$$

non è banale e richiede la densità di probabilità congiunta. Se però le due variabili casuali sono *statisticamente indipendenti* si ha:

$$f_{\xi_1 \xi_2}(x_1, x_2) = f_{\xi_1}(x_1) f_{\xi_2}(x_2)$$

e ne consegue:

$$\begin{aligned} E[\xi_1 \xi_2] &= \int \int x_1 x_2 f_{\xi_1 \xi_2}(x_1, x_2) dx_1 dx_2 \\ &= E[\xi_1] \cdot E[\xi_2] \end{aligned}$$

Due variabili casuali si dicono **scorrelate** se, anche in assenza di indipendenza statistica, vale la relazione:

$$E[\xi_1 \xi_2] = E[\xi_1] \cdot E[\xi_2]$$

Due variabili casuali gaussiane se sono scorrelate sono anche statisticamente indipendenti.

2.1.5 Medie condizionate

Se si deve calcolare $E[g(\xi_1, \xi_2)]$ in generale si ha che:

$$E[g(\xi_1, \xi_2)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{\xi_1 \xi_2}(x, y) dx dy$$

Tuttavia è più spesso utile usare le medie condizionate:

$$\begin{aligned} E[g(\xi_1, \xi_2)] &= E_{\xi_2} \{E_{\xi_1}[g(\xi_1, \xi_2) | \xi_2]\} \\ &= E_{\xi_1} \{E_{\xi_2}[g(\xi_1, \xi_2) | \xi_1]\} \end{aligned}$$

Esempio 2.2 (Calcolo medie condizionate). Sia:

$$\eta = \cos(\alpha + \beta)$$

con α e β variabili casuali uniformi in $[0, \pi]$ e indipendenti. Si trova:

$$\begin{aligned} E[\eta] &= E[\cos(\alpha + \beta)] \\ &= E_{\alpha}[E_{\beta}[\cos(\alpha + \beta) | \alpha]] \\ E_{\beta}[\cos(\alpha + \beta)] &= \int_{-\infty}^{+\infty} \cos(\alpha + \beta) f_{\beta}(\beta) d\beta \\ &= \int_0^{\pi} \cos(\alpha + \beta) \frac{1}{\pi} d\beta \\ &= \frac{1}{\pi} [\sin(\alpha + \beta)]_0^{\pi} \\ &= \frac{1}{\pi} [\sin(\alpha + \pi) - \sin(\alpha)] \\ &= \frac{1}{\pi} [-\sin(\alpha) - \sin(\alpha)] \\ &= -\frac{2}{\pi} \sin(\alpha) \end{aligned}$$

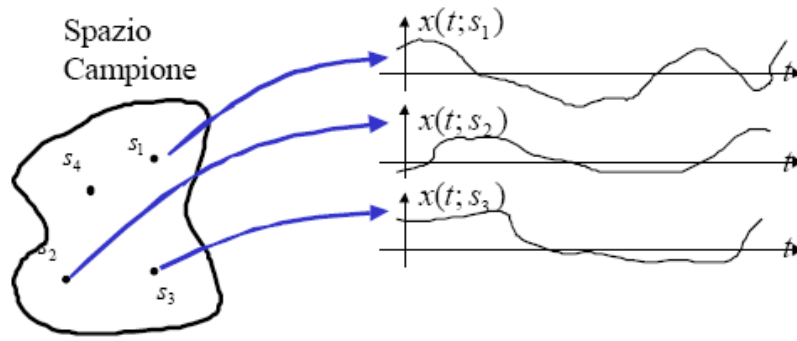


Figura 2.9: Spazio campione dei processi casuali.

$$\begin{aligned}
 E[\eta] &= E_{\alpha} \left[-\frac{2}{\pi} \sin(\alpha) \right] \\
 &= -\frac{2}{\pi} \int_0^{\pi} \sin(\alpha) \frac{1}{\pi} d\alpha \\
 &= -\frac{2}{\pi^2} [-\cos(\alpha)]_0^{\pi} \\
 &= -\frac{2}{\pi^2} [+1 - (-1)] \\
 &= -\frac{4}{\pi^2}
 \end{aligned}$$

■

2.2 Processi casuali

Nella realtà, moltissimi fenomeni fisici generano delle forme d'onda che non sono descrivibili in maniera deterministica. Ad esempio, all'interno di qualunque dispositivo elettronico non alimentato, fluttuazione termiche e quantistiche fanno sì che la corrente e la tensione non siano mai esattamente nulle ma che abbiano delle piccole variazioni intorno allo zero. Questo fenomeno è chiamato *rumore termico* e ha caratteristiche casuali. È chiaro che per descrivere un fenomeno di questo tipo bisogna utilizzare un modello probabilistico.

I *processi casuali* sono modelli probabilistici che si utilizzano per descrivere fenomeni fisici aleatori, quali il rumore termico. Consideriamo la figura 2.9. Immaginiamo di effettuare un esperimento casuale il cui esito non sia un numero ma bensì una forma d'onda. Ciascuna forma d'onda verrà indicata come

$$x(t; \xi_i)$$

e verrà detta *realizzazione* del processo casuale. I processi casuali servono quindi a descrivere matematicamente insiemi di funzioni del tempo. Se le variabili casuali avevano come spazio campione insiemi di numeri reali, qui lo spazio campione è costituito da insiemi di funzioni del tempo (reali). Tutte le funzioni dipendenti dal tempo t e dalla variabile casuale ξ_i vengono dette realizzazione del processo casuale.

Consideriamo ora un altro esempio pratico in cui serve utilizzare i processi casuali. Il modo più semplice di trasmettere dei dati numerici su un canale di comunicazione è

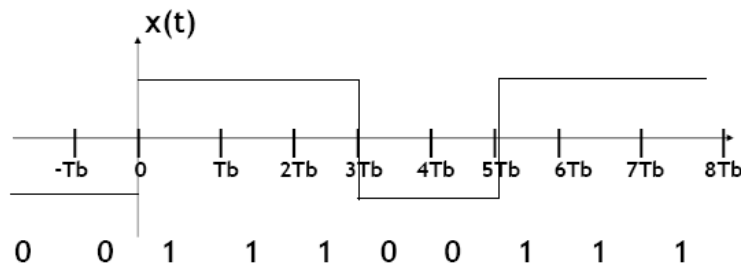


Figura 2.10: Segnale per trasmissione dati.

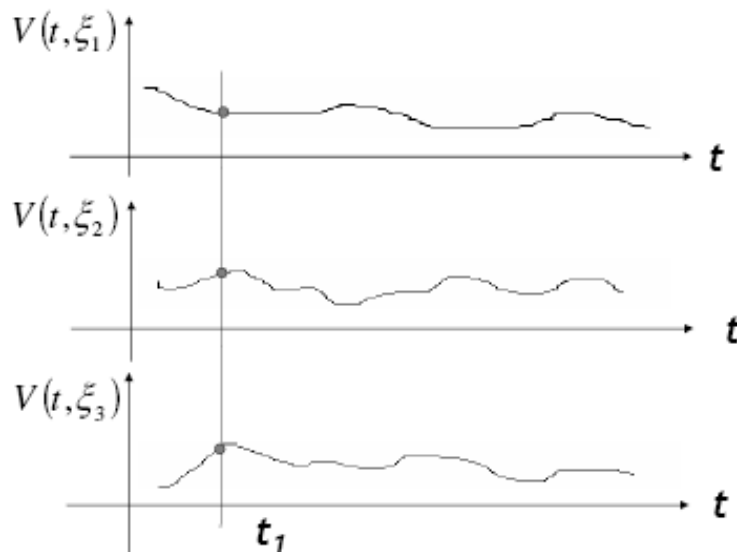


Figura 2.11: Realizzazioni di un processo casuale.

di utilizzare un segnale tipo quello mostrato in figura 2.10. Una singola realizzazione sarebbe trattabile come un segnale deterministico, tuttavia si vuole descrivere in maniera statistica una qualunque uscita di un trasmettitore che riceve una qualunque sequenza di bit. In particolare:

- ▷ Si vogliono poter considerare vari trasmettitori (esempio: le uscite di N modem su PC diversi);
- ▷ Si vogliono considerare per lo stesso modem diversi intervalli di tempo.

La figura 2.11 mostra alcune possibili realizzazioni di un processo casuale:

- ▷ Fissato un certo ξ_N , la funzione $V(t, \xi_N)$ è una funzione deterministica (una realizzazione del processo casuale);
- ▷ Fissato un certo istante di tempo t_1 , i vari valori $V(t_1, \xi_i)$ costituiscono una variabile casuale;
- ▷ $V(t_1, \xi_N)$ è semplicemente un numero.

Si intuisce come un processo casuale possa essere descritto tramite statistiche di variabili casuali. Spesso si indica il processo casuale nel suo complesso senza indicare gli ξ_i , cioè semplicemente come $V(t)$.

2.2.1 Caratterizzazione statistica di un processo casuale

Sia $V_j = V(t_j, \xi_i)$ l'insieme dei valori ottenuti osservando il processo all'istante t_j . La densità di probabilità della risultante variabile casuale V_j dà una caratterizzazione statistica del primo ordine. Una caratterizzazione statistica del primo ordine è data da una densità di probabilità del tipo:

$$f_V(x, t) = \frac{1}{\sqrt{2\pi}\sigma(t)} e^{-\frac{[x-m(t)]^2}{2\sigma^2(t)}}$$

che è un esempio di statistica del primo ordine di tipo gaussiano, con media e varianza dipendenti dal tempo.

Le caratteristiche del primo ordine non sono solitamente sufficienti: non dicono nulla relativamente alla correlazione tra il processo osservato a due istanti diversi $V(t_i)$ e $V(t_j)$. In teoria, dovrebbero essere fornite le densità di probabilità congiunte:

- ▷ del secondo ordine: $f_v(x_1, x_2; t_1, t_2)$;
- ▷ del terzo ordine: $f_v(x_1, x_2, x_3; t_1, t_2, t_3)$;
- ▷ ...;
- ▷ di ordine N .

In pratica, di solito, i parametri che interessa valutare di un processo casuale richiedono la statistica al più del secondo ordine.

Medie di insieme

In queste dispense si faranno soprattutto delle considerazioni sulle medie dei processi casuali (piuttosto che caratterizzazioni complete di ordine $N > 1$). La media di insieme del processo $v(t)$ si indica come $E[v(t)]$ e vale:

$$E[v(t)] = \int_{-\infty}^{+\infty} x f_v(x; t) dx$$

Attenzione: sono medie fatte per "verticali" sulle varie realizzazioni del processo (vedi figura 2.12). Quindi in generale la media di insieme è una funzione del tempo.

Funzione di autocorrelazione La funzione di autocorrelazione di un processo è una media di insieme del secondo ordine definita in questo modo:

$$\begin{aligned} R_v(t_1, t_2) &\triangleq E[v(t_1) v(t_2)] \\ &\triangleq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 f_v(x_1, x_2; t_1, t_2) dx_1 dx_2 \end{aligned}$$

È una quantità che è in generale funzione di due tempi. Estremamente importante, sarà descritta più avanti.

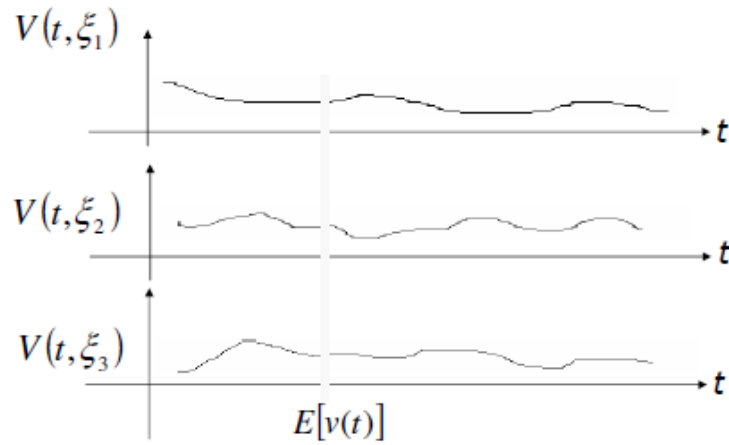


Figura 2.12: Medie di insieme.

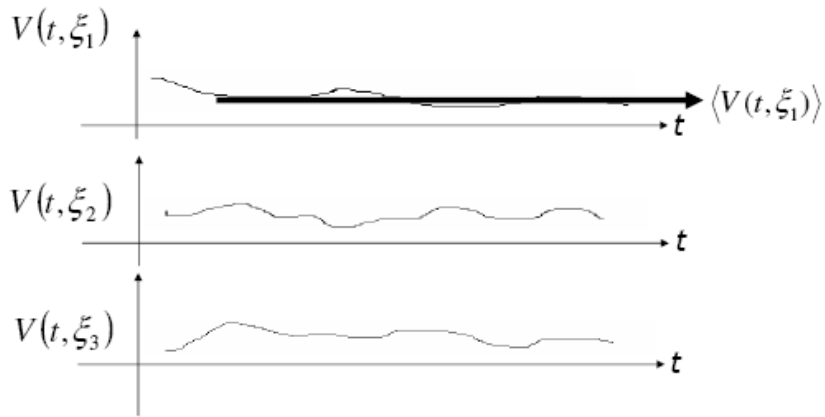


Figura 2.13: Medie temporali.

Medie temporali

Sono le medie fatte:

- ▷ Rispetto al tempo;
- ▷ Su ciascuna realizzazione;
- ▷ Nel senso dei segnali determinati.

Si indicano come:

$$\langle V(t, \xi_i) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} V(t, \xi_i) dt$$

Si noti che mentre le medie di insieme possono dipendere dal tempo, le medie temporali non dipendono dal tempo, ma possono dipendere dalle realizzazioni. Si tratta di medie fatte per "orizzontali" (vedi figura 2.13).

2.2.2 Stazionarietà in senso stretto

Nelle applicazioni pratiche è frequente incontrare processi casuali che abbiano una certa regolarità nel tempo. Questi processi si dicono *stazionari*. Un esempio tipico di processo stazionario è il rumore termico: appare statisticamente uguale anche se ci si sposta da $t = 0$ a $t = t_1$, a patto che non vari la temperatura ambientale.

Un processo casuale si dice *stazionario del primo ordine* se:

$$f_V(x; t) \text{ non dipende dal tempo}$$

Ad esempio, un processo gaussiano non è stazionario se la media e la varianza sono funzioni del tempo.

Un processo casuale si dice *stazionario del secondo ordine* se:

$$f_V(x_1, x_2; t_1, t_2) \text{ dipende solo da } t_1 - t_2$$

e via dicendo si possono definire le stazionarietà di ordine superiore al secondo (che però di solito non si usano).

Esempio 2.3 (Processo casuale quasi determinato). Consideriamo il processo casuale:

$$x(t) = A_p \sin(2\pi f_0 t + \theta)$$

dove:

- ▷ f_0, θ sono valori costanti;
- ▷ A_p è una variabile casuale gaussiana con media m_A e varianza σ_A^2 .

È un processo quasi determinato, poiché è espresso con una formula che dipende da una variabile casuale. È in sostanza un segnale sinusoidale con una incertezza sulle ampiezze.

Vediamo di calcolare la densità di probabilità del primo ordine. Ricordiamo dalla teoria della probabilità che se A_p è gaussiana, $B = k A_p$ è ancora gaussiana con media $k m_A$ e varianza $|k|^2 \sigma_A^2$. Allora il processo $x(t)$ è gaussiano con:

$$\begin{aligned} \text{Varianza} &= \sigma_A^2 |\sin(2\pi f_0 t + \theta)|^2 \\ m_x(t) &= m_A \sin(2\pi f_0 t + \theta) \end{aligned}$$

■

2.2.3 Stazionarietà in senso lato (WSS)

Nelle applicazioni pratiche è spesso molto difficile verificare la stazionarietà di ordine n , anche per piccoli valori di n . Nella maggior parte dei casi è sufficiente verificare la stazionarietà per la media e per l'autocorrelazione.

Un processo si dice *stazionario in senso lato* (Wide Sense Stationary, WSS) se:

- ▷ La media $E[v(t)]$ non dipende dal tempo;
- ▷ La funzione di autocorrelazione $R_v(t_1, t_2)$ dipende solo da $t_1 - t_2$. In tal caso si indica spesso con:

$$R_v(\tau) = E[v(t) v(t + \tau)]$$

con $\tau = t_1 - t_2$.

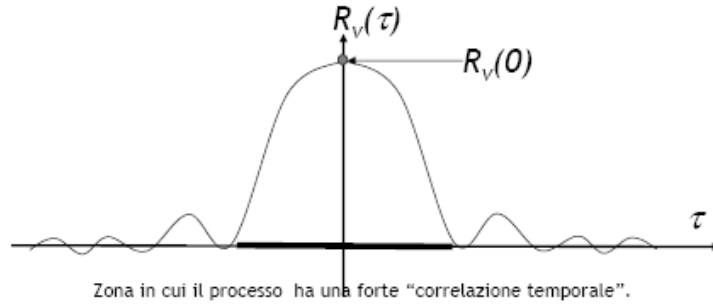


Figura 2.14: Andamento tipico della funzione di autocorrelazione di un processo WSS.

Significato dell'autocorrelazione per processi WSS. Osserviamo intanto alcune proprietà:

- ▷ $R_v(0) = E[v^2(t)]$ = momento di ordine 2 (se $E[v(t)] = 0$ notare che $R_v(0) = \sigma_v^2$);
- ▷ $R_v(\tau) = R_v(-\tau)$ è una funzione pari;
- ▷ $R_v(0) \geq |R_v(\tau)|$ cioè l'autocorrelazione è massima nell'origine.

Per capire il significato di $R_v(\tau)$, supponiamo che il processo sia a media nulla:

$$E[v(t)] = 0$$

Per τ sufficientemente elevati, nell'espressione:

$$R_v(\tau) = E[v(t) v(t + \tau)]$$

si può supporre che $v(t)$ e $v(t + \tau)$ siano scorrelate, cioè il processo, visto a due istanti di tempo sufficientemente distanti, non abbia più memoria (si veda la figura 2.14). Per $t \rightarrow \infty$ si ha:

$$\begin{aligned} R_v(\tau) &= E[v(t) v(t + \tau)] \\ &= E[v(t)] E[v(t + \tau)] \\ &= 0 \end{aligned}$$

Un processo che varia lentamente nel tempo, avrà forte correlazione anche per t_1 e t_2 distanti. Allora funzione di autocorrelazione "larga" vuol dire processo lento, e viceversa.

Esempio 2.4 (Verifica stazionarietà in senso lato). Consideriamo il processo casuale:

$$x(t) = A_p \cos(2\pi f_0 t + \theta)$$

con:

- ▷ A_p e f_0 costanti;
- ▷ θ variabile casuale uniforme in $[-\pi; +\pi]$.

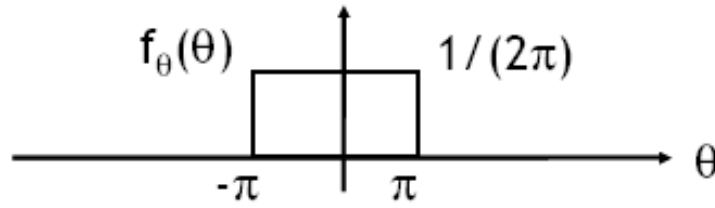


Figura 2.15: Densità di probabilità della fase iniziale θ del segnale sinusoidale dell'esempio 2.4.

Si tratta di un insieme di segnali sinusoidali con fase iniziale θ casuale. Valutiamo se è stazionario in senso lato.

La figura 2.15 mostra la densità di probabilità di θ . Calcoliamo la media:

$$\begin{aligned}
 E[x(t)] &= E[A_p \cos(2\pi f_0 t + \theta)] \\
 &= \int_{-\infty}^{+\infty} A_p \cos(2\pi f_0 t + \theta) f_\theta(\theta) d\theta \\
 &= \int_{-\pi}^{+\pi} A_p \cos(2\pi f_0 t + \theta) \frac{1}{2\pi} d\theta \\
 &= 0
 \end{aligned}$$

La media non dipende dal tempo, quindi il processo è stazionario per la media.

Calcoliamo l'autocorrelazione:

$$\begin{aligned}
 R_x(t, \tau) &= E[x(t) x(t + \tau)] \\
 &= E[A_p \cos(2\pi f_0 t + \theta) A_p \cos(2\pi f_0(t + \tau) + \theta)] \\
 &= A_p^2 E[\cos(2\pi f_0 t + \theta) \cos(2\pi f_0(t + \tau) + \theta)] \\
 &\stackrel{1}{=} \frac{A_p^2}{2} E[\cos(2\pi f_0 \tau) + \cos(4\pi f_0 t + 2\pi f_0 \tau + 2\theta)] \\
 &= \frac{A_p^2}{2} \{E[\cos(2\pi f_0 \tau)] + \underbrace{E[\cos(4\pi f_0 t + 2\pi f_0 \tau + 2\theta)]}_{=0, \text{ stesso conto della media}}\} \\
 &= \frac{A_p^2}{2} \cos(2\pi f_0 \tau)
 \end{aligned}$$

dove in $\stackrel{1}{=}$ si è usata la seguente relazione trigonometrica:

$$\cos(A) \cos(B) = \frac{1}{2} [\cos(B - A) + \cos(B + A)]$$

Si ha dipendenza solo da τ e dunque il processo è stazionario in senso lato. ■

Esempio 2.5. Continuiamo l'esempio 2.3 a pagina 41. Il processo casuale

$$x(t) = A_p \sin(2\pi f_0 t + \theta)$$

con A_p variabile casuale gaussiana con media m_A e varianza σ_A^2 , non è stazionario. Infatti la media dipende dal tempo:

$$E[x(t)] = m_A \sin(2\pi f_0 t + \theta)$$

Sia la media che l'autocorrelazione sono periodiche. In questo caso il processo si dice **ciclostazionario**. ■

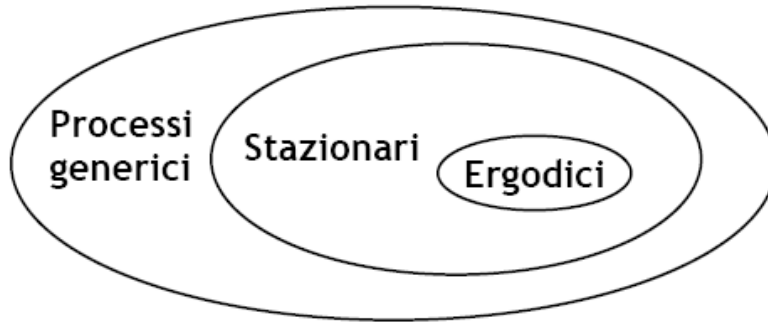


Figura 2.16: Processi casuali stazionari ed ergodici.

2.2.4 Ergodicità

Abbiamo visto che su ciascuna realizzazione di un processo casuale si può definire anche una media temporale (infatti una singola realizzazione è una funzione del tempo). Un processo casuale si dice *ergodico* se le medie temporali coincidono con le medie di insieme:

$$\langle x(t) \rangle = E[x(t)]$$

Un processo ergodico è sempre anche stazionario, ne consegue che un processo ergodico ha le medie temporali uguali su tutte le realizzazioni e costanti nel tempo.

Allora se il processo è ergodico, essendo anche stazionario si ha per la potenza:

$$\begin{aligned} P_x &= \langle x^2(t) \rangle \\ &= E[x^2(t)] \\ &= R_x(0) \end{aligned}$$

Importante: per processi ergodici (e quindi stazionari), la potenza media coincide con il valore quadratico medio come media di insieme, e coincide con la funzione di autocorrelazione calcolata in $\tau = 0$. Se il processo è a media nulla, allora si ha ulteriore coincidenza con la varianza:

$$\begin{aligned} P_x &= \langle x^2(t) \rangle \\ &= E[x^2(t)] \\ &= \sigma_x^2 \end{aligned}$$

Utilizzeremo soprattutto i processi stazionari WSS ed ergodici, che sono tuttavia una piccola classe di tutti i processi casuali, come mostrato in figura 2.16. L'unica e importante eccezione è il segnale per trasmissione dati, che in generale non è stazionario.

Esempio 2.6 (Di ergodicità). Riconsideriamo il processo:

$$x(t) = A_p \cos(2\pi f_0 t + \theta)$$

con θ variabile casuale uniforme in $[-\pi, +\pi]$. Verifichiamo se il processo è ergodico. Avevamo calcolato:

$$\begin{aligned} E[x(t)] &= 0 \\ R_x(\tau) &= \frac{A_p^2}{2} \cos(2\pi f_0 \tau) \end{aligned}$$

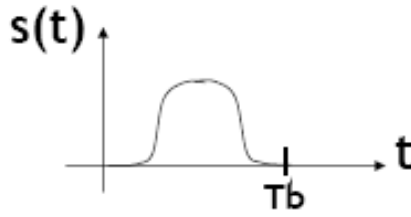


Figura 2.17: Segnale elementare per la trasmissione di dati digitali.

Calcoliamo la media temporale:

$$\begin{aligned} \langle x(t) \rangle &= \langle A_p \cos(2\pi f_0 t + \theta) \rangle \\ &= \frac{A_p}{T} \int_0^T \cos(2\pi f_0 t + \theta) dt \\ &= 0 \end{aligned}$$

Essendo $E[x(t)] = \langle x(t) \rangle$ il processo è ergodico per la media.

Calcoliamo la media quadratica temporale:

$$\begin{aligned} \langle x^2(t) \rangle &= \langle A_p^2 \cos^2(2\pi f_0 t + \theta) \rangle \\ &= \frac{A_p^2}{2} \\ &= \frac{P_{picco}}{2} \end{aligned}$$

Calcoliamo $E[x^2(t)]$:

$$E[x^2(t)] = R_x(0)$$

ma:

$$R_x(\tau) = \frac{A_p^2}{2} \cos(2\pi f_0 \tau)$$

Quindi si ottiene:

$$\begin{aligned} E[x^2(t)] &= R_x(0) \\ &= \frac{A_p^2}{2} \\ &= \langle x^2(t) \rangle \end{aligned}$$

Il processo è ergodico anche per la potenza. ■

Esempio 2.7 (Segnale per trasmissione dati). Il modo più semplice di trasmettere una sequenza di bit in maniera numerica è di utilizzare il seguente segnale:

$$x(t) = \sum_{n=-\infty}^{+\infty} a_n s(t - nT_b) \quad (2.10)$$

dove a_n sono variabili casuali statisticamente indipendenti, che assumono valori ± 1 in maniera equivalente e $s(t)$ è una forma d'onda che solitamente è limitata in $[0, T_b]$ (si veda la figura 2.17). Si tratta di un processo quasi determinato, che però dipende da

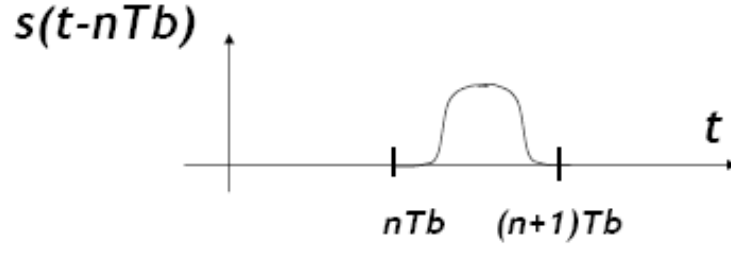


Figura 2.18: Segnale elementare per la trasmissione di dati digitali ritardato di n tempi di bit.

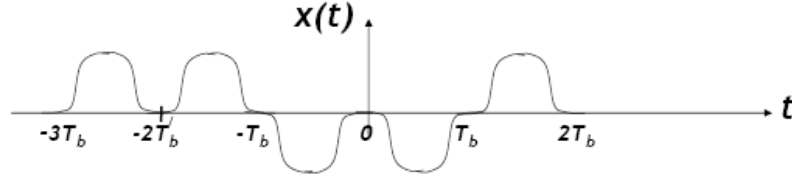


Figura 2.19: Segnale per la trasmissione di dati digitali.

infinite variabili casuali a_n . Il segnale $s(t - nT_b)$ è mostrato in figura 2.18: si tratta di $s(t)$ traslato in nT_b . Allora il segnale (2.10) è quello mostrato in figura 2.19.

Calcolo della media di insieme:

$$\begin{aligned} E[x(t)] &= E \left[\sum_{n=-\infty}^{+\infty} a_n s(t - nT_b) \right] \\ &= \sum_{n=-\infty}^{+\infty} \underbrace{E[a_n]}_{=0} s(t - nT_b) \\ &= 0 \end{aligned}$$

Calcolo dell'autocorrelazione:

$$\begin{aligned} R_x(t, \tau) &= E[x_i(t) x_j(t + \tau)] \\ &= E \left[\sum_{i=-\infty}^{+\infty} a_i s(t - iT_b) \sum_{j=-\infty}^{+\infty} a_j s(t + \tau - jT_b) \right] \\ &= \sum_{j=-\infty}^{+\infty} \sum_{i=-\infty}^{+\infty} E[a_i a_j] s(t - iT_b) s(t + \tau - jT_b) \end{aligned}$$

▷ Se $i = j$:

$$E[a_i a_i] = E[a_i^2] = 1$$

▷ Se $i \neq j$, per indipendenza statistica si ha che:

$$E[a_i a_j] = E[a_i] E[a_j] = 0$$

Quindi l'autocorrelazione vale:

$$R_x(t, \tau) = \sum_{i=-\infty}^{+\infty} s(t - iT_b) s(t + \tau - iT_b)$$

Ne deriva che il processo non è stazionario per l'autocorrelazione, in quanto questa dipende sia da t che da τ . Si può dimostrare che $x(t)$ è ciclostazionario di periodo T_b . ■

2.2.5 Densità spettrale di potenza

In questo paragrafo definiremo lo *spettro di potenza* di un processo casuale $x(t)$. Si tratta di uno strumento matematico estremamente utile che ci permetterà di effettuare un'analisi in frequenza del processo. Questo spettro di potenza dovrà avere le stesse proprietà che si erano ottenute per i segnali deterministici.

Definizione: sia $x_T(t)$ una versione di $x(t)$ troncata a $[-T/2, +T/2]$, cioè:

$$x_T(t) \triangleq p_T(t) x(t)$$

La sua trasformata di Fourier

$$X_T(f) \triangleq F \{x_T(t)\}$$

è ancora un processo casuale, essendo un insieme di forme d'onda.

Lo spettro di potenza $G_x(f)$ del processo viene allora definito in questo modo:

$$G_x(f) \triangleq \lim_{T \rightarrow \infty} \frac{E [|X_T(f)|^2]}{T} \quad (2.11)$$

Questa è una definizione che vale anche per processi non stazionari. La densità spettrale di potenza di un processo casuale $x(t)$ reale ha le seguenti proprietà:

1. $G_x(f)$ è reale e positivo;
2. $G_x(f)$ è una funzione pari;
3. $P_x = \int_{-\infty}^{+\infty} G_x(f) df$.

Teorema di Wiener-Khinchin

Questo teorema mi dice che, se $x(t)$ è stazionario in senso lato (WSS), lo spettro di potenza è uguale alla trasformata di Fourier della funzione di autocorrelazione:

$$\begin{aligned} G_x(f) &= F \{R_x(\tau)\} \\ &= \int_{-\infty}^{+\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau \end{aligned} \quad (2.12)$$

Nella pratica però capita spesso di avere a che fare con processi ciclostazionari: ad esempio il segnale per trasmissione dati che abbiamo considerato nell'esempio 2.7 a pagina 45. In questo caso la (2.12) è ancora valida, ma il teorema è da modificare leggermente: invece di lavorare su $R_x(\tau)$ si deve usare la funzione di autocorrelazione mediata rispetto a t :

$$R_{x'}(\tau) = \langle R_x(t, t + \tau) \rangle$$

quindi lo spettro di potenza diventa:

$$\begin{aligned} G_x(f) &= F \{ \langle R_x(t, t + \tau) \rangle \} \\ &= F \{ R_{x'}(\tau) \} \end{aligned}$$

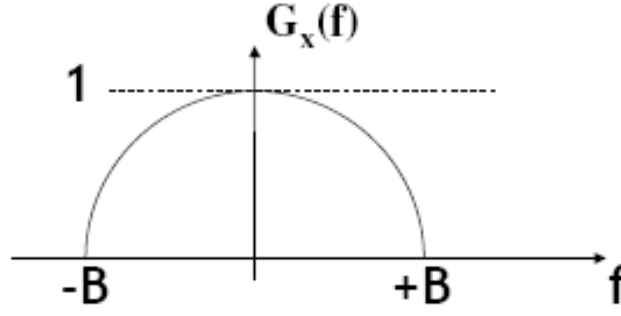


Figura 2.20: Esempio di spettro di potenza di processo casuale in banda base.

Esempio 2.8 (Modulazione di ampiezza). Sia $x(t)$ un processo casuale WSS con densità spettrale di potenza confinata in banda base (vedi figura 2.20). Sia

$$y(t) = x(t) \cos(2\pi f_0 t) \quad (2.13)$$

con $f_0 \gg B$. Calcoliamo $G_y(f)$.

Proviamo a calcolarla passando dall'autocorrelazione:

$$\begin{aligned} R_y(t, \tau) &= E[y(t) y(t + \tau)] \\ &= E[x(t) \cos(2\pi f_0 t) x(t + \tau) \cos(2\pi f_0(t + \tau))] \\ &= E[x(t) x(t + \tau)] \cos(2\pi f_0 t) \cos[2\pi f_0(t + \tau)] \\ &= R_x(\tau) \cos(2\pi f_0 t) \cos[2\pi f_0(t + \tau)] \\ &= \frac{R_x(\tau)}{2} [\cos(2\pi f_0 \tau) + \cos(2\pi f_0 \tau + 4\pi f_0 t)] \end{aligned}$$

Ora calcoliamo $\langle R_y(t, \tau) \rangle$, cioè mediamo rispetto a t :

$$\langle R_y(t, \tau) \rangle = \langle \frac{R_x(\tau)}{2} [\cos(2\pi f_0 \tau) + \cos(2\pi f_0 \tau + 4\pi f_0 t)] \rangle$$

essendo una funzione coseno, si ottiene:

$$\langle R_y(t, \tau) \rangle = \frac{R_x(\tau)}{2} \cos(2\pi f_0 \tau)$$

Quindi lo spettro di potenza vale (vedi figura 2.21):

$$G_y(f) = \frac{1}{4} [G_x(f - f_0) + G_x(f + f_0)]$$

Notare che:

$$\begin{aligned} P_y &= \int_{-\infty}^{+\infty} G_y(f) df \\ &= \frac{1}{2} P_x \end{aligned}$$

Come già dimostrato, moltiplicare per una funzione sinusoidale, dimezza la potenza. ■

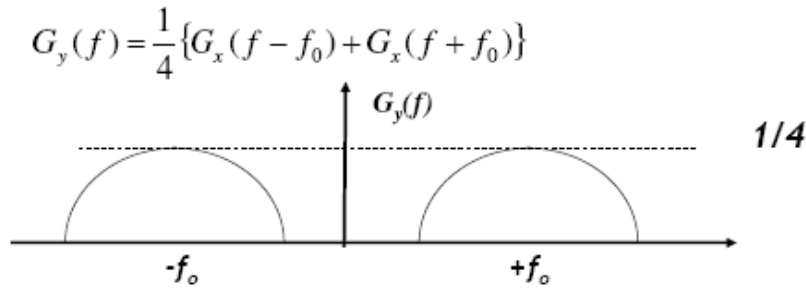


Figura 2.21: Spettro di potenza del segnale modulato in ampiezza (2.13).

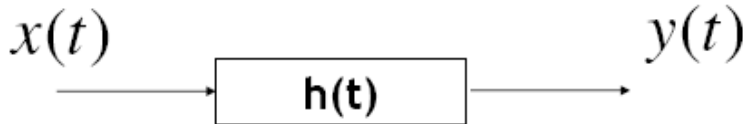


Figura 2.22: Processo casuale WSS che passa attraverso un sistema LTI.

2.2.6 Trasformazioni LTI di processi casuali

Si consideri un processo WSS $x(t)$ che passa attraverso un sistema LTI (vedi figura 2.22). Il processo $y(t)$ sarà anch'esso WSS con densità spettrale di potenza (questa espressione si utilizza molto):

$$G_y(f) = |H(f)|^2 G_x(f)$$

Per la funzione di autocorrelazione si ottiene (si usa poco):

$$R_y(\tau) = h(-\tau) * h(\tau) * R_x(\tau)$$

Trasformazioni LTI di processi casuali gaussiani

Sono processi $x(t)$ per cui le caratteristiche di qualunque ordine sono gaussiane. Hanno svariate proprietà: per noi è importante che dato un sistema lineare e un ingresso $x(t)$ gaussiano, l'uscita è ancora gaussiana (non è vero in generale per nessun altro tipo di processo).

Un esempio tipico di processo gaussiano è il *rumore gaussiano bianco*. È uno dei più importanti processi casuali nell'ambito TLC. È un processo casuale:

- ▷ Gaussiano;
- ▷ Con densità spettrale piatta (per questo si chiama bianco) e convenzionalmente indicata come:

$$G_n(f) = \frac{N_0}{2}$$

La sua funzione di autocorrelazione è quindi una Delta di Dirac:

$$\begin{aligned} R_n(\tau) &= F^{-1} \{G_n(f)\} \\ &= \frac{N_0}{2} \delta(\tau) \end{aligned}$$

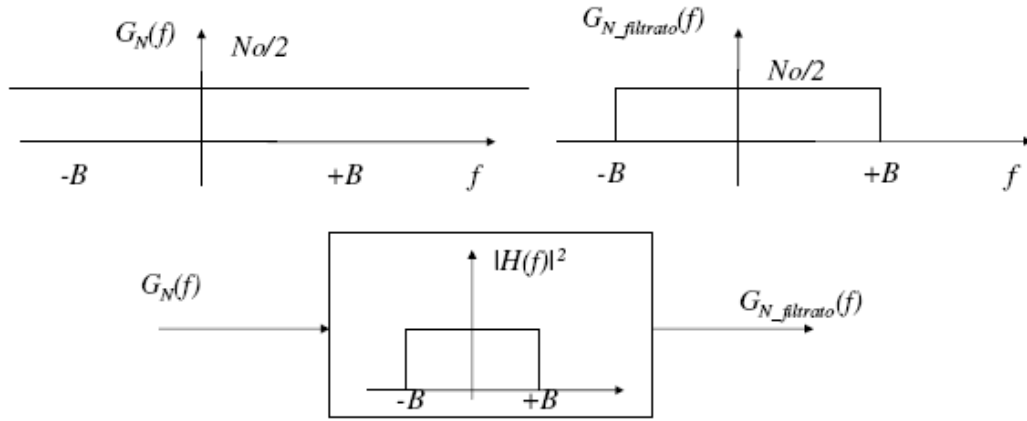


Figura 2.23: Rumore gaussiano binario filtrato.

Tratteremo più in dettaglio il rumore gaussiano bianco nel capitolo 5.

Spesso in pratica si ha a che fare con il rumore gaussiano bianco filtrato con un filtro passa-basso ideale, cioè di tipo rettangolare (vedi figura 2.23). Il rumore gaussiano filtrato (che si chiama colorato) ha:

$$\begin{aligned}
 P_{N \text{ filtrato}} &= \int_{-\infty}^{+\infty} G_{N \text{ filtrato}}(f) df \\
 &= \frac{N_0}{2} 2B \\
 &= N_0 B
 \end{aligned}$$

In generale, per un filtro con una certa $H(f)$:

$$\begin{aligned}
 P_{N\text{-filtrato}} &= \int_{-\infty}^{+\infty} \frac{N_0}{2} |H(f)|^2 df \\
 &= N_0 \int_0^{+\infty} |H(f)|^2 df
 \end{aligned}$$

Banda equivalente di rumore

Per quanto riguarda il rumore, è comodo pensare ad un qualunque filtraggio $H(f)$ un suo equivalente rettangolare ideale (vedi figura 2.24). dove:

▷ massimo del filtro originale $H(f)$:

$$|H_{eq}(f)|^2 = |H(f)|^2|_{max} = A_p$$

▷ Banda equivalente di rumore:

$$B_{eq} \triangleq \frac{1}{|H(f)|^2_{max}} \int_0^{+\infty} |H(f)|^2 df$$

In questo modo un filtro generico, per quel che riguarda le densità spettrali di potenza, è reso equivalente ad un filtro passa-basso ideale.

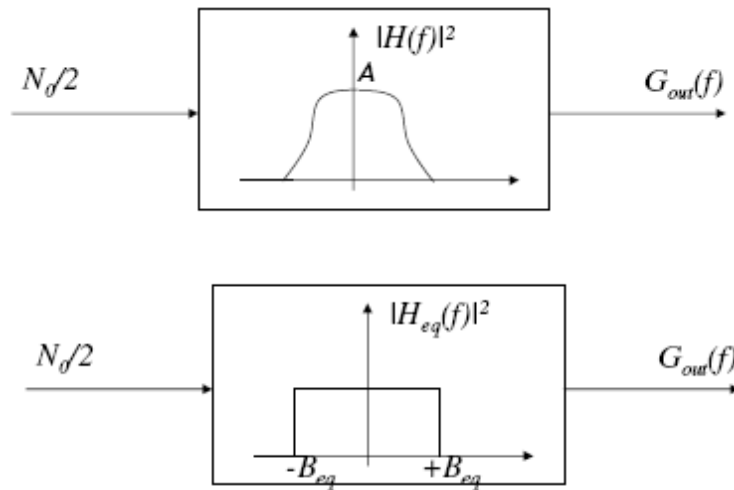


Figura 2.24: Definizione di banda equivalente di rumore.

2.3 Conclusione

In questo capitolo abbiamo richiamato i concetti principali sulla teoria della probabilità e i processi casuali. Il lettore che desiderasse approfondire questi argomenti è invitato a consultare il testo [6]. Si tratta di un libro molto completo sull'argomento.

Un altro testo che consigliamo per la parte sui processi casuali è [2]. Si tratta di un testo introduttivo abbastanza semplice e dal taglio prettamente didattico.

Capitolo 3

Introduzione ai sistemi di trasmissione

Questo capitolo introduce le varie tipologie di segnali e sistemi di trasmissione e le grandezze fondamentali che li caratterizzano. Vengono poi brevemente trattate le equazioni della propagazione nel caso di sistemi via cavo e via etere.

3.1 Tipologie di segnali e di sistemi di trasmissione

I segnali che ci interessano in questo corso sono di tipo elettrico, cioè tensioni, correnti oppure ampiezze di campi elettrici. Principalmente tali segnali possono essere di 2 tipi:

Segnali analogici. Sono segnali (elettrici) il cui contenuto informativo è **continuo** sia nel tempo che nelle ampiezze. Matematicamente sono funzioni reali (o a volte complesse) di variabile reale (il tempo). Esempi di segnali analogici: voce, video, segnali provenienti da sensori, etc.

Segnali digitali/numerici. Sono segnali (elettrici) il cui contenuto informativo è **discretizzato** sia nel tempo che nelle ampiezze. Nella versione più semplice si tratta di stringhe di bit codificati su opportuni livelli logici.

3.1.1 Sistemi analogici

La figura 3.1 mostra lo schema a blocchi di un sistema di trasmissione analogico. Segue la descrizione dei singoli blocchi che lo compongono.

- **Sorgente analogica:** apparato che emette il segnale da trasmettere. Tipicamente è un microfono o un sensore. Il segnale da trasmettere è una funzione reale e continua di variabile reale, e viene indicato con $v_{in}(t)$.
- **TX analogico:** tipicamente è un modulatore analogico, cioè un dispositivo che adatta il segnale $v_{in}(t)$, trasformandolo nel segnale $v_{TX}(t)$, in modo da poterlo trasmettere fisicamente sul canale trasmissivo.
- **Canale di trasmissione:** è il mezzo fisico che "trasporta" il segnale verso l'utente finale. Tipicamente si tratta di cavo, etere, doppino telefonico o fibra ottica. Si noti che in generale il segnale ricevuto $v_{RX}(t)$ è diverso da $v_{TX}(t)$ a causa principalmente di rumore e distorsioni introdotte dal canale.

Descrizione
blocchi
sistema
analogico

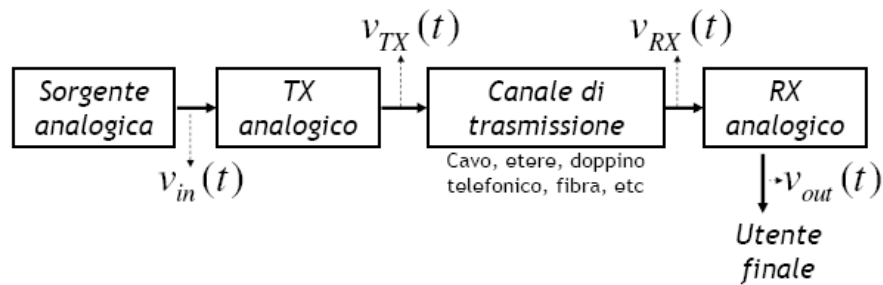


Figura 3.1: Schema a blocchi di un sistema di trasmissione analogico.

- **RX analogico:** tipicamente è un demodulatore analogico, cioè un dispositivo che effettua l'operazione inversa rispetto a *TX analogico*. La sua uscita $v_{out}(t)$ viene inviata all'utente finale.

Scopo sistemi di trasmissione analogici Il concetto importante è che lo scopo della trasmissione analogica è quello di far sì che il segnale di uscita $v_{out}(t)$ verso l'utente finale sia **il più possibile simile** a quello generato dalla sorgente analogica $v_{in}(t)$ (vedremo più avanti nel corso come si quantifica il concetto di "il più possibile simile").

Modulazione Si fa notare che il segnale $v_{TX}(t)$ trasmesso sul canale di trasmissione (cavo, etere, fibra ottica, etc.) può essere significativamente diverso dal segnale generato dalla sorgente $v_{in}(t)$: si tratta appunto delle tecniche di **modulazione**, ampiamente trattate in questo corso. I motivi per cui vengono utilizzate le tecniche di modulazione sono essenzialmente due:

1. Trasportare il segnale nella gamma di frequenze più idonea per la trasmissione sul canale;
2. Inviare sullo stesso mezzo fisico molti segnali contemporaneamente.

Analizzeremo in dettaglio entrambi questi aspetti nel seguito del corso. Si ribadisce che lo scopo fondamentale è assicurare l'integrità di $v_{out}(t)$.

Il progetto di un sistema di trasmissione analogica deve tenere essenzialmente conto di due effetti fisici fondamentali (trattati in questo corso):

- Rumore e distorsioni** degradano il segnale trasmesso
- Il **rumore** introdotto dagli apparati e dal canale di trasmissione.
 - Le **distorsioni** (lineari e non) sul segnale trasmesso. Le distorsioni lineari sono tipicamente generate dalle limitazioni in banda di canale, trasmettitore e ricevitore. Le distorsioni non lineari sono causate principalmente da elementi presenti in trasmettitore o ricevitore (ad esempio amplificatori di potenza).

Altri effetti che causano una degradazione del segnale trasmesso (non trattati in questo corso):

- Interferenze da parte di altri segnali (soprattutto nei sistemi via etere);
- Variabilità nel tempo del canale di trasmissione;
- Etc.

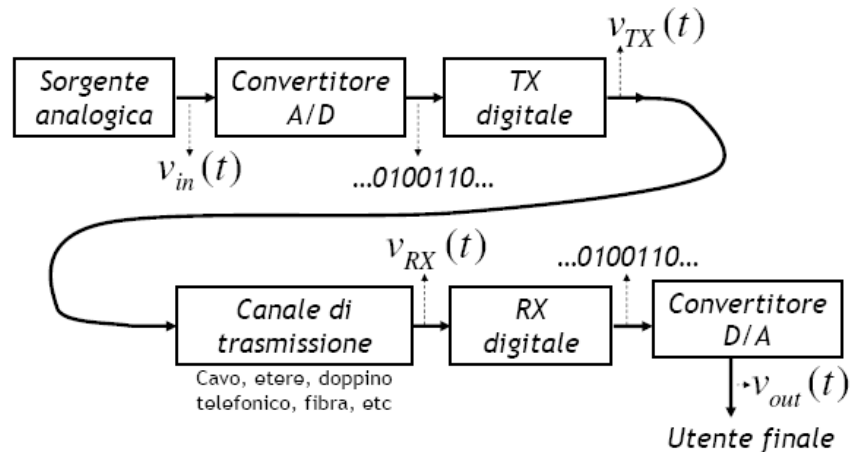


Figura 3.3: Schema a blocchi di un sistema di trasmissione misto analogico-digitale.

- **Convertitore D/A:** effettua l'operazione inversa del convertitore A/D, cioè trasforma la stringa di bit ricevuta nel segnale analogico da inviare all'utente finale.

Le tecniche di conversione analogico-digitali (tecniche *PCM - Pulse Coded Modulation*, trattate in questo corso) costituiscono l'unione tra i due mondi.

Scopo
sistemi di
trasmissione
misti
analogico-
digitali

Si fa notare che, anche nel caso misto, scopo della trasmissione è quello di far sì che il segnale di uscita $v_{out}(t)$ verso l'utente finale sia **il più possibile simile** a quello generato dalla sorgente analogica $v_{in}(t)$. In altre parole, l'utente finale (almeno in linea teorica) non deve percepire alcuna differenza tra un sistema di trasmissione completamente analogico, ed uno misto analogico-digitale. Nei sistemi moderni, la conversione A/D e D/A può avvenire più volte.

Esempio 3.1. Telefonia tradizionale Solitamente nel sistema telefonico:

- Da casa a prima centrale: trasmissione analogica su doppino;
- Nella prima centrale: conversione A/D;
- Nella rete telefonica: trasmissione digitale;
- Nella ultima centrale: conversione D/A;
- Dall'ultima centrale verso l'utente di arrivo: trasmissione analogica su doppino.

Tuttavia, esistono casi (ad esempio cambio di operatore, chiamate intercontinentali, etc.) in cui anche all'interno della rete telefonica è necessario effettuare più volte coppie di conversione A/D e D/A. ■

3.1.4 Segnali audio

La tipologia di sorgente analogica più importante è sicuramente l'audio-voce. Con audio-voce si intende qui il segnale elettrico generato da un trasduttore¹, tipicamente un microfono. Al fine di non "sprecare" risorse, il segnale audio viene "adattato" alle caratteristiche dell'orecchio umano. Ad esempio:

- L'orecchio umano non sente frequenze superiori a 20 kHz (e inferiori a 10 – 20 Hz);
- L'orecchio umano "distingue" tranquillamente la voce di una persona (e la riconosce) anche se il segnale è tagliato attorno ai 4 kHz.

Esempio 3.2. standard per il trasporto dell'audio Tra i moltissimi standard per il trasporto dell'audio, si accennano le principali caratteristiche dei due casi più importanti.

Audio di qualità telefonica:

- Banda tagliata a 4 kHz;
- Dinamica limitata (circa 30 dB);
- Segnale mono (no stereo).

Audio musicale ad alta fedeltà:

- Banda tagliata a 20 kHz;
- Dinamica elevata (più di 70 dB);
- Segnale stereo (cioè, di fatto, due segnali in parallelo).

■

La descrizione di un segnale audio dal punto di vista matematico è fatta tramite la teoria dei processi casuali. Si tratta di un problema molto complesso, in quanto il segnale audio è intrinsecamente molto complicato. In primissima approssimazione, un segnale audio può venire descritto dando:

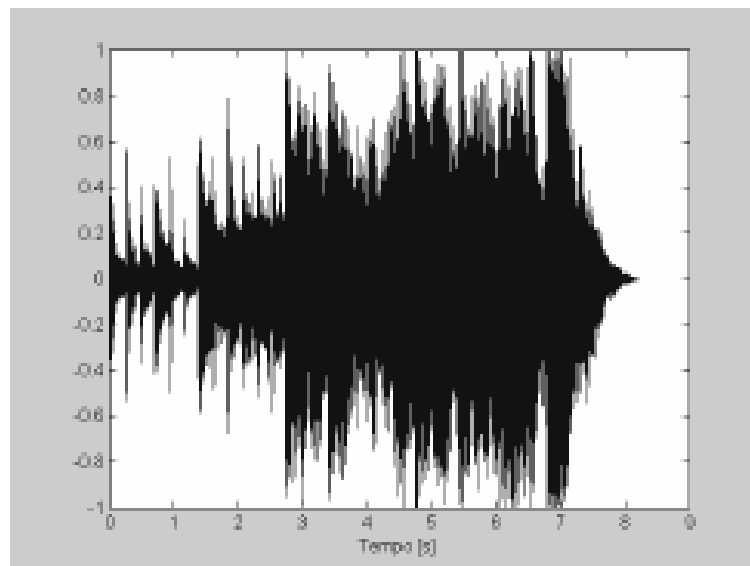
Descrizione
matematica
segnale audio

- La sua densità spettrale di potenza o spesso semplicemente una indicazione della banda che occupa.
- La sua densità di probabilità. Tipicamente, si usa un modello ad esponenziale bilatera.

Esempio 3.3. Segnale musicale La figura 3.4(a) rappresenta l'andamento nel tempo di un breve segnale musicale (circa 8 secondi, file "*KDE_Startup.wav*", disponibile sul sito del corso). La figura 3.4(b) ne mostra due zoom. La figura 3.4(c) mostra l'istogramma e lo spettro del segnale musicale.

■

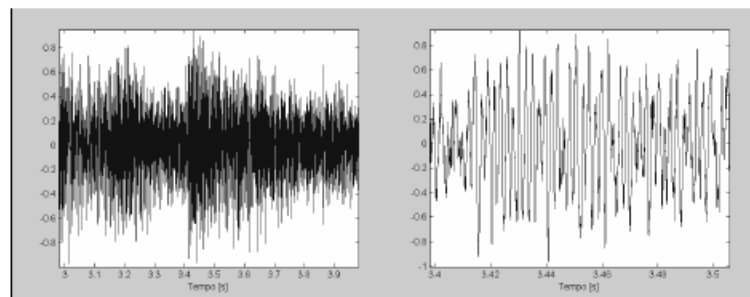
¹Si ricorda che un trasduttore è un dispositivo che converte un segnale meccanico o fisico in un segnale elettrico.



(a) Andamento nel tempo del segnale musicale.

► Zoom su 1 secondo

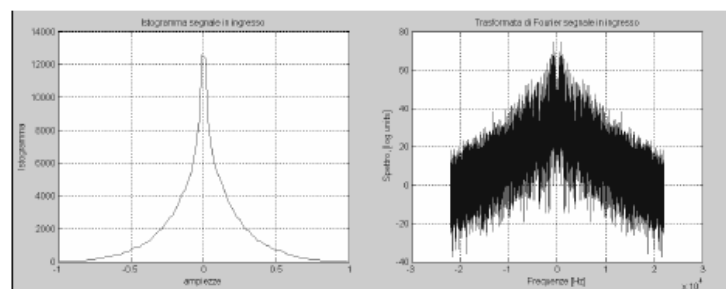
► Zoom su 0.1 secondo



(b) Due zoom del segnale musicale.

► Istogramma

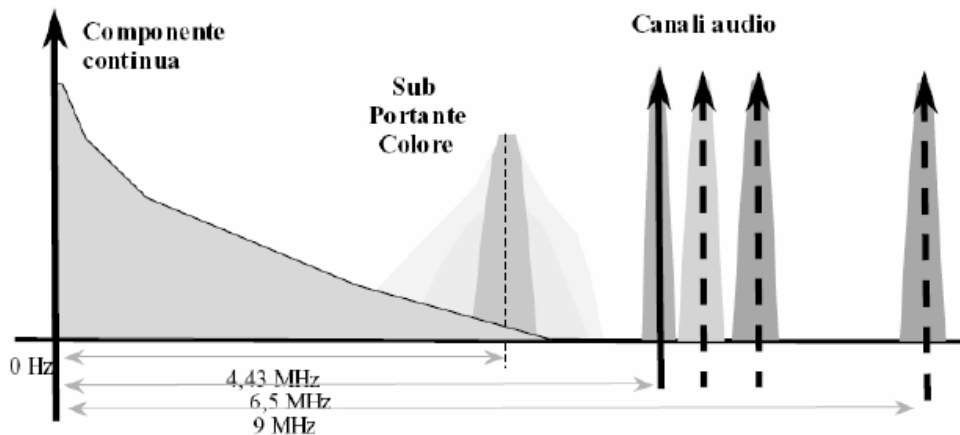
► Spettro



(c) Istogramma (a sinistra) e spettro (a destra) del segnale musicale.

Figura 3.4: Segnale musicale "KDE_Startup.wav", disponibile sul sito del corso.

Spettro del segnale video composito in banda base (standard PAL)



Tratto da: www.dgtvi.it/pdf/DGTVI_Ambrosio.pdf

Figura 3.5: Spettro segnale video PAL descritto nell'esempio 3.4. Si fa notare che lo spettro complessivo occupa circa 8 MHz.

3.1.5 Segnali video

L'altra importante classe di segnali analogici è quella dei segnali video. Si tratta di segnali ancora più complessi di quelli audio. In primissima approssimazione, i segnali video derivano dalla "scansione" per righe dell'immagine, scomposta in termini di intensità sulle tre componenti fondamentali di colore (RGB: Red, Green, Blue).

Esempio 3.4. standard televisivo europeo A puro titolo informativo, si riportano alcune caratteristiche dello standard televisivo per la videodiffusione in Europa:

- 25 quadri al secondo;
- 625 righe per immagine;
- $64 \mu s$ per la scansione di una riga;
- Conversione delle 3 componenti RGB in 1 segnale di luminanza, e 2 segnali di crominanza;
- A questo si aggiungono 2 segnali audio.

Tramite una tecnica di modulazione *VSB* (*Vestigial Side-Band*, non trattata in questo corso), il segnale risultante ha l'organizzazione spettrale mostrata in figura 3.5. ■

3.2 Grandezze fondamentali di un sistema di trasmissione

Un ingegnere delle telecomunicazioni che debba progettare un sistema di TLC dal punto di vista fisico-trasmissivo deve essere in grado di stabilire il migliore trade-off tra le sue grandezze fondamentali:

1. Banda;
2. Potenza;
3. Complessità.

3.2.1 Banda

Fondamentale è il ruolo della banda occupata dal segnale in relazione alla *banda disponibile* sul canale di trasmissione (si veda più avanti per le varie definizioni di banda). Con "banda disponibile" si intende l'intervallo di frequenze su cui si può trasmettere in modo affidabile. Ad esempio, i cavi telefonici hanno banda disponibile [200 Hz - 4 kHz].

La banda è una risorsa molto costosa

La banda disponibile sul canale di trasmissione è una risorsa molto costosa, soprattutto per le trasmissioni via etere. La banda del canale:

- Via etere è limitata per legge da organi internazionali;
- Via cavo è limitata per limiti fisici del cavo.

3.2.2 Potenza

La potenza disponibile in trasmissione è sempre limitata

La potenza utilizzabile all'uscita del trasmettitore è sempre limitata:

- Dai limiti fisici degli apparati;
- Oppure: da imposizioni di legge, ad esempio relative alla massima potenza in uscita da un'antenna (compatibilità elettromagnetica).

Più è alta la potenza disponibile al ricevitore, più il sistema è robusto rispetto alle sorgenti di rumore.

3.2.3 Complessità

Il terzo parametro importante è la "complessità" degli apparati di trasmissione e ricezione. Si vedrà in questo corso e nei prossimi come sia possibile diminuire le richieste di banda e/o potenza aumentando la complessità degli apparati di ricetrasmissione.

Intuitivamente, possiamo dire che i sistemi trasmissivi migliori sono quelli che mantengono la qualità richiesta del segnale, utilizzando la minor banda, potenza e complessità possibili.

3.3 Richiami sui decibel

I dB si usano per rapporti di potenze

I decibel (dB) principalmente si usano per **rapporti di potenze**. Ad esempio il guadagno di un amplificatore:

$$A_{lineare} = \frac{P_{out}}{P_{in}} \Rightarrow A_{dB} \triangleq 10 \log_{10} \left(\frac{P_{out}}{P_{in}} \right)$$

Si ricorda che $\frac{P_{out}}{P_{in}}$ è un numero puro.

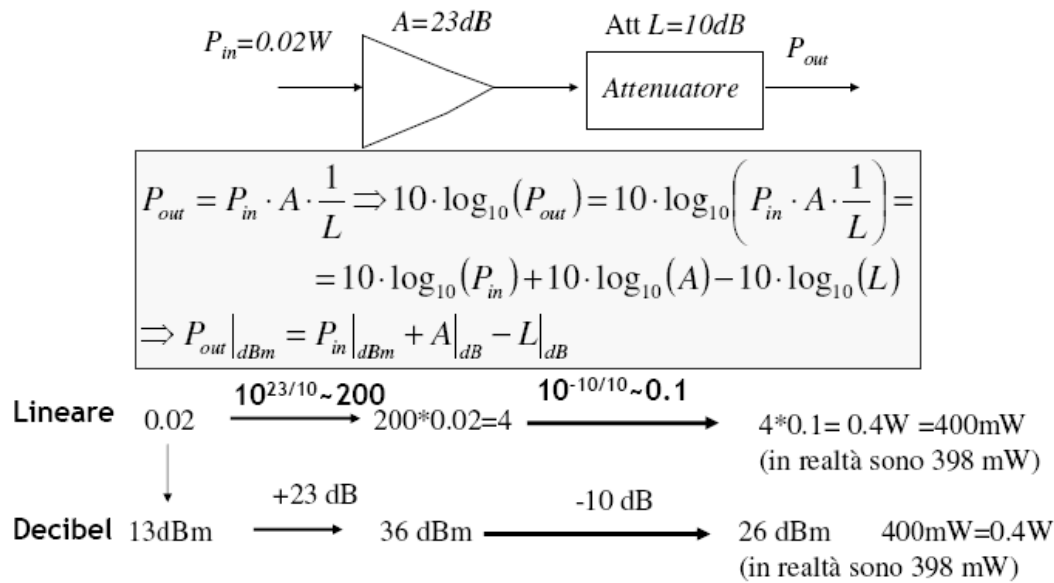


Figura 3.6: Esempio di calcolo svolto sia in lineare che in decibel.

Solo nel caso in cui si vogliano usare tensioni o correnti, si usano:

$$\left(\frac{V_1}{V_2}\right) \Big|_{\text{dB}} = 20 \log_{10} \left(\frac{V_1}{V_2}\right)$$

dB con
tensioni o
correnti

In questo corso useremo solamente rapporti di potenze.

Per comodità si usano spesso le scale logaritmiche anche per misurare potenze assolute. Ad esempio i dBm se si rapportano le potenze a 1 mW:

Importante:
dBm

$$P_{\text{dBm}} \triangleq 10 \log_{10} \left(\frac{P}{1 \text{ mW}}\right)$$

Ad esempio: $P = 0.02 \text{ W} = 20 \text{ mW} \Rightarrow P_{\text{dBm}} = 13 \text{ dBm}$.

Si possono poi introdurre altre quantità, ad esempio:

- dBW: potenze rapportate a 1 Watt;
- dBmV: tensioni rapportate a 1 mV.

La figura 3.6 mostra un esempio di calcolo svolto sia in lineare che in decibel. Si ha una potenza di 0.02 W all'ingresso di un sistema composto dalla cascata di un amplificatore seguito da un attenuatore (tale sistema viene chiamato *tratta* e sarà analizzato in dettaglio durante il corso). La figura riporta i conti, svolti sia in lineare che in decibel, per calcolare la potenza di uscita P_{out} .

3.4 Definizioni di banda

Il termine *banda* (chiamata anche *occupazione spettrale*) si presta a volte ad alcune incomprensioni. **Quando si parla di occupazione spettrale di un segnale reale,**

Banda: solo
frequenze
positive

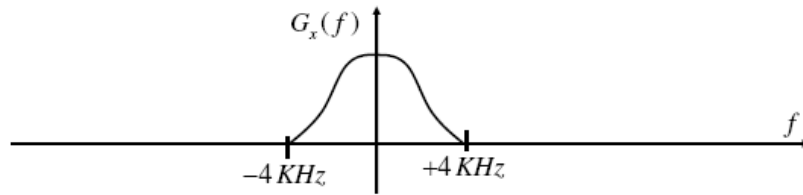


Figura 3.7: Esempio di segnale in banda base. La sua banda vale 4 kHz.

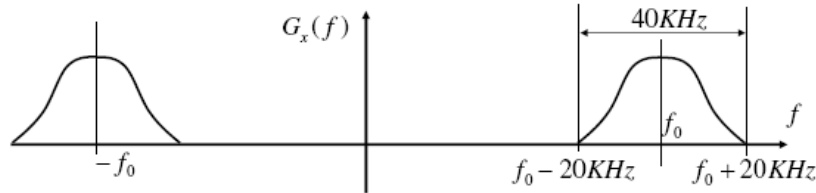


Figura 3.8: Esempio di segnale in banda traslata. La sua banda vale 40 kHz.

ci si riferisce sempre e solo all'asse positivo delle frequenze. Le frequenze negative infatti "esistono" solo nel senso matematico delle trasformate di Fourier, ma non hanno utilità "pratica"². Si ricorda che per un qualunque segnale reale, la trasformata di Fourier ha proprietà di simmetria rispetto a $f = 0$. Conseguentemente, la parte negativa delle frequenze di una trasformata di può ricostruire "per simmetria" da quella positiva.

A parte la (importante) questione relativa alle frequenze positive, un'ulteriore fonte di ambiguità deriva dal fatto che esistono varie definizioni di bande. Vediamo le principali.

Definizioni di banda più usate

Banda assoluta: intervallo di frequenze al di là del quale la densità spettrale di potenza (oppure $H(f)$) è *nulla*. Si tratta di una definizione matematica. In realtà qualunque segnale reale, avendo supporto temporale limitato, ha uno spettro che non va mai a zero. In questo corso la banda assoluta non la utilizzeremo praticamente mai (a parte per i filtri ideali).

Banda a x% di potenza: intervallo di frequenze al cui interno è contenuta x% di potenza del segnale. Spesso si usa B_{95} , cioè la banda che contiene il 95% della potenza del segnale.

Banda equivalente di rumore: di solito questa definizione si applica solo ai filtri. È definita così:

$$B_{EQ} \triangleq \frac{1}{|H(f)|_{MAX}^2} \int_0^{+\infty} |H(f)|^2 df$$

In questo modo un filtro generico, per quel che riguarda le densità spettrali di potenza, è reso equivalente a un filtro passa-basso ideale.

²Nessuno si è mai sognato di sintonizzare la propria radio sulla frequenza di -105.5 MHz.

In questo corso, intenderemo come banda **l'intervallo di frequenze al di là del quale lo spettro del segnale risulta (dal punto di vista ingegneristico) sostanzialmente nullo** (considereremo cioè, in sostanza, sempre la banda a $x\%$ di potenza). Non considereremo mai la banda assoluta di un segnale.

I concetti appena visti si applicano anche ai filtri.

Esempio 3.5. larghezza di banda di alcuni segnali Si consideri il segnale in figura 3.7. Per tale segnale, si intende che l'occupazione spettrale è pari a 4 kHz (e non 8 kHz!!). Per il segnale di figura 3.8 si intende che l'occupazione spettrale è pari a 40 kHz (e non 80 kHz!!). ■

3.5 Definizione di canale

La parola *canale*, nel campo delle trasmissioni, è usata con significati diversi in contesti diversi:

Il termine *canale* ha un duplice significato

- Canale nel senso di risorsa fisica per la trasmissione, cioè canale trasmissivo. Ad esempio etere, cavo, fibra ottica, etc.
- Canale nel senso di singolo segnale, tipicamente (ma non sempre) modulato. Ad esempio in frasi tipo "il terzo canale radio" o "il canale destro di un sistema stereofonico".

Nel seguito, salvo diversa indicazione, intenderemo la prima definizione, che è quella più comune.

Il canale introduce disturbi di varia natura sul segnale ricevuto. Solitamente questi disturbi sono distorsioni e rumore. Il progettista di un sistema di telecomunicazioni deve progettare il sistema di trasmissione in modo da contrastare gli effetti di questi disturbi sul segnale ricevuto.

Il canale introduce disturbi sul segnale

Si osservi che, se non ci fossero fenomeni di rumore e distorsione, sarebbe tecnicamente possibile instaurare trasmissioni altamente affidabili su distanze praticamente illimitate. Sfortunatamente, però, già la sola presenza del rumore limita fortemente l'affidabilità delle comunicazioni.

trasmesso

3.6 Propagazione guidata e in spazio libero

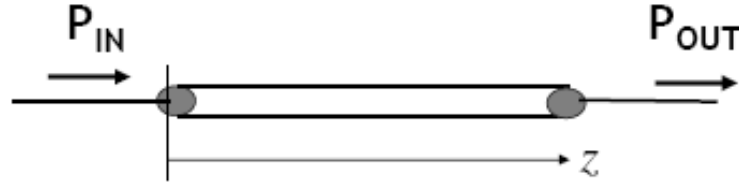
La propagazione di un segnale può essere:

- **Via cavo** (detta anche *propagazione guidata*). Attenuazione lungo il canale del tipo: $e^{-\gamma z}$. L'attenuazione è quindi esponenziale con la distanza.
- **Va etere**: attenuazione lungo il canale del tipo: K/z^2 . In questo caso l'attenuazione è inversamente proporzionale al quadrato della distanza.

3.6.1 Sistemi via cavo

Si consideri la figura 3.9. L'attenuazione di un tratto di cavo di lunghezza z , ad una determinata frequenza, è regolata da:

Attenuazione tratto di cavo

Figura 3.9: Tratto di cavo di lunghezza z .

$$P_{out} = P_{in} e^{-\gamma z} \quad (3.1)$$

dove γ è un coefficiente che dipende da:

- Frequenza di lavoro. γ tipicamente aumenta con la frequenza di lavoro.
- Caratteristiche del cavo. Migliore è il cavo, più piccolo è il valore di γ .

Se esprimiamo la relazione (3.1) in scale logaritmiche otteniamo:

$$10 \log_{10} \left(\frac{P_{out}}{1 \text{ mW}} \right) = 10 \log_{10} \left(\frac{P_{in} e^{-\gamma z}}{1 \text{ mW}} \right)$$

cioè:

$$P_{out}|_{dBm} = P_{in}|_{dBm} + 10 \log_{10} e^{-\gamma z} \quad (3.2)$$

Per l'ultimo termine della (3.2) possiamo applicare la proprietà dei logaritmi $\log_a N = \log_b N (1/\log_b a)$, ottenendo³:

$$10 \log e^{-\gamma z} = 10 \frac{1}{\ln 10} \ln e^{-\gamma z} = -\frac{10}{\ln 10} \gamma z$$

Sostituendo quest'ultimo termine nella (3.2) si trova:

$$\begin{aligned} P_{out}|_{dBm} &= P_{in}|_{dBm} - \frac{10}{\ln 10} \gamma z \\ &= P_{in}|_{dBm} - 4.34 \gamma z \end{aligned}$$

Formula In definitiva, ponendo $\alpha \triangleq 4.34 \gamma$, ottengo:

propagazione
via cavo

$$\boxed{P_{out}|_{dBm} = P_{in}|_{dBm} - \alpha|_{dB/km} \cdot z|_{km}} \quad (3.3)$$

dove α è detta *attenuazione per unità di lunghezza* e si esprime tipicamente in una delle seguenti unità di misura: [dB/m], [dB/100 m], [dB/km], [dB/feet]. In sostanza, l'attenuazione di un cavo lungo D_{km} è esprimibile tramite la formula (vedi figura 3.10):

$$L|_{dB} = \alpha|_{dB/km} \cdot D_{km}$$

α è un parametro che si trova sui data sheet dei cavi.

³Si ricorda che \ln indica il logaritmo naturale, cioè il logaritmo in base e .

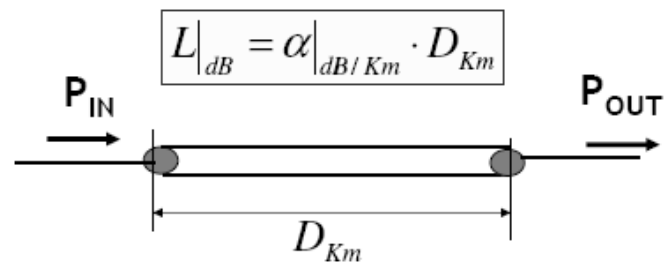


Figura 3.10: Attenuazione di un tratto di cavo di lunghezza D_{km} .

Esempio 3.6. alcuni tipi di cavo **Cavi "twisted pair"** di alta qualità, tipicamente utilizzati per reti locali (tipo 10-100 Mbit Ethernet):

UTP Cat 5 (dB/km) (dati da catalogo RS, cavo LSFN 100 MHz)

- A 1 MHz si ha 21.9 dB/km;
- A 16 MHz si ha 77.0 dB/km;
- A 100 MHz si ha 215.7 dB/km.

Cavi coassiali:

RG 62 A (dB/km)

- A 1 MHz si ha 8.2 dB/km;
- A 10 MHz si ha 27.9 dB/km;
- A 100 MHz si ha 88.6 dB/km.

Cavi coassiali per alte frequenze

RG 402 (dB/m)

- A 1 GHz si ha 0.36 dB/m;
- A 10 GHz si ha 1.37 dB/m.

Fibre ottiche:

- Hanno caratteristiche totalmente diverse dai cavi "elettrici";
- Attenuazione di 0.2 dB/km su bande elevatissime.

■

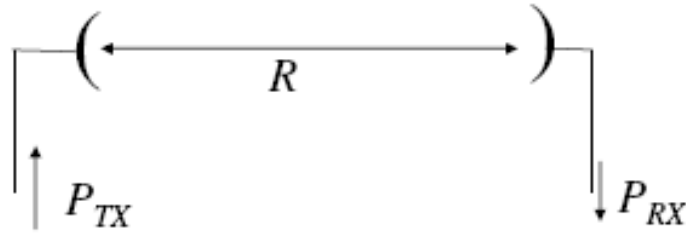


Figura 3.11: Rappresentazione schematica di un sistema di trasmissione via etere. Ci sono un'antenna trasmittente (TX) e un'antenna ricevente (RX) separate da una distanza R . P_{TX} e P_{RX} sono rispettivamente la potenza trasmessa e la potenza ricevuta.

3.6.2 Sistemi via etere

Si consideri ora un sistema di trasmissione via etere, schematizzato in figura 3.11. In questo caso ho un'antenna trasmittente (TX) e un'antenna ricevente (RX) separate da una distanza R .

Antenna

L'antenna è un componente fondamentale dei sistemi di trasmissione via etere, in quanto è l'elemento di interfaccia tra le grandezze elettromagnetiche e quelle elettriche. In trasmissione, essa ha il compito di trasformare la tensione elettrica al suo ingresso in onde elettromagnetiche in grado di propagarsi nell'ambiente. In ricezione svolge il ruolo inverso.

Potenza
ricevuta via
etere

La potenza ricevuta, per un sistema via etere tipo quello di figura 3.11, dipende dalle seguenti grandezze:

- Potenza trasmessa P_{TX} ;
- Distanza tra le antenne R ;
- Caratteristiche delle antenne;
- Frequenza di trasmissione.

Antenna
isotropica

Si consideri un'antenna isotropica: è un'antenna che emette in tutte le direzioni nello stesso modo. Chiaramente si tratta di un'antenna non fisicamente realizzabile, ma utile come riferimento concettuale per descrivere le antenne reali. Data un'antenna isotropica, la potenza per unità di superficie ad una distanza R è data da⁴:

$$\frac{P_{TX}}{4\pi R^2} \quad \left[\frac{W}{m^2} \right]$$

Antenna
reale (non
isotropica)

Un'antenna reale emetterà in maniera diversa a seconda delle direzioni. Fissata una direzione, si definisce il **guadagno d'antenna** G_{TX}^{ant} come guadagno (in potenza) rispetto al radiatore isotropico. Il guadagno d'antenna è un numero puro e solitamente

⁴ Si ricorda che l'area di una circonferenza di raggio R vale $4\pi R^2$.

si esprime in dB. In questo modo la densità di potenza per unità di superficie⁵ trasmessa da un'antenna reale vale:

$$\frac{P_{TX}}{4\pi R^2} G_{TX}^{ant} \quad \left[\frac{W}{m^2} \right]$$

L'antenna ricevente è poi caratterizzata da un'area equivalente A_{RX}^{ant} (che si misura in $[m^2]$). La potenza ricevuta è data dalla densità di potenza sull'antenna ricevente, moltiplicata per la sua area equivalente. La potenza ricevuta sarà allora data dalla seguente espressione, chiamata equazione della trasmissione (o equazione di Friis):

$$P_{RX} = \frac{P_{TX}}{4\pi R^2} G_{TX}^{ant} A_{RX}^{ant} \quad (3.4)$$

Equazione
della
trasmissione

A_{RX}^{ant} è legata a:

Area
equivalente

- Dimensioni fisiche dell'antenna;
- Forma dell'antenna;
- Allineamento/direzione;
- Frequenza di lavoro.

È importante sottolineare che nella (3.4) P_{RX} è la *potenza disponibile* al ricevitore. Si tornerà sulla questione della potenza disponibile nei prossimi capitoli.

Esiste la seguente relazione tra guadagno e area equivalente di un'antenna (a causa della reciprocità TX-RX di un'antenna):

$$\frac{A^{ant}}{G^{ant}} = \frac{\lambda^2}{4\pi} \quad (3.5)$$

Legame tra
guadagno e
area
equivalente

dove $\lambda = c/f$ è la lunghezza dell'onda elettromagnetica, e $c = 3 \cdot 10^8 m/s$ è la velocità della luce nel vuoto.

Utilizzando il legame tra guadagno e area equivalente (3.5), l'equazione della trasmissione (3.4) si può anche scrivere in una delle seguenti due forme:

Altre forme
equazione
della
trasmissione

$$P_{RX} = P_{TX} \frac{A_{TX}^{ant} A_{RX}^{ant}}{(\lambda R)^2} \quad (3.6)$$

$$P_{RX} = P_{TX} \frac{G_{TX}^{ant} G_{RX}^{ant} \lambda^2}{(4\pi R)^2} \quad (3.7)$$

Queste equazioni determinano le **attenuazioni minime** di un generico sistema di propagazione via etere. A questi valori minimi di attenuazione si deve aggiungere l'attenuazione dovuta all'interazione con l'atmosfera, che può aumentare l'attenuazione totale in maniera considerevole, ed è fortemente dipendente dalle frequenze e dalle condizioni ambientali. Spesso la extra-attenuazione rispetto alle equazioni precedenti è espressa tramite parametri caratteristici espressi in [dB/km].

Atmosfera
introduce
un'extra-
attenuazione

⁵La densità di potenza per unità di superficie definita in questa pagina non è da confondersi con la densità spettrale di potenza, che è un concetto totalmente diverso.

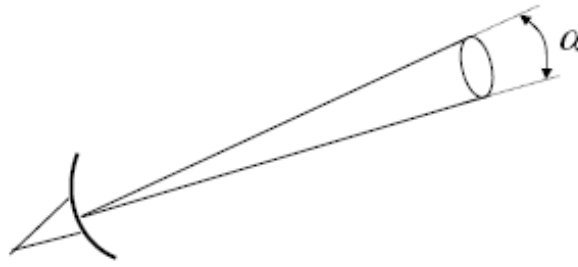


Figura 3.12: Antenna direttiva con apertura angolare α .

Antenna direttiva con diagramma a cono

Si consideri un'antenna che emette su una sezione dello spazio di forma conica (vedi figura 3.12). Per un'antenna di questo tipo si dimostra che:

Guadagno
antenna
direttiva su
cono con
apertura α

$$G^{ant} \approx \frac{52525}{\alpha_{gradi}^2} \quad (3.8)$$

dove α è l'apertura angolare del cono in gradi. Questo tipo di antenna viene spesso indicata come:

- *Antenna direttiva* su un settore angolare α ;
- Oppure: antenna ad elevata direttività.

Esempio 3.7. confronto tra sistemi via etere e via cavo Supponiamo di dover coprire una distanza di 10 km.

Sistema via
cavo

Usando un cavo di tipo RG62A (già considerato in un esempio precedente), e ipotizzando una trasmissione attorno a 100 MHz, avrò:

$$\text{attenuazione} = 88.6 \text{ dB/km}$$

$$\text{attenuazione totale} = 886 \text{ dB}$$

se fosse dato $P_{TX} = 1 \text{ mW}$, avrei:

$$\begin{aligned} P_{RX} &= 10^{-3} \cdot 10^{-886/10} \\ &= 2.5 \cdot 10^{-92} \text{ W!!!!} \end{aligned}$$

Sistema in
fibra ottica

Utilizzando un sistema in fibra ottica, e ipotizzando una trasmissione attorno ai 100 MHz:

$$\text{attenuazione} = 0.2 \text{ dB/km}$$

$$\text{attenuazione totale} = 2 \text{ dB}$$

se fosse dato $P_{TX} = 1 \text{ mW}$, avrei:

$$P_{RX} = 10^{-3} \cdot 10^{-2/10} = 0.631 \text{ mW}$$

Sistema via
etere

Utilizzando un sistema via etere: ipotizzo una trasmissione attorno a 1 GHz e antenne

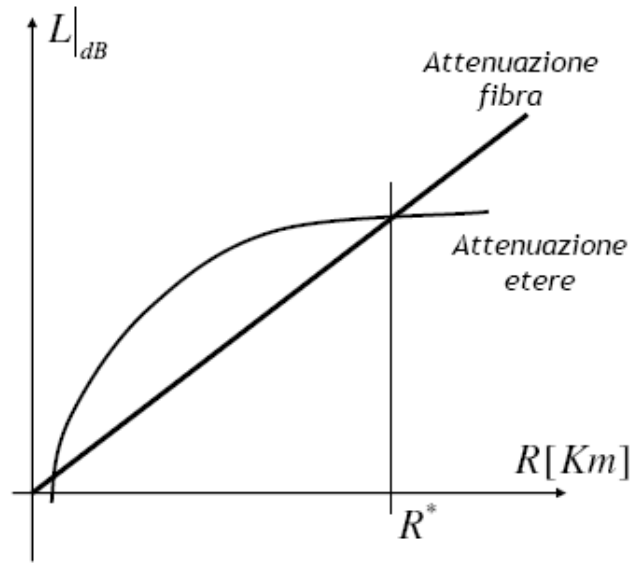


Figura 3.13: Confronto sistema via etere e fibra ottica.

con aree equivalenti di 1 m^2 , $R = 10 \text{ km}$. Avrò:

$$P_{RX} = P_{TX} \frac{A_{TX}^{ant} A_{RX}^{ant}}{(\lambda R)^2} \quad \lambda = c/f = 3 \cdot 10^8 / 10^9 = 0.3 \text{ m}$$

$$\frac{P_{TX}}{P_{RX}} = \frac{(\lambda R)^2}{A_{TX}^{ant} A_{RX}^{ant}} = 9 \cdot 10^6 \Rightarrow 65.9 \text{ dB}$$

se fosse $P_{TX} = 1 \text{ mW}$, avrei:

$$P_{RX} = 10^{-3} \cdot 10^{-69.5/10} = 1.122 \cdot 10^{-10} \text{ W}$$

Confrontiamo le attenuazioni equivalenti di:

- Fibra ottica 0.2 dB/km ;
- Collegamento via etere con le caratteristiche appena viste.

Si ha:

- Fibra: attenuazione $_{dB} = 0.2 \cdot R_{km}$
- Via etere: attenuazione $= 9 \cdot 10^4 \cdot R_{km}^2$,
attenuazione $_{dB} = 49.54 + 20 \cdot \log_{10}(R_{km})$.

Si consideri la figura 3.13. Il punto di incontro tra i due andamenti si ottiene per:

$$R^* = 519.2 \text{ km}$$

Con cavo coassiale (attenuazione 88.6 dB/km) si ha:

$$R^* = 0.489 \text{ km}$$

Confronto
sistema in
fibra ottica
con sistema
via etere

■

3.7 Riassunto formule principali sulla propagazione

Si riportano in questa sezione le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi, e sono tutte presenti anche sul formulario del corso (che si può scaricare dal sito del corso). Si ricorda che il formulario è l'unico materiale che è consentito tenere in sede d'esame.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- Propagazione via cavo:

$$P_{out}|_{dBm} = P_{in}|_{dBm} - \alpha|_{dB/km} \cdot z|_{km}$$

- Propagazione via etere:

$$\begin{aligned} \frac{A^{ant}}{G^{ant}} &= \frac{\lambda^2}{4\pi} \\ P_{RX} &= \frac{P_{TX}}{4\pi R^2} G_{TX}^{ant} A_{RX}^{ant} \\ P_{RX} &= P_{TX} \frac{A_{TX}^{ant} A_{RX}^{ant}}{(\lambda R)^2} \\ P_{RX} &= P_{TX} \frac{G_{TX}^{ant} G_{RX}^{ant} \lambda^2}{(4\pi R)^2} \end{aligned}$$

- Guadagno antenna direttiva su cono con apertura α :

$$G^{ant} \approx \frac{52525}{\alpha_{gradi}^2}$$

Capitolo 4

Trasmissioni analogiche in banda base

4.1 Introduzione

Solitamente, una sorgente analogica genera un segnale in banda base. La trasmissione di questo segnale può poi essere in banda base o in banda traslata. Si faccia riferimento alla figura 8.1.

Trasmissione in banda base: il segnale inviato sul canale trasmissivo ha un contenuto spettrale centrato attorno alla frequenza zero. Non si sono cioè effettuate conversioni significative.

Trasmissione in banda traslata: il segnale inviato sul canale trasmissivo ha un contenuto spettrale centrato attorno ad una frequenza f_0 (diversa da 0). Solitamente in questo caso si usa $f_0 \gg B$, dove B è la banda del segnale da trasmettere. Il segnale in banda traslata si ottiene tramite modulazione. Questa soluzione è di gran lunga la più comune per le trasmissioni analogiche.

Questo capitolo tratta brevemente le trasmissioni analogiche in banda base. Queste sono importanti perché:

1. Saranno utilizzate come riferimento per il confronto con altri schemi di trasmissione più complicati (che esamineremo nei prossimi capitoli);
2. Hanno comunque rilevanza pratica:
 - **Esempio:** in un impianto di amplificazione da concerto, il segnale audio captato dal microfono è gestito tutto in banda base (dai vari amplificatori, cavi, fino alle casse);
 - **Esempio:** vari schemi di trasmissione a bassissimo costo (come il telefono tradizionale sul tratto iniziale, il citofono, etc.) sono direttamente in banda base.

Utilità
trasmissioni
in banda
base

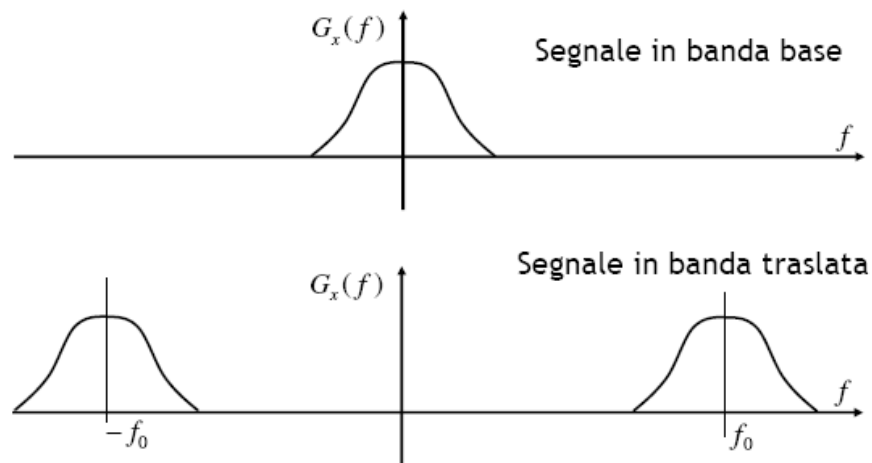


Figura 4.1: Esempio di segnale in banda base (sopra) e segnale in banda traslata (sotto).

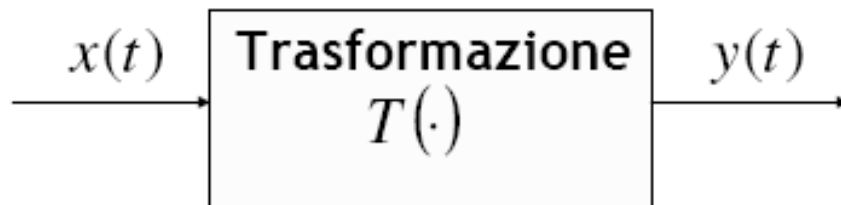


Figura 4.2: Trasmissione di un segnale in banda base su una generica trasformazione $T(\cdot)$.

4.2 Trasmissioni analogiche in banda base

Vogliamo considerare la trasmissione di un segnale in banda base su una generica trasformazione $T(\cdot)$, che in generale può essere affetta da:

- Distorsione lineare (filtraggio, sistema LTI);
- Distorsione non-lineare;
- Rumore (solitamente additivo).

La situazione è rappresentata in figura 4.2.

4.2.1 Canale ideale

Un canale viene detto **ideale** se è affetto solo da:

- Scalamento di ampiezze (fattore di scala o guadagno);
- Ritardo;
- No rumore.

Cioè:

$$y(t) = A_p x(t - t_0) \quad A_p \neq 0$$

corrispondente alla funzione di trasferimento:

$$H(f) = A_p e^{-j2\pi f t_0}$$

cioè un canale a:

- Modulo costante con f ;
- Fase lineare con f ;
- Ritardo di gruppo costante.

Nei sistemi pratici si cerca sempre di avere $H(f)$ "ideale" sulla banda del segnale da trasmettere.

4.2.2 Canale non ideale

Per un generico canale, si potrà sempre scrivere:

$$y(t) = \underbrace{A_p x(t - t_0)}_{\text{Segnale utile}} + \underbrace{m(t)}_{\text{Disturbo}}$$

Il primo termine viene detto *segnale utile* ed è quello che si avrebbe all'uscita di un canale ideale. Il secondo termine $m(t)$ è una generica componente di *disturbo*. Possiamo allora definire il rapporto segnale-disturbo come:

Rapporto
segnale-
disturbo

$$\frac{S}{N} \triangleq \frac{\text{Potenza segnale}}{\text{Potenza disturbo}}$$

Il rapporto segnale-disturbo solitamente si indica come S/N oppure SNR (Signal to Noise Ratio).

Usando le ipotesi di processi casuali stazionari, ergodici, a valor medio nullo e trasformazioni tempo-invarianti, si arriva alla seguente espressione del rapporto segnale-disturbo:

$$\begin{aligned} \frac{S}{N} &= \frac{E[(A_p x(t - t_0))^2]}{E[m^2(t)]} \\ &= \frac{A_p^2 \sigma_x^2}{\sigma_m^2} \end{aligned}$$

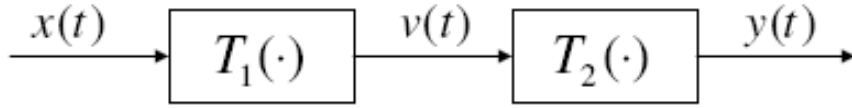
Con alcuni passaggi (non banali) è possibile dimostrare che:

$$\frac{S}{N} = \frac{1}{\frac{1}{\rho^2} - 1}$$

dove:

$$\rho = \max_{t_0} \left\{ \frac{E[y(t) x(t - t_0)]}{\sigma_x \sigma_y} \right\}$$

ρ è il massimo del coefficiente di correlazione tra $x(t)$ e $y(t)$. Il coefficiente va massimizzato rispetto a t_0 .

**Figura 4.3:** Cascata di due trasformazioni.

4.2.3 Cascata di due trasformazioni

Si consideri la cascata di due trasformazioni (vedi figura 4.3). È possibile dimostrare (con passaggi non banali) che per la cascata di due trasformazioni generiche si ottiene:

$$\left(\frac{S}{N}\right)_{TOT} = \frac{\left(\frac{S}{N}\right)_1 \left(\frac{S}{N}\right)_2}{1 + \left(\frac{S}{N}\right)_1 + \left(\frac{S}{N}\right)_2}$$

dove:

$\left(\frac{S}{N}\right)_1$ rapporto segnale/disturbo relativo alla prima trasformazione

$\left(\frac{S}{N}\right)_2$ rapporto segnale/disturbo relativo alla seconda trasformazione

Nell'ipotesi che i rapporti siano sufficientemente elevati, si ottiene:

$$\left(\frac{S}{N}\right)_{TOT} \approx \frac{\left(\frac{S}{N}\right)_1 \cdot \left(\frac{S}{N}\right)_2}{\left(\frac{S}{N}\right)_1 + \left(\frac{S}{N}\right)_2} \quad (4.1)$$

da cui segue:

$$\left(\frac{S}{N}\right)_{TOT} \leq \min \left\{ \left(\frac{S}{N}\right)_1, \left(\frac{S}{N}\right)_2 \right\}$$

S/N cascata
di due tra-
sformazioni

Cioè il rapporto segnale-disturbo complessivo è sempre inferiore al minore dei due rapporti sulle singole trasformazioni. Si invita il lettore a riflettere sul perché di questa disuguaglianza. La (4.1) è la stessa operazione che si effettua in elettrotecnica quando si effettua il parallelo di due resistenze.

4.2.4 Esempio 1: canale ideale + rumore

In questo caso si ha:

$$y(t) = A_p x(t - t_0) + n(t)$$

Dopo alcuni passaggi è possibile ottenere:

$$\begin{aligned} E[y(t) x(t - t_0)] &= A_p \sigma_x^2 \\ \sigma_y^2 &= A_p^2 \sigma_x^2 + \sigma_n^2 \\ \rho^2 &= \frac{A_p^2 \sigma_x^2}{A_p^2 \sigma_x^2 + \sigma_n^2} \end{aligned}$$

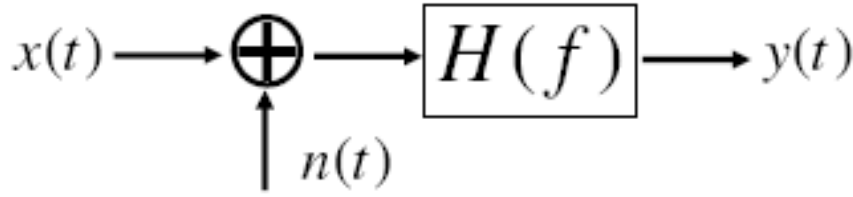


Figura 4.4: Segnale + rumore filtrati.

da cui si ottiene:

$$\begin{aligned} \left(\frac{S}{N} \right) &= \frac{\text{Potenza segnale}}{\text{Potenza rumore}} \\ &= \frac{A_p^2 \sigma_x^2}{\sigma_n^2} \end{aligned}$$

Ad esempio, se il rumore è gaussiano bianco filtrato con un filtro a banda equivalente di rumore B_{eq} , si ottiene:

$$\left(\frac{S}{N} \right) = \frac{\text{Potenza segnale}}{N_0 B_{eq}}$$

Questo sistema è da intendersi come riferimento per i sistemi di trasmissione in banda base (senza modulazione), e verrà preso come riferimento di confronto per tutti i sistemi di trasmissione analogici modulati.

4.2.5 Esempio 2: segnale + rumore gaussiano e filtro

Si consideri il sistema di figura 4.4. I conti sono abbastanza complessi in questo caso, e i risultati dipendono fortemente dalle caratteristiche del filtro e del segnale utile. Daremo solo un risultato qualitativo. Siano:

- B_x = banda del segnale $x(t)$;
- B_F = banda filtro $H(f)$.

La figura 4.5 mostra un risultato qualitativo. Quando il filtraggio sia "ottimizzato", mi riduco al caso dell'esempio precedente e ho che con buona approssimazione:

S/N
trasmissione
in banda
base

$$\boxed{\left(\frac{S}{N} \right) = \frac{\text{Potenza segnale}}{N_0 B_x}} \quad (4.2)$$

dove B_x è la banda del segnale utile.

Si tratta dunque di un sistema equivalente a quello mostrato in figura 4.6.

4.3 Riassunto formule principali sulla trasmissione in banda base

Si riportano in questa appendice le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi, e sono tutte presenti anche sul formulario del corso (che si può scaricare dal sito del corso). Si ricorda



Figura 4.5: Andamento qualitativo del rapporto segnale-disturbo all'uscita del sistema di figura 4.4 (segnale + rumore filtrati).

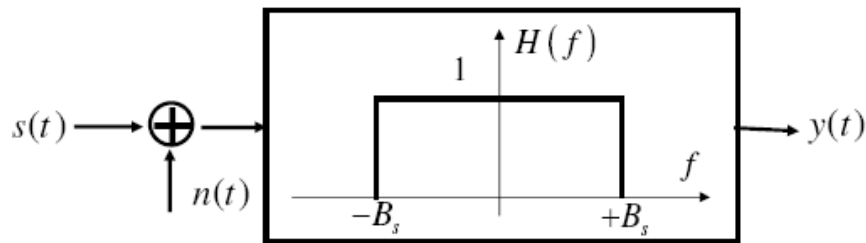


Figura 4.6: Equivalente del sistema della figura 4.4, qualora il filtraggio sia ottimizzato.

che il formulario è l'unico materiale che è consentito tenere in sede d'esame.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- Trasmissione di segnale analogico in banda base:

$$\left(\frac{S}{N}\right) = \frac{P_{RX}}{N_0 B_x}$$

Capitolo 5

Rumore termico e doppi bipoli

In questo capitolo si tratta l'argomento del rumore, in particolare del rumore termico. Si intende qui come **rumore** un qualunque disturbo additivo sul segnale utile che abbia caratteristiche stocastiche.

In particolare, si tratterà il rumore più comunemente riscontrato negli apparati elettrici, cioè il **rumore gaussiano bianco additivo**. Questo genere di rumore è essenziale per gran parte dei sistemi di telecomunicazioni, perché determina le prestazioni finali del sistema stesso.

5.1 Rumore termico

Si consideri un generico resistore con resistenza pari a R , posto a una temperatura T diversa dallo zero assoluto (vedi figura 5.1). Il moto casuale degli elettroni al suo interno genera una certa tensione di rumore $v(t)$, anche se non vengono applicati campi elettrici dall'esterno. Il moto è sempre non nullo tranne che allo zero assoluto. **Tale fenomeno è denominato rumore termico ed è fondamentale per tutte le applicazioni di telecomunicazioni.**

5.1.1 Descrizione probabilistica del rumore termico

La prima domanda che sorge spontanea è quale possa essere la caratterizzazione matematica di questo fenomeno. Poiché la tensione di rumore è generata dal moto *casuale* degli elettroni all'interno del resistore, risulta naturale modellare tale fenomeno con un **processo casuale**.

Il rumore termico è un processo casuale

Essendo la tensione $v(t)$ generata dalla somma dei contributi dei campi elettrici di un gran numero di elettroni, possiamo fare appello al teorema del limite centrale e dire che la densità di probabilità di tale processo è **gaussiana**.

Inoltre, non c'è alcuna ragione che spinga gli elettroni ad andare da una parte piuttosto che da un'altra, quindi possiamo supporre che il processo casuale abbia **valore medio nullo**.

Supponendo poi che le condizioni del resistore e il suo ambiente esterno non varino nel tempo, possiamo infine ipotizzare che tale processo casuale sia **stazionario**.

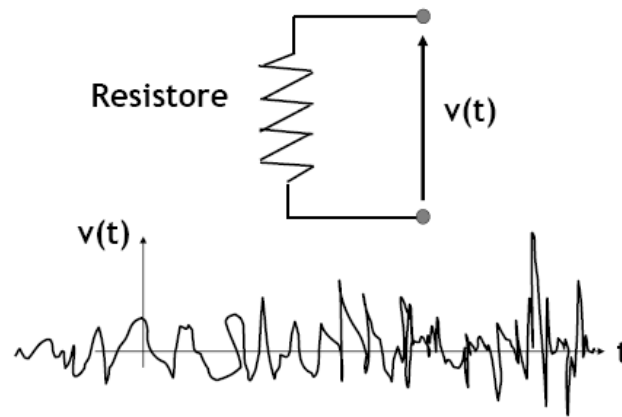


Figura 5.1: Tensione di rumore ai capi di un resistore.

Riassumendo, **il rumore termico può essere modellato con un processo casuale gaussiano, stazionario e a valor medio nullo.**

Resta infine da valutare la densità spettrale di potenza.

Densità spettrale di potenza del rumore termico

Mediante considerazioni di meccanica quantistica che esulano dagli scopi di questo corso, si può calcolare la densità spettrale di potenza del processo $v(t)$ tensione ai capi di un resistore con resistenza R :

Densità
spettrale di
potenza del
rumore

$$G_v(f) = 2R \left(\frac{h|f|}{2} + \frac{h|f|}{e^{\frac{h|f|}{kT}} - 1} \right) \quad \left[\frac{\text{V}^2}{\text{Hz}} \right] \quad (5.1)$$

dove:

- Il termine $\frac{h|f|}{2}$ è rumore dovuto al principio di indeterminazione e il termine $\frac{h|f|}{e^{\frac{h|f|}{kT}} - 1}$ è rumore dovuto al moto termico;
- $k = 1.38 \cdot 10^{-23} \text{ J/}^\circ\text{K}$ è la costante di Boltzmann;
- $h = 6.63 \cdot 10^{-34} \text{ J} \cdot \text{s}$ è la costante di Planck;
- T = temperatura del resistore in $^\circ\text{K}$ (si ricorda che $0^\circ\text{C} = 273^\circ\text{K}$).

Per prima cosa occorre commentare l'unità di misura dell'espressione (5.1). Nel corso di Teoria dei Segnali si consideravano sempre funzioni del tempo adimensionate. Conseguentemente, le densità spettrali di potenza avevano unità di misura $[1/\text{Hz}]$. In questo corso invece le funzioni del tempo avranno spesso unità di misura. Nel caso della (5.1), il segnale nel tempo $v(t)$ è misurato in $[\text{V}]$, quindi la corrispondente densità spettrale di potenza è espressa in $[\text{V}^2/\text{Hz}]$. Si noti come tale grandezza non sia propria-

mente una potenza dal punto di vista dimensionale¹.

Ci chiediamo a questo punto se sia possibile approssimare la (5.1) con un'espressione più semplice da trattare analiticamente. La risposta è la seguente. Per tutte le applicazioni tipiche delle comunicazioni classiche, e cioè per temperature tra 0° e 50° C e per frequenze minori di 1 THz² (nel seguito si considerano solo frequenze positive, quindi si scriverà f invece di $|f|$), si ha che: Approssimazione della (5.1)

$$\frac{hf}{KT} \ll 1 \quad (5.2)$$

Possiamo allora usare lo sviluppo di McLaurin arrestato al primo ordine $e^x \approx 1 + x$ e scrivere:

$$e^{\frac{hf}{KT}} - 1 \approx \frac{hf}{KT}$$

La (5.1) allora diventa:

$$G_v(f) \approx 2R \left(\frac{hf}{2} + KT \right) \left[\frac{V^2}{\text{Hz}} \right]$$

L'espressione (5.2) però mi dice che $hf \ll KT$, quindi posso trascurare il termine $hf/2$. Solitamente quindi si può approssimare la (5.1) come: Rumore bianco

$$G_v(f) \approx 2RKT \left[\frac{V^2}{\text{Hz}} \right] \quad (5.3)$$

L'approssimazione (5.3) viene comunemente fatta nell'ambito delle Comunicazioni Elettriche. Per questo motivo questo tipo di rumore viene anche detto **rumore bianco** (bianco significa spettralmente piatto, infatti l'espressione di $G_v(f)$ non dipende dalla frequenza).

Si noti che tale approssimazione non è indolore da un punto di vista matematico. Dalla (5.3) segue che la potenza del rumore è infinita, il che non è fisicamente realizzabile. Si ricordi però che qualsiasi sistema di comunicazione avrà all'interno del ricevitore un filtro di ricezione³, centrato attorno alla frequenza della portante del segnale che si vuole ricevere, e con banda B sempre molto inferiore al THz. Conseguentemente la densità spettrale di potenza della tensione di rumore dopo il filtro di ricezione sarà sempre piatta e limitata in banda proprio grazie a suddetto filtro. Quindi, in conclusione, l'approssimazione di rumore bianco (5.3) ha senso solamente *a valle del filtro di ricezione*. L'approssimazione (5.3) ha senso solo a valle del filtro di ricezione

Riassumendo, il rumore termico è con ottima approssimazione rappresentabile come un processo stocastico:

- Stazionario;
- Gaussiano;

¹Per tale motivo per la (5.1) il termine più corretto sarebbe *densità spettrale del valore quadratico medio*. Tuttavia spesso si usa più semplicemente il termine *spettro di potenza* o *densità spettrale di potenza*. Riassunto caratteristiche statistiche rumore

²Vanno invece escluse le trasmissioni in fibra ottica, che avvengono a frequenze decisamente superiori, dell'ordine di $10^{14} - 10^{15}$ Hz.

³Questo è già stato visto per le trasmissioni in banda base e si vedrà nei prossimi capitoli che vale anche per la trasmissioni in banda traslata.



Figura 5.2: Circuito equivalente di Thevenin di un resistore rumoroso. Un resistore rumoroso R è equivalente a un resistore non rumoroso R posto in serie a un generatore di rumore $v(t)$.

- Bianco;
- A valor medio nullo;
- Ergodico.

Il valore efficace⁴, cioè la deviazione standard del rumore su una banda B è dato da:

$$\begin{aligned} V_{\text{eff}} &= \sqrt{\langle v^2(t) \rangle} \\ &= \sqrt{\int_{-B}^{+B} G_v(f) df} \\ &= \sqrt{4 R K T B} \quad [\text{V}] \end{aligned}$$

Esempio 5.1. calcolo potenza rumore su banda B Consideriamo una resistenza $R = 100 \, \Omega$ posta alla temperatura $T = 300^\circ \text{K}$. Il valore efficace della potenza di rumore su una banda $B = 1 \text{ MHz}$ vale:

$$\begin{aligned} V_{\text{eff}} &= \sqrt{4 R K T B} \\ &= 1.29 \cdot 10^{-6} \quad [\text{V}] \end{aligned}$$

■

Corrente di
rumore

Supponiamo ora di richiudere il resistore su se stesso. In questo caso la tensione di rumore indurrebbe una *corrente di rumore* la cui densità spettrale di potenza vale⁵:

$$\begin{aligned} G_I(f) &= \frac{G_v(f)}{R^2} \\ &= \frac{2 K T}{R} \left[\frac{\text{A}^2}{\text{Hz}} \right] \end{aligned}$$

Espressione che però in questo corso non utilizzeremo.

Circuito equivalente di un resistore rumoroso

Vogliamo dare ora una descrizione circuitale di un resistore rumoroso R . Poiché si tratta di un bipolo lineare, possiamo applicare il teorema di Thevenin (visto nei corsi di Elettrotecnica), che porta al modello circuitale mostrato in figura 5.2. Il circuito equivalente è costituito da un resistore non rumoroso R in serie a un generatore di tensione di rumore pari a $v(t)$.

Analogamente è possibile ricavare un circuito equivalente di Norton che però in questo corso non ci interessa.

⁴Detto anche RMS - Root Mean Square.

⁵Si invita il lettore a riflettere sul perché.

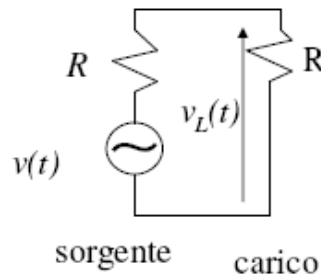


Figura 5.3: Resistenza rumorosa R a temperatura T chiusa su una resistenza ideale R . Siccome la resistenza di carico è uguale alla resistenza della sorgente, il circuito è adattato.

Densità spettrale di potenza disponibile di rumore

Al fine di valutare la densità spettrale di potenza disponibile del rumore, si consideri un resistore rumoroso avente resistenza R e posto alla temperatura T , e lo si chiuda su un resistore ideale⁶ di resistenza R_c . È noto dai corsi di Elettrotecnica che la situazione che consente di avere il massimo trasferimento di potenza dalla sorgente al carico è quella in cui $R_c = R$, cioè il carico abbia la stessa resistenza della sorgente. In tal caso si dice che il circuito è **adattato**. Adattamento energetico

In tal caso, facendo riferimento alla figura 5.3, abbiamo che la tensione di rumore ai capi del carico vale:

$$v_L(t) = \frac{v(t)}{2}$$

e quindi la potenza assorbita dal carico è uguale alla potenza disponibile:

$$\begin{aligned} P_L(t) &= \frac{v_L^2(t)}{R} \\ &= \frac{v^2(t)}{4R} \end{aligned}$$

Allora la densità spettrale di rumore disponibile ai capi di una resistenza è data da:

$$G_d(f) = \frac{G_v(f)}{4R} = \frac{KT}{2} \left[\frac{\text{W}}{\text{Hz}} \right] \quad (5.4)$$

dove il pedice d sta per disponibile. Solitamente si pone $N_0 \triangleq KT$ e si esprime la (5.4) nella forma:

$$G_d(f) = \frac{N_0}{2} \left[\frac{\text{W}}{\text{Hz}} \right]$$

che è già stata vista a Teoria dei Segnali ed è sicuramente quella più familiare per gli studenti.

A questo punto sono doverosi alcuni commenti sulla (5.4).

- Questa è effettivamente una potenza anche dal punto di vista dimensionale.

⁶Cioè non rumoroso.

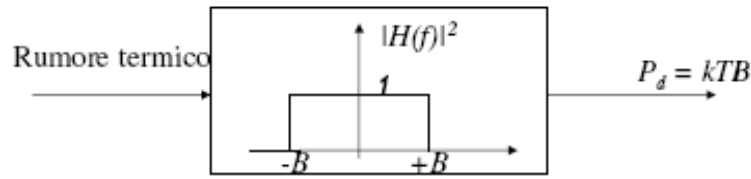


Figura 5.4: Rumore filtrato da filtro con banda equivalente di rumore B .

- Notare che **non dipende dal valore della resistenza**.
- Consideriamo il sistema della figura 5.4, cioè del rumore all'ingresso di un filtro con banda equivalente di rumore B .
 - All'ingresso la densità spettrale vale $G_d(f) = N_0/2$ con $N_0 \triangleq KT$.
 - All'uscita la potenza disponibile è pari a:

$$P_d = KTB \quad (5.5)$$

Come già osservato in precedenza, l'approssimazione di rumore bianco ha senso solo in un caso come questo di rumore filtrato.

Esempio 5.2. calcolo potenza disponibile di rumore su banda B Consideriamo una resistenza $R = 100 \, \Omega$ a temperatura $T = 300^\circ K$. Si pone la resistenza in ingresso al filtro di figura 5.4. Se la banda equivalente di rumore di questo filtro vale $B = 1 \, MHz$, in uscita avrò una potenza disponibile di:

$$\begin{aligned} P_d &= KTB \\ &= 4.14 \cdot 10^{-15} \, [W] \end{aligned}$$

■

Nelle telecomunicazioni ciò che conta è sempre la potenza ricevuta, e non tanto la tensione o la corrente. Inoltre, poiché interessa utilizzare al meglio la potenza ricevuta ⁷, si lavora solitamente in condizioni di adattamento energetico. Dunque ragioneremo sempre in termini di densità spettrale di potenza disponibile. **In seguito useremo sempre circuiti elettrici adattati.**

5.1.2 Temperatura equivalente di rumore di una sorgente

Vogliamo ora caratterizzare una generica sorgente dal punto di vista del rumore. Una sorgente con una certa resistenza interna R genererà una certa quantità di rumore, che può essere solo di tipo termico. Dal punto di vista sistemistico si definisce una *temperatura equivalente di rumore di sorgente*, che chiameremo T_{eq} ⁸. Data una banda di lavoro di riferimento, si tratta della temperatura a cui andrebbe posta una resistenza per generare la stessa potenza disponibile di rumore su quella banda. Facendo riferimento alla (5.5) ⁹, è immediato definire T_{eq} per una sorgente in questo modo:

Temperatura
equivalente
di rumore di
una sorgente

⁷Di segnale utile si intende, non di rumore.

⁸Da non confondere con la temperatura equivalente di un doppio bipolo, di cui parleremo dopo, che è un'altra cosa.

⁹Che valeva per una resistenza.

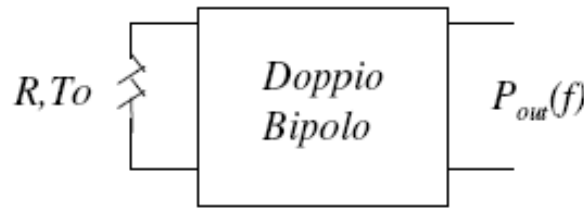


Figura 5.5: Rappresentazione schematica di un doppio bipolo lineare alimentato da una resistenza rumorosa. La resistenza rumorosa serve a modellare tutte le fonti di rumore all'ingresso del doppio bipolo.

$$T_{eq} \triangleq \frac{P_{\text{disponibile su banda } B}}{KB}$$

È una grandezza che può in generale dipendere dalla frequenza. In tal caso si indica come:

$$T_{eq}(f) = \frac{2G_d(f)}{K}$$

dove $G_d(f)$ è la densità spettrale di potenza di rumore disponibile in uscita dalla sorgente. Si noti che:

- T_{eq} è la temperatura fisica della resistenza se la sorgente è costituita da sole resistenze (adattate al carico);
- In generale per una sorgente generica **può non essere la temperatura ambiente** (spesso è superiore se la sorgente contiene dei dispositivi attivi, ma può essere inferiore nel caso in cui la sorgente sia un'antenna).

5.2 Caratterizzazione di doppi bipoli rumorosi

Vogliamo a questo punto capire come il rumore termico intervenga nei sistemi di comunicazione. Per fare questo il primo passo è caratterizzare dal punto di vista del rumore i doppi bipoli lineari, cioè tipicamente gli amplificatori e gli attenuatori (vedi figura 5.5). Dal nostro punto di vista un doppio bipolo si caratterizza con i seguenti parametri:

Caratterizzazione
di un doppio
bipolo

1. **Guadagno di potenza disponibile:**

$$A_d(f) \triangleq \frac{G_{out}(f)}{G_{in}(f)}$$

È il guadagno di *potenza di segnale* supponendo di avere tutto adattato. Attenzione che per la definizione di guadagno disponibile **non si considera l'effetto del rumore**. $G_{out}(f)$ è la densità spettrale di potenza disponibile in uscita e $G_{in}(f)$ è la densità spettrale di potenza disponibile in ingresso.

2. **Rumore.** Per questo si introdurranno la *temperatura equivalente di rumore* e la *cifra di rumore*.

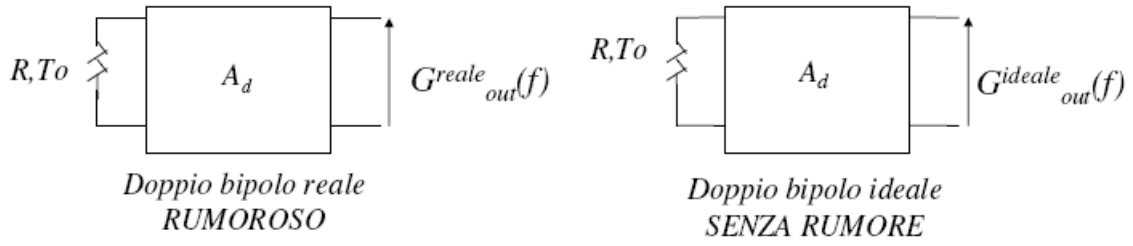


Figura 5.6: Esperimento ideale per determinare la cifra di rumore di un doppio bipolo.

Nel resto del corso, salvo diversamente specificato, si daranno sempre per scontati i seguenti tre punti.

1. I circuiti sono adattati. La resistenza di ingresso del doppio bipolo deve essere uguale alla resistenza della sorgente e la resistenza di uscita del doppio bipolo deve essere uguale alla resistenza di carico.
2. I guadagni e le densità spettrali di potenza sono definiti in questa situazione (e dunque sono delle quantità *disponibili*);
3. Il sistema è all'equilibrio termodinamico.

5.2.1 Cifra di rumore di un doppio bipolo

Per quantificare il rumore introdotto da un doppio bipolo, si fa l'esperimento ideale illustrato in figura 5.6. Si alimenta il doppio bipolo rumoroso con una resistenza rumorosa e si misura la densità spettrale di potenza di rumore disponibile in uscita $G_{out}^{reale}(f)$. Poi si considera il doppio bipolo non rumoroso e lo si alimenta con una resistenza rumorosa e si misura la densità spettrale di potenza di rumore disponibile in uscita $G_{out}^{ideale}(f)$. Si definisce poi la **cifra di rumore**¹⁰ in questo modo:

Cifra di
rumore di un
doppio
bipolo

$$\begin{aligned}
 F(f) &\triangleq \frac{G_{out}^{reale}(f)}{G_{out}^{ideale}(f)} \\
 &= \frac{\frac{KT_0}{2} A_d(f) + G_{interna}(f)}{\frac{KT_0}{2} A_d(f)}
 \end{aligned} \tag{5.6}$$

Importante: la cifra di rumore deve essere definita:

Cifra di
rumore va
definita con
sistema
adattato e
con
resistenza di
ingresso a
 $T = 290^\circ K$

- Con il sistema adattato;
- Con la resistenza di ingresso a $T_0 = 290^\circ K$.

Occorre fare subito alcune osservazioni sulla definizione (5.6).

- $F(f)$, essendo un rapporto di potenze, è un numero puro. Di solito si esprime in dB.

¹⁰Viene anche chiamata *figura di rumore*, in inglese *noise figure*.

- Poiché $G_{interna}(f) \geq 0$, la definizione (5.6) implica che $F(f) \geq 1$. Un doppio bipolo è tanto migliore (dal punto di vista del rumore) quanto più la sua cifra di rumore si avvicina a 1.
- $G_{out}^{reale}(f)$ è la somma della densità spettrale di rumore introdotta dal doppio bipolo ($G_{interna}(f)$) e della densità spettrale di rumore in uscita dovuta alla resistenza R che alimenta il doppio bipolo ($(KT_0/2) A_d(f)$). Le due densità spettrali si sommano perché i due contributi di rumore sono statisticamente indipendenti.

Spesso supporremo che la cifra di rumore sia indipendente dalla frequenza e la indicheremo quindi semplicemente con F .

Dalla definizione (5.6) segue che:

$$G_{out}^{reale}(f) = A_d(f) \frac{K}{2} T_0 F(f) \quad (5.7)$$

Questa espressione tornerà utile più avanti.

5.2.2 Temperatura equivalente di rumore di un doppio bipolo

Un altro modo per quantificare il rumore introdotto da un doppio bipolo è la **temperatura equivalente di rumore $T_{eq}(f)$** , definita implicitamente come:

$$G_{out}^{reale}(f) = A_d(f) \frac{K}{2} (T_0 + T_{eq}(f)) \quad (5.8)$$

Temperatura
equivalente
di rumore di
un doppio
bipolo

La temperatura equivalente di rumore si misura in $^{\circ}K$. È pari all'aumento *ideale* di temperatura che si deve dare al resistore rumoroso in ingresso al doppio bipolo per avere la corretta quantità di rumore in uscita, considerando il doppio bipolo ideale. Ovviamente si ha sempre $T_{eq} > 0^{\circ}K$.

È importante osservare che utilizzando la temperatura equivalente di rumore tutte le sorgenti di rumore vengono trasferite all'ingresso del sistema. Spesso supporremo che $T_{eq}(f)$ sia indipendente dalla frequenza e la indicheremo semplicemente con T_{eq} .

5.2.3 Formule di conversione tra cifra di rumore e temperatura equivalente di rumore

Vediamo come sia possibile passare da cifra di rumore a temperatura equivalente di rumore e viceversa. Supponiamo per semplicità che tutte le quantità in gioco siano indipendenti dalla frequenza, e dunque omettiamo la variabile f . La (5.7) diventa:

$$G_{out}^{reale} = A_d \frac{K}{2} T_0 F \quad (5.9)$$

e la (5.8) diventa:

$$G_{out}^{reale} = A_d \frac{K}{2} (T_0 + T_{eq}) \quad (5.10)$$

Eguagliando la (5.9) e la (5.10), si ricava:

$$T_0 + T_{eq} = T_0 F$$

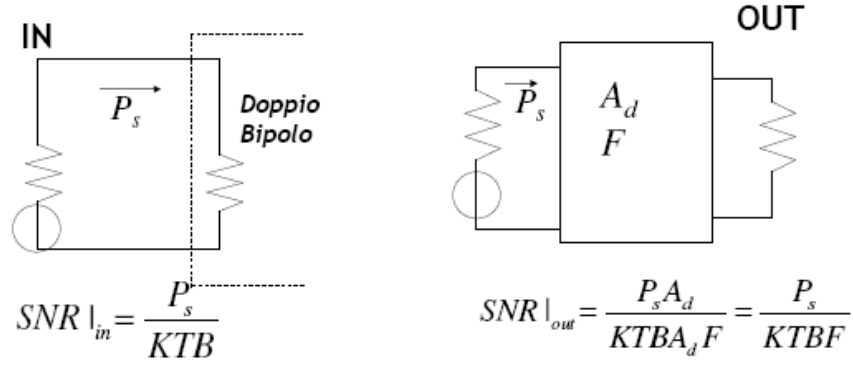


Figura 5.7: Interpretazione fisica della cifra di rumore. È il rapporto tra l'SNR in ingresso e in uscita da un doppio bipolo.

Formule di
conversione

F - T_{eq}

Quindi le formule di conversione tra F e T_{eq} sono:

$$F = 1 + \frac{T_{eq}}{T_0} \quad (5.11)$$

$$T_{eq} = T_0 (F - 1) \quad (5.12)$$

5.2.4 Interpretazione fisica della cifra di rumore

Vogliamo dare ora un'interpretazione fisica della cifra di rumore F . Si consideri la figura 5.7. Supponiamo di avere un doppio bipolo rumoroso con all'ingresso una potenza utile P_s e una potenza di rumore $KT/2$, entrambe su banda B . Il rapporto segnale-rumore in ingresso al doppio bipolo sarà:

$$SNR|_{in} = \frac{P_s}{KTB}$$

e quello in uscita (ricordando la (5.7)) sarà:

$$\begin{aligned} SNR|_{out} &= \frac{P_s A_d}{KTBA_d F} \\ &= \frac{P_s}{KTBF} \end{aligned}$$

A questo punto è chiara l'interpretazione fisica della cifra di rumore:

Interpretazione
fisica cifra di
rumore

$$F = \frac{SNR|_{in}}{SNR|_{out}} \quad (5.13)$$

La cifra di rumore è il rapporto tra il rapporto segnale-rumore in ingresso e in uscita dal doppio bipolo. Poiché $F \geq 1$, la (5.13) ci dice che il rapporto segnale-rumore in ingresso è sempre maggiore o uguale al rapporto segnale-rumore in uscita. Questo è un risultato che intuitivamente potevamo aspettarci.

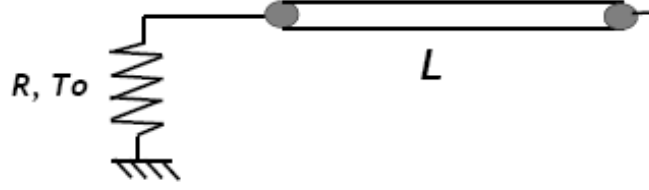


Figura 5.8: Attenuatore alimentato da una resistenza rumorosa.

5.2.5 Calcolo della cifra di rumore di un attenuatore

Consideriamo un attenuatore passivo con attenuazione L (costituito tipicamente da una linea di trasmissione tipo cavo coassiale) alimentato da una resistenza rumorosa (vedi figura 5.8). Il guadagno disponibile è dato da: $A_d = 1/L$. Supponiamo come al solito che il sistema sia adattato e alla temperatura T_0 .

Se tutto è adattato, dall'uscita vedo un sistema completamente passivo alla temperatura T_0 , e di conseguenza la densità spettrale dovrà essere la stessa che si ha all'uscita di una resistenza alla stessa temperatura, cioè:

$$G_{out}(f) = \frac{K}{2} T_0 \quad (5.14)$$

Dalla definizione (5.8) di temperatura equivalente di rumore, e ricordando che l'attenuatore ha $A_d = 1/L$, si ottiene:

$$\begin{aligned} G_{out}(f) &= A_d \frac{K}{2} (T_0 + T_{eq}) \\ &= \frac{K}{2} \frac{(T_0 + T_{eq})}{L} \end{aligned} \quad (5.15)$$

Uguagliando la (5.14) con la (5.15) trovo:

$$T_0 = \frac{T_0 + T_{eq}}{L}$$

quindi:

$$T_{eq} = T_0 (L - 1)$$

Grazie poi alla formula di conversione (5.11) si ottiene:

$$\begin{aligned} F &= 1 + \frac{T_{eq}}{T_0} \\ &= 1 + \frac{T_0 (L - 1)}{T_0} \\ &= L \end{aligned}$$

La cifra di rumore di un attenuatore è quindi pari alla sua attenuazione:

$$F = L$$

Cifra di
rumore di un
attenuatore

Si fa notare che questo è l'unico caso in cui si riescono a calcolare T_{eq} e F con considerazioni fisiche semplici.

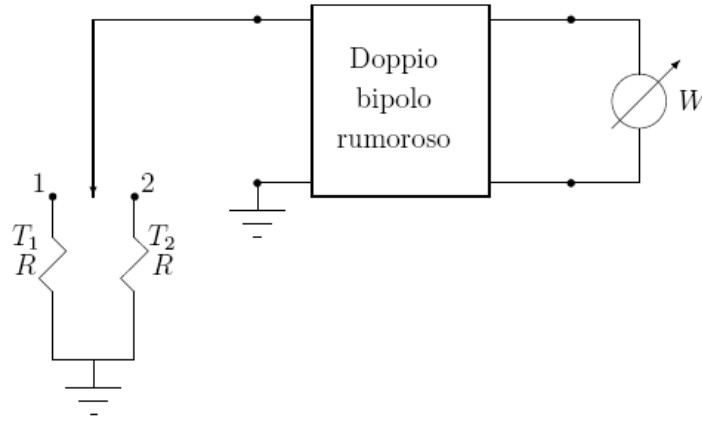


Figura 5.9: Banco di misura per la misura della temperatura equivalente di un doppio bipolo con la tecnica del caldo e freddo.

Esempio 5.3. misura della temperatura equivalente di un doppio bipolo Nella maggior parte dei casi T_{eq} di un doppio bipolo non può essere calcolata analiticamente, ma va misurata sperimentalmente. Un metodo per misurarla è la "tecnica del caldo e freddo". Consideriamo la figura 5.9. Il primo passo è alimentare il doppio bipolo con una resistenza R a temperatura T_1 e misurare la potenza di rumore disponibile P_1 in uscita:

$$P_1 = k (T_1 + T_{eq}) A_d B$$

Poi si alimenta il doppio bipolo con un'altra resistenza con lo stesso valore di R ma posta a una temperatura T_2 differente. Misurando la potenza di rumore disponibile in uscita si ha ora:

$$P_2 = k (T_2 + T_{eq}) A_d B$$

Facendo il rapporto tra P_1 e P_2 si trova:

$$\frac{P_1}{P_2} = \frac{T_1 + T_{eq}}{T_2 + T_{eq}}$$

Quindi T_{eq} vale:

$$T_{eq} = \frac{T_1 P_2 - T_2 P_1}{P_1 - P_2}$$

Dalla questa espressione si vede che, per avere una buona accuratezza della misura, T_1 e T_2 devono essere note con precisione e molto diverse tra di loro. Solitamente una resistenza si immerge nell'azoto liquido e l'altra in acqua bollente. ■

5.3 Cascate di doppi bipoli

Spesso nelle applicazioni si ha a che fare con cascate di doppi bipoli. Una cascata di doppi bipoli alimentati da un resistore rumoroso è un ottimo modello di un ricevitore per telecomunicazioni o di un sistema di trasmissione via cavo. Una cascata di doppi bipoli è, naturalmente, un doppio bipolo e pertanto per caratterizzarla dal punto di vista del rumore occorre definire la sua temperatura equivalente di rumore o, equivalentemente, la sua cifra di rumore.

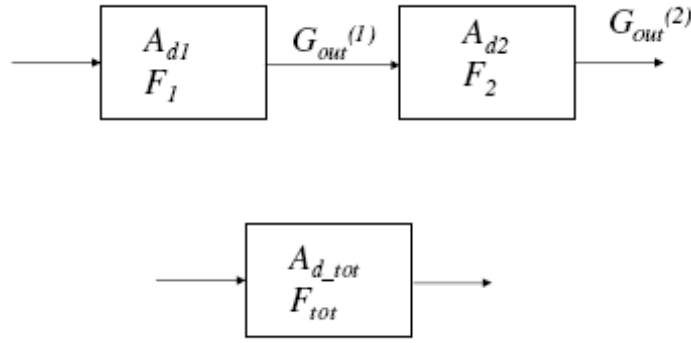


Figura 5.10: Cascata di due doppi bipoli (sopra) e doppio bipolo equivalente (sotto).

5.3.1 Calcolo della cifra di rumore e della temperatura equivalente di rumore di una cascata di doppi bipoli

Iniziamo a calcolare la temperatura equivalente di rumore dei due stadi di figura 5.10, supponendo che tutto sia indipendente dalla frequenza. Bisogna calcolare le densità spettrali di potenza di rumore nelle varie sezioni della cascata di doppi bipoli. Usando la (5.8), e indicando con $T_{eq}^{(1)}$ la temperatura equivalente del primo doppio bipolo, si trova che la densità spettrale di potenza di rumore alla sua uscita vale:

$$G_{out}^{(1)}(f) = A_{d1} \frac{K}{2} (T_0 + T_{eq}^{(1)})$$

Pertanto, dal punto di vista del secondo doppio bipolo, è come se questo fosse alimentato da un resistore rumoroso con temperatura $A_{d1}(T_0 + T_{eq}^{(1)})$. Si trova quindi che la densità spettrale di rumore all'uscita del secondo stadio vale:

$$\begin{aligned} G_{out}^{(2)}(f) &= A_{d2} \frac{K}{2} [A_{d1}(T_0 + T_{eq}^{(1)}) + T_{eq}^{(2)}] \\ &= A_{d1} A_{d2} \frac{K}{2} \left(T_0 + T_{eq}^{(1)} + \frac{T_{eq}^{(2)}}{A_{d1}} \right) \end{aligned}$$

Da questa espressione, confrontandola con la definizione di temperatura equivalente di rumore (5.8), si trova che la cascata di due doppi bipoli ha:

$$\begin{cases} T_{eq}^{(TOT)} = T_{eq}^{(1)} + \frac{T_{eq}^{(2)}}{A_{d1}} \\ A_{dTOT} = A_{d1} A_{d2} \end{cases}$$

Procedendo con conti analoghi, per N doppi bipoli in cascata si trova che:

$$T_{eq}^{(TOT)} = T_{eq}^{(1)} + \frac{T_{eq}^{(2)}}{A_{d1}} + \frac{T_{eq}^{(3)}}{A_{d1} A_{d2}} + \dots + \frac{T_{eq}^{(N)}}{A_{d1} A_{d2} \dots A_{d(N-1)}} \quad (5.16)$$

$$A_{dTOT} = A_{d1} \cdot A_{d2} \dots A_{dN}$$

Alcuni commenti sulla (5.16) sono doverosi.

Temperatura
equivalente
di rumore di
N doppi
bipoli in
cascata

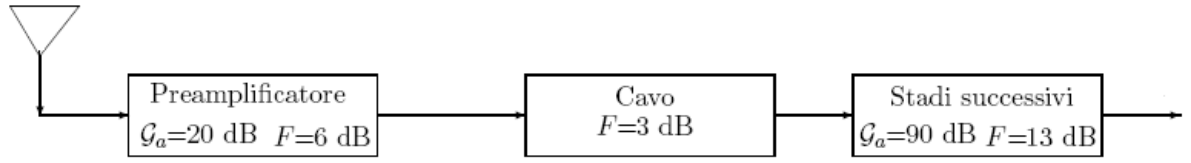


Figura 5.11: Schema a blocchi di un ricevitore televisivo.

1. La temperatura equivalente di rumore dello stadio i -esimo è divisa per i guadagni degli stadi precedenti.
2. In presenza di guadagni positivi (in dB) risultano dunque maggiormente critici per il rumore i primi stadi. Se si considera ad esempio un ricevitore per telecomunicazioni, ne segue che, per avere una temperatura equivalente di rumore bassa, i primi stadi devono essere progettati con estrema cura, in modo da avere bassa cifra di rumore, magari a scapito del guadagno. Gli ultimi stadi, invece, sono progettati per avere alte amplificazioni, anche se questo porta ad avere alte cifre di rumore.
3. In generale è sempre meglio prima amplificare e poi attenuare, non viceversa.

Con conti simili a quelli appena fatti, si può ottenere una formula analoga per la cifra di rumore:

Cifra di
rumore di N
doppi bipoli
in cascata

$$F_{TOT} = F_1 + \frac{F_2 - 1}{A_{d1}} + \frac{F_3 - 1}{A_{d1} A_{d2}} + \dots + \frac{F_N - 1}{A_{d1} A_{d2} \dots A_{d(N-1)}} \quad (5.17)$$

Attenzione alla presenza del -1 rispetto alla formula precedente.

Esempio 5.4. La figura 5.11 rappresenta lo schema a blocchi di un ricevitore televisivo. Sul tetto subito dopo l'antenna è posto un pre-amplificatore a basso rumore. Poi c'è il cavo coassiale che porta il segnale fino al televisore. Infine c'è il televisore. I dati numerici sono tutti mostrati in figura. Chiamiamo F_1 la cifra di rumore del pre-amplificatore, F_2 la cifra di rumore del cavo e F_3 la cifra di rumore degli stadi successivi.

Vogliamo determinare la cifra di rumore totale del sistema.

La cifra di rumore totale vale:

$$\begin{aligned} F_{TOT} &= F_1 + \frac{F_2 - 1}{G_a} + \frac{F_3 - 1}{\frac{G_a}{L}} \\ &= 4 + \frac{2 - 1}{100} + \frac{20 - 1}{100 \frac{1}{2}} \\ &= 4.39 \\ &= 6.4 \text{ dB} \end{aligned}$$

■

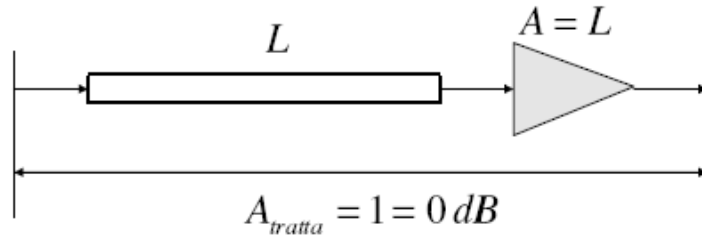


Figura 5.12: Tratta amplificata. L'amplificatore compensa esattamente l'attenuazione del cavo.

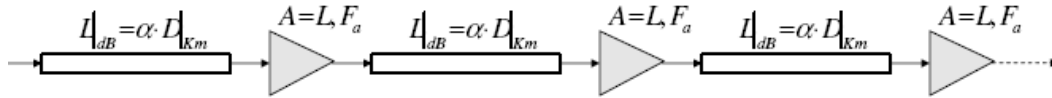


Figura 5.13: Sistema multitratta.

5.3.2 Sistemi multitratta

In un sistema di telecomunicazioni, per coprire lunghe distanze, spesso si spezza un collegamento in varie tratte. Si faccia riferimento alla figura 5.12. Una **tratta** è la cascata di un cavo seguito da un amplificatore che ne compensa esattamente le perdite. L'amplificatore ha guadagno A e cifra di rumore F_a . Il cavo ha attenuazione $L = A$ e cifra di rumore L . Il guadagno di potenza di una tratta è dunque unitario per definizione.

Tratte amplificate

Per una tratta posso applicare la (5.17) e calcolare la sua cifra di rumore:

Cifra di rumore di una tratta singola

$$\begin{aligned}
 F_{tratta} &= F_1 + \frac{F_2 - 1}{A_1} \\
 &= L + \frac{F_a - 1}{\frac{1}{L}} \\
 &= F_a L
 \end{aligned} \tag{5.18}$$

che in dB diventa, ricordando che $L|_{dB} = \alpha|_{km} \cdot D|_{dB/km}$:

$$F_{tratta}|_{dB} = F_a|_{dB} + \alpha|_{km} \cdot D|_{dB/km}$$

Quindi la cifra di rumore di tratta, in dB, è linearmente proporzionale alla lunghezza del cavo. **La cifra di rumore in unità lineari cresce dunque esponenzialmente con la lunghezza del cavo.**

Solitamente nella pratica i collegamenti a lunga distanza sono divisi in sotto-tratte di lunghezza più piccola (vedi figura 5.13). Vediamo cosa succede in questo caso. Consideriamo la cascata di N tratte identiche. Consideriamo ogni tratta come un doppio bipolo con guadagno unitario e cifra di rumore $F_a L$ (come calcolato nella (5.18)). Con queste ipotesi, possiamo applicare la (5.17) per calcolare la cifra di rumore

Analisi sistemi multitratta

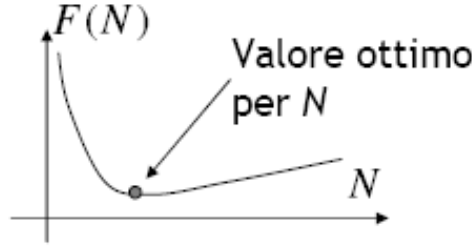


Figura 5.14: Andamento della cifra di rumore di un sistema multitratta in funzione del numero di tratte. Come si vede dalla figura, la funzione ha un minimo assoluto.

complessiva della cascata di N tratte:

$$\begin{aligned} F(N) &= F_a L + \underbrace{(F_a L - 1) + \dots + (F_a L - 1)}_{(N-1) \text{ termini}} \\ &= F_a L + (N - 1)(F_a L - 1) \end{aligned}$$

Cifra di
rumore
cascata di N
tratte

cioè:

$$\boxed{F(N) = N L F_a - (N - 1)} \quad (5.19)$$

Ottimizzazione sistemi multitratta

A questo punto sorge spontanea una domanda. Dato un collegamento, ci chiediamo se sia possibile calcolare il numero di tratte che minimizza la cifra di rumore totale, cioè che minimizza l'espressione (5.19). La figura 5.14 mostra l'andamento di $F(N)$ in funzione di N . Come si vede, si tratta di una funzione concava con un minimo assoluto. È possibile quindi calcolare il numero ottimo di tratte da utilizzare, cioè il numero di tratte che minimizza la cifra di rumore complessiva. Le ipotesi sono le seguenti.

- Distanza complessiva da coprire: D_{TOT} .
- Attenuazione complessiva in dB: $L_{TOT} = \alpha \cdot D_{TOT} [dB]$.
- Lunghezza di ogni tratta $D = D_{TOT}/N$.
- Attenuazione di ogni tratto di cavo (unità lineari): $L = L_{TOT}^{\frac{1}{N}}$.

Sostituendo questi valori nella (5.19) trovo che:

$$F(N) = N L_{TOT}^{\frac{1}{N}} F_a - (N - 1) \quad (5.20)$$

Per trovare il minimo devo calcolare la derivata di $F(N)$ e porla uguale a zero. Trascurando il termine $(N - 1)$ e derivando la (5.20) rispetto a N , si trova:

$$\begin{aligned} \frac{dF(N)}{dN} &\approx F_a L_{TOT}^{\frac{1}{N}} + F_a N \left[L_{TOT}^{\frac{1}{N}} \ln(L_{TOT}) (-N^{-2}) \right] \\ &= F_a L_{TOT}^{\frac{1}{N}} \left[1 - \frac{\ln(L_{TOT})}{N} \right] \end{aligned}$$

che, uguagliata a zero, fornisce il valore ottimo di N :

$$\boxed{N_{ottimo} \approx \ln(L_{TOT}) \approx 0.23 \cdot L_{TOT}|_{dB}} \quad (5.21)$$

Numero
ottimo di
tratte

Il risultato va sempre arrotondato all'intero più vicino.

Sostituendo la (5.21) nella (5.20) si trova la cifra di rumore complessiva. Dopo qualche passaggio algebrico il risultato è:

Cifra di
rumore
sistema
multitratta
ottimizzato

$$\begin{aligned} F_{ottimo} &= F_a \cdot \ln(L_{TOT}) \cdot e \quad (\text{in lineare}) \\ F_{ottimo}|_{dB} &= F_a|_{dB} + 10 \log_{10}(L_{TOT}|_{dB}) - 2.0378 \text{ dB} \quad (\text{in unità logaritmiche}) \end{aligned}$$

Confronto multitratta con singola tratta

Confrontiamo i sistemi singola tratta e multitratta (ottimizzato) sulla **stessa distanza complessiva**. Dato che l'attenuazione totale è la stessa, si ha:

$$\begin{aligned} \text{Singola tratta} \quad F_{singola} &= F_a \cdot L_{TOT} \\ \text{Multitratta} \quad F_{multi} &= F_a \cdot \ln(L_{TOT}) \cdot e \end{aligned}$$

- Nel sistema singola tratta la cifra di rumore è proporzionale alla attenuazione totale, e dunque esponenziale con la distanza.
- Nel sistema multitratta la cifra di rumore è logaritmica con l'attenuazione, e dunque lineare con la distanza.
- Si noti come la soluzione multitratta abbia prestazioni decisamente migliori a fronte di costi molto maggiori.

5.4 Conclusione

In questo capitolo abbiamo trattato i concetti fondamentali sul rumore termico. Innanzitutto ne abbiamo dato un modello probabilistico come processo casuale, poi abbiamo visto come caratterizzare dal punto di vista del rumore i dispositivi che costituiscono un sistema di telecomunicazioni, cioè i doppi bipoli.

Il rumore termico è così importante perché è una fonte di disturbo sempre presente nei sistemi di telecomunicazioni. Vedremo nei prossimi capitoli come una misura della qualità di un sistema di comunicazione sia il rapporto segnale-rumore cioè il rapporto tra il segnale utile e il rumore in uscita dal ricevitore.

Aggiungiamo inoltre che i risultati ottenuti in questo capitolo valgono anche per sistemi di comunicazione digitali. In tali sistemi (come si vedrà nel corso di Trasmissione Numerica) la qualità di un collegamento è data dalla probabilità di errore sul bit. Questa quantità dipende in qualche modo dal rapporto segnale-rumore in uscita dal ricevitore.

5.5 Riassunto formule principali sul rumore termico

Si riportano in questa appendice le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi, e sono tutte presenti anche sul formulario del corso (che si può scaricare dal sito del corso). Si ricorda

che il formulario è l'unico materiale che è consentito tenere in sede d'esame.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- Rumore termico: densità spettrale di potenza di rumore ai capi di una resistenza:

$$G_d(f) = \frac{KT}{2} \left[\frac{W}{Hz} \right]; \quad K = 1.38 \cdot 10^{-23} \left[\frac{J}{^\circ K} \right]$$

- Rumore termico: densità spettrale di potenza di rumore in uscita da un doppio bipolo:

$$G_{out}^{reale}(f) = A_d(f) \frac{K}{2} (T_0 + T_{eq}(f)) = A_d(f) \frac{K}{2} T_0 F(f)$$

- Formule di conversione tra cifra di rumore e temperatura equivalente di un doppio bipolo:

$$\begin{aligned} F &= 1 + \frac{T_{eq}}{T_0} \\ T_{eq} &= T_0 (F - 1) \end{aligned}$$

- Cascate di doppi bipoli:

$$\begin{aligned} T_{eq}^{(TOT)} &= T_{eq}^{(1)} + \frac{T_{eq}^{(2)}}{A_{d1}} + \frac{T_{eq}^{(3)}}{A_{d1} A_{d2}} + \cdots + \frac{T_{eq}^{(N)}}{A_{d1} A_{d2} \cdots A_{d(N-1)}} \\ F_{TOT} &= F_1 + \frac{F_2 - 1}{A_{d1}} + \frac{F_3 - 1}{A_{d1} A_{d2}} + \cdots + \frac{F_N - 1}{A_{d1} A_{d2} \cdots A_{d(N-1)}} \end{aligned}$$

- Tratte periodiche amplificatore + attenuatore:

$$\begin{aligned} F(N) &= NLF_a - (N - 1) \\ N_{ottimo} &\approx \ln(L_{TOT}) \approx 0.23 \cdot L_{TOT}|_{dB} \end{aligned}$$

Capitolo 6

Segnale analitico e inviluppo complesso

In questo capitolo verranno introdotti i concetti di segnale analitico e inviluppo complesso. Si tratta di due tecniche matematiche molto diffuse nello studio dei sistemi di telecomunicazioni per rappresentare segnali e sistemi di tipo passa-banda. I risultati ottenuti in questo capitolo saranno utilizzati a fondo nel resto del corso e nei prossimi corsi trasmissivi.

6.1 Classificazione spettrale dei segnali

In base al loro spettro i segnali si possono classificare in:

Segnali in banda base: segnali con componenti spettrali concentrate attorno a $f = 0$, e nulle al di fuori di un certo range (vedi figura 6.1 in alto). Esempi:

- Uscita audio di un microfono;
- Segnali logici (ad esempio TTL¹).

Segnali in banda traslata: segnali con componenti spettrali concentrate attorno alla frequenza "centrale" $f_c \neq 0$ (vedi figura 6.1 in basso). Esempi:

- Segnale televisivo broadcast;
- Segnali radio FM, AM;
- Segnale per comunicazioni via satellite.

I segnali in banda traslata sono i tipici segnali che vengono generati all'uscita di un sistema che si basa su modulazione AM, PM o FM. Queste tre modulazioni verranno ampiamente trattate nel resto del corso.

I segnali in banda traslata sono così importanti che è utile introdurre una rappresentazione matematica alternativa, detta del segnale analitico, che è lo scopo di questo capitolo.

¹TTL sta per Transistor-Transistor Logic. Si tratta di una classe di circuiti digitali integrati costruiti con l'utilizzo di transistor a giunzione bipolare (BJT) e resistori. È chiamata "logica transistor-transistor" perché (a differenza di altre tecnologie), sono svolte dal transistor sia la funzione di porta logica che quella di amplificazione.

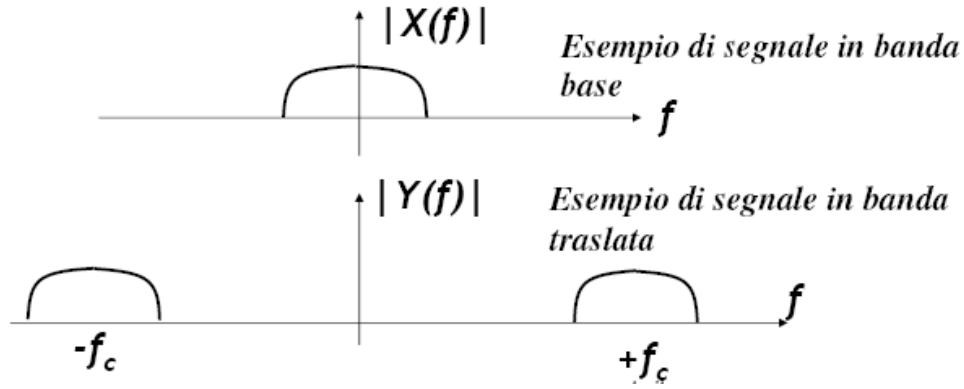


Figura 6.1: Esempio di segnale in banda base (in alto) e di segnale in banda traslata (in basso).

6.1.1 Accenno alla modulazione

Sia $x(t)$ un segnale in banda base, che chiameremo *segnale modulante*. Sia poi:

$$y(t) = x(t) \cos(2\pi f_c t)$$

Il segnale $y(t)$ viene detto *segnale modulato*². $y(t)$, come visto nel corso di Teoria dei Segnali, è un segnale in banda traslata. La figura 6.1 mostra un esempio qualitativo degli spettri di $x(t)$ e $y(t)$. Se B è la banda di $x(t)$, solitamente nelle applicazioni si usa $f_c \gg B$.

Gran parte delle trasmissioni (ad esempio tutte quelle via etere) sono modulate, e dunque in banda traslata. Per questo, è utile introdurre un formalismo ad hoc per questo tipo di segnali. Parleremo nel dettaglio dei vari schemi di modulazione nei prossimi capitoli.

6.2 Segnale analitico

Alla base della definizione del segnale analitico c'è la seguente idea di base. È noto che un segnale reale $v(t)$ ha una trasformata di Fourier $V(f)$ con modulo pari e fase dispari. Per questo motivo, la parte di $V(f)$ per $f > 0$ contiene tutta l'informazione necessaria per ricostruire $v(t)$.

Definizione
di segnale
analitico

Dato un segnale $v(t)$ (solitamente in banda traslata, ma non necessariamente), si definisce un segnale associato a $v(t)$, detto segnale analitico, e indicato come:

$$\overset{\circ}{v}(t) \quad \text{segnale analitico}$$

nel modo seguente. Si considera il filtro:

$$H(f) \triangleq 2u(f) \tag{6.1}$$

²Questo è un caso particolare di *modulazione di ampiezza*, come vedremo in uno dei prossimi capitoli.

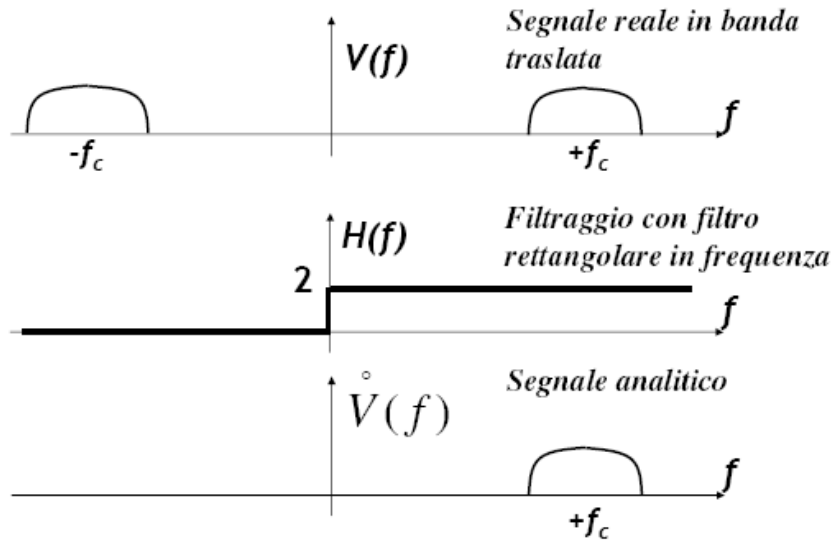


Figura 6.2: Rappresentazione grafica costruzione segnale analitico. Filtro il segnale in modo da eliminare tutte le componenti a frequenza negativa.

dove $u(f)$ è la funzione gradino (o funzione di Heaviside). Il segnale analitico e il suo spettro sono definiti come:

$$\begin{aligned}
 \overset{\circ}{V}(f) &\triangleq H(f) \cdot V(f) \\
 \overset{\circ}{v}(t) &\triangleq F^{-1} \left\{ \overset{\circ}{V}(f) \right\} \\
 &= F^{-1} \{ H(f) \cdot V(f) \} \\
 &= h(t) * v(t)
 \end{aligned} \tag{6.2}$$

La figura 6.2 mostra come si costruisce il segnale analitico graficamente. Si usa $u(f)$ moltiplicata per 2 così $V(f)$ e $\overset{\circ}{V}(f)$ hanno la stessa energia.

In pratica, **il segnale analitico è un segnale che contiene solo le frequenze positive del segnale di partenza**. Si noti che, avendo una trasformata di Fourier asimmetrica in f , è sicuramente un segnale complesso nel dominio del tempo.

Legame tra segnale analitico e segnale reale

Vediamo che legame c'è tra il segnale analitico $\overset{\circ}{v}(t)$ e il segnale reale $v(t)$. Dalla (6.1) ho che (usando le tavole delle trasformate):

$$\begin{aligned}
 h(t) &= F^{-1} \{ 2u(f) \} \\
 &= \delta(t) + j \frac{1}{\pi t}
 \end{aligned}$$

Andando a sostituire questa espressione nella definizione di segnale analitico (6.2) trovo:

$$\begin{aligned}
 \overset{\circ}{v}(t) &= h(t) * v(t) \\
 &= \left[\delta(t) + j \frac{1}{\pi t} \right] * v(t) \\
 &= v(t) + j \left[v(t) * \frac{1}{\pi t} \right] \\
 &= v(t) + j \hat{v}(t)
 \end{aligned} \tag{6.3}$$

avendo definito:

$$\hat{v}(t) \triangleq v(t) * \frac{1}{\pi t} \tag{6.4}$$

Si osservi che $\hat{v}(t)$ è sicuramente un segnale reale. Abbiamo quindi trovato che la rappresentazione in termini di segnale analitico permette di ritornare al segnale di partenza. Dalla (6.3) ho infatti che:

$$v(t) = \text{Re} \left[\overset{\circ}{v}(t) \right] \tag{6.5}$$

Legame tra
segnale
analitico e
segnale reale

In alternativa posso usare la costruzione grafica nel dominio delle frequenze.

Trasformata di Hilbert

Nella (6.4) si è definito il segnale:

$$\hat{v}(t) \triangleq v(t) * \frac{1}{\pi t}$$

Questo segnale è detto *trasformata di Hilbert* del segnale $v(t)$. Si tratta di un segnale (reale se $v(t)$ è reale) che si ottiene *in teoria* dal filtraggio di $v(t)$ con un filtro che abbia risposta all'impulso:

$$h_{\text{Hilbert}}(t) = \frac{1}{\pi t}$$

Si noti che la risposta all'impulso non è causale, quindi in teoria non è fisicamente realizzabile. Se la si vuole implementare in pratica bisognerà approssimarla opportunamente.

Esempio 6.1. Segnale analitico di un segnale sinusoidale Consideriamo l'importante caso particolare di un segnale sinusoidale:

$$x(t) = A \cos(2\pi f_c t + \varphi)$$

e calcoliamone il corrispondente segnale analitico. La sua trasformata di Fourier vale:

$$X(f) = \frac{A}{2} [\delta(f - f_c) e^{j\varphi} + \delta(f + f_c) e^{-j\varphi}]$$

Si passa al segnale analitico moltiplicando per 2 e tenendo solo la parte positiva delle frequenze di $X(f)$:

$$\overset{\circ}{X}(f) = A \delta(f - f_c) e^{j\varphi}$$

e antitrasformando:

Segnale
analitico di
un segnale
sinusoidale

$$\begin{aligned}\overset{\circ}{x}(t) &= A e^{j2\pi f_c t} e^{j\varphi} \\ &= A e^{j(2\pi f_c t + \varphi)}\end{aligned}$$

Abbiamo dunque ottenuto:

$$x(t) = A \cos(2\pi f_c t + \varphi) \quad \Rightarrow \quad \overset{\circ}{x}(t) = A e^{j(2\pi f_c t + \varphi)}$$

Importante: per un segnale sinusoidale il segnale analitico coincide con il concetto di **fasore** tipico di altri corsi (ad esempio Elettrotecnica). Si noti che a volte il fasore è definito come $A e^{j(2\pi f_c t + \varphi)}$, altre volte come $A e^{j\varphi}$. ■

Conclusione sul segnale analitico

In conclusione, il segnale analitico è un modo per "estendere" il concetto di fasore (di per sé valido solo per segnali sinusoidali) a generici segnali in banda traslata. Ad esempio, sarà utile per tutti i segnali del tipo:

$$x(t) = A_c a(t) \cos[2\pi f_c t + \varphi(t)]$$

di cui parleremo a fondo nei prossimi capitoli.

6.3 Inviluppo complesso

Un generico segnale $v(t)$ in banda traslata attorno alla frequenza f_c può essere ulteriormente rappresentato tramite il suo *inviluppo complesso*, definito come:

Definizione
di inviluppo
complesso

$$\tilde{v}(t) \triangleq \overset{\circ}{v}(t) e^{-j2\pi f_c t} \quad (6.6)$$

Anche questo segnale è, in generale, complesso. L'operazione svolta nella (6.6) serve a prendere lo spettro del segnale analitico e a riportarlo in banda base. Graficamente l'inviluppo complesso si costruisce come in figura 6.3.

Dalla (6.5) e dalla (6.6) si ottiene dunque che, dato l'inviluppo complesso e conoscendo la frequenza centrale f_c , si può ritornare al segnale reale tramite la formula:

Legame tra
inviluppo
complesso e
segnale reale

$$v(t) = \text{Re} \left[\tilde{v}(t) e^{j2\pi f_c t} \right] \quad (6.7)$$

Si noti che dunque, in generale, **lo spettro dell'inviluppo complesso si trova in banda base. È questo il motivo principale del suo utilizzo.** Cioè l'inviluppo complesso permette di rappresentare un segnale in banda traslata tramite un segnale corrispondente in banda base (anche se complesso).

L'inviluppo complesso del segnale $v(t)$ può essere anche scritto mettendo in evidenza la sua parte reale e la sua parte immaginaria:

$$\tilde{v}(t) = x(t) + j y(t) \quad (6.8)$$

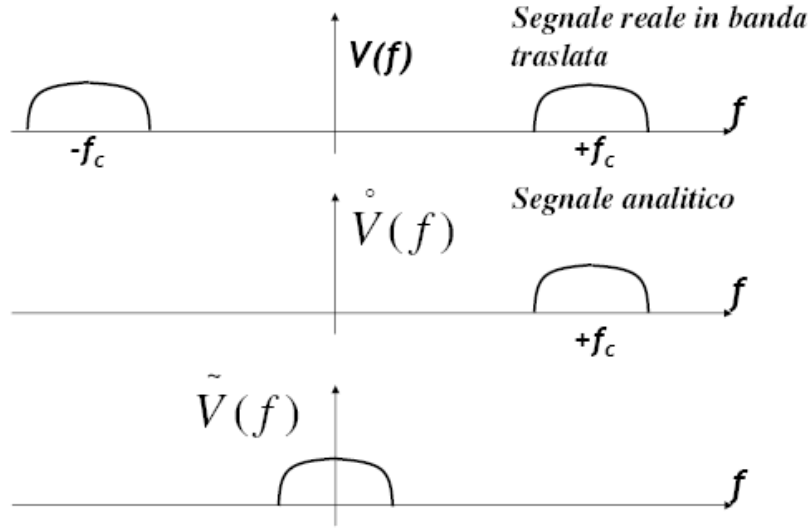


Figura 6.3: Costruzione grafica inviluppo complesso. Si calcola il segnale analitico e poi lo si riporta in banda base. L'inviluppo complesso è sempre un segnale in banda base.

dove $x(t)$ viene detta *componente in fase* e $y(t)$ viene detta *componente in quadratura*. $x(t)$ e $y(t)$ sono segnali reali in banda base. Sostituendo l'espressione (6.8) nella (6.7) trovo:

$$\begin{aligned} v(t) &= \operatorname{Re} \{ [x(t) + j y(t)] [\cos(2\pi f_c t) + j \sin(2\pi f_c t)] \} \\ &= x(t) \cos(2\pi f_c t) - y(t) \sin(2\pi f_c t) \end{aligned} \quad (6.9)$$

Riassumendo, si ha che, data una funzione $v(t)$ reale in banda traslata:

- $\tilde{v}^o(t)$ è una funzione complessa in banda traslata;
- $\tilde{v}(t)$ è una funzione complessa in banda base.

6.3.1 Rappresentazione geometrica di un segnale in banda traslata

L'inviluppo complesso del segnale $v(t)$ può essere anche scritto mettendo in evidenza il suo modulo e la sua fase:

$$\tilde{v}(t) = r(t) e^{j\theta(t)} \quad (6.10)$$

dove $r(t)$ e $\theta(t)$ sono segnali reali e in banda base. Sostituendo la (6.10) nella (6.7) trovo:

$$\begin{aligned} v(t) &= \operatorname{Re} \{ [r(t) e^{j\theta(t)}] e^{j2\pi f_c t} \} \\ &= r(t) \cos[2\pi f_c t + \theta(t)] \end{aligned}$$

Questa espressione mi dice che un segnale in banda traslata può essere rappresentato come un segnale sinusoidale modulato in ampiezza e in fase. Si tornerà su questo punto nei prossimi capitoli.

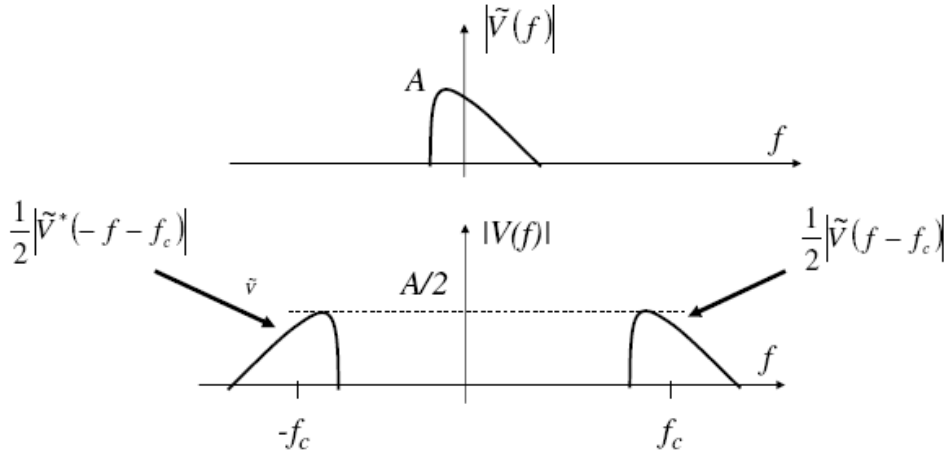


Figura 6.4: Esempio di spettro di un segnale in banda traslata (sotto) e spettro del suo involucro complesso (sopra). La figura vuole evidenziare il legame tra i due spettri.

Confrontando la (6.10) con la (6.8) abbiamo infine che:

$$\begin{aligned} r(t) &= \sqrt{x^2(t) + y^2(t)} \\ \theta(t) &= \text{Fase}[\tilde{v}(t)] \end{aligned} \quad (6.11)$$

6.4 Spettri

Si consideri un segnale determinato $v(t)$ in banda traslata, reale e a energia finita:

Segnali
determinati

$$v(t) = \text{Re} \left[\tilde{v}(t) e^{j2\pi f_c t} \right]$$

Tra poco dimostreremo che la sua trasformata di Fourier vale:

$$\begin{aligned} V(f) &= F \{v(t)\} \\ &= F \left\{ \text{Re} \left[\tilde{v}(t) e^{j2\pi f_c t} \right] \right\} \\ &= \frac{1}{2} \left[\tilde{V}(f - f_c) + \tilde{V}^*(-f - f_c) \right] \end{aligned} \quad (6.12)$$

Questa espressione mette in relazione lo spettro di un segnale con lo spettro del suo involucro complesso. La figura 6.4 mostra un esempio di questi due spettri.

Se invece $v(t)$ è un processo casuale WSS, si ha che la sua densità spettrale di potenza vale:

Processi
casuali

$$G_v(f) = \frac{1}{4} [G_{\tilde{v}}(f - f_c) + G_{\tilde{v}}(-f - f_c)] \quad (6.13)$$

La figura 6.5 mostra un esempio di questi due spettri.

Dimostrazione della (6.12) relativa alle trasformate. La (6.7) mi dice che:

$$v(t) = \text{Re} \left[\tilde{v}(t) e^{j2\pi f_c t} \right] \quad (6.14)$$

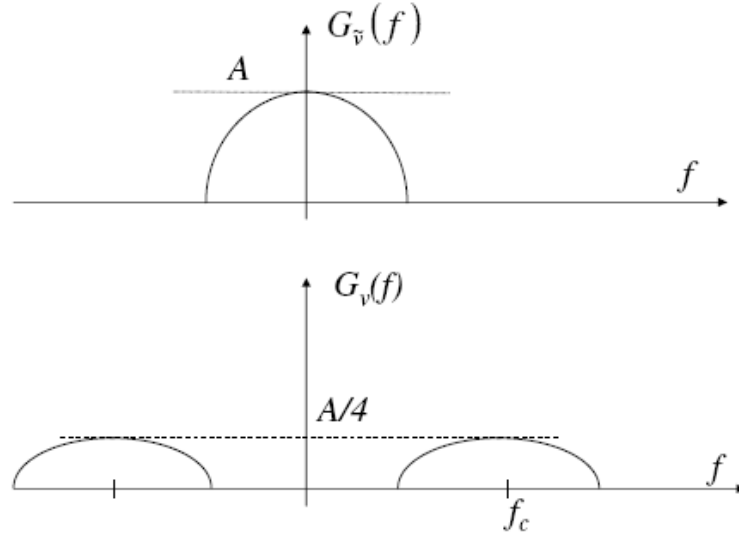


Figura 6.5: Esempio di densità spettrale di potenza di un segnale in banda traslata (sotto) e densità spettrale di potenza del suo inviluppo complesso (sopra). La figura vuole evidenziare il legame tra i due spettri.

ma, dato un generico numero o funzione complessa z , si ricorda che:

$$\operatorname{Re}(z) = \frac{1}{2}(z + z^*)$$

Quindi la (6.14) diventa:

$$v(t) = \frac{1}{2} \tilde{v}(t) e^{j2\pi f_c t} + \frac{1}{2} \tilde{v}^*(t) e^{-j2\pi f_c t}$$

Effettuandone la trasformata di Fourier trovo:

$$V(f) = F \left\{ \frac{1}{2} \tilde{v}(t) e^{j2\pi f_c t} \right\} + F \left\{ \frac{1}{2} \tilde{v}^*(t) e^{-j2\pi f_c t} \right\} \quad (6.15)$$

Ma:

$$\begin{aligned} F \{ \tilde{v}(t) \} = \tilde{V}(f) &\Rightarrow F \left\{ \frac{1}{2} \tilde{v}(t) e^{j2\pi f_c t} \right\} = \frac{1}{2} \tilde{V}(f - f_c) \\ F \{ \tilde{v}^*(t) \} = \tilde{V}^*(-f) &\Rightarrow F \left\{ \frac{1}{2} \tilde{v}^*(t) e^{-j2\pi f_c t} \right\} = \frac{1}{2} \tilde{V}^*(-f - f_c) \end{aligned}$$

Sostituendo queste due espressioni nella (6.15), trovo:

$$V(f) = \frac{1}{2} \left[\tilde{V}(f - f_c) + \tilde{V}^*(-f - f_c) \right]$$

Abbiamo così dimostrato la (6.12). Notare che, essendo $\tilde{V}(f)$ in banda base, $\tilde{V}^*(-f - f_c)$ è centrato in $f = -f_c$. ■

Dimostrazione della (6.13) relativa alle densità spettrali. È necessario passare dal calcolo dell'autocorrelazione:

$$R_v(\tau) = \langle v(t) v(t + \tau) \rangle$$

da intendere come **media temporale rispetto al tempo t** (o indifferentemente in termini di media di insieme se il segnale è in effetti un processo casuale ergodico). Attenzione al fatto che $\tilde{v}(t)$ è in generale una funzione complessa. Nel nostro caso, ricordando la (6.7) ho che:

$$\begin{aligned} R_v(\tau) &= \langle v(t) v(t + \tau) \rangle \\ &= \langle \operatorname{Re} [\tilde{v}(t) e^{j2\pi f_c t}] \cdot \operatorname{Re} [\tilde{v}(t + \tau) e^{j2\pi f_c (t + \tau)}] \rangle \end{aligned} \quad (6.16)$$

Ricordando che per due qualsiasi numeri o funzioni complesse z_1 e z_2 vale:

$$\operatorname{Re} [z_1] \cdot \operatorname{Re} [z_2] = \frac{1}{2} (\operatorname{Re} [z_1^* \cdot z_2] + \operatorname{Re} [z_1 \cdot z_2])$$

la (6.16) diventa:

$$R_v(\tau) = \frac{1}{2} \langle \operatorname{Re} \left\{ \tilde{v}^*(t) e^{-j2\pi f_c t} \tilde{v}(t + \tau) e^{j2\pi f_c (t + \tau)} \right\} \rangle + \frac{1}{2} \langle \operatorname{Re} \left\{ \tilde{v}(t) e^{j2\pi f_c t} \tilde{v}(t + \tau) e^{j2\pi f_c (t + \tau)} \right\} \rangle$$

Essendo gli operatori $\langle \cdot \rangle$ e $\operatorname{Re}\{\cdot\}$ lineari, possono essere scambiati:

$$R_v(\tau) = \frac{1}{2} \operatorname{Re} \left\{ \langle \tilde{v}^*(t) \tilde{v}(t + \tau) e^{j2\pi f_c \tau} \rangle \right\} + \frac{1}{2} \operatorname{Re} \left\{ \langle \tilde{v}(t) \tilde{v}(t + \tau) e^{j4\pi f_c t + j2\pi f_c \tau} \rangle \right\}$$

Osserviamo che le medie sono rispetto a t , e dunque:

$$\begin{aligned} R_v(\tau) &= \frac{1}{2} \operatorname{Re} \left\{ \underbrace{\langle \tilde{v}^*(t) \tilde{v}(t + \tau) \rangle}_{R_{\tilde{v}}(\tau)} e^{j2\pi f_c \tau} \right\} + \frac{1}{2} \operatorname{Re} \left\{ \underbrace{\langle \tilde{v}(t) \tilde{v}(t + \tau) e^{j4\pi f_c t} \rangle}_{=0 (*)} e^{j2\pi f_c \tau} \right\} \\ &= \frac{1}{2} \operatorname{Re} \{ R_{\tilde{v}}(\tau) e^{j2\pi f_c \tau} \} \end{aligned} \quad (6.17)$$

Il termine media temporale $(*)$ è approssimativamente nullo se la frequenza $2f_c$ è molto più alta dello spettro dell'involuppo complesso. Infatti in tal caso l'integrale va fatto su periodi di sinusoidi ad ampiezza "costante" e dunque con integrali nulli.

Trasformando secondo Fourier la (6.17) e applicando la proprietà (6.12), troviamo:

$$\begin{aligned} G_v(f) &= F\{R_v(\tau)\} \\ &= \frac{1}{2} F\left\{ \operatorname{Re}[R_{\tilde{v}}(\tau) e^{j2\pi f_c \tau}] \right\} \\ &= \frac{1}{4} [G_{\tilde{v}}(f - f_c) + G_{\tilde{v}}(-f - f_c)] \end{aligned} \quad (6.18)$$

Abbiamo quindi dimostrato anche la (6.13). ■

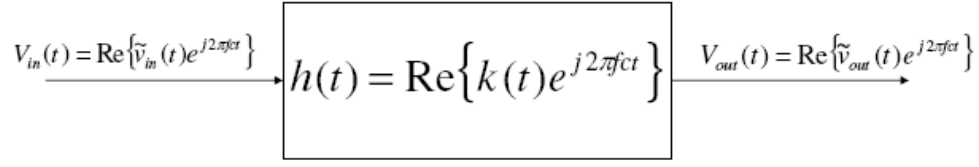


Figura 6.6: Rappresentazione di un sistema passa-banda.

Calcolo della potenza media del segnale reale in funzione della potenza media dell'inviluppo complesso

Per esercizio, calcoliamo la potenza media del segnale reale in funzione della potenza media dell'inviluppo complesso. Usando la (6.18) trovo che:

$$\begin{aligned}
 P_v &= \langle v^2(t) \rangle \\
 &= \int_{-\infty}^{+\infty} G_v(f) df \\
 &= \frac{1}{4} \left[\int_{-\infty}^{+\infty} G_{\tilde{v}}(f - f_c) df + \int_{-\infty}^{+\infty} G_{\tilde{v}}(-f - f_c) df \right]
 \end{aligned}$$

ma:

$$\begin{aligned}
 \int_{-\infty}^{+\infty} G_{\tilde{v}}(f - f_c) df &= P_{\tilde{v}} \\
 \int_{-\infty}^{+\infty} G_{\tilde{v}}(-f - f_c) df &= P_{\tilde{v}}
 \end{aligned}$$

Quindi:

$$\begin{aligned}
 P_v &= \frac{1}{4} (P_{\tilde{v}} + P_{\tilde{v}}) \\
 &= \frac{1}{2} P_{\tilde{v}}
 \end{aligned}$$

Il segnale $v(t)$ ha quindi una potenza pari a metà del corrispondente inviluppo complesso $\tilde{v}(t)$.

6.5 Equivalente passa-basso di un sistema passa-banda

Consideriamo un segnale $v_{in}(t)$ in banda traslata, centrato attorno alla frequenza f_c . Supponiamo di far passare questo segnale in un sistema LTI passa-banda e di ottenere in uscita il segnale $v_{out}(t)$. La situazione è rappresentata schematicamente in figura 6.6. Chiamiamo $k(t)$ l'inviluppo complesso di $h(t)$. Vogliamo ora vedere che legame c'è tra $\tilde{v}_{in}(t)$, $\tilde{v}_{out}(t)$ e $k(t)$. Il segnale analitico di $v_{out}(t)$ vale:

$$\begin{aligned}
 \overset{\circ}{V}_{out}(f) &= 2u(f) V_{out}(f) \\
 &= 2u(f) V_{in}(f) V_{out}(f) \\
 &= \frac{1}{2} \overset{\circ}{V}_{in}(f) \overset{\circ}{V}_{out}(f)
 \end{aligned}$$

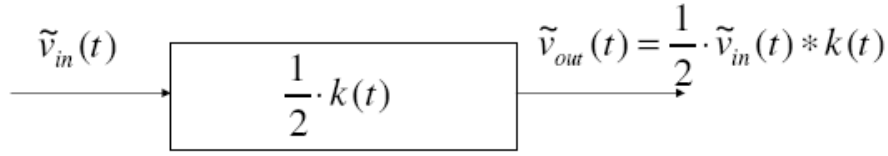


Figura 6.7: Equivalente passa-basso del sistema passa-banda di figura 6.6.

A questo punto posso traslare tutto a sinistra di f_c , ponendo $f = f + f_c$. Ottengo così:

$$\underbrace{\overset{\circ}{V}_{out}(f + f_c)}_{\tilde{V}_{out}(f)} = \frac{1}{2} \underbrace{\overset{\circ}{V}_{in}(f + f_c)}_{\tilde{V}_{in}(f)} \underbrace{\overset{\circ}{H}(f + f_c)}_{K(f)}$$

Ho cioè ottenuto che, in termini di trasformate di Fourier:

$$\tilde{V}_{out}(f) = \frac{1}{2} \tilde{V}_{in}(f) K(f)$$

che nel dominio del tempo diventa:

$$\tilde{v}_{out}(t) = \tilde{v}_{in}(t) * \left[\frac{1}{2} \tilde{h}(t) \right]$$

Questo risultato è fondamentale in quanto permette di descrivere un sistema passa-banda con un sistema equivalente passa-basso (vedi figura 6.7). Si noti che nel sistema passa-basso equivalente i segnali di ingresso-uscita e la risposta all'impulso del filtro sono in generale complessi. Un'importante applicazione di questo risultato è la simulazione di sistemi di trasmissione in banda passante, tramite un sistema corrispondente in banda base.

Esempio 6.2. Simulazione di un sistema di trasmissione Si consideri un segnale modulato attorno ad una frequenza centrale di 1.8 GHz, e con una banda di 2 MHz. La sua massima frequenza è dunque pari a 1.801 GHz. Se si dovesse simulare numericamente, sarebbe necessario campionare a:

$$\begin{aligned} f_{sample} &= 2 f_{max} \\ &= 2 \cdot 1.801 \text{ [GHz]} \\ &\approx 3.6 \left[\frac{\text{Gsample}}{\text{s}} \right] \end{aligned}$$

Il corrispondente involuppo complesso avrebbe invece una banda di 1 MHz, e sarebbe dunque campionabile a:

$$f_{sample} = 2 \text{ [M samples/s]}$$

Si noti l'enorme vantaggio del secondo approccio. AGGIUNGERE QUALCOSA.. VEDERE SE IN ESEMPI GAUDINO CE N'è UNO CHE VA BENE. ■

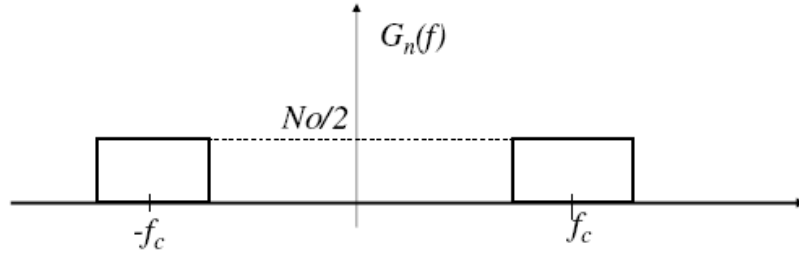


Figura 6.8: Spettro di rumore gaussiano bianco filtrato con filtro passa-banda.

6.6 Rumore in banda traslata

Anche per il rumore gaussiano bianco è utile avere una rappresentazione in banda traslata. Ci serve soprattutto la rappresentazione di rumore gaussiano bianco filtrato con filtro passa-banda (vedi figura 6.8). Un processo di questo tipo può essere sempre scritto tramite l'inviluppo complesso come:

$$n(t) = \text{Re} \left[\tilde{n}(t) e^{j2\pi f_c t} \right]$$

dove:

$$\tilde{n}(t) = n_1(t) + j n_2(t)$$

ottenendo:

$$n(t) = \text{Re} \left\{ [n_1(t) + j n_2(t)] e^{j2\pi f_c t} \right\}$$

I processi casuali $n_1(t)$ e $n_2(t)$ sono le parti in fase e quadratura del processo $n(t)$ e sono particolarmente importanti. $n(t)$, ricordando la (6.9), può essere anche scritto come:

$$n(t) = n_1(t) \cos(2\pi f_c t) - n_2(t) \sin(2\pi f_c t)$$

Risultato importante: si può dimostrare (non lo facciamo) che $n_1(t)$ e $n_2(t)$ hanno le stesse proprietà, cioè:

$$\begin{aligned} R_{n_1}(\tau) &= R_{n_2}(\tau) \\ G_{n_1}(f) &= G_{n_2}(f) \end{aligned}$$

Si ha inoltre che $n_1(t)$ e $n_2(t)$ sono processi reali e tra di loro statisticamente indipendenti. La funzione di autocorrelazione di $n_1(\tau)$ è uguale alla funzione di autocorrelazione di $n_2(\tau)$:

$$R_{n_1}(\tau) = R_{n_2}(\tau) = N_0 \delta(\tau)$$

e dunque:

$$G_{n_1}(f) = G_{n_2}(f) = N_0$$

vedi figura 6.9. Attenzione ai fattori 2.

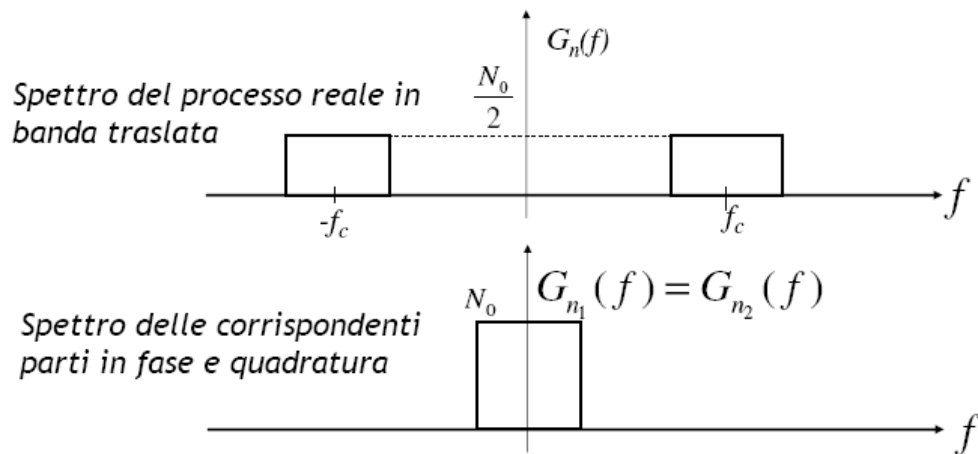


Figura 6.9: Spettro del rumore gaussiano bianco filtrato con filtro passa-banda (sopra) e spettro delle sue componenti in fase e quadratura (sotto).

6.7 Conclusione

Questo capitolo ha trattato il segnale analitico e l'involuppo complesso. Si tratta di due tecniche ampiamente utilizzate nei sistemi di telecomunicazioni per descrivere segnali e sistemi di tipo passa-banda. Verranno utilizzati molto in questo corso e negli altri corsi a carattere trasmissivo.

Capitolo 7

Modulazioni di ampiezza: AM

La modulazione di ampiezza (*AM*, *Amplitude Modulation*) la più semplice forma di modulazione. È stata la prima ad essere utilizzata (trasmissioni radio a modulazione d'ampiezza). Segue una breve cronologia.

- 1895: primi esperimenti radio di Marconi.
- 1901: primo messaggio telegrafico transatlantico (Marconi).
- 1906: prime vere trasmissioni AM in broadcast.
- 1920: prima radio commerciale AM.

7.1 Introduzione

Modulare significa partire da un segnale $m(t)$ in banda base, detto **segnale modulante**, e "modificarlo" in modo da:

1. Spostare lo spettro attorno ad una certa frequenza f_c , detta **portante** (o **carrier**).
2. Essere in grado di ricostruire il segnale di informazione $m(t)$, in ricezione.

Le modulazioni analogiche si dividono in due grandi famiglie:

1. Modulazioni di ampiezza (*AM*, *Amplitude Modulation*).
2. Modulazioni angolari. Queste si suddividono a loro volta in:
 - (a) Modulazioni di frequenza (*FM*, *Frequency Modulation*);
 - (b) Modulazioni di fase (*PM*, *Phase Modulation*).

In questo capitolo analizzeremo diverse modulazioni di ampiezza:

1. AM tradizionale;
2. AM a portante soppressa (*AM-DSB-SC*, *Amplitude Modulation Double Side-Band Suppressed Carrier*);
3. AM-SSB (*Single SideBand*).

Il significato dei vari nomi sarà chiaro quando ne parleremo.

Le modulazioni di frequenza e di fase saranno invece esaminate nel prossimo capitolo.

7.2 Modulazione AM standard

Iniziamo a dare la definizione di segnale modulato in ampiezza. Utilizzeremo il formalismo del segnale analitico. Dato un segnale modulante $m(t)$, si impone:

$$\tilde{s}(t) \triangleq A_c [1 + m(t)]$$

in termini di inviluppo complesso. Il *segnale modulato* sarà dato dal corrispondente segnale reale, cioè:

$$s(t) = \text{Re} \left[\tilde{s}(t) e^{j2\pi f_c t} \right]$$

Definizione
di segnale
modulato
AM

da cui:

$$s(t) = A_c [1 + m(t)] \cos(2\pi f_c t)$$

Cerchiamo di capire l'andamento nel tempo del segnale modulato AM. Consideriamo la figura 7.1. Innanzitutto supponiamo che $m(t)$ sia normalizzato in modo da assumere valori in $[-1, +1]$ (caso tipico per questo tipo di modulazione). Di conseguenza, $[1+m(t)]$ assumerà valori in $[0, 2]$. Il segnale modulato $s(t)$ avrà allora un andamento compreso nell'intervallo $[-2A_c, +2A_c]$.

7.2.1 Calcolo densità spettrale di potenza per AM standard

Calcoliamo lo spettro di potenza di un segnale modulato in ampiezza. A tale scopo sfruttiamo un risultato che avevamo dimostrato per gli spettri degli inviluppi complessi, cioè:

$$G_s(f) = \frac{1}{4} [G_{\tilde{s}}(f - f_c) + G_{\tilde{s}}(-f - f_c)] \quad (7.1)$$

Bisogna quindi prima calcolare $G_{\tilde{s}}(f)$, cioè la densità spettrale di potenza dell'inviluppo complesso. Poi si ottiene $G_s(f)$. Occorre fare delle ipotesi sul segnale modulante. Consideriamo il segnale modulante $m(t)$ come un processo casuale con le seguenti caratteristiche:

- Stazionario;
- Ergodico;
- A valor medio nullo;
- Con densità spettrale di potenza nota e in banda base.

Densità
spettrale di
potenza
segnale AM

$G_s(f)$ in questo caso vale:

$$G_s(f) = \frac{A_c^2}{4} [\delta(f - f_c) + G_m(f - f_c) + \delta(f + f_c) + G_m(f + f_c)] \quad (7.2)$$

Dimostrazione. Iniziamo a calcolare la densità spettrale di potenza dell'inviluppo complesso $G_{\tilde{s}}(f)$. Partiamo dall'inviluppo complesso:

$$\tilde{s}(t) = A_c [1 + m(t)]$$

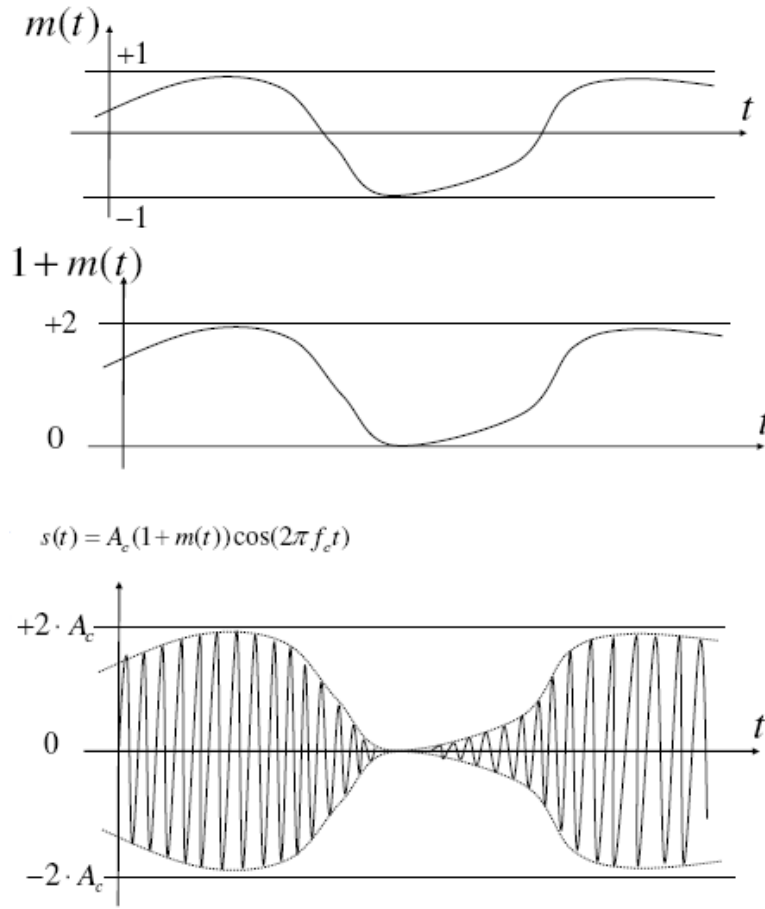


Figura 7.1: Esempio di segnale modulato in ampiezza. In alto c'è il segnale modulante $m(t)$, che assume valori in $[-1, +1]$. In mezzo c'è $[1 + m(t)]$, che assume quindi valori in $[0, 2]$. In basso c'è il segnale modulato $A_c[1 + m(t)] \cos(2\pi f_c t)$ che assume valori in $[-2A_c, +2A_c]$.

e calcoliamone la funzione di autocorrelazione. Ricordando che $E\{\cdot\}$ è un operatore lineare, si ottiene:

$$\begin{aligned}
 R_{\tilde{s}}(\tau) &= E[\tilde{s}(t) \tilde{s}(t + \tau)] \\
 &= A_c^2 \cdot E\{[1 + m(t)][1 + m(t + \tau)]\} \\
 &= A_c^2 \cdot E[1 + m(t + \tau) + m(t) + m(t)m(t + \tau)] \\
 &= A_c^2 + A_c^2 \underbrace{E[m(t + \tau)]}_{=0} + \underbrace{E[m(t)]}_{=0} + \underbrace{E[m(t)m(t + \tau)]}_{=R_m(\tau)} \quad (7.3)
 \end{aligned}$$

Per ipotesi di stazionarietà e valor medio nullo:

$$E[m(t + \tau)] = E[m(t)] = 0$$

e inoltre per definizione:

$$E[m(t)m(t + \tau)] = R_m(\tau) \quad \text{funzione di autocorrelazione di } m(t)$$

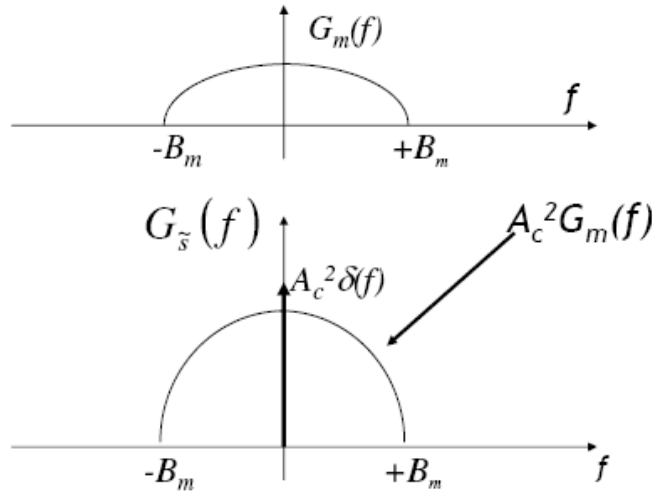


Figura 7.2: Spettro di potenza di un generico segnale modulante in banda base $m(t)$ (sopra) e spettro di potenza dell'involuppo complesso di $m(t)$ modulato in ampiezza (sotto).

Quindi la (7.3) diventa:

$$R_{\tilde{s}}(\tau) = A_c^2 [1 + R_m(\tau)]$$

Dall'autocorrelazione posso calcolare lo spettro di potenza:

$$\begin{aligned} G_{\tilde{s}}(f) &= F[R_{\tilde{s}}(\tau)] \\ &= F\{A_c^2 [1 + R_m(\tau)]\} \\ &= A_c^2 [\delta(f) + G_m(f)] \end{aligned} \quad (7.4)$$

Dato dunque un generico spettro del segnale modulante $m(t)$ in banda base, del tipo di quello in alto in figura 7.2, si ottiene uno spettro dell'involuppo complesso del tipo di quello in basso in figura 7.2. Si noti che, essendo $m(t)$ reale, $G_m(f)$ è una funzione pari. Vediamo allora dalla (7.4) che anche $G_{\tilde{s}}(f)$ è pari. Quindi:

$$G_{\tilde{s}}(-f - f_c) = G_{\tilde{s}}(f + f_c)$$

Sostituendo quest'ultima espressione nella (7.1) trovo che:

$$\begin{aligned} G_s(f) &= \frac{1}{4} [G_{\tilde{s}}(f - f_c) + G_{\tilde{s}}(-f - f_c)] \\ &= \frac{1}{4} [G_{\tilde{s}}(f - f_c) + G_{\tilde{s}}(f + f_c)] \\ &= \frac{A_c^2}{4} [\delta(f - f_c) + G_m(f - f_c) + \delta(f + f_c) + G_m(f + f_c)] \end{aligned}$$

Abbiamo quindi dimostrato la (7.2). ■

Commenti su
spettro
segnale AM

La figura 7.3 mostra un andamento qualitativo di questo spettro. Si noti che:

1. Il contenuto spettrale di $m(t)$ è stato spostato attorno alle frequenze f_c ;
2. L'occupazione spettrale è raddoppiata (si è passati da B_m in banda base a $2B_m$ in banda traslata);
3. In $+f_c$ e $-f_c$ sono presenti due righe spettrali che *non portano informazione*.

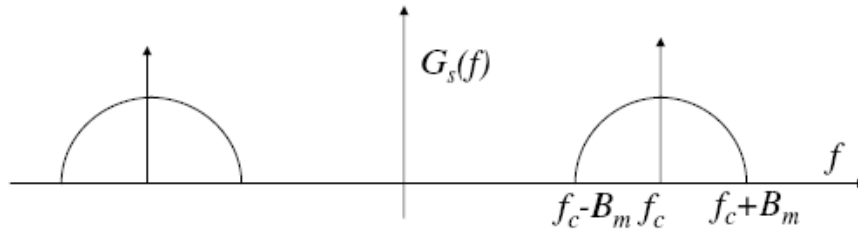


Figura 7.3: Spettro di potenza di generico segnale modulato AM. Si noti che ci sono due righe in $+f_c$ e $-f_c$ e che il segnale modulato ha una occupazione di banda doppia rispetto al segnale in banda base.

Calcolo potenza media segnale AM

Vogliamo ora calcolare la potenza media di un segnale modulato AM. Dalla teoria sul segnale analitico avevamo in generale:

$$P_s = \frac{1}{2} P_{\tilde{s}} \quad (7.5)$$

Nel caso di modulazione AM abbiamo:

$$\tilde{s}(t) = A_c [1 + m(t)]$$

Calcoliamo dunque la potenza dell'involuppo complesso, cioè $P_{\tilde{s}}$. Ricordando la linearità del valor medio e che il processo è ergodico si ottiene:

$$\begin{aligned} P_{\tilde{s}} &= A_c^2 \langle |1 + m(t)|^2 \rangle \\ &= A_c^2 [1 + \underbrace{\langle m^2(t) \rangle}_{P_m} + 2 \underbrace{\langle m(t) \rangle}_{=0}] \end{aligned}$$

Come abbiamo ipotizzato prima, $m(t)$ è a media nulla, quindi $\langle m(t) \rangle = 0$. Allora:

$$P_{\tilde{s}} = A_c^2 (1 + P_m)$$

che, sostituito nella (7.5) dà finalmente il valore di P_s :

$$\begin{aligned} P_s &= \frac{P_{\tilde{s}}}{2} \\ &= \frac{A_c^2}{2} (1 + P_m) \end{aligned}$$

Potenza
media
segnale
modulato
AM

Si dovranno fare spesso calcoli di questo tipo. Verifichiamo dunque per esercizio la congruenza tra le varie formule ottenute:

1. Calcoli nel dominio delle frequenze:

$$\begin{aligned} P_s &= \int_{-\infty}^{+\infty} G_s(f) df \\ &= 2 \int_0^{+\infty} \frac{A_c^2}{4} [\delta(f - f_c) + G_m(f - f_c)] df \\ &= \frac{A_c^2}{2} \left[\int_0^{+\infty} \delta(f - f_c) df + \int_0^{+\infty} G_m(f - f_c) df \right] \\ &= \frac{A_c^2}{2} (1 + P_m) \end{aligned}$$

2. Calcoli nel dominio del tempo:

$$\begin{aligned}
 P_s &= \langle |s(t)|^2 \rangle \\
 &= A_c^2 \langle [1 + m(t)]^2 \cdot \cos^2(2\pi f_c t) \rangle \\
 &= A_c^2 \langle [1 + m^2(t) + 2m(t)] \cdot \cos^2(2\pi f_c t) \rangle \\
 &\stackrel{1}{=} A_c^2 \left[\underbrace{\langle \cos^2(2\pi f_c t) \rangle}_{=\frac{1}{2}} + 2 \underbrace{\langle m(t) \cos^2(2\pi f_c t) \rangle}_{=0} + \underbrace{\langle m^2(t) \cos^2(2\pi f_c t) \rangle}_{=\frac{1}{2}P_m} \right] \\
 &= A_c^2 \left(\frac{1}{2} + \frac{1}{2}P_m \right) \\
 &= \frac{A_c^2}{2} (1 + P_m)
 \end{aligned}$$

I risultati tornano.

7.2.2 Ricevitore a rivelazione di inviluppo

In questo paragrafo vogliamo mostrare il modo più semplice per demodulare un segnale AM. Prima è però utile definire le *percentuali di modulazione*.

Percentuali di modulazione

Si definiscono in questo modo:

- Percentuale di modulazione totale $\triangleq \frac{A_{max} - A_{min}}{2A_c} 100 = \frac{\max[m(t)] - \min[m(t)]}{2} \cdot 100$
- Percentuale di modulazione negativa $\triangleq \frac{A_c - A_{min}}{A_c} 100 = -\min[m(t)] \cdot 100$
- Percentuale di modulazione positiva $\triangleq \frac{A_{max} - A_c}{A_c} 100 = \max[m(t)] \cdot 100$

Riscriviamo qui per comodità l'espressione di un segnale modulato AM:

$$s(t) = A_c[1 + m(t)] \cos(2\pi f_c t)$$

Se $\max[m(t)] = +1$ e $\min[m(t)] = -1$ si dice che il segnale AM è modulato al 100%. Solitamente si vuole che:

$$1 + m(t) > 0, \text{ cioè } m(t) > -1$$

Condizione
di assenza di
sovramodula-
zione

per *non avere sovramodulazione*. In questo caso in ricezione è possibile usare un semplice rivelatore di inviluppo, come descritto tra poco. Notare che questo richiede che la percentuale di modulazione negativa sia al più il 100%, mentre in teoria quella positiva può essere $> 100\%$.

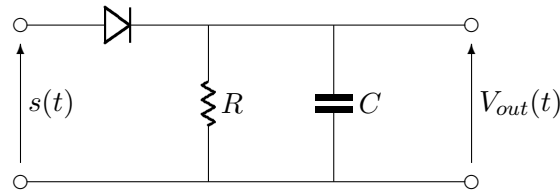


Figura 7.4: Circuito rivelatore di inviluppo.

Circuito rivelatore di inviluppo

Il rivelatore di inviluppo è mostrato in figura 7.4. Si tratta di un circuito che segue l'inviluppo del segnale in ingresso. Matematicamente, dato in ingresso un segnale in banda traslata $s(t)$, il rivelatore di inviluppo estrae all'uscita un segnale $V_{out}(t)$ proporzionale al modulo dell'inviluppo complesso di $s(t)$:

Funzionamento
rivelatore di
inviluppo

$$V_{out}(t) = |\tilde{s}(t)|$$

Ad esempio per una modulazione AM si ha:

$$\begin{aligned} V_{out}(t) &= |\tilde{s}(t)| \\ &= |A_c [1 + m(t)]| \end{aligned}$$

In assenza di sovrarmodulazione, cioè se per qualsiasi t ho che $[1 + m(t)] > 0$:

$$V_{out}(t) = A_c [1 + m(t)]$$

cioè si ricostruisce perfettamente $m(t)$.

Se invece si è in presenza di sovrarmodulazione:

$$1 + m(t) < 0 \text{ per qualche } t$$

si ha un segnale che crea sovrarmodulazione. Si nota infatti che se $1 + m(t) < 0$, il segnale ricostruito tramite una rivelazione di inviluppo *non è più il segnale di partenza*. Si ha una distorsione significativa del segnale di partenza, in quanto:

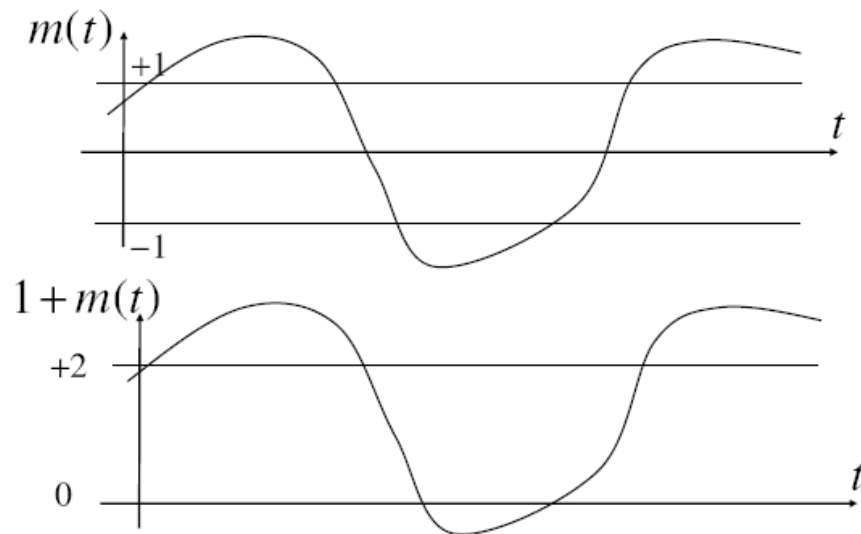
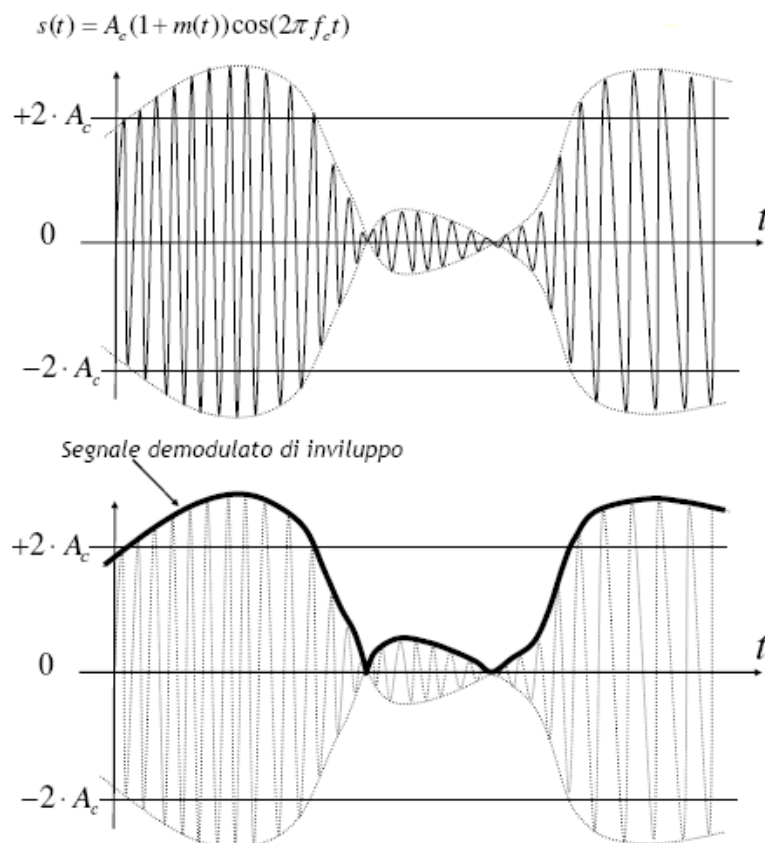
In caso di
sovrarmodula-
zione
ricevitore di
inviluppo
non funziona

$$\begin{aligned} V_{out}(t) &= |A_c [1 + m(t)]| \\ &\neq A_c [1 + m(t)] \end{aligned}$$

Dato che nella stragrande maggioranza dei casi pratici il segnale $m(t)$ presenta una sorta di "simmetria" sui valori positivi e negativi delle ampiezze, in sostanza la condizione di assenza di sovrarmodulazione è semplicemente:

$$|m(t)| \leq 1$$

La figura 7.5(a) mostra un esempio di segnale $m(t)$ che crea sovrarmodulazione. La figura 7.5(b) mostra l'effetto della demodulazione di inviluppo di tale segnale. Si noti come il segnale demodulato non sia una copia fedele di $m(t)$. Il motivo è che ci sono dei valori di t per cui $m(t) < 0$.

(a) Esempio di segnale $m(t)$ che crea sovramodulazione.(b) Segnale $m(t)$ della figura (a) modulato in ampiezza (sopra) e risultato della demodulazione di inviluppo (sotto). Si noti come a causa della sovramodulazione, il segnale demodulato non sia una copia fedele di quello di ingresso.**Figura 7.5:** Esempio di segnale che crea sovramodulazione.

Efficienza di modulazione

Abbiamo visto che la potenza media di un segnale modulato AM vale:

$$\begin{aligned} P_s &= \langle s^2(t) \rangle \\ &= \underbrace{\frac{1}{2}A_c^2}_{P_{portante}} + \underbrace{\frac{1}{2}A_c^2 \langle m^2(t) \rangle}_{P_{utile}} \end{aligned}$$

dove:

- La prima parte indica la *potenza relativa alla riga della portante*;
- La seconda parte è la *potenza utile che porta informazione*.

La potenza della riga di portante è in un certo senso una potenza "sprecata". Si definisce allora l'*efficienza di modulazione* come il rapporto tra la potenza utile e la potenza totale:

$$E \triangleq \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle}$$

Efficienza di
modulazione

Notare che, *se non si vuole sovramodulare*, cioè se $\max[|m(t)|] \leq 1$

- allora $\max[m^2(t)] \leq 1$
- e dunque $\max[\langle m^2(t) \rangle] \leq 1$

$\langle m^2(t) \rangle = 1$ corrisponde al caso limite di un'onda quadra che assume valori ± 1 . In tal caso l'efficienza di modulazione è del 50%. In tutti gli altri casi, l'efficienza di modulazione è inferiore.

La massima efficienza di modulazione che si può ottenere se non si vuole la sovramodulazione è quindi pari al 50%. Nei casi pratici l'efficienza è inferiore, come mostrato nell'esempio che segue.

Esempio 7.1. Efficienza di modulazione di un segnale sinusoidale Consideriamo un segnale $m(t)$ sinusoidale:

$$m(t) = \cos(2\pi f_m t)$$

Calcoliamo l'efficienza di modulazione supponendo di non avere sovramodulazione. Avrò:

$$\begin{aligned} \langle m^2(t) \rangle &= \langle \cos^2(2\pi f_m t) \rangle \\ &= \frac{1}{2} \end{aligned}$$

quindi

$$\begin{aligned} E &= \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle} \\ &= \frac{\frac{1}{2}}{1 + \frac{1}{2}} \\ &= \frac{1}{3} \end{aligned}$$

Efficienza di
modulazione
segnale
sinusoidale

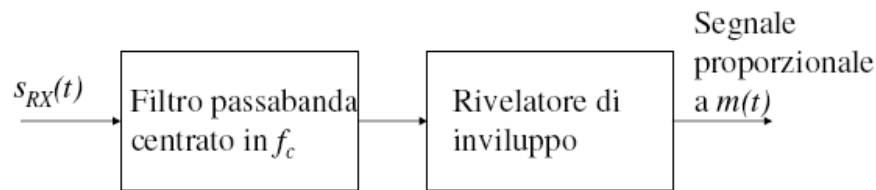


Figura 7.6: Schema a blocchi demodulatore AM a rivelazione di inviluppo.

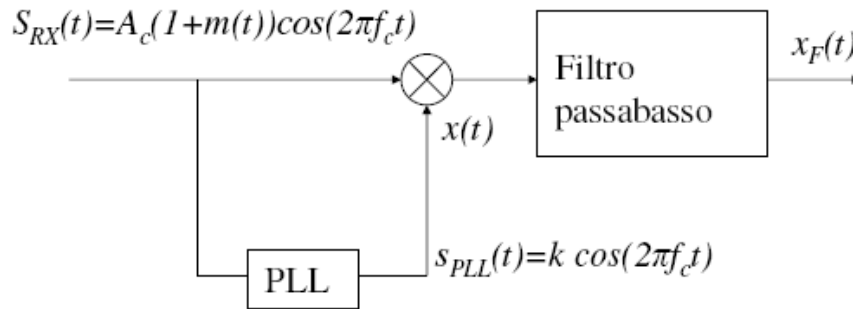


Figura 7.7: Demodulatore AM coerente. Il PLL recupera esattamente la frequenza e la fase della portante. Il filtro passa-basso elimina le due componenti centrati attorno a $+2f_0$ e $-2f_0$.

Per un segnale sinusoidale l'efficienza è del 33%. Questo è un valore tipico per segnali reali, tipo segnale musicale. ■

Conclusioni: AM e sovramodulazione

Per la *modulazione AM tradizionale*, cioè per:

$$s(t) = A_c[1 + m(t)] \cos(2\pi f_c t)$$

se si dimensiona il segnale $m(t)$ in modo da non avere sovramodulazione, si può usare un ricevitore molto semplice, a rivelazione di inviluppo, il cui schema a blocchi è quello mostrato in figura 7.6. L'efficienza di modulazione è tuttavia al più del 50%, e tipicamente inferiore. Cioè si spreca *almeno* il 50% della potenza per trasmettere il segnale di portante. Si tratta del classico esempio di scelta tra semplicità del ricevitore e utilizzo efficace della potenza.

7.3 Modulazione AM-DSB-SC

Obiettivo:
ridurre
potenza di
portante

Ci poniamo ora l'obiettivo di ridurre o annullare la potenza relativa alla portante per la modulazione AM, accettando sovramodulazione. Ipotizziamo di poter ricostruire al ricevitore un segnale con modulo e fase uguale alla portante (ad esempio tramite un *PLL*, *Phase Locked Loop*, in italiano *anello ad aggancio di fase*). Un ricevitore di questo tipo si chiama *coerente*. La figura 7.7 mostra lo schema a blocchi di un ricevitore AM coerente. Analizziamone ora il funzionamento. $s_{PLL}(t)$ è un segnale in fase con la

portante (e di ampiezza generica K). Ottengo:

$$\begin{aligned} x(t) &= s_{RX}(t) \cdot s_{PLL}(t) \\ &= kA_c [1 + m(t)] [\cos^2(2\pi f_c t)] \\ &= \frac{kA_c}{2} [1 + m(t)] [1 + \cos(4\pi f_c t)] \end{aligned}$$

$x(t)$ viene poi filtrato con un filtro passa-basso con banda pari alla banda del segnale modulante. Il filtro passa-basso elimina il segnale a frequenza doppia $2f_c$, ottenendo:

$$x_F(t) = \frac{kA_c}{2} [1 + m(t)]$$

In uscita si ottiene quindi un segnale identico (salvo costanti moltiplicative) al segnale modulante. Notare che adesso da $x_F(t)$ è possibile ricostruire perfettamente $m(t)$, senza nessun tipo di richiesta su $m(t)$. Al limite si può anche avere:

$$s(t) = A_c m(t) \cos(2\pi f_c t)$$

Questo tipo di modulazione viene detta *modulazione AM-DSB-SC*.

Modulazione
AM-DSB-SC

- *DSB* = *double-sideband* (spettro bilaterale, si tornerà su questa questione più avanti);
- *SC* = *suppressed carrier* (cioè non si ha la riga della portante).

Rifacendo conti del tutto analoghi a quelli appena fatti (basta togliere un $+1$), si ottiene che per una demodulazione coerente di un segnale modulato AM-DSB-SC:

$$x_F(t) = \frac{kA_c}{2} m(t)$$

cioè si riesce a ricostruire esattamente il segnale di partenza.

Facendo calcoli del tutto analoghi a quelli già svolti per AM standard, si ottiene la seguente densità spettrale di potenza per AM-DSB-SC:

$$G_s(f) = \frac{A_c^2}{4} [G_m(f - f_c) + G_m(f + f_c)]$$

Densità
spettrale
AM-DSB-SC

Si noti che ora non compare nessuna delta in frequenza. Questo significa che non si spreca potenza per la portante, cioè si raggiunge un'efficienza di modulazione del 100% ($E = 1$):

AM-DSB-SC
ha efficienza
pari a 1

$$\begin{aligned} E &= \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle} \\ &= \frac{\langle m^2(t) \rangle}{\langle m^2(t) \rangle + 1} \\ &= 1 \end{aligned}$$

Conclusione: il ricevitore coerente permette di demodulare correttamente un segnale anche con forte sovramodulazione o (al limite) di tipo AM-DSB-SC, ottenendo un migliore utilizzo della potenza (fino a $E = 1$). Questo risultato si ottiene a scapito di

un aumento di complessità del ricevitore rispetto al ricevitore di inviluppo.

Considerazioni pratiche: Un circuito di aggancio di fase (PLL) è facile da realizzare con la tecnologia attuale, ma non lo era all'inizio delle trasmissioni radio. Gli standard pratiche su per la trasmissione AM radio (attorno a 1 MHz) sono tuttavia basati su AM standard realizzazione (con portante e senza sovr modulazione) e sono dunque compatibili con un ricevitore di PLL inviluppo. Tuttavia in tutti i ricevitori moderni si usano comunque i PLL, anche se il segnale è AM standard, in quanto si riescono ad ottenere prestazioni migliori.

Indice di modulazione

Si vuole ora introdurre una notazione che permetta di descrivere sia AM standard che AM-DSB-SC. Sia dunque:

$$s(t) = A_c [1 + m(t)] \cos(2\pi f_c t)$$

e sia $x(t)$ il segnale che porta informazione, supposto normalizzato in modo che $\max [|m(t)|] = 1$. Si introduce il concetto di indice di modulazione:

$$m(t) = m x(t)$$

dove m è detto *indice di modulazione*:

- $m = 1$ per AM standard senza sovr modulazione.
- $m \gg 1$ per AM-DSB-SC.

Ci possono essere casi intermedi, in cui $m > 1$, ma si lascia volutamente una parte della portante per facilitare il circuito di recupero di fase.

7.4 Modulazione SSB

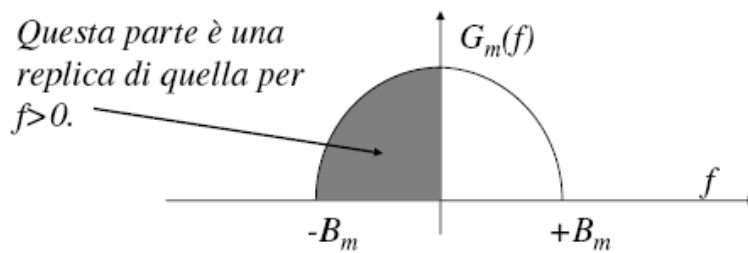
Nei paragrafi precedenti abbiamo visto che le modulazioni AM¹ occupano sul canale il doppio della banda del segnale modulante. L'allocazione di banda (soprattutto nelle trasmissioni via etere) è molto costosa. Cerchiamo ora un metodo per limitare il più possibile la banda utilizzata sul canale.

Idea di base modulazione SSB Esiste una tecnica che permette di dimezzare la banda occupata sul canale, basandosi sul fatto che un segnale reale ha spettro simmetrico rispetto all'asse delle ordinate (funzione pari). Consideriamo un generico segnale modulante reale $m(t)$ in banda base. In quanto segnale reale, il suo spettro è simmetrico in frequenza. Osserviamo che, per lo meno in linea di principio, la conoscenza delle sole componenti a frequenza positiva permette di conoscere tutto il segnale. Posso infatti ricostruire la parte delle frequenze negative per simmetria (vedi figura 7.8(a)). Consideriamo una normale modulazione AM-DSB-SC:

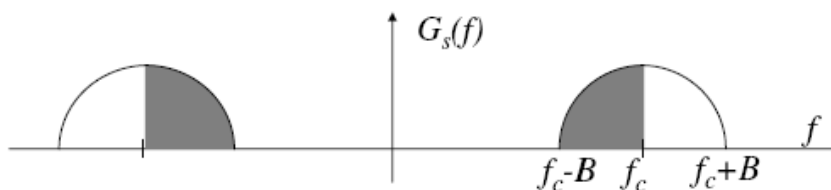
$$s(t) = A_c m(t) \cos(2\pi f_c t)$$

Almeno in linea teorica, la parte evidenziata in grigio nella figura 7.8(b) è inutile per la trasmissione del segnale, poiché è simmetrica rispetto all'altra metà (le modulazioni

¹Sia AM standard che AM-DSB-SC.



(a) Esempio di spettro di segnale $m(t)$ reale. Ogni segnale reale ha spettro pari in frequenza. Pertanto le frequenze negative (parte evidenziata in grigio) non portano nessuna informazione.



(b) Esempio di spettro di segnale modulato. Poiché i due spettri sono simmetrici le parti evidenziate in grigio sono inutili per la trasmissione del segnale.

Figura 7.8: Idea alla base della modulazione SSB.

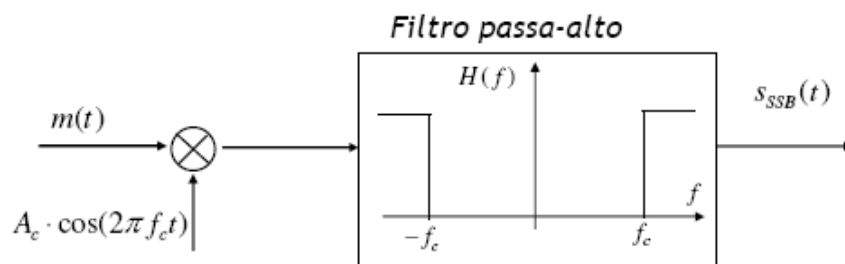


Figura 7.9: Schema di principio del trasmettitore SSB.

AM precedenti sono dette DSB perché vi sono entrambe le "sideband"). **Allora è possibile eliminare metà della banda** (ad esempio con un filtro opportuno), **senza perdere informazione**. Si ottiene la modulazione detta *AM-SSB*, (*Single Side-Band*, detta anche *a banda laterale unica*). Un possibile schema del trasmettitore è mostrato in figura 7.9.

Esistono due possibili schemi:

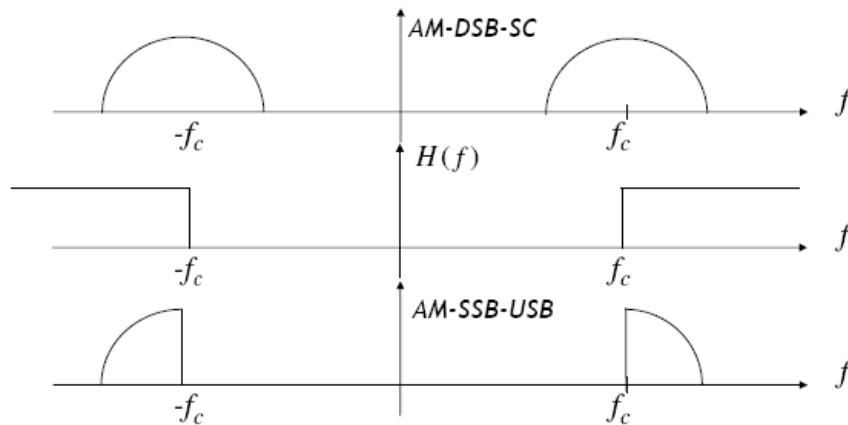
Varianti:
USB e LSB

- AM-SSB-USB (*USB = Upper Side Band*) (vedi figura 7.10(a));
- AM-SSB-LSB (*LSB = Lower Side Band*) (vedi figura 7.10(b)).

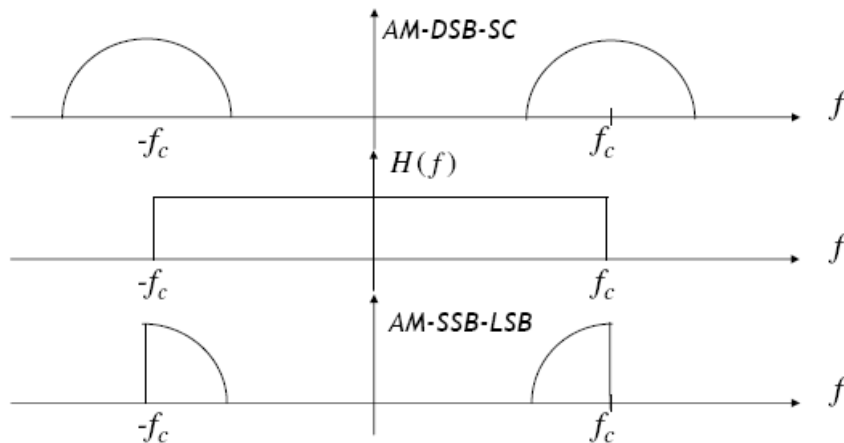
La modulazione SSB ha vantaggi e svantaggi rispetto a AM e AM-DSB-SC:

Vantaggi e
svantaggi

- Vantaggio: occupa metà banda rispetto a AM e a AM-DSB-SC (e si vedrà che le prestazioni in termini di SNR sono le stesse di AM-DSB-SC).



(a) Illustrazione costruzione segnale AM-USB nel dominio della frequenza. Si filtra il segnale AM-DSB-SC con un filtro passa-alto con frequenza di taglio f_c .



(b) Illustrazione costruzione segnale AM-LSB nel dominio della frequenza. Si filtra il segnale AM-DSB-SC con un filtro passa-basso con frequenza di taglio f_c .

Figura 7.10: Illustrazioni costruzione segnali SSB.

- Svantaggio: gli apparati TX e RX sono più complicati. In particolare, non è possibile usare il rivelatore di inviluppo, quindi è necessaria una demodulazione coerente.

La modulazione AM-SSB è usata ad esempio per:

- Radioamatori;
- Sistemi radio per la navigazione.

7.4.1 Schema alternativo per il modulatore SSB

Analizziamo ora le modulazioni SSB in maniera più rigorosa nel dominio del tempo. Questo approccio servirà a capire come può essere fatto il ricevitore di un sistema SSB.

Consideriamo un segnale modulante reale in banda base $m(t)$. Introduciamo il segnale analitico corrispondente a $m(t)$. Notare che questo è un approccio "anomalo",

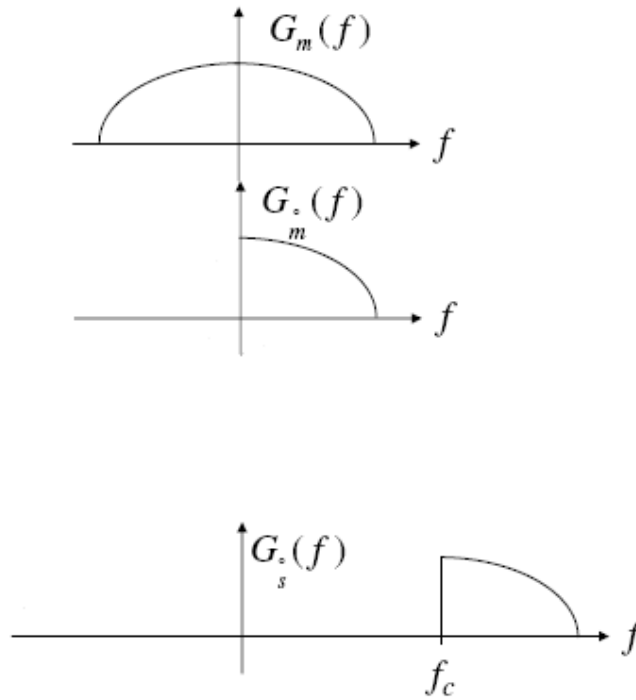


Figura 7.11: Spettro di potenza di un generico segnale $m(t)$ in banda base (sopra), spettro di potenza del corrispondente segnale analitico (in mezzo) e spettro di potenza del nuovo segnale analitico $\overset{\circ}{s}(t)$ (in basso).

in quanto $m(t)$ è in banda base, e dunque solitamente non si usa il segnale analitico. Tuttavia è possibile definire il segnale analitico per *qualunque segnale*, anche in banda base, semplicemente applicando la definizione.

Si faccia riferimento alla figura 7.11. In generale, dato $m(t)$ sappiamo che il corrispondente segnale analitico può essere scritto come:

$$\overset{\circ}{m}(t) = m(t) + j \hat{m}(t)$$

Si consideri adesso il nuovo segnale analitico definito come:

$$\begin{aligned} \overset{\circ}{s}(t) &\triangleq A_c \overset{\circ}{m}(t) e^{j2\pi f_c t} \\ &= A_c [m(t) + j \hat{m}(t)] e^{j2\pi f_c t} \end{aligned}$$

ed il corrispondente spettro (vedi figura 7.11, in basso). Consideriamo ora il corrispondente segnale reale:

$$\begin{aligned} s(t) &= \text{Re} [\overset{\circ}{s}(t)] \\ &= \text{Re} \left\{ A_c [m(t) + j \hat{m}(t)] e^{j2\pi f_c t} \right\} \\ &= A_c [m(t) \cos(2\pi f_c t) - \hat{m}(t) \sin(2\pi f_c t)] \end{aligned} \quad (7.6)$$

Il suo spettro è mostrato in figura 7.12. Allora il segnale reale così ottenuto:

$$s_{USB}(t) = A_c [m(t) \cos(2\pi f_c t) - \hat{m}(t) \sin(2\pi f_c t)]$$

Segnale
SSB-USB nel
dominio del
tempo

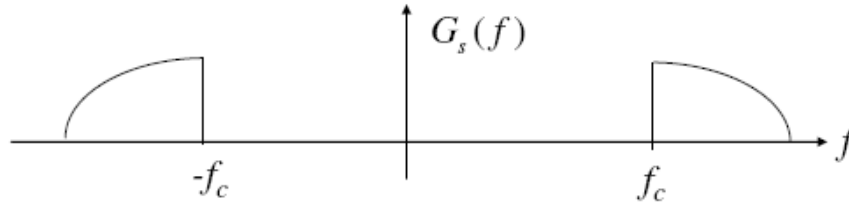


Figura 7.12: Spettro di potenza del segnale (7.6). Si tratta di un segnale SSB-USB

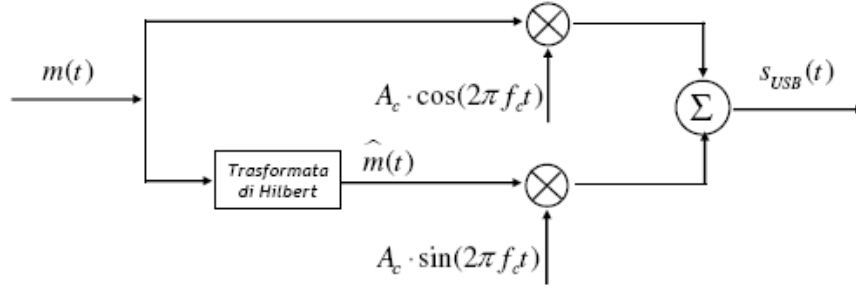


Figura 7.13: Schema di modulatore SSB-USB nel dominio del tempo.

è effettivamente il segnale SSB-USB corrispondente al segnale modulante $m(t)$. Abbiamo dunque trovato un modo per generare un segnale SSB-USB lavorando nel dominio del tempo. Lo schema di questo modulatore è mostrato in figura 7.13.

Filtro di
Hilbert

Ricordando quanto ottenuto nel capitolo sul segnale analitico, la trasformata di Hilbert del segnale $m(t)$ si ottiene come segue:

$$\hat{m}(t) = m(t) * \frac{1}{\pi t}$$

cioè tramite un filtraggio con risposta all'impulso:

$$h_{Hilbert}(t) = \frac{1}{\pi t}$$

Il filtraggio non è di per sé fisicamente realizzabile (non è causale) e dunque deve essere opportunamente approssimato.

7.4.2 Demodulazione di segnali SSB

Si usa un de-
modulatore
coerente

Dimostriamo ora che un segnale AM-SSB può essere demodulato tramite un *demodulatore coerente* (del tutto analogo a quello visto per AM-DSB). Si consideri infatti l'espressione (ad esempi per USB):

$$s_{USB}(t) = A_c [m(t) \cos(2\pi f_c t) - \hat{m}(t) \sin(2\pi f_c t)]$$

Moltiplicando per $\cos(2\pi f_c t)$ si ottiene:

$$A_c [m(t) \cos^2(2\pi f_c t) - \hat{m}(t) \sin(2\pi f_c t) \cos(2\pi f_c t)]$$

che può essere riscritto come:

$$\frac{A_c}{2} \{m(t) [1 + \cos(4\pi f_c t)] - \hat{m}(t) \sin(4\pi f_c t)\}$$

Dopo un filtraggio passa-basso sulla banda del segnale modulante, si cancellano i termini attorno alla frequenza $2f_c$, ottenendo:

$$\frac{A_c}{2} m(t)$$

e dunque si ricostruisce esattamente il segnale modulante $m(t)$.

Non è invece possibile usare un demodulatore di inviluppo. Infatti per un demodulatore di inviluppo si ha:

$$V_{\text{inviluppo}}(t) = |\tilde{s}(t)|$$

ma in questo caso $\tilde{s}(t) = \tilde{s}_{USB}(t)$, quindi:

$$\tilde{s}_{USB}(t) = A_c [m(t) + j \hat{m}(t)] \quad \Rightarrow \quad V_{\text{inviluppo}}(t) = |A_c [m(t) + j \hat{m}(t)]|$$

$$V_{\text{inviluppo}}(t) = A_c \sqrt{m^2(t) + \hat{m}^2(t)}$$

che non è una ricostruzione del segnale modulante $m(t)$.

Con SSB non si può usare demodulatore di inviluppo

7.5 Prestazioni modulazioni AM in presenza di rumore

Iniziamo ora l'importante argomento dello studio delle prestazioni di sistemi di modulazione in presenza di rumore. Con il termine *prestazioni* si intende l'andamento del rapporto segnale-rumore (Signal to Noise Ratio, SNR o S/N) all'uscita del demodulatore. Il rapporto segnale rumore è cioè un indice della qualità della comunicazione.

In questo paragrafo vogliamo calcolare l'SNR delle modulazioni AM considerate in questo capitolo e confrontarlo con l'SNR di un sistema non modulato in banda base.

Prestazioni sistemi in banda base

Per confrontare i vari sistemi di modulazione, si prenderanno come riferimento le prestazioni di un sistema di trasmissione non modulato in banda base, il cui schema a blocchi è mostrato in figura 7.14. Le caratteristiche fondamentali di questo sistema di trasmissione sono:

- Potenza di segnale ricevuta P_{RX} ;
- Banda occupata dal segnale B_m ;
- Densità spettrale di rumore all'ingresso del ricevitore $N_0/2$;
- Filtraggio al ricevitore ideale (passa-basso ideale con banda B_m).

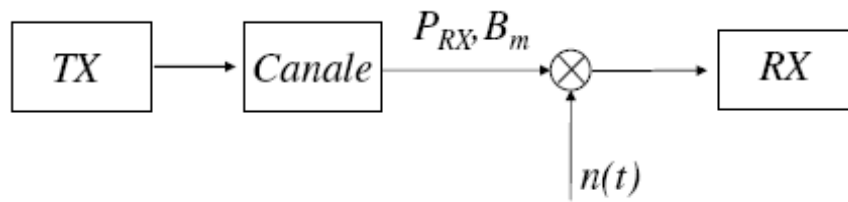


Figura 7.14: Schema a blocchi di un sistema di trasmissione non modulato in banda base. Verrà usato come riferimento per il calcolo delle prestazioni dei vari tipi di modulazione.

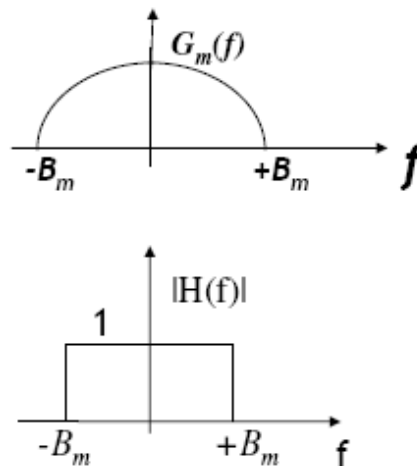


Figura 7.15: Esempio di spettro di potenza di segnale in banda base (sopra) e risposta in frequenza del filtro passa-basso ideale usato nel ricevitore (sotto). Si noti che il filtro ha la stessa banda del segnale di informazione.

Allora il rapporto segnale-rumore del sistema in banda base è dato da:

SNR sistema
di
riferimento
in banda
base

$$\left(\frac{S}{N} \right)_{BB} = \frac{P_{RX}}{N_0 B_m}$$

Questo valore verrà usato come riferimento per tutti i sistemi AM.

La figura 7.15 mostra un esempio di spettro di segnale trasmesso (in alto) e il relativo filtro di ricezione (in basso). Si tratta di un filtro passa-basso ideale con banda pari a quella del segnale trasmesso. Questa ipotesi sul filtro passa-basso di ricezione è la stessa che si farà anche in tutti i sistemi di ricezione per segnali modulati che studieremo nelle prossime pagine.

Schema di riferimento per modulazioni analogiche in banda traslata

Lo schema di principio di un sistema di comunicazione analogico in banda traslata è mostrato in figura 7.16. Questo schema di riferimento sarà adottato sia per i sistemi AM che per i successivi FM/PM (trattati nel prossimo capitolo). Le grandezze che lo caratterizzano sono:

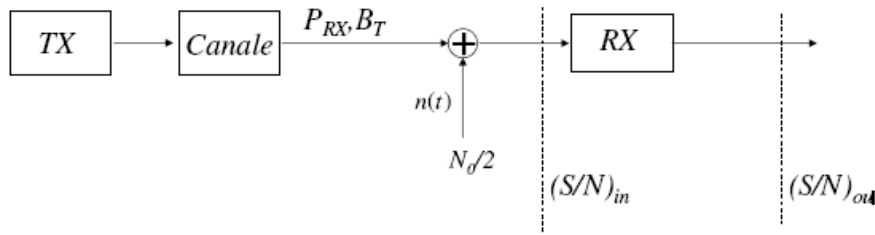


Figura 7.16: Schema di riferimento catena di trasmissione e ricezione sistema analogico.

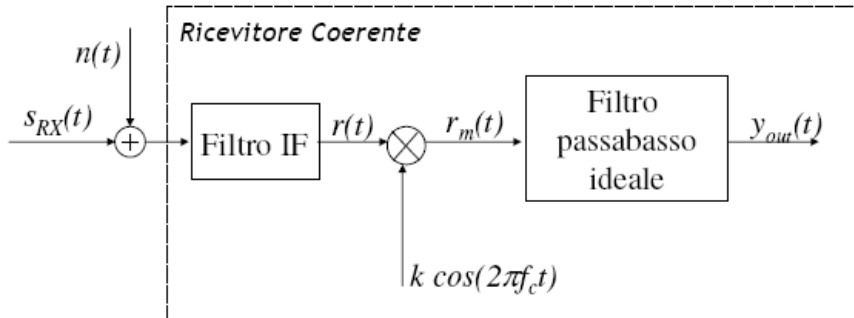


Figura 7.17: Schema ricevitore coerente di un sistema AM.

- P_{RX} : potenza del segnale modulato all'ingresso del ricevitore;
- B_m : banda del segnale modulante, cioè del segnale $m(t)$ che porta informazione;
- B_T : banda del segnale modulato;
- Rumore additivo $n(t)$: gaussiano bianco con densità spettrale di potenza $N_0/2$;
- $(S/N)_{in}$: SNR all'ingresso del ricevitore, calcolato sulla banda del segnale modulato B_T (attenzione!!);
- $(S/N)_{out}$: SNR all'uscita del ricevitore, calcolato sulla banda B_m . È il parametro più importante, poiché definisce la *qualità* del segnale per l'utente finale.

7.5.1 Calcolo SNR per sistemi AM tradizionali con ricezione coerente

Si inizia ora a studiare le prestazioni di un sistema di modulazione AM tradizionale con ricevitore coerente. A tale scopo consideriamo lo schema di figura 7.17. Vedremo che un ricevitore di questo tipo può essere usato anche per modulazioni AM-DSB-SC e SSB.

Il filtro IF (Intermediate Frequency) è un filtro passa-banda che, idealmente, lascia passare solo le frequenze del segnale modulato (vedi figura 7.18). Ha due diverse funzioni:

- Limitare la banda di rumore;
- Selezionare il canale voluto all'interno di un set di canali in multiplexazione di frequenza (*FDM, Frequency Division Multiplexing*), argomento sul quale si tornerà più avanti.

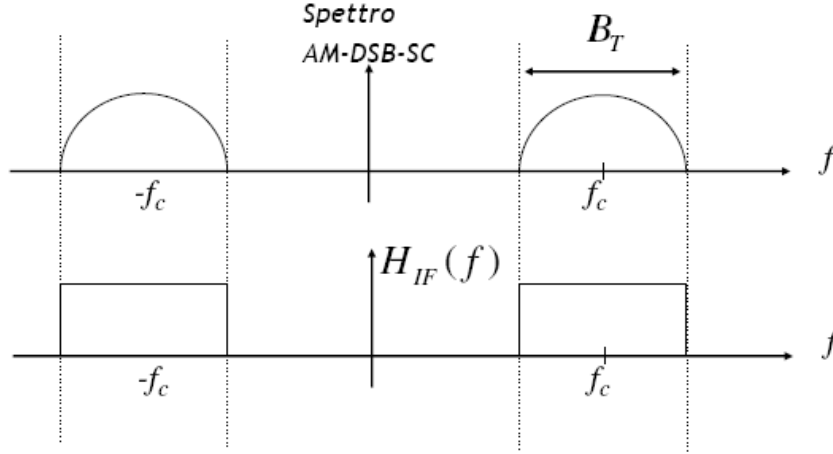


Figura 7.18: Filtro IF. Serve a far passare il segnale di informazione e a eliminare il rumore fuori banda e eventuali altri segnali su altre frequenze.

Esprimiamo i vari segnali in gioco:

$$\begin{aligned} s_{RX}(t) &= A_c[1 + m(t)] \cos(2\pi f_c t) \\ n(t) &= x_n(t) \cos(2\pi f_c t) - y_n(t) \sin(2\pi f_c t) \\ r(x) &= s_{RX}(t) + n(t) \end{aligned}$$

Importante:

- $n(t)$ è un rumore gaussiano bianco con densità spettrale di potenza $N_0/2$;
- $x_n(t)$ e $y_n(t)$ hanno densità spettrale di potenza pari a N_0 (si riveda la teoria sulla rappresentazione in fase e quadratura del rumore gaussiano in banda traslata);
- Il filtro IF ha banda uguale a quella del segnale modulato, ovvero B_T .

Il rapporto segnale-rumore in ingresso, calcolato sulla banda di trasmissione B_T , è dato da:

$$\begin{aligned} \left(\frac{S}{N}\right)_{in} &= \frac{P_{RX}}{N_0 B_T} \\ &= \frac{P_{RX}}{2 N_0 B_m} \\ &= \frac{1}{2} \left(\frac{S}{N}\right)_{BB} \end{aligned}$$

È in sostanza il rapporto segnale-rumore in ingresso al ricevitore, calcolato subito dopo il filtro IF. È una quantità che tornerà utile solo più avanti. A noi in realtà interesserà $(S/N)_{out}$ sul segnale demodulato, cioè il rapporto segnale-rumore sul segnale fornito all'utente del sistema di trasmissione.

Il segnale dopo il mixer è poi dato da:

$$r_m(t) = r(t) k \cos(2\pi f_c t) \quad (7.7)$$

dove:

$$r(t) = A_c [1 + m(t)] \cos(2\pi f_c t) + x_n(t) \cos(2\pi f_c t) - y_n(t) \sin(2\pi f_c t) \quad (7.8)$$

Sostituendo la (7.8) nella (7.7) si trova:

$$\begin{aligned} r_m(t) &= \{A_c[1 + m(t)] + x_n(t)\} k \cos^2(2\pi f_c t) - y_n(t) k \cos(2\pi f_c t) \sin(2\pi f_c t) \\ &= \frac{k}{2} \{A_c[1 + m(t)] + x_n(t)\} [1 + \cos(4\pi f_c t)] - \frac{k}{2} y_n(t) \sin(4\pi f_c t) \end{aligned}$$

Dobbiamo ora calcolare:

$$y_{out}(t) = r_m(t) * h(t)$$

dove $h(t)$ è la risposta all'impulso del filtro passa-basso ideale dopo il mixer. Tale filtro ha banda B_m . Si suppone che:

- Il filtro elimini tutte le componenti a $2f_c$;
- Il filtro faccia passare il segnale utile $m(t)$ senza distorsione;
- Il filtro limiti il rumore $x_n(t)$.

L'effetto del filtro è quindi il seguente:

$$\frac{k}{2} \{A_c[1 + m(t)] + x_n(t)\} \underbrace{[1 + \cos(4\pi f_c t)]}_{=0} - \underbrace{\frac{k}{2} y_n(t) \sin(4\pi f_c t)}_{=0}$$

Quindi il segnale in uscita dal ricevitore vale:

$$y_{out}(t) = \frac{k}{2} \{A_c[1 + m(t)] + x_n(t) * h(t)\} \quad (7.9)$$

Dato dunque il segnale all'uscita del sistema, calcoliamone il rapporto segnale-rumore. Il rapporto segnale-rumore in uscita dal ricevitore è la potenza della componente di segnale utile diviso la potenza della componente di rumore.

Componente utile: è la componente relativa al *solo* segnale che porta informazione, e cioè, tralasciando la parte continua che *non* porta in formazione, si ottiene il seguente segnale:

$$\frac{k}{2} A_c m(t)$$

che ha potenza:

$$\text{Potenza segnale utile demodulato} = \frac{k^2}{4} A_c^2 \langle m^2(t) \rangle \quad (7.10)$$

Componente di rumore:

$$\frac{k}{2} [x_n(t) * h(t)]$$

Si presti attenzione su questo punto. Consideriamo la figura 7.19. $x_n(t)$ ha densità spettrale di potenza N_0 su una banda B_T più grande di B_m . $x_n(t) * h(t)$ ha dunque potenza $2N_0B_m$. La potenza associata alla componente di rumore vale dunque

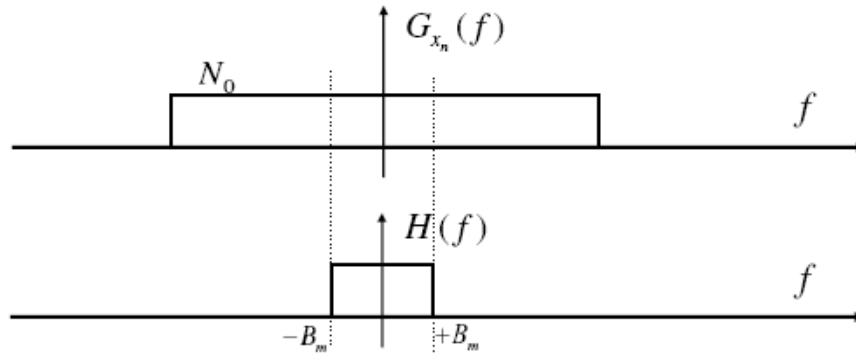


Figura 7.19: Densità spettrale di potenza di $x_n(t)$ (in alto) e risposta in frequenza del filtro passa-basso ideale (in basso). Dalla figura si vede bene che questo filtro limita ulteriormente la potenza di rumore.

complessivamente:

$$\begin{aligned}
 \text{Potenza rumore} &= \left\langle \left\{ \frac{k}{2} [x_n(t) * h(t)] \right\}^2 \right\rangle \\
 &= \frac{k^2}{4} \langle [x_n(t) * h(t)]^2 \rangle \\
 &= \frac{k^2}{4} 2 N_0 B_m \\
 &= \frac{k^2 N_0 B_m}{2}
 \end{aligned} \tag{7.11}$$

Il rapporto segnale rumore in uscita è quindi dato dal rapporto tra la (7.10) e la (7.11):

$$\begin{aligned}
 \left(\frac{S}{N} \right)_{out} &= \frac{\text{Potenza segnale utile demodulato}}{\text{Potenza rumore}} \\
 &= \frac{\frac{k^2}{4} A_c^2 \langle m^2(t) \rangle}{\frac{k^2}{4} 2 N_0 B_m} \\
 &= \frac{A_c^2 \langle m^2(t) \rangle}{2 N_0 B_m}
 \end{aligned} \tag{7.12}$$

Vogliamo ora esprimere il rapporto SNR in uscita in funzione della potenza di segnale ricevuto all'ingresso del ricevitore. Questa abbiamo visto che è data da:

$$P_{RX} = \frac{A_c^2}{2} (1 + \langle m^2(t) \rangle)$$

da cui possiamo ottenere:

$$A_c^2 = \frac{2 P_{RX}}{1 + \langle m^2(t) \rangle}$$

SNR AM
standard con
ricevitore
coerente

che sostituito nella (7.12) dà finalmente:

$$\boxed{\left(\frac{S}{N} \right)_{out} = \frac{P_{RX}}{N_0 B_m} \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle}} \tag{7.13}$$

Abbiamo dunque ottenuto l'importante formula che dà le prestazioni di un sistema AM standard con ricevitore coerente in presenza di rumore.

Riassumendo, abbiamo ottenuto:

Riassunto

- Sistema di riferimento in banda base:

$$\left(\frac{S}{N}\right)_{BB} = \frac{P_{RX}}{N_0 B_m}$$

- Sistema AM standard:

$$\begin{aligned} \left(\frac{S}{N}\right)_{out} &= \frac{P_{RX}}{N_0 B_m} \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle} \\ &= \left(\frac{S}{N}\right)_{BB} \underbrace{\frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle}}_E \end{aligned}$$

dove si ricorda che E è l'efficienza di modulazione.

Abbiamo quindi trovato che il sistema AM standard a ricezione coerente ha prestazioni peggiori del sistema in banda base per un fattore pari all'efficienza di modulazione (che si ricorda essere sempre inferiore a 1). Inoltre valgono le seguenti relazioni:

$$\begin{aligned} \left(\frac{S}{N}\right)_{in} &= \frac{1}{2} \left(\frac{S}{N}\right)_{BB} \\ \left(\frac{S}{N}\right)_{in} &= \frac{1}{2} \frac{1}{E} \left(\frac{S}{N}\right)_{out} \end{aligned}$$

dove si ricorda che:

$$\left(\frac{S}{N}\right)_{BB} = \frac{P_{RX}}{N_0 B_m}$$

7.5.2 SNR sistemi AM-DSB-SC con ricezione coerente

Passiamo ora a considerare un sistema di trasmissione AM-DSB-SC. Ripetendo esattamente gli stessi calcoli appena fatti, in presenza di rumore e di ricezione coerente si ottiene che il rapporto segnale-rumore in uscita vale semplicemente:

SNR sistemi
AM-DSB-SC
con ricevitore
coerente

$$\boxed{\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} = \left(\frac{S}{N}\right)_{BB}}$$

cioè è pari a quello del sistema di riferimento in banda base.

7.5.3 SNR sistemi AM-SSB con ricezione coerente

Si dimostra che anche per sistemi SSB, sempre con ricezione coerente:

SNR sistemi
SSB con
ricevitore
coerente

$$\left(\frac{S}{N}\right)_{out} = \left(\frac{S}{N}\right)_{in} = \left(\frac{S}{N}\right)_{BB}$$

Dunque i sistemi AM-SSB hanno le stesse prestazioni del formato AM-DSB-SC in termini di SNR, ma occupano metà della banda sul canale. Questo è ottenuto a costo di una maggiore complessità di trasmettitore e ricevitore.

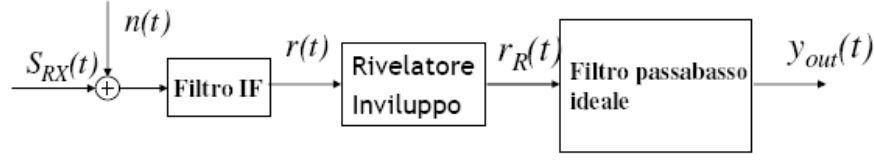


Figura 7.20: Ricevitore non coerente per modulazione AM.

7.5.4 Calcolo prestazioni con ricevitore di inviluppo

Passiamo ora a considerare la ricezione non coerente, cioè la ricezione di un segnale AM standard, senza sovr modulazione, tramite rivelatore di inviluppo. Lo schema del ricevitore è mostrato in figura 7.20. Facendo riferimento alla figura, si ha che:

$$\begin{aligned} s_{RX}(t) &= A_c [1 + m(t)] \cos(2\pi f_c t) \\ n(t) &= x_n(t) \cos(2\pi f_c t) - y_n(t) \sin(2\pi f_c t) \\ r(t) &= s_{RX}(t) + n(t) \end{aligned}$$

Quindi il segnale in ingresso al rivelatore di inviluppo vale:

$$r(t) = \{A_c [1 + m(t)] + x_n(t)\} \cos(2\pi f_c t) - y_n(t) \sin(2\pi f_c t)$$

dove $x_n(t)$ e $y_n(t)$ sono i processi di rumore in fase e quadratura, con una potenza legata alla banda del filtro IF.

Come visto in precedenza, il rivelatore di inviluppo genera un segnale proporzionale al modulo dell'inviluppo complesso di $r(t)$, cioè in questo caso di:

$$\tilde{r}(t) = A_c [1 + m(t)] + x_n(t) + j y_n(t)$$

dunque in uscita dal rivelatore di inviluppo si ha:

$$\begin{aligned} r_R(t) &= k |\tilde{r}(t)| \\ &= k |A_c [1 + m(t)] + x_n(t) + j y_n(t)| \end{aligned} \quad (7.14)$$

Primo caso: SNR_{in} elevato

Se SNR_{in} è elevato, dimostreremo tra poco che si può approssimare il segnale all'uscita del ricevitore di inviluppo come:

$$r_R(t) \approx k A_c [1 + m(t)] + k x_n(t) \quad (7.15)$$

Questo segnale viene filtrato poi dal filtro passa-basso ideale $h(t)$ ottenendo $y_{out}(t)$. Utilizzando le ipotesi consuete, e cioè che $h(t)$ non modifica il segnale di informazione e limita la banda del rumore, si ottiene:

$$y_{out}(t) \approx k A_c [1 + m(t)] + k x_n(t) * h(t)$$

Questa espressione, a parte un fattore $1/2$ assolutamente irrilevante, è analoga all'espressione (7.9) che avevamo trovato a pagina 131 per la ricezione coerente.

In conclusione possiamo dire che, **in presenza di SNR_{in} elevato, il ricevitore di inviluppo ha le stesse prestazioni del ricevitore coerente.**

Dimostrazione della (7.15). A partire dalla (7.14) si può scrivere:

$$\begin{aligned} r_R(t) &= k |A_c[1 + m(t)] + x_n(t) + j y_n(t)| \\ &= k \sqrt{\{A_c[1 + m(t)] + x_n(t)\}^2 + [y_n(t)]^2} \\ &= k A_c \sqrt{[1 + m(t)]^2 + 2[1 + m(t)] \frac{x_n(t)}{A_c} + \left[\frac{x_n(t)}{A_c}\right]^2 + \left[\frac{y_n(t)}{A_c}\right]^2} \end{aligned}$$

Se si ipotizza che il rapporto segnale-rumore in ingresso sia elevato, le quantità:

$$\left[\frac{x_n(t)}{A_c}\right]^2 \text{ e } \left[\frac{y_n(t)}{A_c}\right]^2$$

risultano essere trascurabili rispetto al resto dell'espressione e si possono semplificare. Allora possiamo scrivere:

$$\begin{aligned} r_R(t) &\approx k A_c \sqrt{[1 + m(t)]^2 + 2[1 + m(t)] \frac{x_n(t)}{A_c}} \\ &= k A_c [1 + m(t)] \sqrt{1 + \frac{2}{[1 + m(t)]} \frac{x_n(t)}{A_c}} \end{aligned}$$

Approssimiamo l'espressione sotto radice con lo sviluppo di McLaurin arrestato al primo ordine:

$$\sqrt{1 + x} \approx 1 + \frac{x}{2}$$

da cui troviamo infine:

$$\begin{aligned} r_R(t) &\approx k A_c [1 + m(t)] \left[1 + \frac{1}{1 + m(t)} \frac{x_n(t)}{A_c}\right] \\ &= k A_c [1 + m(t)] + k x_n(t) \end{aligned}$$

Abbiamo dunque dimostrato la (7.15). ■

Graficamente, utilizzando la notazione dei fasori, la situazione è mostrata in figura 8.17. Dalla figura si vede bene che, se n è piccolo rispetto a S_{RX}^2 , la lunghezza (cioè l'inviluppo) del vettore risultante dalla somma vettoriale del segnale più rumore è principalmente influenzata dalla componente di rumore in fase, mentre quella in quadratura genera un effetto trascurabile.

Secondo caso: SNR_{in} basso

Se SNR_{in} è basso, non posso fare l'approssimazione precedente e conviene allora scrivere il rumore gaussiano tramite una notazione in modulo e fase che vedremo nel dettaglio quando parleremo di FM:

$$\begin{aligned} \tilde{n}(t) &= x_n(t) + j y_n(t) \\ &= R_n(t) e^{j\theta(t)} \end{aligned}$$

²I fasori hanno gli stessi nomi delle rispettive funzioni del tempo, senza però indicata la dipendenza dal tempo.

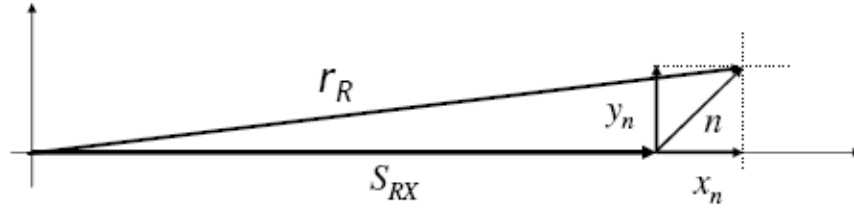


Figura 7.21: Rappresentazione tramite fasori della (7.15). Se n è piccolo rispetto a S_{RX} , la lunghezza del vettore risultante dalla somma vettoriale del segnale più rumore è principalmente influenzata dalla componente di rumore in fase, mentre quella in quadratura genera un effetto trascurabile.

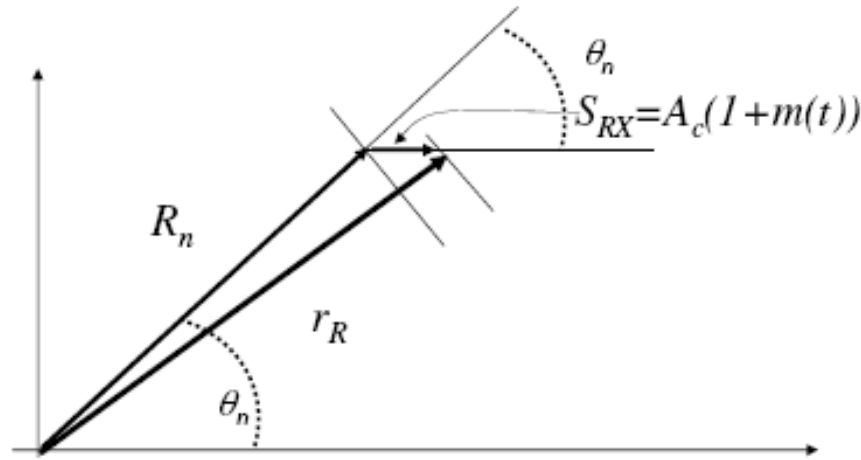


Figura 7.22: Rappresentazione tramite fasori della (7.16). Serve a calcolare l'uscita del rivelatore di inviluppo nel caso di SNR_{in} basso.

Allora il rivelatore di inviluppo fornisce il segnale:

$$r_R(t) = k |A_c [1 + m(t)] + R_n(t) e^{j\theta(t)}| \quad (7.16)$$

L'ipotesi SNR_{in} basso equivale a dire sostanzialmente che $R_n(t)$ è confrontabile, o addirittura superiore ad A_c . Siamo dunque graficamente nella situazione di figura 7.22, e dunque si può scrivere:

$$r_R(t) \approx k \{A_c [1 + m(t)] \cos[\theta_n(t)] + R_n(t)\}$$

Da questa espressione si nota che, oltre a comparire nuovamente un termine di rumore "additivo" dato da $R_n(t)$ (che dunque verrà filtrato dal seguente filtro passa-basso) compare anche un termine moltiplicativo sul segnale, dato da $\cos[\theta_n(t)]$. Questo termine non viene direttamente filtrato dal seguente passa-basso, e peggiora drasticamente le prestazioni.

Conclusioni su prestazioni con ricevitore di inviluppo

In conclusione, la ricezione incoerente (di inviluppo):

- Ha le stesse prestazioni del caso coerente quando il rapporto segnale-rumore di ingresso è sufficientemente elevato;
- Ha prestazioni inaccettabili quando il rapporto segnale-rumore di ingresso scende sotto ad una determinata soglia. In particolare, è richiesto $SNR_{in} \gg 1$ affinché il sistema lavori "sopra soglia".

Ricordiamoci che:

$$\left(\frac{S}{N}\right)_{in} = \frac{1}{2} \frac{1}{E} \left(\frac{S}{N}\right)_{out}$$

7.6 Confronto tra le varie modulazioni di ampiezza

In questo capitolo abbiamo studiato le seguenti modulazioni di ampiezza:

- AM standard;
- AM-DSB-SC (AM a portante soppressa);
- AM-SSB (AM Single Side-Band).

Possiamo a questo punto confrontare tra di loro questi schemi di modulazione. I risultati andranno ancora confrontati con FM e con PCM (che vedremo nei prossimi capitoli). Il confronto può essere fatto da diversi punti di vista. Qui consideriamo i tre più importanti: *efficienza in potenza*, *efficienza in banda* e *facilità di implementazione*. È ovvio che nessuna delle modulazioni è migliore delle altre da tutti i punti di vista. Ognuna avrà i suoi vantaggi e i suoi svantaggi e starà al progettista del sistema di telecomunicazioni scegliere lo schema di modulazione stabilendo il miglior trade-off caso per caso.

Efficienza in potenza. La tabella 10.1 riassume le prestazioni dei sistemi AM visti in questo capitolo. Ecco le varie grandezze che compaiono nella tabella:

- B_m è la banda occupata dal segnale modulante $m(t)$;
- Si ipotizza che il ricevitore disponga di un filtro passa-basso ideale anch'esso con banda B_m ;
- B_T è la banda occupata dal segnale modulato;
- $\frac{P_{RX}}{N_0 B_m}$ è $(S/N)_{BB}$, cioè il rapporto segnale-rumore del sistema equivalente in banda base;
- $\left(\frac{S}{N}\right)_{out}$ è il rapporto segnale-rumore in uscita dal ricevitore.

Un criterio per valutare l'efficienza in potenza delle varie modulazioni è quello di confrontare le rispettive curve di SNR. Con "curva di SNR" intendiamo un grafico che riporti l'andamento di $(S/N)_{out}$ in funzione di $(S/N)_{BB}$. Curve di SNR

Seguono tre figure che riassumono le prestazioni dei vari sistemi AM. Nelle figure non è indicata la modulazione AM-SSB. questa ha la stessa curva di SNR di AM-DSB-SC.

- La figura 7.23 mostra le prestazioni dei vari sistemi AM. Dalla figura si vede bene come i sistemi AM standard con ricezione incoerente abbiano un comportamento a soglia.

Tipo di trasmettitore	Tipo di ricevitore	Banda B_T	$\left(\frac{S}{N}\right)_{out}$
AM standard	inviluppo o coerente	$2B_m$	$\frac{P_{RX}}{N_0 B_m} \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle}$
AM-DSB-SC	solo coerente	$2B_m$	$\frac{P_{RX}}{N_0 B_m}$
AM-SSB	solo coerente	B_m	$\frac{P_{RX}}{N_0 B_m}$

Tabella 7.1: Riassunto prestazioni dei vari sistemi AM.

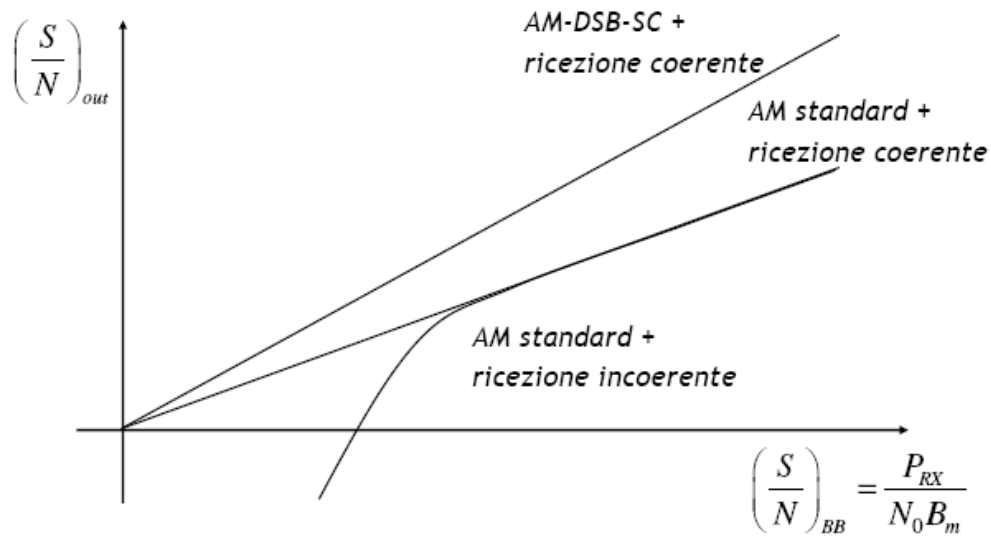


Figura 7.23: Riassunto prestazioni modulazioni AM.

- La figura 7.24 mostra lo stesso grafico ma con SNR espressi in dB.
- La figura 7.25 mostra la penalità di AM standard rispetto a AM-DSB-SC. Infatti, fissato un certo $(S/N)_{out}$ e lavorando con la stessa potenza ricevuta e con la stessa potenza di rumore, AM standard ha bisogno di una potenza ricevuta maggiore per ottenerlo.

Dai grafici appena esaminati si vede che AM-DSB-SC e AM-SSB sono quelle con la più alta efficienza in potenza. AM standard ha una efficienza in potenza minore perché "spreca" parte della potenza per trasmettere la portante.

Efficienza in banda. Lo schema di modulazione che utilizza meglio la banda è AM-SSB. Un segnale SSB ha una occupazione di banda $B_T = B_m$. Per questo motivo questo sistema è ampiamente utilizzato in applicazioni critiche per l'utilizzo della banda, come trasmissione di voce su rete wireless e sistemi di comunicazione in aree congestionate.

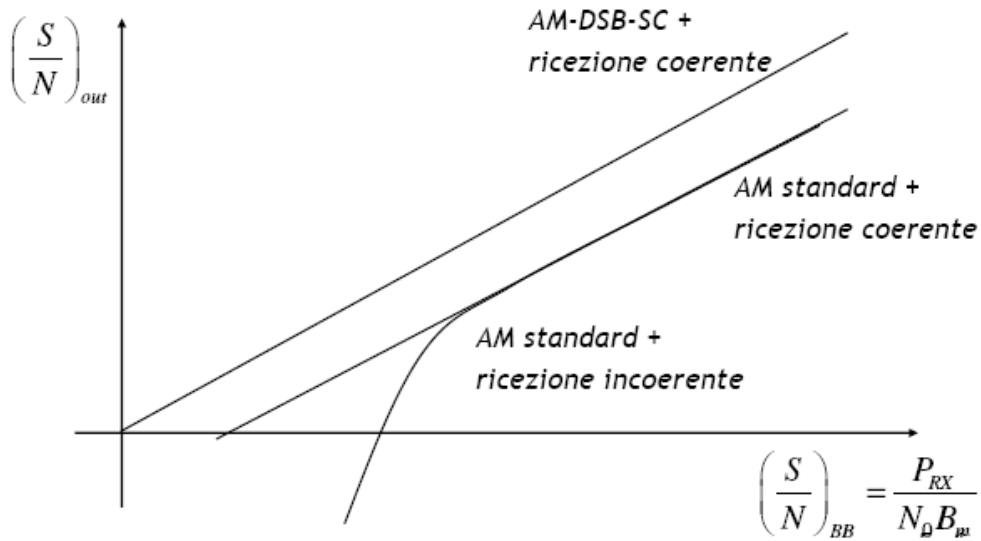


Figura 7.24: Riassunto prestazioni modulazioni AM con SNR espressi in dB.

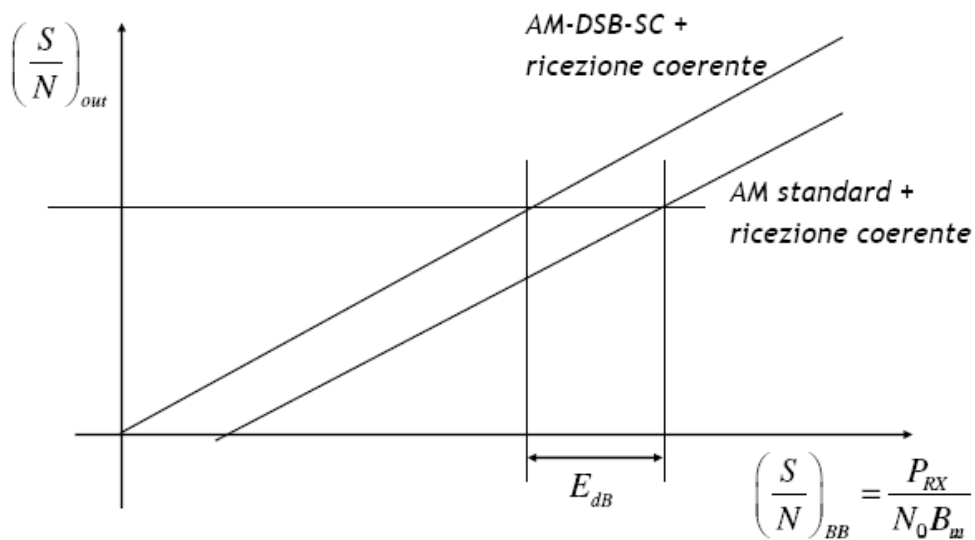


Figura 7.25: Penalità di AM standard rispetto a AM-DSB-SC.

AM standard e AM-DSB-SC hanno un'occupazione di banda doppia rispetto a AM-SSB. Occupano una banda $B_T = 2 B_m$. Per questo motivo non vengono utilizzate in applicazioni critiche dal punto di vista dell'occupazione di banda.

Facilità di implementazione. La struttura di ricevitore più semplice è quella di AM standard, che utilizza un rivelatore di involuppo.

AM-DSB-SC e AM-SSB hanno bisogno invece di un ricevitore coerente che è più difficile da implementare. AM-SSB è comunque quella con la complessità maggiore, per via dell'elevata complessità del trasmettitore.

Come si è visto, ogni modulazione ha i suoi pregi e i suoi difetti. Il compito del progettista è scegliere la modulazione più adatta per soddisfare le specifiche del sistema

che si intende progettare.

7.7 Riassunto formule principali su AM

Si riportano in questa appendice le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi delle esercitazioni.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- AM standard, densità spettrale segnale modulato:

$$G_s(f) = \frac{A_c^2}{4} [\delta(f - f_c) + G_m(f - f_c) + \delta(f + f_c) + G_m(f + f_c)]$$

- Efficienza di modulazione:

$$E = \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle}$$

- Prestazioni AM standard con ricevitore coerente ideale:

$$\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} \frac{\langle m^2(t) \rangle}{1 + \langle m^2(t) \rangle} = \left(\frac{S}{N}\right)_{BB} \cdot E$$

- Prestazioni AM-DSB-SC con ricevitore coerente ideale:

$$\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} = \left(\frac{S}{N}\right)_{BB}$$

Capitolo 8

Modulazioni angolari: FM e PM

8.1 Introduzione

Questo capitolo tratta le modulazioni angolari. Verranno considerate le modulazioni di frequenza (*Frequency Modulation*, FM) e le modulazioni di fase (*Phase Modulation*, PM). Si tratta di modulazioni con caratteristiche completamente diverse dalle modulazioni di ampiezza.

Analizzeremo i seguenti punti sulle modulazioni angolari:

1. Definizioni e caratteristiche dei segnali modulati;
2. Spettri;
3. Implementazione di modulatori e demodulatori;
4. Calcolo del rapporto segnale-rumore;
5. Confronto con altri sistemi di modulazione precedentemente studiati.

Sommario
del capitolo

8.2 Rappresentazione di segnali FM e PM

In questo paragrafo definiamo le modulazioni FM e PM e introduciamo alcune grandezze fondamentali che le caratterizzano.

Definizioni di FM e PM

Le modulazioni di fase (PM) e di frequenza (FM) operano sul termine di fase dell'involuppo complesso, cioè lavorano su un involuppo complesso del tipo:

$$\tilde{s}(t) = A_c e^{j\theta(t)} \quad (8.1)$$

Ricordando che il segnale reale può essere ottenuto a partire dall'involuppo complesso in questo modo:

$$s(t) = \operatorname{Re} \left\{ \tilde{s}(t) e^{j2\pi f_c t} \right\}$$

si trova che il segnale reale trasmesso è del tipo (vedi figura 8.1):

$$s(t) = A_c \cos[2\pi f_c t + \theta(t)]$$

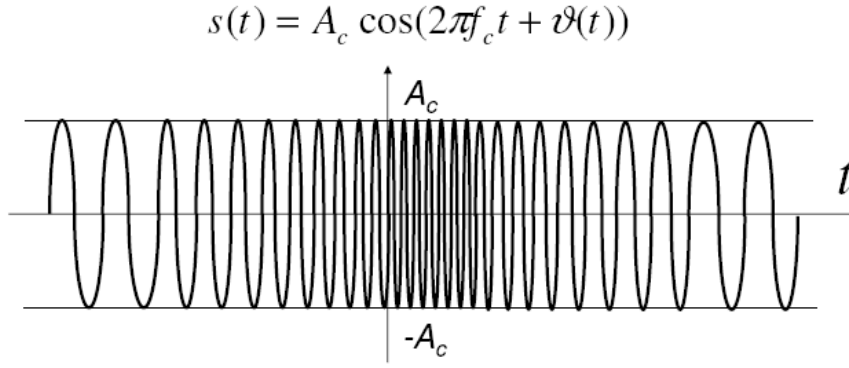


Figura 8.1: Esempio di andamento nel dominio del tempo di segnale con modulazione angolare. Si noti come il segnale abbia inviluppo costante.

dove f_c è la *frequenza di portante* e $\theta(t)$ è legato al segnale di informazione da trasmettere.

Dalla (8.1) si vede che si tratta di **modulazioni ad inviluppo costante**, in quanto abbiamo:

$$|\tilde{s}(t)| = A_c = \text{costante}$$

Si tratta cioè di una situazione diametralmente opposta a quella delle modulazioni di ampiezza. **Si parla di modulazioni angolari in quanto l'informazione $m(t)$ è affidata al termine di fase $\theta(t)$ del segnale modulato, cioè al suo "angolo".**

Modulazione PM Nelle modulazioni PM si impone la seguente relazione:

$$\theta(t) = D_p m(t)$$

dove $m(t)$ è il segnale modulante e D_p è una costante di proporzionalità detta *sensibilità del modulatore di fase*. L'unità di misura di D_p è [rad/V] (ovviamente se $m(t)$ è espresso in [V]).

Modulazione FM Nelle modulazioni FM si impone la seguente relazione:

$$\theta(t) = D_f \int_{-\infty}^t m(\tau) d\tau \quad (8.2)$$

dove $m(t)$ è il segnale modulante e D_f è una costante di proporzionalità detta *sensibilità del modulatore di frequenza*. L'unità di misura di D_f è [rad/V/s] (ovviamente se $m(t)$ è espresso in [V]). Il significato esatto della (8.2) verrà spiegato tra poco.

Frequenza istantanea

Definizione di frequenza istantanea Introduciamo ora il concetto di *frequenza istantanea*, che sarà molto utile in seguito. Si consideri il seguente segnale sinusoidale reale:

$$s(t) = R \cos[\Psi(t)]$$

dove $\Psi(t)$ viene detta *fase istantanea* del segnale sinusoidale, ed è una generica funzione reale del tempo. Il corrispondente segnale analitico vale:

$$\overset{\circ}{s}(t) = R e^{j\Psi(t)}$$

Si definisce la **frequenza istantanea** come:

$$f_i(t) \triangleq \frac{1}{2\pi} \frac{d}{dt} \Psi(t)$$

Per comprendere meglio il significato di questa definizione, calcoliamo ora la frequenza istantanea di un segnale sinusoidale e poi di un segnale modulato angolarmente.

Si consideri un segnale sinusoidale:

$$s(t) = A_c \cos(2\pi f_c t + \theta)$$

$\Psi(t)$ in questo caso vale:

$$\Psi(t) = 2\pi f_c t + \theta$$

e quindi la corrispondente frequenza istantanea è data da:

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi} \frac{d}{dt} \Psi(t) \\ &= \frac{1}{2\pi} \frac{d}{dt} (2\pi f_c t + \theta) \\ &= f_c \end{aligned}$$

Cioè in questo caso in concetto di frequenza istantanea coincide con il concetto tradizionale di frequenza di un segnale sinusoidale.

Si consideri ora un segnale modulato angolarmente (sia PM che FM):

$$s(t) = A_c \cos[2\pi f_c t + \theta(t)]$$

In questo caso $\Psi(t)$ vale:

$$\Psi(t) = 2\pi f_c t + \theta(t)$$

Calcoliamo la frequenza istantanea:

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi} \frac{d}{dt} \Psi(t) \\ &= f_c + \frac{1}{2\pi} \frac{d\theta(t)}{dt} \end{aligned}$$

In questo caso chiaramente la frequenza istantanea risulta essere una funzione del tempo.

Vediamo ora di capire in che senso la definizione di FM (8.2) corrisponde a una modulazione di frequenza. Per FM posso scrivere:

$$\theta(t) = D_f \int_{-\infty}^t m(\tau) d\tau$$

Calcolo
frequenza
istantanea
segnale
sinusoidale

Calcolo
frequenza
istantanea
modulazione
angolare

Calcolo
frequenza
istantanea
segnale FM

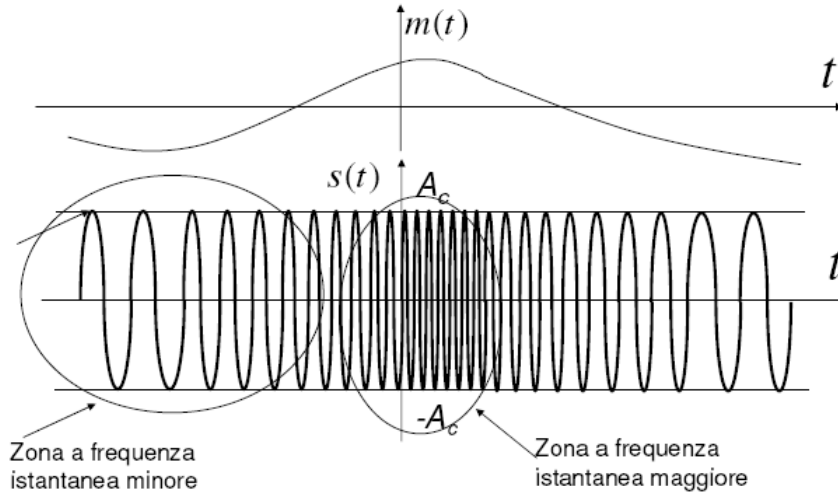


Figura 8.2: Interpretazione della (8.3) nel dominio del tempo. La frequenza istantanea è proporzionale al segnale modulante $m(t)$.

$\Psi(t)$ vale quindi:

$$\Psi(t) = 2\pi f_c t + D_f \int_{-\infty}^t m(\tau) d\tau$$

Quindi la frequenza istantanea vale:

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi} \frac{d}{dt} \left(2\pi f_c t + D_f \int_{-\infty}^t m(\tau) d\tau \right) \\ &= f_c + \frac{1}{2\pi} D_f m(t) \end{aligned} \quad (8.3)$$

Risulta dunque che **per una modulazione FM la frequenza istantanea è proporzionale al segnale modulante $m(t)$** . Per questo motivo la modulazione viene chiamata *di frequenza*. La figura 8.2 mostra un'interpretazione di questo risultato nel dominio del tempo.

Legame tra FM e PM

Le modulazioni FM e PM sono strettamente legate tra di loro. Infatti, riassumendo quanto scritto prima, ho che nei due casi il segnale trasmesso vale:

$$\begin{aligned} s(t) &= A_c \cos[2\pi f_c t + \theta(t)] \\ &= \begin{cases} A_c \cos[2\pi f_c t + D_p m(t)] & \text{(PM)} \\ A_c \cos\left[2\pi f_c t + D_f \int_{-\infty}^t m(\tau) d\tau\right] & \text{(FM)} \end{cases} \end{aligned}$$

da cui si vede bene che:

- Un segnale modulato FM può essere ottenuto da un modulatore di fase aggiungendo al suo ingresso un integratore;
- Un segnale modulato PM può essere generato da un modulatore di frequenza inserendo al suo ingresso un derivatore.

Per questo motivo le due modulazioni presentano aspetti molto simili e quindi saranno analizzate insieme. Alcuni aspetti saranno discussi solo per FM ma i risultati per PM sono sempre analoghi.

Definizioni di altri parametri utili

Vediamo ora le definizioni di altre grandezze che torneranno utili in seguito.

Si definisce come **deviazione di frequenza** rispetto alla frequenza centrale di portante f_c :

$$f_d(t) \triangleq f_i(t) - f_c = \frac{1}{2\pi} \frac{d\theta}{dt}$$

Si definisce infine la **deviazione di frequenza di picco** come:

$$\Delta F \triangleq \max \{f_d(t)\}$$

Nel caso FM, dalle definizioni appena date e dalla (8.3), risulta che:

$$\begin{aligned} \Delta F &= \frac{1}{2\pi} D_f \max \{m(t)\} \\ &= \frac{1}{2\pi} D_f V_p \end{aligned} \quad (8.4)$$

dove abbiamo definito:

$$V_p \triangleq \max \{m(t)\}$$

Anche nel caso di modulazione di fase è utile definire una **deviazione di fase di picco** come:

$$\Delta\theta \triangleq \max \{\theta(t)\} = D_p V_p$$

dove abbiamo definito:

$$V_p \triangleq \max \{m(t)\}$$

A partire dai parametri appena introdotti, è utile definire (sia per PM che per FM) due parametri molto importanti detti *indici di modulazione*. Il loro significato (e dunque anche la loro definizione) apparirà più chiaro nei prossimi paragrafi in cui si parlerà di spettri e SNR. Indici di modulazione

Indice di modulazione di fase (PM):

$$\beta_p \triangleq \Delta\theta = D_p V_p \quad (8.5)$$

Indice di modulazione di frequenza (FM):

$$\beta_f \triangleq \frac{\Delta F}{B_m} \quad (8.6)$$

dove B_m è la banda del segnale modulante $m(t)$.

8.3 Spettri di modulazioni angolari

Il calcolo degli spettri per le modulazioni angolari è in generale molto complicato. In particolare, risulta difficile legare in maniera esatta il segnale modulante, l'indice di modulazione e la corrispondente occupazione spettrale del segnale modulato.

Formulazione
problema
calcolo
spettro

Formuliamo il problema utilizzando l'involuppo complesso. In generale, dato:

$$s(t) = \text{Re} \left\{ \tilde{s}(t) e^{j2\pi f_c t} \right\}$$

sappiamo che, per le proprietà dell'involuppo complesso (vedi capitolo su segnale analitico e involuppo complesso), la risultante trasformata di Fourier è data da:

$$S(f) = \frac{1}{2} \left\{ \tilde{S}(f - f_c) + \tilde{S}^*(-f - f_c) \right\} \quad (8.7)$$

Il problema è calcolare:

$$\begin{aligned} \tilde{S}(f) &= F \{ \tilde{s}(t) \} \\ &= F \left\{ A_c e^{j\theta(t)} \right\} \end{aligned}$$

Se per AM è stato facile scrivere $S(f)$ (spettro del segnale modulato) in funzione di $M(f)$ (spettro del segnale modulante), la situazione è decisamente più complicata per le modulazioni angolari. Questo poiché non esiste una proprietà della trasformata di Fourier che permetta di "gestire" in generale espressioni del tipo:

$$F \left\{ A_c e^{j\theta(t)} \right\}$$

con $\theta(t)$ generica funzione del tempo.

Quindi, per studiare le caratteristiche spettrali dei segnali angolari, si parte dallo studio di casi particolari, in cui il segnale modulante è molto semplice e si riesce a calcolarne lo spettro in forma chiusa o mediante approssimazioni. I risultati che si ottengono sono poi estesi a segnali modulanti più complicati.

Si considererà dunque ora un caso particolare, corrispondente ad un segnale modulante di tipo sinusoidale.

8.3.1 Spettro di segnale angolare con modulante sinusoidale

Consideriamo il caso in cui il segnale modulante sia sinusoidale:

$$\theta(t) = \beta \sin(2\pi f_m t)$$

Il segnale modulato è quindi:

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)] \quad (8.8)$$

Le figure 8.3 e 8.4 rappresentano l'andamento nel tempo rispettivamente di $\theta(t)$ e di $s(t)$.

La notazione introdotta $\theta(t) = \beta \sin(2\pi f_m t)$ permette di studiare contemporaneamente:

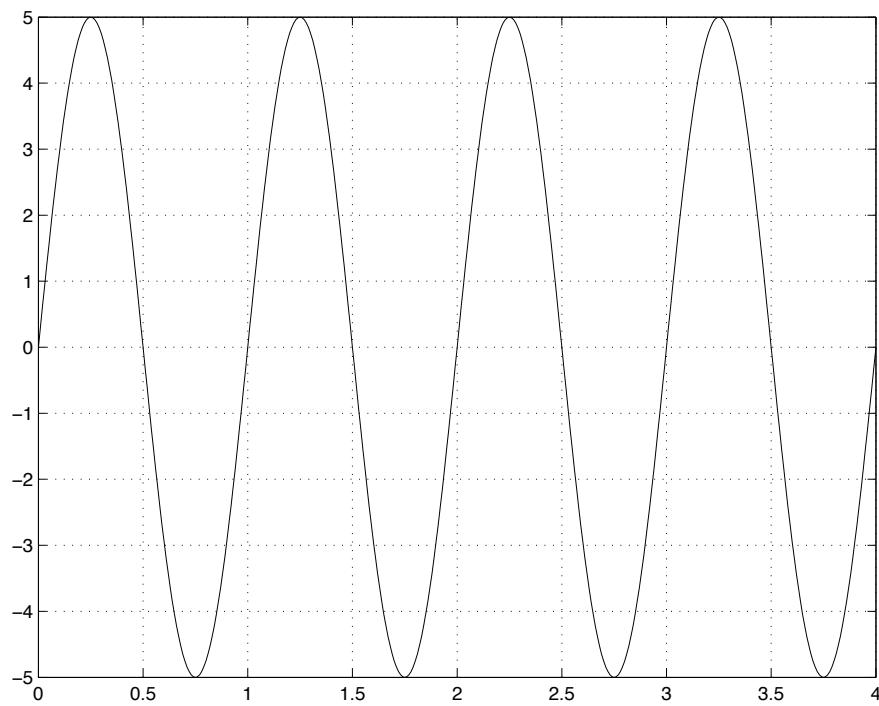


Figura 8.3: Modulante sinusoidale $\theta(t)$.

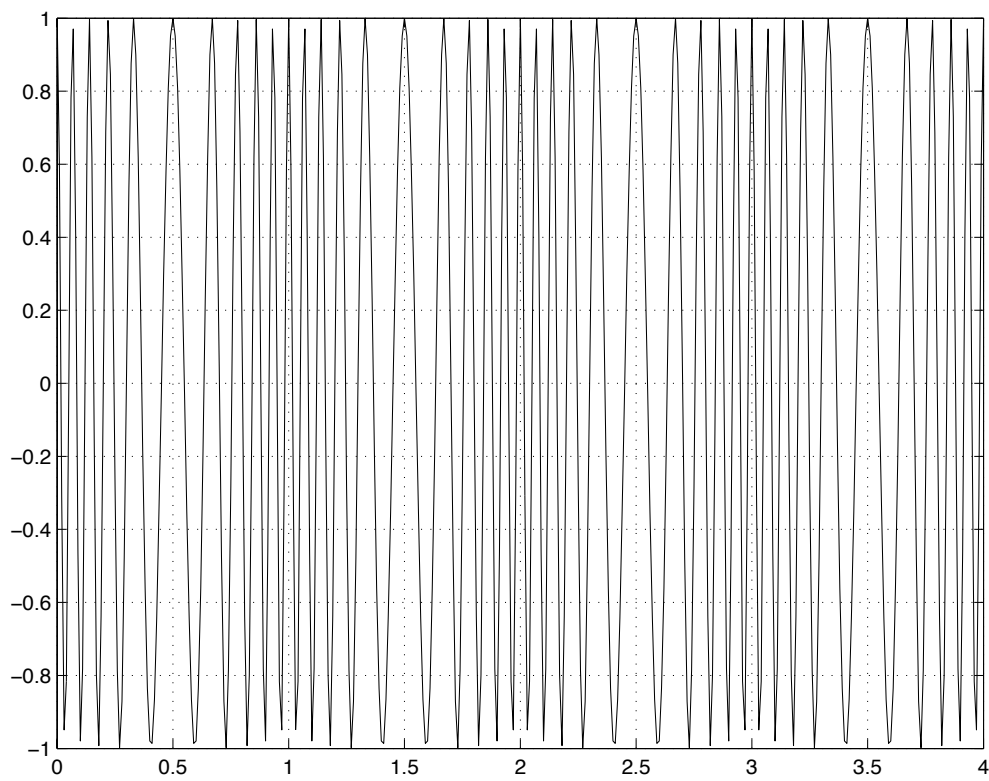


Figura 8.4: Segnale $s(t)$ modulato FM/PM con modulante sinusoidale.

- Una modulazione PM con segnale modulante:

$$m(t) = \sin(2\pi f_m t)$$

e con indice di modulazione di fase:

$$\beta_p = \beta$$

- Oppure: una modulazione FM con segnale modulante:

$$\begin{aligned} m(t) &= \frac{d\theta(t)}{dt} \\ &= \frac{d}{dt}[\beta \sin(2\pi f_m t)] \\ &= \beta 2\pi f_m \cos(2\pi f_m t) \end{aligned}$$

e con deviazione di frequenza istantanea:

$$\begin{aligned} f_d(t) &= \frac{1}{2\pi} \frac{d\theta(t)}{dt} \\ &= \frac{1}{2\pi} \beta 2\pi f_m \cos(2\pi f_m t) \\ &= \beta f_m \cos(2\pi f_m t) \end{aligned}$$

La deviazione di frequenza di picco vale:

$$\Delta F = \beta f_m$$

Quindi, ricordando che $B_m = f_m$ si trova l'indice di modulazione di frequenza:

$$\begin{aligned} \beta_f &= \frac{\Delta F}{B_m} \\ &= \frac{\beta f_m}{f_m} \\ &= \beta \end{aligned}$$

La (8.8) va quindi bene sia per analizzare una modulazione PM che FM. Nei due casi cambia solo leggermente il segnale modulante, ma si tratta sempre di un segnale sinusoidale con frequenza f_m .

Con alcuni passaggi matematici si ottiene che lo spettro dell'involuppo complesso di $s(t)$ vale:

$$\tilde{S}(f) = A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) \delta(f - n f_m) \quad (8.9)$$

Spettro del-
l'involuppo
complesso di
segnale
FM/PM con
modulante
sinusoidale

dove $J_n(\beta)$ è la *funzione di Bessel* del primo tipo di ordine n . Lo spettro del segnale dipende quindi dalle funzioni di Bessel, che vedremo più in dettaglio tra poco.

Si noti che si tratta di uno spettro a righe.

Dimostrazione della (8.9)

L'involuppo complesso:

$$\begin{aligned} \tilde{s}(t) &= A_c e^{j\theta(t)} \\ &= A_c e^{j\beta \sin(2\pi f_m t)} \end{aligned}$$

è un segnale periodico di periodo:

$$T_m = \frac{1}{f_m}$$

e dunque (in quanto segnale periodico) sicuramente può essere espresso in serie di Fourier:

$$\begin{aligned} \tilde{s}(t) &= \sum_{n=-\infty}^{+\infty} c_n e^{j2\pi n f \frac{t}{T_m}} \\ &= \sum_{n=-\infty}^{+\infty} c_n e^{j2\pi n f_m t} \end{aligned} \quad (8.10)$$

dove i coefficienti della serie di Fourier sono dati da:

$$\begin{aligned} c_n &= \frac{1}{T_m} \int_{-\frac{T_m}{2}}^{+\frac{T_m}{2}} \tilde{s}(t) e^{-j2\pi n f_m t} dt \\ &= \frac{A_c}{T_m} \int_{-\frac{T_m}{2}}^{+\frac{T_m}{2}} e^{j\beta \sin(2\pi f_m t)} e^{-j2\pi n f_m t} dt \end{aligned}$$

Con un cambio di variabile ($u = 2\pi f_m t$) si può passare a:

$$\begin{aligned} c_n &= \frac{A_c}{2\pi} \int_{-\pi}^{+\pi} e^{j\beta \sin(u)} e^{-jnu} dt \\ &= A_c \underbrace{\left(\frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{j[\beta \sin(u) - nu]} dt \right)}_{J_n(\beta)} \end{aligned} \quad (8.11)$$

L'espressione tra parentesi tonde è esattamente uguale alla definizione delle funzioni di Bessel del primo tipo, di ordine n , calcolate in β :

$$J_n(\beta) \triangleq \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{j[\beta \sin(u) - nu]} dt \quad (8.12)$$

Si tratta di un integrale che non può essere risolto in forma chiusa. Mettendo insieme la (8.11) e la (8.12) trovo che:

$$c_n = A_c J_n(\beta)$$

che, sostituito nella (8.10) dà:

$$\tilde{s}(t) = A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) e^{j2\pi n f_m t}$$

A questo punto è semplice calcolare $\tilde{S}(f)$:

$$\begin{aligned} \tilde{S}(f) &= F\{\tilde{s}(t)\} \\ &= F\left\{A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) e^{j2\pi n f_m t}\right\} \\ &= A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) \delta(f - n f_m) \end{aligned} \quad (8.13)$$

Abbiamo così completato la dimostrazione della (8.9). Si noti come, essendo $\tilde{s}(t)$ un segnale periodico, il suo spettro sia a righe (come si è visto nel corso di Teoria dei Segnali).

◇

Commenti sulle funzioni di Bessel

Nella (8.12) abbiamo introdotto le funzioni di Bessel, che probabilmente non erano state incontrate nei corsi precedenti a questo. Vediamo di aggiungere qualche commento in proposito.

Le funzioni di Bessel sono una famiglia di "funzioni speciali". Per funzioni speciali si intendono funzioni che non hanno un'espressione diretta in termini di funzioni "standard" quali $\sin()$, $\cos()$, $\exp()$, etc. Si tratta solitamente di funzioni definite tramite un'espressione integrale, in cui l'integrale non è risolvibile in forma chiusa ma deve essere calcolato mediante metodi numerici. Solitamente si utilizzano:

- Opportune tavole o grafici;
- Sviluppi in serie;
- Software di calcolo (Matlab, Maple, Mathematica, etc.).

In Matlab, si utilizza la funzione:

$$\text{besselj}(nu, x)$$

A titolo di curiosità:

- Le funzioni di Bessel furono introdotte per studiare problemi di vibrazione/propagazione in strutture a simmetria cilindrica. Ad esempio i modi fondamentali di vibrazione di una membrana circolare sono legate alle funzioni di Bessel.

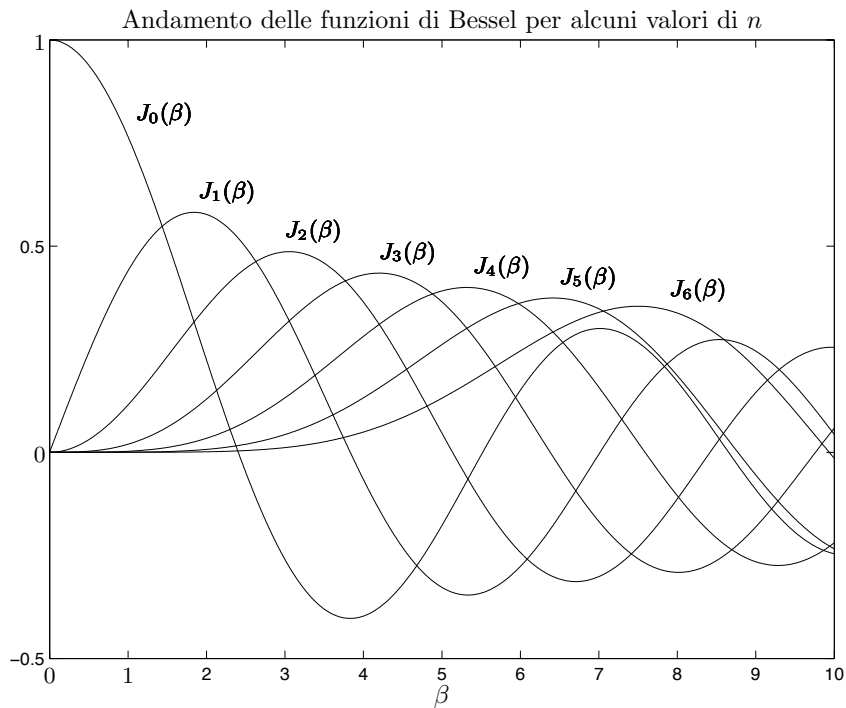


Figura 8.5: Andamento delle funzioni di Bessel per alcuni valori di n .

- Il logo di Matlab è una rappresentazione tridimensionale di funzioni di Bessel.

La figura 8.5 mostra l'andamento delle funzioni di Bessel $J_n(\beta)$ per diversi valori di n . Il grafico è stato generato tramite il semplice programma Matlab:

```
x=(0:0.01:10);
y = besselj(0:6, x');
for i=1:7
    plot(x, y(:, i));
hold on;
end
```

Ecco alcune proprietà delle funzioni di Bessel:

1. Sono "dispari" in n , cioè:

$$J_{-n}(\beta) = (-1)^n \cdot J_n(\beta)$$

2. Fissato un certo β , $J_n(\beta)$ decresce al crescere di n , inoltre:

$$\lim_{n \rightarrow +\infty} J_n(\beta) = 0$$

Caratteristiche dello spettro

Tornando al segnale modulato, la figura 8.6 mostra lo spettro $\tilde{S}(f)$ del suo segnale analitico.

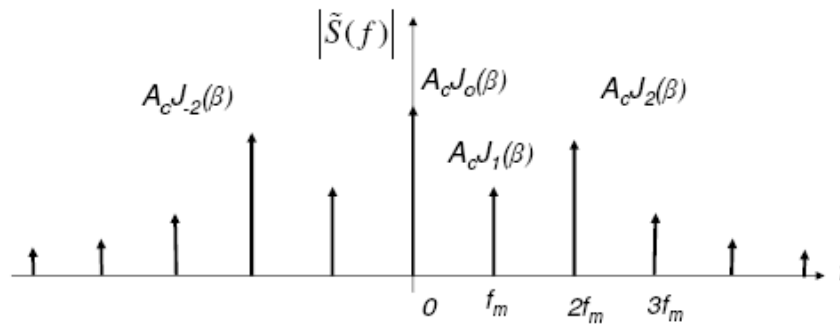


Figura 8.6: Trasformata di Fourier del segnale analitico di un segnale sinusoidale modulato FM/PM.

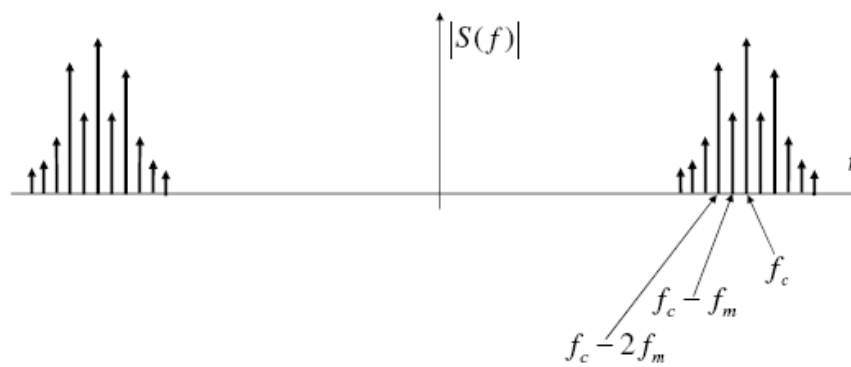


Figura 8.7: Trasformata di Fourier di un segnale sinusoidale modulato FM/PM. Si tratta di uno spettro in banda traslata centrato attorno alla frequenza f_c .

Lo spettro $S(f)$ del segnale reale si calcola poi con la (8.7). La figura 8.7 mostra $S(f)$. Si noti come si tratti di uno spettro in banda traslata centrato attorno alla frequenza f_c .

Occorre fare diversi commenti importanti su questi spettri:

Osservazioni
sugli spettri

- Lo spettro risultante contiene varie righe spaziate di f_m ;
- L'ampiezza della n -esima riga dipende da $J_n(\beta)$;
- L'altezza delle righe dipende dunque fortemente dall'indice di modulazione;
- All'aumentare di β , aumenta la banda dove le righe sono significative;
- Dato un certo β , al crescere di n le corrispondenti righe spettrali diminuiscono di ampiezza, tendendo a zero per n tendente a infinito. Si tratta di uno spettro che *matematicamente*¹ ha estensione illimitata.

Si nota quindi che **per le modulazioni angolari esiste un parametro (l'indice di modulazione) che permette di variare la banda del segnale modulato, dato lo stesso segnale modulante**. Questa è una caratteristica che differenzia fortemente le modulazioni angolari da quelle di ampiezza. Infatti nelle modulazioni di ampiezza lo spettro del segnale modulato è direttamente legato a quello del segnale modulante, senza nessun "grado di libertà". Ritorneremo su questa caratteristica quando parleremo di SNR.

Importante

La figura 8.8 mostra gli spettri di un segnale sinusoidale modulato PM con quattro diversi indici di modulazione. Il segnale modulante è sempre lo stesso. Dalla figura si può verificare facilmente che la banda del segnale modulato aumenta con β_p . I quattro grafici sono stati ottenuti con un programma Matlab analogo a quello sul sito del corso. Chi volesse visualizzare qualche spettro di modulazione angolare, sia con modulanti sinusoidali che generici, può scaricare dal sito del corso gli esempi Matlab "Esempi di spettri relativi a modulazioni analogiche PM o FM" e farli girare. Sarebbe un lavoro istruttivo.

Esempi
Matlab
spettri
FM/PM

8.3.2 Occupazione spettrale di segnali FM/PM: regola di Carson

Nel caso esaminato nel paragrafo precedente di segnale modulato FM/PM con modulante sinusoidale, la banda occupata (al 98% di energia) risulta essere circa pari a:

$$B_T = 2(\beta + 1)f_m \quad (8.14)$$

dove β è l'indice di modulazione (indifferentemente di frequenza o di fase). La (8.14) è nota con il nome di *regola di Carson*.

Regola di
Carson per
segnali
FM/PM con
modulante
sinusoidale

La regola di Carson può essere usata in prima approssimazione anche per un generico segnale modulante, intendendola come:

$$B_T \approx 2(\beta + 1)B_m$$

Regola di
Carson per
segnali
FM/PM con
modulante
generico

¹Si ricorda che in questo corso non ci interessa la banda assoluta dei segnali.

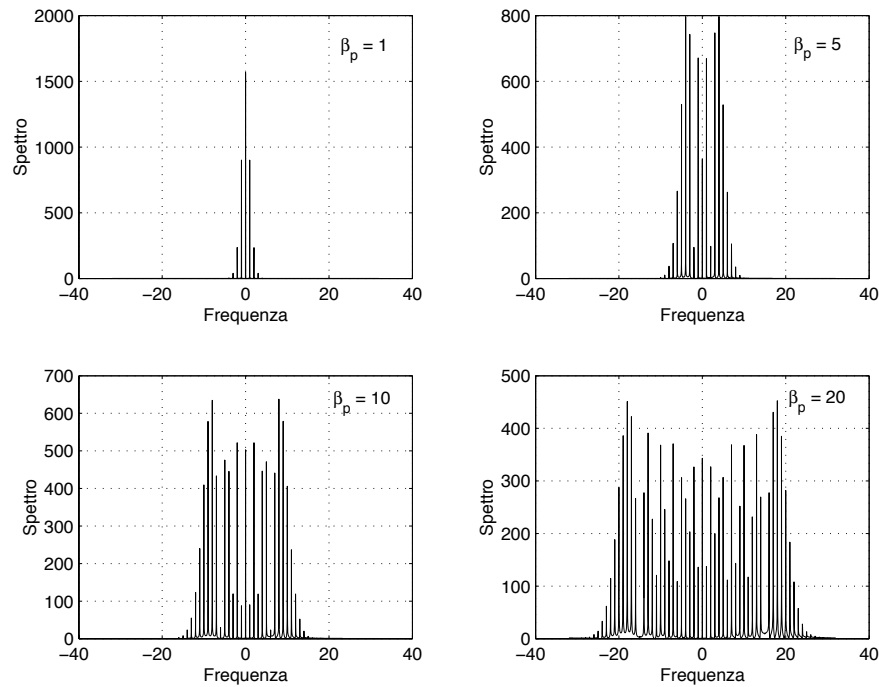


Figura 8.8: Spettri di segnale sinusoidale modulato PM con quattro diversi indici di modulazione. Il segnale modulante è sempre lo stesso. Si noti come la banda del segnale modulato aumenti aumentando β_p .

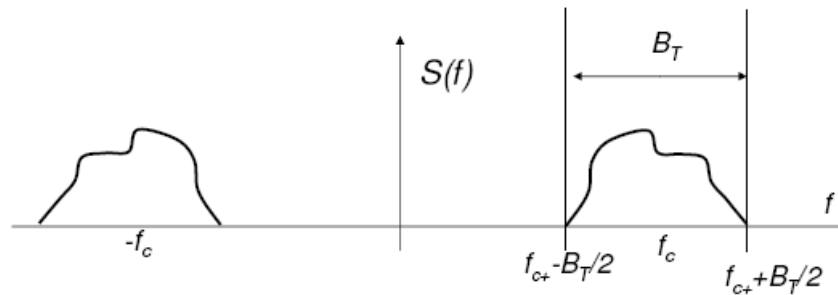


Figura 8.9: Rappresentazione dell'estensione di banda di un segnale modulato FM/PM. La banda B_T può essere calcolata usando la regola di Carson.

dove B_m è la banda del segnale modulante. La figura 8.9 rappresenta l'estensione di banda di un segnale modulato FM/PM.

Dalla regola di Carson si vede che la banda occupata dal segnale modulato cresce al crescere dell'indice di modulazione.

Cerchiamo ora di commentare il significato pratico della regola di Carson. Concentriamoci sul caso FM, analizzando due casi estremi: alto indice di modulazione e basso indice di modulazione.

Alto indice di modulazione. Se si dimensiona il sistema in modo tale che $\beta_f \gg 1$,

allora:

$$\begin{aligned} B_T &= 2(\beta_f + 1)B_m \\ &\approx 2\beta_f B_m \end{aligned}$$

ma dalla (8.6) so che l'indice di modulazione di frequenza vale:

$$\beta_f = \frac{\Delta F}{B_m}$$

Quindi in definitiva:

$$B_T \approx 2\Delta F$$

Banda
segnale FM
con alto
indice di
modulazione

Questo risultato ha la seguente interpretazione pratica:

- La banda occupata da modulazioni FM, per alti indici di modulazione, è pari a 2 volte la deviazione di frequenza di picco;
- Consideriamo che la frequenza istantanea varia nell'intervallo $f_c \pm \Delta F$. Allora è intuibile che il range di frequenze occupate sia pari a $B_T \approx 2\Delta F$.

Basso indice di modulazione. Consideriamo ora il caso opposto, e dunque supponiamo che $\theta(t)$ sia piccolo per qualsiasi t . Possiamo allora utilizzare lo sviluppo di McLaurin arrestato al primo ordine $e^x \approx 1 + x$ e scrivere:

$$\begin{aligned} \tilde{s}(t) &= A_c e^{j\theta(t)} \\ &\approx A_c [1 + j\theta(t)] \end{aligned} \quad (8.15)$$

Allora il segnale reale vale:

$$\begin{aligned} s(t) &= \operatorname{Re} \left\{ \tilde{s}(t) e^{j2\pi f_c t} \right\} \\ &\approx A_c \cos(2\pi f_c t) - A_c \theta(t) \sin(2\pi f_c t) \end{aligned} \quad (8.16)$$

Si tratta in qualche modo di una situazione simile a AM standard, salvo il fatto che il segnale modulato è sfasato di 90° rispetto alla portante (cioè la riga di portante è in coseno, la modulante in seno).

Spettro
segnale FM
con basso
indice di
modulazione

Il calcolo dello spettro è semplice in questo caso, e porta a:

$$S(f) \approx \frac{A_c}{2} \{ \delta(f - f_c) + \delta(f + f_c) + j[\theta(f - f_c) - \theta(f + f_c)] \}$$

Banda
segnale FM
con basso
indice di
modulazione

Quindi in questo caso la banda del segnale modulato è circa uguale alla banda di un segnale AM:

$$B_T \approx 2B_m$$

Come vedremo quando parleremo di rapporto segnale-rumore, l'uso di un basso indice di modulazione non fornisce una migliore immunità al rumore rispetto ai sistemi AM. Per questo motivo raramente viene usato nei sistemi di trasmissione. Vedremo però tra poco che segnali a basso indice di modulazione possono essere usati negli stadi intermedi per la generazione di segnali ad alto indice.

8.3.3 Altro risultato

Esiste un ulteriore risultato, che accenniamo brevemente, relativamente alle densità spettrali di potenza di segnali FM ad alto indice. Per una modulazione di frequenza ad alto indice, partendo da un segnale modulante $m(t)$ inteso come processo casuale si ottiene una densità spettrale di potenza pari a:

$$G(f) = \frac{\pi A_c^2}{2D_f} \left[\underline{f_m} \left(\frac{2\pi}{D_f} (f - f_c) \right) + \underline{f_m} \left(\frac{2\pi}{D_f} (-f - f_c) \right) \right]$$

dove $\underline{f_m}(x)$ è la *densità di probabilità* di $m(t)$.

Spiegazione intuitiva: quando $m(t)$ assume un certo valore istantaneo x_c , allora la frequenza istantanea risulta pari a:

$$f_{ist}(t) = f_c + \frac{1}{2\pi} D_f x_c$$

La "probabilità" che $m(t) = x_c$ è legata alla densità di probabilità valutata in x_c , cioè $\underline{f_m}(x_c)$. Allora la "probabilità" di una componente spettrale alla frequenza:

$$f = f_c + \frac{1}{2\pi} D_f x_c$$

è data da $\underline{f_m}(x_c)$.

8.4 Implementazione di modulatori e demodulatori per sistemi FM

In questo paragrafo vediamo alcuni metodi per modulare e demodulare segnali FM. Gli stessi metodi possono essere usati anche per PM, con piccole varianti. Alcuni metodi hanno più rilevanza pratica, altri hanno un valore più storico e concettuale.

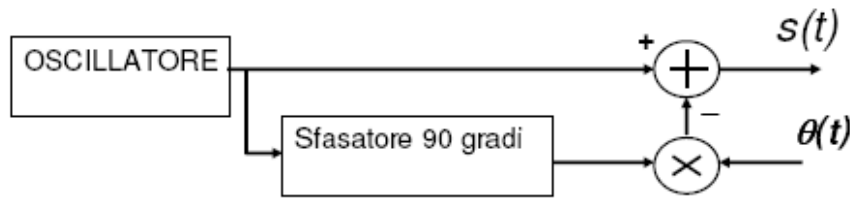


Figura 8.10: Circuito di Armstrong: genera segnali angolari a basso indice di modulazione. Il segnale di uscita $s(t)$ contiene ancora una modulazione di ampiezza spuria, che può essere eliminata mediante l'uso di un limitatore di ampiezza.

8.4.1 Modulatori FM

Vediamo due tipi di modulatori:

Modulatore diretto basato su VCO. Oggi, grazie ai processi delle tecnologie elettroniche, il metodo più comune per generare segnali FM è quello basato sulla modulazione diretta di un VCO (Voltage Controlled Oscillator). Il VCO è un oscillatore la cui frequenza di uscita dipende da una tensione di controllo in ingresso. Spesso, per aumentare la precisione, i VCO sono inseriti in circuiti PLL (Phase Locked Loop).

In questo corso non esaminiamo in dettaglio questo tipo di modulatori. Il lettore interessato può fare riferimento a [1] o [2].

Modulatore indiretto (basato su circuito di Armstrong). In passato, vista la difficoltà ad avere VCO stabili, si utilizzava un altro approccio, chiamato anche *metodo indiretto*. Oramai questo metodo ha più un valore storico e concettuale.

Si generava prima un segnale a basso indice di modulazione e poi lo si convertiva in un segnale ad alto indice di modulazione. Per bassi indici di modulazione abbiamo visto che il segnale modulato può essere approssimato con l'espressione (8.16):

$$s(t) \approx A_c \cos(2\pi f_c t) - A_c \theta(t) \sin(2\pi f_c t)$$

una cui possibile implementazione è mostrata in figura 8.10. Questo tipo di modulatore viene chiamato *circuito di Armstrong* e richiede l'utilizzo di soli oscillatori (fissi) e moltiplicatori.

In realtà, il circuito di Armstrong non genera un segnale esattamente di tipo PM/FM, in quanto il segnale risultante non è ad inviluppo costante, cioè presenta una modulazione di ampiezza spuria:

$$\begin{aligned} \tilde{s}(t) &\approx A_c [1 + j\theta(t)] \\ |\tilde{s}(t)|^2 &\approx A_c^2 [1 + \theta^2(t)] \end{aligned}$$

Il motivo è che nella (8.15) è stata fatta un'approssimazione, utilizzando lo sviluppo di McLaurin arrestato al primo ordine. Il circuito di Armstrong prevedeva dunque un limitatore di ampiezza (ad esempio amplificatore in saturazione) per rimuovere la modulazione di ampiezza spuria.

A questo punto è molto semplice generare un segnale angolare ad alto indice.

Generazione
segnale ad
alto indice

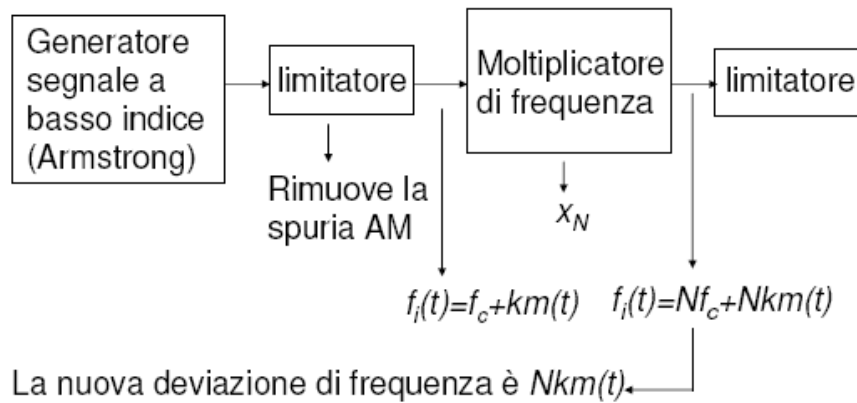


Figura 8.11: Generatore di segnali angolari ad alto indice di modulazione.

Basta aggiungere un moltiplicatore di frequenza, utilizzando lo schema di figura 8.11. Come mostrato nella figura, il segnale a basso indice entra nel moltiplicatore di frequenza, che ne moltiplica la frequenza istantanea per N . Scegliendo N opportunamente, il risultato è un segnale angolare ad alto indice di modulazione.

8.4.2 Demodulatori FM

La funzionalità base che un demodulatore FM deve implementare è una conversione del segnale FM in un corrispondente segnale in banda base proporzionale al segnale modulante. A tale scopo, esistono due diverse categorie di demodulatori FM:

Demodulatori FM basati su PLL. Il segnale di "errore" di un PLL è proporzionale alla differenza di fase tra i due segnali in ingresso al comparatore di fase, ed è dunque legato al segnale modulante. La maggior parte dei demodulatori moderni è di questo tipo.

In questo corso non esaminiamo in dettaglio questo tipo di demodulatori. Il lettore interessato può fare riferimento a [1] o [2].

Demodulatori FM basati su filtraggio e rivelazione di inviluppo. Questo schema ha più un valore storico e concettuale.

Il segnale FM (ad alto indice) è inviato ad un opportuno filtro che genera una conversione da FM a AM, e poi ad un rivelatore di inviluppo. Lo schema di questo demodulatore è mostrato in figura 8.12. Il filtro per effettuare la conversione da FM a AM deve avere il modulo della funzione di trasferimento all'incirca rettilineo, almeno all'interno della banda del segnale da demodulare (vedi figura 8.13).

Esaminiamo ora un caso particolare, che permette di svolgere i conti in dettaglio. Un possibile filtro per effettuare la conversione FM-AM è un derivatore ideale, cioè un filtro avente funzione di trasferimento:

$$H(f) = j 2\pi f$$

Infatti, se consideriamo un segnale con modulazione FM:

$$s(t) = A_c \cos \left[2\pi f_c t + \int_{-\infty}^t m(\tau) d\tau \right]$$

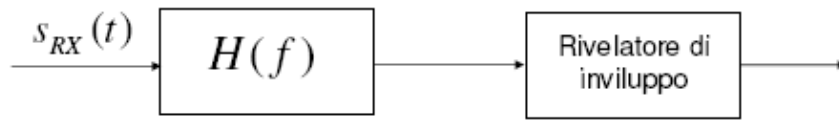


Figura 8.12: Schema a blocchi demodulatore FM basato su filtraggio e rivelazione di inviluppo.

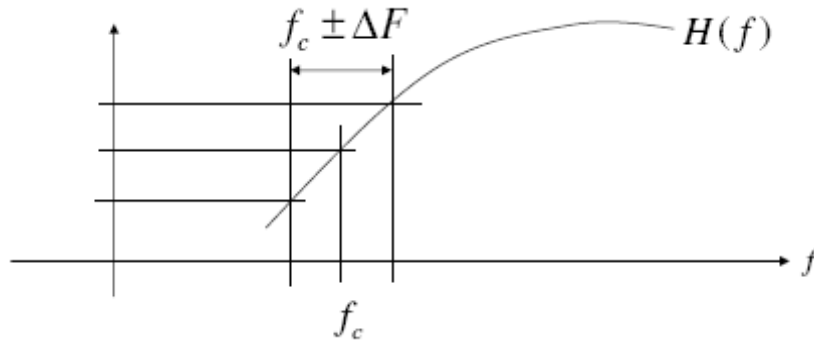


Figura 8.13: Filtro che converte modulazione FM in AM. Il filtro deve avere il modulo della risposta in frequenza lineare all'interno della banda del segnale da demodulare.

effettuandone la derivata si trova:

$$s'(t) = -A_c [2\pi f_c + m(t)] \sin \left[2\pi f_c t + \int_{-\infty}^t m(\tau) d\tau \right]$$

$s'(t)$ è un segnale modulato AM a frequenza variabile. Passando questo segnale in un rivelatore di inviluppo si può ricavare infine il segnale modulante $m(t)$. Si noti infatti che il fatto che il segnale sia a frequenza variabile non influenza il funzionamento del rivelatore di inviluppo.

8.5 Effetto del rumore sulle modulazioni angolari

Affrontiamo in questo paragrafo l'argomento dell'analisi delle prestazioni dei sistemi PM/FM in presenza di rumore.

Demodulatore di riferimento per il calcolo del rapporto segnale-rumore

Si consideri la figura 8.14, che mostra la struttura del demodulatore di riferimento per il calcolo del rapporto segnale-rumore. Si considerano nel seguito solo circuiti con rivelazione FM o PM ideale. I blocchi presenti nello schema sono i seguenti.

Filtro IF B_T . Si tratta di un filtro passa-banda con banda B_T (banda del segnale modulato) e centrato attorno alla frequenza f_c (IF sta per frequenza intermedia, in inglese *Intermediate Frequency*). Il suo compito è quello di ridurre la banda del rumore senza distorcere il segnale modulato.

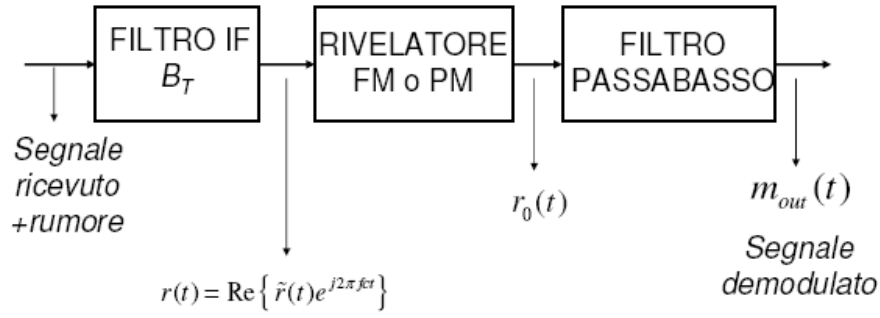


Figura 8.14: Schema a blocchi demodulatore FM/PM di riferimento per calcolo rapporto segnale-rumore.

Rivelatore FM o PM. Si tratta di un rivelatore FM o PM ideale. Per rivelatori FM o PM ideali si intendono dei blocchi funzionali che svolgano in maniera esatta le seguenti conversioni:

$$\begin{aligned} \text{per PM: } r_0(t) &= \arg[\tilde{r}(t)] \\ \text{per FM: } r_0(t) &= \frac{d}{dt} \arg[\tilde{r}(t)] \end{aligned}$$

cioè che forniscano idealmente dei segnali proporzionali a:

- Per PM: la fase dell'involuppo complesso del segnale ricevuto;
- Per FM: la frequenza istantanea dell'involuppo complesso del segnale ricevuto.

Filtro passabasso. Filtro passa-basso con banda B_m (banda del segnale modulante). Serve a limitare ulteriormente la banda del rumore. Non distorce il segnale di informazione.

Prima di procedere col calcolo del rapporto segnale-rumore, è necessario introdurre alcuni risultati relativi alla rappresentazione del rumore gaussiano in banda traslata.

Rumore gaussiano in coordinate polari

Consideriamo un rumore gaussiano bianco $n(t)$ in banda traslata (vedi figura 8.15). In base a quanto abbiamo visto nel capitolo su segnale analitico e involuppo complesso, $n(t)$ può essere rappresentato come:

$$\tilde{n}(t) = x_n(t) + j y_n(t)$$

dove $x_n(t)$ è la parte in fase e $y_n(t)$ è la parte in quadratura.

Un'ulteriore rappresentazione è quella in modulo e fase:

$$\tilde{n}(t) = R_n(t) e^{j\theta_n(t)} \quad (8.17)$$

da cui si ricava facilmente la rappresentazione in coordinate polari:

$$n(t) = R_n(t) \cos[2\pi f_c t + \theta_n(t)]$$

Per i dettagli su queste rappresentazioni, consultare il capitolo su segnale analitico e involuppo complesso.

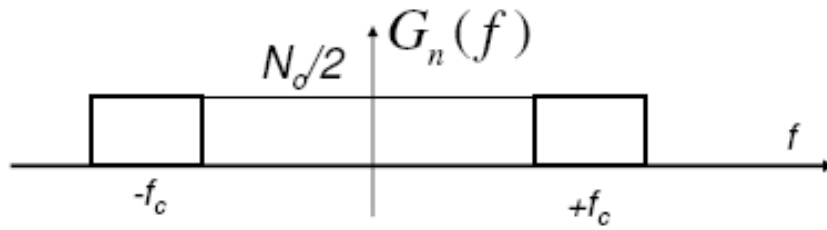


Figura 8.15: Spettro di potenza di rumore gaussiano bianco in banda traslata.

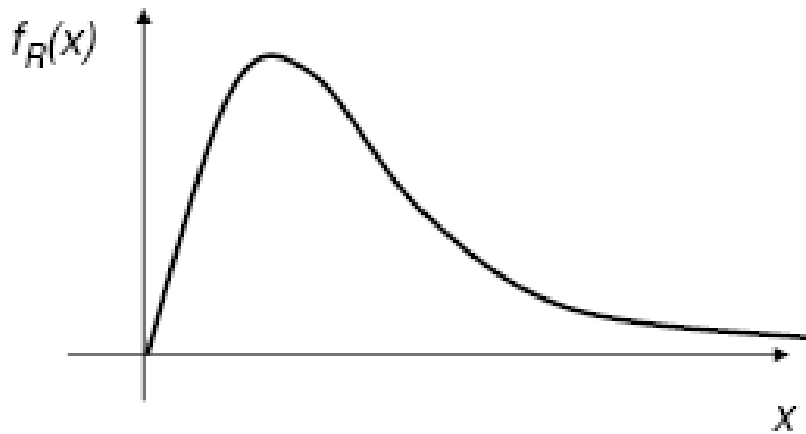


Figura 8.16: Densità di probabilità di Rayleigh.

Si può scrivere, a partire dalle componenti in fase e quadratura del rumore gaussiano:

$$\begin{cases} \theta_n(t) = \arg \{x_n(t) + jy_n(t)\} \\ R_n(t) = \sqrt{x_n^2(t) + y_n^2(t)} \end{cases}$$

Con la notazione $\arg \{\cdot\}$ si intende la fase del numero complesso tra parentesi graffe.

Si può dimostrare (non lo facciamo) che:

1. $\theta_n(t)$ è un processo casuale (stazionario ed ergodico) uniformemente distribuito in $[0, 2\pi]$;
2. $R_n(t)$ è un processo casuale (stazionario ed ergodico) con densità di probabilità di Rayleigh (vedi figura 8.16):

$$f_R(x) = \begin{cases} 0 & \text{per } x < 0 \\ \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & \text{per } x > 0 \end{cases}$$

dove σ^2 = varianza del processo gaussiano reale.

8.5.1 Calcolo rapporto segnale-rumore per sistemi PM

Calcoliamo ora l'SNR in uscita dal demodulatore per sistemi PM. Il calcolo dell'SNR per sistemi FM è più complicato e verrà effettuato nel prossimo paragrafo.

Facciamo riferimento alla figura 8.14. Partiamo dall'espressione del segnale all'uscita del filtro IF, espresso in termini di inviluppo complesso:

$$\tilde{r}(t) = \underbrace{A_c e^{j\theta_s(t)}}_{\text{Segnale utile}} + \underbrace{x_n(t) + j y_n(t)}_{\text{Rumore}} \quad (8.18)$$

La prima parte indica il segnale utile con una modulazione di fase legata a $\theta_s(t)$. La seconda parte indica il rumore filtrato su banda B_T . Per fare i calcoli, è utile riscrivere la (8.18) in termini di rumore in modulo e fase (si ricorda la (8.17)):

$$\tilde{r}(t) = A_c e^{j\theta_s(t)} + R_n(t) e^{j\theta_n(t)}$$

All'uscita del rivelatore di fase ideale si ha:

$$\begin{aligned} r_0(t) &= k \arg \{ \tilde{r}(t) \} \\ &= k \arg \left\{ A_c e^{j\theta_s(t)} + R_n(t) e^{j\theta_n(t)} \right\} \end{aligned} \quad (8.19)$$

dove k è un generico fattore di scala.

Per andare avanti coi conti occorre fare delle approssimazioni. Nel seguito di questa trattazione assumeremo che il rapporto segnale-rumore all'uscita del filtro IF (cioè all'ingresso del rivelatore di fase) sia elevato, cioè che si lavori per:

Approssimazione lavoro con $(S/N)_{in}$ elevato

$$(S/N)_{in} \text{ elevato}$$

Introduciamo una rappresentazione in termini di vettori nel piano complesso, e omettiamo nel seguito la dipendenza dal tempo:

$$\begin{aligned} \tilde{s}_{RX} &= A_c e^{j\theta_s} && \text{Segnale utile} \\ \tilde{n} &= R_n e^{j\theta_n} && \text{Rumore} \end{aligned}$$

Su questi numeri complessi, l'ipotesi $(S/N)_{in}$ elevato significa in pratica:

$$|\tilde{s}_{RX}| \gg |\tilde{n}| \Rightarrow A_c \gg R_n$$

La figura 8.17 presenta una costruzione grafica che permette di calcolare l'uscita del rivelatore di fase sotto l'approssimazione:

$$A_c \gg R_n$$

Si tratta di una costruzione puramente geometrica per calcolare la fase del vettore risultante dalla somma del vettore di segnale con quello di rumore, cioè la fase di:

$$\tilde{r}(t) = \underbrace{A_c e^{j\theta_s(t)}}_{\tilde{s}_{RX}} + \underbrace{R_n(t) e^{j\theta_n(t)}}_{\tilde{n}} \quad (8.20)$$

Si vedano le figure 8.18, 8.19, 8.20 e 8.21. L'obiettivo di queste figure è il calcolo dell'angolo θ_r , cioè la fase di $\tilde{r}(t)$. Le figure svolgono dei semplici calcoli geometrici.

Soffermiamoci a spiegare la figura 8.21. Visto che $A_c \gg R_n$, si ha che $\Delta\theta$ è un angolo piccolo. Il segmento di lunghezza $R_n \sin(\theta_n - \theta_s)$ è allora approssimabile come un arco di circonferenza. Il raggio di tale circonferenza è circa uguale ad A_c , visto che

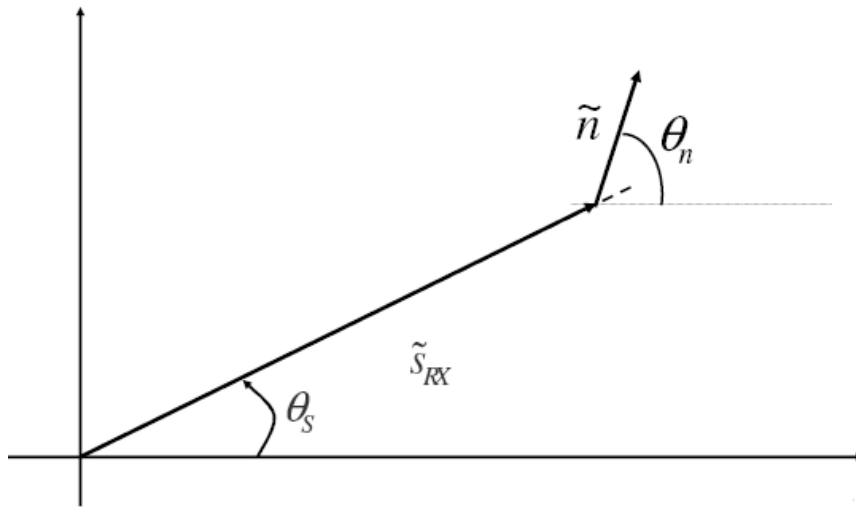


Figura 8.17: Costruzione geometrica per il calcolo della fase del vettore $\tilde{r}(t)$ della (8.20). Tale vettore è risultante dalla somma del segnale utile \tilde{s}_{RX} e del rumore \tilde{n} . Si ricorda che $|\tilde{s}_{RX}| \gg |\tilde{n}|$.

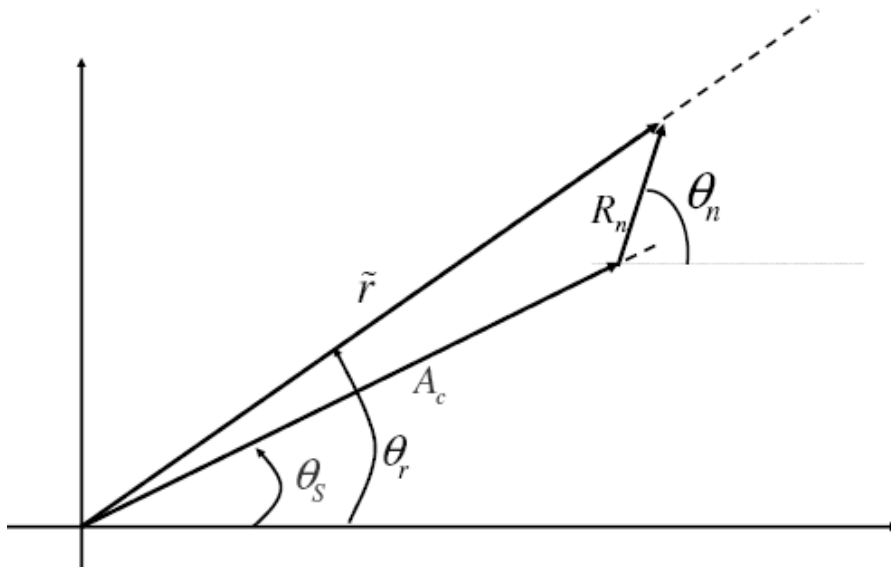


Figura 8.18: Figura 1 di 4 per il calcolo della fase della (8.20).

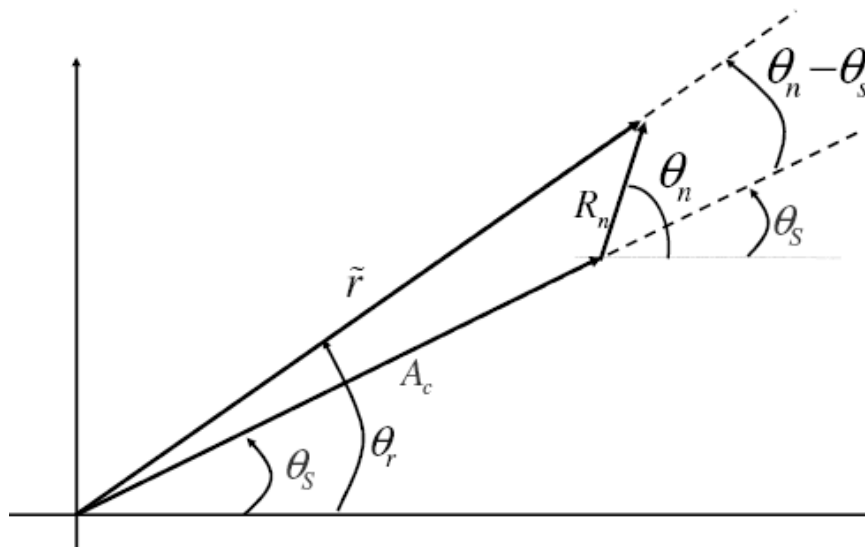


Figura 8.19: Figura 2 di 4 per il calcolo della fase della (8.20).

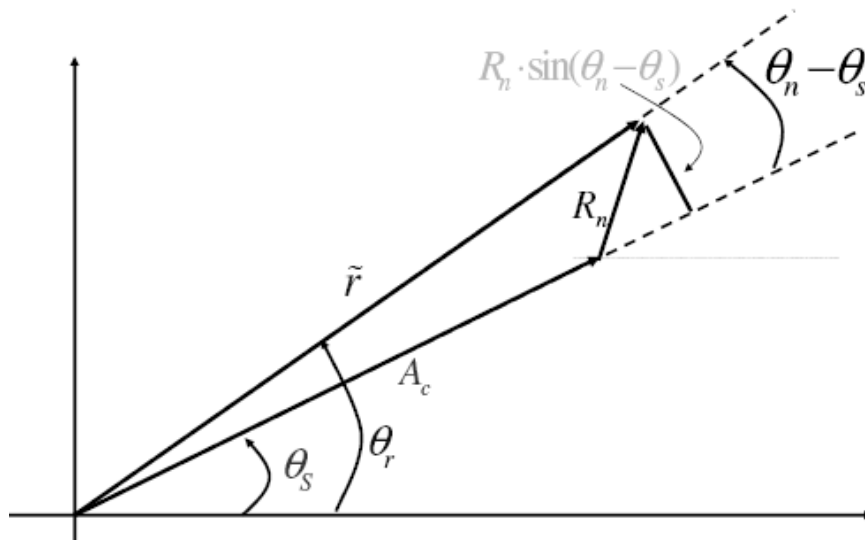


Figura 8.20: Figura 3 di 4 per il calcolo della fase della (8.20).

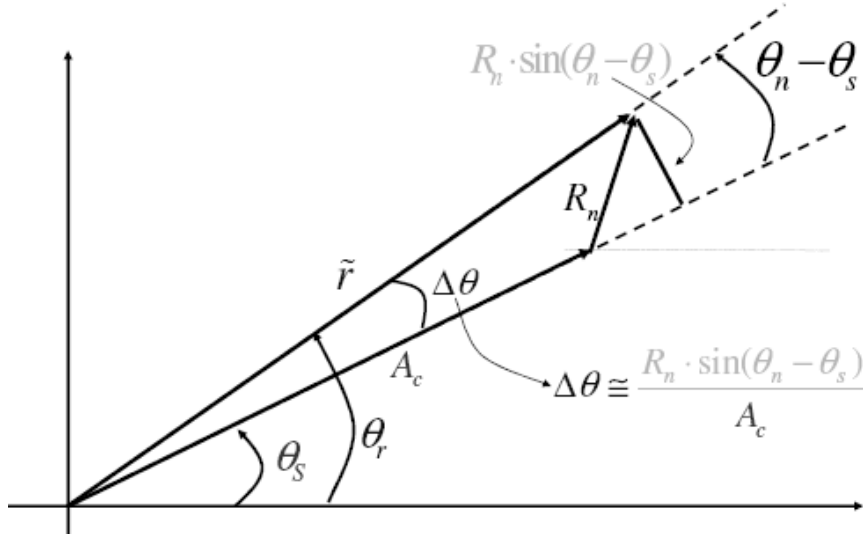


Figura 8.21: Figura 4 di 4 per il calcolo della fase della (8.20).

$A_c \gg R_n$. Come è noto, moltiplicando il valore di un angolo in radianti per il valore del raggio della circonferenza, si ottiene la lunghezza dell'arco di circonferenza sotteso da quell'angolo.

θ_r vale pertanto:

$$\begin{aligned}\theta_r &= \theta_s + \Delta\theta \\ &\approx \theta_s + \frac{R_n \sin(\theta_n - \theta_s)}{A_c}\end{aligned}$$

Quindi:

$$\begin{aligned}r_0(t) &= k \theta_r(t) \\ &= k \left\{ \theta_s(t) + \frac{1}{A_c} R_n(t) \sin [\theta_n(t) - \theta_s(t)] \right\}\end{aligned}\quad (8.21)$$

Osserviamo ora che in questa espressione compare $\theta_n(t) - \theta_s(t)$. Questa quantità sarà ancora un processo casuale uniformemente distribuito sull'intervallo $[0, 2\pi]$. Dunque introducendo il nuovo processo casuale:

$$\theta'_n(t) = \theta_n(t) - \theta_s(t)$$

questo avrà ancora una densità di probabilità uniforme in $[0, 2\pi]$.

Allora posso riscrivere la (8.21) in questo modo:

$$r_0(t) = k \left\{ \theta_s(t) + \frac{1}{A_c} R_n(t) \sin [\theta'_n(t)] \right\}\quad (8.22)$$

Consideriamo ora il termine:

$$R_n(t) \sin [\theta'_n(t)]$$

Questo termine è assolutamente analogo dal punto di vista dei processi casuali alla parte in quadratura di un rumore gaussiano. Allora posso scrivere:

$$R_n(t) \sin [\theta'_n(t)] = y_n(t)$$

Uscita del
rivelatore di
fase in demodulatore
PM

e in conclusione posso riscrivere la (8.22) così:

$$\begin{aligned} r_0(t) &= k \left[\theta_s(t) + \frac{1}{A_c} y_n(t) \right] \\ &= \underbrace{k \theta_s(t)}_{\text{Segnale utile}} + \underbrace{\frac{k}{A_c} y_n(t)}_{\text{Rumore}} \end{aligned}$$

Questo risultato ci dice che, sotto le ipotesi fatte, il segnale $r_0(t)$ all'uscita del rivelatore di fase è composto da:

1. La fase del segnale utile ($k \theta_s(t)$), cioè il termine proporzionale al segnale modulante;
2. Un termine di rumore additivo ($k y_n(t)/A_c$), coincidente (dal punto di vista dei processi casuali) con una componente di rumore in quadratura. È importante osservare che **questa componente di rumore è inversamente proporzionale all'ampiezza A_c del segnale modulante**. Quindi, maggiore è il livello del segnale utile, minore sarà il livello del rumore. Questo risultato è completamente diverso da quello che succedeva nelle modulazioni di ampiezza, in cui la componente di rumore era indipendente dal segnale utile.

Importante

Reintroducendo i parametri del modulatore, abbiamo dunque ottenuto:

$$r_0(t) = k D_p m(t) + \frac{k}{A_c} y_n(t) \quad (8.23)$$

dove:

- D_p : sensibilità modulazione di fase;
- $m(t)$: segnale modulante che porta informazione;
- $y_n(t)$: rumore con densità spettrale di potenza N_0 .

Ora procedo
con gli stessi
conti fatti
per AM

Notiamo a questo punto che per proseguire nella trattazione, possiamo di qui in poi seguire gli stessi passaggi fatti per la ricezione AM, in quanto ora siamo giunti ad una espressione in banda base che comprende il segnale modulante e rumore gaussiano additivo.

Il termine $r_0(t)$ della (8.23) viene poi ulteriormente filtrato con un passa-basso ideale $h(t)$ su banda B_m (vedi figura 8.22) per dare luogo al segnale demodulato (uscita del demodulatore di figura 8.14):

$$m_{out}(t) = \underbrace{k D_p m(t)}_{\text{Segnale utile}} + \underbrace{\frac{k}{A_c} y_n(t) * h(t)}_{\text{Rumore}}$$

avendo qui fatto le stesse ipotesi fatte nel caso AM. In quest'ultima espressione sono ben chiare le componenti di segnale e di rumore, e possiamo dunque calcolare il rapporto segnale-rumore in uscita dal demodulatore (cioè il parametro al quale volevamo arrivare):

$$\begin{aligned} \left(\frac{S}{N} \right)_{out} &= \frac{k^2 D_p^2 < m^2(t) >}{\frac{k^2}{A_c^2} 2N_0 B_m} \\ &= \frac{A_c^2 D_p^2 < m^2(t) >}{2N_0 B_m} \end{aligned} \quad (8.24)$$

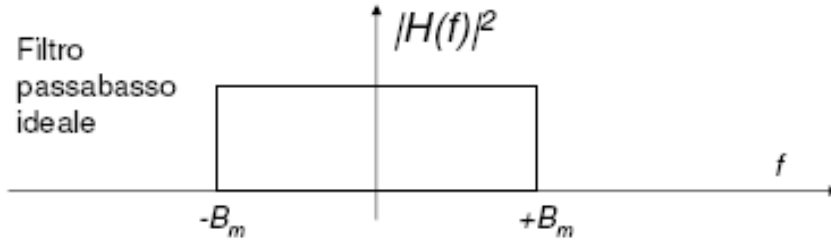


Figura 8.22: Filtro passa-basso ideale con banda B_m .

Così come abbiamo fatto per AM, vogliamo ora collegare questo risultato a:

$$\left(\frac{S}{N}\right)_{BB} \quad \left(\frac{S}{N}\right)_{in}$$

A tale scopo, reintroduciamo intanto l'indice di modulazione di fase introdotto nella (8.5) a pagina 145:

$$\beta_p = D_p V_p$$

e ricordando che:

$$P_{RX} = \frac{A_c^2}{2}$$

la (8.24) si può riscrivere così:

$$\begin{aligned} \left(\frac{S}{N}\right)_{out} &= \frac{A_c^2}{2} \frac{\beta_p^2 \langle \frac{m^2(t)}{V_p^2} \rangle}{N_0 B_m} \\ &= \frac{P_{RX}}{N_0 B_m} \beta_p^2 \langle \frac{m^2(t)}{V_p^2} \rangle \end{aligned}$$

Ricordando ora che per il sistema di riferimento in banda base si ha:

$$\left(\frac{S}{N}\right)_{BB} = \frac{P_{RX}}{N_0 B_m}$$

Si ottiene il risultato importante:

Formula
(S/N)_{out}
sistemi PM

$$\boxed{\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} \beta_p^2 \langle \frac{m^2(t)}{V_p^2} \rangle = \left(\frac{S}{N}\right)_{BB} \beta_p^2 \langle \left(\frac{m(t)}{V_p}\right)^2 \rangle} \quad (8.25)$$

Si ricorda che questa espressione è stata ottenuta sotto l'ipotesi di SNR elevato all'ingresso del rivelatore di fase.

Calcoliamo ora anche il rapporto segnale-rumore all'ingresso del rivelatore di fase:

$$\begin{aligned} \left(\frac{S}{N}\right)_{in} &= \frac{A_c^2}{2} \frac{1}{\frac{N_0}{2} 2B_T} \\ &= \frac{A_c^2}{2N_0 B_T} \end{aligned}$$

Se inseriamo la formula di Carson $B_T = 2(\beta_p + 1)B_m$ otteniamo:

$$\left(\frac{S}{N}\right)_{in} = \frac{A_c^2}{4N_0(\beta_p + 1)B_m}$$

$(S/N)_{out}$ in
funzione di
 $(S/N)_{in}$ per
PM

E dunque, sostituendo, si ottiene l'interessante relazione:

$$\left(\frac{S}{N}\right)_{out} = \left(\frac{S}{N}\right)_{in} 2\beta_p^2 (\beta_p + 1) < \left(\frac{m(t)}{V_p}\right)^2 > \quad (8.26)$$

Dalla (8.25) e dalla (8.26) possiamo fare diverse osservazioni importanti sulla modulazione PM. Vedremo che le stesse osservazioni varranno anche per FM.

Osservazioni
importanti
su PM e FM

1. **Le modulazioni angolari possono dare $(S/N)_{out}$ migliori del caso in banda base per indici di modulazione β sufficientemente elevati.** Dalle formule precedenti si può infatti notare che il processo di demodulazione aumenta di molto $(S/N)_{out}$ per β elevati. Nelle modulazioni AM una cosa di questo tipo non era possibile.
2. **Le modulazioni angolari permettono di migliorare SNR a spese di una maggiore occupazione di banda.** Notare che questa possibilità di trade-off è assente per le modulazioni di ampiezza.

Queste sono le caratteristiche fondamentali delle modulazioni d'angolo rispetto a quelle AM. Almeno in linea di principio, il rapporto segnale-rumore in uscita può essere incrementato a piacere aumentando l'indice di modulazione, con però la conseguenza di aumentare la banda B_T richiesta sul canale di trasmissione. Torneremo su questo punto nel paragrafo 8.5.3 a pagina 172 quando parleremo di effetto soglia.

8.5.2 Calcolo rapporto segnale-rumore per sistemi FM

Calcoliamo ora il rapporto segnale-rumore dopo la demodulazione per FM. Facciamo anche qui riferimento alla figura 8.14. La modulazione FM è di gran lunga più utilizzata della PM. Tuttavia si è preferito introdurre prima la trattazione del comportamento in presenza di rumore per PM in quanto la trattazione per PM è matematicamente più semplice.

I passaggi sono simili al caso PM, salvo che ora la fase del segnale utile è data da:

$$\theta_s(t) = D_f \int_{-\infty}^t m(\tau) d\tau$$

Eseguendo passaggi del tutto simili a PM, e per $(S/N)_{in}$ elevato, per l'uscita del rivelatore di frequenza si arriva alla seguente espressione:

Approssimazione
lavoro con
 $(S/N)_{in}$
elevato

$$\begin{aligned} r_0(t) &= \frac{k}{2\pi} \frac{d}{dt} \arg \{ \tilde{r}(t) \} \\ &= \frac{k}{2\pi} \frac{d\theta_r(t)}{dt} \end{aligned}$$

Infatti adesso si considera un rivelatore ideale di frequenza, che fornisce la derivata della fase del segnale ricevuto. Con approssimazioni analoghe a PM, è possibile arrivare a scrivere il segnale all'uscita del rivelatore di frequenza come:

$$r_0(t) = s_0(t) + n_0(t)$$

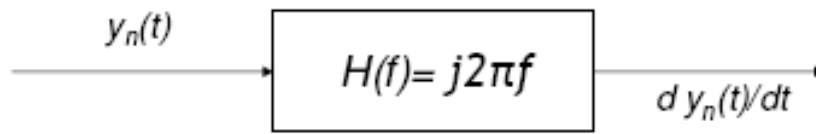
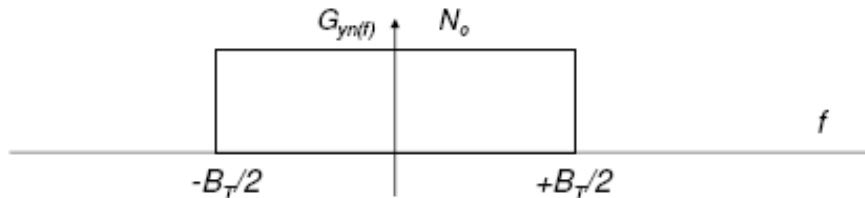


Figura 8.23: Filtro derivatore.

Figura 8.24: Densità spettrale di potenza del rumore $y_n(t)$ all'uscita del filtro IF per demodulazione FM.

dove:

$$\begin{cases} s_0(t) = \frac{k}{2\pi} \frac{d}{dt} \theta_s(t) = \frac{kD_f}{2\pi} m(t) \\ n_0(t) = \frac{k}{2\pi A_c} \frac{d}{dt} y_n(t) \end{cases}$$

Anche qui, come in PM, si ha che **la potenza di rumore è inversamente proporzionale all'ampiezza del segnale utile A_c** . In pratica, l'unica differenza rispetto a PM sta nella presenza della derivata sul processo di rumore di $y_n(t)$. Ne consegue che il rumore su $r_0(t)$ è ancora additivo, ma con uno spettro non piatto, in quanto l'operazione di derivata è analoga ad un filtraggio con funzione di trasferimento (vedi figura 8.23):

$$H(f) = j 2\pi f$$

Ne consegue dunque che lo spettro di potenza della componente di rumore $n_0(t)$ è dato da:

$$G_{n_0}(f) = \left(\frac{k}{2\pi A_c} \right)^2 (j2\pi f)^2 G_{y_n}(f)$$

dove $G_{y_n}(f)$ è una componente di rumore filtrato dal filtro IF, e dunque con la densità spettrale mostrata in figura 8.24. $G_{y_n}(f)$ è piatto e vale N_0 nell'intervallo $[-B_T/2, +B_T/2]$. Allora:

$$G_{n_0}(f) = \begin{cases} \frac{k^2}{A_c^2} N_0 f^2 & \text{per } |f| < \frac{B_T}{2} \\ 0 & \text{altrimenti} \end{cases}$$

$G_{n_0}(f)$ è rappresentata in figura 8.25.

Quindi, nella modulazione FM, l'effetto del rumore è più rilevante ad alta frequenza e meno rilevante a bassa frequenza. $G_{n_0}(f)$ è ulteriormente filtrata dal passa-basso ideale presente al fondo del ricevitore, la cui funzione di trasferimento è quella già considerata sia per AM che per PM (figura 8.26). Facendo le ipotesi già fatte più volte, e cioè:

In FM viene
introdotta
più rumore
ad alta
frequenza

- Il filtraggio non modifica il segnale utile;
- Il filtraggio è significativo solo sulla componente di rumore.

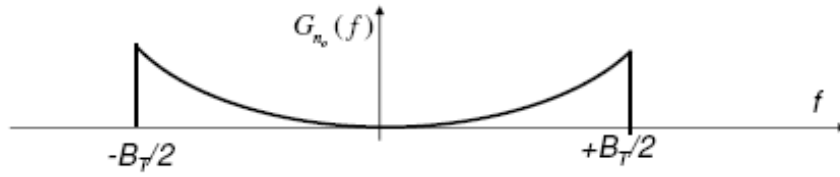


Figura 8.25: Densità spettrale di potenza del rumore in uscita dal rivelatore di frequenza per demodulazione FM.

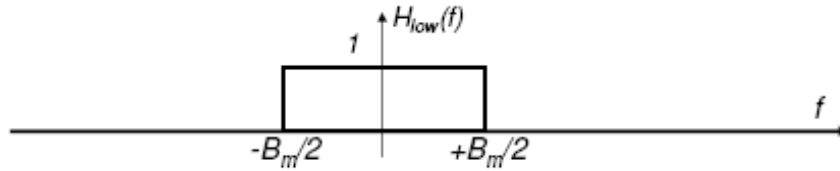


Figura 8.26: Risposta in frequenza filtro passa-basso alla fine del demodulatore FM.

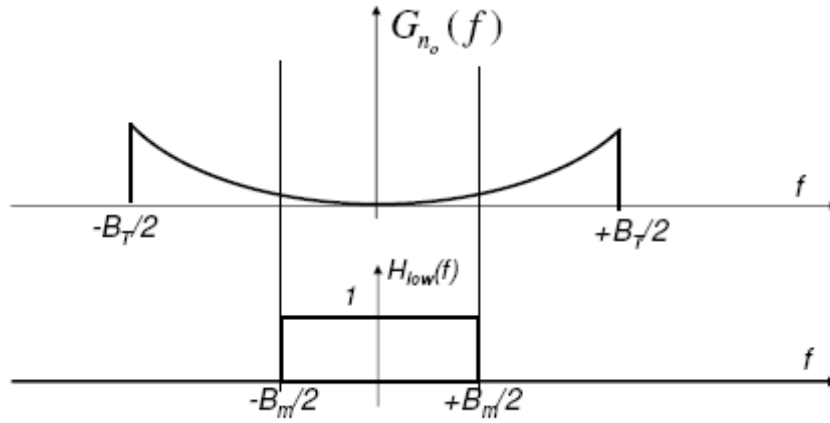


Figura 8.27: Rappresentazione nel dominio della frequenza della (8.27).

All'uscita del filtro passa-basso di ideale si ha dunque:

$$m_{out}(t) = \frac{kD_f}{2\pi} m(t) + n_{F,0}(t)$$

dove $n_{F,0}(t)$ è il processo $n_0(t)$ filtrato dal passa-basso ideale.
In formule:

$$n_{F,0}(t) = n_0(t) * h_{LOW}(t) \quad (8.27)$$

che corrisponde a un prodotto nel dominio della frequenza (vedi figura 8.27).

Per arrivare al rapporto segnale-rumore in uscita, calcoliamo dunque la potenza del

rumore filtrato:

$$\begin{aligned}
 P_{\tilde{n}_{F,0}} &= \int_{-B_m}^{+B_m} G_{\tilde{n}_{F,0}}(f) df \\
 &= \frac{k^2}{A_c^2} N_0 \int_{-B_m}^{+B_m} f^2 df \\
 &= \frac{2}{3} \left(\frac{k}{A_c} \right)^2 N_0 B_m^3
 \end{aligned}$$

e conseguentemente il rapporto segnale-rumore vale:

$$\begin{aligned}
 \left(\frac{S}{N} \right)_{out} &= \frac{\left(\frac{k D_f}{2\pi} \right)^2 \langle m^2(t) \rangle}{\frac{2}{3} \left(\frac{k}{A_c} \right)^2 N_0 B_m^3} \\
 &= \frac{3 A_c^2 \left(\frac{D_f}{2\pi B_m} \right)^2 \langle m^2(t) \rangle}{2 N_0 B_m}
 \end{aligned}$$

Reintroducendo ora l'indice di modulazione di frequenza definito nella (8.6) (pagina 145) e ricordando la (8.4):

$$\frac{D_f}{2\pi B_m} = \frac{\beta_f}{V_p}$$

si trova:

$$\left(\frac{S}{N} \right)_{out} = \frac{3 A_c^2 \beta_f^2}{2 N_0 B_m} \langle \left(\frac{m(t)}{V_p} \right)^2 \rangle$$

Otteniamo infine la formula fondamentale per le prestazioni dei sistemi FM:

Formula
(S/N)_{out}
sistemi FM

$$\boxed{\left(\frac{S}{N} \right)_{out} = \frac{P_{RX}}{N_0 B_m} 3 \beta_f^2 \langle \left(\frac{m(t)}{V_p} \right)^2 \rangle = \left(\frac{S}{N} \right)_{BB} 3 \beta_f^2 \langle \left(\frac{m(t)}{V_p} \right)^2 \rangle} \quad (8.28)$$

Si ricorda che questa espressione è stata ottenuta sotto l'ipotesi di SNR elevato all'ingresso del rivelatore di frequenza.

Deduciamo anche le formule per (S/N)_{in}:

$$\begin{aligned}
 \left(\frac{S}{N} \right)_{in} &= \frac{A_c^2}{4 N_0 B_m (\beta_f + 1)} \\
 &= \frac{P_{RX}}{N_0 B_m} \frac{1}{2(\beta_f + 1)}
 \end{aligned} \quad (8.29)$$

Quindi (S/N)_{out} in funzione di (S/N)_{in} vale:

(S/N)_{out} in
funzione di
(S/N)_{in} per
FM

$$\left(\frac{S}{N} \right)_{out} = \left(\frac{S}{N} \right)_{in} 6 \beta_f^2 (\beta_f + 1) \langle \left(\frac{m(t)}{V_p} \right)^2 \rangle \quad (8.30)$$

Osservando le espressioni (8.28) e (8.30) e confrontandole con le espressioni (8.25) e (8.26) precedentemente ottenute per PM, notiamo che le formule sono praticamente identiche, fatto salvo per un fattore 3 in FM.

Possiamo ripetere quindi anche per FM le osservazioni fatte a pagina 168 per PM.

Stesse
osservazioni
PM

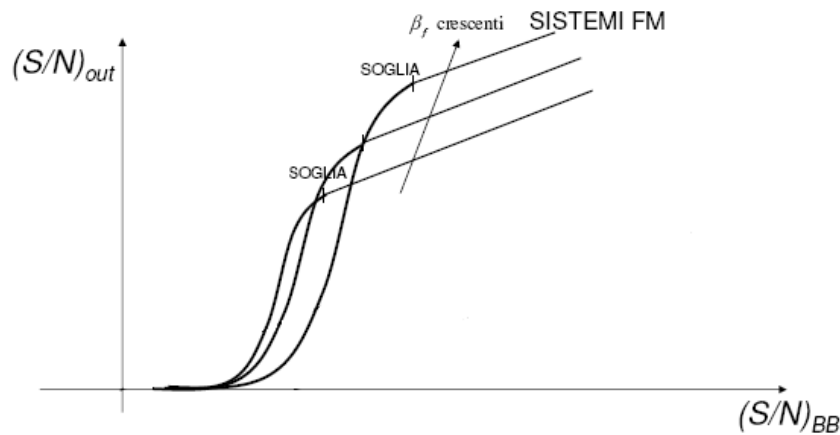


Figura 8.28: Prestazioni effettive FM in funzione dell'indice di modulazione. Si noti come il valore di soglia aumenti per β_f crescenti.

8.5.3 Effetto soglia nelle modulazioni angolari

In questo paragrafo analizziamo l'effetto soglia per le modulazioni angolari. Effettueremo l'analisi per FM ma i risultati si applicano identici anche a PM.

Dall'espressione (8.28) sembra che le prestazioni in termini di $(S/N)_{out}$ debbano migliorare in maniera costante al crescere di β_f . In realtà, la (8.28) è valida solo nell'ipotesi di SNR elevato all'ingresso del rivelatore di frequenza. Esaminando l'espressione (8.29) di $(S/N)_{in}$ si vede che, a parità di $(S/N)_{BB}$, aumentando l'indice di modulazione diminuisce $(S/N)_{in}$. Quindi l'approssimazione $(S/N)_{in} \gg 1$ oltre un certo limite non è più valida.

La discussione matematica dell'effetto soglia e il calcolo del β_f di soglia sono piuttosto complicati e esulano dagli scopi di questo corso. Il risultato importante è che la soglia è più alta per β_f crescenti.

La soglia è
più alta per
 β_f crescenti

La figura 8.28 mostra le prestazioni effettive dei sistemi FM. La figura può essere letta in due modi diversi, che sono due facce della stessa medaglia:

- Fissato un certo $(S/N)_{BB}$, quello che succede è che esiste un valore di soglia dell'indice di modulazione, al di sopra del quale le prestazioni crollano drasticamente.
- La stessa cosa si può vedere anche in questo modo: fissato un certo β_f , esisterà un valore di soglia di $(S/N)_{BB}$ (che si ricorda è legato a $(S/N)_{in}$) al di sopra del quale vale la (8.28). Al di sotto di questo valore le prestazioni del sistema precipitano.

Conclusione su FM

Possiamo a questo punto trarre le conclusioni su FM. **FM è usatissima in quanto, sopra soglia, permette di ottenere dei grossi vantaggi rispetto alla trasmissione**

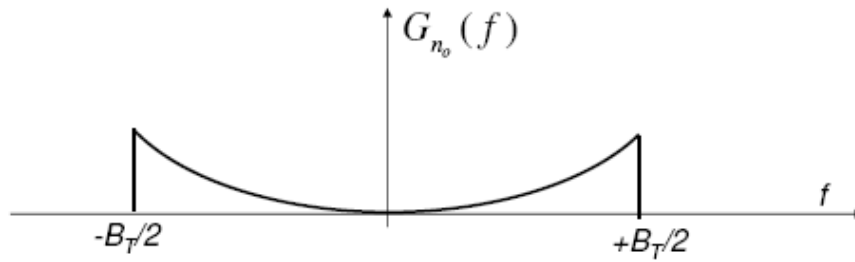


Figura 8.29: Spettro di potenza del rumore all'uscita del rivelatore di frequenza FM.

di riferimento in banda base. Tali vantaggi sono legati al termine:

$$\left(\frac{S}{N}\right)_{out} = \left(\frac{S}{N}\right)_{BB} 3\beta_f^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

Cioè si ha un guadagno proporzionale a:

$$3\beta_f^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

Questo è il motivo fondamentale di utilizzo delle modulazioni FM.

Un ulteriore incremento delle prestazioni può essere ottenuto con la tecnica illustrata nel prossimo paragrafo.

8.5.4 Tecniche di pre-enfasi e de-enfasi per la modulazione FM

Vediamo ora un metodo per migliorare le prestazioni della modulazione FM.

Iniziamo con un'osservazione di partenza. Come abbiamo già visto, il rumore all'uscita del rivelatore di frequenza è additivo, ma non è piatto (vedi figura 8.29). Questa è una situazione nettamente diversa rispetto alle modulazioni AM o PM. A causa di questo fatto, le componenti ad alta frequenza del segnale utile sono più degradate dal rumore, rispetto a quelle a bassa frequenza.

Si può allora pensare di filtrare maggiormente con il passa-basso in ricezione. Si parla in questo caso di *filtro di de-enfasi*. Si usano spesso filtri semplici, quali passa-basso ad un polo come in figura 8.30. Sia $H_{LOW}(f)$ la risposta in frequenza del filtro e sia f_1 la sua frequenza di taglio a 3 dB. Per avere una significativa riduzione di rumore, il taglio deve essere decisamente all'interno della banda del segnale modulante, cioè deve essere:

$$f_1 < B_m$$

A questo punto però il filtro taglierà anche in parte il segnale utile.

Si usa allora in trasmissione un altro filtro, detto di *pre-enfasi*, con funzione di trasferimento $1/H_{LOW}(f)$ (vedi figura 8.31). Si lavora cioè sullo schema a blocchi complessivo di figura 8.32. Almeno in prima approssimazione:

1. Il segnale modulante non viene distorto, in quanto passa attraverso i due filtri di pre-enfasi e di de-enfasi, che hanno funzioni di trasferimento con prodotto unitario;

Filtro di
de-enfasi
(in RX)

Filtro di
pre-enfasi
(in TX)

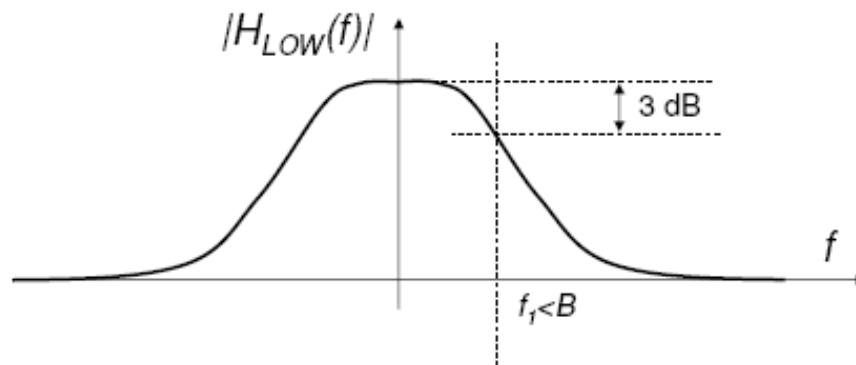


Figura 8.30: Esempio di funzione di trasferimento di filtro di de-enfasi a un polo per sistema FM.

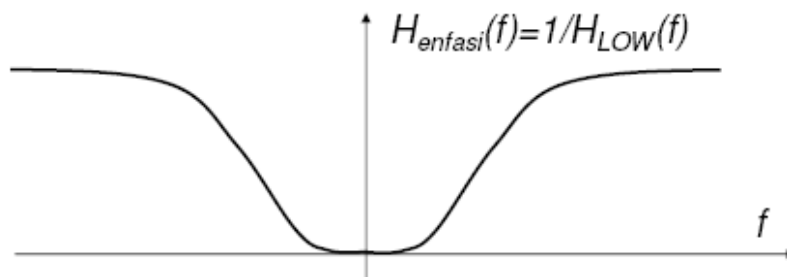


Figura 8.31: Esempio di funzione di trasferimento di filtro di pre-enfasi a un polo per sistema FM.

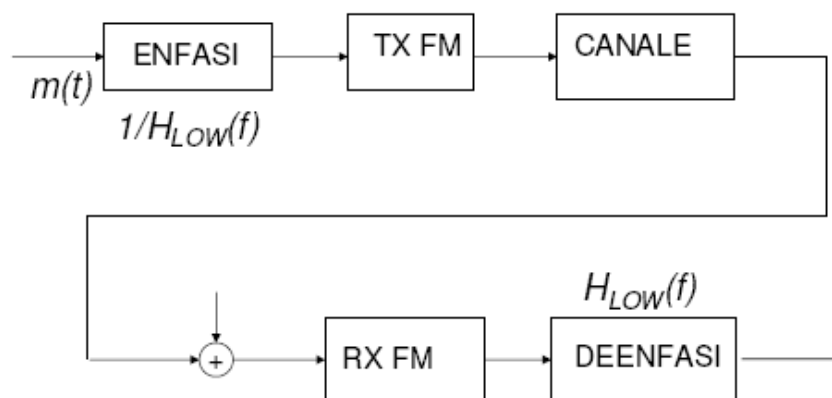


Figura 8.32: Schema di sistema di trasmissione FM con filtri di pre-enfasi e de-enfasi.

2. Il rumore viene invece filtrato maggiormente rispetto al caso di ricevitore FM standard, grazie al filtro di de-enfasi.

Sia dunque f_1 la frequenza del polo del filtro di de-enfasi. Grazie a questa metodologia si ottiene, tramite alcune approssimazioni (che saranno viste ad esercitazione):

$$\left(\frac{S}{N}\right)_{out} \approx \left(\frac{S}{N}\right)_{BB} \beta_f^2 \left(\frac{B_m}{f_1}\right)^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

Il guadagno che si ottiene con questa tecnica è dunque dato da:

$$\frac{\left(\frac{S}{N}\right)_{OUT}^{FM_PRE-ENFASI e DE-ENFASI}}{\left(\frac{S}{N}\right)_{OUT}^{FM_CLASSICO}} = \left(\frac{B_m}{f_1}\right)^2 \frac{1}{3}$$

Prestazioni
FM +
tecniche di
enfasi/de-
enfasi
(approssima-
zione)

In teoria, riducendo f_1 si aumentano a piacere le prestazioni. In pratica, non potrà utilizzare filtri con f_1 piccola a piacere, perché avrei difficoltà nella progettazione.

Esempio di standard. Trasmissione radio FM broadcast italiana:

- $\beta_f = 5$;
- $B_m = 15$ kHz;
- $f_1 = 2.1$ kHz;
- Guadagno = 17 = 12.3 dB.

8.6 Confronto tra sistemi di modulazione

Abbiamo a questo punto studiato:

- Vari schemi di rice-trasmissione AM (AM standard, AM-DSB-SC, AM-SSB);
- PM;
- FM e FM con pre-enfasi/de-enfasi.

Possiamo a questo punto confrontare tra di loro questi schemi di modulazione. I risultati andranno ancora confrontati col PCM.

Il confronto può essere fatto da diversi punti di vista. Qui consideriamo i tre più importanti: efficienza in potenza, efficienza in banda, facilità di implementazione. È ovvio che nessuna delle modulazioni è migliore delle altre da tutti i punti di vista. Ognuna avrà i suoi pregi e i suoi difetti e starà al progettista del sistema di telecomunicazioni scegliere lo schema di modulazione stabilendo il miglior trade-off caso per caso.

Efficienza in potenza. Un criterio per confrontare l'efficienza in potenza dei vari sistemi è quello di confrontare le rispettive curve di SNR. In figura 8.33 confrontiamo qualitativamente le prestazioni dei vari sistemi in termini di SNR, riferendoli alle prestazioni di un sistema di riferimento in banda base. PM non sarà considerato, in quanto ha le stesse prestazioni di FM.

Dalla figura si vede bene che:

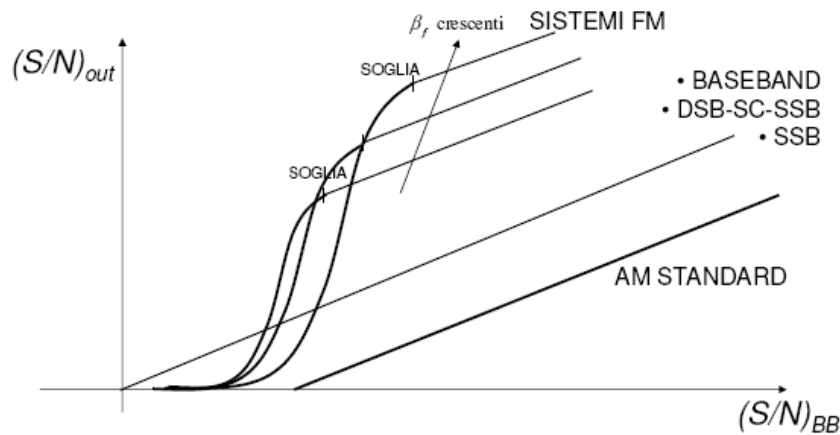


Figura 8.33: Confronto prestazioni sistemi FM/AM/banda base.

1. I sistemi FM sotto soglia non funzionano;
2. Per i sistemi FM la soglia aumenta per β_f crescenti;
3. Per i sistemi FM le prestazioni sopra-soglia migliorano aumentando β_f ;
4. Lavorando sopra soglia e con indice di modulazione elevato, FM ha prestazioni migliori di AM (tutte) e trasmissioni in banda base.

FM è quindi la modulazione con migliore efficienza in potenza. Viene largamente utilizzata in collegamenti critici dal punto di vista dell'utilizzo della potenza, come ad esempio i collegamenti via satellite.

AM standard è quella con la peggiore efficienza in potenza. Il motivo è che spreca potenza per trasmettere la portante.

Efficienza in banda. Lo schema di modulazione che utilizza meglio la banda è la AM-SSB. Questo sistema è ampiamente utilizzato in applicazioni critiche per l'utilizzo della banda, come trasmissione di voce su rete wireless e sistemi di comunicazione in aree congestionate.

FM è quella con peggiore efficienza in banda e non viene utilizzata in applicazioni critiche dal punto di vista dell'occupazione di banda.

Facilità di implementazione. La struttura di ricevitore più semplice è quella di AM standard (che utilizza un rivelatore di inviluppo, come visto nel capitolo su AM).

I ricevitori FM sono anche semplici da implementare (ma meno di quelli per AM standard).

AM-SSB richiede una demodulazione sincrona e quindi ricevitori con struttura molto più complicata.

8.7 Conclusione

Conclusioni Questo capitolo ha trattato le modulazioni angolari FM e PM. Si tratta di modulazioni in cui l'informazione da trasmettere è affidata alla fase della portante di un segnale sinusoidale.

Di tutti i risultati che si sono ottenuti, il più importante è questo: tali modulazioni consentono una migliore immunità al rumore rispetto alle modulazioni AM, a costo di una maggiore occupazione di banda e di una maggiore complessità di trasmettitore e ricevitore.

Chi fosse interessato a visualizzare un po' di spettri di segnali modulati PM/FM, Esempi può scaricare dal sito del corso i programmi Matlab contenuti nella sezione "*Esempi di spettri relativi a modulazioni analogiche PM o FM*". Matlab Con questi programmi si possono visualizzare sia gli spettri di segnali modulati FM/PM con modulante sinusoidale, che con modulante preso da file audio. Sarebbe un esperimento interessante per capire meglio le cose.

8.8 Riassunto formule principali sulle modulazioni angolari

Si riportano in questa sezione le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi, e sono tutte presenti anche sul formulario del corso (che si può scaricare dal sito del corso). Si ricorda che il formulario è l'unico materiale che è consentito tenere in sede d'esame.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- Definizioni:

$$\begin{aligned}
 s(t) &= A_c \cos[2\pi f_c t + \theta(t)] \\
 \theta(t) &= D_p m(t) \\
 \theta(t) &= D_f \int_{-\infty}^t m(\tau) d\tau \\
 f_d(t) &= f_i(t) - f_c = \frac{1}{2\pi} \frac{d\theta}{dt} \\
 \Delta F &= \max \{f_d(t)\} \\
 \Delta\theta &= \max \{\theta(t)\} = D_p V_p \quad \text{dove: } V_p = \max \{m(t)\} \\
 \beta_p &= \Delta\theta = D_p V_p \quad \text{Indice di modulazione PM} \\
 \beta_f &= \frac{\Delta F}{B_m} \quad \text{Indice di modulazione FM}
 \end{aligned}$$

- Formula di Carson:

$$B_T = 2(\beta + 1)f_m$$

- Prestazioni PM:

$$\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} \beta_p^2 < \frac{m^2(t)}{V_p^2} > = \left(\frac{S}{N}\right)_{BB} \beta_p^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

- Prestazioni FM:

$$\left(\frac{S}{N}\right)_{out} = \frac{P_{RX}}{N_0 B_m} 3 \beta_f^2 < \left(\frac{m(t)}{V_p}\right)^2 > = \left(\frac{S}{N}\right)_{BB} 3 \beta_f^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

- Prestazioni FM + tecniche di enfasi/de-enfasi (approssimazione):

$$\left(\frac{S}{N}\right)_{out} \approx \left(\frac{S}{N}\right)_{BB} \beta_f^2 \left(\frac{B_m}{f_1}\right)^2 < \left(\frac{m(t)}{V_p}\right)^2 >$$

Capitolo 9

Sistemi PCM

Questo capitolo tratta i sistemi PCM. Il PCM (*Pulse Code Modulation*) è il sistema più usato per trasmettere un segnale analogico in modo digitale, ed è di fatto un altro modo per indicare il meccanismo di conversione Analogico-Digitale (A/D) e Digitale-Analogico (D/A).

Il capitolo si focalizzerà sulla quantizzazione. Verranno trattate:

1. Quantizzazione uniforme;
2. Quantizzazione non uniforme.

Per approfondire questi argomenti si consiglia di consultare [5] oppure [7].

9.1 Introduzione

La figura 9.1 mostra lo schema a blocchi di un sistema PCM. Si tratta del sistema di trasmissione che nella prima lezione avevamo chiamato "misto analogico-digitale". Si ricorda che lo scopo di un sistema di questo tipo è la trasmissione affidabile del segnale analogico. Un sistema PCM progettato bene deve dunque fornire in uscita un segnale $V_{out}(t)$ che sia il più possibile simile al segnale di ingresso $V_{in}(t)$.

Sistemi PCM

In questo capitolo ci concentreremo sui blocchi *Convertitore A/D* (convertitore analogico-digitale) e *Convertitore D/A* (convertitore digitale-analogico). Le operazioni base della conversione A/D sono le seguenti:

Operazioni
base
conversione
A/D

1. Campionamento nel tempo;
2. Quantizzazione sulle ampiezze;
3. Codifica su uno stream seriale di bit.

Questa suddivisione serve soprattutto dal punto di vista teorico. Nella pratica, nella stragrande maggioranza dei casi, le tre operazioni sono svolte da un unico chip (convertitore A/D). In questo corso ci focalizzeremo principalmente sul punto 2, cioè sulla quantizzazione. Il punto 1 è infatti già stato trattato nel corso di Teoria dei Segnali. Il lettore interessato a ripassarlo può consultare [1]. Per il punto 3, considereremo solo un caso molto semplice, in quanto la sua trattazione generale è complessa e esula dagli scopi di questo corso. Chi fosse interessato ad approfondire l'argomento 3 può far riferimento a [7].

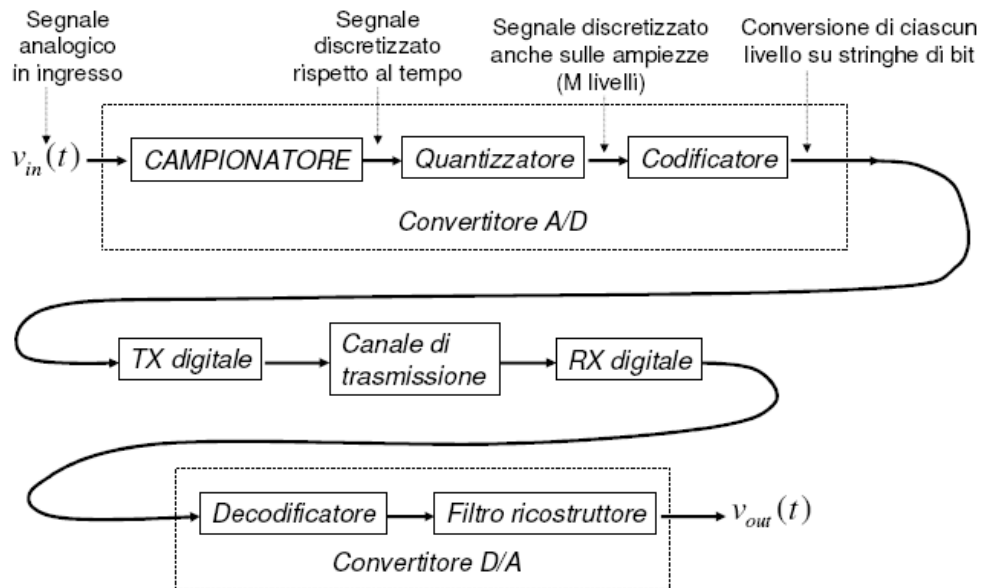


Figura 9.1: Schema a blocchi sistema PCM (misto analogico-digitale).

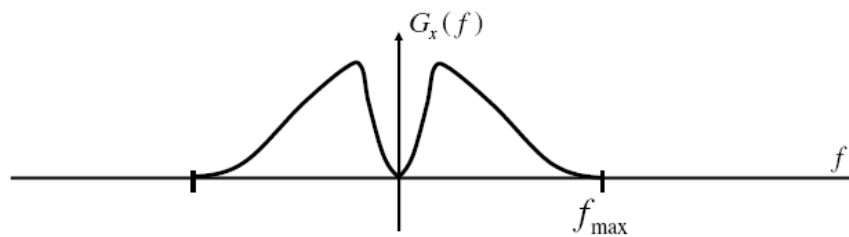


Figura 9.2: Esempio di spettro di segnale da campionare. Lo spettro ha frequenza massima pari a f_{max} . Il segnale deve essere quindi campionato con una frequenza di campionamento $f_c \geq 2 f_{max}$.

Campionamento e quantizzazione

Teorema del
campionamento

Il primo blocco del convertitore A/D è un campionatore. Ricordiamo quindi brevemente il teorema del campionamento. Si desidera campionare un segnale il cui spettro abbia frequenza massima pari a f_{max} (vedi figura 9.2). In tal caso devo campionare almeno a:

$$f_c \geq 2 f_{max}$$

Per ciascun istante di campionamento, si deve poi quantizzare il segnale. Si deve cioè mappare la tensione corrispondente V_{in} (che assume valori continui) su numero finito di tensioni V_{out} , detti "livelli".

Differenza
tra campio-
namento e
quantizzazio-
ne

Occorre essere ben consapevoli della differenza tra campionamento e quantizzazione:

Campionamento: significa passare da $V(t)$ continuo alla sequenza di numeri reali $V(iT_c)$. Questa operazione non comporta perdita di informazione, almeno in teoria.

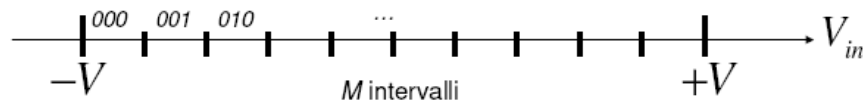


Figura 9.3: Quantizzatore uniforme. Ho un intervallo $[-V, +V]$ suddiviso in M sotto-intervalli tutti della stessa lunghezza. Ad ogni sotto-intervallo viene associata una stringa di bit. A ogni valore campionato V_{in} viene poi associata la stringa di bit relativa all'intervallo a cui appartiene V_{in} .



Figura 9.4: Convertitore A/D seguito da convertitore D/A.

Quantizzazione: significa passare da ciascuno dei numeri reali $V(iT_c)$ ad un insieme discreto e finito di ampiezze. Questa operazione comporta sempre una perdita di informazione.

9.2 Quantizzazione uniforme

Iniziamo a esaminare il caso più semplice e più comune, la *quantizzazione uniforme*. Consideriamo la figura 9.3. Supponiamo che il segnale di ingresso campionato V_{in} sia compreso nell'intervallo $[-V, +V]$. Dal lato A/D, il range $[-V, +V]$ viene suddiviso in M intervalli uguali, e a ciascun intervallo viene associata una stringa di bit. Per ragioni pratiche, M è sempre scelto pari ad una potenza di 2:

$$M \triangleq 2^{n_{bit}}$$

dove n_{bit} rappresenta il numero di bit necessario a "contare" in binario il numero M di intervalli. A ogni valore campionato V_{in} viene poi associata la stringa di bit relativa all'intervallo a cui appartiene V_{in} .

Dal lato D/A, per ciascuna stringa di bit si ricostruisce un valore di tensione. Per ragioni che saranno chiare tra poco, si sceglie come valore di tensione il punto centrale dell'intervallo di partenza.

Si consideri ora l'effetto complessivo della cascata di un A/D ed un D/A, assumendo ideale il sistema di trasmissione (vedi figura 9.4). Ipotizziamo di lavorare con:

- V_{in} nel range $[-8, +8]$;
- Quantizzazione su $M = 8$ livelli.

Effetto
complessivo
di A/D e
D/A

La cascata di una conversione A/D e D/A è equivalente ad un sistema (non lineare, senza memoria) che opera la trasformazione ingresso-uscita rappresentata in figura 9.5. Si noti che un intero intervallo in ingresso viene "mappato" in uscita su un singolo valore. Ad esempio, se V_{in} appartiene a $[0, 2]$, ne deriva che $V_{out} = 1$. Conseguentemente, V_{in}

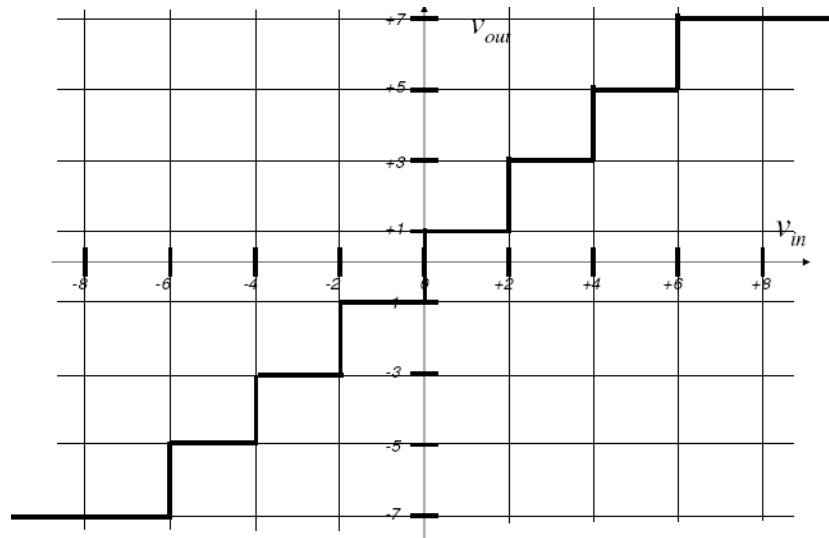


Figura 9.5: Caratteristica ingresso-uscita del sistema di figura 9.4, cioè della cascata di un convertitore A/D e di un convertitore D/A.

sarà sempre (salvo casi eccezionali) diverso da V_{out} .

Mentre il campionamento non introduce perdita di informazione (almeno in teoria), **la quantizzazione introduce sempre un errore di quantizzazione** che va calcolato e tenuto sotto controllo.

9.2.1 Calcolo del rapporto segnale-rumore di quantizzazione

Vogliamo ora quantificare il degrado delle prestazioni dovuto alla cascata di A/D e D/A, dovuta al solo effetto della quantizzazione¹. Questa operazione è fondamentale, in quanto permetterà di ottenere dei criteri per dimensionare il convertitore A/D. In particolare, si troverà un criterio per scegliere il numero di livelli M .

Definizione di errore di quantizzazione

Facendo riferimento alla figura 9.5, definiamo l'errore di quantizzazione come:

$$e_q \triangleq V_{in} - V_{out}$$

Se chiamiamo Δ la lunghezza di ciascun intervallo del quantizzatore, avremo che per il quantizzatore uniforme:

$$\max |e_q| = \frac{\Delta}{2}$$

Per quanto riguarda l'errore di quantizzazione, possiamo allora modellare il sistema di figura 9.4 come in figura 9.6. Nell'ipotesi di utilizzare un numero elevato di livelli di quantizzazione, e_q è modellabile come un processo casuale stazionario bianco e indipendente dal segnale di ingresso. Queste caratteristiche statistiche sono analoghe a quelle del rumore termico, per tale ragione possiamo pensare di modellare l'errore di quantizzazione come un rumore, che nel seguito chiameremo **rumore di quantizzazione**.

Calcolo rapporto segnale-rumore di quantizzazione

Definiamo il rapporto segnale-rumore di quantizzazione, che nel seguito indicheremo

¹Non si considerano cioè gli errori di trasmissione sul canale.

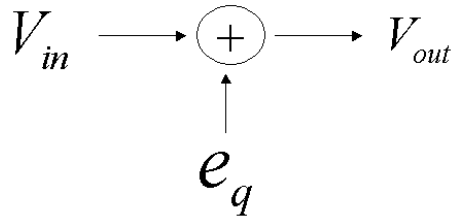


Figura 9.6: Equivalente di cascata di convertitore A/D e D/A, dal punto di vista dell'errore di quantizzazione.

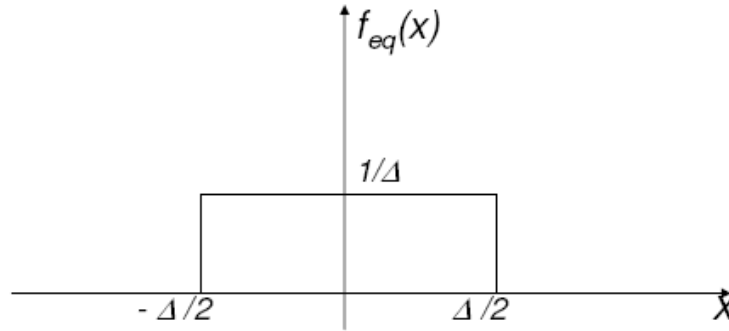


Figura 9.7: Densità di probabilità dell'errore di quantizzazione con quantizzatore uniforme e segnale di ingresso con densità di probabilità uniforme con la stessa dinamica del quantizzatore.

con $(S/N)_Q$, come la potenza media del segnale utile divisa per la potenza media dell'errore di quantizzazione. Per calcolarlo serve fare le seguenti ipotesi:

1. Quantizzazione uniforme (gradini tutti uguali) in $[-V, +V]$ a M livelli: $\Delta = \frac{2V}{M}$.
2. Segnale con densità di probabilità uniforme all'interno di un certo range uguale a quello del quantizzatore: $[-V, +V]$.
3. Segnale a valor medio nullo.

Per le ipotesi fatte, si avrà che $e_q = V_{in} - V_{out}$ avrà una distribuzione uniforme tra $-\Delta/2$ e $+\Delta/2$, essendo Δ la larghezza dell'intervallo (vedi figura 9.7). Notiamo che, su un certo intervallo, l'errore di quantizzazione è a valor medio nullo. Inoltre, la statistica è la stessa su qualunque intervallo.

Possiamo allora valutare il rapporto segnale-rumore di quantizzazione come rapporto delle varianze:

$$\left(\frac{S}{N}\right)_Q = \frac{\sigma_{V_{in}}^2}{\sigma_{e_q}^2} \quad (9.1)$$

Facendo riferimento alla figura 9.7 e ricordando che $\Delta = 2V/M$, si trova che la varianza

dell'errore di quantizzazione vale:

$$\begin{aligned}
 \sigma_{e_q}^2 &= E[e_q^2] \\
 &= \int_{-\infty}^{+\infty} x^2 f_{eq}(x) dx \\
 &= \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} x^2 \frac{1}{\Delta} dx \\
 &= \frac{1}{\Delta} \left[\frac{x^3}{3} \right]_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \\
 &= \frac{\Delta^2}{12} \\
 &= \frac{V^2}{3M^2}
 \end{aligned} \tag{9.2}$$

A questo punto è chiaro perché nella conversione D/A si è scelto di prendere come valore di tensione in uscita il punto centrale dell'intervallo di partenza. Questa scelta permette infatti di minimizzare la potenza dell'errore di quantizzazione.

Ricordando poi che il segnale di ingresso ha densità di probabilità uniforme in $[-V, +V]$, si trova che la sua varianza vale:

$$\begin{aligned}
 \sigma_{V_{in}}^2 &= E[V_{in}^2] \\
 &= \int_{-V}^{+V} x^2 \frac{1}{2V} dx \\
 &= \frac{1}{2V} \left[\frac{x^3}{3} \right]_{-V}^{+V} \\
 &= \frac{V^2}{3}
 \end{aligned} \tag{9.3}$$

Possiamo ora sostituire le espressioni (9.2) e (9.3) nella (9.1) e trovare il rapporto segnale-rumore relativo alla sola quantizzazione:

$(S/N)_Q$ con
quantizzato-
re
uniforme

$$\left(\frac{S}{N} \right)_Q = M^2 \tag{9.4}$$

Mappando i valori di tensione su parole digitali di n bit si ha che $M = 2^n$. Allora:

$$\left(\frac{S}{N} \right)_Q = 2^{2n}$$

Passando ai dB:

$$\begin{aligned}
 \left(\frac{S}{N} \right)_Q \Big|_{dB} &= 10 \log_{10} (2^{2n}) \\
 &= 2n \cdot 10 \log_{10}(2)
 \end{aligned}$$

$(S/N)_Q$ in dB con quantizzatore uniforme che svolgendo i conti diventa:

$$\left(\frac{S}{N} \right)_Q \Big|_{dB} \approx 6n \text{ [dB]} \quad (9.5)$$

In generale succede che, anche in situazioni più complesse, aumentare di 1 bit il numero di livelli, fa aumentare di 6 dB le prestazioni in termini di $(S/N)_Q$. Questo famoso risultato è solitamente indicato come "6-dB law". Indica il fatto che ogni bit di quantizzazione in più migliora di 6 dB le prestazioni.

Esempio 9.1. PCM telefonico Segnale vocale, qualità "telefonica". Si "forza" una occupazione spettrale da 300 a 3400 Hz.

- Frequenza di campionamento minima: $f_c = 2 \cdot 3400 = 6.8 \text{ kHz}$. Lasciando un margine, è stato standardizzato: $f_c = 8 \text{ kHz}$.
- Numero di bit: per avere valori accettabili: $n = 8$ (256 livelli). Il bit-rate complessivo è dunque dato da $R = 8f_c = 64 \text{ kbit/s}$.
- Il risultante valore di $(S/N)_Q = 48 \text{ dB}$.

Nella realtà, per il PCM telefonico non si usa una quantizzazione uniforme (vedi prossimo paragrafo sulla quantizzazione non uniforme). A causa della sua enorme diffusione, il PCM telefonico ha "standardizzato" anche in altri contesti i seguenti due parametri:

1. Campionamento a 8 kHz (periodicità 125 μs).
2. Quantizzazione a 8 bit.

■

Esempio 9.2. Segnale musicale Standard CD audio: segnale fino a 20 kHz.

- È stato scelto: $f_c = 44.1 \text{ kHz}$.
- Quantizzazione: $n = 16 \text{ bit}$ (65000 livelli).
- Bit-rate complessivo = 705.6 kbit/s (per canale, si raddoppia per l'effetto stereofonico: 1411.2 kbit/s = 1.4 Mbit/s).
- $(S/N)_Q = 96 \text{ dB}$. Lo standard base per il PCM audio prevede una quantizzazione uniforme.

Notare che è necessario, senza compressione, un bit-rate estremamente elevato. Proviamo infatti a calcolare il bit-rate netto:

$$\text{Bit-rate} = 44.1 \text{ [ksamples/s]} \cdot 32 \text{ [bit/sample]} = 1.41 \text{ [Mbit/s]}$$

Di fatto, tutte le moderne soluzioni che prevedono lo scambio di segnali musicali su rete, utilizzano qualche forma di compressione. Ad esempio, con il formato MP3 si ottengono compressioni elevate, tipicamente con fattori da 12 a 16.

Consideriamo la quantità di dati memorizzati su un CD audio da 80 minuti:

$$\begin{aligned} & 44.1 \left[\frac{\text{ksample}}{\text{sec}} \right] \cdot 32 \left[\frac{\text{bit}}{\text{sample}} \right] \cdot 60 \text{ [sec]} \cdot 80 \text{ [min]} = \\ & = 6.774 \cdot 10^9 \text{ [bit]} = 846.7 \cdot 10^6 \text{ [byte]} \approx 807.5 \text{ [Mbyte]} \end{aligned}$$

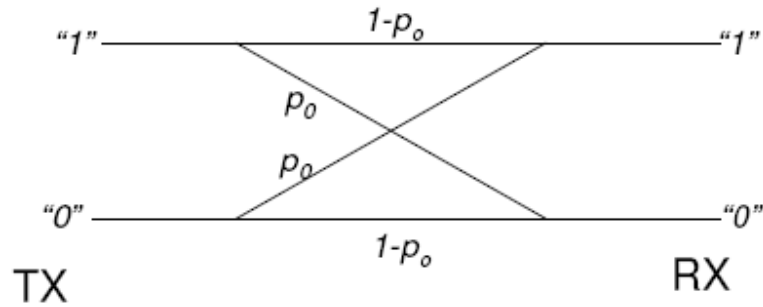


Figura 9.8: Rappresentazione schematica di un BSC.

Curiosità: tipicamente un CD-R è specificato per 80 minuti di audio, e per 700 Mbyte di dati. Abbiamo appena visto che a 80 minuti corrispondono circa 800 Mbyte. La discrepanza si spiega con il fatto che per i dati si usa uno standard diverso che per l'audio PCM. In particolare, si usa lo standard CD-ROM, che usa una protezione agli errori più "forte" rispetto allo standard CD-AUDIO, e dunque diminuisce lo "spazio" equivalente disponibile. ■

9.2.2 Prestazioni del PCM in presenza di errori sul canale

L'espressione (9.4) mi dice quanto vale il rapporto segnale-rumore dovuto alla sola quantizzazione. Tuttavia, un'altra fonte di degrado delle prestazioni di un sistema PCM riguarda gli errori sui bit durante la trasmissione. In questo contesto, ci interessa caratterizzare il sistema di trasmissione digitale nel modo più semplice possibile. Questo è possibile tramite il modello del *canale binario simmetrico* (BSC, Binary Symmetric Channel).

Cenni al canale binario simmetrico (BSC)

Un sistema di trasmissione digitale è essenzialmente caratterizzabile, dal punto di vista sistemistico, come un sistema che trasmette dei bit con una certa probabilità di errore. Rivediamo dunque brevemente il modello del canale binario simmetrico, già introdotto nel corso di Teoria dei Segnali.

La figura 9.8 mostra una rappresentazione schematica del BSC. p_0 è detta *probabilità di transizione* del canale BSC, ed è definita come:

$$p_0 \triangleq P(0 \text{ RX} | 1 \text{ TX}) = P(1 \text{ RX} | 0 \text{ TX})$$

Calcolo
probabilità
di errore sul
bit per
canale BSC

La probabilità di errore complessiva sul bit $P_b(e)$ può essere calcolata col teorema della probabilità totale nel seguente modo:

$$P_b(e) = P(e | 1 \text{ TX}) P(1 \text{ TX}) + P(e | 0 \text{ TX}) P(0 \text{ TX})$$

Se anche la sorgente è simmetrica, cioè se emette in maniera equiprobabile "1" e "0", allora:

$$P(1 \text{ TX}) = P(0 \text{ TX}) = \frac{1}{2}$$

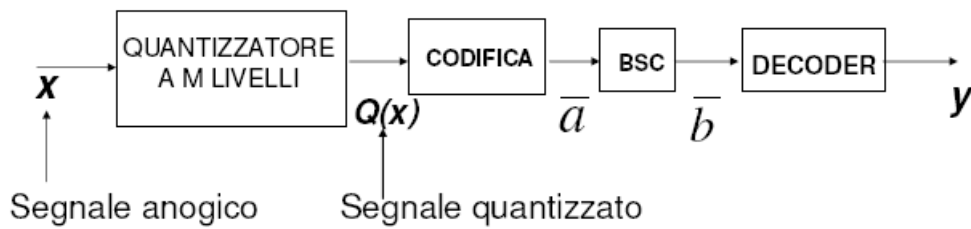


Figura 9.9: Trasmissione di segnale quantizzato su canale BSC.

e conseguentemente:

$$P_b(e) = p_0$$

La probabilità di errore complessiva sul bit è anche indicata come P_e oppure *BER* (*Bit Error Rate*). Di fatto, la probabilità di errore sul bit è l'unico parametro che utilizzeremo nel seguito di questo capitolo per caratterizzare il sistema di trasmissione digitale.

Le modulazioni digitali saranno trattate in dettaglio nei corsi di Trasmissione Numerica, Laboratorio di Telecomunicazioni e in altri corsi della Laurea Specialistica. Il lettore interessato può fare riferimento a [3] oppure a [4].

Calcolo del rapporto segnale-rumore dovuto ai soli errori sui bit sul canale di trasmissione

Vogliamo adesso calcolare il valore di S/N risultante dalla trasmissione di uno stream di bit generato con un sistema PCM su un canale BSC con una certa probabilità di errore sul bit P_e . La deduzione della formula di S/N dovuto agli errori sui bit è abbastanza lunga, ma porterà ad un risultato molto semplice e compatto. La dimostrazione è utile per capire più nel dettaglio alcuni aspetti importanti del PCM.

Facciamo riferimento alla figura 9.9. Iniziamo a dare un'espressione che lega il segnale quantizzato $Q(x)$ (cioè quello su M livelli e usando n bit per campione) e la corrispondente n -upla di bit che saranno da inviare sul sistema di trasmissione digitale. Indichiamo la n -upla di bit come:

$$\bar{a} \triangleq [a_1, \dots, a_n]$$

Per comodità, assumiamo che ciascun elemento della stringa valga $+1$ o -1 (e non "0" e "1"). La codifica più semplice è data da:

$$Q(x) = V \cdot \sum_{j=1}^n a_j \left(\frac{1}{2}\right)^j \quad (9.6)$$

La formula (9.6) va compresa bene. Consideriamo dunque la figura 9.10. La figura 9.10 riporta il caso $M = 8$, $n = 3$, $\Delta = 2$, $[-V, +V] = [-8, +8]$. Si supponga di contare i livelli di uscita in decimale, e poi di convertirli in binario puro: Verifica della formula (9.6)

$$\begin{aligned} 0 &\rightarrow [000] \\ 1 &\rightarrow [001] \\ &\dots \\ 7 &\rightarrow [111] \end{aligned}$$

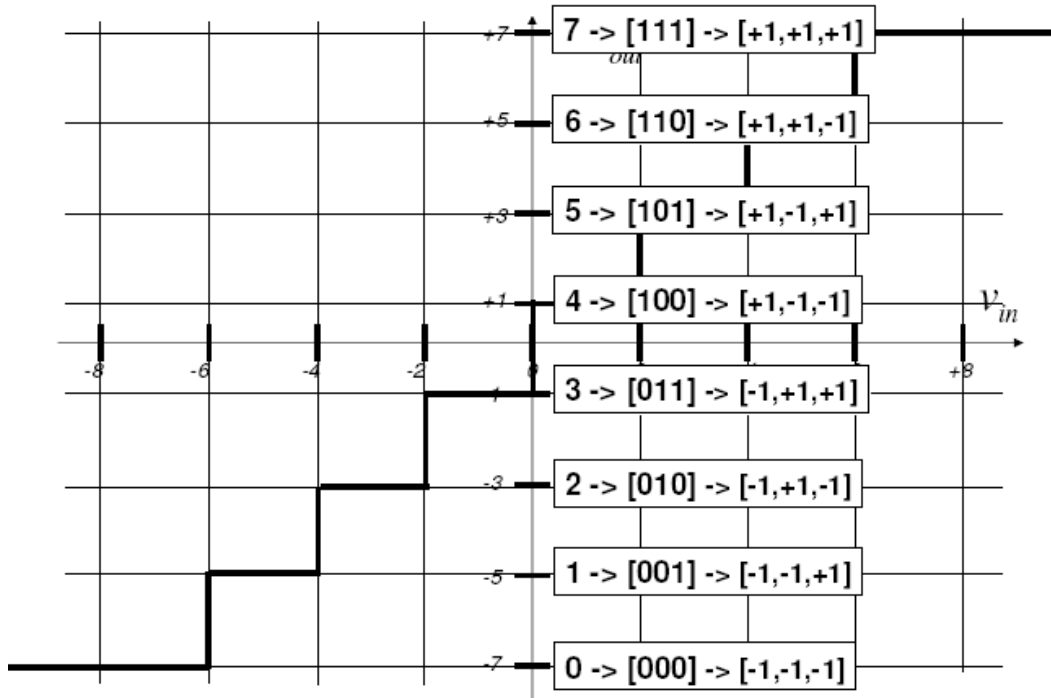


Figura 9.10: Caratteristica ingresso-uscita di un quantizzatore uniforme con $M = 8$ e $[-V, V] = [-8, +8]$. Nei riquadri sono specificate le codifiche assegnate ai vari intervalli.

Si convertano poi gli "1" in +1 e gli "0" in -1. Considerando un valore di n generico, se $a = [1, 1, \dots, 1]$, la (9.6) mi dà:²

$$\begin{aligned}
 Q(x) &= V \cdot \sum_{j=1}^n \left(\frac{1}{2}\right)^j \\
 &= V \left[1 - \left(\frac{1}{2}\right)^n \right] \\
 &= V - \frac{V}{2^n}
 \end{aligned} \tag{9.7}$$

In generale, ricordando che $\Delta = \frac{2V}{M}$ e che $M = 2^n$, si ottiene che:

$$\begin{aligned}
 \frac{V}{2^n} &= \frac{V}{M} \\
 &= \frac{\Delta}{2}
 \end{aligned} \tag{9.8}$$

Nel caso particolare di figura 9.10:

- $\Delta = 2$;

²Nei calcoli che seguono si usa questa ridotta della serie geometrica:

$$\sum_{k=k_0}^{k_1} z^k = \frac{z^{k_0} - z^{k_1+1}}{1 - z}$$

- $V = 8$;
- Il livello massimo si può calcolare con la (9.7) e la (9.8) e vale: $8 - 8/2^3 = 7$. Questo è in effetti il valore che si vede anche in figura.

Per gli altri livelli si possono trovare i valori con conti analoghi. Il lettore interessato può provare a svolgere i conti e troverà esattamente i valori mostrati in figura 9.10.

Il vettore di bit \bar{a} passa sul canale BSC, dando luogo ad un vettore \bar{b} . Tuttavia, a causa degli errori sul canale di trasmissione digitale, si avrà in generale $\bar{b} \neq \bar{a}$. È questo l'effetto che ci proponiamo di studiare. Il segnale ricostruito sarà poi semplicemente:

$$y = V \cdot \sum_{j=1}^n b_j \left(\frac{1}{2}\right)^j \quad (9.9)$$

Dobbiamo allora confrontare le espressioni di $Q(x)$ e di y per valutare il "rumore" dovuto agli errori sui bit. Questo "rumore" lo definisco come:

$$e_b \triangleq y - Q(x)$$

Calcolo
potenza
rumore
dovuto a
errori sui bit

Usando la (9.6), la (9.9) e la linearità del valor medio, si trova che il valore quadratico medio di e_b vale:

$$\begin{aligned} E[e_b^2] &= E[(y - Q(x))^2] \\ &= E\left[V^2 \left(\sum_{j=1}^n (b_j - a_j) \left(\frac{1}{2}\right)^j\right)^2\right] \\ &= V^2 \cdot E\left[\sum_{j=1}^n (b_j - a_j) 2^{-j} \cdot \sum_{k=1}^n (b_k - a_k) 2^{-k}\right] \\ &= V^2 \cdot \sum_{j=1}^n \sum_{k=1}^n 2^{-j-k} (E[(b_j - a_j)(b_k - a_k)]) \\ &= V^2 \cdot \sum_{j=1}^n \sum_{k=1}^n 2^{-j-k} (E[b_j b_k] - E[a_j b_k] - E[b_j a_k] + E[a_j a_k]) \quad (9.10) \end{aligned}$$

Per tutti i termini con j diverso da k , grazie al fatto che le medie sono nulle:

$$E[b_j] = E[b_k] = E[a_j] = E[a_k] = 0$$

e grazie all'indipendenza statistica, si ha:

$$\begin{aligned} E[b_j b_k] &= E[b_j] \cdot E[b_k] = 0 \\ E[a_j a_k] &= E[a_j] \cdot E[a_k] = 0 \\ E[b_j a_k] &= E[b_j] \cdot E[a_k] = 0 \end{aligned}$$

Invece, per $j = k$ si ottiene:

$$E[a_j a_j] = E[b_j b_j] = 1 \quad (9.11)$$

a_j	b_j	$P(a_j b_j)$	$a_j b_j$
+1	+1	$\frac{1}{2}(1 - P_e)$	+1
+1	-1	$\frac{1}{2}P_e$	-1
-1	+1	$\frac{1}{2}P_e$	-1
-1	-1	$\frac{1}{2}(1 - P_e)$	+1

Tabella 9.1: Valori per risolvere i termini misti della (9.10) con $j = k$.

Allora nella (9.10) rimane solo una singola sommatoria dove $j = k$ che va da 1 a n . Inoltre sono importanti, per $j = k$ i termini misti:

$$E[a_j b_j]$$

Dobbiamo esaminare 4 combinazioni. Usando i valori della tabella 9.1 si trova:

$$\begin{aligned} E[a_j b_j] &= (+1) \frac{1}{2} (1 - P_e) + (-1) \frac{1}{2} P_e + (-1) \frac{1}{2} P_e + (+1) \frac{1}{2} (1 - P_e) \\ &= (1 - P_e - P_e) \\ &= 1 - 2P_e \end{aligned} \quad (9.12)$$

Mettiamo ora insieme i risultati appena trovati. Sostituiamo la (9.11) e la (9.12) nella (9.10), ricordando che gli altri termini misti valgono 0. Otteniamo:

$$\begin{aligned} E[e_b^2] &= V^2 \cdot \sum_{j=1}^n \sum_{k=1}^n 2^{-j-k} (E[b_j b_k] - E[a_j b_k] - E[b_j a_k] + E[a_j a_k]) \\ &= V^2 \cdot \sum_{j=1}^n 2^{-2j} (E[b_j^2] - 2E[a_j b_j] + E[a_j^2]) \\ &= V^2 \cdot \sum_{j=1}^n 2^{-2j} (1 - 2 + 4P_e + 1) \\ &= 4V^2 P_e \cdot \sum_{j=1}^n \left(\frac{1}{4}\right)^j \end{aligned}$$

Utilizzando la formula per la serie geometrica ricordata a pagina 188 si ottiene infine:

$$E[e_b^2] = \frac{4}{3} V^2 P_e \frac{M^2 - 1}{M^2} \quad (9.13)$$

Calcolo
potenza di
segnale utile

Consideriamo ora la potenza di segnale per arrivare alla formula sul rapporto segnale-rumore. Ricordando che V_{in} ha densità di probabilità uniforme in $[-V, +V]$ si trova:

$$\begin{aligned} E[V_{in}^2] &= \int_{-V}^{+V} x^2 \frac{1}{2V} dx \\ &= \frac{1}{2V} \left[\frac{x^3}{3} \right]_{-V}^{+V} \\ &= \frac{V^2}{3} \end{aligned} \quad (9.14)$$

Mettiamo ora insieme i due risultati (9.13) e (9.14), calcolando il rapporto segnale-rumore dovuto agli errori sui bit:

$$\begin{aligned} \left(\frac{S}{N}\right)_e &\triangleq \frac{E[V_{in}^2]}{E[e_b^2]} \\ &= \frac{\frac{V^2}{3}}{\frac{4}{3} V^2 P_e \frac{M^2-1}{M^2}} \\ &= \frac{1}{4 P_e \frac{M^2-1}{M^2}} \end{aligned}$$

Notare che, almeno per $M \gg 1$ (come succede quasi sempre nelle applicazioni), si ha in sostanza:

$$\boxed{\left(\frac{S}{N}\right)_e \approx \frac{1}{4P_e}} \quad (9.15) \quad \begin{array}{l} \text{Rapporto} \\ \text{segnale-} \\ \text{rumore} \\ \text{dovuto ai soli} \\ \text{errori sui bit} \end{array}$$

Questo risultato è molto importante: il rapporto segnale-rumore "di errore" dipende solo da P_e .

Riassumendo: abbiamo trovato finora due cause di degrado delle prestazioni nei sistemi PCM: Riassunto

1. L'effetto della quantizzazione: $\left(\frac{S}{N}\right)_Q = M^2$;
2. L'effetto degli errori sui bit: $\left(\frac{S}{N}\right)_e \approx \frac{1}{4P_e}$.

Calcolo del rapporto segnale-rumore complessivo dovuto alla quantizzazione e agli errori sui bit

Vogliamo ora riportare le due formule ad un'unica formula che tenga conto dell'insieme dei due effetti. Chiamiamo l'errore totale e_{out} . È ragionevole supporre che l'errore introdotto dalla quantizzazione e quello introdotto dagli errori di trasmissione siano indipendenti, essendo dovuti a cause diverse. Possiamo allora sommarne le varianze e trovare così la varianza dell'errore totale. Dalla (9.13) che abbiamo appena ricavato e dalla (9.2) di pagina 184, si trova che:

$$\begin{aligned} E[e_{out}^2] &= E[e_b^2] + E[e_q^2] \\ &= \frac{4}{3} V^2 P_e \frac{M^2-1}{M^2} + \frac{V^2}{3M^2} \\ &= \frac{V^2}{3M^2} [4P_e(M^2-1) + 1] \end{aligned}$$

Ricordando inoltre la (9.14), cioè che:

$$E[V_{in}^2] = \frac{V^2}{3}$$

Facendo $\frac{E[V_{in}^2]}{E[e_{out}^2]}$ si ottiene il rapporto segnale-rumore totale:

$$\boxed{\left(\frac{S}{N}\right)_{out} = \frac{M^2}{1 + 4P_e(M^2-1)}} \quad (9.16) \quad \begin{array}{l} (S/N)_{out} \\ \text{complessivo} \\ \text{sistemi PCM} \end{array}$$

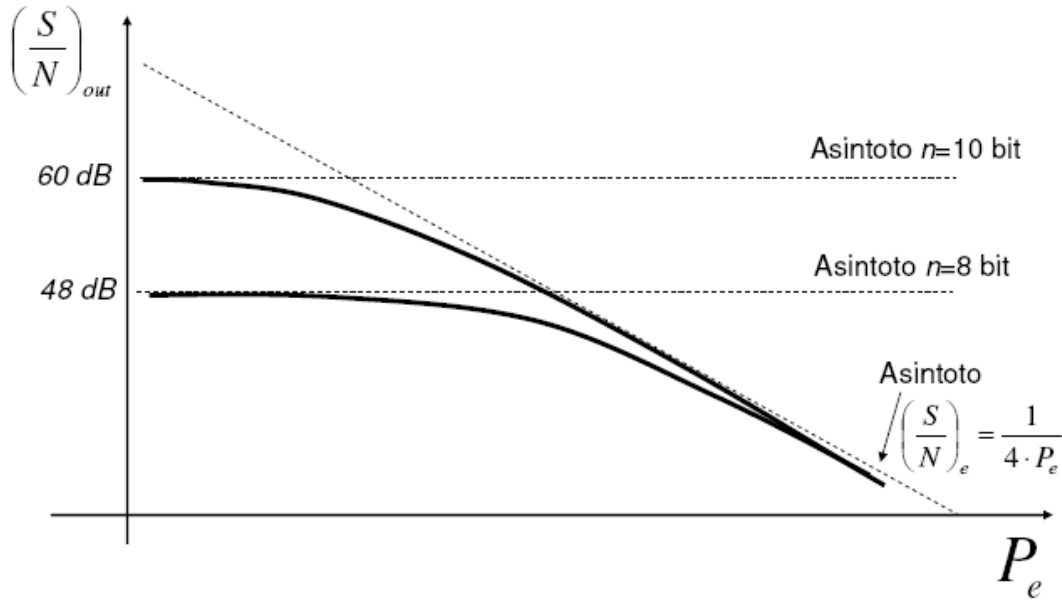


Figura 9.11: Andamento di $(S/N)_{out}$ complessivo di sistema PCM, dovuto a quantizzazione e errori sul canale, in funzione della probabilità di errore sul bit P_e . Sono considerati i due casi $n = 8$ bit e $n = 10$ bit. Si osservi come per bassi valori di P_e le prestazioni dipendano solo da n , mentre per elevati valori di P_e dipendono solo da P_e .

Soglia

sistemi PCM

La figura 9.11 presenta due curve di $(S/N)_{out}$ in funzione di P_e . Una curva è per $n = 8$ bit, l'altra per $n = 10$ bit. Osservando le due curve si nota che i sistemi PCM tendono a comportarsi "a soglia" rispetto alla probabilità di errore sul bit, cioè:

1. Per una P_e inferiore ad una soglia critica, le prestazioni non dipendono da P_e e inoltre coincidono con quelle dovute alla sola quantizzazione.
2. Per una P_e superiore ad una soglia critica, le prestazioni tendono a dipendere solo da P_e e inoltre crollano drasticamente al crescere di P_e . Sostanzialmente, in queste condizioni in sistema PCM non funziona.

Definiamo ora la probabilità di errore di soglia P_e^* come il valore di P_e che dà un peggioramento di 3 dB (1/2 in lineare) rispetto al $(S/N)_{out}$ massimo. Considerando la figura 9.12 si ottiene:

$$\left(\frac{S}{N} \right)_{out} \Big|_{P_e^*} = \frac{1}{2} \left(\frac{S}{N} \right)_Q$$

cioè:

$$\frac{M^2}{1 + 4(M^2 - 1)P_e^*} = \frac{1}{2} M^2$$

Probabilità
di errore di
soglia per
sistemi PCM

e infine:

$$P_e^* = \frac{1}{4(M^2 - 1)} \quad (9.17)$$

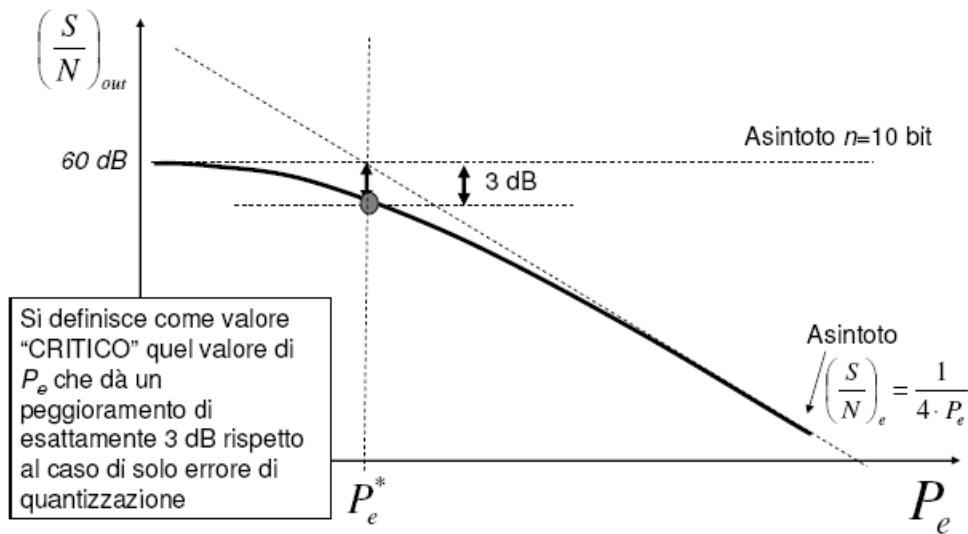


Figura 9.12: Definizione di probabilità di errore di soglia per sistemi PCM.

Si noti che il valore della soglia diminuisce all'aumentare del numero di livelli M . Questo si vede anche dalla figura 9.11.

Il comportamento "a soglia" dei sistemi PCM, è tipico di gran parte dei sistemi digitali. Tendono cioè a funzionare perfettamente fino a un certo valore di soglia, al di là del quale le prestazioni crollano drasticamente. In conclusione, un sistema PCM è detto:

- "Sopra soglia" se $P_e < P_e^*$;
- "Sotto soglia" se $P_e > P_e^*$.

Valori tipici di P_e^* :

- 8 bit $\Rightarrow 3.8 \cdot 10^{-6}$.
- 16 bit $\Rightarrow 5.8 \cdot 10^{-11}$.

9.2.3 Confronto con i sistemi analogici

Dal momento che lo scopo dei sistemi PCM è la trasmissione affidabile di un segnale analogico, come lo era per i sistemi AM e FM/PM, ha senso confrontare tra di loro questi tipi di sistemi. Occorre però prestare attenzione a una differenza. Per AM/FM/PM il termine di disturbo (il "rumore" del rapporto segnale-rumore) è costituito solamente dal rumore introdotto dal sistema di trasmissione.

Per i sistemi PCM le cose funzionano diversamente, e comunque le cause di disturbo sono 2:

1. L'errore di quantizzazione.
2. Il rumore introdotto dal sistema di trasmissione ma in modo diverso dai sistemi analogici. Abbiamo infatti visto che, lavorando sopra soglia, le prestazioni sono

Cause
disturbo in
sistemi PCM

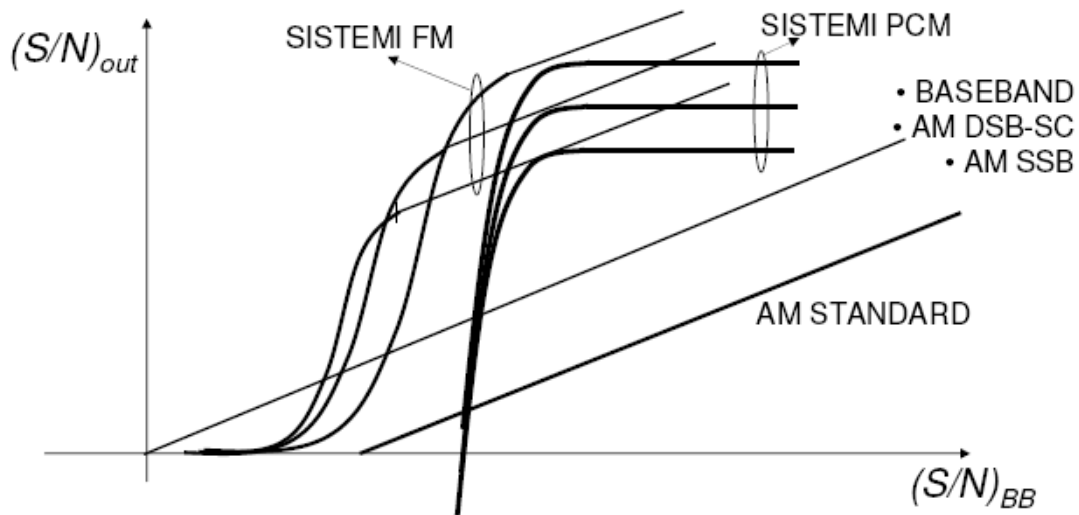


Figura 9.13: Confronto prestazioni sistemi PCM con sistemi AM e FM. Il grafico è puramente qualitativo.

indipendenti da $P(e)$, e dipendenti solo dal numero di bit usati per la quantizzazione. Nel corso di *Trasmissione Numerica* si dimostrerà che, dato un certo sistema di trasmissione digitale, la $P(e)$ dipende dal rapporto segnale-rumore sul canale di trasmissione cioè in un certo senso da $(S/N)_{BB}$.

Confronto PCM con AM/FM/PM La figura 9.13 mostra un confronto tra le prestazioni dei sistemi PCM e quelle dei sistemi AM, PM e FM. Si noti tuttavia che il grafico è puramente qualitativo. Per il PCM i valori "reali" della curva dipendono fortemente dal sistema digitale utilizzato per il trasporto dei bit, come si vedrà nel successivo corso *Trasmissione Numerica*.

Dal grafico si vede bene che i sistemi PCM, FM/PM hanno un comportamento a soglia, i sistemi AM invece no. Possiamo quindi concludere che per valori elevati di rapporto segnale-rumore i sistemi PCM, FM/PM hanno prestazioni migliori. Per bassi valori di SNR invece con AM si ottengono prestazioni migliori.

9.3 Quantizzazione non uniforme

Nel paragrafo precedente abbiamo analizzato dettagliatamente la quantizzazione uniforme. Il segnale di ingresso aveva densità di probabilità uniforme in $[-V, +V]$ e il quantizzatore aveva come dinamica di ingresso l'intervallo $[-V, +V]$, che veniva poi diviso in M sotto-intervalli tutti della stessa lunghezza.³

Molti segnali di interesse pratico hanno tuttavia una *densità di probabilità non uniforme*. La voce ad esempio ha una densità di probabilità concentrata attorno al valor medio (vedi figura 9.14). In questi casi, può essere di interesse una quantizzazione

³Nella trattazione che segue, si presti particolare attenzione al fatto che si devono separare bene i concetti di:

- Densità di probabilità del segnale di ingresso (ad esempio, uniforme);
- Tipo di schema di quantizzazione adottato (ad esempio, uniforme);

che sono due cose ben distinte.

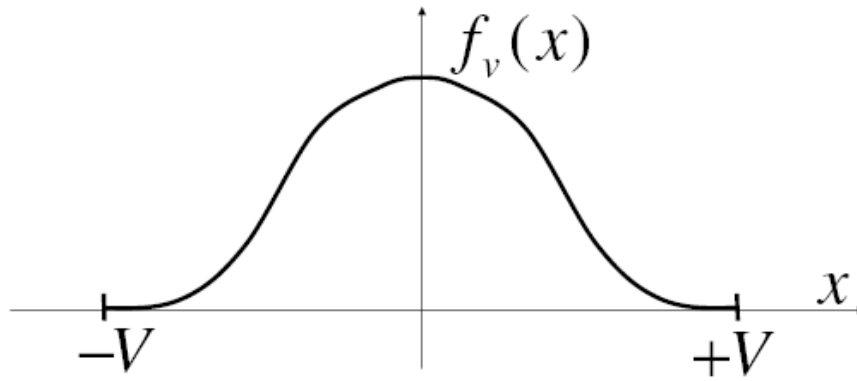


Figura 9.14: Esempio di densità di probabilità non uniforme. In questo caso i valori più probabili sono concentrati intorno al valor medio. La voce ha una densità di probabilità di questo tipo.

di tipo *non uniforme*.

Si può dimostrare che in tutti i casi in cui si utilizza un quantizzatore uniforme con un segnale con densità di probabilità non uniforme, il rapporto segnale-rumore è peggiore che nel caso di segnale con distribuzione uniforme. Matematicamente, per il caso di quantizzazione uniforme si dimostra che:

1. Se si usa un numero elevato di livelli M , la varianza dell'errore di quantizzazione è indipendente dalla densità di probabilità del segnale di ingresso;
2. A parità di range, tuttavia, la potenza del segnale utile è inferiore.

Supponiamo cioè di confrontare l'effetto della quantizzazione uniforme su due segnali:

1. Segnale con densità di probabilità uniforme in $[-V, +V]$.
2. Segnale con densità di probabilità non uniforme in $[-V, +V]$.

Si otterrà che nell'espressione:

$$\left(\frac{S}{N}\right)_Q = \frac{\sigma_{V_{in}}^2}{\sigma_{e_q}^2}$$

la varianza di errore è uguale nei due casi:

$$\sigma_{e_q}^2|_1 = \sigma_{e_q}^2|_2$$

tuttavia il primo segnale ha una varianza maggiore:

$$\sigma_{V_{in}}^2|_1 > \sigma_{V_{in}}^2|_2$$

e dunque:

$$\left(\frac{S}{N}\right)_Q|_1 > \left(\frac{S}{N}\right)_Q|_2$$

Per capire meglio, vediamo ora un esempio semplice in cui si riescono a fare i conti.

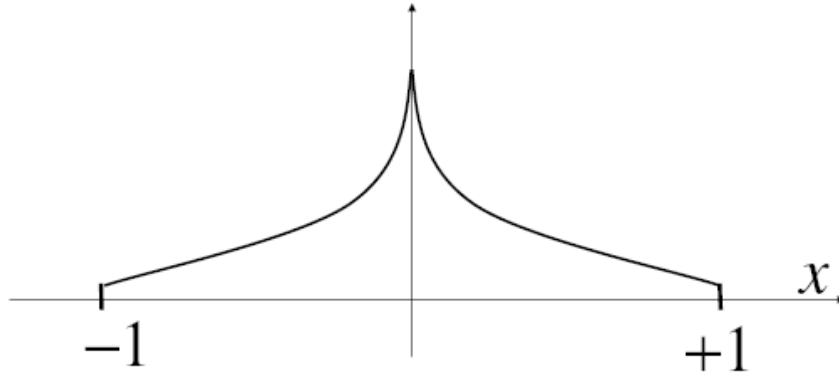


Figura 9.15: Densità di probabilità esponenziale bilatera troncata tra -1 e $+1$.

Esempio 9.3. Segnale con densità di probabilità esponenziale bilatera troncata Consideriamo un segnale con densità di probabilità esponenziale bilatera troncata in $[-1, +1]$ (vedi figura 9.15):

$$f_v(x) = k \cdot e^{-\alpha|x|}$$

dove k e α sono due parametri. Affinché $f_v(x)$ sia una densità di probabilità, deve essere $\int_{-1}^{+1} f_v(x) dx = 1$. Si ottiene quindi:

$$k = \frac{\alpha}{2(1 - e^{-\alpha})}$$

da cui:

$$f_v(x) = \frac{\alpha}{2(1 - e^{-\alpha})} \cdot e^{-\alpha|x|}$$

Svolgendo i calcoli, si ottiene che il valore quadratico medio (che è uguale alla varianza, visto che il valor medio è nullo) di questa densità di probabilità vale:

$$E[V^2] = \frac{1}{\alpha^2(1 - e^{-\alpha})} \cdot [2 - e^{-\alpha}(\alpha^2 + 2\alpha + 2)] \quad (9.18)$$

Il parametro α serve quindi a cambiare la varianza della densità di probabilità. La (9.18) ha l'andamento mostrato in figura 9.16. Dalla figura si vede che per qualsiasi valore di α è sempre valida la seguente disuguaglianza:

$$E[V^2] \leq \frac{1}{3}$$

Si noti ora che, sullo stesso range $[-1, +1]$ una densità di probabilità uniforme ha valore quadratico medio (che è uguale alla varianza, visto che il valor medio è nullo) pari a $1/3$. Si vede dunque in questo caso che, a parità di range, qualunque esponenziale bilatera troncata ha potenza inferiore alla distribuzione uniforme. ■

Per le considerazioni appena fatte, ha interesse studiare leggi di quantizzazione diverse dalla quantizzazione uniforme. L'idea base che andremo a dimostrare è che conviene quantizzare più fittamente nelle zone dove il segnale è più probabile.

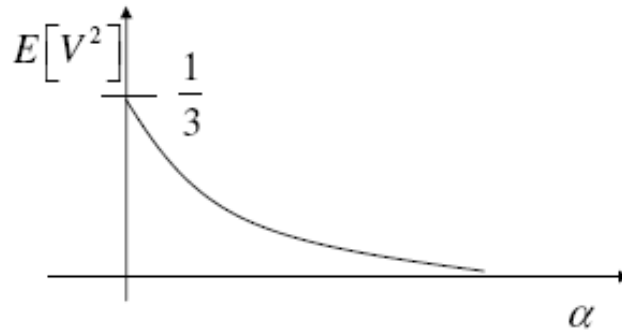


Figura 9.16: Grafico della funzione (9.18): varianza dell'esponenziale bilatera troncata, in funzione di α .

Segnale con dinamica minore della dinamica del quantizzatore

Si consideri un primo caso semplice:

- Quantizzazione uniforme su range $[-V, +V]$;
- Segnale V_{in} uniforme nel range $[-W, +W]$, con $W < V$.

$(S/N)_Q$ si calcola con la (9.2) e la (9.3) mettendo però in quest'ultima W al posto di V (visto che ora il segnale di ingresso è uniforme in $[-W, +W]$). Si ottiene:

$$\left(\frac{S}{N}\right)_Q = M^2 \left(\frac{W}{V}\right)^2$$

che, passando ai dB diventa (ricordando la (9.5)):

$$\begin{aligned} \left(\frac{S}{N}\right)_Q &\approx 6n + 10 \log_{10} \left(\frac{W}{V}\right)^2 \\ &= 6n + 20 \log_{10} \left(\frac{W}{V}\right) \end{aligned}$$

Essendo $W < V$, la quantità $20 \log_{10} \left(\frac{W}{V}\right)$ è negativa, e rappresenta la "perdita" che si ha a causa del fatto che la dinamica del segnale $[-W, +W]$ è minore della dinamica del quantizzatore. Se ad esempio $W = V/2$, ho $20 \log_{10} \left(\frac{W}{V}\right) = 20 \log_{10}(1/2) = -6$ dB e ottengo che:

$$\left(\frac{S}{N}\right)_Q \Big|_{dB} \approx 6(n - 1)$$

che equivale ad avere "perso" 1 bit di quantizzazione. Il risultato è comprensibile: in sostanza, non si usa un bit del convertitore A/D.

9.3.1 Calcolo dell'errore di quantizzazione

Fatte le premesse del paragrafo precedente, affrontiamo ora il problema nel caso generale. Ci proponiamo cioè di valutare la varianza dell'errore di quantizzazione nel caso generale di:

$(S/N)_Q$
quantizzazio-
ne uniforme
su range
diverso da
quello del
segnale

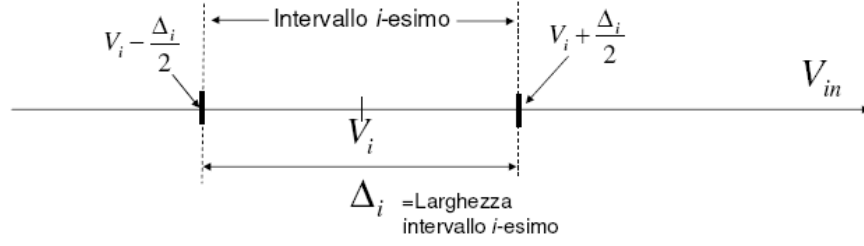


Figura 9.17: Rappresentazione grafica intervallo i -esimo per un quantizzatore non uniforme. Ogni intervallo ha una larghezza diversa.

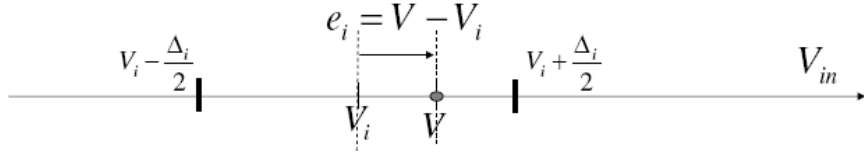


Figura 9.18: Rappresentazione grafica dell'errore di quantizzazione per quantizzatore non uniforme, nel caso in cui il segnale di ingresso V cada all'interno dell'intervallo i -esimo.

- Quantizzatore non uniforme;
- Segnale di ingresso con densità di probabilità generica.

Supponiamo sempre di avere $M = 2^n$ intervalli sul range di ingresso $[-V, +V]$. La "novità" è che ora supporremo ciascun intervallo di quantizzazione di larghezza arbitraria. I parametri relativi all' i -esimo intervallo di quantizzazione sono (vedi figura 9.17):

- Δ_i = larghezza i -esimo intervallo;
- V_i = punto centrale i -esimo intervallo.

Se V è un generico segnale in ingresso che cade nell'intervallo i -esimo, l'errore di quantizzazione è dato da (vedi figura 9.18):

$$e_i \triangleq V - V_i$$

e il suo valore quadratico medio vale:

$$E[e_i^2] = \int_{V_i - \frac{\Delta_i}{2}}^{V_i + \frac{\Delta_i}{2}} (V - V_i)^2 \cdot f_V(V | V \in i\text{-esimo}) dV \quad (9.19)$$

Se gli intervalli sono abbastanza piccoli, posso considerare la densità di probabilità $f_V(x)$ pressoché costante all'interno di un intervallo (vedi figura 9.19). In particolare, per la densità di probabilità condizionata all'intervallo i -esimo si ha:

$$f_V(V | V \in i\text{-esimo}) \approx \frac{1}{\Delta_i} \quad (9.20)$$

Posso allora sostituire la (9.20) nella (9.19) e trovare:⁴

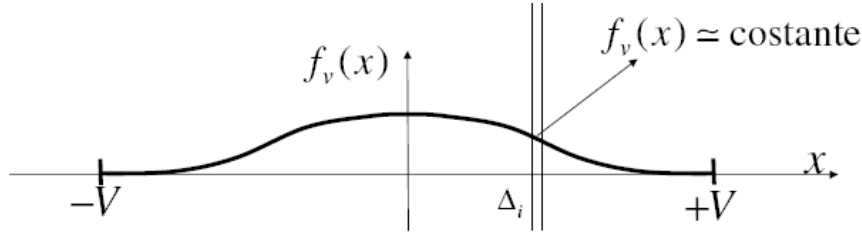


Figura 9.19: Nel caso che gli intervalli di quantizzazione Δ_i siano sufficientemente piccoli, posso approssimare la densità di probabilità $f_V(x)$ del segnale di ingresso all'interno di ogni intervallo come costante.

$$\begin{aligned}
 E[e_i^2] &= \int_{V_i - \frac{\Delta_i}{2}}^{V_i + \frac{\Delta_i}{2}} (V - V_i)^2 \cdot f_V(V | V \in i\text{-esimo}) dV \\
 &= \int_{V_i - \frac{\Delta_i}{2}}^{V_i + \frac{\Delta_i}{2}} (V - V_i)^2 \frac{1}{\Delta_i} dV \\
 &= \int_{-\frac{\Delta_i}{2}}^{+\frac{\Delta_i}{2}} x^2 \frac{1}{\Delta_i} dx \\
 &= \frac{1}{\Delta_i} \left[\frac{x^3}{3} \right]_{-\Delta_i/2}^{+\Delta_i/2} \\
 &= \frac{\Delta_i^2}{12}
 \end{aligned}$$

L'errore di quantizzazione totale sarà legato a tutti i vari $E[e_i^2]$ pesati per la probabilità che V cada nell'intervallo i -esimo:

Formula per
l'errore di
quantizzazio-
ne nel caso
generico

$$\begin{aligned}
 E[e_Q^2] &= \sum_{i=1}^M P(V \in i\text{-esimo}) \frac{\Delta_i^2}{12} \\
 &\approx \sum_{i=1}^M f_V(V_i) \Delta_i \frac{\Delta_i^2}{12} \\
 &= \sum_{i=1}^M f_V(V_i) \frac{\Delta_i^3}{12}
 \end{aligned}$$

Si dimostra che, a parità di M , $(S/N)_Q$ viene minimizzato nella situazione in cui i termini della sommatoria tendono ad essere uguali. La dimostrazione di questo risultato è difficile, ed esula dagli scopi di questo corso. In sostanza si cerca di ottenere:

$$f_V(V_i) \frac{\Delta_i^3}{12} = \text{Costante}$$

⁴Nei conti che seguono si passa dalla seconda alla terza riga effettuando il cambio di variabile: $x = V - V_i$, che rende l'integrale più semplice da risolvere.

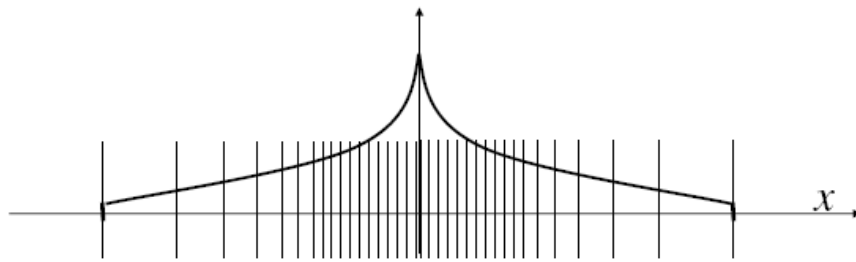


Figura 9.20: Segnale con densità di probabilità non uniforme. In questi casi conviene quantizzare più fittamente nelle zone dove il segnale è più probabile.

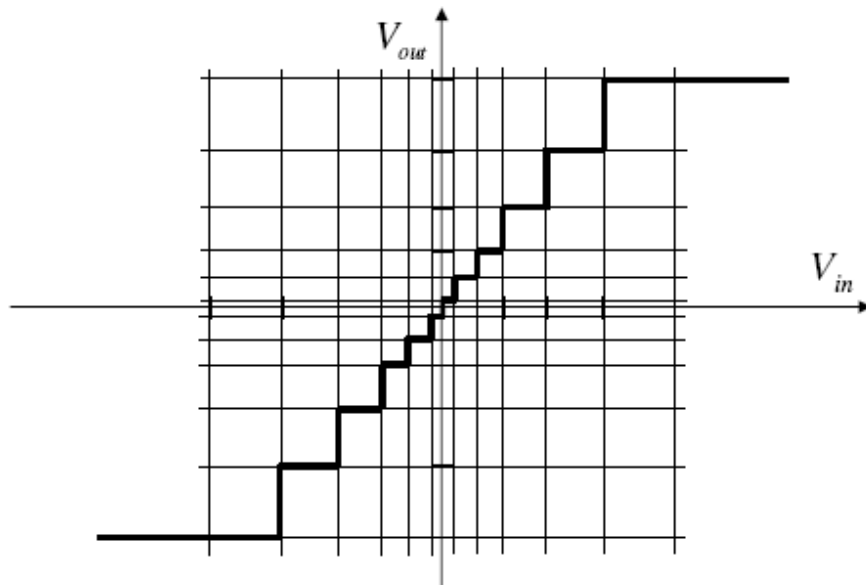


Figura 9.21: Curva di quantizzazione per la densità di probabilità della figura 9.20.

Questa espressione mi dice che con la quantizzazione non uniforme le prestazioni sono ottimizzate se si quantizza più fittamente dove $f_V(V_i)$ è più elevato. Questo risultato è rappresentato graficamente in figura 9.20. La curva di quantizzazione per la densità di probabilità di figura 9.20 diventa qualitativamente quella di figura 9.21.

9.3.2 Tecniche di companding

Gran parte dei convertitori A/D e D/A commerciali (in chip) sono di tipo uniforme. Per ottenere un quantizzatore non uniforme partendo da un quantizzatore uniforme si usa solitamente una tecnica detta di *companding*. Si fa cioè precedere il convertitore A/D uniforme da un dispositivo con curva di ingresso-uscita non lineare scelta opportunamente, come mostrato in figura 9.22. La figura 9.23 mostra come con tale tecnica si ottenga un quantizzatore non uniforme.

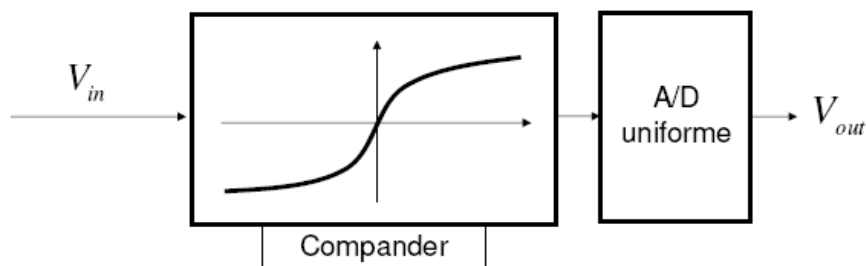


Figura 9.22: Schema realizzativo di convertitore A/D non uniforme utilizzando tecnica di companding. Si fa precedere il convertitore A/D uniforme da un dispositivo con curva di ingresso-uscita non lineare scelta opportunamente.

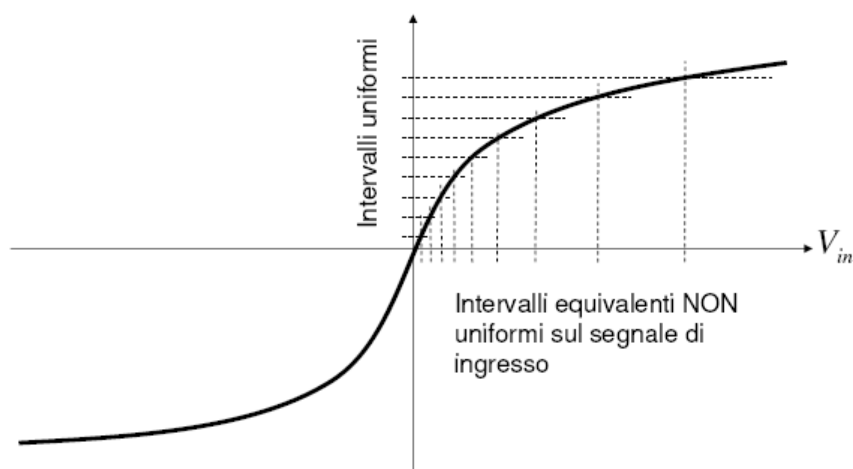


Figura 9.23: Illustrazione di come lo schema di figura 9.22 equivalga a realizzare un quantizzatore non uniforme.

Dal lato del ricevitore si dovrà poi utilizzare un dispositivo non lineare con la curva complementare per ricostruire il segnale di partenza.

Le tecniche di companding sono utili anche per controbattere un problema che per ora non avevamo affrontato. In particolare, nei casi pratici il range del segnale in ingresso non è noto a priori. Ad esempio, il segnale vocale all'uscita di un microfono ha una densità di probabilità esponenziale bilatera troncata (in prima approssimazione) ma su un range che dipende sia dall'intensità del suono emesso, sia dalle caratteristiche del microfono.

Si segue allora il seguente approccio. Si sceglie una legge di companding che renda $(S/N)_Q$ il più possibile indipendente dal segnale di ingresso. I passaggi matematici per risolvere questo problema vanno al di là degli scopi di questo corso. Il lettore interessato può fare riferimento a [Khalin].

PCM telefonico: le leggi " μ -law" e "A-law"

Il PCM telefonico utilizza *sempre* una tecnica di companding. Le "leggi di companding" utilizzate sono diverse tra Europa, Stati Uniti e Giappone, ma il principio è lo stesso.

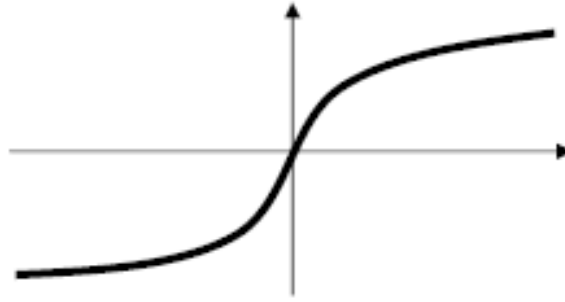


Figura 9.24: Caratteristica ingresso-uscita di un compander generico. Si noti come per elevati valori dell'ingresso, l'uscita tenda a saturare.

In tutti i casi, la legge è indicativamente del tipo di figura 9.24. Si usa cioè una curva che "satura" per ingressi crescenti.

A-law

- "A-law" Europa:

$$|V_{out}| = \begin{cases} \frac{A|V_{in}|}{1+\log(A)} & \text{per } |V_{in}| \in [0, \frac{1}{A}] \\ \frac{1+\log(A|V_{in}|)}{1+\log(A)} & \text{altrove} \end{cases}$$

dove: $V_{in} \in [-1, +1]$ e $A = 87.56$.

μ -law

- " μ -law" (USA):

$$|V_{out}| = \frac{\log(1 + \mu|V_{in}|)}{\log(1 + \mu)}$$

dove: $V_{in} \in [-1, +1]$ e $\mu = 255$.

Si può dimostrare (anche se esula dagli scopi di questo corso) che per qualunque legge di companding "ragionevole" si ha una legge di questo tipo:

$(S/N)_Q$ per
legge di
companding
generica

$$\left(\frac{S}{N} \right)_Q \Big|_{dB} \approx 6n - a$$

dove a è un parametro che dipende dal segnale in ingresso e dalla legge scelta. Rimane cioè ancora vera la "6 dB law", cioè il fatto che ogni bit di quantizzazione in più migliora di 6 dB le prestazioni.

9.4 Conclusione

In questo capitolo abbiamo trattato i sistemi PCM, con particolare attenzione alle tecniche di quantizzazione. Abbiamo prima trattato la quantizzazione uniforme, calcolando il rapporto segnale-rumore dovuto alla sola quantizzazione e il rapporto segnale rumore nel caso di trasmissione su canale digitale.

Abbiamo poi trattato la quantizzazione non uniforme. Questo tipo di quantizzazione è più adatto per segnali con densità di probabilità non uniforme. Abbiamo calcolato

l'errore di quantizzazione nel caso di quantizzazione non uniforme. Abbiamo infine visto come realizzare un quantizzatore non uniforme utilizzando le cosiddette tecniche di companding.

La quantizzazione è una tecnica ampiamente utilizzata nell'ambito della compressione dati, come si vedrà nel corso di Elaborazione di Immagini e Video. Durante questo corso verranno ripresi e trattati più a fondo alcuni argomenti di questo capitolo.

9.5 Riassunto formule principali sul PCM

Si riportano in questa sezione le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi delle esercitazioni.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- $(S/N)_Q$ quantizzatore uniforme e segnale con densità di probabilità uniforme:

$$\left(\frac{S}{N}\right)_Q = M^2 \approx 6n \text{ [dB]}$$

- $(S/N)_Q$ dovuto agli errori sui bit:

$$\left(\frac{S}{N}\right)_e \approx \frac{1}{4P_e}$$

- $(S/N)_{out}$ complessivo in presenza di errori sul canale e probabilità di errore alla soglia P_e^* :

$$\begin{aligned} \left(\frac{S}{N}\right)_{out} &= \frac{M^2}{1 + 4(M^2 - 1)P_e} \\ P_e^* &= \frac{1}{4(M^2 - 1)} \end{aligned}$$

- $(S/N)_Q$ quantizzatore uniforme, ma su range diverso da quello del segnale:

$$\left(\frac{S}{N}\right)_Q = M^2 \left(\frac{W}{V}\right)^2$$

- Errore di quantizzazione per quantizzatore non uniforme:

$$\begin{aligned} E[e_Q^2] &= \sum_{i=1}^M P(V \in i\text{-esimo}) \frac{\Delta_i^2}{12} \\ &\approx \sum_{i=1}^M f_V(V_i) \Delta_i \frac{\Delta_i^2}{12} \\ &= \sum_{i=1}^M f_V(V_i) \frac{\Delta_i^3}{12} \end{aligned}$$

Capitolo 10

Teoria dell'Informazione per sorgenti

10.1 Introduzione

La *Teoria dell'Informazione* costituisce la base matematica delle trasmissioni numeriche moderne. Si divide in due grandi ambiti:

Codifica di Sorgente: è alla base delle tecniche di compressione. Data una certa sorgente di informazione digitale, cerca di ridurre al massimo il bit-rate da trasmettere. L'idea di base che si sfrutta è quella di togliere ridondanza alla sorgente.

Codifica di Canale: cerca di rendere la trasmissione più resistente agli errori, cioè di diminuire la $P_b(e)$ effettiva sul canale. L'idea di base che sfrutta è quella di aggiungere ridondanza al messaggio da trasmettere.

La figura 10.1 mostra la struttura di un trasmettitore digitale. Il flusso di dati viene prima inviato al codificatore di sorgente che effettua la compressione e poi al codificatore di canale che rende lo stream di bit più resistente agli errori sul canale. L'uscita del codificatore di canale è poi inviata al modulatore digitale.

In questo capitolo parleremo di *Teoria dell'Informazione per sorgenti* (cioè la codifica di sorgente). Si tratta di una teoria atta a:

1. Caratterizzare le sorgenti di informazione in maniera statistica.
2. Introdurre i concetti di "compressione" del flusso informativo.

Si tratta di una teoria matematica molto complessa e articolata, pertanto una sua trattazione completa esula dagli scopi di questo corso. In questo capitolo vedremo solo alcuni concetti introduttivi e un metodo di codifica.

Non tratteremo invece la codifica di canale, che sarà trattata in altri corsi, a partire da Trasmissione Numerica.

10.2 Codifica di sorgente

Come detto in precedenza, lo scopo della codifica di sorgente è di ridurre il bit-rate in uscita da una sorgente, ed è molto efficiente su:

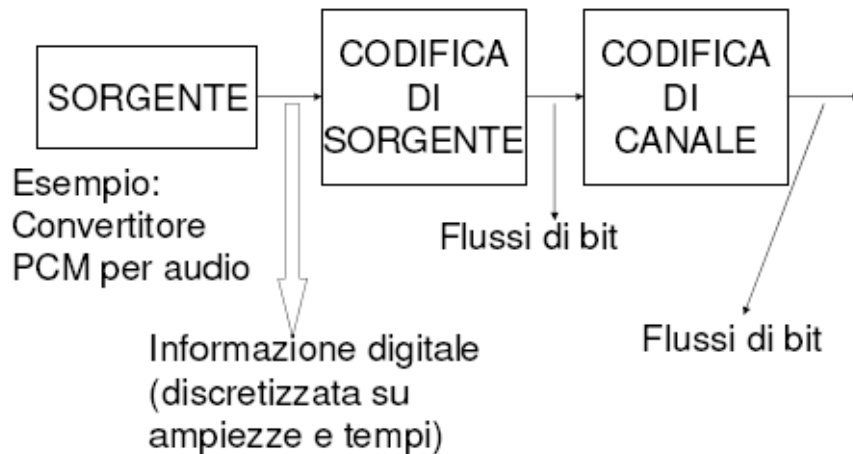


Figura 10.1: Schema a blocchi di un trasmettitore digitale.

- Segnali analogici convertiti in digitale con PCM (ad esempio stream audio MPEG 3 o video MPEG-II);
- File di tipo testuale.

Vediamo ora un esempio che mostra che in alcuni casi è assolutamente necessario comprimere il segnale di informazione.

Esempio 10.1. Video digitale Consideriamo un segnale video con queste caratteristiche:

- 1000 * 800 pixel;
- 16 bit per pixel (circa 65000 colori);
- 50 refresh al secondo.

Il bit-rate risultante senza compressione è di $640 \text{ Mbit/s} = 0.64 \text{ Gbit/s}$. Questo valore è assolutamente inaccettabile per le applicazioni pratiche. Con la tecnica di compressione MPEG-II si riesce a trasmettere un buon segnale video con un bit-rate attorno a 4-6 Mbit/s. ■

Esempio 10.2. Audio PCM Audio digitalizzato, qualità telefonica. Dai 64 kbit/s tipici del PCM non quantizzato si passa a:

- Circa 13 kbit/s per il cellulare GSM;
- Si riesce a scendere a qualche kbit/s per le tecniche tipiche delle applicazioni Internet per lo streaming audio (Netmeeting, NetPhone). ■

Classificazione tecniche di codifica di sorgente

Facendo riferimento alla figura 10.2, possiamo dividere le tecniche di codifica di sorgente in 2 grandi famiglie:

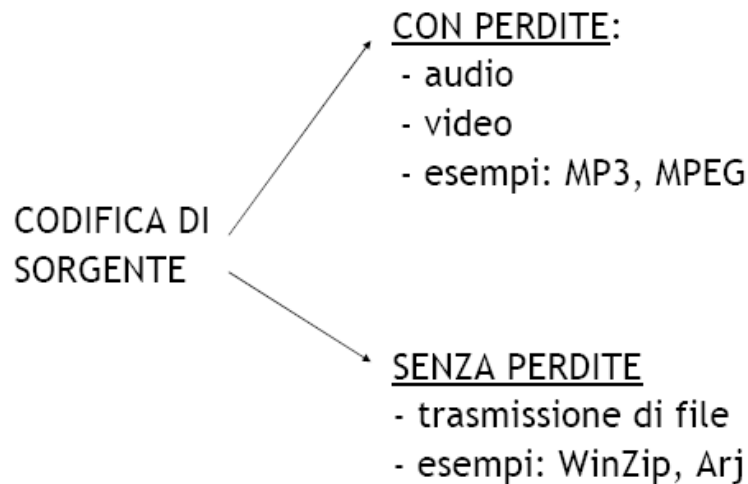


Figura 10.2: Classificazione tecniche di codifica di sorgente: senza perdite e con perdite.

Compressione senza perdite (lossless): la cascata dei processi di compressione-decompressione porta esattamente allo stesso flusso digitale che si ha in ingresso. Ci sono infatti molte applicazioni in cui è importante che il processo di compressione-decompressione non generi perdita di informazione. La compressione lossless si usa solitamente per comprimere dati "discreti", come testo, dati generati da computer e alcuni tipi di immagini e video.

Compressione con perdite (lossy): la cascata dei processi di compressione-decompressione porta in uscita ad un flusso digitale diverso da quello in ingresso. Tuttavia, si fa in modo che la "perdita" sia tollerabile (o addirittura irrilevante) per l'utente finale. La compressione con perdite ha il vantaggio che può ottenere tassi di compressione molto più elevati rispetto a quella senza perdite. La compressione lossy si usa solitamente per comprimere audio o video, in applicazioni dove un'esatta ricostruzione dei dati non è fondamentale.

In questo corso, ci occupiamo solo del primo caso (lossless) dove la cascata di codificatore-decodificatore (di sorgente) possono ricostruire in maniera ideale il segnale. L'argomento è molto complesso, e noi daremo solo i concetti base e un metodo di codifica.

Osservazione generale: i metodi senza perdite si basano sulle caratteristiche statistiche della sorgente; vedremo che in alcuni casi non è possibile ottenere una significativa compressione.

10.3 Teoria dell'Informazione per sorgenti: premesse matematiche

Facendo riferimento alla figura 10.1, supponiamo di avere una sorgente caratterizzabile come segue:

- Discretizzata su tempi ed ampiezze;

- Per ogni intervallo di tempo, emette un "simbolo" preso da un alfabeto di M simboli.

Un semplice esempio di questo tipo è un convertitore PCM per audio (cioè un dispositivo che campiona e quantizza un segnale analogico). Un altro esempio è una tastiera che emette caratteri ASCII.

Indichiamo i simboli dell'alfabeto di sorgente con l'insieme:

$$X = \{m_1, \dots, m_M\}$$

Caratterizziamo statisticamente la sorgente con le probabilità di emissione p_1, \dots, p_M di ciascun simbolo, dove p_i è la probabilità che sia emesso il simbolo m_i , cioè $p_i = P(m_i)$. Supponiamo che la sorgente emetta un flusso di simboli che, in istanti diversi, siano *statisticamente indipendenti* tra di loro. L'ipotesi di indipendenza è detta ipotesi di *sorgente senza memoria*. Questa ipotesi non è vera in moltissimi casi. La facciamo solo per semplificare la trattazione. Molti sistemi di codifica traggono vantaggio proprio dalla correlazione temporale tra i vari simboli (ad esempio: frame successivi di un segnale video).

In questo contesto i "bit" delle stringhe di codifica saranno in realtà chiamati "digit". Nella Teoria dell'Informazione alcuni termini assumono significati diversi da altri ambiti. Il termine "bit" ha un altro significato, che vedremo tra poco.

Le idee di base che andremo a utilizzare sono le seguenti:

1. Sfruttare il fatto che i simboli non sono in generale equiprobabili.
2. Associare ad ogni simbolo una opportuna stringa di valori binari "01001..." con una associazione a lunghezza variabile.
3. Associare codifiche più corte per i simboli di sorgente più probabili.

Codifica a
lunghezza
variabile

La teoria che presenteremo si basa sull'uso di una "codifica a lunghezza variabile". Si associano codifiche più corte (in numero di digit) per i simboli di sorgente più probabili. La codifica a lunghezza variabile è il nocciolo della questione, e andrà confrontata con la soluzione più semplice di codifica a lunghezza fissa, su:

$$INT_{superiore} [\log_2(M)] = \lceil \log_2(M) \rceil$$

Esempio: se ho 256 simboli, la codifica più ovvia è quella a lunghezza fissa su:

$$\log_2(256) = 8 \text{ digit/simbolo}$$

Se i simboli sono da emettere ogni T_m secondi, il *digit-rate* necessario è dato da:

$$\frac{1}{T_m} \lceil \log_2(M) \rceil$$

Lo scopo della Teoria dell'Informazione per sorgenti è di cercare di ridurre questa quantità passando a codifica a lunghezza variabile.

Si ricorda la formula per il calcolo di logaritmi in base 2:

$$\log_2(x) = \frac{\log_{10}(x)}{\log_{10}(2)} = \frac{\log_e(x)}{\log_e(2)}$$

e inoltre che:

$$\log_2(2^x) = x$$

10.3.1 Entropia di sorgenti senza memoria

Imponendo due semplici condizioni, intuitivamente desiderabili, ovvero che l'informazione portata da un simbolo sia tanto maggiore quanto meno il simbolo è probabile; e che l'informazione di una coppia di simboli indipendenti sia la somma delle rispettive quantità di informazione, Shannon riconobbe che si deve definire la **quantità di informazione** di uno specifico simbolo m_i avente probabilità p_i come:

Quantità di
informazione

$$I(m_i) \triangleq \log_2 \left(\frac{1}{p_i} \right) \quad (10.1)$$

La quantità di informazione si misura in "bit". Vediamo alcune caratteristiche della quantità di informazione definita dalla (10.1), ricordando che $0 < p_i < 1$:

1. $I(m_i) > 0$, cioè è sempre una quantità non negativa (è un numero reale, e può essere anche minore di 1).
2. Al crescere di p_i si ha che $I(m_i)$ decresce.
Esempio: Se $p_i = P(m_i) = 0.5 \Rightarrow I(m_i) = 1 \text{ bit}$.
Se $p_i = P(m_i) = 1 \Rightarrow I(m_i) = 0 \text{ bit}$.
Interpretazione: un simbolo "sicuro" non porta informazione.
3. Se m_1 e m_2 sono due simboli indipendenti, si ha che $I(m_1 m_2) = I(m_1) + I(m_2)$.
Infatti:

$$\begin{aligned} I(m_1 m_2) &= \log_2 \left[\frac{1}{P(m_1 m_2)} \right] \\ &= \log_2 \left[\frac{1}{P(m_1) P(m_2)} \right] \\ &= \log_2 \left[\frac{1}{P(m_1)} \right] + \log_2 \left[\frac{1}{P(m_2)} \right] \\ &= I(m_1) + I(m_2) \end{aligned}$$

Si definisce come *entropia di sorgente* la quantità di informazione media di tutta la sorgente:

Entropia di
sorgente

$$H(x) \triangleq E[I_i] = \sum_{i=1}^M p_i \log_2 \left(\frac{1}{p_i} \right) \left[\frac{\text{bit}}{\text{simbolo}} \right] \quad (10.2)$$

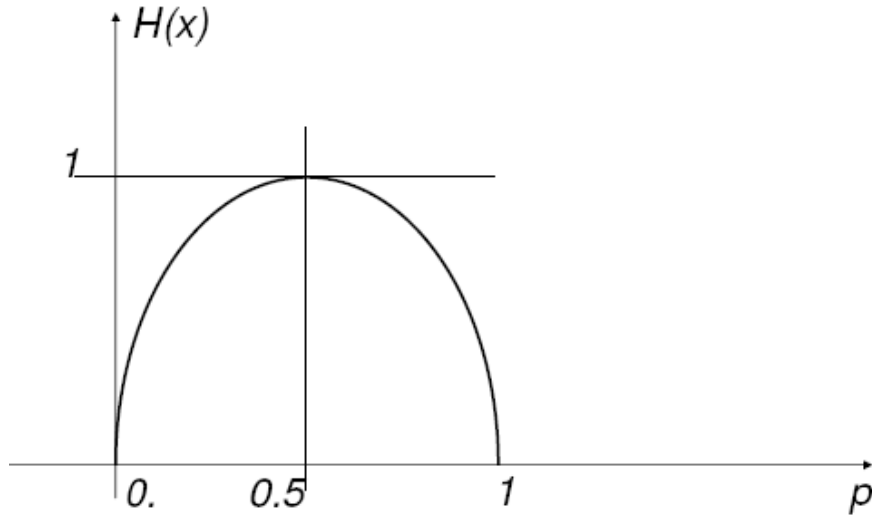


Figura 10.3: Entropia di sorgente con 2 simboli.

Entropia di sorgente con due elementi

Consideriamo una sorgente con due soli simboli $\{m_1, m_2\}$ con probabilità $p_1 = p$ e $p_2 = 1 - p$. Usando la definizione (10.2) si trova che la sua entropia vale:

$$H(x) = p \log_2 \left(\frac{1}{p} \right) + (1 - p) \log_2 \left(\frac{1}{1 - p} \right)$$

La figura 10.3 mostra $H(x)$ in funzione di p . L'entropia della sorgente è massima (e in questo caso vale 1 bit/simbolo) per $p = 0.5$, cioè per simboli equiprobabili. L'entropia tende a zero per $p \rightarrow 0$ e per $p \rightarrow 1$.

Entropia di sorgente con M elementi

Il precedente risultato si estende anche a sorgenti con $M > 2$ simboli. In particolare si dimostra (non lo facciamo) che l'entropia è massima per simboli equiprobabili, cioè per:

$$p_i = \frac{1}{M} \quad \forall p_i$$

e in tal caso l'entropia vale:

$$H(x) = \log_2(M) \quad \left[\frac{\text{bit}}{\text{simbolo}} \right]$$

In generale si ha sempre il seguente bound superiore per l'entropia di una sorgente con M simboli:

$$H(x) \leq \log_2(M) \quad \left[\frac{\text{bit}}{\text{simbolo}} \right] \quad (10.3)$$

Bound
sull'entropia
di sorgente
con M
simboli

10.3.2 Codifica di sorgenti senza memoria

Sia n_i la lunghezza della codifica del simbolo m_i in numero di digit. Si definisce come **lunghezza media di codifica**:

Lunghezza
media di
codifica

$$\bar{n} \triangleq E[n_i] = \sum_{i=1}^M p_i n_i \quad (10.4)$$

La lunghezza media di codifica è un parametro essenziale per la codifica di sorgente. Infatti, il digit-rate medio necessario dopo la codifica è dato da:

$$\frac{1}{T_m} \cdot \bar{n} \quad \left[\frac{\text{digit}}{\text{secondo}} \right]$$

Il digit-rate medio è un parametro fondamentale. Data una certa sorgente digitale, dopo la compressione, dovrò "comprare" un sistema di trasmissione che abbia una velocità di trasmissione pari ad (almeno) il digit-rate medio. Allora una codifica è tanto migliore quanto più \bar{n} è inferiore alla codifica a lunghezza fissa, dove si ha:

$$\bar{n} = \lceil \log_2(M) \rceil$$

Il caso a lunghezza fissa è cioè da considerare come un riferimento a cui confrontare le prestazioni del caso codificato a lunghezza variabile.

Teorema di Shannon della codifica di sorgente

Presentiamo ora un risultato fondamentale della Teoria dell'Informazione. Per qualunque possibile codifica:

$$\bar{n} \geq H(x) \quad (10.5)$$

Inoltre, per qualunque sorgente, esiste una codifica tale per cui:

Teorema di
Shannon

$$H(x) \leq \bar{n} \leq H(x) + 1 \quad (10.6)$$

della codifica
di sorgente

Quest'ultimo risultato è noto anche come *Teorema di Shannon per la codifica di sorgente*. La dimostrazione dei risultati (10.5) e (10.6) esula dagli obiettivi di questo corso. Questi risultati sono essenziali poiché danno un limite teorico alla lunghezza media di codifica, cioè a quanto si può comprimere (senza perdite). Questo limite è strettamente legato all'entropia della sorgente, cioè alle sue caratteristiche statistiche.

Esempio: simboli equiprobabili. Se $M = 256$ e i simboli sono equiprobabili:

$$H(x) = 8 \text{ bit/simbolo}$$

una codifica a lunghezza fissa richiede 8 digit/simbolo ed è dunque la migliore possibile.

Simbolo	Codifica
A	0
B	01
C	10
D	1

Tabella 10.1: Esempio codifica a lunghezza variabile.



Figura 10.4: Due possibili modi per decodificare la stringa $ABBCA$ codificata col codificatore di figura 10.1. Si noti che entrambe le decodifiche sono errate.

Codici a lunghezza variabile a prefisso

Nei codici a lunghezza variabile non si vuole mettere un demarcatore di separazione tra le codifiche di ciascun simbolo. Questo pone delle richieste sulle codifiche.

Esempio 10.3. Codice non univocamente decodificabile Consideriamo una codifica come in tabella 10.1. Supponiamo che la stringa da codificare sia:

$$A B B C A \Rightarrow 00101100$$

Il decodificatore può indifferentemente decodificare in uno dei due modi mostrati in figura 10.4. È chiaro che un sistema di questo tipo non funziona. Nessuna delle due stringhe decodificate è uguale a quella di partenza, pur avendo utilizzato la stessa tabella di decodifica. Il motivo è che il codice non è univocamente decodificabile: infatti 01 può essere decodificato come B oppure come due simboli A e D . ■

Regola del
prefisso

Per risolvere il problema, si devono usare codici che siano univocamente decodificabili. Data cioè una sequenza di digit deve essere possibile decodificarla univocamente. In questo corso, per risolvere questo problema, useremo la **regola del prefisso: nessuna codifica può essere il prefisso della codifica di un altro simbolo**. In questo modo siamo sicuri che un codice è univocamente decodificabile¹. Nel seguito, chiameremo i codici che seguono la regola del prefisso *codici a prefisso*.

Esempio 10.4. Codice a prefisso Consideriamo il codice di tabella 10.2. Si tratta di un codice a prefisso. Supponiamo che la stringa da decodificare sia:

$$A B B C A \Rightarrow 010101100$$

In questo caso il decodificatore opera correttamente, come mostrato in figura 10.5. Si noti come in questo caso non ci sia ambiguità nella decodifica. ■

¹Si noti che esistono dei codici univocamente decodificabili che non seguono la regola del prefisso ma in questo contesto non ci interessano.

Simbolo	Codifica
A	0
B	10
C	110
D	111

Tabella 10.2: Esempio codice a lunghezza variabile a prefisso.

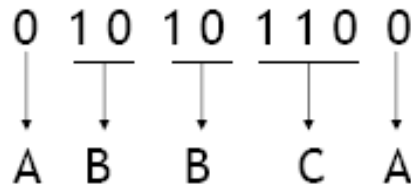


Figura 10.5: Decodifica codice tabella 10.2. Si noti come in questo caso la decodifica sia corretta, trattandosi di un codice a prefisso.

10.4 La codifica di Huffman

Presentiamo ora il più semplice esempio di codifica di sorgente (senza perdite). Si tratta della codifica di Huffman (dal nome del suo inventore). Si tratta del semplice algoritmo descritto qui di seguito.

STEP (ALGORITMO DI HUFFMAN)

1. Ordinare i simboli per probabilità decrescenti.
2. Raggruppare i due simboli con probabilità più bassa, e considerarli come un unico simbolo.
3. GOTO 2.

Alla fine, assegnare opportunamente i digit sulla struttura ad albero risultante. Su ciascuna ramificazione, assegnare "0" verso il basso e "1" verso l'alto.

Vediamo il funzionamento dell'algoritmo con un esempio.

Esempio 10.5. Codifica di Huffman Si consideri l'alfabeto di tabella 10.3. Con una codifica a lunghezza fissa devo usare 3 digit per simbolo. Calcolo dell'entropia:

$$H(x) = \sum_{i=1}^5 p_i \log_2 \left(\frac{1}{p_i} \right) = 1.29 \frac{\text{bit}}{\text{simbolo}}$$

Costruiamo ora una codifica di Huffman per questo alfabeto. L'albero per la codifica

Simbolo	Probabilità
A	0,2
B	0,05
C	0,04
D	0,01
E	0,7

Tabella 10.3: Alfabeto di sorgente per l'esempio 10.5.

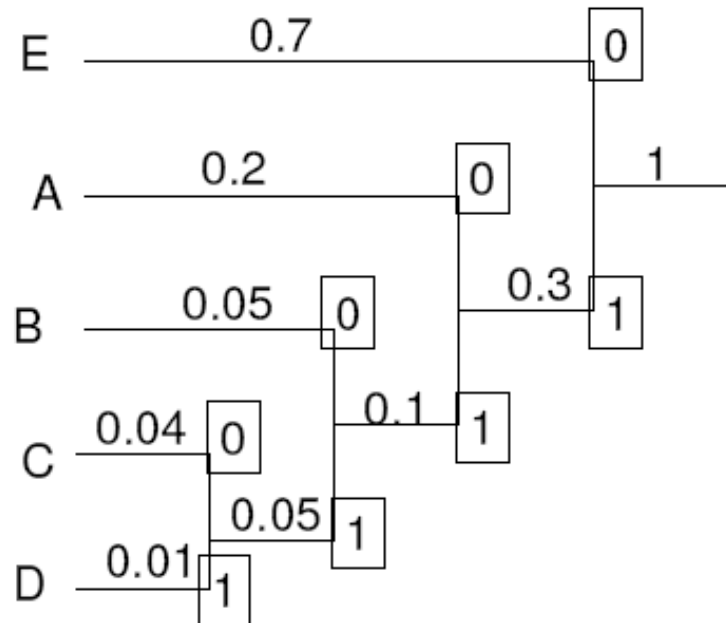


Figura 10.6: Costruzione albero per l'esempio 10.5 sulla codifica di Huffman.

è mostrato in figura 10.6. Una volta assegnati i digit, per trovare le parole di codice bisogna percorrere l'albero da destra verso sinistra, dalla "root" fino alle foglie. Ad esempio per il simbolo C si ottiene la codifica "1110" 10.7.

Procedendo analogamente con gli altri simboli, si ottengono i risultati schematizzati in tabella 10.4. Si ha:

$$\bar{n} = 1.45 \frac{\text{digit}}{\text{simbolo}}$$

Notare che si è ottenuta una compressione rispetto al caso a lunghezza fissa pari a:

$$\frac{1.45}{3} = 0.483 \approx 48\% \text{ di compressione}$$

Se la sorgente emette simboli a 1 kbaud²:

- Con codifica a lunghezza fissa sono necessari:

$$3 \frac{\text{kdigit}}{\text{s}}$$

²1 baud = 1 simbolo/secondo.

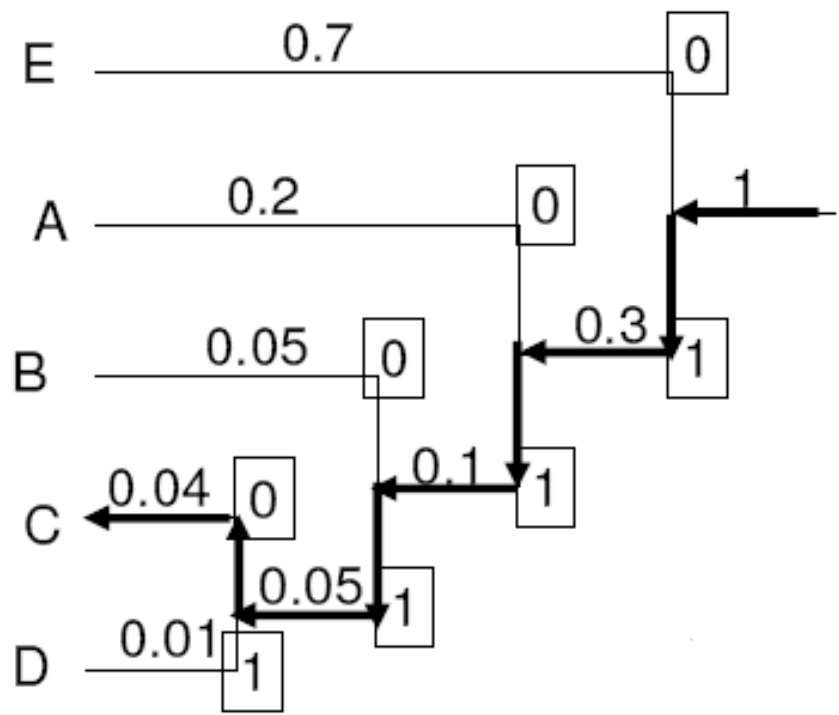


Figura 10.7: Percorso per trovare la parola di codice relativa al simbolo C dell'esempio 10.5 sulla codifica di Huffman.

- Con questa codifica si ha:

$$1.45 \frac{\text{kdigit}}{\text{s}}$$

■

In conclusione possiamo dire che l'algoritmo di Huffman ha un'ottima efficienza, pur essendo uno dei più semplici. Tuttavia richiede una conoscenza completa delle statistiche della sorgente prima di iniziare il processo di codifica. Ad esempio, per codificare un file di testo richiede due passaggi: un primo passaggio per stimare le statistiche e stabilire la tabella di codifica. Un secondo passaggio per effettuare la vera e propria codifica.

Esistono altri algoritmi più complessi, ad esempio i metodi a dizionario, che in particolare:

Altri
algoritmi

SIMBOLI	p_i	CODIFICA	n_i
A	0.2	10	2
B	0.05	110	3
C	0.04	1110	4
D	0.01	1111	4
E	0.7	0	1

Tabella 10.4: Codifica Huffman per l'esempio 10.5.

- Creano la tabella di codifica "in tempo reale" in contemporanea al processo di codifica stessa;
- Tengono conto della correlazione dei simboli (memoria della sorgente).

I metodi a dizionario sono ampiamente utilizzati in prodotti commerciali. Esempi di metodi a dizionario:

- Metodo di Lempel-Ziv (alla base di svariati software di compressione freeware);
- Metodo ZIP (alla base del software commerciale WinZip).

Si tratta di un argomento che non trova posto in questo corso. Sarà affrontato nei corsi di compressione dati. Il lettore interessato ai metodi a dizionario può comunque consultare [1].

10.5 Conclusione

Questo capitolo ha presentato una breve introduzione alla Teoria dell'Informazione, con particolare attenzione alla codifica di sorgente senza perdite. Si tratta di un argomento molto vasto e complicato, la cui trattazione esaustiva esula dagli scopi di questo corso. Il lettore interessato ad approfondimenti può consultare il testo [7].

10.6 Riassunto formule principali sulla codifica di sorgenti

Si riportano in questa sezione le formule principali che sono state ottenute in questo capitolo. Tali formule possono essere utili per lo svolgimento degli esercizi delle esercitazioni.

Si ricorda che è fondamentale, ed è compito dello studente, conoscere a fondo sia le ipotesi sotto le quali sono state ottenute le varie formule, sia il significato esatto dei vari parametri.

Per facilitare il ritrovamento di queste formule all'interno del capitolo, quando sono state ottenute, sono state messe all'interno di un riquadro per evidenziarle meglio.

- Quantità di informazione:

$$I(m_i) = \log_2 \left(\frac{1}{p_i} \right) \text{ [bit]}$$

- Entropia di sorgente:

$$H(x) = E[I_i] = \sum_{i=1}^M p_i \log_2 \left(\frac{1}{p_i} \right) \left[\frac{\text{bit}}{\text{simbolo}} \right]$$

$$H(x) \leq \log_2(M) \left[\frac{\text{bit}}{\text{simbolo}} \right]$$

- Lunghezza media di codifica:

$$\bar{n} = E[n_i] = \sum_{i=1}^M p_i n_i \text{ [digit]}$$

- Relazioni fondamentali:

$$H(x) \leq \bar{n}$$

$$H(x) \leq \bar{n} \leq H(x) + 1$$

Bibliografia

- [1] Letizia Lo Presti, Fabio Neri, *L'Analisi dei Segnali*, Seconda edizione, CLUT, Torino, 1992.
- [2] Letizia Lo Presti, Fabio Neri, *Introduzione ai Processi Casuali*, CLUT, Torino, 1992.
- [3] Roberto Garelo, *Tutorial on digital modulations*, Torino, 2010 (disponibile sul sito Internet <http://www.tlc.polito.it/garelo/tutorial/tutorial.html>).
- [4] John Proakis, Masoud Salehi, *Communication Systems Engineering*, Prentice Hall.
- [5] Massimiliano Laddomada, Marina Mondin, *Elaborazione numerica dei segnali*, Pearson, 2007.
- [6] Athanasios Papoulis, S. Unnikrishna Pillai, *Probability, Random Variables and Stochastic Processes*, Quarta edizione, McGraw-Hill, 2002.
- [7] Khalid Sayood, *Introduction to Data Compression*, Morgan Kaufmann Publishers In.