

Appunti di Telecomunicazioni

Stefano Rossini

26 agosto 2025

Introduzione

Appunti ordinati, con approfondimenti passo-passo, del corso di Telecomunicazioni per il corso di laurea in Ingegneria Elettronica presso l'Università Politecnica delle Marche per l'A.A. 2024/2025.

Per quelli che seguono il nuovo corso di Ingegneria elettronica:

https://www.univpm.it/Entra/Offerta_formativa_1/CORSO_di_laurea_trienNALE_in_Ingegneria_Elettronica_e_delle_Tecnologie_Digitali

cioè la Laurea in Ingegneria Elettronica e delle Tecnologie Digitali, il corso prende il nome di Telecomunicazioni digitali.

Le fonti degli appunti sono le seguenti:

- Slide del corso del prof Franco Chiaraluce "Telecomunicazioni digitali A.A. 2024/2025"
- Appunti personali e quelli di Damiano Gollinucci presi a lezione
- Communication Systems Engineering, Second Edition, Proakis - Salehi ISBN 0130617938

Chiaraluce è un professore con la P maiuscola: è coinvolgente, esaustivo e chiaro.

Come dice a lezione, nel corso ci saranno delle salite e poi delle discese, un po' come la vita.

Venite a lezione, se potete, e chiedete almeno un ricevimento al prof: in tempo massimo una settimana avrete modo di chiedergli qualsiasi dubbio avete in mente.

Venite a lezione anche per le esercitazioni di Matlab, perchè si potranno capire alcuni argomenti, specialmente quelli finali: dal metodo "carta e penna" al metodo iterativo e di programmazione.

Sfruttate il ricevimento il più possibile, perchè, giustamente, a lezione non vi verranno tutti i tre miliardi di dubbi sul corso, specialmente quelli in cui il prof lascia le slide senza spiegazioni e solo con la formula.

La cosa più importante dell'orale è come ci si arriva alla formula.

È un po' come la vita: l'importante è il viaggio, non dove si arriva.

È altamente consigliato studiare e superare prima l'esame di teoria dei segnali, o come si chiama adesso "Segnali Determinati e Aleatori", perchè non capirete nulla sul corso, ve lo posso assicurare mettendo la mano sul fuoco.

Ogni concetto scritto in queste dispense, fa riferimento al corso precedente, o come lo chiama il professore a "Telecomunicazioni 1", perchè questa è "Telecomunicazioni 2".

Se vuoi dare un'occhiata ai miei vecchi appunti del corso di Teoria dei Segnali:

<https://github.com/ciccio25/appunti-teoria-dei-segnali>

I concetti importanti degli altri corsi precedenti che non sono quelli di Chiaralce sono importanti ma non fondamentali: ci serviranno i concetti base.

Lascerò dei link a video e/o spiegazioni esterne per ulteriori approfondimenti.

Per qualsiasi domanda, se trovi qualche errore grammaticale o hai bisogno di chiarimenti, scrivimi a rossini.stefano.appunti@gmail.com o contattami su Whatsapp se hai il mio numero di cellulare personale.

Se trovi qualche errore e/o vuoi contribuire agli appunti, puoi aprire un Issue sulla repository degli appunti di GitHub.

Buono studio e buona lettura

Indice

1 Rumore nei sistemi lineari	9
1.1 Banda equivalente di rumore	10
1.1.1 Osservazioni riguardo la Banda equivalente di rumore	14
1.2 Temperatura equivalente di rumore di una rete 2-porte lineare	16
1.3 Cifra di rumore di una rete 2-porte lineare	18
1.4 Cifra di rumore di una rete 2-porte lineare e temperatura equivalente di rumore generalizzata	21
1.5 2-porte in cascata	23
1.5.1 Cifra di rumore e temperatura equivalente di rumore di una cascata di reti 2-porte lineare	25
1.6 2-porte attenuatrice: rumorosità	26
1.7 Rumorosità di un sistema ricevente radio	28
1.8 Conclusioni riguardo la temperatura equivalente	29
1.9 Da segnale analogico a segnale digitale	30
1.9.1 Calcolo numerico di n tratte per un segnale analogico rispetto al segnale digitale .	31
2 Modulazioni analogiche	33
2.1 Rappresentazioni di segnali in modulazione	34
2.2 Modulazione analogica di una portante sinusoidale	36
2.3 Modulazione di ampiezza	39
2.3.1 Banda Laterale Doppia Portante Soppressa (DSB-SC)	40
2.3.1.1 Demodulazione per la DSB-SC	44
2.3.2 Modulazione di ampiezza convenzionale - AM	48
2.3.3 Banda Laterale Unica (SSB)	52
2.3.3.1 Demodulazione SSB	55
2.3.3.2 Modulazione di ampiezza con portanti in quadratura	57
2.3.4 Banda laterale ridotta - VSB	58
2.3.5 Altri schemi di modulatori di ampiezza	61
2.4 Modulazione angolare	64
2.4.1 Indici della modulazione angolare	68
2.4.1.1 Modulazione angolare a basso indice	70
2.4.2 Banda di una modulazione angolare	72
2.4.2.1 Banda di una modulazione angolare per un segnale modulante qualsiasi .	74
2.4.3 Mo/demodulatori angolari	75
2.4.3.1 Demodulatori angolari	77
3 Qualità nelle modulazioni analogiche	79
3.1 Sistema in banda base	80
3.2 Qualità della DSB-SC	82
3.3 Qualità della SSB-SC	87
3.4 Qualità della AM convenzionale	88
3.4.1 Qualità della AM convenzionale con demodulazione coerente	89
3.4.2 Qualità della AM convenzionale con demodulazione d'inviluppo	90
3.5 PLL per il recupero della portante	92
3.6 Effetto del rumore nelle modulazioni angolari	94
3.6.1 Demodulazione di segnali PM e FM	96

4 Conversione AD	99
4.1 Quantizzazione	100
4.2 Quantizzazione scalare	102
4.2.1 Funzione di quantizzazione Q	103
4.2.1.1 Rapporto segnale-rumore di quantizzazione SQNR	106
4.2.2 Quantizzazione uniforme	107
4.2.2.1 Criterio a minima distanza	109
4.2.3 Quantizzazione non uniforme	111
4.2.4 Ottimizzazione della quantizzazione non uniforme	112
4.3 Condizioni di Lloyd-Max	114
4.3.0.1 Quantizzazione uniforme e non uniforme utilizzando le tabelle	115
4.4 Quantizzazione vettoriale	117
4.4.1 Quantizzazione vettoriale ottima	118
4.5 PCM Pulse Code Modulation	119
4.5.1 PCM uniforme	120
4.5.2 PCM non uniforme	123
4.5.3 Soluzioni per ridurre la banda occupata	124
5 Teoria dell'informazione	127
5.1 Modello matematico per una sorgente di informazione	128
5.2 Sorgente discreta	129
5.2.1 Sorgente d.m.s.	130
5.3 Misura dell'informazione	131
5.4 Auto-informazione	133
5.4.1 Contenuto informativo di una d.m.s.	134
5.4.1.1 Entropia binaria	135
5.4.1.2 Entropia massima	136
5.4.1.3 Entropia congiunta	140
5.4.1.4 Entropia condizionata	142
5.4.2 Legame tra entropia condizionata ed entropia congiunta	144
5.4.2.1 Entropy rate H di una sorgente stazionaria	146
5.5 Tributo a Claude Shannon	148
5.5.1 Teorema della codifica di sorgente (Source Coding Theorem)	150
5.5.2 Efficienza di codifica	151
5.5.2.1 Codifica di sorgente con sequenze codificate di lunghezza fissa	152
5.5.3 Sequenza di lunghezza variabile	153
5.5.4 Proprietà del prefisso (prefix condition)	156
5.5.5 Algoritmo di Huffman	157
5.5.5.1 Algoritmo di Huffman: implementazione scalare e binaria	158
5.5.5.2 Algoritmo di Huffman: implementazione vettoriale	160
5.5.6 Algoritmo di Lempel-Ziv	162
5.5.6.1 Algoritmo di Lempel-Ziv: esempio	163
5.6 Rate-distortion theory	165
5.6.1 Misura della distorsione	166
5.6.1.1 Distanza di Hamming	167
5.6.1.2 Distanza in media quadratica	168
5.6.1.3 Distanza tra vettori	169
5.6.1.4 Distanza nel caso di distanza d_H	170
6 Codifica di linea	171
6.1 Cosa è la codifica di linea	172
6.1.1 Segnali unipolari: NRZ e RZ	173
6.1.1.1 Confronto spettro di potenza tra NRZ e RZ	177
6.1.2 Codifica AMI	178
6.1.3 Codifica HBDn	180
6.1.4 Altri codici di linea	181
6.2 Cenni sul problema della sincronizzazione	183

7 Formati di trasmissione numerica	185
7.1 Procedura di ortogonalizzazione di Gram-Schmidt	186
7.1.1 Due segnali ortogonali e ortonormali	188
7.1.2 Segnali ortogonali e ortonormali	189
7.1.3 Segnali come combinazione lineare della base N-dimensione	192
7.2 Esempi di segnali in spazi N-dimensional	195
7.3 Segnali in spazi di dimensione N = 1	196
7.3.0.1 Esempi di modulazioni con N = 1	200
7.4 Segnali in spazi di dimensione N = 2	203
7.4.1 Esempio di modulazione con N = 2: la PPM	208
7.4.2 Segnali bi-ortogonali	209
7.4.3 Segnali nella stessa circonferenza	211
7.4.4 Segnali in diverse circonferenze	213
7.4.5 Costellazioni in segnali passa-banda	215
7.4.5.1 Modulazioni con costellazioni di segnali passa-banda	217
7.5 Segnali ortogonali, transagonali e bi-ortogonali	222
7.5.1 PPM (Pulse Position Modulation)	224
7.5.1.1 PPM con base di impulsi rettangolari	225
7.5.2 FSK (Frequency Shift Keying)	226
7.5.3 Forme d'onda bi-ortogonali	230
7.5.4 Forme d'onda transagonali	231
7.5.4.1 Il ruolo del coefficiente di correlazione e sulla qualità della trasmissione	232
7.5.5 Dai segnali alle sequenze	234
7.6 Alcune considerazioni sullo spettro e la banda occupata	237
8 Ricevitore numerico	239
8.1 Ricevitore ottimo	240
8.2 Demodulatore (ricevitore) a correlatore	242
8.2.1 Caratteristiche del rumore additivo	244
8.2.2 Caratteristiche del vettore in uscita dal correlatore	248
8.2.2.1 M-PAM con demodulatore a correlatore	250
8.3 Demodulatore (ricevitore) a filtri adattati (Matched-Filter)	252
8.3.1 Filtro adattato	255
8.3.1.1 Proprietà del filtro adattato	256
8.3.1.2 SNR in uscita dal filtro adattato	257
8.4 Decisore	259
8.4.1 Criterio MAP: Maximum A Posteriori probability	261
8.4.2 Criterio ML: Maximum Likelihood	262
8.4.3 Distanza Euclidea minima	263
9 Prestazioni dei sistemi numerici	265
9.1 Probabilità di errore per formati antipodali	266
9.2 Probabilità di errore per un generico formato binario con simboli equiprobabili	271
9.3 Probabilità di errore per M-PAM	274
9.4 Probabilità di errore per M-ASK	277
9.5 Probabilità di errore per OOK	278
9.6 Probabilità di errore per M-PSK	280
9.7 Probabilità di errore per DPSK	281
9.8 Probabilità di errore per M-QAM	283
9.9 QAM vs. PSK	284
9.10 Probabilità di errore formati ortogonali (ricevitore a correlatore)	286
9.10.1 Limite di Shannon	287
9.11 Probabilità di errore formati bi-ortogonali e trans-ortogonali	288
9.12 Probabilità di errore M-FSK ortogonali (ricevitore non coerente)	289
9.13 Confronto tra le modulazioni fissato una probabilità di errore sul simbolo	290

10 Codifica di canale	291
10.1 Cosa è una codifica di canale	292
10.1.1 Codice a bit di parità	293
10.1.2 Codice a ripetizione	294
10.2 Caratteristiche di una codifica di canale	295
10.2.1 Distanza di Hamming	296
10.2.2 Sottospazio di una decodifica	297
10.3 Probabilità di un decodificatore di correggere un errore	299
10.4 Espansione della banda con un codice	302
10.5 Effetti del codice sull'SNR	304
10.6 Codici BCH	306
10.7 Codici MDS	308
11 Generazione di variabili aleatorie	309
11.1 Definizioni utili per le sequenze casuali	310
11.1.1 Definizioni di sequenze	311
11.1.2 Definizioni di grandezze per le sequenze casuali	312
11.1.3 Proprietà delle sequenze random	314
11.2 Generazione di sequenze pseudo-random mediante LFSR	315
11.2.1 Proprietà di una sequenza con N massimo da LFSR	318
11.2.2 Collegamenti in un LFSR per avere periodo massimo	319
11.3 Rappresentazione polinomiale di un LFSR	320
11.4 PSL di un LFSR	321
11.5 Generazione di variabili casuali	322
11.5.1 Generazione di una variabile causale con distribuzione uniforme	323
11.5.2 Generazione di una variabile aleatoria secondo una assegnata statistica	324
11.5.2.1 Dimostrazione dell'ammissibilità della trasformazione	325
11.5.2.2 Metodo di trasformazione	326
11.5.2.3 Algoritmo Box-Muller	327
12 Spettro espanso	329
12.1 Tecniche a spettro espanso	330
12.1.1 Teorema di Shannon	331
12.2 Vantaggi e modello base di un sistema SS	333
12.3 Direct-Sequence Spread-Spectrum (DS-SS)	334
12.3.1 DS-SS: da bit rate a chip rate, formulazione matematica, e spreading	335
12.3.2 DS-SS: despreading, demodulazione e interferenze a banda stretta	338
12.4 Guadagno di processo (Process Gain)	340
12.5 Ruolo delle sequenze PN	341
12.6 Probabilità di errore in ricezione	342
12.6.1 Interferente monocromatico	345
12.6.2 Interference margin	346
12.6.2.1 Interference margin: esempio numerico	347
12.6.3 Interferente a banda larga	348
12.7 DS-SS: applicazioni	350
12.7.1 Low detectability signal transmission	351
12.7.2 CDMA	352
12.7.2.1 CDMA: valutazione di massima delle prestazioni (worst case scenario)	354
12.7.2.2 CDMA: Trasmettitore e ricevitore	355
12.7.3 CDMA: valutazione più accurata	356
12.7.4 Auto-correlazione, cross-correlazione, sequenza PN, prestazioni CDMA	359
12.7.5 Communications over multipath channels	361
12.7.6 Pulsed interference	362
12.7.6.1 Pulsed interference: worst case scenario	363
12.7.6.2 Pulsed interference: prestazioni	364
12.7.6.3 Pulsed interference: miglioramento delle prestazioni tramite interlever	365
12.8 Generazione di sequenze PN	366
12.8.1 Maximum-length sequences (m-sequences)	367
12.8.1.1 Proprietà delle m-sequences	369
12.8.2 Cross-correlazione	370
12.9 Gold-sequences	371

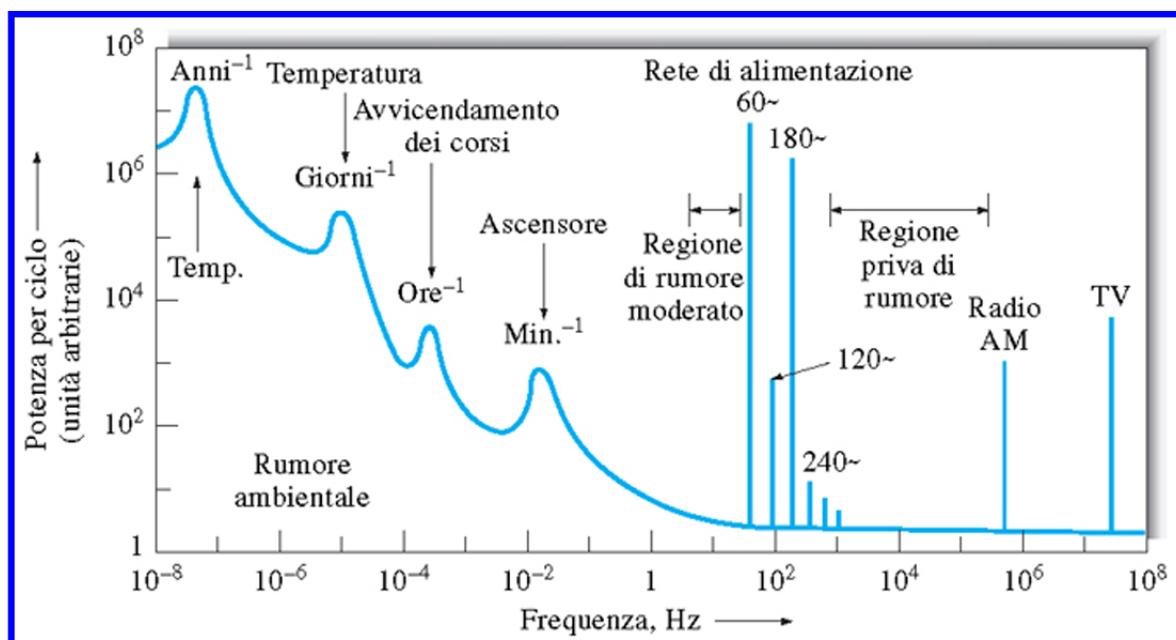
12.10Lower bound	373
12.11Kasami sequences (small set)	374
12.12Kasami sequences (large set)	376
12.13Sequenze di Barker	377
12.14Frequency-Hopping Spread-Spectrum (FH-SS)	378
12.14.1 Classificazione dei sistemi FH-SS	380

Tabella lettere greche utili per il corso

Lettera maiuscola	Lettera minuscola	Nome lettera greca
A	α	Alfa
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ε	Epsilon
H	η	Eta
Θ	θ	Theta
K	κ	Kappa
Λ	λ	Lambda
M	μ	Mu
Π	π	Pi
P	ρ	Rho
Σ	σ	Sigma
T	τ	Tau
Φ	ϕ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Capitolo 1

Rumore nei sistemi lineari



1.1 Banda equivalente di rumore

1

Dato un sistema lineare:



dove $x(t)$ è il segnale di ingresso e $y(t)$ è il segnale di uscita dal sistema lineare, dal corso di Teoria dei Segnali, è noto che una generica rete 2-porte, anche contenente elementi attivi, è responsabile dell'introduzione di un contributo di rumore termico, che si sovrappone al segnale utile (i.e. segnale di ingresso) e ne degrada le caratteristiche.

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 5 - Sistemi lineari - pag 41

Allo stesso tempo, un rumore termico presente all'ingresso di una rete lineare, viene modificato dal transito di essa, e si ritrova in uscita amplificato (o attenuato) oltre che, eventualmente, con una minore occupazione spettrale (determinata dalle caratteristiche filtranti della rete resistiva).

Un sistema lineare, a differenza di un sistema non lineare, può solo togliere le componenti in frequenza dal segnale di ingresso iniziale e non ne aggiunge di nuove

Questi aspetti devono essere combinati per poter analizzare come la potenza di rumore viene alterata dal passaggio attraverso un sistema lineare e, soprattutto, come tale potenza aumenta quando più sistemi lineari vengono posti in cascata.

Un primo concetto utile, nel contesto delineato, è quello di banda equivalente di rumore.

Una rete 2-porte lineare è caratterizzata dalla sua funzione di trasferimento $H(f)$.

Quando un rumore termico, con densità spettrale di potenza bilatera $\frac{\eta}{2}$ (costante sia per le frequenze positive che per quelle negative) è applicato all'ingresso della $H(f)$, la potenza di rumore in uscita sarà data da:

¹Slide del prof — Parametri di rumore sistema lineare — pag 1
Appunti di Damiano — pag 1

$$\begin{aligned}
< n^2 > &= \int_{-\infty}^{+\infty} p_{n_h}(f) df \\
&= \int_{-\infty}^{+\infty} p_h(f) \cdot p_n(f) df \\
&= \int_{-\infty}^{+\infty} |H(f)|^2 \cdot |N(f)|^2 df \\
&= \int_{-\infty}^{+\infty} |H(f)|^2 \cdot \frac{\eta}{2} df \\
&= \frac{\eta}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df
\end{aligned} \tag{1.1}$$

dove:

- $p_{n_h}(f)$ si intende la potenza del rumore (noise dall'inglese) dopo il sistema lineare
- $p_h(f)$ si intende la potenza del sistema lineare
- $p_n(f)$ si intende la potenza del rumore

Per ripassare i sistemi lineari

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 5.2 Densità spettrale di energia e di potenza di un filtro — pag 44

Dato $x(t)$ il segnale di ingresso e $y(t)$ il segnale di uscita di un sistema lineare, possiamo definire le loro rappresentazioni in Fourier come $X(\omega)$ e $Y(\omega)$.

Essendo:

$$Y(\omega) = H(\omega)X(\omega) \quad (1.2)$$

le densità spettrali di energia (ove applicabili) dei segnali in ingresso e in uscita sono tra loro legate dalla seguente relazione:

$$|Y(\omega)|^2 = |H(\omega)X(\omega)|^2 = |H(\omega)|^2 \cdot |X(\omega)|^2 \quad (1.3)$$

Una relazione analoga vale per gli spettri di potenza, indicati con $p_x(\omega)$ (per l'ingresso) e $p_y(\omega)$ (per l'uscita):

$$p_y(\omega) = |H(\omega)|^2 p_x(\omega) \quad (1.4)$$

■

Dunque, la densità in uscita si ottiene moltiplicando la densità in ingresso per il modulo al quadrato della funzione di trasferimento.

Per ripassare il rumore

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 14.1 Un esempio di processo stocastico: il rumore termico — pag 162 - 163

Ponendo T come il valore della temperatura assoluta a cui si trova il sistema lineare in kelvin, dove la relazione tra temperatura espressa in kelvin e gradi celsius è la seguente:

$$0 \text{ K} = -273.15 \text{ }^\circ\text{C} \quad (1.5)$$

Sapendo che:

$$K = 1.38 \cdot 10^{-23} \frac{J}{\text{K}} \quad (1.6)$$

è la costante di Boltzmann.

Sapendo che il rumore è un processo "stocastico", possiamo esprimere la potenza del rumore indipendente dal tempo come:

$$\langle P_n \rangle = KTB \quad (1.7)$$

dove P_n è la "potenza disponibile".

La densità spettrale di potenza sarà:

$$\begin{aligned} p_n(f) &= \frac{KT}{2} \\ &= \frac{\eta}{2} \end{aligned} \quad (1.8)$$

L'espressione:

$$\langle n^2 \rangle = \frac{\eta}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df \quad (1.9)$$

mette in evidenza che, una volta assegnato il valore di $\frac{\eta}{2}$, la potenza di rumore in uscita dipende solo dall'andamento della funzione di trasferimento del sistema lineare.

Sapendo che:

$$\frac{1}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df = |H_0|^2 B_N \quad (1.10)$$

dove:

- $|H_0|$ è l'ampiezza del sistema lineare
- B_N è la banda del rumore di ingresso al sistema lineare

Isolando B_N dalla formula precedente, con dei semplici passaggi algebrici:

$$\begin{aligned} \frac{1}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df &= |H_0|^2 B_N \\ \Downarrow \\ B_N &= \frac{1}{2|H_0|^2} \int_{-\infty}^{+\infty} |H(f)|^2 df \end{aligned} \quad (1.11)$$

Grazie alla relazione di B_N , possiamo riscrivere la potenza del segnale di rumore come:

$$\begin{aligned} \langle n^2 \rangle &= \frac{\eta}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df \\ &= \eta \frac{1}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df \\ &= \eta |H_0|^2 B_N \\ &= |H_0|^2 \eta B_N \end{aligned} \quad (1.12)$$

$|H_0|^2$, in questa relazione di $\langle n^2 \rangle$, ha il significato di guadagno in potenza, che possiamo esprimere anche come guadagno disponibile G_d della rete 2-ponte.

B_N prende il nome di "banda equivalente di rumore" della rete 2-ponte.

B_N rappresenta la larghezza di banda di un sistema lineare con funzione di trasferimento costante e pari a $|H_0|$, che fornisce in uscita la stessa potenza di rumore che si ha in uscita da $H(f)$.

Dal punto di vista della potenza di rumore in uscita, i due sistemi sono perfettamente equivalenti.

1.1.1 Osservazioni riguardo la Banda equivalente di rumore

²

D'altro canto, per un dato sistema, l'integrale della formula di B_N :

$$B_N = \frac{1}{2|H_0|^2} \int_{-\infty}^{+\infty} |H(f)|^2 df \quad (1.13)$$

può essere calcolato una volta e può essere utilizzato ai fini della valutazione della potenza di rumore in uscita:

$$\langle n^2 \rangle = |H_0|^2 \eta B_N \quad (1.14)$$

Quindi passeremo da un complicato integrale da svolgere:

$$\langle n^2 \rangle = \frac{\eta}{2} \int_{-\infty}^{+\infty} |H(f)|^2 df \quad (1.15)$$

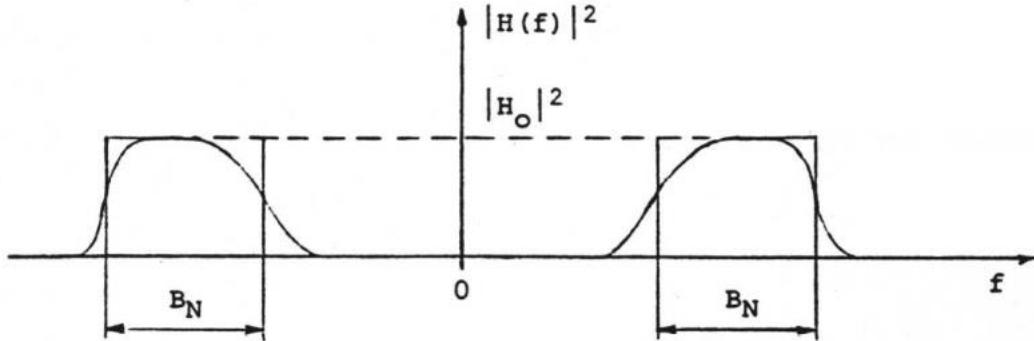
a una semplice moltiplicazione.

In effetti, il valore di B_N viene normalmente specificato insieme alle altre caratteristiche del sistema, o comunque predeterminato per uno specifico servizio.

Ad esempio, nel caso di un segnale televisivo analogico, la banda equivalente di rumore del ricevitore può essere posta uguale a 5 MHz, vale a dire la massima frequenza del segnale modulante in banda base.

La conoscenza della relazione tra potenza del rumore dopo un sistema lineare $\langle n^2 \rangle$ e B_N , consente di semplificare il calcolo della potenza di rumore in uscita da una rete 2-porte quando il rumore in ingresso è a spettro piatto.

Notando un grafico di esempio riguardo alla valutazione della banda equivalente di rumore, per un generico andamento di $H(f)$:



Se ci troviamo il caso uno spettro di rumore "colorato" come in figura (vale a dire a spettro non uniforme), B_N non ha alcuna implicazione pratica diretta, in quanto lo spettro in uscita si ottiene dalla combinazione del modulo al quadrato della funzione di trasferimento e delle spettri di potenza in ingresso, ed è sempre il risultato del loro prodotto a dover essere integrato per ricavare la potenza in uscita.

In altre parole, per calcolare la potenza del rumore nel caso ideale del segnale rettangolare, abbiamo:

$$\begin{aligned} \langle n^2 \rangle &= \frac{\eta}{2} \cdot |H_0|^2 \cdot 2B \\ &= \eta B |H_0|^2 \end{aligned} \quad (1.16)$$

²Slide del prof — Parametri di rumore sistema lineare — pag 1 - 2
Appunti di Damiano — pag 1 - 2

e possiamo dire che per un determinato B_N , la partenza del sistema equivalente è uguale a quello del sistema.

1.2 Temperatura equivalente di rumore di una rete 2-porte lineare

3

Si consideri ora una rete 2-porte rumorosa, al cui ingresso si connette un resistore rumoroso a temperatura T_i .

All'uscita della rete 2-porte, si ha il rumore di ingresso amplificato (o attenuato) ed il rumore introdotto dalla rete 2-porte rumorosa.

Le fonti di rumore interne della rete sono, di regola, incorrelate rispetto al rumore generato dal resistore connesso in ingresso (i.e. sono indipendenti ed hanno covarianza nulla).

Pertanto, la potenza di rumore disponibile all'uscita della rete è uguale alla somma delle potenze di rumore disponibile dovute dalla:

- rumorosità della rete
- rumore di ingresso amplificato (o attenuato) dalla rete

La stessa considerazione vale per le densità di potenza di rumore:

$$\frac{\eta'}{2} = G_d \frac{\eta}{2} + \frac{\delta\eta}{2} \quad (1.17)$$

dove:

- $\frac{\eta'}{2}$ è il segnale in uscita dalla rete 2-porte rumorosa, i.e. la densità di potenza in uscita
- G_d è il rumore generato dal sistema lineare
- $\frac{\eta}{2}$ è il segnale di ingresso con rumore
- $\frac{\delta\eta}{2}$ è la densità spettrale bilatera dovuta alla rete, i.e. la densità spettrale bilatera dovuta alla rete

Sapendo che:

$$\begin{cases} \eta = K \cdot T_i \\ \eta' = K \cdot T_u \end{cases} \quad (1.18)$$

dove:

- T_i è la temperatura in ingresso dal sistema lineare
- T_u è la temperatura in uscita dal sistema lineare

L'equazione della densità di potenza in uscita $\frac{\eta'}{2}$ diventa:

$$\begin{aligned} \frac{\eta'}{2} &= G_d \frac{\eta}{2} + \frac{\delta\eta}{2} \\ &\downarrow \\ \frac{K \cdot T_u}{2} &= G_d \frac{K \cdot T_i}{2} + \frac{\delta\eta}{2} \\ &= \frac{1}{2}(G_d \cdot K \cdot T_i + \delta\eta) \\ K \cdot T_u &= G_d \cdot K \cdot T_i + \delta\eta \end{aligned} \quad (1.19)$$

³Slide del prof — Parametri di rumore sistema lineare — pag 2 - 4
Appunti di Damiano — pag 2 - 4

Sapendo che la densità spettrale alla rete $\delta\eta$ si trova ad una temperatura T_e , possiamo esprimere la come:

$$\delta\eta = G_d \cdot K \cdot T_e \quad (1.20)$$

l'espressione della densità di potenza in uscita $\frac{\eta'}{2}$ possiamo continuare a semplificare ancora come:

$$\begin{aligned} K \cdot T_u &= G_d \cdot K \cdot T_i + \delta\eta \\ &\downarrow \\ K \cdot T_u &= G_d \cdot K \cdot T_i + G_d \cdot K \cdot T_e \\ &= G_d \cdot K(T_i + T_e) \end{aligned} \quad (1.21)$$

T_e prende il nome di "temperatura equivalente di rumore" della rete 2-ponte.

T_e rappresenta l'incremento di temperatura che occorre dare al resistore connesso all'ingresso della rete 2-ponte supposta non rumorosa, per avere all'uscita la stessa densità di potenza disponibile di rumore che si ha con il generatore (di rumore) in ingresso a temperatura T_i e rete 2-ponte rumorosa.

O, spiegato in un'altra maniera, usando i sistemi equivalenti, la temperatura equivalente sposta tutto il rumore nell'ingresso, azzerando il rumore del sistema lineare.

Quindi, T_e è un parametro caratteristico della rete e ne quantifica in maniera semplice ed esplicita la rumorosità.

È sufficiente aumentare di T_e la temperatura di rumore in ingresso per poter considerare la rete 2-ponte non rumorosa, pur ottenendo in uscita la stessa densità di potenza (e quindi la stessa potenza) di rumore del sistema reale.

T_u rappresenta la temperatura equivalente di rumore complessiva in uscita, e può essere utilizzata per caratterizzare una cascata di reti 2-ponte.

Quando si descrive la rumorosità del singolo dispositivo, si è soliti riferirsi alla sua sezione in ingresso, assegnando quindi direttamente il valore di T_e .

1.3 Cifra di rumore di una rete 2-porte lineare

⁴

In alternativa, alla temperatura equivalente di rumore, la rumorosità introdotta da una rete 2-porte può essere assegnata specificando la cosiddetta "cifra di rumore" (o "fattore di rumore") F .

F è definita come rapporto tra la densità di potenza disponibile di rumore in uscita da una rete 2-porte rumorosa, al cui ingresso sia connesso un generatore di rumore a temperatura standard T_0 :

$$\begin{aligned} T_0 &= 290^\circ K \\ &= (290 - 273.15)^\circ C \\ &= 16.85^\circ C \\ &\approx 17^\circ C \end{aligned} \tag{1.22}$$

e la densità di potenza disponibile di rumore dovuta all'amplificazione del rumore immesso in ingresso dal generatore.

In formula, sulla base della notazione precedente:

$$\begin{aligned} F &= \frac{\eta' \Big|_{T_i=T_0}}{G_d \cdot K \cdot T_0} \\ &= \frac{G_d \cdot K(T_i + T_e) \Big|_{T_i=T_0}}{G_d \cdot K \cdot T_0} \\ &= \frac{G_d \cdot K(T_0 + T_e)}{G_d \cdot K \cdot T_0} \\ &= \frac{T_0 + T_e}{T_0} \\ &= \frac{T_0(1 + \frac{T_e}{T_0})}{T_0} \\ &= 1 + \frac{T_e}{T_0} \end{aligned} \tag{1.23}$$

Generalmente, e nella vita pratica, F cifra di rumore è sempre maggiore di 1, ma a livello teorico valgono le seguenti considerazioni teoriche.

La relazione di F con T_e e T_0 :

$$F = 1 + \frac{T_e}{T_0} \tag{1.24}$$

stabilisce un legame esplicito tra temperatura equivalente di rumore e cifra di rumore della rete 2-porte, a conferma dell'equivalenza "operativa" delle due funzioni.

In certi casi, può essere utile esplicitare il legame inverso, cioè:

⁴Slide del prof — Parametri di rumore sistema lineare — pag 3 -
Appunti di Damiano — pag 3 -

$$\begin{aligned}
 F &= 1 + \frac{T_e}{T_0} \\
 &\downarrow \\
 F - 1 &= \frac{T_e}{T_0} \\
 T_0(F - 1) &= T_e \\
 T_e &= T_0(F - 1)
 \end{aligned} \tag{1.25}$$

D'altro canto, moltiplicando numeratore e denominatore dall'equazione di F per la banda equivalente di rumore della rete 2-porte, la cifra di rumore F può anche essere riscritta come:

$$\begin{aligned}
 F &= 1 + \frac{T_e}{T_0} \\
 &\downarrow \\
 F &= \frac{\langle n'^2 \rangle}{G_d \cdot \langle n^2 \rangle}
 \end{aligned} \tag{1.26}$$

dove:

- $\langle n^2 \rangle$ è la potenza di rumore dovuta al solo resistore a temperatura T_0 in ingresso (che abbiamo precedentemente calcolato)
- $\langle n'^2 \rangle$ è la potenza di rumore totale in uscita (comprensivo del contributo dovuto alla rete 2-porte)

Grazie a questa ultima relazione di F :

$$F = \frac{\langle n'^2 \rangle}{G_d \cdot \langle n^2 \rangle} \tag{1.27}$$

la cifra di rumore F esprime il rapporto tra il rumore in uscita dal 2-ponte e quello che si avrebbe (sempre in uscita, stante la moltiplicazione di $\langle n^2 \rangle$ per G_d) se il 2-ponte fosse ideale, cioè non introduceisse alcuna rumorosità aggiuntiva.

È evidente che:

$$F \geq 1 \tag{1.28}$$

e, quanto più F risulta prossimo a 1, tanto più il comportamento del sistema lineare risulta migliore ai fini della rumorosità.

Tenendo conto che un segnale utile di potenza S_i in ingresso si ritrova amplificato (o attenuato) e pari a $G_d \cdot S_i$ in uscita, si può semplificare ulteriormente numeratore e denominatore a secondo membro della funzione di F per S_i , ottenendo così:

$$\begin{aligned}
F &= \frac{\langle n'^2 \rangle}{G_d \cdot \langle n^2 \rangle} \\
&= \frac{S_i \cdot \langle n'^2 \rangle}{S_i \cdot G_d \cdot \langle n^2 \rangle} \\
&= \frac{\frac{S_i}{\langle n'^2 \rangle}}{\frac{S_i \cdot G_d}{\langle n^2 \rangle}} \quad (1.29) \\
&= \frac{\left(\frac{S}{N}\right)_i}{\left(\frac{S}{N}\right)_u}
\end{aligned}$$

$$= \frac{SNR_i}{SNR_u}$$

dove:

- SNR_i esprime il rapporto segnale-rumore in ingresso, in presenza del solo resistore a temperatura T_0
- SNR_u esprime il rapporto segnale-rumore in uscita, includendo anche la rumorosità della rete a 2-ponte

In condizioni ideali, per la rete a 2-ponte, questo rapporto varrebbe 1, poiché segnale utile e rumore in ingresso verrebbero amplificati (o attenuati) della stessa quantità G_d .

In condizioni reali (cioè nel caso di un rete 2-ponte rumorosa) F è sempre maggiore di 1 ed esprime il peggioramento nel rapporto segnale-rumore dovuto al 2-ponte.

1.4 Cifra di rumore di una rete 2-porte lineare e temperatura equivalente di rumore generalizzata

5

È opportuno precisare che la definizione di cifra di rumore e temperatura equivalente di rumore si estende anche nel caso in cui la rumorosità introdotta dal sistema lineare è caratterizzata da uno spettro di potenza non uniforme e il guadagno disponibile è variabile con la frequenza.

In questo caso la F dipenderà dalla frequenza f, quindi passeremo al caso costante per ogni frequenza al caso generalizzato:

$$\begin{aligned} F &= 1 + \frac{T_e}{T_0} \\ &\downarrow \\ F(f) &= 1 + \frac{\delta\eta(f)}{\eta G_d(f)} \end{aligned} \tag{1.30}$$

Analogamente, possiamo generalizzare T_e per ogni frequenza f:

$$\begin{aligned} T_e &= T_0(F - 1) \\ &\downarrow \\ T_e(f) &= T_0 \cdot [(F(f) - 1)] \end{aligned} \tag{1.31}$$

Nell'ottica di caratterizzare con un unico numero la rete 2-ponte, anche in questa situazione più generale, F(f) e $T_e(f)$ possono essere mediati sulla banda passante del rete 2-ponte, in tal modo ottenendo:

$$\begin{aligned} \bar{F} &= \frac{\int_{-\infty}^{+\infty} F(f) \cdot G_d(f) df}{\int_{-\infty}^{+\infty} G_d(f) df} \\ &= \frac{1}{2|H_0|^2 B_N} \int_{-\infty}^{+\infty} F(f) \cdot G_d(f) df \end{aligned} \tag{1.32}$$

e

$$\begin{aligned} \bar{T}_e &= \frac{\int_{-\infty}^{+\infty} T_e(f) \cdot G_d(f) df}{\int_{-\infty}^{+\infty} G_d(f) df} \\ &= \frac{1}{2|H_0|^2 B_N} \int_{-\infty}^{+\infty} T_e(f) \cdot G_d(f) df \\ &= T_0 \cdot (\bar{F} - 1) \end{aligned} \tag{1.33}$$

⁵Slide del prof — Parametri di rumore sistema lineare — pag 4 - 5
Appunti di Damiano — pag 4 - 5

Le relazioni di \overline{F} e $\overline{T_e}$ sono valide se:

$$G_d(f) = |H(f)|^2 \quad (1.34)$$

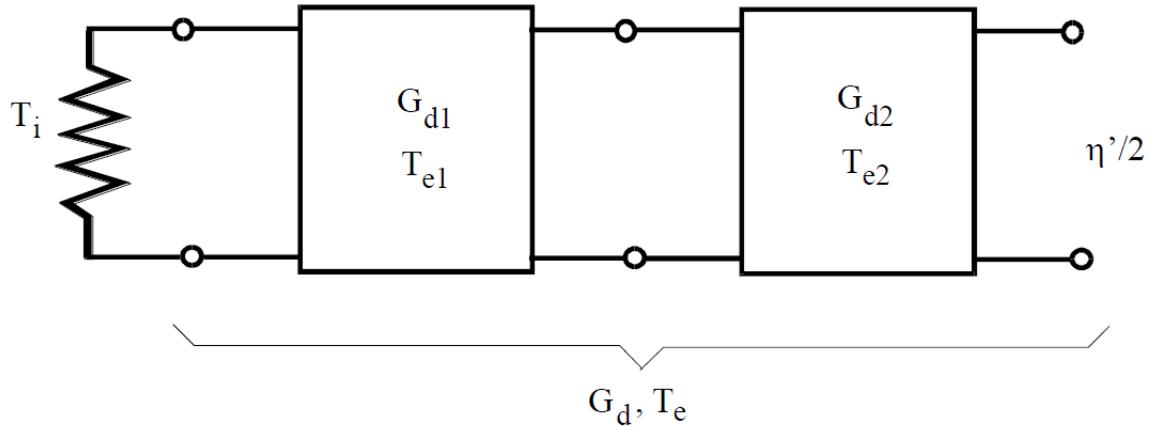
D'ora in poi non consideriamo la dipendenza dalla frequenza dei parametri per semplicità di notazione, ma bisogna ricordarci che è possibile che esista per la teoria

1.5 2-porte in cascata

6

Le espressioni sin qui introdotto sono sufficienti per descrivere la rumorosità di una singola rete 2-porte.

Cosa succede quando un certo numero di reti 2-porte vengono connessi in cascata come in figura?



La resistenza in ingresso, che schematizza una generica sorgente di rumore termico, si trova a temperatura T_i .

Ciascuna rete 2-porte è caratterizzata dal proprio guadagno disponibile, rispettivamente G_{d1} e G_{d2} , e dalla propria temperatura equivalente di rumore, rispettivamente T_{e1} e T_{e2} .

Se:

$$G_{d1} > 1 \quad (1.35)$$

la rete a 2-porte viene identificato come amplificatore.

Se:

$$G_{d1} < 1 \quad (1.36)$$

la rete a 2-porte viene identificato come attenuatrice.

Considerando solo il primo sistema lineare (quello a sinistra della figura), possiamo esprimere la temperatura di rumore complessiva T_{u1} del primo sistema lineare (come studiato nelle scorse sezioni) come:

$$\begin{aligned} K \cdot T_{u1} &= G_d \cdot K \cdot (T_i + T_{e1}) \\ T_{u1} &= G_d \cdot (T_i + T_{e1}) \end{aligned} \quad (1.37)$$

T_{u1} può essere interpretata come la temperatura equivalente in ingresso alla seconda rete 2-porte, in uscita dalla quale si avrà allora:

$$\begin{aligned} T_{u2} &= G_{d2} \cdot (T_{u1} + T_{e2}) \\ &= G_{d2} \cdot [G_{d1}(T_i + T_{e1}) + T_{e2}] \\ &= G_{d1} \cdot G_{d2} \cdot T_i + G_{d1} \cdot G_{d2} \cdot T_{e1} + G_{d2} \cdot T_{e2} \end{aligned} \quad (1.38)$$

⁶Slide del prof — Parametri di rumore sistema lineare — pag 5 - 6
Appunti di Damiano — pag 5 - 6

Tenendo conto che il guadagno disponibile della cascata G_d vale:

$$\begin{aligned} G_d &= G_{d1} \cdot G_{d2} \\ &\Downarrow \\ G_{d2} &= \frac{G_d}{G_{d1}} \end{aligned} \quad (1.39)$$

e che, per definizione di temperatura equivalente, si deve poter scrivere:

$$T_{u2} = G_d \cdot (T_i + T_e) \quad (1.40)$$

con T_e temperatura equivalente della cascata, per confronto si ricava:

$$\begin{aligned} T_{u2} &= G_{d1} \cdot G_{d2} \cdot T_i + G_{d1} \cdot G_{d2} \cdot T_{e1} + G_{d2} \cdot T_{e2} \\ G_d \cdot (T_i + T_e) &= G_d \cdot T_i + G_d \cdot T_{e1} + G_{d2} \cdot T_{e2} \\ G_d \cdot T_i + G_d \cdot T_e &= \\ G_d \cdot T_e &= G_d \cdot T_{e1} + G_{d2} \cdot T_{e2} \\ &= G_d \cdot T_{e1} + \frac{G_d}{G_{d1}} \cdot T_{e2} \\ &= G_d \cdot \left(T_{e1} + \frac{T_{e2}}{G_{d1}} \right) \\ T_e &= T_{e1} + \frac{T_{e2}}{G_{d1}} \end{aligned} \quad (1.41)$$

La formula di T_e appena scritta esplicita che la temperatura equivalente di due sistemi lineari in cascata si ottiene sommando alla temperatura equivalente del primo stadio T_{e1} la temperatura equivalente del secondo stadio T_{e2} per il guadagno del primo stadio G_{d1} .

Inoltre, la formula di T_e fornisce un'altra informazione quantitativa estremamente importante: doven-
do scegliere la collocazione relativa, ad esempio, di due stadi amplificatori, tendenzialmente si metterà in testa quello meno rumoroso, in quanto la rumorosità dell'altro stadio (maggiore) sarà ridotta del fattore G_{d1} .

Se il primo stadio è un attenuatore, cioè $G_{d1} < 1$, la rumorosità del secondo stadio verrebbe amplifi-
cata.

In questo caso, è meglio, se possibile, porre l'attenuatore dopo l'amplificatore.

È per questo motivo per cui il pre-amplificatore, per un sistema di ricezione televisivo, viene posto a
contatto con l'antenna e prima del cavo che porta il segnale al cinescopio (la cosiddetta "discesa d'an-
tenna").

1.5.1 Cifra di rumore e temperatura equivalente di rumore di una cascata di reti 2-porte lineare

⁷

In generale, la disposizione relativa dei componenti di una cascata va ottimizzata, tenendo conto dei valori relativi del guadagno e della temperatura equivalente di rumore, in modo da minimizzare la formula di T_e :

$$T_e = T_{e1} + \frac{T_{e2}}{G_{d1}} \quad (1.42)$$

La procedura che ha consentito di ricavare la temperatura equivalente di rumore per una coppia di reti 2-porte, può essere estratta ad un numero n qualsiasi di sistemi lineari connessi in cascata.

La temperatura equivalente di rumore globale si ottiene dalla seguente espressione:

$$T_e = T_{e1} + \frac{T_{e2}}{G_{d1}} + \frac{T_{e3}}{G_{d1} \cdot G_{d2}} + \cdots + \frac{T_{en}}{G_{d1} \cdot G_{d2} \cdot \cdots \cdot G_{dn-1}} \quad (1.43)$$

Una eguale espressione compatta, sapendo il legame tra temperatura equivalente e cifra di rumore, si ricava come:

$$F = F_1 + \frac{F_2 - 1}{G_{d1}} + \frac{F_3 - 1}{G_{d1} \cdot G_{d2}} + \cdots + \frac{F_n - 1}{G_{d1} \cdot G_{d2} \cdot \cdots \cdot G_{dn-1}} \quad (1.44)$$

Questa ultima formula di F è nota come "formula di Friis".

Grazie alla formula di Friis, possiamo fare le stesse considerazioni sulla cascata delle n porte partendo dalla cascata dei due sistemi lineari.

La rumorosità del primo stadio, qui espressa da F_1 , è, in generale, la più importante nell'ambito della cascata (e va dunque adeguatamente controllata).

Si è già detto in precedenza che, uno o più stadi della cascata, possono, in realtà, presentare un comportamento da attenuatore, nel caso in cui:

$$G_d = \frac{1}{A_d} < 1 \quad (1.45)$$

e, più che il guadagno, risulta significativo considerare l'attenuazione disponibile A_d , che risulta maggiore di 1.

Frequentemente, sia G_d che A_d , sono espresse in unità logaritmiche (dB), in cui avere un numero in unità assolute, maggiore di 1 (o minore di 1), corrisponde a una misura in dB positiva (o negativa).

⁷Slide del prof — Parametri di rumore sistema lineare — pag 6 -
Appunti di Damiano — pag 6 -

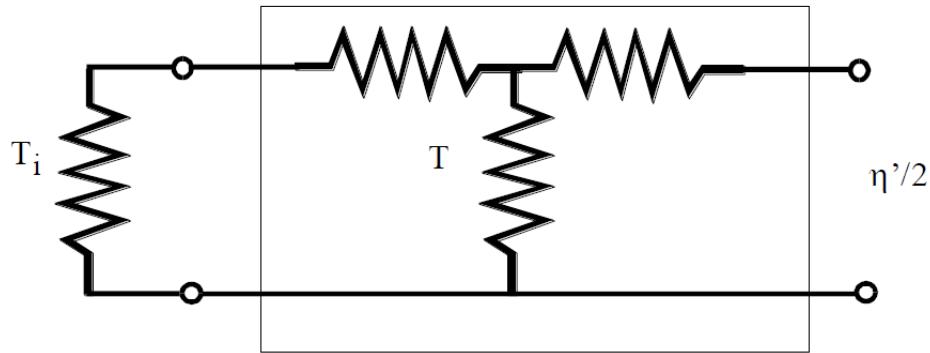
1.6 2-porte attenuatrice: rumorosità

8

È interessante chiedersi quanto valga la rumorosità di una rete 2-porte attenuatrice.

Essa è, tipicamente, realizzata utilizzando un attenuatore resistivo, in cui l'attenuazione è il risultato della dissipazione di potenza elettrica immensa nelle resistenze.

Dal punto di vista grafico, possiamo schematizzarlo in questa maniera:



Si suppone che l'attenuatore sia collegato in ingresso ad un generatore di rumore a temperatura T_i , in generale diversa dalla temperatura T dell'attenuatore.

Si può pensare che la densità di potenza disponibile nel sistema considerato sia divisa in due parti:

- la prima, pari a $K \cdot \frac{T_i - T}{2}$ è relativa all'eccesso di temperatura T_i rispetto a T
- la seconda, pari a $K \cdot \frac{T}{2}$ sarebbe invece, la sola densità di potenza disponibile nel caso in cui il resistore fosse alla stessa temperatura T dell'attenuatore

Se:

$$T_i = T \quad (1.46)$$

l'insieme del resistore esterno all'attenuatore e dei resistori interni costituirebbero un'unica resistenza a temperatura T .

La densità spettrale di potenza di rumore $\langle P_n \rangle$ disponibile è indipendente dal valore della resistenza.

Se si considera solo f positive, $\langle P_n \rangle$ sarà uguale a:

$$\begin{aligned} \langle p_n \rangle &= \int_0^{+B} \eta(f) df \\ &= K \cdot T \end{aligned} \quad (1.47)$$

Se si considera solo f positive e f negative, $\langle P_n \rangle$ sarà uguale a:

$$\begin{aligned} \langle p_n \rangle &= \int_{-B}^{+B} \eta(f) df \\ &= K \cdot \frac{T}{2} \end{aligned} \quad (1.48)$$

⁸Slide del prof — Parametri di rumore sistema lineare — pag 6 - 7
Appunti di Damiano — pag 6 - 7

L'incremento di R dovuto all'attenuatore, non modificherebbe in alcun modo la rumorosità in ingresso.

Il contributo $K \cdot \frac{T}{2}$ resta tale qualunque sia la sezione in cui si colloca (in particolare in uscita dell'attenuatore).

Il contributo $K \cdot \frac{T_i - T}{2}$, invece, che è relativo all'eccesso di temperatura del resistore in ingresso, deve essere moltiplicato per il guadagno, ovvero, diviso per l'attenuazione disponibile.

In definitiva, la densità di potenza di rumore nella sezione di uscita vale:

$$\begin{aligned} \frac{\eta'}{2} &= \frac{K \cdot T}{2} + \frac{K \cdot (T_i - T)}{2 \cdot A_d} \\ &= \frac{K \cdot T_i}{2 \cdot A_d} + \frac{K \cdot T}{2} \cdot \left(1 - \frac{1}{A_d}\right) \end{aligned} \quad (1.49)$$

Indicando con T_e la temperatura equivalente dell'attenuatore, possiamo riscrivere la densità di potenza di rumore nella sezione di uscita come:

$$\begin{aligned} \frac{\eta'}{2} &= \frac{K \cdot (T_i - T_e)}{2 \cdot A_d} \\ &= \frac{K \cdot T_i}{2 \cdot A_d} + \frac{K \cdot T_e}{2 \cdot A_d} \end{aligned} \quad (1.50)$$

Uguagliando le due formule di $\frac{\eta'}{2}$ e svolgendo alcuni passi algebrici, possiamo ricavare T_e come:

$$T_e = T \cdot (A_d - 1) \quad (1.51)$$

Questa ultima espressione di T_e esprime la temperatura equivalente di un attenuatore resistivo.

Di conseguenza, la cifra di rumore che corrisponde ad un attenuatore resistivo, vale:

$$F = 1 + \frac{T}{T_0} \cdot (A_d - 1) \quad (1.52)$$

Nel caso particolare, ma importante, in cui l'attenuatore si trovi a temperatura standard T_0 , la sua cifra di rumore uguaglia l'attenuazione disponibile, cioè:

$$F = A_d \quad (1.53)$$

Perciò, maggiore è l'attenuazione, maggiore è la rumorosità introdotta dall'attenuatore.

Le considerazioni svolte per gli attenuatori resistivi valgono per tronchi di guida d'onda o di linee che introducono attenuazione.

1.7 Rumorosità di un sistema ricevente radio

⁹

Nel caso in cui si debba caratterizzare la rumorosità di un sistema ricevente per collegamenti radio, in cui l'insieme degli apparati che costituiscono il ricevitore (cavi, amplificatori, mixer, ...) sia chiuso in ingresso dall'antenna ricevente, la temperatura di ingresso T_i deve essere assunta la temperatura di antenna T_A .

La somma della temperatura equivalente del radioricevitore prende il nome di temperatura di sistema T_S , che si ottiene come:

$$T_S = T_A + T_e \quad (1.54)$$

T_A non è la temperatura fisica dell'antenna, ma dipende da dove viene puntata.

Quando $T_A \approx T_0$, T_S si può esprimere come:

$$T_S \approx T_0 \cdot F_r \quad (1.55)$$

dove F_r è la cifra di rumore del ricevitore.

⁹Slide del prof — Parametri di rumore sistema lineare — pag 8
Appunti di Damiano — pag 8

1.8 Conclusioni riguardo la temperatura equivalente

¹⁰

In chiusura, si vuole ribadire l'importanza applicativa della definizione di temperatura equivalente (o cifra di rumore a cui essa è legata).

Trasferendo la rumorosità dovuta alla rete 2-ponte (o alla cascata di reti 2-ponte) in un incremento della temperatura che schematizza la rumorosità in ingresso, essa consente di considerare la rete stessa non rumorosa.

Come conseguenza, il rapporto segnale-rumore è costante in qualunque sezione del sistema lineare in cui venga calcolato.

Segnale utile e rumore sono trattati allo stesso modo da blocchi ideali.

Ovviamente, ciò che resta invariato è un rapporto segnale-rumore in cui la potenza di rumore è onnicomprensiva e tiene conto, in qualunque sezione, anche della rumorosità dei blocchi posti a valle.

D'altro canto, è questa la rumorosità che interessa ai fini del dimensionamento del sistema e della valutazione delle sue prestazioni.

Una potenza di rumore che escluda alcuni contributi può essere utile per capire la dinamica del disturbo, ma non ha alcuna utilità della caratterizzazione del sistema complessivo.

Il rapporto segnale-rumore può essere riferito e calcolato in una qualunque sezione.

Tipicamente ci si colloca in ingresso, ma ciò non rappresenta una scelta obbligata.

Il fatto che l'SNR resti invariato nel sistema equivalente non implica che la potenza di segnale utile o quella di rumore, considerati individualmente, restino pure invariati.

Al contrario, la potenza S , ad esempio, viene moltiplicata, nel passaggio dall'ingresso all'uscita, per il guadagno disponibile G_d , lo stesso si verifica per la potenza di rumore.

Di conseguenza, anche la temperatura equivalente della cascata (cui la potenza di rumore risulta proporzionale) può essere riportata in una qualunque sezione, semplicemente moltiplicando o dividendo per gli opportuni guadagni (dei blocchi che separano le sezioni considerate).

¹⁰Slide del prof — Parametri di rumore sistema lineare — pag 8
Appunti di Damiano — pag 8

1.9 Da segnale analogico a segnale digitale

¹¹

In un sistema analogico, le potenze di rumore introdotte dalle singole tratte, supposte tra loro indipendenti, si sommano.

La potenza di rumore, in analogico, è approssimativamente ad n volte il risultato che si ottiene considerando la tratta singola.

Ciò si traduce in una significativa degradazione del rapporto segnale-rumore il cui valore, è inversamente proporzionale al numero delle tratte considerate.

Il rapporto segnale rumore è il parametro di merito da utilizzare per esprimere la qualità di un segnale analogico.

La situazione è radicalmente diversa quando si considera un sistema numerico (sempre organizzato in n tratte).

Qui il parametro di merito diventa la probabilità di errore e si dimostra che (sarà dimostrato nei successivi capitoli) che, in questo caso, sono proprio le probabilità di errore a dover essere sommate, nella cascata di n tratte, per ottenere la probabilità di errore complessiva che influenza l'intero collegamento.

Questo risultato, peraltro, è conseguenza di una diversa organizzazione di sistema di trasmissione.

Mentre nel caso analogico i ripetitori posti alla fine di ogni tratta sono essenzialmente degli amplificatori, nel caso numerico può essere più conveniente effettuare ad ogni stazione intermedia una "rigenerazione" del segnale, cioè estrarre i dati ed effettuare una nuova trasmissione nella tratta successiva.

È proprio questa procedura che consente di evitare l'accumulazione dei disturbi.

Per fissare le idee, si può supporre che, in ciascuna tratta, la trasmissione avvenga con le modalità di un canale binario simmetrico caratterizzato da una probabilità di transizione errata sul simbolo binario uguale a p .

Si ha, allora, un errore finale sul simbolo trasmesso se avviene in un numero dispari di tratte: con un numero pari di errori: infatti, si ha una compensazione e, in uscita, si riottiene il simbolo corretto.

Ad esempio se si ha il simbolo 1, se le tratte sono pari, il simbolo ritorna a 1; se le tratte sono dispari, il simbolo o può essere 1 o 0, che è diverso da 1.

Ma la probabilità di errore su 3, 5, ... tratte è in genere (se p è piccola) trascurabile, rispetto alla probabilità di sbagliare su una sola delle n tratte.

Quest'ultima probabilità, che approssima, nel senso precisato, la probabilità di errore globale P_E^n vale:

$$\binom{n}{1} p(1-p)^{n-1} \approx n \cdot p \approx P_E^{(n)} \quad (1.56)$$

Questa ultima formula esplicita che la probabilità di errore si sommano nella cascata di n tratte.

Questa formula è nettamente favorevole rispetto a quella valida nel caso di sistema analogici.

In particolare, tenendo conto che la probabilità di errore decresce molto velocemente al crescere del rapporto segnale-rumore (tipicamente con legge esponenziale), si verifica che, aumentando anche considerevolmente il numero di tratte la variazione corrispondente nel rapporto segnale-rumore, si mantiene costante.

¹¹Slide del prof — Parametri di rumore sistema lineare — pag 13
Appunti di Damiano — pag 13

1.9.1 Calcolo numerico di n tratte per un segnale analogico rispetto al segnale digitale

¹²

Ad esempio, nel caso di un segnale analogico, considerando:

$$n = 10 \quad (1.57)$$

tratte, il rapporto segnale e rumore deve essere aumentato di 10 dB per mantenere la qualità che si avrebbe con una tratta singola.

Considerando invece un segnale numero in banda base con forma d'onda anti-podali, e volendo conseguire una probabilità di errore globale:

$$P_E^{(n)} = 10^{-5} \quad (1.58)$$

sappiamo che, considerando $n = 1$:

$$\binom{n}{1} p (1 - p)^{n-1} \approx n \cdot p \approx P_E^{(n)} \quad (1.59)$$

quindi:

$$\begin{aligned} n \cdot p &\approx P_E^{(n)} \\ &\downarrow \\ 1 \cdot p &\approx P_E^{(n)} \\ p &= P_E^{(n)} = 10^{-5} \end{aligned} \quad (1.60)$$

se invece $n = 10$:

$$\begin{aligned} n \cdot p &\approx P_E^{(n)} \\ &\downarrow \\ n \cdot p &\approx P_E^{(n)} = 10^{-5} \\ p &= \frac{P_E^{(n)}}{n} = \frac{10^{-5}}{10} = 10^{-6} \end{aligned} \quad (1.61)$$

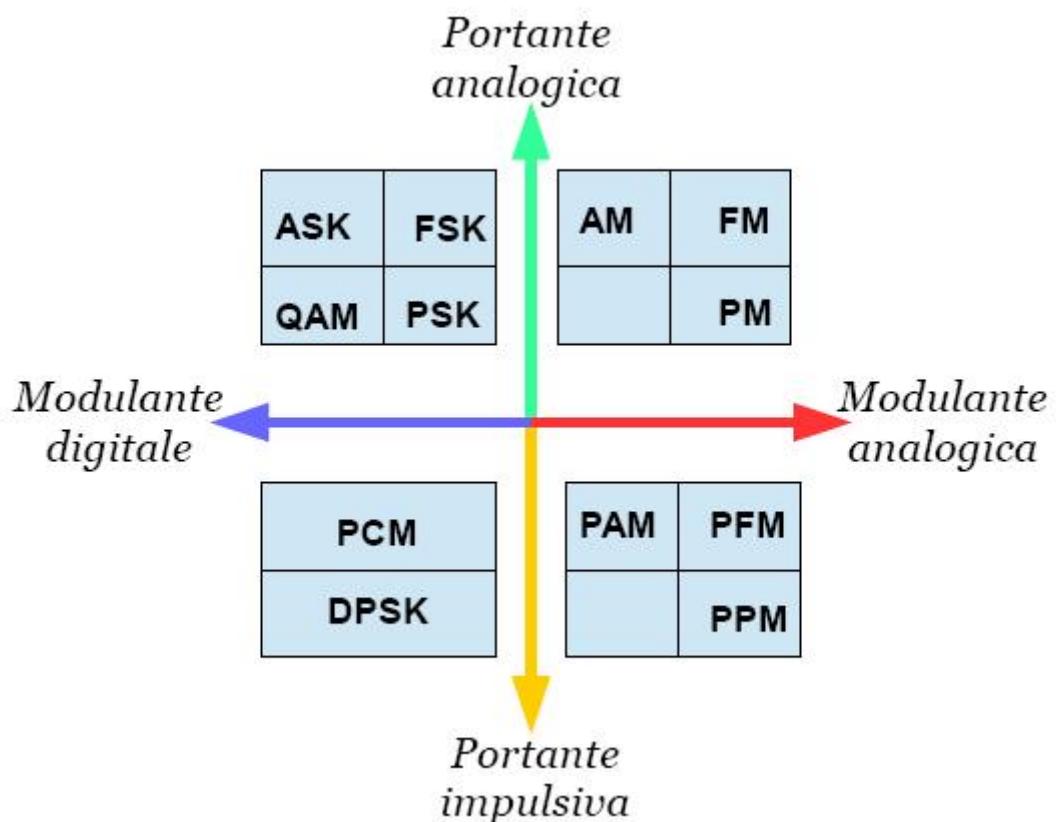
Come si dimostrerà, trattando la qualità delle trasmissioni numeriche, in particolare per il caso di segnali antipodali, il rapporto segnale-rumore quando si passa da $n = 1$ a $n = 10$, deve essere aumentato di circa 1 dB.

Un valore significativamente minore rispetto a quanto richiesto nel caso di collegamento analogico.

¹²Slide del prof — Parametri di rumore sistema lineare — pag 13
Appunti di Damiano — pag 13

Capitolo 2

Modulazioni analogiche



2.1 Rappresentazioni di segnali in modulazione

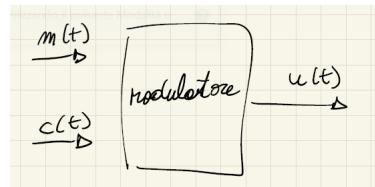
1

Un significativo esempio di applicazione della Teoria dei Segnali, si rinviene nella rappresentazione nel dominio del tempo e della frequenza dei segnali modulati.

Nell'ambito delle telecomunicazioni, il termine modulazione sta ad indicare l'operazione (o in generale, il complesso delle operazioni) che consente di trasferire l'informazione da trasmettere (chiamato segnale modulante) in uno o più parametri di un segnale di portante.

Il risultato di questo "trasferimento" è un nuovo segnale chiamato segnale modulato.

Possiamo schematizzare la modulazione in questa maniera:



dove:

- $m(t)$ è il segnale modulante, dall'inglese modulating signal
- $c(t)$ è il segnale portante, dall'inglese carrier
- $u(t)$ è il segnale modulato

Il segnale modulato è quello più idoneo alla trasmissione nel canale e può essere inviato nel canale con le modalità che si sono scelte a priori.

Benché la modulazione possa essere riferita anche a segnali passa-basso (ad esempio nel mixer di uno strumento musicale), "storicamente" la modulazione nasce per trasformare segnali passa-basso (cioè in banda base) in segnali passa-banda (quindi che dalla banda base li troveremo in alta frequenza, i.e. traslati nello spettro).

Questa traslazione in frequenza è dovuta principalmente a tre punti:

1. Necessità di utilizzare antenne di dimensioni ridotte nelle trasmissioni radio: la frequenza di trasmissione dipende dalla lunghezza d'onda del segnale, che deve essere uguale alle dimensioni dell'antenna
2. Necessità di utilizzare al meglio la funzione di trasferimento del canale trasmittivo: la modulazione permette di modulare il segnale modulante, che si trova in BB (Banda Base) in intervalli di frequenze in cui il canale distorce meno il segnale
3. Necessità di trasmettere contemporaneamente più segnali che originariamente occupano lo stesso intervallo di frequenze: si pensi ad esempio al segnale musicale in BB che deve essere trasmesso nella radio commerciale

¹Slide del prof — Modulazioni analogiche — pag 1 - 2
 Appunti di Damiano — pag 1 - 2
 Appunti — Modulazioni analogiche — pag 1 - 2
 Appunti — 2025-02-28 — pag 2

È vero che adesso esistono i podcast, ma ricordati di ascoltare un po' la radio.

Prima di diventare anche tu zanzaroso e accendere la radio per ascoltare La Zanzara con David Parenzo e Cruciani, ascoltati altri programmi interessanti nel palinsesto della "Radio di Confindustria".

Te ne consiglio due dal lunedì al venerdì:

- Due di Denari, con Mauro Meazza e Debora Rosciani, dalle 11 alle 12
<https://www.radio24.ilsole24ore.com/conduttori/debora-rosciani/programmi/due-denari>
- Focus economia, con Sebastiano Barisoni, dalle 17 alle 18:30, specialmente l'immancabile episodio di venerdì con la "poco invidiabile classifica degli sprechi"
<https://www.radio24.ilsole24ore.com/programmi/focus-economia>

Prima il dovere e poi il piacere.

AVANTI TIGRE

Oltre ai vantaggi appena elencati, ce ne sono molti altri che non stiamo qui ad elencare.

Un grosso vantaggio della modulazione è l'allargamento dello spettro, che permette di migliorare l'immunità della trasmissione ai diversi agenti di disturbi di norma presenti.

Generalmente il disturbo principale che avremo a che fare è il nostro "caro amico" rumore termico.

Al nostro caro amico rumore termico gli dedico questa canzone:

Toy Story - Un amico in me

https://youtu.be/00_W0Nxv5p0?si=Hp4Aiu40BY5xu6pT

2.2 Modulazione analogica di una portante sinusoidale

2

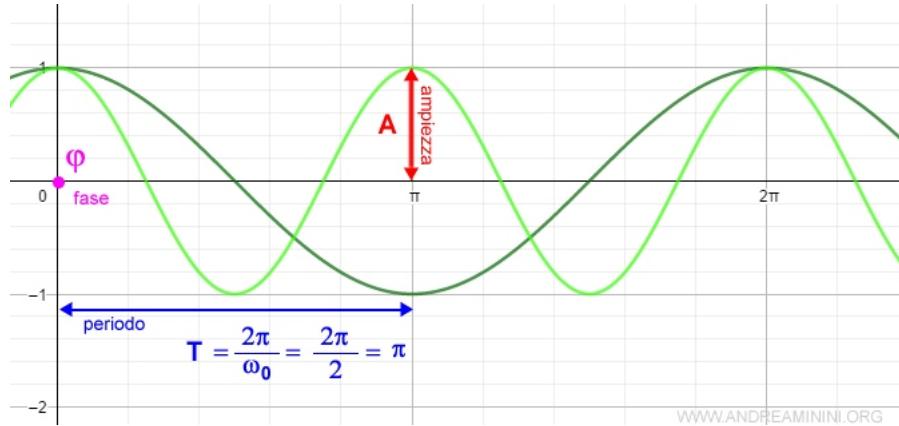
Un segnale $c(t)$ (carrier, cioè segnale di portante) sinusoidale può essere descritto dalla seguente equazione:

$$c(t) = A_c \cdot \cos(2\pi f_c t + \phi_c) \quad (2.1)$$

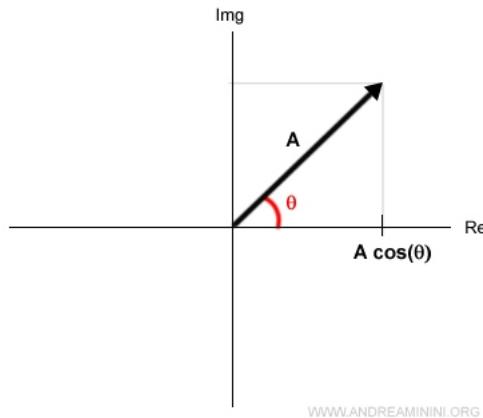
dove:

- A_c è l'ampiezza della sinusoide
- f_c è la frequenza, cioè l'inverso del periodo T di una sinusoide
- t è il tempo
- ϕ_c è la fase iniziale

Da un punto di vista grafico, nel tempo t possiamo visualizzare i parametri della sinusoide con la seguente figura:



oppure nel piano fasoriale (che ci sarà molto utile nelle future trattazioni):

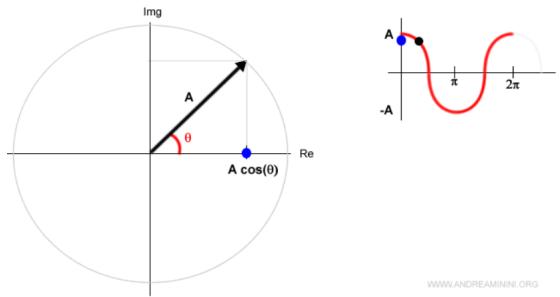


dove:

- $A \cdot \cos(\theta)$ è la proiezione del vettore \vec{A} sull'asse reale
- $A \cdot \sin(\theta)$ è la proiezione del vettore \vec{A} sull'asse immaginario

²Slide del prof — Modulazioni analogiche — pag 3
 Appunti di Damiano — pag 3
 Appunti — Modulazioni analogiche — pag 3
 Appunti — 2025-02-28 — pag 3

Con la seguente figura:



possiamo capire la relazione tra l'andamento nella sinusoide nel tempo e il piano fasoriale.

Sembra banale ripassare i concetti basilari di una sinusoide, ma meglio essere tutti sulla stessa
punto di inizio che far finta di capire qualcosa.

Per approfondire:

<https://www.andreaminini.org/telecomunicazioni/il-segnale-sinusoidale>

Ritornando alla formula di $c(t)$:

$$c(t) = A_c \cdot \cos(2\pi f_c t + \phi_c) \quad (2.2)$$

è evidente che è caratterizzata da tre gradi di libertà (cioè parametri che possiamo cambiare):

- l'ampiezza A_c
- la frequenza f_c
- la fase iniziale ϕ_c

Rispetto al corso di Segnali Determinati e Aleatori, o Teoria dei segnali per quelli del vecchio ordinamento, in questo corso si predilige l'uso della frequenza f piuttosto che dei radianti ω .

La relazione tra radianti e frequenza è la seguente:

$$\omega = 2\pi f \quad (2.3)$$

Questi gradi di libertà sono costanti e invariabili per qualunque istante di tempo nel segnale $c(t)$.

Con la modulazione si vuole far dipendere uno o più di questi parametri di $c(t)$, detto segnale di portante, a dei parametri di un altro segnale $m(t)$, detto appunto segnale modulante.

Si parlerà di:

- modulazione di ampiezza (Amplitude Modulation o AM) quando l'ampiezza A_c verrà resa dipendente da $m(t)$
- modulazione in frequenza (Frequency Modulation o FM) quando la frequenza f_c verrà resa dipendente da $m(t)$
- modulazione di fase (Phase Modulation o PM) quando la fase iniziale ϕ_c verrà resa dipendente da $m(t)$

In realtà, si avrà modo di precisare che sia la modulazione di frequenza e sia la modulazione di fase sono tra loro intimamente legate: per cui, ove si realizzi una modulazione di frequenza, si realizza anche

una modulazione di fase e viceversa.

Per questo motivo, sia la modulazione di fase che di frequenza, possono essere definite anche come modulazioni angolari, perché il segnale modulante agisce sull'angolo della funzione sinusoidale.

2.3 Modulazione di ampiezza

³

Ritornando al segnale $c(t)$:

$$c(t) = A_c \cdot \cos(2\pi f_c t + \phi_c) \quad (2.4)$$

andiamo a spiegare diverse modulazioni in cui l'ampiezza A_c verrà resa dipendente da $m(t)$, segnale modulante.

³Slide del prof — Modulazioni analogiche — pag 3
Appunti di Damiano — pag 3
Appunti — Modulazioni analogiche — pag 3
Appunti — 2025-02-28 — pag 3

2.3.1 Banda Laterale Doppia Portante Soppressa (DSB-SC)

4

Nella modulazione di ampiezza con banda laterale doppia portante soppressa (o più frequentemente DSB-SC: Double Side Band - Suppressed Carrier), il segnale modulato $u(t)$ è ottenuto semplicemente moltiplicando il segnale $c(t)$ per il segnale modulante $m(t)$.

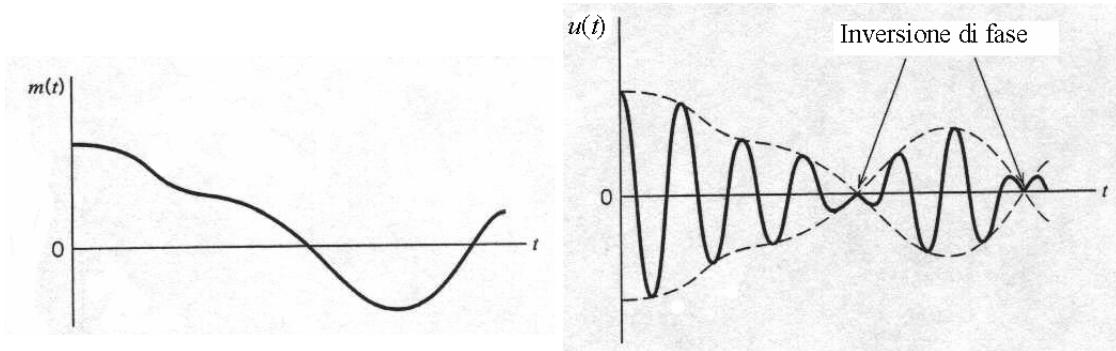
In formula:

$$\begin{aligned} u(t) &= m(t) \cdot c(t) \\ &= m(t) \cdot [A_c \cdot \cos(2\pi f_c t + \phi_c)] \\ &= [A_c \cdot m(t)] \cdot \cos(2\pi f_c t + \phi_c) \end{aligned} \quad (2.5)$$

Un esempio di segnale modulato $u(t)$:

$$u(t) = [A_c \cdot m(t)] \cdot \cos(2\pi f_c t + \phi_c) \quad (2.6)$$

partendo da un segnale $m(t)$ modulante è raffigurato nelle seguenti figure:



È interessante l'inversione di fase del segnale modulato in corrispondenza dei passaggi per lo zero del segnale modulante.

Inoltre, nella figura di $u(t)$, è tratteggiato l'andamento di $m(t)$.

Utilizzando le proprietà della trasformata di Fourier, cioè quella della trasformazione in frequenza e quella del prodotto, $u(t)$ ha una semplice rappresentazione in frequenza:

$$\begin{aligned} u(t) &= [A_c \cdot m(t)] \cdot \cos(2\pi f_c t + \phi_c) \\ &\downarrow \\ U(f) &= \frac{A_c}{2} [M(f - f_c)e^{j\theta_c} + M(f + f_c)e^{-j\theta_c}] \end{aligned} \quad (2.7)$$

⁴Slide del prof — Modulazioni analogiche — pag 3 - 6
 Appunti di Damiano — pag 3 - 6
 Appunti — Modulazioni analogiche — pag 3 - 6
 Appunti — 2025-02-28 — pag 3 - 5
 Appunti — 2025-03-03 — pag 2

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 3.3 Proprietà della trasformata di Fourier — pag 22 - 23

Proprietà della traslazione in frequenza

Considerando ω_1 la frequenza traslata, possiamo dire che:

$$s'(t) = s(t)e^{j\omega_1 t} \leftrightarrow S'(\omega) = S(\omega - \omega_1) \quad (2.8)$$

Proprietà del prodotto

$$s'(t) = s_1(t)s_2(t) \leftrightarrow S'(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_1(\theta)S_2(\omega - \theta)d\theta \quad (2.9)$$

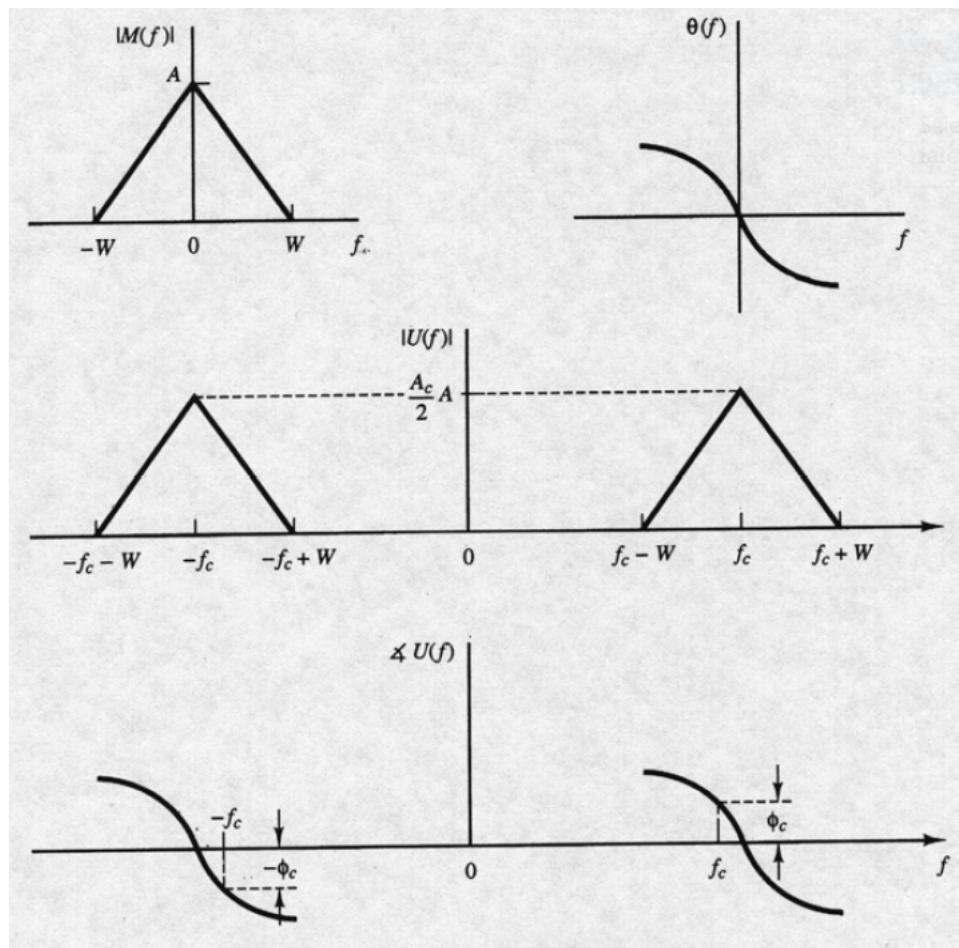
L'integrazione viene svolta su θ e ω viene visto come parametro nella formula di integrazione.

■

Da <https://www.andreaminini.org/telecomunicazioni/il-fasore>

$e^{-j\theta_c}$ è un $\cos(\omega t + \theta_c)$ nell'asse dei tempi

In frequenza i segnali $m(t)$ e $u(t)$ possiamo rappresentarli come in figura:



Questa figura mette in evidenza tutte le caratteristiche peculiari della modulazione DSB-SC.

La modulazione trasla lo spettro del segnale di $|M(f)|$ da $[-W, +W]$ a $|U(f)|$ in $[-f_c - W, -f_c + W]$ e $[f_c - W, f_c + W]$.

Il modulo di $|M(f)|$ in $|U(f)|$ resta inalterato (a meno di un fattore moltiplicativo che è $\frac{A_c}{2}$).

L'unico parametro che cambia è quello della fase che passa da $\theta(f) = 0$ in f a $-\theta_c$ a $-f_c$ e θ_c a f_c .

In altre parole, la fase iniziale viene incrementata di θ_c nelle frequenze positive e viene ridotta di θ_c per quanto riguarda le frequenze negative.

A meno di questo sfasamento (che peraltro è nullo quando $\theta_c = 0$), lo spettro del segnale non viene modificato dalla modulazione.

Per questo motivo, la modulazione DSB-SC, e più in generale la modulazione in ampiezza, viene classificata come "lineare" anche se non è una modulazione svolta con componenti lineari.

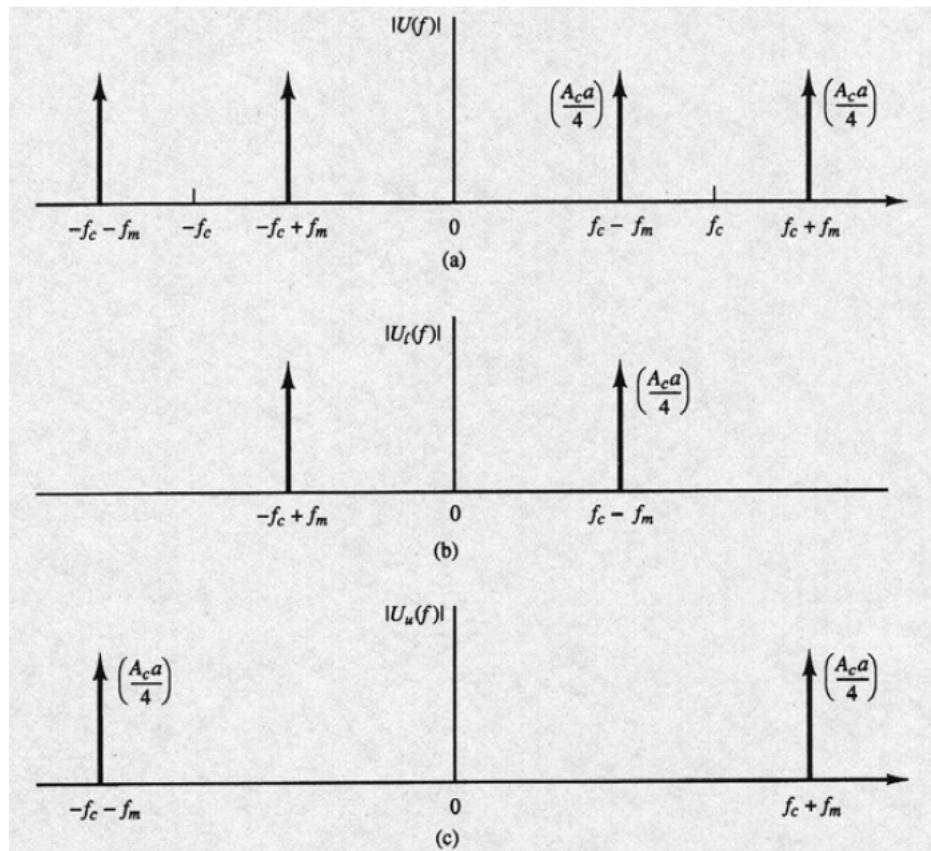
Un sistema lineare, può eliminare (cioè si comporta da filtro) o equalizzare (cioè svolge un equalizzatore di ampiezza o di fase) totalmente o in parte le frequenze che gli si presentano in ingresso, ma non ne potrà mai introdurre di nuove.

Dalla figura degli spettri del segnale modulato rispetto al segnale modulante, $M(f)$ occupa una banda $[-W, +W]$, invece $U(F)$ occupa $[-f_c - W, -f_c + W]$ e $[f_c - W, f_c + W]$, che è il doppio della banda occupata da $M(f)$.

Prende il nome di banda laterale superiore di $U(f)$ (o come è stato definito a lezione Banda Laterale Superiore BLSup) la parte dello spettro per $|f| > f_c$.

Prende il nome di banda laterale inferiore di $U(f)$ (o come è stato definito a lezione Banda Laterale Inferiore BLInf) la parte dello spettro per $|f| < f_c$.

Il seguente schema ci può far capire meglio la differenza tra banda laterale superiore e banda laterale inferiore:



Se la modulante è sinusoidale, grazie alle proprietà di Fourier, avremo come segnale modulato una serie di righe nello spettro in frequenza (come si vede in figura da $|U(f)|$).

$|U_u(f)|$ è lo spettro della banda superiore (u dall'inglese si intende upper, superiore in italiano), $|U_l(f)|$ è lo spettro della banda inferiore (l dall'inglese si intende lower, inferiore in italiano)

Quindi, come si vede dalla figura, in $|U_u(f)|$ ci sono solo le frequenze a $[-f_c + f_m]$ e $[f_c - f_m]$, in $|U_l(f)|$ ci sono solo le frequenze a $[-f_c - f_m]$ e $[f_c + f_m]$.

Grazie al concetto dei processi stocastici, in particolare della teoria dei processi ciclo-stazionari, i segnali determinati e le loro proprietà si estendono anche ai segnali di spettro di potenza, cioè quelli che abbiamo visto in figura.

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>
Capitolo 13 Processi stocastici — pag 151 - 160

2.3.1.1 Demodulazione per la DSB-SC

5

Un ultimo aspetto riguarda la demodulazione, cioè la capacità di riottenere il segnale modulante a partire dal segnale modulato.

In un sistema di comunicazioni, la demodulazione avviene al lato ricevente quando si deve recuperare l'informazione $m(t)$ per renderla utilizzabile dal destinatario.

Si pensi ad un segnale telefonico: la voce umana sta nella banda da 300 Hz a 3400 Hz.

Nella trasmissione sul canale, ad esempio il cavo di rame, il segnale voce viene modulato ad alta frequenza in modo da essere trasportato nella rete di telecomunicazioni, e alla fine viene ricevuto dal destinatario che riporta il segnale modulato in alta frequenza alla frequenza originale, così da poter ascoltare la voce del mittente.

In assenza di rumore, distorsione o altre cause di disturbo, il segnale ricevuto $r(t)$ coincide con $m(t)$: il modo concettuale per riottenere il segnale $m(t)$ non è quello di dividere per la portante, bensì moltiplicare $r(t)$ per un segnale sinusoidale.

Il segnale sinusoidale sarà proprio il segnale di portante, che è stata generata localmente dal ricevitore, con la stessa frequenza e con la stessa fase della portante ricevuta.

In formule, se si è ricevuto $r(t)$ e grazie alle proprietà trigonometriche:

$$\begin{aligned}
 r(t) \cdot \cos(2\pi f_c t + \phi_c) &= A_c \cos^2(2\pi f_c t + \phi_c) \\
 &= A_c m(t) \cdot \frac{1 + \cos[2 \cdot (2\pi f_c t + \phi_c)]}{2} \\
 &= \frac{1}{2} A_c m(t) + \frac{1}{2} A_c m(t) \cos[2 \cdot (2\pi f_c t + \phi_c)] \\
 &= \frac{1}{2} A_c m(t) + \frac{1}{2} A_c m(t) \cos(4\pi f_c t + 2\phi_c)
 \end{aligned} \tag{2.10}$$

Ripasso delle formule trigonometriche che non fa mai male:

<https://www.youmath.it/formulari/65-formulari-di-trigonometria-logaritmi-esponenziali/159-identita-trigonometriche-formule-di-prostaferesi-formule-di-werner.html>

In questo corso le utilizzeremo spesso, quindi può essere comodo averle sempre sotto mano

Osservando la formula appena scritta, possiamo fare le seguenti considerazioni:

- Aggiungendo un filtro passa basso con frequenza di taglio a $2\pi f_c$, possiamo togliere tutta la componente in alta frequenza
- $\frac{1}{2}A_c$ è solo un fattore moltiplicativo

Con queste osservazioni, possiamo semplificare l'equazione del demodulatore come:

$$\begin{aligned}
 r(t) \cdot \cos(2\pi f_c t + \phi_c) &= \frac{1}{2} A_c m(t) + \frac{1}{2} A_c m(t) \cos(4\pi f_c t + 2\phi_c) \\
 &\downarrow \\
 r(t) \cdot \cos(2\pi f_c t + \phi_c) &= \frac{1}{2} A_c m(t)
 \end{aligned} \tag{2.11}$$

⁵Slide del prof — Modulazioni analogiche — pag 6 - 8

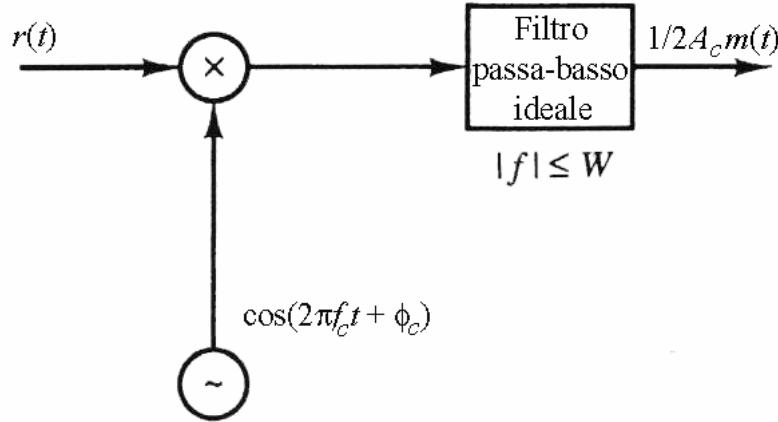
Appunti di Damiano — pag 6 - 8

Appunti — Modulazioni analogiche — pag 6 - 8

Appunti — 2025-03-03 — pag 3 - 11

Cioè il segnale ricevuto e de-modulato è proprio il segnale originale in banda base, a meno del fattore moltiplicativo $\frac{1}{2}A_c$.

Possiamo schematizzare il demodulatore con la seguente figura:



Se invece consideriamo una differenza di fase tra la portante ricevuta e la portante generata localmente, avremo la seguente equazione:

$$r(t) \cos(2\pi f_c t + \phi) = \frac{1}{2}A_c m(t) \cos(\phi_c - \phi) + \frac{1}{2}A_c m(t) \cos(4\pi f_c t + \phi_c - \phi) \quad (2.12)$$

A valle del filtro passa-basso del demodulatore, e per le considerazioni svolte precedentemente, avremo il segnale di uscita $y_l(t)$:

$$\begin{aligned} r(t) \cos(2\pi f_c t + \phi) &= \frac{1}{2}A_c m(t) \cos(\phi_c - \phi) + \frac{1}{2}A_c m(t) \cos(4\pi f_c t + \phi_c - \phi) \\ &\downarrow \\ y_l(t) &= \frac{1}{2}A_c m(t) \cos(\phi_c - \phi) \end{aligned} \quad (2.13)$$

Questa ultima espressione di $y_l(t)$ mette in evidenza che, rispetto al caso ideale:

$$y_l'(t) = \frac{1}{2}A_c m(t) \quad (2.14)$$

il segnale utile in uscita è ridotto del fattore $\cos(\theta_c - \theta)$.

Se:

$$\theta_c - \theta = 90^\circ \quad (2.15)$$

$y_l(t)$ diventa:

$$\begin{aligned}
 y_l(t) &= \frac{1}{2}A_c m(t) \cos(\phi_c - c) \\
 &\downarrow \\
 y_l(t) &= \frac{1}{2}A_c m(t) \cos(90^\circ) \\
 &= \frac{1}{2}A_c m(t) \cdot 0 \\
 &= 0
 \end{aligned} \tag{2.16}$$

cioè $y_l(t)$ si annulla.

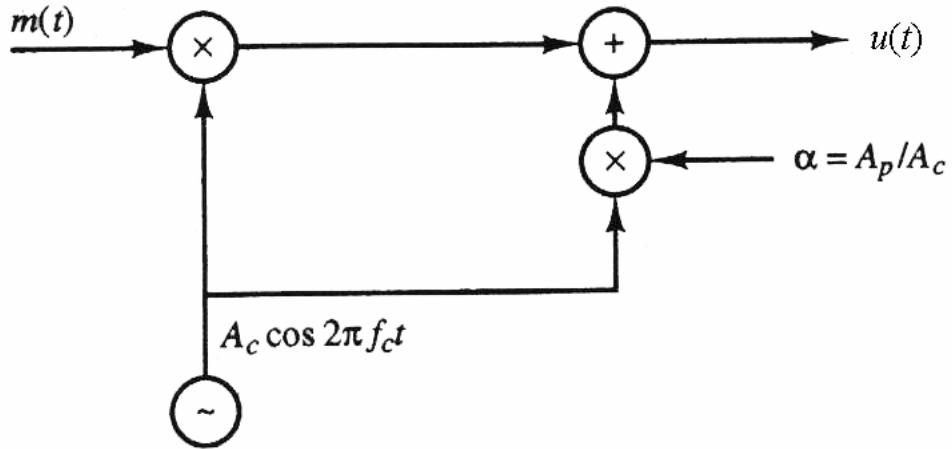
In caso di un segnale utile sovrapposto ad un contributo di rumore (ad esempio il rumore termico), la rivelazione imperfetta riduce la potenza del segnale utile, rendendo quest'ultimo più vulnerabile agli effetti del rumore.

Il demodulatore raffigurato precedentemente è detto anche ricevitore sincrono (o coerente).

Viene definito ricevitore coerente perché, per ricavare il segnale modulante, si moltiplica il segnale modulato per un coseno.

Il problema della ricostruzione della fase θ_c della portante è uno dei problemi di maggior importanza nell'ambito della ricezione sincrona e viene tipicamente risolto con i PLL (Phase-Locked Loop).

In alternativa, la demodulazione può essere semplificata sovrapponendo alla trasmissione del segnale DSB-SC la trasmissione di un residuo di portante, utilizzando il seguente schema:

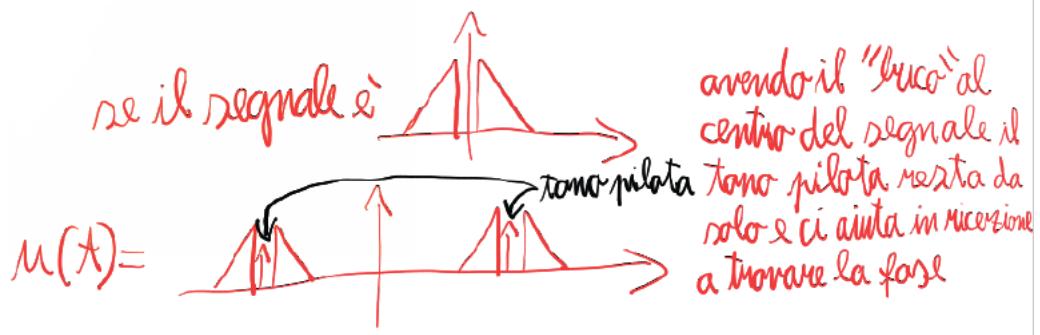


Rispetto al caso precedente, $u(t)$ sarà:

$$u(t) = \left[A_c m(t) + \frac{A_p}{A_c} \right] \cdot \cos(2\pi f_c t) \tag{2.17}$$

dove A_p viene definito come residuo di portante, la cui ampiezza è una frazione di α di A_c : A_c viene definito come tono pilota.

In frequenza avremo:



Con questo schema non vi è necessità di generare localmente la portante al ricevitore, ma il prezzo che si paga è in termini di potenza trasmessa addizionale e di intervallo di frequenza non utilizzabile perché necessario alla trasmissione della portante.

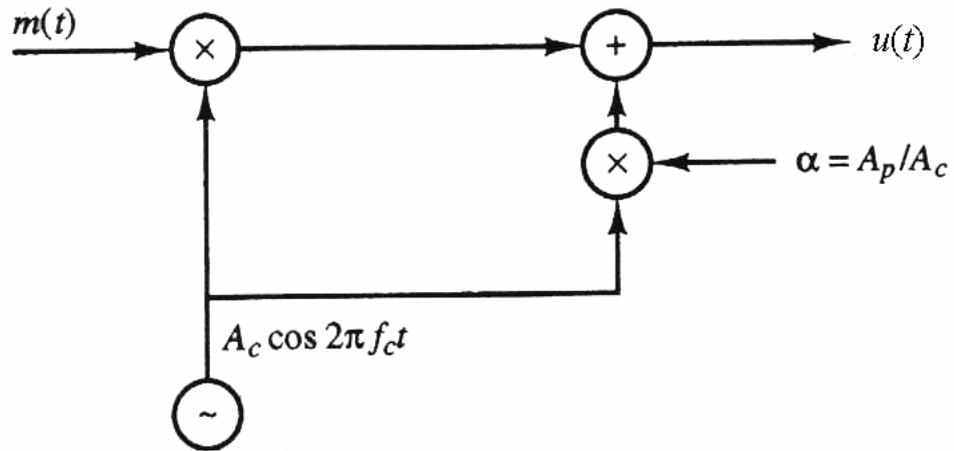
In questo senso, rispetto al caso senza tono pilota, lo schema DSB-SC costituisce una soluzione "ibrida", che permette di avere un demodulatore più semplice ed a costo economico ridotto, ma sprecando nettamente più energia in trasmissione.

2.3.2 Modulazione di ampiezza convenzionale - AM

6

Nella sezione precedente, cioè riguardo alla DSB-SC, abbiamo discusso e analizzato una DSB-SC con una portante quando si va in trasmissione per semplificare il ricevitore: questo caso particolare è proprio la tecnica AM (Amplitude Modulation) convenzionale della modulazione DSB-SC.

Riportando lo schema del modulatore AM:



il segnale modulato in uscita $u(t)$ sarà:

$$u(t) = A_c[1 + \alpha \cdot m_n(t)] \cos(2\pi f_c t + \phi_c) \quad (2.18)$$

dove:

$$\begin{cases} m_n(t) = \frac{m(t)}{\max |m(t)|} \\ 0 < \alpha \leq 1 \end{cases} \quad (2.19)$$

$m_n(t)$ è definito come "segnale modulante normalizzato", invece α è definito come "indice di modulazione di ampiezza".

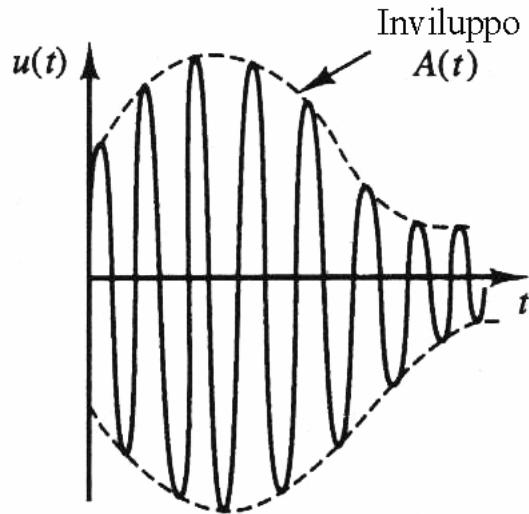
Grazie alle definizioni di $m_n(t)$ e α , possiamo dire che $u(t)$ non diventa mai negativa nel tempo.

Considerando il caso particolare in cui:

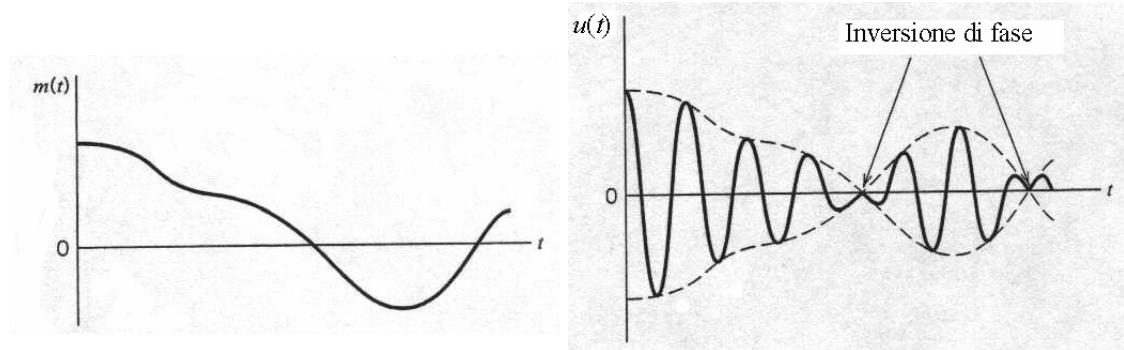
$$\begin{cases} m_n(t_0) = -1 \\ \alpha = 1 \end{cases} \quad (2.20)$$

⁶Slide del prof — Modulazioni analogiche — pag 8 - 11
Appunti di Damiano — pag 8 - 11
Appunti — Modulazioni analogiche — pag 8 - 11
Appunti — 2025-03-03 — pag 11 - 13

l'andamento del segnale modulato $u(t)$ è il seguente:



e confrontando il segnale modulato in AM con quello a DSB-SC:



notiamo che in AM, l'inviluppo mantiene l'andamento del segnale modulante (cioè quello tratteggiato in figura): è proprio per questa caratteristica che la demodulazione del segnale AM convenzionale può essere particolarmente semplice.

Inoltre, nei passaggi sullo zero, in AM non ci sono inversioni della fase.

Per inviluppo si intende che la sequenza dei massimi tra segnale modulato e modulante è la stessa.

La trasformata di Fourier di un segnale in AM convenzionale diventa:

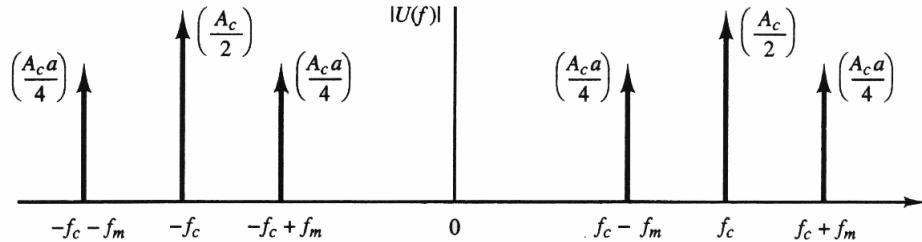
$$u(t) = A_c[1 + \alpha \cdot m_n(t)] \cos(2\pi f_c t + \phi_c)$$

\downarrow

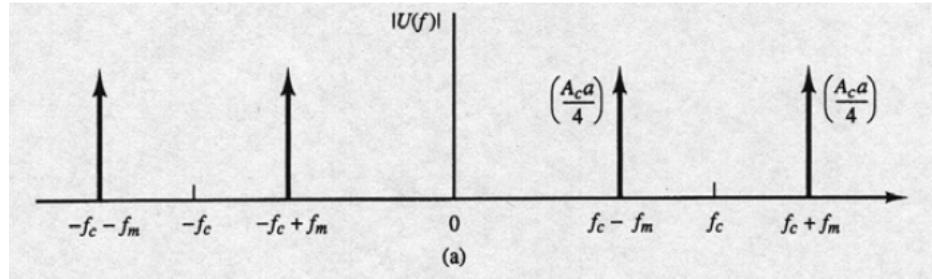
$$U(f) = \frac{A_c}{2} [e^{j\phi_c} a M_n(f - f_c) + e^{j\phi_c} \delta(f - f_c) + e^{-j\phi_c} a M_n(f + f_c) + e^{-j\phi_c} \delta(f + f_c)] \quad (2.21)$$

dove $M_n(f)$ è la trasformata di Fourier del segnale modulante normalizzato.

Un esempio di andamento di $U(f)$, nel caso di modulante sinusoidale, è mostrato nella seguente figura:



Rispetto allo spettro in frequenza di un segnale modulato in DSB-SC:



notiamo che, in AM convenzionale, è presente la portante a $|f_c|$.

Inoltre, si nota dalla figura, che, se ci troviamo nel caso di $\alpha = 1$, l'ampiezza della portante resta doppia di quella di ciascuno dei segnali assegnati alle bande laterali (in questo caso particolare, a loro volta sinusoidali).

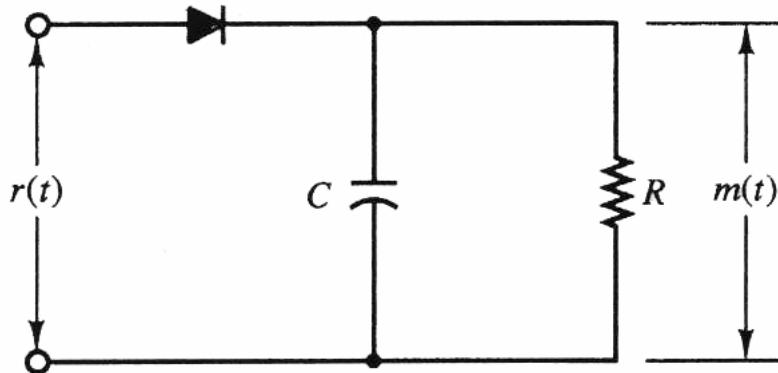
Di conseguenza, anche nel caso limite di modulazione al 100 %, cioè $\alpha = 1$, i $\frac{2}{3}$ della potenza complessiva del segnale modulato sono associati alla portante e solo la frazione rimanente, cioè $\frac{1}{3}$ è associata alle bande laterali.

Con valori di $\alpha < 1$, questo "squilibrio" si enfatizza, specialmente nei casi in cui $\alpha < \frac{1}{2}$, che sono la grande maggioranza dei segnali di interesse.

Il fatto di usare la maggior parte dell'energia di un segnale modulato per la portante è molto inefficiente dal punto di vista energetico, ma è molto utile per realizzare un circuito di demodulazione molto semplice.

Essendo l'informazione contenuta nell'inviluppo del segnale, è sufficiente ricostruire al ricevitore proprio l'inviluppo.

Quindi non è necessario un ricevitore coerente come nella DSB-SC (anche se tecnicamente lo si potrebbe utilizzare), ma possiamo utilizzare questo semplice circuito del demodulatore a inviluppo:

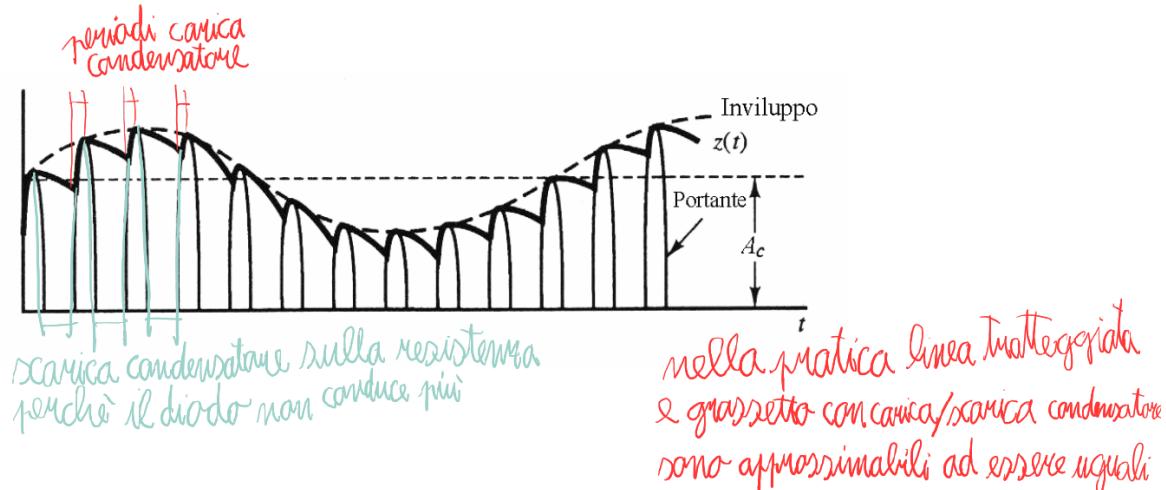


Questo demodulatore si definisce non coerente perché non si moltiplica $r(t)$ per un coseno.

I componenti di questo demodulatore sono i seguenti:

- diodo, che è un componente non lineare che elimina le componenti $r(t)$ quando quest'ultima diventa negativa
- una capacità C
- un resistore R

Attraverso una scelta adeguata dei componenti R-C (che realizzano un filtraggio passa-basso), si ottiene un andamento del segnale come la seguente figura:



La scelta di semplificare il ricevitore è decisiva in tutte quelle applicazioni in cui il numero dei trasmettitori è ridotto, mentre è elevato il numero dei ricevitori.

Per questo motivo l'utilizzo della AM convenzionale, o altre tecniche simili, ha costituito la soluzione privilegiata (negli anni precedenti prima dell'avvento del digitale e dei componenti a basso costo) per i servizi commerciali di tipo diffusivo come radio e televisione.

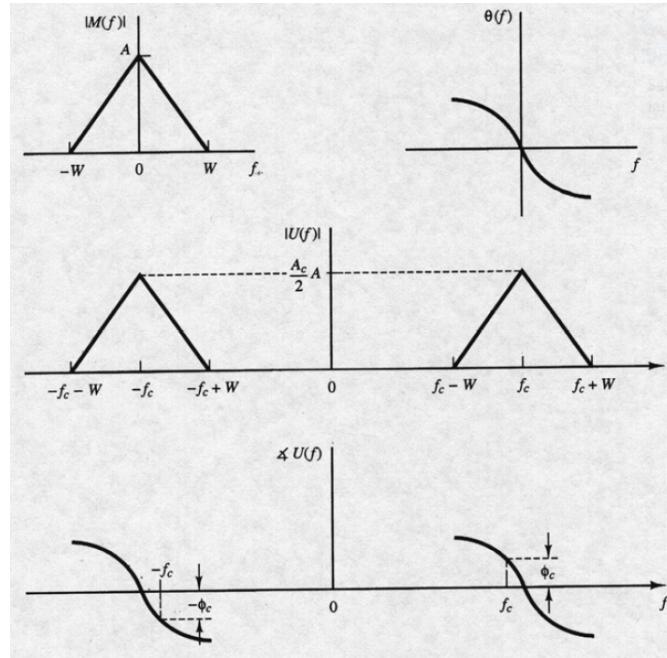
2.3.3 Banda Laterale Unica (SSB)

7

L'obiettivo della modulazione di ampiezza a banda laterale unica (acronimo italiano BLU o, in inglese SSB Single Side Band) è quello di migliorare l'efficienza spettrale.

Nelle tecniche precedenti trattate, cioè DSB-SC e AM convenzionale, la banda occupata dal segnale modulato era doppia rispetto a quella del segnale modulante.

Osservando alla forma dello spettro del segnale modulato, ad esempio quello in DSB-SC, e al modo in cui è stato ottenuto per traslazione dello spettro della banda base:



ci si convince facilmente che la ricostruzione del segnale è possibile anche conoscendo soltanto una delle due bande laterali.

L'andamento temporale del segnale SSB è del tipo:

$$u(t) = A_c m(t) \cos(2\pi f_c t) \mp A_c \tilde{m}(t) \sin(2\pi f_c t) \quad (2.22)$$

dove $\tilde{m}(t)$ è la trasformata di Hilbert del segnale modulante.

Come al solito, un altro argomento che abbiamo incontrato e studiato al precedente corso.

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 8 Trasformata di Hilbert — pag 83 - 85

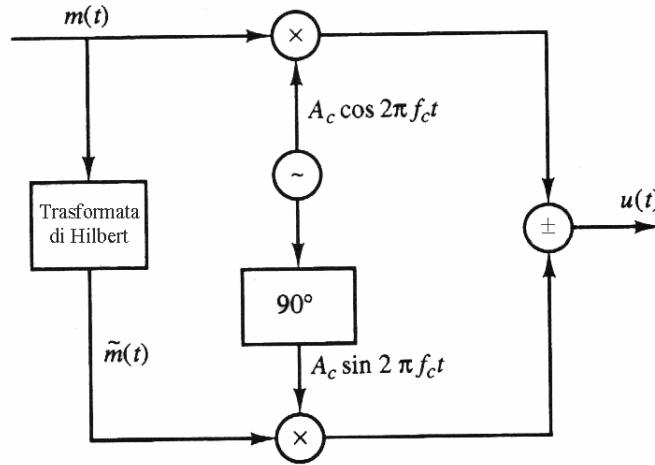
Per semplicità, si è trascurata la fase iniziale della portante ϕ_c .

Nella formula è presente un segno \mp perché si utilizza uno dei due casi:

- il segno negativo - se si è deciso di trasmettere la banda laterale superiore
- il segno positivo + se si è deciso di trasmettere la banda laterale inferiore

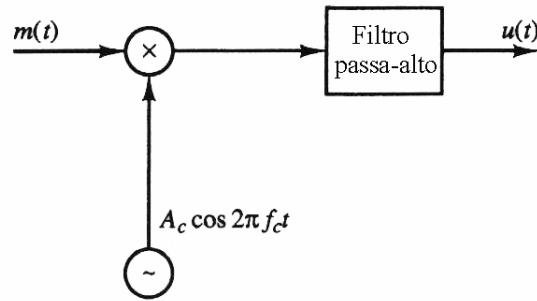
⁷Slide del prof — Modulazioni analogiche — pag 11 - 13
Appunti di Damiano — pag 11 - 13
Appunti — Modulazioni analogiche — pag 11 - 13
Appunti — 2025-03-03 — pag 13

Dalla formula di $u(t)$ possiamo ricavare uno schema a blocchi del modulatore, che è il seguente:



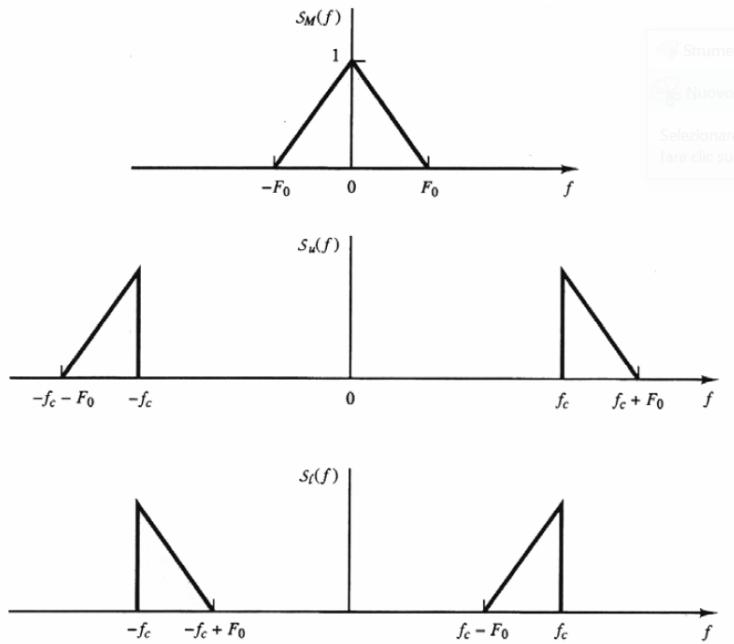
Dalle proprietà della trasformata di Hilbert, ricordiamo che $m(t)$ non deve avere un contenuto significativo alle basse frequenze, perché altrimenti si perderà informazione.

In alternativa, la stessa modulazione SSB può essere applicata con il seguente schema:



dove è presente un filtro passa-alto, con frequenza di taglio in prossimità della frequenza di portante f_c , del segnale $m(t)$ modulato in DSB-SC.

Un esempio di spettro di segnale SSB, nel caso di segnale modulante stocastico è riportato nella seguente figura:



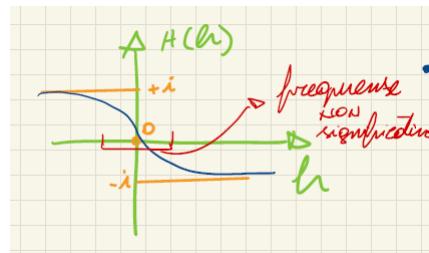
dove:

- $S_M(f)$ è il segnale modulante in Banda Base
- $S_u(f)$ è il segnale modulato in SSB con banda laterale superiore
- $S_l(f)$ è il segnale modulato in SSB con banda laterale inferiore

Sapendo che il segnale modulante $S_M(f)$ si estende in banda da $[-F_0, +F_0]$, trasmettendo solo in banda superiore o solo in banda inferiore, si occupa la metà della banda rispetto alla DSB-SC e la AM-convenzionale, mantenendo la stessa qualità.

O se si applica il filtraggio passa-alto dalla modulazione della DSB-SC o si utilizza il metodo con la trasformata di Hilbert, la SSB è assente di componenti significativi intorno alla frequenza d'origine.

Se infatti idealmente il filtraggio di Hilbert o quello passa-banda può avvenire con transizione brusca (o a frequenza nulla per la trasformata di Hilbert o a f_c per il filtro passa-alto), come si può visualizzare nel seguente andamento del filtro di Hilbert:



nella pratica un filtro reale presenterà una regione di transizione che, inevitabilmente, distorcerà, se presenti, le componenti armoniche al suo interno.

Questo giustifica il fatto che la modulazione SSB può essere utilizzata per il segnale telefonico, perché lo spettro significativo del parlato inizia dai 300 Hz in su, mentre non si può applicare al segnale televisivo, in cui il contenuto armonico include anche le bassissime frequenze (in rapporto alla larghezza di banda del segnale televisivo).

2.3.3.1 Demodulazione SSB

8

La demodulazione del segnale modulato in SSB:

$$u(t) = A_c m(t) \cos(2\pi f_c t) \mp A_c \tilde{m}(t) \sin(2\pi f_c t) \quad (2.23)$$

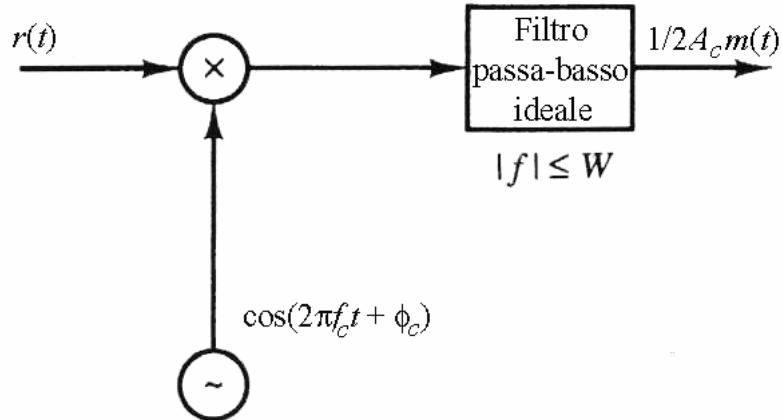
deve, per forza, necessariamente avvenire con tecnica coerente.

Come dice il Darione:



Si pongono le stesse criticità già evidenziate per la tecnica DSB-SC, ma aggravate dal fatto che l'eventuale non coincidenza tra la fase ϕ_c della portante trasmessa e quella ϕ , della portante generata localmente, si traduce qui non una semplice attenuazione (come per la DSB-SC), ma in una distorsione del segnale demodulato.

Per esempio, posto per semplicità $\phi_c = 0$ e supponendo di aver trasmesso la banda laterale superiore, ove sia $\phi \neq 0$, utilizzando lo stesso schema circuitale della DSB-SC, cioè un filtro coerente:



dove avevamo nel caso DSB-SC:

$$y_l(t) = \frac{1}{2} A_c m(t) \cos(\phi_c - c) \quad (2.24)$$

nel caso SSB in banda laterale superiore con $\phi_c = 0$, avremo che:

⁸Slide del prof — Modulazioni analogiche — pag 13
 Appunti di Damiano — pag 13
 Appunti — Modulazioni analogiche — pag 13
 Appunti — 2025-03-03 — pag 13 - 14

$$\begin{aligned}
 u(t) &= A_c m(t) \cos(2\pi f_c t) - A_c \tilde{m}(t) \sin(2\pi f_c t) \\
 &\downarrow \\
 y_l(t) &= \frac{1}{2} A_c m(t) \cos(\phi) + \frac{1}{2} A_c \tilde{m}(t) \sin(\phi)
 \end{aligned} \tag{2.25}$$

Dalla formula di $y_l(t)$ notiamo che il segnale utile $\frac{1}{2} A_c m(t) \cos(\phi)$ è sovrapposto alla sua trasformata di Hilbert, che è ovviamente diversa dal segnale originale.

Ecco perché è importante realizzare un controllo molto accurato della fase ϕ al ricevitore (ad esempio impiegando i PLL), in modo da evitare che il problema evidenziato si presenti.

Per problema evidenziato della fase ϕ si intende che, nel caso peggiore in cui:

$$\phi = 0^\circ \tag{2.26}$$

allora:

$$\begin{aligned}
 y_l(t) &= \frac{1}{2} A_c m(t) \cos(\phi) + \frac{1}{2} A_c \tilde{m}(t) \sin(\phi) \\
 &\downarrow \\
 y_l(t) &= \frac{1}{2} A_c m(t) \cos(0^\circ) + \frac{1}{2} A_c \tilde{m}(t) \sin(0^\circ) \\
 &= 0 + \frac{1}{2} A_c \tilde{m}(t) \\
 &= \frac{1}{2} A_c \tilde{m}(t)
 \end{aligned} \tag{2.27}$$

cioè è presente solo la trasformata di Hilbert di $m(t)$.

In modo da evitare questo problema, come ribadito nella DSB-SC, si può trasmettere un tono pilota a f_c , sprecando sulla potenza e sulla efficienza notevole della modulazione SSB.

2.3.3.2 Modulazione di ampiezza con portanti in quadratura

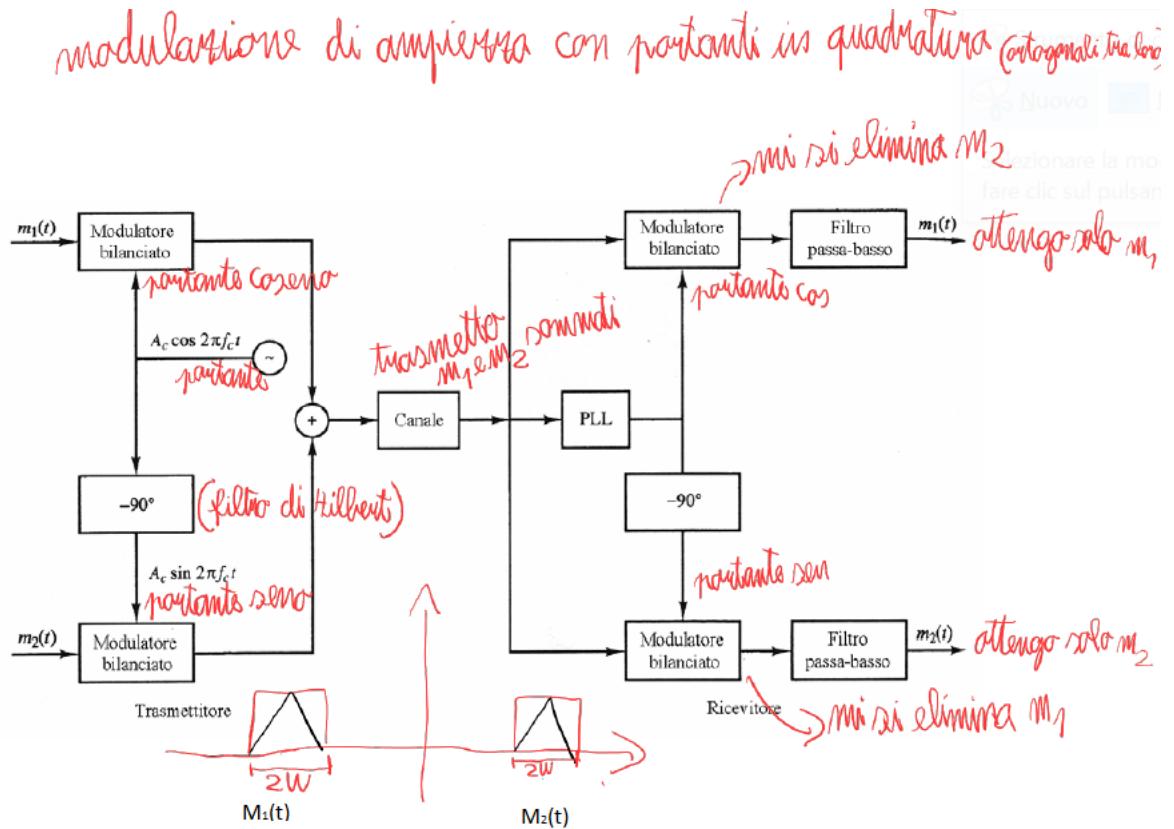
9

A chiusura di questo sezione, vogliamo evidenziare che la medesima efficienza spettrale del segnale SSB può essere conseguita utilizzando due portanti in quadratura, alla stessa frequenza, a ciascuna delle quali sia associato un diverso segnale modulante.

La funzione del tempo che ne risulta può essere scritta come:

$$u(t) = A_c m_1(t) \cos(2\pi f_c t) + A_c m_2(t) \cos(2\pi f_c t) \quad (2.28)$$

La seguente figura mostra a sinistra la generazione del segnale al lato trasmettitore, mentre a destra della figura si recupera $m_1(t)$ e $m_2(t)$ al lato ricevitore:



Questo tipo di schema ci permette di dire che è possibile realizzare delle modulazioni DSB-SC che in ricezione è possibile separare, grazie alla ortogonalità delle portanti.

Pertanto, in questo caso, si pongono gli stessi problemi di ricostruzione della fase, e se ciò non avviene a dovere, il segnale, ad esempio $m_1(t)$, sarà sovrapposto a un residuo dell'altro segnale $m_2(t)$: questa sovrapposizione è definita come cross-talk.

La modulazione di ampiezza con portanti in quadratura è l'archetipo di una delle modulazioni più importanti: la cosiddetta QAM Quadrature Amplitude Modulation, che avremo di approfondire nell'ambito delle modulazioni digitali.

⁹ Slide del prof — Modulazioni analogiche — pag 13 - 14
 Appunti di Damiano — pag 13 - 14
 Appunti — Modulazioni analogiche — pag 13 - 14
 Appunti — 2025-03-03 — pag 13 - 15

2.3.4 Banda laterale ridotta - VSB

¹⁰

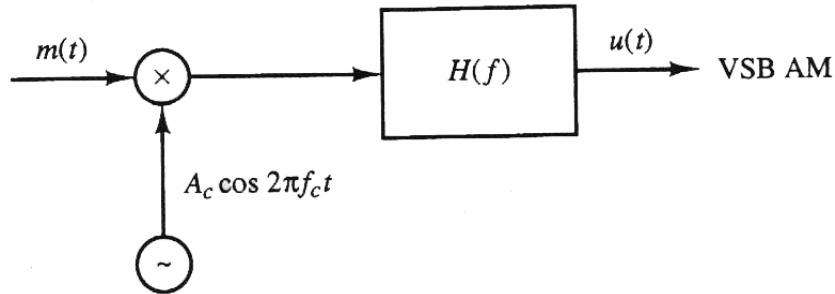
Per i segnali le cui caratteristiche spettrali non consentono l'utilizzo della modulazione SSB, si può migliorare l'efficienza spettrale utilizzando la modulazione di ampiezza a banda laterale ridotta (BLR in italiano, o in inglese VSB Vestigal Side Band).

Con la frase "i segnali le cui caratteristiche spettrali non consentono l'utilizzo della modulazione SSB" si intende che la SSB non è indicata per segnali che hanno una informazione significativa alle basse frequenze.

Letteralmente dall'inglese Vestigal significa residuo

L'idea è quella di trasmettere un segnale che contenga una delle due bande laterali, ma anche una porzione (un vestiglio, un residuo) dell'altra banda laterale.

Il risultato è ottenuto con un opportuno filtraggio ad alta frequenza, come spiegato in questa figura:



Dallo schema notiamo che, nel tempo:

$$u(t) = A_c m(t) \cos(2\pi f_c t) \otimes h(t) \quad (2.29)$$

dove:

- $A_c m(t) \cos(2\pi f_c t)$ è il segnale modulato in DSB-SC
- il segno \otimes indica la convoluzione
- $h(t)$ è la risposta impulsiva del filtro

Dalle proprietà tra tempo e Fourier, sappiamo che, se facciamo una convoluzione nel tempo, avremo un prodotto in frequenza.

Quindi $u(t)$ in Fourier diventa:

$$\begin{aligned} u(t) &= A_c m(t) \cos(2\pi f_c t) \otimes h(t) \\ U(f) &= \frac{A_c}{2} [M(f - f_c) + M(f + f_c)] \cdot H(f) \end{aligned} \quad (2.30)$$

Il filtro introduce massimo un'attenuazione di $\frac{1}{2}$: l'ampiezza della portante eventualmente trasmessa viene dimezzata.

¹⁰Slide del prof — Modulazioni analogiche — pag 14 - 15
Appunti di Damiano — pag 14 - 15
Appunti — Modulazioni analogiche — pag 14 - 15
Appunti — 2025-03-04 — pag 2 - 4
Appunti — 2025-07-08 Ricevimento — pag 10

Nelle applicazioni classiche della VSB (ad esempio al broadcasting televisivo), si utilizza la trasmissione della portante per semplificare il demodulatore, il quale può essere d'inviluppo.

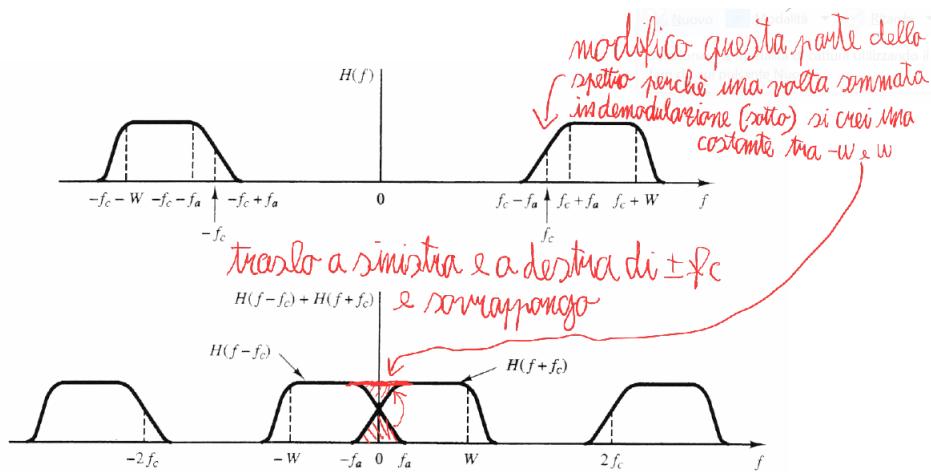
Il filtro $H(f)$ indicato in figura deve soddisfare la seguente condizione:

$$H(f - f_c) + H(f + f_c) = \text{costante} \quad (2.31)$$

Questa relazione deve essere valida per $|f| \leq W$.

Se questa relazione è soddisfatta, la demodulazione del segnale modulante, con banda W , avviene senza distorsione.

Come spiegato dalle seguenti figure:



nella modulazione VSB, bisogna traslare una volta verso destra e una volta verso sinistra, della stessa quantità f_c , la funzione di trasferimento del filtro e di verificare che la somma delle due funzioni risultanti dalla due traslazione sia costante per $-W \leq f \leq +W$.

La banda occupata dal segnale VSB sarà di $W + f_a$, per cui il risparmio in banda può essere significativo.

De-modulando in modo coerente, consideriamo $v(t)$ il segnale de-modulato nel tempo:

$$v(t) = u(t) \cos(2\pi f_c t) \quad (2.32)$$

In Fourier $v(t)$ diventa:

$$\begin{aligned} v(t) &= u'(t) \cos(2\pi f_c t) \\ &\downarrow \\ V(f) &= \frac{1}{2} [U(f - f_c) + U(f + f_c)] \end{aligned} \quad (2.33)$$

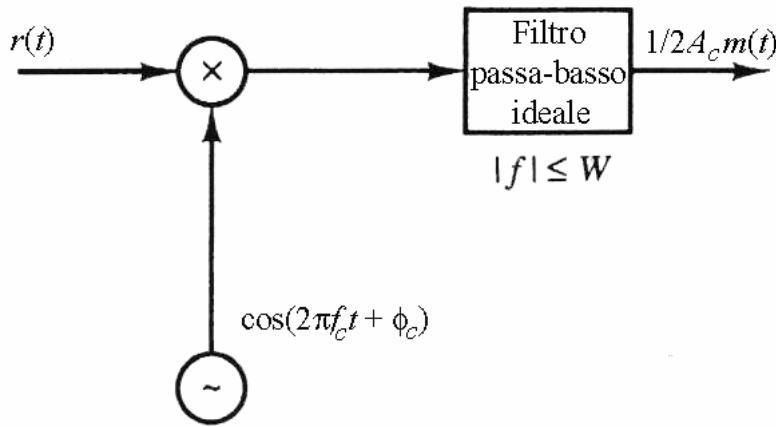
Sostituendo il valore di $U(f)$ a $V(f)$, $V(f)$ diventa:

$$V(f) = \frac{1}{2} [U(f - f_c) + U(f + f_c)]$$

↓

$$V(f) = \frac{A_c}{4} [M(f - 2f_c) + M(f)] \cdot H(f - f_c) + \frac{A_c}{4} [M(f) + M(f + 2f_c)] \cdot H(f + f_c) \quad (2.34)$$

Riportando lo schema del demodulatore coerente:



i fattori $M(f - 2f_c)$ e $M(f + 2f_c)$ vengono eliminati, quindi $V(f)$ si semplifica come:

$$\begin{aligned} V(f) &= \frac{A_c}{4} [M(f - 2f_c) + M(f)] \cdot H(f - f_c) + \frac{A_c}{4} [M(f) + M(f + 2f_c)] \cdot H(f + f_c) \\ &\downarrow \\ V(f) &= \frac{A_c}{4} [0 + M(f)] \cdot H(f - f_c) + \frac{A_c}{4} [M(f) + 0] \cdot H(f + f_c) \\ &= \frac{A_c}{4} M(f) [H(f - f_c) + H(f + f_c)] \end{aligned} \quad (2.35)$$

Siccome il nostro obiettivo è recuperare $M(f)$, cioè $m(t)$, deve essere valida la relazione che abbiamo citato precedentemente:

$$H(f - f_c) + H(f + f_c) = \text{costante} \quad (2.36)$$

e per $|f| \leq W$.

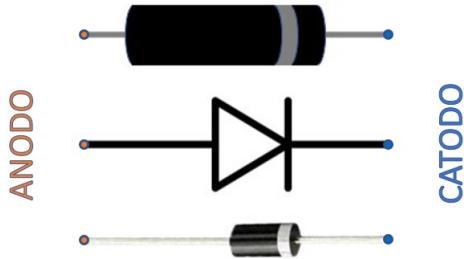
2.3.5 Altri schemi di modulatori di ampiezza

¹¹

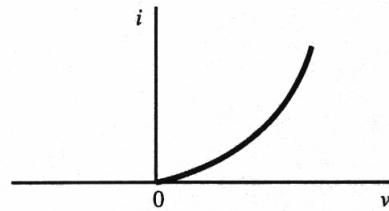
Oltre all'architettura delle modulazioni, è necessario fare qualche esempio riguardo la loro implementazione circuitale.

L'operazione di modulazione richiede, necessariamente, la presenza di componenti non lineari, i soli in grado di produrre le nuove frequenze richieste dalla conversione spettrale.

Un componente tipico è il diodo:



in cui la caratteristica funzione tensione-corrente ha questo tipo di andamento:



In generale, è sufficiente la presenza di un elemento non lineare, la cui caratteristica ingresso-uscita presenta una componente quadratica.

Ragionando in tensione, ad esempio, deve avversi:

$$v_0(t) = a_1 v_i(t) + a_2 v_i^2(t) \quad (2.37)$$

in cui:

- $v_i(t)$ è il segnale in ingresso
- $v_0(t)$ è il segnale in uscita

Applicando in ingresso al dispositivo non lineare una funzione:

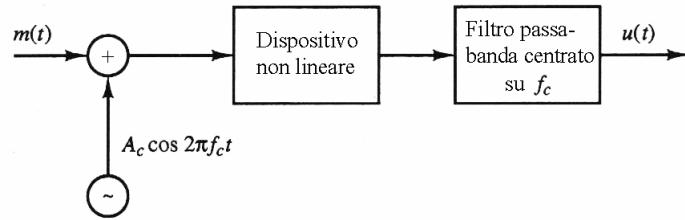
$$v_i(t) = m(t) + A_c \cos(2\pi f_c t) \quad (2.38)$$

e filtrando con un filtro passa-banda centrato su f_c , come in figura:
l'uscita $u(t)$ risulta, con le dovute semplificazioni:

$$u(t) = A_c a_1 \left[1 + \frac{2a_2}{a_1} m(t) \right] \cos(2\pi f_c t) \quad (2.39)$$

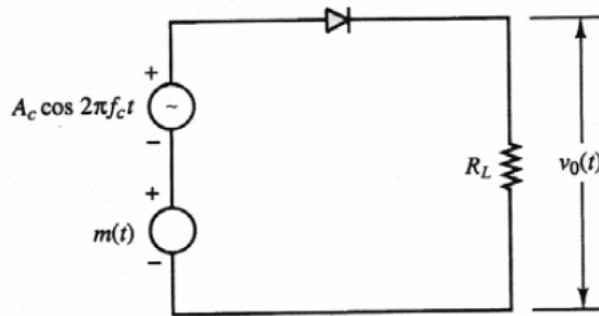
$u(t)$ ha le caratteristiche di un segnale modulato in ampiezza in modo convenzionale, i.e. AM convenzionale.

¹¹Slide del prof — Modulazioni analogiche — pag 15 - 18
 Appunti di Damiano — pag 15 - 18
 Appunti — Modulazioni analogiche — pag 15 - 18
 Appunti — 2025-03-04 — pag 5 - 7

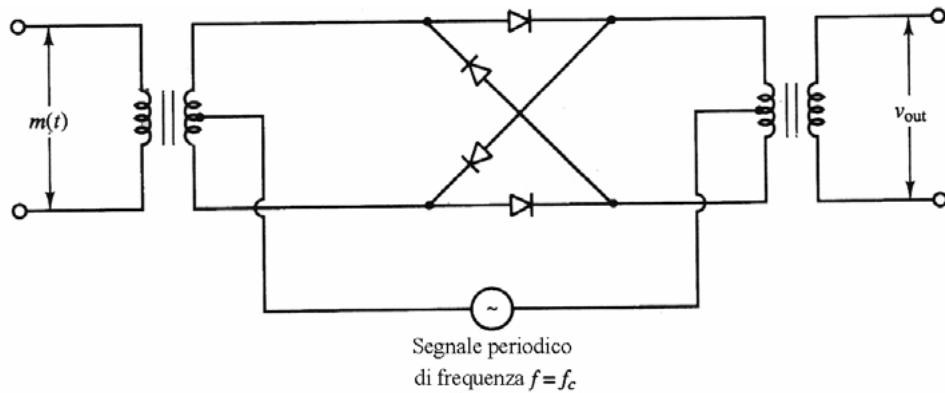


Inoltre, sapendo che $u(t)$ è un segnale convenzionale, il termine $1 + \frac{2a_2}{a_1}m(t)$ non può essere negativo.

Grazie alla funzione caratteristica del diodo, che appunto è un dispositivo non lineare, è possibile realizzare uno "Switching Modulator", da cui si ottiene una modulazione AM-convenzionale, che ha il seguente schema circuitale:



oppure è possibile realizzare un "Ring Modulator", da cui si ottiene una modulazione DSB-SC, che ha il seguente schema circuitale:

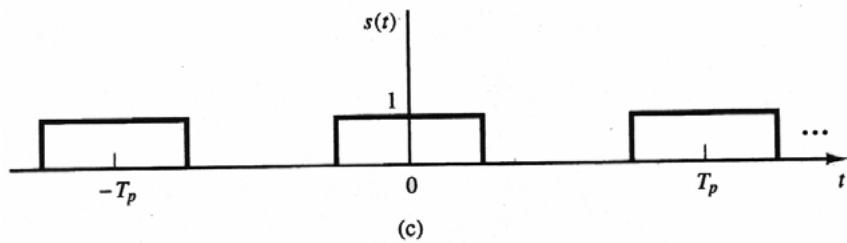


Per lo schema dello "Switching Modulator" e del "Ring Modulator", il principio di funzionamento è quello di moltiplicare il segnale modulante $m(t)$ per una funzione periodica $p(t)$ di periodo $\frac{1}{f_c}$.

$p(t)$, sviluppata in serie di Fourier, conterrà anche la componente di interesse a frequenza f_c .

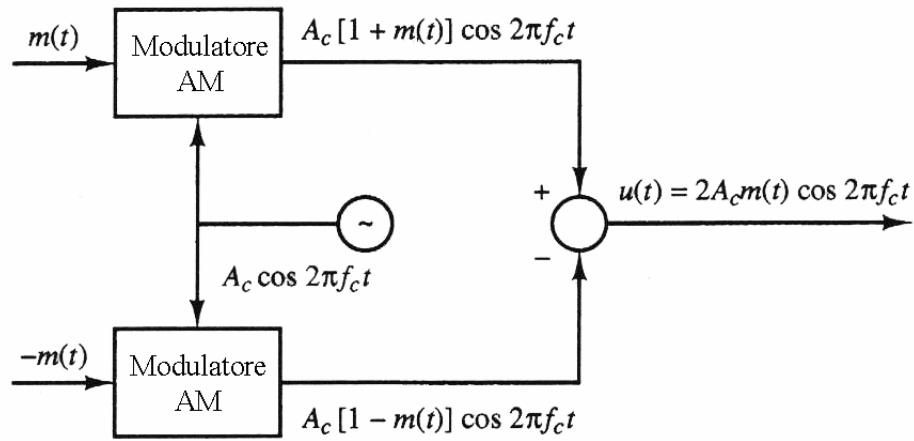
Ad esempio l'uscita del circuito del Ring Modulator, che nel tempo ha questo andamento: può essere scritta matematicamente come:

$$\begin{aligned} v_0(t) &= m(t) \cdot p(t) \\ &= m(t) \frac{4}{\pi} \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{2n-1} \cos[2\pi f_c(2n-1)] \end{aligned} \quad (2.40)$$



Come al solito, la componente di interesse può essere isolata utilizzando, a valle, un opportuno filtro passa-banda.

Un segnale DSB-SC può anche essere ottenuto combinando due modulatori identici AM convenzionali in uno schema bilanciato, come illustrato con la seguente figura:



Chiaraluce potrebbe tranquillamente, a mio parere, far insegnare un corso di Elettronica Analogica. Fine

2.4 Modulazione angolare

¹²

Ricordando che un segnale $c(t)$ (carrier, cioè segnale di portante) sinusoidale può essere descritto dalla seguente equazione:

$$c(t) = A_c \cdot \cos(2\pi f_c t + \phi_c) \quad (2.41)$$

il segnale della portante $c(t)$ può, non solo modificare l'ampiezza del segnale modulante $m(t)$, ma anche modificare la fase ϕ_c o la frequenza f_c .

In particolare, si parla di modulazione in fase PM (Phase Modulation) se il segnale modulato risulta:

$$\begin{aligned} u(t) &= A_c \cos(\theta(t)) \\ &= A_c \cos(2\pi f_c t + \phi(t)) \end{aligned} \quad (2.42)$$

in cui:

$$\phi(t) = k_p m(t) \quad (2.43)$$

dove k_p è definito come costante di deviazione di fase.

Nella modulazione di fase, viene modificata ϕ in funzione del segnale modulante $m(t)$ in maniera proporzionale.

Viene definita la frequenza istantanea del segnale $u(t)$ come:

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi} \frac{d\theta(t)}{dt} \\ &= \frac{1}{2\pi} \frac{d}{dt}(2\pi f_c t + \phi(t)) \\ &= f_c + f_c \frac{1}{2\pi} \frac{d\phi(t)}{dt} \end{aligned} \quad (2.44)$$

$f_i(t)$ viene definita come modulazione di frequenza FM (Frequency Modulation) se risulta:

$$f_i(t) - f_c = k_f m(t) \quad (2.45)$$

dove k_f è definita come costante di deviazione di frequenza.

In altri termini, viene definita modulazione di frequenza quando la differenza rispetto a $f_c(t)$ è proporzionale a $m(t)$ segnale modulante.

Se:

$$m(t) = 0 \quad (2.46)$$

nella modulazione di frequenza si trasmette solo $f_c(t)$, cioè la portante.

¹²Slide del prof — Modulazioni analogiche — pag 18 - 21
 Appunti di Damiano — pag 18 - 21
 Appunti — Modulazioni analogiche — pag 18 - 21
 Appunti — 2025-03-04 — pag 8 - 9
 Appunti — 2025-03-07 — pag 2

La modulazione FM e PM sono strettamente legate: dove esiste una esiste l'altra e viceversa.

In effetti, integrando l'espressione di $f_i(t)$ e tenendo conto dell'espressione di $k_f m(t)$, possiamo esprimere la fase $\phi(t)$ come:

$$\phi(t) = 2\pi k_f \int_{-\infty}^t m(\tau) d\tau \quad (2.47)$$

A sua volta, derivando:

$$\dot{\phi}(t) = k_p m(t) \quad (2.48)$$

e tenendo conto dell'espressione di $f_i(t)$, si conclude che la PM è esprimibile come FM se:

$$f_i(t) - f_c = \frac{1}{2\pi} k_p \frac{d}{dt} m(t) \quad (2.49)$$

In base a queste due espressioni, possiamo dire che la FM può essere vista come una PM in cui $m(t)$ viene integrato, invece la PM può essere vista come una FM in cui $m(t)$ viene derivata.

Ricapitolando si parla di FM o PM quando:

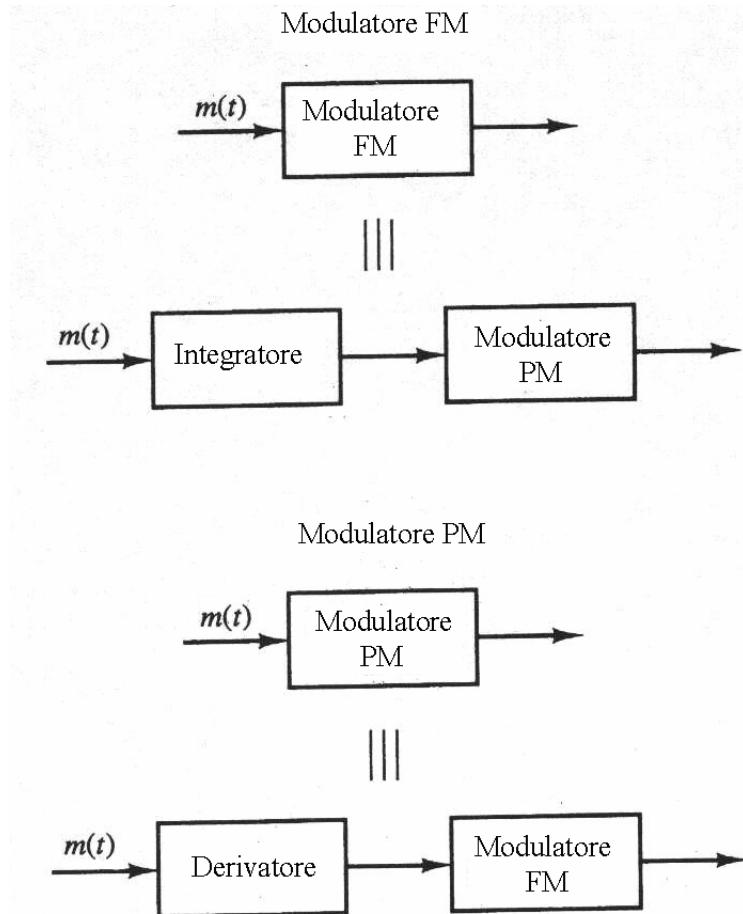
$$\phi(t) = \begin{cases} k_p m(t) & \text{PM} \\ 2\pi k_f \int_{-\infty}^t m(\tau) d\tau & \text{FM} \end{cases} \quad (2.50)$$

e

$$f_i(t) - f_c = \begin{cases} \frac{1}{2\pi} k_p \frac{d}{dt} m(t) & \text{PM} \\ k_f m(t) & \text{FM} \end{cases} \quad (2.51)$$

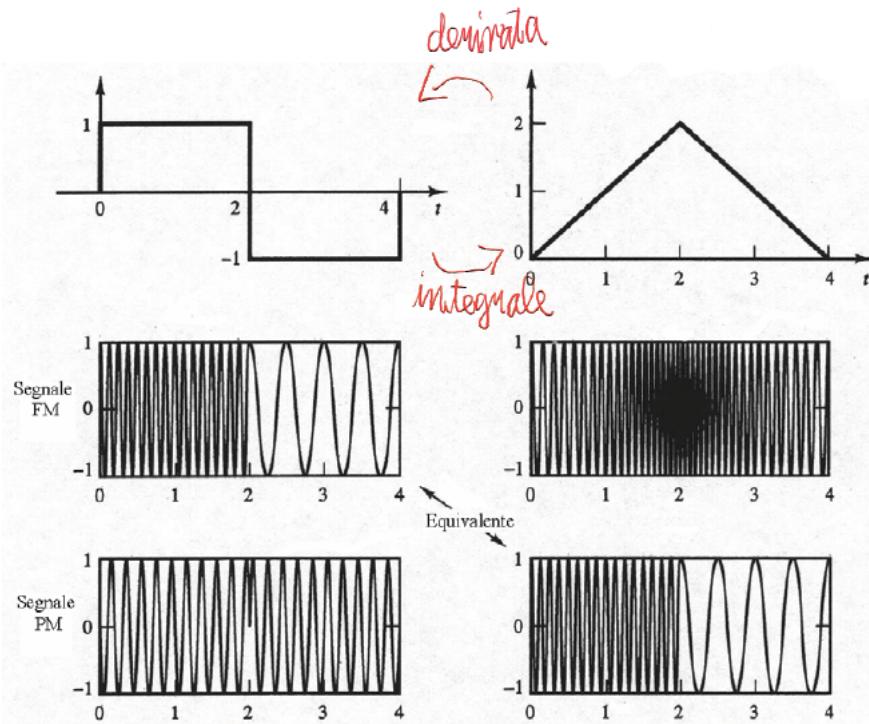
Queste considerazioni danno un'indicazione esplicita su come realizzare il modulatore.

Come viene spiegato dalle seguenti figure:



con l'aggiunta di un circuito o integratore o derivatore, è possibile realizzare un modulatore FM o PM partendo dal suo reciproco.

Inoltre, visualizzando i segnali di ingresso e quelli di uscita dei modulatori:



possiamo osservare che i segnali FM e PM, con le osservazioni svolte, sono equivalenti.

Stabilita questa corrispondenza tra le due tecniche di modulazione, si può parlare genericamente di modulazione angolare.

2.4.1 Indici della modulazione angolare

¹³

Per il fatto che il segnale modulante agisce sull'argomento della portante, la modulazione angolare è più difficile da trattare rispetto alla modulazione in ampiezza.

La modulazione di ampiezza è molto semplice: basta traslare lo spettro del segnale in BB a frequenza di portante.

Invece questo fatto, nelle angolazioni angolari, non avviene; anche la modulazione angolare è definita come modulazione non lineare, a differenza di quella di ampiezza.

Per questo motivo, molti dei calcoli relativi alla modulazione angolare sono stati preventivati considerando, come segnale modulante, il segnale sinusoidale.

Per una modulante sinusoidale:

$$m(t) = a \cos(2\pi f_m t) \quad (2.52)$$

si definisce come indici di modulazione, rispettivamente di fase β_p e di frequenza β_f , come segue:

$$\beta_p = k_p a \quad (2.53)$$

$$\beta_f = \frac{k_f a}{f_m} \quad (2.54)$$

Grazie a questi indici, possiamo esprimere il segnale modulato $u(t)$ come:

$$u(t) = \begin{cases} A_c \cos [2\pi f_c t + \beta_p \cos(2\pi f_m t)] & \text{PM} \\ A_c \cos [2\pi f_c t + \beta_f \sin(2\pi f_m t)] & \text{FM} \end{cases} \quad (2.55)$$

È immediato verificare che $k_p a$ rappresenta la massima deviazione di fase $\Delta\phi_{max}$:

$$\Delta\phi_{max} = k_p \max |m(t)| \quad (2.56)$$

nella PM.

Invece nella FM, $k_f a$ rappresenta la massima deviazione di frequenza Δf_{max} :

$$\Delta f_{max} = k_f \max |m(t)| \quad (2.57)$$

Grazie a queste definizioni, possiamo riscrivere β_p e β_f come:

$$\beta_p = \Delta\phi_{max} \quad (2.58)$$

$$\beta_f = \frac{\Delta f_{max}}{f_m} \quad (2.59)$$

Originariamente introdotte per il caso di segnale modulante sinusoidale, le definizioni di β_p e β_f possono essere estese ad un segnale modulante qualsiasi, a patto di sostituire f_m con la sua larghezza di

¹³Slide del prof — Modulazioni analogiche — pag 21 - 23

Appunti di Damiano — pag 21 - 23

Appunti — Modulazioni analogiche — pag 21 - 23

Appunti — 2025-03-07 — pag 2 - 3

banda W.

Formalmente, per la PM, per un segnale qualsiasi modulante di larghezza di banda W, rimane uguale a prima, quindi:

$$\beta_p = \Delta\phi_{max} \quad (2.60)$$

invece, per la FM, diventa:

$$\beta_f = \frac{\Delta f_{max}}{W} \quad (2.61)$$

Se un segnale avesse più frequenze, bisogna considerare la banda totale del segnale modulante W.

2.4.1.1 Modulazione angolare a basso indice

¹⁴

L'indice di modulazione β è un parametro fondamentale della modulazione angolare e svolge un ruolo chiave nella qualità di trasmissione quando il segnale è affetto da rumore.

Si parla di modulazione angolare a basso indice (o modulazione angolare a banda stretta) quando la variazione di $\phi(t)$:

$$\phi(t) = \begin{cases} k_p m(t) & \text{PM} \\ 2\pi k_f \int_{-\infty}^t m(\tau) d\tau & \text{FM} \end{cases} \quad (2.62)$$

risulta molto piccola, cioè:

$$\phi(t) \ll 1 \quad (2.63)$$

Sotto questa ipotesi, il segnale modulato con modulazione angolare da:

$$\begin{aligned} u(t) &= A_c \cos(\theta(t)) \\ &= A_c \cos(2\pi f_c t + \phi(t)) \end{aligned} \quad (2.64)$$

diventa:

$$\begin{aligned} u(t) &= A_c \cos(2\pi f_c t + \phi(t)) \\ &= A_c \cos(2\pi f_c t) \cos(\phi(t)) - A_c \sin(2\pi f_c t) \sin(\phi(t)) \\ &\approx A_c \cos(2\pi f_c t) - A_c \sin(2\pi f_c t) \phi(t) \\ &= A_c \cos(2\pi f_c t) - A_c \phi(t) \sin(2\pi f_c t) \end{aligned} \quad (2.65)$$

Non dobbiamo impararci le formule a memoria.

Dobbiamo solo sapere che l'informazione è contenuta nella fase $\phi(t)$.

Dall'analisi matematica 1, si è potuto approssimare $\sin(\phi(t))$ a $\phi(t)$ grazie alla teoria degli o-piccoli.

Ripassino al volo che non fa mai male:

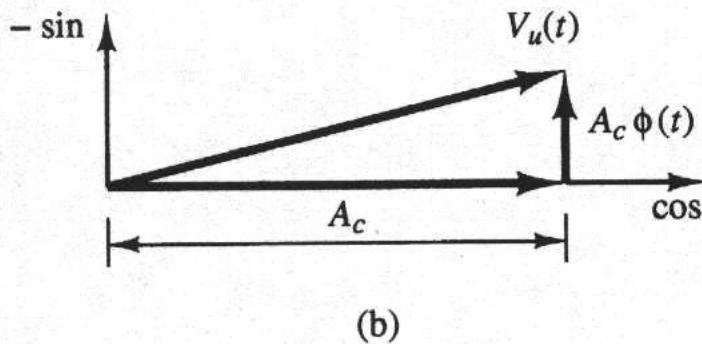
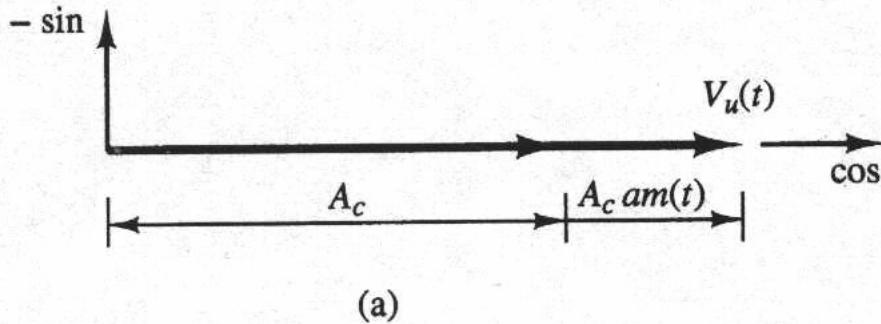
<https://www.youmath.it/lezioni/analisi-matematica/limiti-continuita-e-asintoti/3247-o-piccolo.html>

Questa espressione di $u(t)$ esplicita la somiglianza con l'espressione di un segnale modulato in AM convenzionale:

$$\begin{aligned} u(t) &= A_c [1 + \alpha \cdot m_n(t)] \cos(2\pi f_c t + \phi_c) \\ &= A_c \cos(2\pi f_c t + \phi_c) + A_c \alpha \cdot m_n(t) \cos(2\pi f_c t + \phi_c) \end{aligned} \quad (2.66)$$

¹⁴Slide del prof — Modulazioni analogiche — pag 23 - 24
 Appunti di Damiano — pag 23 - 24
 Appunti — Modulazioni analogiche — pag 23 - 24
 Appunti — 2025-03-07 — pag 3 - 4
 Appunti — 2025-07-08 Ricevimento — pag 10

Utilizzando il piano dei fasori, possiamo graficare le due espressione di $u(t)$:



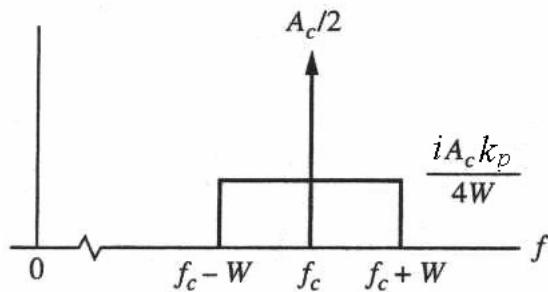
dove la prima in alto è $u(t)$ con modulazione AM-convenzionale, invece la seconda in basso è $u(t)$ con modulazione angolare a basso indice.

Si può svolgere la seguente osservazione: nella AM convenzionale, il modulo di $V_u(t)$ varia in base alla modulazione (vedo $m(t)$ che si trova sul piano del coseno) e il suo angolo rimane costante; invece, nella modulazione angolare, abbiamo che il modulo di $V_u(t)$ rimane costante, ma varia il suo angolo (vedi $\phi(t)$ sul piano del seno).

Per fissare bene le idee, consideriamo al caso PM in cui il segnale modulato vale:

$$u(t) \approx A_c \cos(2\pi f_c t) - A_c k_p m(t) \sin(2\pi f_c t) \quad (2.67)$$

Facendo la trasformata di Fourier di $u(t)$ e considerando $m(t)$ reale pari, sapendo le proprietà tra trasformata di Fourier nel tempo e in frequenza anche $M(f)$ sarà a sua volta reale pari, possiamo visualizzare $U(f)$ come:



dove in figura si considerano solo le frequenze positive.

Si parla di modulazione angolare a banda stretta perché, come si visualizza in figura, $2W$ è la minima occupazione spettrale conseguibile in modulazione angolare, quindi PM e FM, e si ottiene sono nell'ipotesi di:

$$\phi(t) \ll 1 \quad (2.68)$$

2.4.2 Banda di una modulazione angolare

15

Assumendo per il segnale modulante $m(t)$ una funzione cosinusoidale con frequenza f_m , si può scrivere:

$$u(t) = A_c \cos(2\pi f_c t + \beta \cos(2\pi f_m t)) \quad (2.69)$$

con $\beta = \beta_p$ se si fa una PM o $\beta = \beta_f$ se si fa una FM.

Visto che analiticamente è difficile svolgere un coseno dentro ad un coseno, $u(t)$ può essere riscritta come:

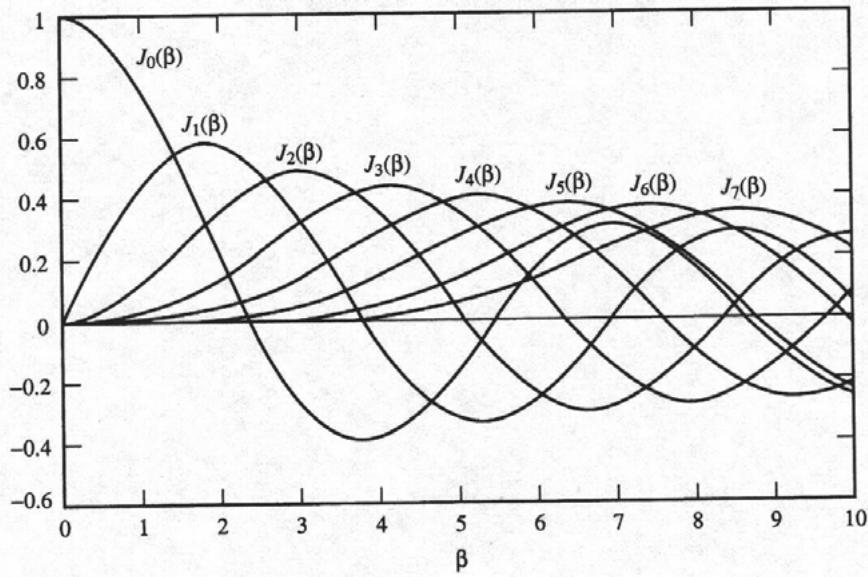
$$u(t) = \sum_{n=-\infty}^{+\infty} A_c J_n(\beta) \cos [2\pi(f_c + n f_m)t] \quad (2.70)$$

dove $J_n(\beta)$ è la funzione di Bessel di prima specie di ordine n .

Cosa sono e perché si utilizzano in casi particolari le funzioni di Bessel?

https://it.wikipedia.org/wiki/Equazioni_di_Bessel

Alcuni andamenti di $J_n(\beta)$:



Sulla base della formula di $u(t)$, notiamo che $u(t)$ è esprimibile con una sommatoria di infiniti termini, in cui avremo, nello spettro, infinite armoniche, quindi infinite righe ogni frequenza multipla di f_m .

Per questo motivo la modulazione angolare non viene definita lineare come la modulazione in ampiezza, perché, in questo caso, lo spettro in BB non viene semplicemente traslato alla frequenza della portante.

Inoltre, sapendo che la potenza del segnale modulato è costante a pari a $\frac{A_c^2}{2}$ (dalle proprietà di ciclo-stazionarietà dello spettro in frequenza), si può dire che la potenza delle funzioni di Bessel vale:

$$\sum_{n=-\infty}^{+\infty} [J_n(\beta)]^2 = 1 \quad (2.71)$$

¹⁵Slide del prof — Modulazioni analogiche — pag 25 - 26
Appunti di Damiano — pag 25 - 26
Appunti — Modulazioni analogiche — pag 25 - 26
Appunti — 2025-03-07 — pag 4

Questa è una proprietà delle funzioni di Bessel di prima specie.

L'occupazione illimitata della banda non consentirebbe un utilizzo pratico del sistema.

Se però si visualizza l'andamento delle funzioni di Bessel, si nota che, per ogni valore di β , c'è un numero limitato di funzioni di Bessel che danno un contributo significativo.

In particolare, risulta applicabile la regola secondo cui le funzioni di Bessel di ordine n:

$$n \leq \beta + 1 \quad (2.72)$$

sono apprezzabilmente diverse da zero.

Invece le funzioni di Bessel di ordine n:

$$n > \beta + 1 \quad (2.73)$$

sono trascurabili.

Sulla base di queste considerazioni, si può concludere che delle infinite armoniche nominalmente presenti nello spettro del segnale modulato danno, in realtà, contributo significativo allo spettro solo quelle che distano dalla portante non più di $(\beta + 1)$ volte f_m .

Tenendo conto della distribuzione bilatera rispetto alla portante, ne risulta una banda "effettiva" occupata dalla modulazioni angolari di:

$$B_c = 2(\beta + 1)f_m \quad (2.74)$$

B_c contiene il 98 % della potenza totale del segnale.

Sapendo la relazione tra β e B_c , possiamo esprimere B_c per le due modulazioni angolari studiate:

$$B_c = \begin{cases} 2(k_p a + 1)f_m & \text{PM} \\ 2(k_f a + f_m) & \text{FM} \end{cases} \quad (2.75)$$

dove a è l'ampiezza del segnale modulante.

Da queste due formule di B_c per PM e FM si nota che all'aumentare di a, aumenta B_c sia nella PM che nella FM.

Invece, f_m nella PM moltiplica il termine $2(k_p a + 1)$, invece nella FM, f_m ha un ruolo additivo al termine $k_f a$.

Quindi, all'aumentare di f_m , la banda B_c aumenta nettamente di più nella PM rispetto alla FM.

Queste formule della banda B_c sono valide per un segnale modulante cosinusoidale.

2.4.2.1 Banda di una modulazione angolare per un segnale modulante qualsiasi

¹⁶

Nel caso in cui si considera un generico segnale periodico (che possiamo generalizzare come caso di modulante cosinusoidale) e un processo gaussiano (esemplificativo di una funzione stocastica), si possono adattare i parametri visti precedente a questo caso.

La formula di B_c vista precedentemente è valida se a f_m si sostituisce la frequenza massima significativa W dello spettro del segnale modulante:

$$B_c = \begin{cases} 2(k_p a + 1)f_m & \text{PM} \\ 2(k_f a + f_m) & \text{FM} \end{cases}$$

↓

$$B_c = \begin{cases} 2(k_p a + 1)W & \text{PM} \\ 2(k_f a + W) & \text{FM} \end{cases}$$
(2.76)

Lo stesso principio vale per l'altra formula di B_c :

$$B_c = 2(\beta + 1)f_m$$

↓

$$B_c = 2(\beta + 1)W$$
(2.77)

Queste formula di B_c costituisce una eccellente approssimazione della banda occupata da un segnale in modulazione angolare.

Questa formula è universalmente nota come "regola di Carson".

Ma, è bene precisare che, la formula di Carson nel caso FM, così calcolata, è una sottostima di B_c se il valore di β è compreso tra 2 e 10.

Quindi la formula di Carson, in questo caso particolare diventa:

$$B_c = 2(\beta_f + 2)W \text{ FM per } 2 < \beta_f < 10$$
(2.78)

¹⁶Slide del prof — Modulazioni analogiche — pag 26 - 27
 Appunti di Damiano — pag 26 - 27
 Appunti — Modulazioni analogiche — pag 26 - 27
 Appunti — 2025-03-07 — pag 4 - 5

2.4.3 Mo/demodulatori angolari

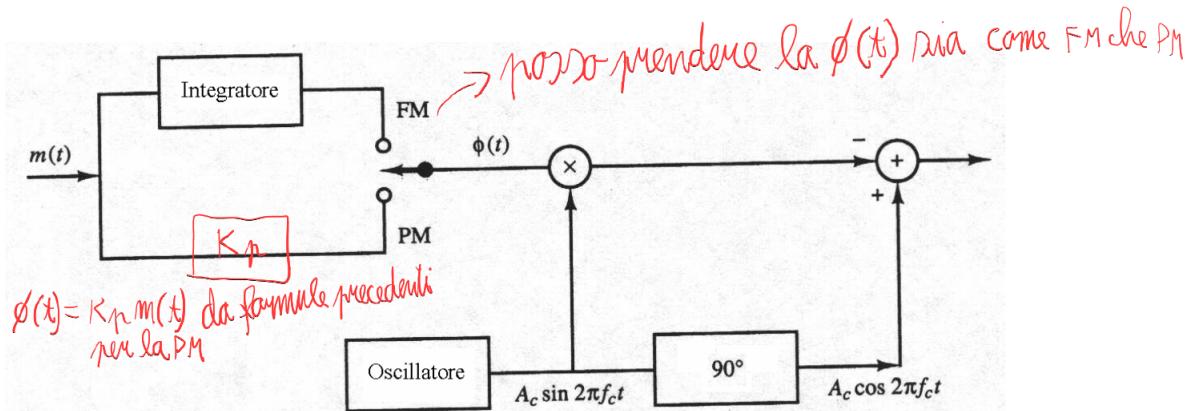
17

Dalla formula della modulazione angolare PM a basso indice $u(t)$:

$$u(t) \approx A_c \cos(2\pi f_c t) - A_c k_p m(t) \sin(2\pi f_c t) \quad (2.79)$$

possiamo dire che la modulazione angolare può essere vista come una modulazione AM-convenzionale con uno shift di 90° (il seno rispetto al coseno differiscono di 90°) della portante.

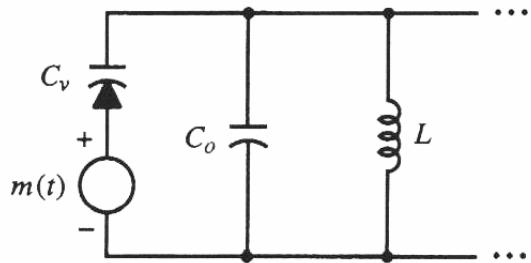
Grazie a questa osservazione, possiamo costruire il seguente schema del modulatore angolare:



Questo tipo di modulatore angolare è definito come modulatore di Armstrong.

Invece, per un segnale FM, non necessariamente a basso indice, si ottiene utilizzando un diodo varactor o chiamato anche oscillatore in tensione (VCO: Voltage Controlled Oscillator).

Un possibile circuito per modulare un segnale FM a basso indice:

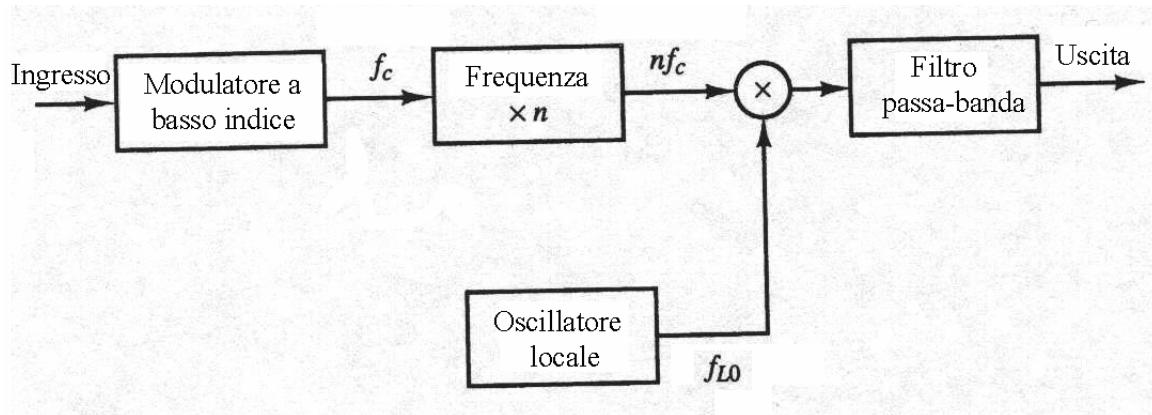


Quel componente segnato come C_v che sembra sia un diodo che una capacità è un condensatore variabile che cambia il suo valore di capacità in base alla tensione ai suoi capi, in questo circuito C_v cambia la sua capacità in base alla tensione di $m(t)$

Un altro circuito per generare un segnale FM o PM con indice di modulazione β generico è quello del narrowband angle modulator, che fa riferimento al modulatore di Armstrong ma a banda stretta.

¹⁷Slide del prof — Modulazioni analogiche — pag 27 - 31
 Appunti di Damiano — pag 27 - 31
 Appunti — Modulazioni analogiche — pag 27 - 31
 Appunti — 2025-03-10 — pag 2 - 4
 Appunti — 2025-07-08 Ricevimento — pag 10

Di seguito, il disegno dell'architettura:



L'introduzione dell'oscillatore locale a frequenza f_{LO} è dovuta alla necessità di allocare il segnale trasmesso alla frequenza di portante desiderata, ove tale non fosse il risultato della moltiplicazione di $n \cdot f_c$.

L'oscillatore locale serve per una conversione in frequenza.

Il filtro passa-banda toglie tutte le componenti che non sono necessarie.

La tecnica che applica questo tipo di architettura circuitale prende il nome di metodo indiretto.

2.4.3.1 Demodulatori angolari

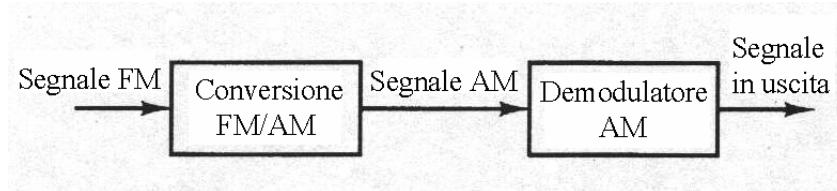
¹⁸

Benché, in linea di principio, la demodulazione di un segnale PM o FM possa essere realizzata con tecniche coerenti (quindi moltiplicando per delle funzioni cosinusoidali), in grado di estrarre la fase del segnale (in cui è associata l'informazione), in pratica si preferisce di adottare tecniche non coerenti, in modo da semplificare drasticamente la struttura del ricevitore.

Consideriamo il caso solo della demodulazione in FM, sapendo che il demodulatore FM può essere anche un demodulatore PM.

L'idea è quella di convertire il segnale FM in segnale AM-convenzionale prima di procedere alla demodulazione.

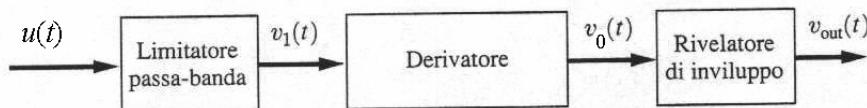
Lo schema del demodulatore FM è il seguente:



in cui il convertitore FM/AM è un banale derivatore.

Ora andiamo a calcolare analiticamente la demodulazione.

Un possibile schema di un convertitore FM/AM:



Consideriamo il segnale in ingresso $u(t)$ modulato:

$$u(t) = A_c \cos \left(2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau \right) \quad (2.80)$$

La funzione di trasferimento $H(f)$ di un blocco derivatore è caratterizzato dalle seguenti relazioni:

$$|H(f)| = V_0 + k(f - f_c) \text{ per } |f - f_c| \leq \frac{B_c}{2} \quad (2.81)$$

Dalle relazioni tra dominio del tempo e frequenza, una derivata nel tempo, dobbiamo avere un'onda sinusoidale nell'intorno della frequenza da considerare, in questo caso nell'intorno di f_c .

Per realizzare questo filtro derivatore, si usano dei circuiti risonanti e si trova il loro tratto lineare.

Essendo circuiti risonanti composti da induttori, questi permettono di svolgere delle derivate nel tempo, grazie alla loro funzione caratteristica.

Per il resto, si tratta di una funzione puramente immaginaria, e dunque asimmetrica rispetto all'origine.

¹⁸Slide del prof — Modulazioni analogiche — pag 28
Appunti di Damiano — pag 28
Appunti — 2025-03-10 — pag 2
Appunti — 2025-07-08 Ricevimento — pag 9

Il segnale che esce fuori dopo il filtro è:

$$v_0(t) = -A_c [V_0 + kk_f m(t)] \sin \left(2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau \right) \quad (2.82)$$

Il rivelatore di inviluppo estrae dalla funzione di $v_0(t)$ l'argomento $-A_c [V_0 + kk_f m(t)]$

Capitolo 3

Qualità nelle modulazioni analogiche



3.1 Sistema in banda base

1

Stai molto attento alla terminologia utilizzata in questo capitolo perché può differire dagli altri capitoli, quindi prima leggi sopra la descrizione e la spiegazione della formula, poi applicala al compito scritto / negli esercizi

Consideriamo, al momento, solo il sistema in banda base.

Il ricevitore è costituito da un filtro passa-basso con banda W , dove W è la larghezza di banda del segnale utile.

Se il filtro passa-basso ha una banda W minore della banda del segnale ricevuto, il segnale ricevuto verrà distorto.

Se il filtro passa-basso ha una banda W maggiore della banda del segnale ricevuto, il segnale ricevuto non verrà distorto, ma il filtro aggiungerà del rumore sulle frequenze aggiunte.

Considera la banda del filtro passa-basso come l'altezza di una porta.

Se io sono una persona molta alta (segnale ricevuto a banda larga) ed ho una porta molto alta, posso tranquillamente entrare nella stanza (segnale ricevuto non distorto) rimanendo intatto, ma avrai dei tacchi in modo da avere la stessa altezza della porta (aggiunta del rumore).

Se io sono una persona molto alta ed ho una porta molto bassa, o non ci passo o dovrei "tagliarmi" in più pezzettini (segnale ricevuto distorto).

In formule, il $(\frac{S}{N})_b$ è:

$$\left(\frac{S}{N}\right)_b = \frac{P_R}{N_0 \cdot W} \quad (3.1)$$

dove:

- $(\frac{S}{N})_b$ è il rapporto segnale rumore in banda base (dall'inglese base-band Signal-to-Noise ratio)
- P_R è la potenza utile ricevuto
- N_0 è la potenza di rumore termico in uscita

Come abbiamo studiato nei precedenti capitoli, la potenza di rumore termico in uscita da un filtro ideale:

$$\begin{aligned} P_{n_o} &= \int_{-W}^{+W} \frac{N_0}{2} df \\ &= N_0 \cdot W \end{aligned} \quad (3.2)$$

dove $\frac{N_0}{2}$ è la densità spettrale di rumore termico bilatera in ingresso.

Il nostro obiettivo è quello di esprimere il rapporto segnale rumore per le varie tecniche di modulazione, in termini del rapporto segnale-rumore in banda base.

¹Slide del prof — Qualità nelle modulazioni analogiche — pag 1.1
Appunti di Damiano — pag 1.1
Slide — Qualità nelle modulazioni analogiche — pag 1.1
Appunti — 2025-03-10 — pag 5 - 6
Appunti — 2025-07-08 Ricevimento — pag 8.2

In formule:

$$\left(\frac{S}{N}\right)_o = A \cdot \left(\frac{S}{N}\right)_b \quad (3.3)$$

dove:

- $\left(\frac{S}{N}\right)_o$ è il rapporto segnale rumore dopo la modulazione
- $\left(\frac{S}{N}\right)_b$ è il rapporto segnale rumore in banda base
- A è un coefficiente moltiplicativo di qualità della modulazione

Il coefficiente moltiplicativo A della qualità della modulazione può assumere diversi valori.

Se $A = 1$, la modulazione ha la stessa qualità del segnale in banda base, cioè:

$$\left(\frac{S}{N}\right)_o = \left(\frac{S}{N}\right)_b \quad (3.4)$$

Se $A < 1$, la modulazione ha una prestazione inferiore rispetto al segnale in banda base, cioè:

$$\left(\frac{S}{N}\right)_o < \left(\frac{S}{N}\right)_b \quad (3.5)$$

Se $A > 1$, la modulazione ha una prestazione migliore rispetto al segnale in banda base, cioè:

$$\left(\frac{S}{N}\right)_o > \left(\frac{S}{N}\right)_b \quad (3.6)$$

Come vedremo in seguito, alcune modulazioni avranno un SNR peggiore in modulazione rispetto all'SNR in banda base, ma verranno scelte quelle modulazioni non per la qualità ma per altri pregi.

Ad esempio l'AM-convenzionale avrà un SNR peggiore in modulazione rispetto all'SNR in banda base, ma verrà scelta perché il demodulatore ha un basso costo economico di progettazione.

Quando si fa un progetto, generalmente si fissa l'SNR e poi si realizza il progetto.

Il progettista del sistema di telecomunicazioni fissa un SNR in banda base che deve essere in uscita dal ricevitore.

Lo stadio in cui si analizza e quantifica il rapporto segnale-rumore in banda base dipende dal progetto.

Il termine ricevitore, generalmente (ma può variare a caso a caso) si intende a valle, cioè in uscita, del ricevitore, in particolare lo stadio dopo la demodulazione.

Ad esempio, se ci troviamo in radio-frequenza, possiamo fissare e progettare l'SNR in banda base dopo lo stadio di ingresso del ricevitore (cioè il segnale dopo l'antenna, dopo il cavo dall'antenna alla radio, dopo la radio al demodulatore).

Ma in altri casi lo si può fissare dopo il quantizzatore.

Come tutti gli episodi della vita, dipende.

https://www.youtube.com/watch?v=cf_ITVX7qCg

JARABE DE PALO - "DIPENDE" (Video Oficial en italiano)

3.2 Qualità della DSB-SC

2

Focalizziamoci adesso sulla DSB-SC.

Dai capitoli precedenti, sappiamo che il segnale modulato in DSB-SC $u(t)$ ha la seguente formula (considerando $\phi_c = 0$ per semplicità di notazione):

$$\begin{aligned}
 u(t) &= m(t) \cdot c(t) \\
 &= m(t) \cdot [A_c \cdot \cos(2\pi f_c t + \phi_c)] \\
 &= [A_c \cdot m(t)] \cdot \cos(2\pi f_c t + \phi_c) \\
 &= A_c m(t) \cos(2\pi f_c t + \phi_c) \\
 &\quad \downarrow \\
 u(t) &= A_c m(t) \cos(2\pi f_c t + 0) \\
 &= A_c m(t) \cos(2\pi f_c t)
 \end{aligned} \tag{3.7}$$

dove:

- $u(t)$ è il segnale modulato in DSB-SC
- $m(t)$ è il segnale modulante
- $c(t)$ è il segnale della portante

Considerando un rumore additivo $n(t)$ perché è AWGN, allora il segnale che riceviamo $r(t)$ è uguale a:

$$\begin{aligned}
 r(t) &= u(t) + n(t) \\
 &= [A_c m(t) \cos(2\pi f_c t)] + [n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t)]
 \end{aligned} \tag{3.8}$$

dove:

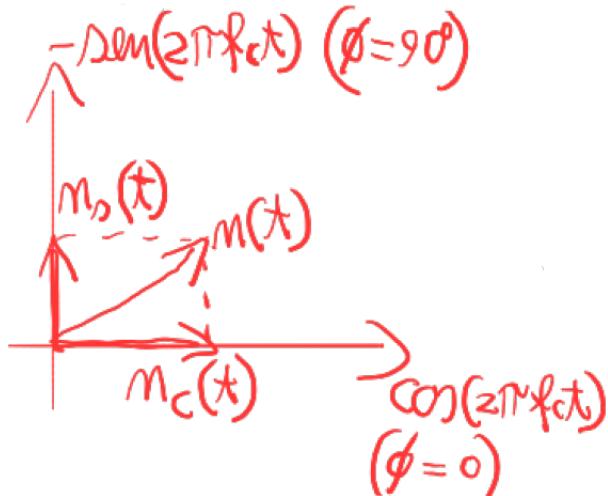
- $n_c(t)$ è la componente del rumore in fase
- $n_s(t)$ è la componente di rumore in quadratura

²Slide del prof — Qualità nelle modulazioni analogiche — pag 1.2 - 3.1
Appunti di Damiano — pag 1.2 - 3.1
Slide — Qualità nelle modulazioni analogiche — pag 1.2 - 3.1
Appunti — 2025-03-10 — pag 6 - 8

Utilizzando il piano dei fasori, possiamo visualizzare $n(t)$:

$$n(t) = n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t) \quad (3.9)$$

come:



Quindi il rumore $n(t)$ è un vettore, in cui $n_c(t)$ è la proiezione di $n(t)$ dell'asse dei coseni, $n_s(t)$ è la proiezione del rumore $n(t)$ nell'asse dei seni.

Ma, dalla teoria dei Segnali Determinati Aleatori, sappiamo che:

$$\langle n_c(t) \rangle = \langle n_c(t) \rangle = 0 \quad (3.10)$$

essendo un rumore AWGN, il valore medio del rumore $n(t)$ è nullo.

Invece, per quanto riguarda la varianza del rumore $n(t)$:

$$\langle n_c^2(t) \rangle = \langle n_c^2(t) \rangle = \langle n^2(t) \rangle = \sigma_n^2 \quad (3.11)$$

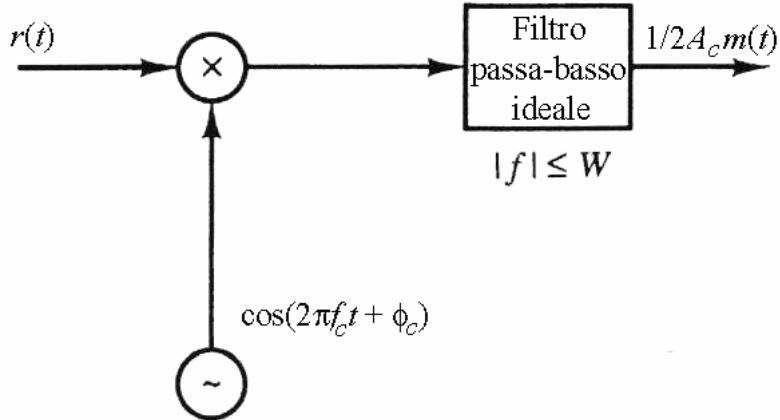
cioè, in parole e per esteso, il valore medio quadratico del rumore $\langle n^2(t) \rangle$ è uguale alla varianza del rumore σ_n^2 .

Per la demodulazione, scegliamo un demodulatore coerente, ciò significa moltiplicare $r(t)$ per un coseno con la frequenza f_c della portante.

In formule step-by-step, se consideriamo $y(t)$ il segnale demodulato in modo coerente:

$$\begin{aligned} y(t) &= r(t) \cdot \cos(2\pi f_c t) \\ &= A_c m(t) \cos^2(2\pi f_c t) + n_c(t) \cos^2(2\pi f_c t) - n_s(t) \sin(2\pi f_c t) \cos(2\pi f_c t) \\ &= \frac{A_c m(t)}{2} + \frac{A_c m(t) \cos(4\pi f_c t)}{2} + \frac{n_c(t)}{2} + \frac{n_c(t) \cos(4\pi f_c t)}{2} - \frac{n_s(t) \sin(4\pi f_c t)}{2} \end{aligned} \quad (3.12)$$

Ricordando lo schema di una demodulazione coerente:



togliamo dal segnale demodulato tutte le componenti oltre a $2\pi f_c t$, che fisicamente verranno tolte dal filtro passa-basso ideale:

$$\begin{aligned}
 y(t) &= \frac{A_c m(t)}{2} + \frac{A_c m(t) \cos(4\pi f_c t)}{2} + \frac{n_c(t)}{2} + \frac{n_c(t) \cos(4\pi f_c t)}{2} - \frac{n_s(t) \sin(4\pi f_c t)}{2} \\
 &\downarrow \\
 y(t) &= \frac{A_c m(t)}{2} + 0 + \frac{n_c(t)}{2} + 0 - 0 \\
 &= \frac{A_c m(t)}{2} + \frac{n_c(t)}{2}
 \end{aligned} \tag{3.13}$$

Quindi, dopo la demodulazione, ciò che ci resta è:

$$y(t) = \frac{A_c m(t)}{2} + \frac{n_c(t)}{2} \tag{3.14}$$

Dalla Teoria dei Segnali Determinati sappiamo che, essendo il segnale di potenza ergodico, per calcolare la potenza del segnale dobbiamo calcolare il valore quadratico medio del segnale, cioè, in pratica elevare i termini alla seconda.

Calcolando P_o il segnale utile in uscita da $y(t)$ abbiamo:

$$\begin{aligned}
 P_o &= \left\langle \left(\frac{A_c m(t)}{2} \right)^2 \right\rangle \\
 &= \left\langle \left(\frac{A_c}{2} \right)^2 \right\rangle \cdot \left\langle (m(t))^2 \right\rangle \\
 &= \frac{A_c^2}{4} \cdot P_m
 \end{aligned} \tag{3.15}$$

dove P_m è la potenza del segnale modulante.

Invece, la potenza in uscita è uguale a:

$$\begin{aligned}
P_{n_o} &= \left\langle \left(\frac{n_c(t)}{2} \right)^2 \right\rangle \\
&= \frac{1}{4} \cdot \left\langle n_c(t)^2 \right\rangle \\
&= \frac{1}{4} \cdot P_n
\end{aligned} \tag{3.16}$$

dove P_n è la potenza del rumore.

Sapendo che la densità spettrale di rumore termico bilatera è uguale a $\frac{N_0}{2}$, possiamo calcolarci la potenza del rumore P_n :

$$\begin{aligned}
P_n &= \int_{-f_c-W}^{-f_c+W} \frac{N_0}{2} df + \int_{f_c-W}^{f_c+W} \frac{N_0}{2} df \\
&= \frac{N_0}{2} \int_{-f_c-W}^{-f_c+W} df + \frac{N_0}{2} \int_{f_c-W}^{f_c+W} df \\
&= \frac{N_0}{2} \cdot 2W + \frac{N_0}{2} \cdot 2W \\
&= N_0 \cdot W + N_0 \cdot W \\
&= 2 \cdot N_0 \cdot W \\
&= 2N_0 W
\end{aligned} \tag{3.17}$$

in cui, ricordo che, W è la banda del segnale.

Ora, sapendo quanto vale la potenza del segnale utile P_0 e la potenza del rumore P_{n_o} dopo la demodulazione coerente, possiamo calcolarci il rapporto segnale-rumore in uscita $(\frac{S}{N})_o$:

$$\begin{aligned}
\left(\frac{S}{N} \right)_o &= \frac{P_o}{P_{n_o}} \\
&= \frac{\frac{A_c^2}{4} \cdot P_m}{\frac{1}{4} \cdot P_n} \\
&= \frac{A_c^2 \cdot P_m}{2 \cdot N_o \cdot W}
\end{aligned} \tag{3.18}$$

Invece adesso calcoliamo la potenza del segnale modulato sommato al rumore $n(t)$, cioè $r(t)$:

$$r(t) = A_c m(t) \cos(2\pi f_c t) + n(t) \tag{3.19}$$

Calcolando da $r(t)$ la potenza del segnale, abbiamo:

$$\begin{aligned}
P_r &= \left\langle [A_c m(t) \cos(2\pi f_c t)]^2 \right\rangle \\
&= \dots \\
&= \frac{A_c^2}{2} \cdot P_m
\end{aligned} \tag{3.20}$$

Ho messo i puntini perché ci dovrebbero essere dei passaggi algebrici in cui sicuramente si dovrebbero applicare delle formule trigonometriche.

Lascio a voi, futuri editor, di scrivere i passaggi step-by-step: per adesso credeteci.

P_n cioè la potenza del rumore in banda base è calcolata nell'intervallo $[f_c - W, f_c + W]$, quindi anche in banda base:

$$\begin{aligned} P_n &= \int_{f_c-W}^{f_c+W} \frac{N_0}{2} df \\ &= \frac{N_0}{2} \int_{f_c-W}^{f_c+W} df \\ &= \frac{N_0}{2} \cdot 2W \\ &= N_0 \cdot W \end{aligned} \quad (3.21)$$

Quindi, in banda base, il $\left(\frac{S}{N}\right)_b$ è uguale a:

$$\begin{aligned} \left(\frac{S}{N}\right)_b &= \frac{P_r}{P_n} \\ &= \frac{\frac{A_c^2}{2} \cdot P_m}{N_0 \cdot W} \\ &= \frac{A_c^2 \cdot P_m}{2 \cdot N_0 \cdot W} \end{aligned} \quad (3.22)$$

Quindi il rapporto segnale-rumore in banda base è uguale del rapporto segnale-rumore in modulazione.

Scriviamo in formulo con l'indica A, cioè il coefficiente di qualità della modulazione:

$$\begin{aligned} \left(\frac{S}{N}\right)_o &= \left(\frac{S}{N}\right)_b \\ &= 1 \cdot \left(\frac{S}{N}\right)_b \\ &= A \cdot \left(\frac{S}{N}\right)_b \end{aligned} \quad (3.23)$$

Come si può vedere dalla precedente uguaglianza, $A = 1$, quindi non si guadagna né si perde nulla con la modulazione.

Se devo per forza trasmettere in DSB-SC sappiamo che non ci perdiamo nulla in qualità rispetto alla banda base.

Ora che abbiamo capito come si calcola analiticamente il rapporto segnale rumore in banda base e in modulazione con l'aggiunta di rumore, nelle prossime sezioni non verranno elencati tutti i passaggi matematici, ma verrà data la formula così come è.

Sarà, come aveva detto il prof a lezione, una "lista della spesa" e di credi, perché la trattazione matematica sarebbe molto lunga.

3.3 Qualità della SSB-SC

³

Dato il segnale ricevuto in modulazione dalla SSB-SC $r(t)$ con l'aggiunta di rumore $n(t)$:

$$r(t) = A_c m(t) \cos(2\pi f_c t) \mp A_c \tilde{m}(t) \sin(2\pi f_c t) + n_s(t) \cos(2\pi f_c t) - n_o(t) \sin(2\pi f_c t) \quad (3.24)$$

sapendo che $\tilde{m}(t)$ è la trasformata di Hilbert di $m(t)$, avremo come il caso precedente, cioè:

$$\left(\frac{S}{N} \right)_o = \left(\frac{S}{N} \right)_b \quad (3.25)$$

cioè $A = 1$.

Quindi non si guadagna né si perde nulla con la modulazione.

Se devo per forza trasmettere in SSB-SC sappiamo che non ci perdiamo nulla in qualità rispetto alla banda base.

³Slide del prof — Qualità nelle modulazioni analogiche — pag 3.2
Appunti di Damiano — pag 3.2
Slide — Qualità nelle modulazioni analogiche — pag 3.2
Appunti — 2025-03-10 — pag 8

3.4 Qualità della AM convenzionale

⁴

Come scritto nel capitolo della AM convenzionale, possiamo de-modulare con uno tra i due metodi:

- con la demodulazione coerente, quindi moltiplicando il segnale modulante con $\cos(2\pi f_c t)$
- con la demodulazione di inviluppo, che viene scelta non per la qualità ma perché costa poco economicamente realizzare un demodulatore

⁴Slide del prof — Qualità nelle modulazioni analogiche — pag 4.1
Appunti di Damiano — pag 4.1
Slide — Qualità nelle modulazioni analogiche — pag 4.1
Appunti — 2025-03-11 — pag 2 - 4

3.4.1 Qualità della AM convenzionale con demodulazione coerente

5

Consideriamo per l'AM convenzionale una demodulazione coerente.

$r(t)$, cioè il segnale modulante in modo coerente e con l'aggiunta di rumore $r(t)$, sarà uguale a:

$$r(t) = A_c [1 + a \cdot m_n(t)] \cos(2\pi f_c t) + n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t) \quad (3.26)$$

dove:

$$\begin{cases} m_n(t) = \frac{m(t)}{\max |m(t)|} \\ 0 < \alpha \leq 1 \end{cases} \quad (3.27)$$

Ora confrontiamo il rapporto segnale rumore tra la banda base e dopo la modulazione:

$$\begin{aligned} \left(\frac{S}{N}\right)_o &= \frac{a^2 \cdot P_{m_n}}{1 + a^2 P_{m_n}} \left(\frac{S}{N}\right)_b \\ &= \eta \left(\frac{S}{N}\right)_b \end{aligned} \quad (3.28)$$

dove:

- P_{m_n} è la potenza di $m_n(t)$
- η è l'efficienza di modulazione

Qui non abbiamo A come efficienza di modulazione ma è η : è solo una notazione differente

Sapendo che a^2 e P_{m_n} sono sempre positivi, sappiamo che:

$$\eta = \frac{a^2 \cdot P_{m_n}}{1 + a^2 P_{m_n}} < 1 \quad (3.29)$$

Essendo η minore di 1, considerando solo ed esclusivamente la qualità, non conviene modulare in modo coerente in AM convenzionale.

⁵Slide del prof — Qualità nelle modulazioni analogiche — pag 4.1
 Appunti di Damiano — pag 4.1
 Slide — Qualità nelle modulazioni analogiche — pag 4.1
 Appunti — 2025-03-11 — pag 2 - 4

3.4.2 Qualità della AM convenzionale con demodulazione d'inviluppo

6

Consideriamo per l'AM convenzionale una demodulazione d'inviluppo.

$r(t)$, cioè il segnale modulante in modo coerente e con l'aggiunta di rumore $r(t)$, sarà uguale a:

$$r(t) = A_c [1 + a \cdot m_n(t)] \cos(2\pi f_c t) + n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t) \quad (3.30)$$

Abbiamo che il rapporto segnale-rumore in ingresso al demodulatore d'inviluppo $(\frac{S}{N})_i$:

$$\left(\frac{S}{N}\right)_i = \frac{A_c^2 \cdot (1 + a^2 \cdot P_{m_n})}{4 \cdot N_0 \cdot W} \quad (3.31)$$

Confrontando il rapporto segnale-rumore in ingresso al demodulatore d'inviluppo $(\frac{S}{N})_i$ e il rapporto segnale-rumore in banda base, abbiamo la seguente relazione:

$$\left(\frac{S}{N}\right)_i = \frac{1}{2} \cdot \left(\frac{S}{N}\right)_b \quad (3.32)$$

State attenti ai pedici di $(\frac{S}{N})$:

- se c'è la lettera i , cioè $(\frac{S}{N})_i$, si intende il rapporto segnale-rumore in ingresso al demodulatore
- se c'è la lettera b , cioè $(\frac{S}{N})_b$, si intende il rapporto segnale-rumore in uscita dal demodulatore, in questo caso di inviluppo

Possiamo distinguere due casi:

- $(\frac{S}{N})_i \gg 1$
- $(\frac{S}{N})_i < 1$

Nel primo caso, cioè $(\frac{S}{N})_i \gg 1$, abbiamo che:

$$\begin{aligned} \left(\frac{S}{N}\right)_o &= \eta \cdot \left(\frac{S}{N}\right)_b \\ &= \frac{a^2 \cdot P_{m_n}}{1 + a^2 P_{m_n}} \cdot \left(\frac{S}{N}\right)_b \end{aligned} \quad (3.33)$$

Nel secondo caso, cioè $(\frac{S}{N})_i < 1$, cioè un $(\frac{S}{N})_i$ basso, le prestazioni degradano rispetto al caso precedente.

Schwart, Bennet e Stein hanno dimostrato che, se $(\frac{S}{N})_i < 1$, allora in uscita ad un demodulatore di inviluppo avremo il seguente rapporto segnale-rumore:

⁶Slide del prof — Qualità nelle modulazioni analogiche — pag 4.2 - 5.1
Appunti di Damiano — pag 4.2 - 5.1
Slide — Qualità nelle modulazioni analogiche — pag 4.2 - 5.1
Appunti — 2025-03-11 — pag 4 - 9

$$\begin{aligned} \left(\frac{S}{N}\right)_o &\approx \left[\left(\frac{S}{N}\right)_i\right]^2 \\ &= \left[\frac{A_c^2 \cdot (1 + a^2 \cdot P_{m_n})}{4 \cdot N_0 \cdot W}\right]^2 \end{aligned} \quad (3.34)$$

Negli esercizi questi calcoli svolti riguardo alla qualità sono indicati in dB.

Quindi ricordiamo al volo come convertire il numero non in dB a dB.

Se abbiamo a che fare con delle potenze:

$$\left(\frac{S}{N}\right)_{dB} = 10 \cdot \log_{10} \left(\frac{S}{N}\right) \quad (3.35)$$

Per maggiori approfondimenti rimando alle conversione e alle proprietà delle unità logaritmiche:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 10 - Unità logaritmiche - pag 97

3.5 PLL per il recupero della portante

7

Siccome nel recupero della portante è necessario sapere il ϕ_c della portante, si può utilizzare, come scritto precedente, il PLL.

Considerando il segnale ricevuto modulato in DSB-SC $u(t)$ con l'aggiunta del rumore $n(t)$, cioè $r(t)$:

$$\begin{aligned} r(t) &= u(t) + n(t) \\ &= A_c m(t) \cos(2\pi f_c t + \phi_c) + n(t) \end{aligned} \quad (3.36)$$

Se consideriamo $r^2(t)$:

$$\begin{aligned} r^2(t) &= [u(t) + n(t)]^2 \\ &= \frac{1}{2} A_c^2 + \frac{1}{2} A_c^2 m^2(t) \cos(4\pi f_c t + 2\phi_c) + \text{contributi di rumore} \end{aligned} \quad (3.37)$$

Possiamo implementare il PLL nella DSB-SC in questa posizionandolo in questa maniera:

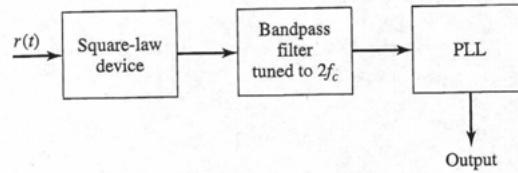


Figure 5.1 System for carrier-phase estimation.

Gli elementi basilare di un PLL sono i seguenti:

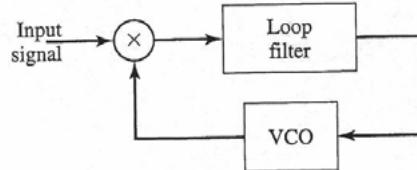


Figure 5.2 Basic elements of a PLL.

Quindi possiamo vedere l'esploso dello schema precedente (non quello del PLL), come:

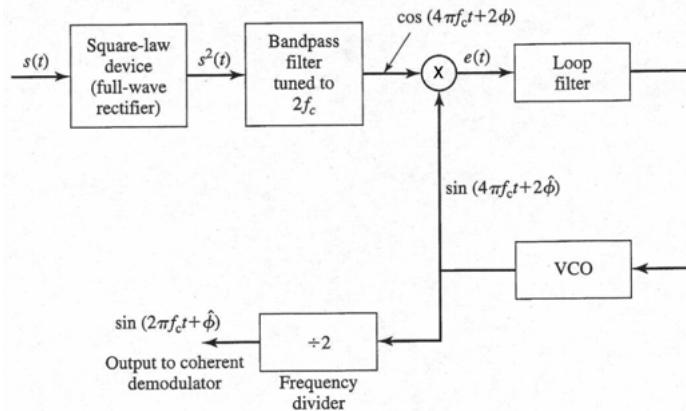


Figure 5.9 Carrier recovery using a square-law device.

⁷Slide del prof — Qualità nelle modulazioni analogiche — pag 5.2 - 7.1
 Appunti di Damiano — pag 5.2 - 7.1
 Slide — Qualità nelle modulazioni analogiche — pag 5.2 - 7.1
 Appunti — 2025-03-11 — pag 9 - 10

Oppure possiamo utilizzare questo tipo di architettura:

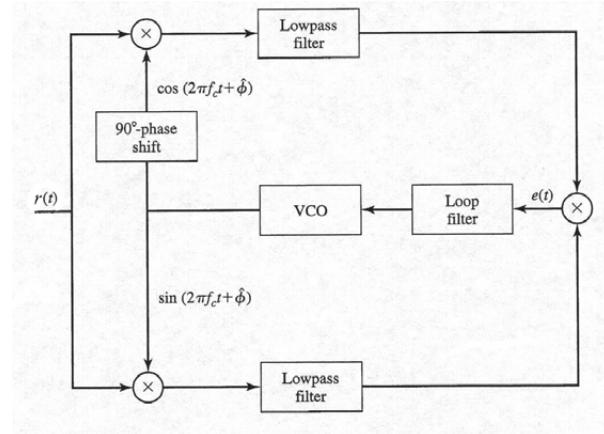


Figure 5.10 Block diagram of Costas loop.

3.6 Effetto del rumore nelle modulazioni angolari

8

Come rappresentato dai seguenti andamenti delle funzioni con modulazione angolare sovrapposta a del rumore:

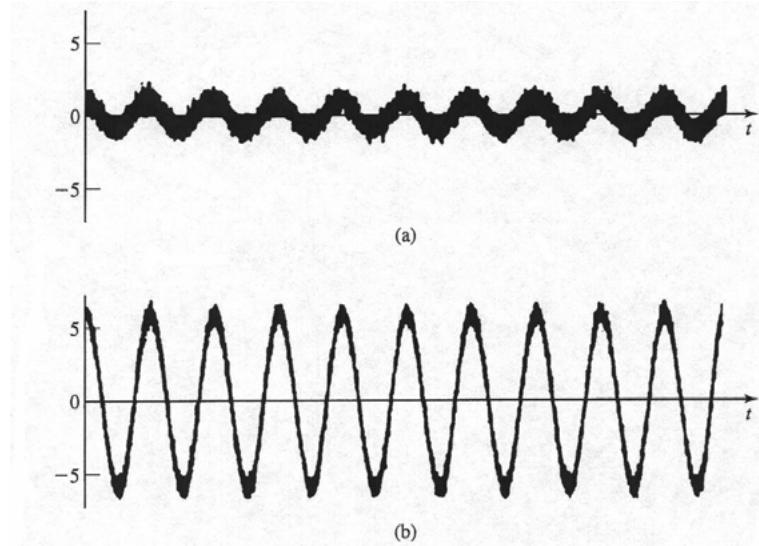


Figure 5.11 Effect of noise on the zero crossings of (a) low-power and (b) high-power modulated signals.

in particolare in alto un segnale a bassa potenza ed in basso un segnale ad alta potenza.

Il rumore cambia l'ampiezza, ma, nelle modulazioni angolari, l'informazione sta nei passaggi sullo zero.

Considerando dal punto di vista analitico il segnale modulato angolare $u(t)$ con l'aggiunta del rumore $n(t)$, abbiamo il segnale $r(t)$ e distinguendo per la PM e la FM:

$$\begin{aligned} r(t) &= u(t) + n(t) \\ &= A_c \cos(2\pi f_c t + \phi(t)) \\ &= \begin{cases} A_c \cos(2\pi f_c t + k_p m(t)) + n(t) & \text{PM} \\ A_c \cos \left(2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau \right) + n(t) & \text{FM} \end{cases} \quad (3.38) \end{aligned}$$

⁸Slide del prof — Qualità nelle modulazioni analogiche — pag 7.2
 Appunti di Damiano — pag 7.2 - 8.2
 Slide — Qualità nelle modulazioni analogiche — pag 7.2 - 8.2
 Appunti — 2025-03-11 — pag 10 - 12
 Appunti — 2025-07-08 Ricevimento — pag 8.1

Utilizzando la rappresentazione dei fasori, possiamo visualizzare il contributo del rumore $\phi_n(t)$ sulla fase del segnale $\phi(t)$:

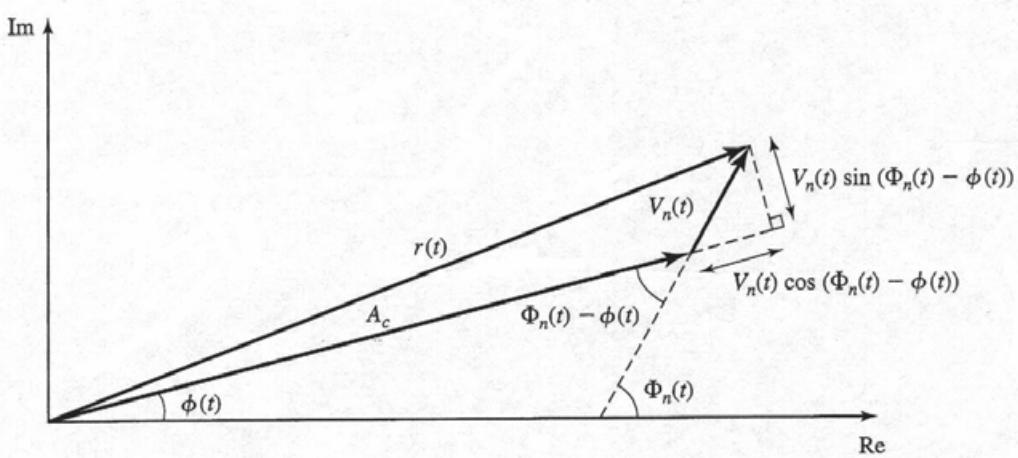


Figure 5.13 Phasor diagram of signal and noise in an angle-modulated system.

Quindi il vettore $r(t)$, senza rumore, dovrebbe coincidere con il vettore A_c , ma, a causa del vettore rumore $V_n(t)$, $r(t)$ è uguale a:

$$\overrightarrow{r(t)} = \overrightarrow{A_c} + \overrightarrow{V_n(t)} \quad (3.39)$$

quindi, il ricevitore, piuttosto che visualizzare $\phi(t)$ visualizza $\phi(t) + \phi_n(t)$, cioè $\phi_t(t)$:

$$\phi_t(t) = \phi(t) + \phi_n(t) \quad (3.40)$$

Facendo i calcoli (che sono omessi), la differenza tra quello che dovremmo ricevere in modulazione angolare $\phi(t)$ e quello che riceviamo realmente con il rumore, avremo $\gamma(t)$:

$$\begin{aligned} \gamma(t) &= \phi_n(t) - \phi(t) \\ &\arctan \left[\frac{V_n \sin(\phi_n(t) - \phi(t))}{A_c + V_n(t) \cos(\phi_n(t) - \phi(t))} \right] \end{aligned} \quad (3.41)$$

Non è necessario impararsi a memoria la formula di $\gamma(t)$.

Basta sapere che il rumore con la sua fase $\phi_n(t)$ fa variare la fase in ricezione $\phi(t)$ e l'idea è quella di misurare la differenza tra le due fasi.

Idealmente vorremmo avere $\phi_n(t)$ nullo, ma ciò, purtroppo, non capita

3.6.1 Demodulazione di segnali PM e FM

9

Considerando il caso in cui il rapporto segnale-rumore in ricezione elevato:

$$\Pr\{V_n(t) \ll A_c^2\} \approx 1 \quad (3.42)$$

Essendo il rumore un processo stocastico, quindi probabilistico, formalmente non possiamo dire che si ha un rumore basso rispetto al segnale modulato.

Formalmente non possiamo scrivere:

$$V_n(t) \ll A_c^2 \quad (3.43)$$

anche se di fatto si intende ciò con la formula di probabilità $\Pr\{V_n(t) \ll A_c^2\} \approx 1$.

Considerando questo tipo di sistema:

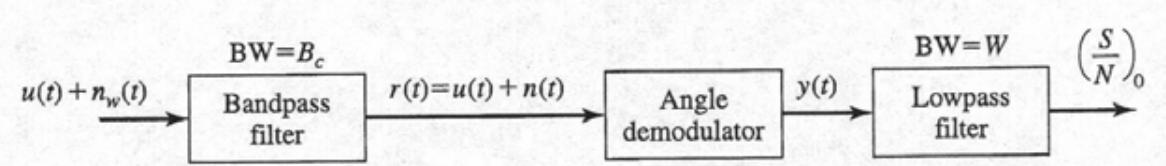


Figure 5.12 Block diagram of receiver for a general angle-demodulated signal.

il segnale che esce dal demodulatore angolare $y(t)$ è il seguente a seconda di PM e FM:

$$y(t) = \begin{cases} k_p m(t) + \frac{V_n(t)}{A_c} \sin(\phi_n(t) - \phi(t)) & \text{PM} \\ k_f m(t) + \frac{1}{2\pi} \frac{d}{dt} \left[\frac{V_n(t)}{A_c} \sin(\phi_n(t) - \phi(t)) \right] & \text{FM} \end{cases} \quad (3.44)$$

⁹Slide del prof — Qualità nelle modulazioni analogiche — pag 9.1 - 11.2
 Appunti di Damiano — pag 9.1 - 11.2
 Slide — Qualità nelle modulazioni analogiche — pag 9.1 - 12.2
 Appunti — 2025-03-11 — pag 13
 Appunti — 2025-07-08 Ricevimento — pag 4.2 - 7

Di seguito la densità spettrale di rumore all'ingresso del filtro:

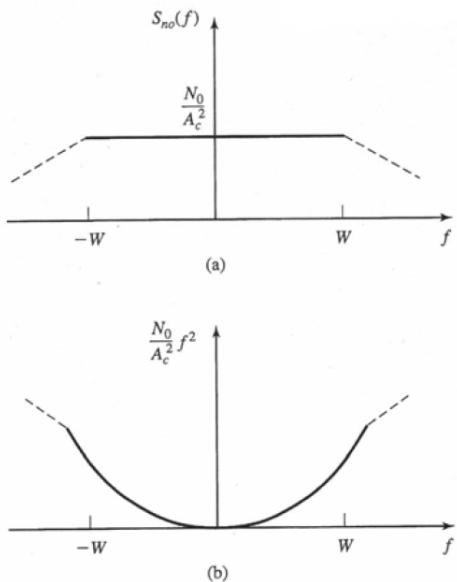


Figure 5.15 Noise power spectrum at demodulator output in (a) PM and (b) FM.

Questo grafico ci dimostra la densità spettrale del rumore in base alla modulazione utilizzata.

La figura in alto, quella indicata con (a), è la densità di rumore della PM, che rimane costante alle basse frequenze, cioè nell'intervallo tra $-W$ e $+W$.

Oltre quell'intervallo, la densità di rumore diminuisce.

Questo significa che la PM è consigliata per le alte frequenze perché si avrà meno rumore al ricevitore.

Invece, notando la seconda figura in basso, quella indicato con (b), è la densità di rumore della FM, che ha un effetto parabolico e poi diventa quasi lineare oltre all'intervallo $-W$ e $+W$.

Questo significa che la FM è consigliata per le basse frequenze perché si avrà meno rumore al ricevitore.

Per questo motivo, la FM è utilizzata nella radio commerciale (quella che ascolti tutti i giorni in macchina).

Nella realtà, grazie ai pregi e ai difetti delle due modulazioni angolari, si cercherà di fare una modulazione mista: una FM alle basse frequenze, una PM alle alte frequenze.

È possibile combinare le modulazioni angolari se abbiamo un sistema a banda molto larga; se la banda è molto stretta, quindi poco larga, si può utilizzare una tra le due.

Per questo ultimo motivo citato, ad esempio, la radio commerciale è a basse frequenze ed ha una banda molto stretta, quindi si utilizza solo la FM.

Quindi, dopo tutte queste considerazioni, possiamo esprimere le seguenti formule in caso di rapporto segnale-rumore all'ingresso del demodulatore sufficientemente elevato:

$$\left(\frac{S}{N}\right)_o = \begin{cases} \beta_p^2 P_{m_n} \left(\frac{S}{N}\right)_b & \text{PM} \\ 3\beta_f^2 P_{m_n} \left(\frac{S}{N}\right)_b & \text{FM} \end{cases} \quad (3.45)$$

Utilizzando l'effetto soglia, possiamo considerare questo grafico per confrontare il rapporto segnale-rumore scelto un b_f rispetto all'SNR della modulazione DSB e SSB:

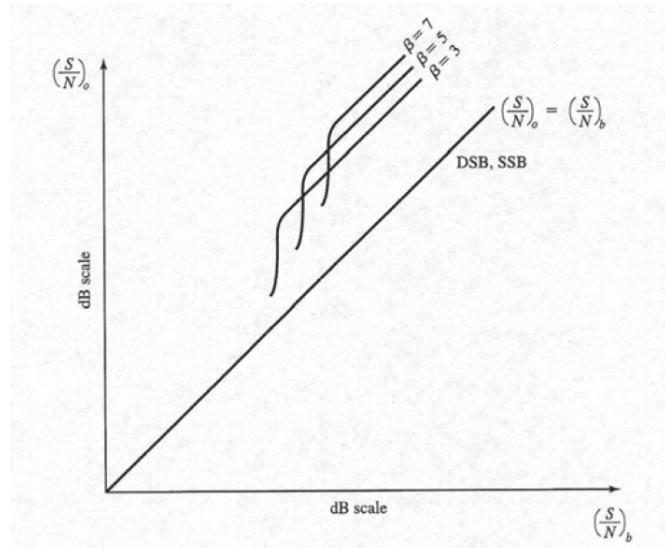


Figure 5.16 Output SNR vs. baseband SNR in an FM system for various values of β .

Utilizzando l'effetto soglia, in FM possiamo calcolarci il rapporto segnale-rumore:

$$\left(\frac{S}{N}\right)_{b,th} = 20(\beta + 1) \quad (3.46)$$

$(\frac{S}{N})_{b,th}$ c'è il pedice th, perché dall'inglese si intende "threshold", cioè soglia.

Dato un certo valore di β , il valore $(\frac{S}{N})_{b,th}$ è un valore approssimato della gobba della curva nella figura che si trova qui sopra.

Se abbiamo un $(\frac{S}{N})_b$ minore della soglia, cioè $(\frac{S}{N})_{b,th}$, non bisogna utilizzare la modulazione angolare perché il rapporto segnale-rumore in banda base sarà troppo piccolo, i.e. ci sarà più rumore rispetto al segnale utile e quest'ultimo non sarà ben reinterpretato dal ricevitore

Le caratteristiche del filtro per svolgere la pre-enfasi e la de-enfasi è la seguente:

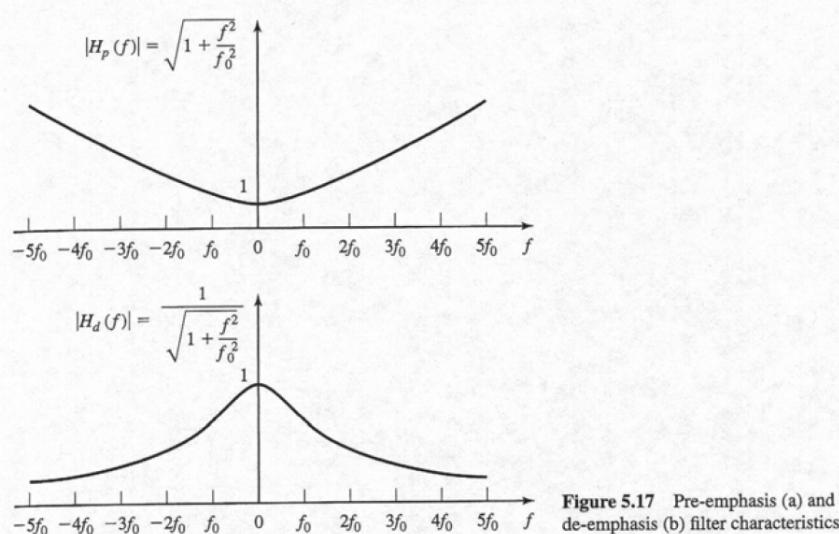
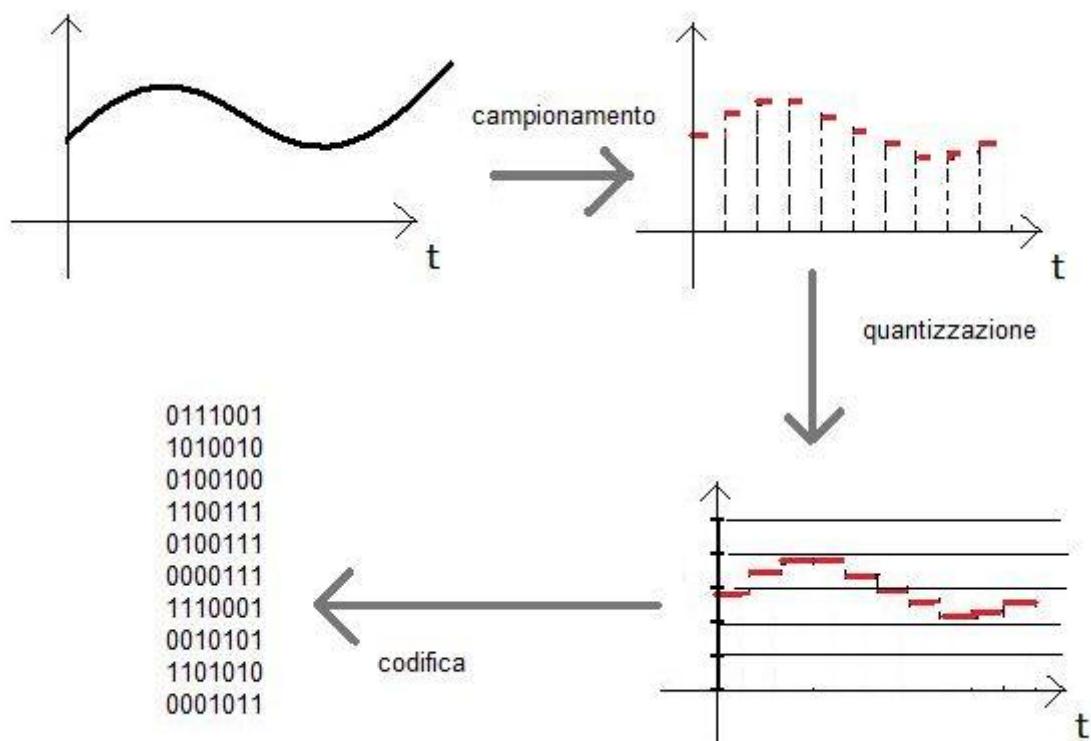


Figure 5.17 Pre-emphasis (a) and de-emphasis (b) filter characteristics.

Capitolo 4

Conversione AD



4.1 Quantizzazione

1

Data una funzione analogica, cioè una funzione con infiniti valori sia nel tempo che nelle ascisse, la si vuole discretizzare in modo da essere reinterpretata dai circuiti digitali.

Per discretizzarla sull'asse dei tempi, alla funzione deve essere applicato il campionamento.

Argomento ripreso da quasi tutti i corsi di ingegneria elettronica.

Da <https://github.com/ciccio25/appunti-misure-elettroniche>

Capitolo 7.4 Teorema fondamentale del campionamento — pag 112

Per il campionamento, è stato rivoluzionario il teorema del campionamento scritto dal matematico Claude.

Di seguito il teorema fondamentale del campionamento:

Theorem 1: If a function $f(t)$ contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced $1/2 W$ seconds apart.

Shannon pubblicò il suo teorema nella seguente rivista: "Claude E. Shannon Communication in the Presence of Noise Proceedings of the IRE, vol. 37, no. 1, pp. 10-21, Jan. 1949" consultabile al seguente link

<https://webusers.imj-prg.fr/~antoine.chambert-loir/enseignement/2020-21/shannon/shannon1949.pdf>

Visto che il teorema è stato pubblicato nel 1949, alcune terminologie sono diverse da quello al giorno d'oggi.

Le seguenti terminologie sono:

- cps è un'abbreviazione di cycles per second, nei giorni odierni utilizziamo l'indicazione di Hz
- per ordinates si intendono i valori istantanei

La traduzione italiana, con i seguenti adattamenti, del teorema fondamentale del campionamento è la seguente:

Se la funzione $f(t)$ contiene nessuna frequenza maggiore di W Hz, è completamente determinata dai suoi valori istantanei spaziati da una serie di punti spaziati di $\frac{1}{2W}$ secondi tra di loro.

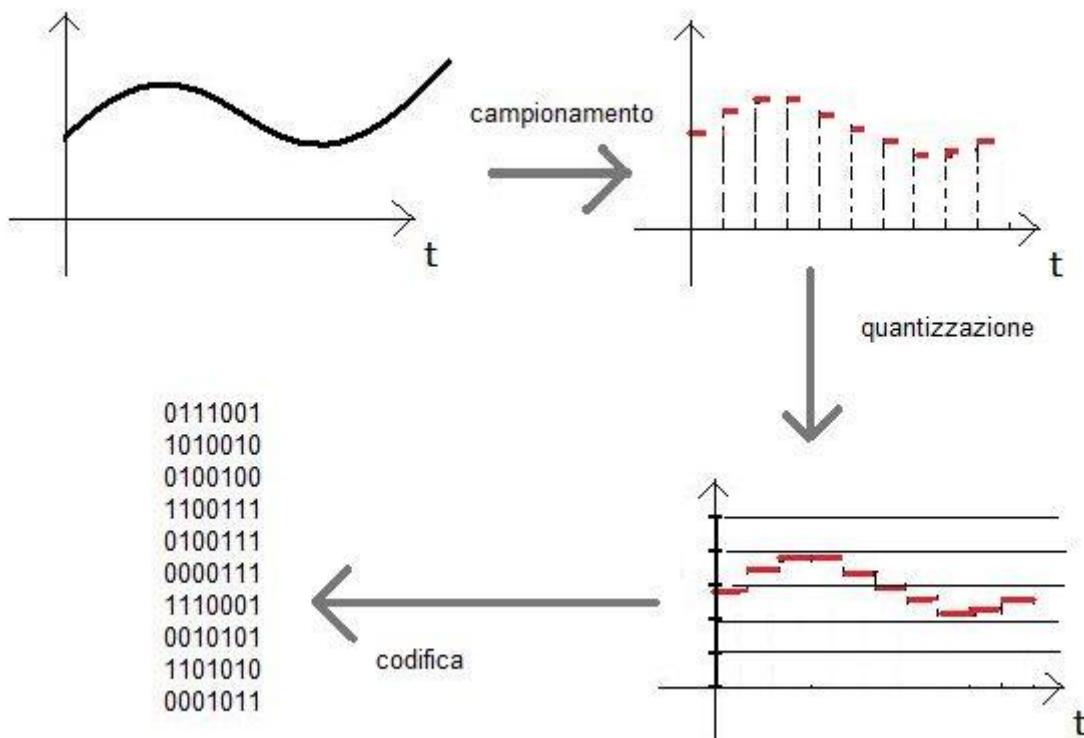
Questo teorema ci dice che, se gli istanti di campionamento sono opportunamente individuati, la funzione che si ottiene andando a prelevare il valore istantaneo del segnale solo in quegli istanti, è completamente determinata, quindi non si ha perdita di informazione.

Una volta che la funzione è campionata, quindi è discretizzata nell'asse dei tempi, la si può discretizzare nell'asse delle ascisse.

La discretizzazione nell'asse delle ascisse prende il nome di quantizzazione.

¹Slide del prof — Conversione AD — pag 1.1

Possiamo graficare i processi di discretizzazione svolti con questa figura:



Per chi come me ha svolto prima o ha seguito le lezioni del terzo anno di Misure Elettroniche della mitica prof.ssa Spinsante, alcune terminologie sono differenti, ma alla fine andiamo a finire negli stessi argomenti sotto l'occhio non di un misurista, bensì di un telecomunicazionista

La discretizzazione può essere di due tipi, in base a come si divide l'asse delle ordinate:

- quantizzazione scalare, ogni intervallo è costante
- quantizzazione non uniforme, ogni intervallo non è costante rispetto agli altri

Per tutta la quantizzazione, rimando al corso di Misure elettroniche dove è spiegato benissimo.

Da <https://github.com/ciccio25/appunti-misure-elettroniche>
Capitolo 7.8 Fase 2: Quantizzazione — pag 122 - 124

Nel corso di Misure Elettroniche, abbiamo affrontato solo la quantizzazione scalare.

4.2 Quantizzazione scalare

²

Data una funzione già campionata, possiamo applicare una quantizzazione scalare.

In questo tipo di quantizzazione, si associa un valore x a un valore \hat{x}_i , che a sua volta ricade in un valore nelle ordinate a_i .

Come si visualizza dalla seguente figura:

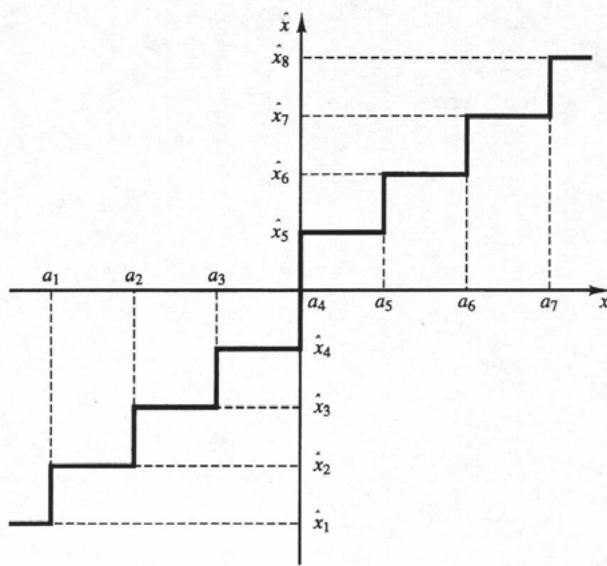


FIGURE 4.10. Example of an 8-level quantization scheme.

Consideriamo N il numero dei livelli di quantizzazione (in figura $N = 8$).

²Slide del prof — Conversione AD — pag 1.1
Slide — Conversione AD — pag 1.1

4.2.1 Funzione di quantizzazione Q

3

Consideriamo Q la funzione di quantizzazione:

$$Q(x) = \hat{x}_i \quad (4.1)$$

dove:

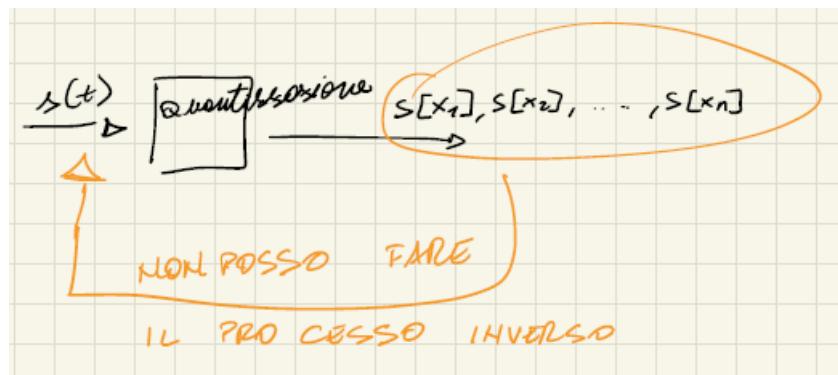
- x è il valore "analogico" della funzione campionata
- \hat{x}_i è il valore a cui è stato assegnato il valore analogico x

Per quantizzazione si intende associare un elemento sull'asse delle ascisse, ad un intervallo ben determinato.

La funzione di quantizzazione Q è una funzione non lineare e non invertibile, cioè, una volta svolta la quantizzazione, non si può ritornare al valore analogico campionato x.

In termini matematici, la funzione di quantizzazione Q è una funzione univoca e non biunivoca.

Lo si può spiegare meglio con il seguente grafico:



Quindi, con la quantizzazione, a differenza del campionamento, parte dell'informazione viene persa nel processo di quantizzazione e non è più recuperabile.

Proprio perché nella quantizzazione Q viene assegnato un valore analogico x ad un valore discreto x_i , avviene una distorsione, e si vuole misurare quest'ultima in media quadratica.

La misura delle distorsione in media quadratica sarà:

$$d(x, \hat{x}) = (x - Q(x))^2 = \tilde{x}^2 \quad (4.2)$$

dove:

- $d(x, \hat{x})$ è la distanza x analogica al valore assegnato discreto \hat{x}
- \tilde{x}^2 è la media quadratica tra x e la funzione di quantizzazione Q(x)

Siccome il valore analogico x, in realtà è una variabile aleatoria, non si scrive x piccolo bensì X maiuscolo.

Quindi, riscriviamo la formula precedente dal punto di vista aleatorio:

³Slide del prof — Conversione AD — pag 1.2 - 2.1
Slide — Conversione AD — pag 1.2 - 2.1

$$\begin{aligned}
 d(x, \hat{x}) &= (x - Q(x))^2 = \tilde{x}^2 \\
 &\downarrow \\
 D &= \mathbb{E} [d(X, \hat{X})]^2 = \mathbb{E} [(X - Q(x))^2]
 \end{aligned} \tag{4.3}$$

in cui con il simbolo \mathbb{E} , dall'inglese expected value o expectation, si intende la media della variabile aleatoria contenuta nelle parentesi quadre.

Avevamo incontrato la funzione \mathbb{E} già al corso di Teoria dei Segnali.

Da:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.10.1 - Valore medio - pag 131

Per una variabile aleatoria X con densità di probabilità $f_X(x)$, il valore medio di $f_X(x)$ è fornito dal seguente integrale:

$$m_X = \int_{-\infty}^{\infty} x \cdot f_X(x) dx \tag{4.4}$$

In alcuni casi, si indica il valore medio con la lettera, con la lettera E (dall'inglese Expectation), oppure con le parentesi $<>$:

$$\langle X \rangle = \mathbb{E}\{X\} = m_X \tag{4.5}$$

Il valore medio ha, essenzialmente, il significato di "baricentro" attorno al quale si distribuiscono i valori della variabile aleatoria X .

■

Per maggiori approfondimenti:

https://it.wikipedia.org/wiki/Valore_atteso

Come notazione, per quanto riguarda la statistica, si utilizzano le lettere maiuscole.

Quindi, se x è una variabile aleatoria, non si scrive probabilità di x , bensì probabilità di X

La quantizzazione scalare, a sua volta, può essere divisa in due tipi:

- quantizzazione silenziata, cioè il valore zero è un livello di quantizzazione
- quantizzazione non silenziata, cioè il valore zero non è un livello di quantizzazione

Ritornando alla quantizzazione scalare $N = 8$, possiamo confrontarla con una quantizzazione silenziata:

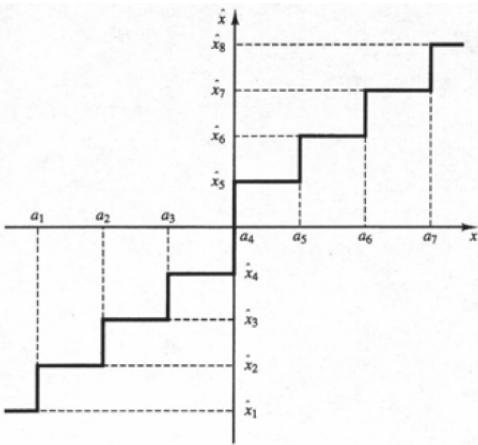


FIGURE 4.10. Example of an 8-level quantization scheme.

*N = 8, caratteristica
“non silenziata”*

*N = 7, caratteristica
“silenziata”*

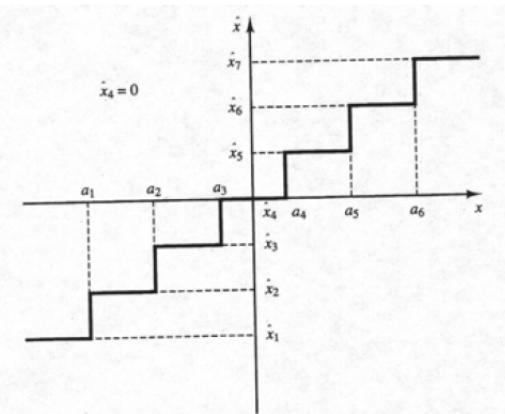


FIGURE 4.11. Seven-level uniform quantizer.

in cui si perderà un livello di quantizzazione (si passa da $N = 8$ a $N = 7$).

4.2.1.1 Rapporto segnale-rumore di quantizzazione SQNR

⁴

Considerando il rumore come qualsiasi elemento che causa incertezza, e che quindi fa variare il valore X in un altro, anche la quantizzazione genera rumore.

Quindi, come nel caso dei segnali studiati nei capitoli precedenti, anche qui possiamo esprimere un rapporto segnale-rumore, in particolare un rapporto segnale-rumore di quantizzazione (dall'inglese SQNR Signal Quantization Noise Ratio):

$$SQNR = \frac{\mathbb{E}[X^2]}{\mathbb{E}[(X - Q(X))^2]} = \frac{P_X}{P_{\hat{X}}} \quad (4.6)$$

dove:

- P_X è la potenza del segnale utile
- $P_{\hat{X}}$ è la potenza del rumore di quantizzazione

Si è potuto esprimere questa relazione perché facciamo l'ipotesi che la funzione X(t) sia stazionario.

Queste dispense assomigliano al corso di Analisi Matematica 2, dove ogni teorema è: "Dato il teorema a una variabile studiato ad Analisi Matematica 1 ...".

Ritornando a noi.

Per spiegare perché la potenza P_X si può esprimere così, si rimanda al corso di Teoria dei segnali.

Da:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 13.4 - Processo ergodico e il teorema di Wiener-Khintchine - pag 156

Nel caso di un processo ergodico, le medie di insieme (vale a dire le medie statistiche) coincidono con le medie temporali e possono essere calcolate a partire da un'unica e generica realizzazione del processo.

Per questo tipo di processi, possiamo enunciare il seguente teorema, il teorema di Wiener-Khintchine: se il processo x(t) è stazionario ed ergodico, almeno nella sua autocorrelazione, lo spettro di potenza del processo ad esso associato risulta la trasformata di Fourier della sua autocorrelazione statistica $R(\tau)$.

Sempre in virtù dell'ergodicità e delle proprietà ad esse associate, si può osservare che:

$$\begin{aligned} P &= \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} x^2(t) dt \\ &= \overline{R(0)} \\ &= R(0) \\ &= \langle x_1^2 \rangle \end{aligned} \quad (4.7)$$

⁴Slide del prof — Conversione AD — pag 2.2
Slide — Conversione AD — pag 2.2

4.2.2 Quantizzazione uniforme

5

Ritornando alla figura di una quantizzazione uniforme scalare:

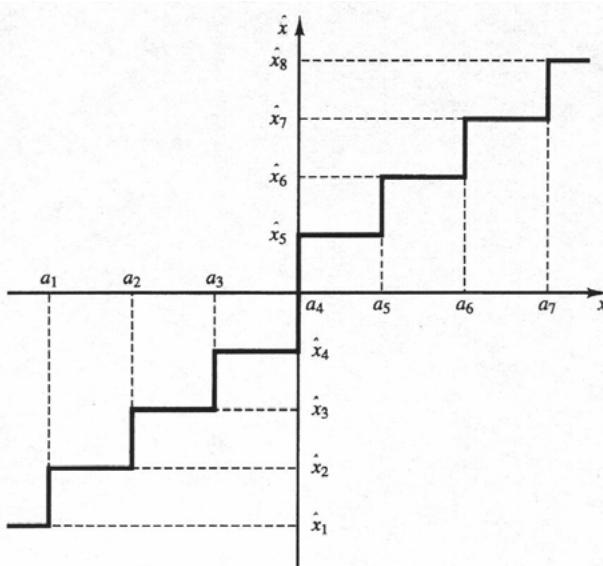


FIGURE 4.10. Example of an 8-level quantization scheme.

L'asse reale (valori della X) è partizionato in N regioni di quantizzazione.

Leggere il grafico è molto semplice.

Consideriamo x un valore da quantizzare e poniamolo nell'asse delle x.

Se x si trova in un intervallo, ad esempio si trova tra a_1 e a_2 , gli viene assegnato il relativo valore di \hat{x} , in questo esempio \hat{x}_2 .

Considerando un altro esempio, x si trova tra a_5 e a_6 , gli viene assegnato il valore quantizzato di \hat{x}_6 .

Ad eccezione delle regioni $(-\infty, a_1]$ e $(a_{N-1}, +\infty)$, l'ampiezza di ciascuna regione è pari a Δ .

Quindi, dopo tutte queste considerazioni, abbiamo $N + 2$ gradi di libertà, cioè:

- \hat{x}_i con $i = 1, 2, \dots, N$
- a_i
- Δ l'ampiezza dell'intervallo tra le due ampiezze contigue (cioè quelle attaccate)

L'obiettivo è quello di ottimizzare i gradi di libertà, in modo che la distorsione sia minima.

Teoricamente, per avere distorsione di quantizzazione nulla, dovremmo tendere a infinito i livelli di quantizzazione N, in modo tale che l'intervallo Δ tenda a zero, quindi il segnale analogico e quello quantizzato coincidono.

Ma, questo non si può fare perché, come abbiamo studiato a Teoria dei segnali, se abbiamo infiniti valori nel tempo, ciò richiederebbe infinita banda in frequenza, il quale non è realizzabile nei dispositivi reali.

⁵Slide del prof — Conversione AD — pag 3.1
Slide — Conversione AD — pag 3.1
Appunti — 2025-03-17 — pag 2 - 3

Da:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 3.4 - Altre caratteristiche sulla dualità tempo-frequenza - pag 24

N livelli di quantizzazioni di solito non è un grado di libertà perché N ha una relazione con la banda del segnale (si spiegherà questo concetto a breve).

4.2.2.1 Criterio a minima distanza

[6](#)

In questa sotto-sezione, non vi dovete ricordare le formule matematica passo passo. Ricordatevi, piuttosto quello che scrivo per esteso, anche se è cosa buona e giusta ricordarsi le formule matematiche, specialmente all'esame orale

Ritornando all'equazione della distanza D di X con il valore quantizzato \hat{X} :

$$D = \mathbb{E} [d(X, \hat{X})]^2 = \mathbb{E} [(X - Q(x))^2] \quad (4.8)$$

Possiamo espandere la formula per tutti gli intervalli di quantizzazione, in questo caso in una quantizzazione uniforme.

$$\begin{aligned} D &= \mathbb{E} [d(X, \hat{X})]^2 \\ &= \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 \cdot f_X(x) dx \\ &\quad + \int_{a_1}^{a_2} (x - \hat{x}_2)^2 \cdot f_X(x) dx + \dots + \int_{a_{N-2}}^{a_{N-1}} (x - \hat{x}_{N-1})^2 \cdot f_X(x) dx \\ &\quad + \int_{a_{N-1}}^{+\infty} (x - \hat{x}_N)^2 \cdot f_X(x) dx \end{aligned} \quad (4.9)$$

$$\begin{aligned} &= \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 \cdot f_X(x) dx \\ &\quad + \sum_{i=1}^{N-2} \left[\int_{a_i}^{a_{i+1}} (x - \hat{x}_2)^2 \cdot f_X(x) dx \right] \\ &\quad + \int_{a_{N-1}}^{+\infty} (x - \hat{x}_N)^2 \cdot f_X(x) dx \end{aligned}$$

dove $f_X(x)$ è la densità di probabilità della variabile aleatoria X.

Cioè, per calcolare la distanza D tra X e i valori centrali \hat{X} , si calcola la varianza di X in ogni intervallo di quantizzazione.

Siccome il nostro obiettivo è ottimizzare D in base all'intervalllo a_i , vogliamo ci sia una distanza minima tra X e il suo valore a_i .

Quindi possiamo scrivere:

$$\frac{\partial D}{\partial a_i} = 0 \quad (4.10)$$

e questa relazione è valida solo se a_i vale:

⁶Slide del prof — Conversione AD — pag 4.1
Slide — Conversione AD — pag 4.1
Appunti — 2025-03-17 — pag 3 - 4

$$\begin{aligned}
 a_i &= \frac{\Delta}{2} \\
 &= \frac{\hat{x}_i + \hat{x}_{i+1}}{2}
 \end{aligned} \tag{4.11}$$

Ad essere puntigliosi, in realtà, se volessimo calcolare il punto di minimo della distanza D, dovremmo calcolare la derivata seconda di D su a_i , ma il prof a lezione ha più volte ribadito che lo diamo per scontato che quel punto sia un minimo, così da fare meno calcoli matematici.

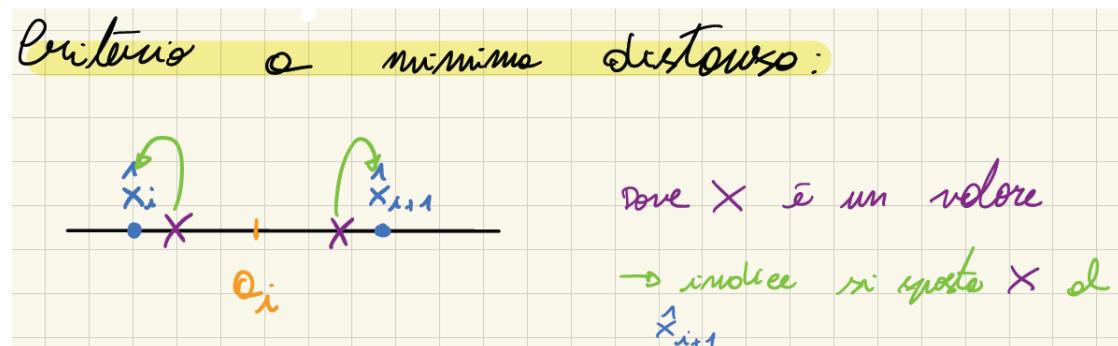
Se vuoi farti un bel ripasso da Analisi Matematica 1 e il teorema che dimostra come calcolare i punti di massimi e di minimi in una funzione:

Il teorema di Weierstrass, dei valori intermedi e di esistenza degli zeri - by Schooltoonchannel
https://youtu.be/k0AIKTr07Fw?si=FNlg_2yLgPTgs6Xf

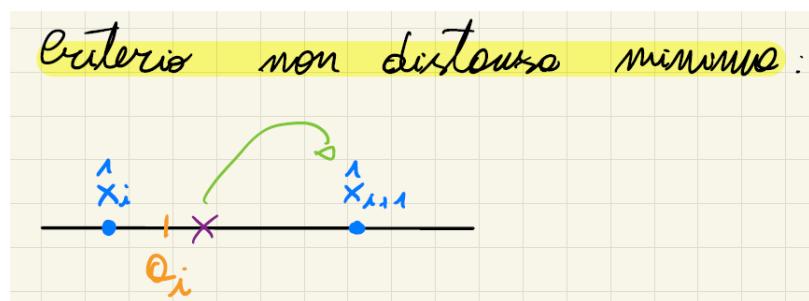
Quindi, i confini delle regioni di quantizzazione sono nel punto medio di livelli di quantizzazione adiacenti: ad ogni x si assegna il livello di quantizzazione più vicino.

Questo criterio prende il nome di criterio a minima distanza.

Da un punto di vista grafico, possiamo vederlo con questo esempio:



Invece se non è applicato il criterio a minima distanza, avremo questo caso:



4.2.3 Quantizzazione non uniforme

⁷

In questo caso, rispetto alla quantizzazione uniforme, l'ampiezza delle regioni in cui è partizionato l'asse delle X (segnaletico in ingresso al quantizzatore) è variabile.

Rispetto alla quantizzazione uniforme, abbiamo $2N - 1$ gradi di libertà, che sono:

- \hat{x}_i in cui $i = 1, 2, \dots, N$
- a_i in cui $i = 1, 2, \dots, N - 1$

L'obiettivo, anche nella quantizzazione non uniforme, è quello di ottimizzare i gradi di libertà, in modo che la distorsione sia minima.

Teoricamente, per avere distorsione di quantizzazione nulla, dovremmo tendere a infinito il numero dei livelli di quantizzazione N, in modo tale che l'intervallo Δ tenda a zero, quindi il segnale analogico e quello quantizzato coincidono.

Ma, questo non si può fare perché, come abbiamo studiato a Teoria dei segnali, se abbiamo infiniti valori nel tempo, ciò richiederebbe infinita banda in frequenza, il quale non è realizzabile nei dispositivi reali.

⁷Slide del prof — Conversione AD — pag 3.2
Slide — Conversione AD — pag 3.2
Appunti — 2025-03-17 — pag 3

4.2.4 Ottimizzazione della quantizzazione non uniforme

⁸

Come il caso della quantizzazione uniforme, svolgiamo gli stessi procedimenti matematici.

Quindi dal campionatore, si avranno dei valori reali che poi saranno quantizzati.

Ritornando all'equazione della distanza D di X con il valore quantizzato \hat{X} :

$$D = \mathbb{E} [d(X, \hat{X})]^2 = \mathbb{E} [(X - Q(x))^2] \quad (4.12)$$

Possiamo espandere la formula per tutti gli intervalli di quantizzazione, in questo caso in una quantizzazione uniforme.

$$\begin{aligned} D &= \mathbb{E} [d(X, \hat{X})]^2 \\ &= \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 \cdot f_X(x) dx \\ &\quad + \int_{a_1}^{a_2} (x - \hat{x}_2)^2 \cdot f_X(x) dx + \dots + \int_{a_{N-2}}^{a_{N-1}} (x - \hat{x}_{N-1})^2 \cdot f_X(x) dx \\ &\quad + \int_{a_{N-1}}^{+\infty} (x - \hat{x}_N)^2 \cdot f_X(x) dx \\ \\ &= \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 \cdot f_X(x) dx \\ &\quad + \sum_{i=1}^{N-2} \left[\int_{a_i}^{a_{i+1}} (x - \hat{x}_i)^2 \cdot f_X(x) dx \right] \\ &\quad + \int_{a_{N-1}}^{+\infty} (x - \hat{x}_N)^2 \cdot f_X(x) dx \end{aligned} \quad (4.13)$$

dove $f_X(x)$ è la densità di probabilità della variabile aleatoria X.

Considerando la derivata di \hat{x}_i sulla distanza D:

$$\frac{\partial D}{\partial \hat{x}_i} = - \int_{a_{i-1}}^{a_i} 2(x - \hat{x}_i) \cdot f_X(x) dx \quad (4.14)$$

Poniamo:

$$\frac{\partial D}{\partial \hat{x}_i} = 0 \quad (4.15)$$

perché vogliamo massimizzare i \hat{x}_i , che è un nostro grado di libertà nella quantizzazione non uniforme.

⁸Slide del prof — Conversione AD — pag 4.2
Slide — Conversione AD — pag 4.2
Appunti — 2025-03-17 — pag 4

Questa relazione è vera se:

$$\begin{aligned}
 \hat{x}_i &= \frac{\int_{a_{i-1}}^{a_i} x \cdot f_X(x) dx}{\int_{a_{i-1}}^{a_i} f_X(x) dx} \\
 &= \frac{\int_{a_{i-1}}^{a_i} x \cdot f_X(x) dx}{\Pr(a_{i-1} < X \leq a_i)} \\
 &= \int_{a_{i-1}}^{a_i} \frac{x \cdot f_X(x) dx}{\Pr(a_{i-1} < X \leq a_i)} \tag{4.16} \\
 &= \int_{a_{i-1}}^{a_i} x \cdot f_X(x | a_{i-1} < X \leq a_i) dx \\
 &= \mathbb{E}[x | a_{i-1} < X \leq a_i]
 \end{aligned}$$

Per gli indicatori statistici utilizzati in queste equazioni, si rimanda al capitolo della probabilità studiato a Teoria dei Segnali.

Da:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.7 - Distribuzione di probabilità cumulativa e densità di probabilità - pag 121 - 123
 Capitolo 12.13 - Densità di probabilità condizionata - pag 139

Grazie a queste formule matematiche, abbiamo dimostrato che, se vogliamo che $\frac{\partial D}{\partial x_i} = 0$, il livello di quantizzazione della i-esima regione corrisponde non al centro della regione, come nella quantizzazione uniforme, bensì al centroide di quella regione.

4.3 Condizioni di Lloyd-Max

9

Ho trovato una spiegazione migliore su wikipedia:
https://it.wikipedia.org/wiki/Algoritmo_di_Lloyd

Sapendo che:

$$a_i = \frac{\hat{x}_i + \hat{x}_{i+1}}{2} \quad (4.17)$$

$$\hat{x}_i = \mathbb{E}[x | a_{i-1} < X \leq a_i] \quad (4.18)$$

queste sono le due condizioni di Lloyd-Max che servono per calcolare il baricentro da assegnare al quantizzatore.

È una procedura iterativa che segue questi step:

1. Si parte con un insieme (arbitrario) di regioni di quantizzazione e si calcolano gli \hat{x}_i usando la seconda condizione $\hat{x}_i = \mathbb{E}[x | a_{i-1} < X \leq a_i]$
2. Si usano i livelli di quantizzazione calcolati al punto 1 per determinare nuove regioni di quantizzazione usando la prima condizione $a_i = \frac{\hat{x}_i + \hat{x}_{i+1}}{2}$
3. Si usano le nuove regioni di quantizzazione nella seconda condizione cioè $\hat{x}_i = \mathbb{E}[x | a_{i-1} < X \leq a_i]$
4. Si ripetono alternativamente le condizioni 1, cioè a_i , e la seconda condizione 2, cioè \hat{x}_i

Possiamo vedere la quantizzazione scalare come una quantizzazione vettoriale con $n = 1$, i.e. la quantizzazione scalare è una condizione particolare della quantizzazione vettoriale

4.3.0.1 Quantizzazione uniforme e non uniforme utilizzando le tabelle

[10](#)

Piuttosto che applicare ogni volta le condizioni di Llyod-Max, generalmente si utilizzano le seguenti tabelle.

Per una quantizzazione uniforme:

TABLE 4-2 OPTIMAL UNIFORM QUANTIZER FOR A GAUSSIAN SOURCE

No. output levels <i>N</i>	Output-level spacing Δ	Mean-squared error <i>D</i>	Informational entropy $H(\hat{x})$
1	—	1.000	0.0
2	1.596	0.3634	1.000
3	1.224	0.1902	1.536
4	0.9957	0.1188	1.904
5	0.8430	0.08218	2.183
6	0.7334	0.06065	2.409
7	0.6508	0.04686	2.598
8	0.5860	0.03744	2.761
9	0.5338	0.03069	2.904
10	0.4908	0.02568	3.032
11	0.4546	0.02185	3.148
12	0.4238	0.01885	3.253
13	0.3972	0.01645	3.350
14	0.3739	0.01450	3.440
15	0.3534	0.01289	3.524
16	0.3352	0.01154	3.602
17	0.3189	0.01040	3.676
18	0.3042	0.009430	3.746
19	0.2909	0.008594	3.811
20	0.2788	0.007869	3.874
21	0.2678	0.007235	3.933
22	0.2576	0.006678	3.990
23	0.2482	0.006185	4.045
24	0.2396	0.005747	4.097
25	0.2315	0.005355	4.146
26	0.2240	0.005004	4.194
27	0.2171	0.004687	4.241
28	0.2105	0.004401	4.285
29	0.2044	0.004141	4.328
30	0.1987	0.003905	4.370
31	0.1932	0.003688	4.410
32	0.1881	0.003490	4.449
33	0.1833	0.003308	4.487
34	0.1787	0.003141	4.524
35	0.1744	0.002986	4.560
36	0.1703	0.002843	4.594

Per una quantizzazione non uniforme:

TABLE 4-3 OPTIMAL NONUNIFORM QUANTIZER FOR A GAUSSIAN SOURCE

<i>N</i>	$\pm a_i$	$\pm \hat{x}_i$	<i>D</i>	$H(\hat{X})$
1	—	0	1	0
2	0	0.7989	0.3634	1
3	0.6120	0.1224	0.1902	1.536
4	0.9916	0.4528, 1.510	0.1175	1.911
5	0.3823, 1.244	0.07646, 1.724	0.07994	2.203
6	0.6589, 1.447	0.3177, 1.000, 1.894	0.05798	2.443
7	0.2803, 0.8744, 1.611	0.05606, 1.188, 2.033	0.04400	2.647
8	0.5006, 1.050, 1.748	0.2451, 0.7560, 1.344, 2.152	0.03454	2.825
9	0.2218, 0.6812, 1.198, 1.866	0.04436, 0.9188, 1.476, 2.255	0.02785	2.983
10	0, 0.4047, 0.8339, 1.325, 1.968	0.1996, 0.6099, 1.058, 1.591,	0.02293	3.125
		2.345		
11	0.1837, 0.5599, 0.9656, 1.436, 2.059	0.3675, 0.7524, 1.179, 1.693,	0.01922	3.253
12	0, 0.3401, 0.6943, 1.081, 1.534, 2.141	0.1684, 0.5119, 0.8768, 1.286,	0.01634	3.372
13	0.1569, 0.4760, 0.8126, 1.184, 1.623, 2.215	0.3138, 0.6383, 0.9870, 1.381,	0.01406	3.481
14	0, 0.2935, 0.5959, 0.9181, 1.277, 1.703, 2.282	0.1457, 0.4413, 0.7505, 1.086, 1.468, 1.939, 2.625	0.01223	3.582
15	0.1369, 0.4143, 0.7030, 1.013, 1.361, 1.776, 2.348	0.02739, 0.5548, 0.8512, 1.175,	0.01073	3.677
16	0, 0.2582, 0.5224, 0.7996, 1.099, 1.437, 1.844, 2.401	0.1284, 0.3881, 0.6568, 0.9424, 1.256, 1.618, 2.069, 2.733	0.009497	3.765
17	0.1215, 0.3670, 0.6201, 0.8875, 1.178, 1.508, 1.906, 2.454	0.02430, 0.4909, 0.7493, 1.026,	0.008463	3.849
18	0, 0.2306, 0.4653, 0.7091, 0.9680, 1.251, 1.573, 1.964, 2.504	0.1148, 0.3464, 0.5843, 0.8339, 1.102, 1.400, 1.746, 2.181, 2.826	0.007589	3.928
19	0.1092, 0.3294, 0.5551, 0.7908, 1.042, 1.318, 1.634, 2.018, 2.55	0, 0.2184, 0.4404, 0.6698, 0.9117,	0.006844	4.002
20	0, 0.2083, 0.4197, 0.6375, 0.8661, 1.111, 1.381, 1.690, 2.068, 2.594	0.1038, 0.3128, 0.5265, 0.7486, 0.9837, 1.239, 1.524, 1.857, 2.279, 2.908	0.006203	4.074
21	0.09918, 0.2989, 0.5027, 0.7137, 2.116, 2.635	0, 0.1984, 0.3994, 0.6059, 0.8215, 1.051, 1.300, 1.579, 1.908, 2.324, 2.946	0.005648	4.141
22	0, 0.1900, 0.3822, 0.5794, 0.7844, 1.001, 1.235, 1.495, 1.793, 2.160, 2.674	0.09469, 0.2852, 0.4793, 0.6795, 0.8893, 1.113, 1.357, 1.632, 1.935, 2.366, 2.982	0.005165	4.206
23	0.09085, 0.2736, 0.4594, 0.6507, 0.8504, 1.062, 1.291, 1.546, 1.841, 2.203, 2.711	0, 0.1817, 0.3654, 0.5534, 0.7481, 0.9327, 1.172, 1.411, 1.682, 2.000, 2.406, 3.016	0.004741	4.268
24	0, 0.1746, 0.3510, 0.5312, 0.7173, 0.9122, 1.119, 1.344, 1.595, 1.885, 2.243, 2.746	0.08708, 0.2621, 0.4399, 0.6224, 0.8122, 1.012, 1.227, 1.462, 1.728, 2.042, 2.444, 3.048	0.004367	4.327

¹⁰Slide del prof — Conversione AD — pag 5.2 - 7.1

Slide — Conversione AD — pag 5.2 - 7.1

Appunti — 2025-03-17 — pag 4 - 6

Oppure:

TABLE 4-3 CONTINUED

<i>N</i>	$\pm a_i$	$\pm \hat{x}_i$	<i>D</i>	$H(\hat{x})$
25	0.08381, 0.2522, 0.4231, 0.5982, 0.7797, 0.9702, 1.173, 1.394, 1.641, 1.927, 2.281, 2.779	0, 0.1676, 0.3368, 0.5093, 0.6870, 0.8723, 1.068, 1.279, 1.510, 1.772, 2.083, 2.480, 3.079	0.004036	4.384
26	0, 0.1616, 0.3245, 0.4905, 0.6610, 0.8383, 1.025, 1.224, 1.442, 1.685, 1.968, 2.318, 2.811	0.08060, 0.2425, 0.4066, 0.5743, 0.7477, 0.9289, 1.121, 1.328, 1.556, 1.814, 2.121, 2.514, 3.109	0.003741	4.439
27	0.07779, 0.2340, 0.3921, 0.5587, 0.7202, 0.8936, 1.077, 1.273, 1.487, 1.727, 2.006, 2.352, 2.842	0, 0.1556, 0.3124, 0.4719, 0.6354, 0.8049, 0.9824, 1.171, 1.374, 1.599, 1.854, 2.158, 2.547, 3.137	0.003477	4.491
28	0, 0.1503, 0.3018, 0.04556, 0.6132, 0.7760, 0.9460, 1.126, 1.319, 1.529, 1.766, 2.042, 2.385, 2.871	0.07502, 0.2256, 0.3780, 0.5333, 0.6930, 0.8589, 1.033, 1.118, 1.419, 1.640, 1.892, 2.193, 2.578, 3.164	0.003240	4.542

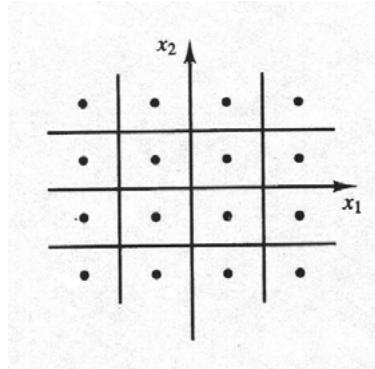
4.4 Quantizzazione vettoriale

¹¹

Rispetto alla quantizzazione scalare in cui un campione si quantizza uno alla volta, nella quantizzazione vettoriale si quantizzano n campioni alla volta, dove n campione vale:

$$n \geq 2 \quad (4.19)$$

Un esempio grafico con n campioni uguale a due:



Come scritto nella quantizzazione scalare, anche nella quantizzazione vettoriale si hanno dei gradi di libertà, tra cui, come si nota dalla figura, la forma della regione.

In questo caso, ogni regione ha una forma di un parallelepipedo.

La forma della regione, generalmente, cambia in base a come il progettista del quantizzatore voglia ridurre il rumore dovuto alla quantizzazione.

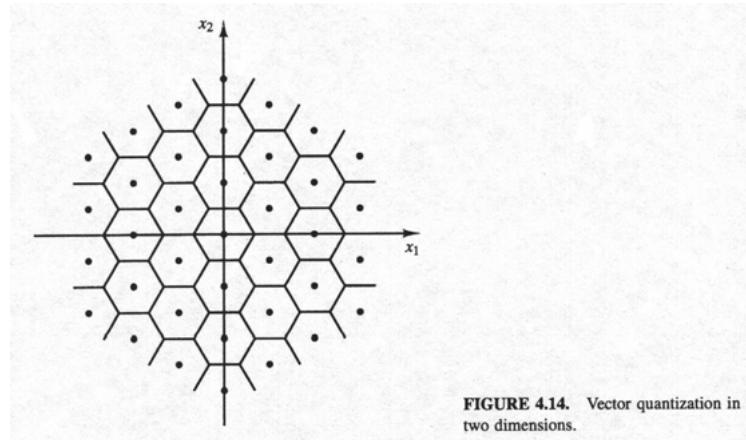
Una possibile implementazione per ridurre il rumore dovuto alla quantizzazione è quella di impiegare una quantizzazione vettoriale ottima.

¹¹Slide del prof — Conversione AD — pag 7.2
Slid — Conversione AD — pag 7.2
Appunti — 2025-03-17 — pag 6

4.4.1 Quantizzazione vettoriale ottima

¹²

Nella quantizzazione vettoriale ottima, come si vede dalla seguente figura:



si passa dalla regione rettangolare della quantizzazione vettoriale a una regione a nido d'api.

Si sceglie questa forma della regione per sottostare al teorema della distanza minima, cioè lo stesso studiato nella quantizzazione scalare uniforme.

Andando nello specifico della quantizzazione vettoriale ottima, la regione di quantizzazione i-esima è il luogo dei punti, nello spazio n dimensionale, che si trovano più vicini ad \hat{x}_i che ad ogni altro \hat{x}_j per $j \neq i$.

\hat{x}_i viene definito il centroide della i-esima regione di quantizzazione.

¹²Slide del prof — Conversione AD — pag 8.1
Slide — Conversione AD — 8.1
Appunti — 2025-03-17 — pag 6 - 7

4.5 PCM Pulse Code Modulation

¹³

La PCM è una modulazione impiegata in molti ambiti, ad esempio il canale voce negli anni 80, oppure nella lettura e scrittura dei DVD.

Se vuoi farti una cultura:

https://it.wikipedia.org/wiki/Modulazione_a_impulsi_codificati

Uno schema a blocchi di una modulazione PCM è la seguente:

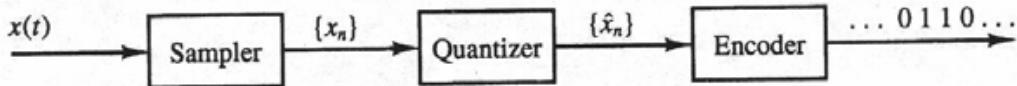


FIGURE 4.15. Block diagram of a PCM system.

Soffermandoci sul quantizzatore, esistono due varianti di PCM:

- PCM uniforme
- PCM non uniforme

¹³Slide del prof — Conversione AD — pag 8.2
 Slide — Conversione AD — pag 8.2
 Appunti — 2025-03-17 — pag 7

4.5.1 PCM uniforme

¹⁴

Facciamo le seguenti ipotesi:

- il segnale di ingresso è compreso nell'intervallo $[-x_{max}, +x_{max}]$
- il numero di livelli di quantizzazione è $N = 2^v$
- ciascun campione è rappresentato con v impulsi binario

Allora l'intervallo di quantizzazione Δ è:

$$\begin{aligned}\Delta &= \frac{|-x_{max}| + |x_{max}|}{N} \\ &= \frac{2 \cdot x_{max}}{2^v} \\ &= \frac{x_{max}}{2^{v-1}}\end{aligned}\tag{4.20}$$

Ipotizzando che:

- il numero dei livelli N è molto elevato e l'intervallo di varietà x_{max} è piccolo, allora Δ sarà piccolo.
- il livello di quantizzazione al centro della regione di quantizzazione è al centro della regione di quantizzazione

allora l'errore è approssimabile con una distribuzione uniforme tra $-\frac{\Delta}{2}$ e $+\frac{\Delta}{2}$.

Come abbiamo con l'ottimizzazione della quantizzazione uniforme, calcoliamoci la potenza del segnale in PCM.

Sappiamo che \tilde{x} è la differenza tra x valore analogico e $Q(x)$ funzione di quantizzazione.

In formule:

$$\tilde{x} = x - Q(x)\tag{4.21}$$

e questa relazione vale in tutto l'intervallo di quantizzazione, quindi in $2 \cdot x_{max}$.

Inoltre, sappiamo che \tilde{x} è una variabile aleatoria nell'intervallo $(-\frac{\Delta}{2}, +\frac{\Delta}{2}]$.

Nella PCM uniforme vogliamo calcolarci la potenza di \tilde{x} in ogni Δ .

Ricordando che, nel campionamento, consideriamo \tilde{x} il rumore perché il rumore fa variare il valore campionato x , allora calcoliamoci la potenza del rumore, cioè la potenza di \tilde{x} .

Essendo \tilde{x} una variabile aleatoria, allora:

¹⁴Slide del prof — Conversione AD — pag 9.1 - 10.1
Slide — Conversione AD — pag 9.1 - 10.1
Appunti — 2025-03-17 — pag 7

$$\begin{aligned}
\mathbb{E} [\tilde{X}]^2 &= \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \frac{1}{\Delta} \cdot \tilde{x}^2 dx \\
&= \frac{1}{\Delta} \cdot \frac{\tilde{x}^3}{3} \Big|_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \\
&= \dots \\
&= \frac{\Delta^2}{12} \\
&= \frac{x_{max}^2}{3N^2} \\
&= \frac{x_{max}^2}{3 \cdot 2^{2v}} \\
&= \frac{x_{max}^2}{3 \cdot 4^v}
\end{aligned} \tag{4.22}$$

Ora che sappiamo quanto vale la potenza del rumore di quantizzazione, possiamo calcolarci l'SQNR (Signal to Quantization Noise Ratio):

$$\begin{aligned}
SQNR &= \frac{\mathbb{E}[X^2]}{\mathbb{E}[\tilde{X}^2]} \\
&= \frac{P_X}{\mathbb{E}[\tilde{X}^2]} \\
&= 3 \cdot 4^v \cdot \frac{P_x}{x_{max}^2} \\
&= 3 \cdot 4^v \cdot P_{X_n}
\end{aligned} \tag{4.23}$$

dove:

- P_X è la potenza del segnale utile prima della quantizzazione
- v sono i bit della quantizzazione
- x_{max} è il valore massimo che il segnale può assumere nel quantizzatore
- P_{X_n} è la potenza P_X normalizzata, cioè $\frac{P_x}{x_{max}^2}$

Molte volte può essere utile esprimere l'SQNR in dB.

Quindi la relazione di prima approssimata in dB diventa:

$$\begin{aligned}
SQNR &= 3 \cdot 4^v \cdot P_{X_n} \\
&\downarrow \\
SQNR|_{dB} &\approx P_{X_n}|_{dB} + 6 \cdot v + 4.8
\end{aligned} \tag{4.24}$$

Dalla formula dell'SQNR in dB si può affermare che, per ogni incremento di 1 nel valore di v , l'SQNR aumenta di 6 dB.

Anche questa osservazione ha senso da un punto di vista logico, perché se aumentassimo a infinito i valori di quantizzazione, il rumore di quantizzazione diminuirebbe, cioè l'SQNR migliora perché, in questa maniera, la quantizzazione distorce di meno il segnale di ingresso.

4.5.2 PCM non uniforme

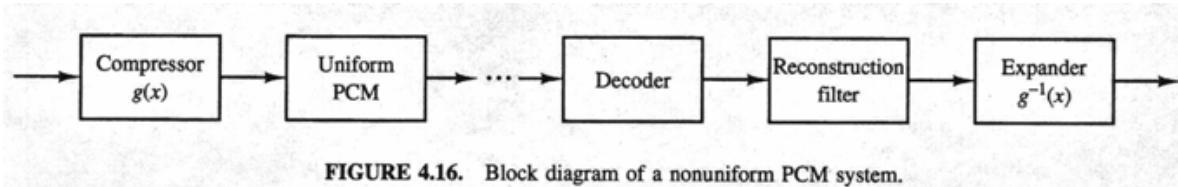
15

Al posto della PCM uniforme, quindi l'uso della quantizzazione uniforme, si può impiegare una PCM non uniforme, quindi l'uso della quantizzazione non uniforme.

La PCM non uniforme è molto efficace quando $f_X(x)$ differisce significativamente dall'andamento uniforme, cioè ci sono dei valori nella quantizzazione che sono più probabili rispetto ad altri.

Un esempio di segnale in cui i valori quantizzati sono più probabili rispetto ad altri è la voce umana perché i valori bassi saranno quelli più probabili rispetto a quelli alti.

Lo schema di una PCM non uniforme è la seguente:



Come si nota a sinistra della figura, per rendere una PCM da uniforme a non uniforme, si può aggiungere prima della PCM uniforme un compressore, che ha questo tipo di andamento:

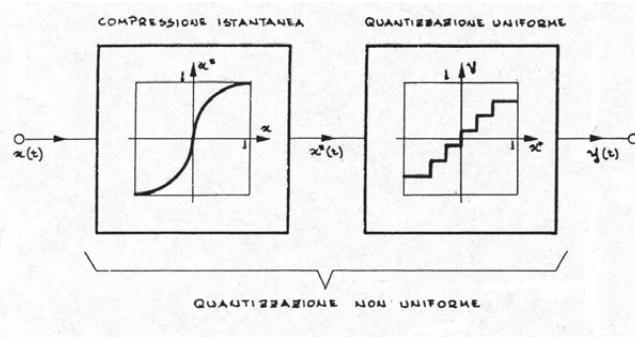


Fig. III.7 – Realizzazione di principio di una quantizzazione non uniforme

Per motivi storici, la legge di compressione della voce tra Stati Uniti e Europa sono diversi (ma nei nuovi standard adesso questa differenza non esiste più).

¹⁵Slide del prof — Conversione AD — pag 10.2 - 12.2
Slide — Conversione AD — pag 10.2 - 12.2
Appunti — 2025-03-17 — pag 8

4.5.3 Soluzioni per ridurre la banda occupata

16

Il problema più grosso della PCM è quello che, all'aumentare degli N livelli di quantizzazione, bisogna anche aumentare la banda del segnale modulato: per questo motivo, sono stati ideati e realizzati altre tecniche per diminuire la banda occupata dal segnale dopo la modulazione.

Una prima soluzione è l'uso della DPCM (Differential Pulse Code Modulation).

Nella DPCM si quantizza e si codifica la differenza tra due campioni successivi, cioè $X_n - X_{n-1}$.

Questa differenza tra i campioni successivi crea una dipendenza, in particolare una correlazione, tra i campioni.

Ma se la differenza tra i due campioni successiva è piccola, non c'è nessun problema nell'impiegare la DPCM.

Come si visualizza dal seguente schema di una DPCM:

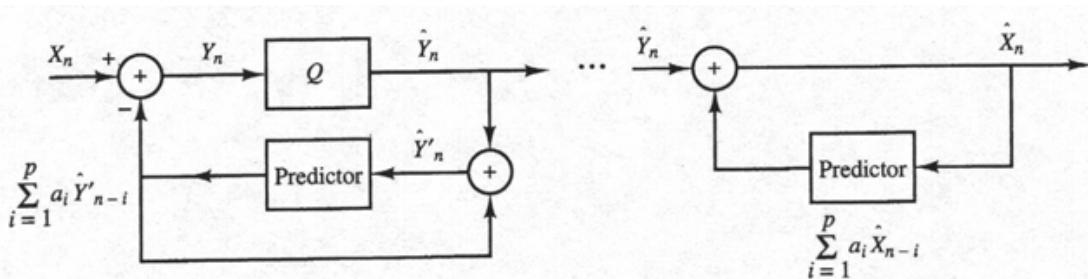


FIGURE 4.21. A general DPCM scheme.

si può notare il blocco del predittore (Predictor) nello schema della DPCM.

Questo perché nella DPCM il valore corrente (cioè quello di adesso) viene "predetto" sulla base di p valori quantizzati precedenti e poi si quantizza la differenza tra il valore corrente e la sua predizione.

Oppure, al posto di utilizzare una DPCM, si utilizza una modulazione Delta, che è un caso limite della DPCM, in cui transita solo un bit si segno (se è più grande o più piccolo di un bit rispetto al valore precedente).

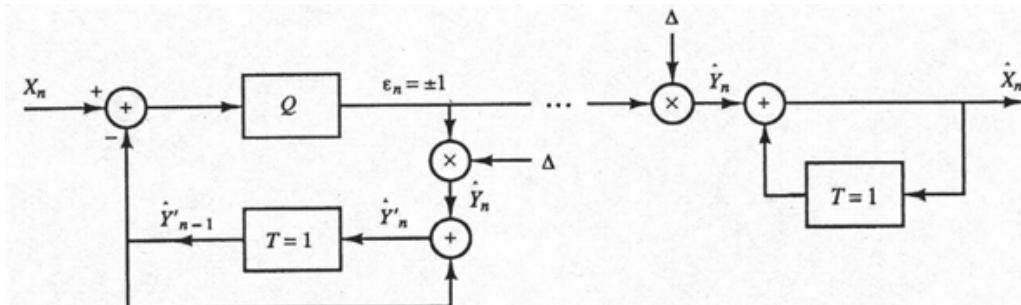


FIGURE 4.22. Delta modulation.

In questo caso si utilizza una frequenza di campionamento molto elevata, maggiore di quella fissata dalla condizione di Nyquist.

La riduzione della banda è minore di quanto prevedibile, ma pure sempre consistente.

¹⁶Slide del prof — Conversione AD — pag 13.1 - 14.2
Slide — Conversione AD — pag 13.1 - 14.2

Per la modulazione delta si può utilizzare anche questo tipo di architettura:

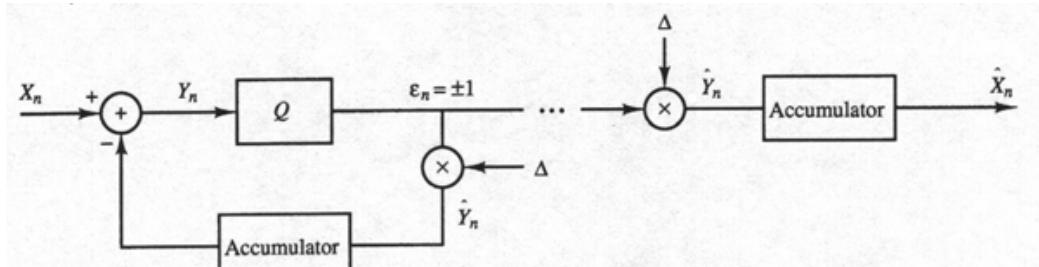


FIGURE 4.23. Delta modulation with integrators.

Degli ADC sigma-delta e ADC e convertitori ad inseguimento, che seguono lo stesso principio della modulazione delta ne ho approfondito qui.

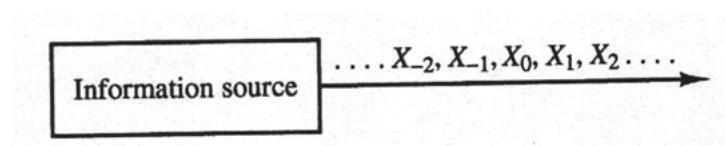
<https://github.com/ciccio25/appunti-misure-elettroniche>
Capitolo 9.8 - Convertitore ad inseguimento - pag 176 - 181

Ma, nel caso di un telecomunicazionista, non si filtra in banda base, bensì si filtra nella banda in cui bisogna trasmettere il segnale.

<https://github.com/ciccio25/appunti-misure-elettroniche>
Capitolo 9.9 - Sigma-Delta Σ – Δ - pag 182 - 183

Capitolo 5

Teoria dell'informazione



5.1 Modello matematico per una sorgente di informazione

¹

Le sorgenti di informazione possono essere modellate come processi stocastici, le cui proprietà dipendono dalla natura e dalle caratteristiche della sorgente.

Ripassati al volo il capitolo sui processi stocastici da Teoria dei segnali perché sarà molto utile.

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 13 - Processi stocastici - pag 151 - 158

Ad esempio questo è lo spettro di un segnale di potenza di una voce umana:

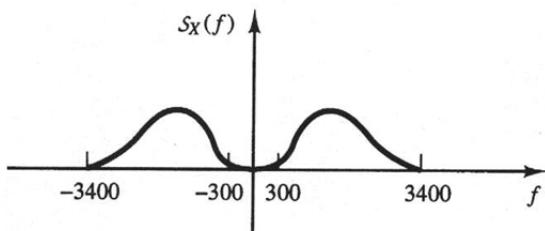


FIGURE 4.1. Typical power spectrum of speech signal.

¹Slide del prof — Teoria dell'informazione — pag 1.1
Slide — Teoria dell'informazione — pag 1.1

5.2 Sorgente discreta

2

Consideriamo la seguente sorgente:

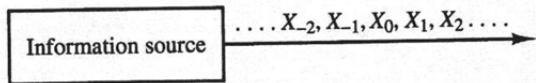


FIGURE 4.2. Mathematical model for a discrete-time information source.

cioè una sorgente discreta, cioè che ha un alfabeto di valori finiti e non infiniti, in particolare una sorgente d.m.s., cioè una sorgente discreta senza memoria (discrete memoryless sources).

d.m.s. non d.r.s. della Formula 1

²Slide del prof — Teoria dell'informazione — pag 1.2
Slide — Teoria dell'informazione — pag 1.2

5.2.1 Sorgente d.m.s.

³

In una sorgente d.m.s., i simboli sono generati indipendente a partire da una assegnata (e costante, se la sorgente è stazionario) distribuzione di probabilità.

Ricordiamo che essendo indipendenti, ogni elemento dell'alfabeto ha covarianza nulla con tutti gli altri elementi.

La sorgente è caratterizzata da un alfabeto A:

$$A = a_1, a_2, \dots, a_N \quad (5.1)$$

di valori possibili e una distribuzione di probabilità discrete p_i :

$$p_i = \Pr\{X = a_i\} \quad \forall i = 1, 2, \dots, N \quad (5.2)$$

Considerando come esempio una sorgente binario, avremo che l'alfabeto A è uguale a:

$$A = 0, 1 \quad (5.3)$$

Dalla teoria della probabilità, si sa che la probabilità dell'alfabeto è:

$$\Pr\{A\} = 1 \quad (5.4)$$

quindi, essendo presenti un due elementi equiprobabili:

$$\begin{aligned} \Pr\{X = 1\} &= \Pr\{A\} - \Pr\{X = 0\} \\ &= 1 - \Pr\{X = 0\} = p \end{aligned} \quad (5.5)$$

Questa equazione è vera solo se:

$$\Pr\{X = 1\} = \Pr\{X = 0\} = p = 0.5 \quad (5.6)$$

Quando $p = 0.5$, si dice che è una sorgente binaria simmetrica (o b.s.s.), che è un caso particolare.

³Slide del prof — Teoria dell'informazione — pag 2
Slide — Teoria dell'informazione — pag 2

5.3 Misura dell'informazione

4

Prima di parlare e discutere di informazione, dobbiamo specificare cosa si intende per informazione.



Per informazione si intende la riduzione di incertezza che si può calcolare a priori sul verificarsi di un evento.

L'informazione associata ad un evento dipende dalla probabilità dell'evento: più la probabilità è bassa, più l'informazione è elevata.

Immagina di conoscere una coppia che è fidanzata da anni.

Per te è più una informazione che ieri si sono lasciati o che ieri hanno comprato il gelato?

Ovviamente, l'informazione è che si sono lasciati (se sei uno come me, sarebbe più interessato al gelato perché vorrebbe sapere se il gelato è buono e dove lo hanno comprato, ma questo è un altro discorso) perché, prima di oggi, sapevi che stavano bene insieme e non avevano mai litigato.

Un evento certo non "porta" informazione.

Se io accendo il fornello, è una informazione farsi male toccando il fuoco con la mano?

No, non lo è (a meno che non sei masochista, ma qui, come scritto prima, non consideriamo i casi particolari), perché sicuramente ti farai male.

Quindi toccare con una mano il fuoco non è una informazione perché è certa.

L'informazione non cambia in modo significativo se la probabilità non varia in modo significativo.

Se sai che domani pioverà e piuttosto che piovere per 30 minuti, l'app del meteo ti manda una notifica in cui ti informa che domani pioverà per 35 minuti.

La notifica dell'app del meteo è informativa?

Sì, lo è, ma 35 minuti di pioggia non sono tanto diversi dai 30 minuti che sapevi prima della notifica.

⁴Slide del prof — Teoria dell'informazione — pag 3.1
Slide — Teoria dell'informazione — pag 3.1

Una proprietà molto importante di due eventi indipendenti, e che quindi hanno covarianza nulla, l'informazione ad essi associata è uguale alla somma delle informazioni degli eventi singoli.

Prima di andare avanti, ripassati al volo la teoria assiomatica di Kolmogorov, che è la base della teoria della probabilità.

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.2 - Teoria assiomatica di Kolmogorov - pag 114 - 115

In particolare, uno dei punti della Teoria assiomatica di Kolmogorov è la seguente.

Dati due eventi A e B mutuamente esclusivi (ovvero incompatibili, cioè tali che non possano verificarsi contemporaneamente) la probabilità dell'evento unione è data dalla somma delle probabilità dei singoli eventi, si ha cioè:

$$A \cup B = \emptyset \rightarrow \Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} \quad (5.7)$$

5.4 Auto-informazione

5

Dato un evento (ad esempio la generazione di un simbolo a_j da parte di una sorgente) con probabilità p_j , l'auto-informazione viene indicata con $I(p_j)$.

L'auto-informazione gode delle seguenti proprietà:

- è una funzione continua di p_j
- è una funzione decrescente del suo argomento
- se $p_j = p_{j1} \cdot p_{j2}$ allora $I(p_j) = I(p_{j1}) + I(p_{j2})$

Si può dimostrare che l'unica funzione in grado di soddisfare queste proprietà è la funzione logaritmo.

In particolare, da una probabilità p_j , possiamo esprimere la sua auto-informazione come:

$$I(p_j) = -\log(p_j) \quad (5.8)$$

La base del logaritmo non è importante, ma definisce l'unità con cui si esprime l'informazione.

Se la base del logaritmo è 2, l'informazione è espressa in bit.

Se la base del logaritmo è in base e, cioè ln, l'informazione è espressa in nat.

Siccome abbiamo a che fare con sistemi digitali, abbiamo a che fare sempre con:

$$I(p_j) = -\log_2(p_j) \quad (5.9)$$

Nelle slide sarà solo indicato log.

La proprietà dell'auto-informazione che esprime che "una funzione decrescente del suo argomento" si intende l'argomento del logaritmo.

Sapendo che l'informazione di p_j possiamo scriverla come, applicando le proprietà del logaritmo:

$$\begin{aligned} I(p_j) &= -\log_2(p_j) \\ &= \log_2(p_j)^{-1} \\ &= \log_2\left(\frac{1}{p_j}\right) \end{aligned} \quad (5.10)$$

All'aumentare di p_j , l'argomento del logaritmo diminuisce e tende a zero per p_j che tende ad infinito.

⁵Slide del prof — Teoria dell'informazione — pag 3.2
 Slide — Teoria dell'informazione — pag 3.2
 Appunti — 2025-07-08 Ricevimento — pag 1.2

5.4.1 Contenuto informativo di una d.m.s.

6

Consideriamo una d.m.s. (discrete memoryless source):

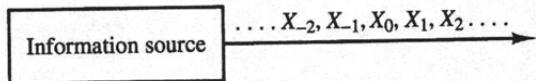


FIGURE 4.2. Mathematical model for a discrete-time information source.

Consideriamo la media "pesata" delle auto-informazioni dei simboli della sorgente.

$$\sum_{i=1}^N p_i \cdot I(p_i) \quad (5.11)$$

La media-pesata delle auto-informazioni prende il nome di entropia, in questo caso la entropia della sorgente.

Andando a segnare con $H(X)$ l'entropia della sorgente d.m.s. come:

$$\begin{aligned} H(X) &= \sum_{i=1}^N p_i \cdot I(p_i) \\ &= \sum_{i=1}^N p_i \cdot [-\log_2(p_i)] \\ &= -\sum_{i=1}^N p_i \cdot \log_2(p_i) \\ &= \sum_{i=1}^N p_i \cdot \log_2\left(\frac{1}{p_i}\right) \end{aligned} \quad (5.12)$$

⁶Slide del prof — Teoria dell'informazione — pag 4.1
Appunti di Damiano — pag 4.1
Slide — Teoria dell'informazione — pag 4.1

5.4.1.1 Entropia binaria

7

Per una sorgente d.m.s. binaria, abbiamo che l'entropia $H(X)$ è uguale a:

$$\begin{aligned}
 H(X) &= \sum_{i=1}^N p_i \cdot I(p_i) \\
 &= - \sum_{i=1}^2 p_i \cdot \log_2(p_i) \\
 &= -p \cdot \log_2(p) - (1-p) \cdot \log_2(1-p)
 \end{aligned} \tag{5.13}$$

Posiamo visualizzare $H(X)$ in questa maniera:

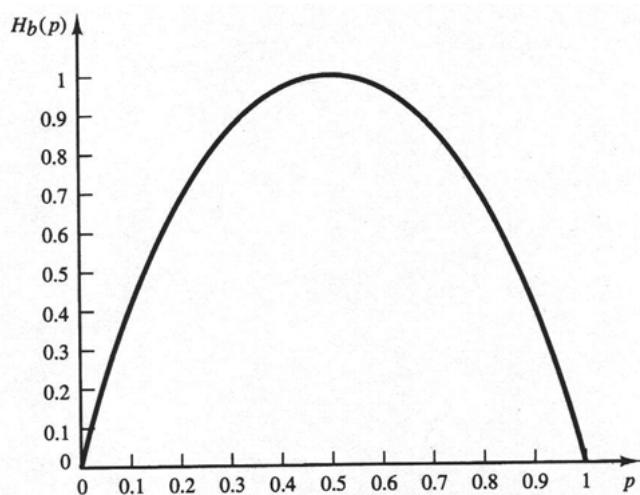


FIGURE 4.3. The binary entropy function.

$H(X)$ è una funzione che ha le seguenti caratteristiche:

- concava
- simmetrica
- ha un massimo per $p = \frac{1}{2}$

Se 0 e 1 sono equiprobabili, la sequenza binaria è una sequenza di bit (informativi).

Se p si discosta da $\frac{1}{2}$, alcuni elementi binari non portano informazioni, quindi non tutti gli 0 e 1 sono bit (informativi).

Per bit si intendono solo quelli che portano una informazione e questa è la definizione nella teoria dell'informazione.

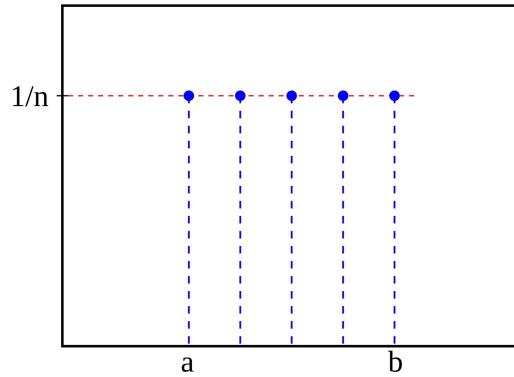
Nella realtà, il termine bit viene spesso usato come sinonimo di 0 e 1.

⁷Slide del prof — Teoria dell'informazione — pag 4.2
Appunti di Damiano — pag 4.2
Slide — Teoria dell'informazione — pag 4.2

5.4.1.2 Entropia massima

8

Graficamente una funzione di distribuzione discreta uniforme, ha il seguente andamento:



dove N è il numero degli elementi dell'alfabeto A .

Questo perché, data la formula di $H(X)$, sappiamo che:

$$\begin{aligned} H(X) &= \sum_{i=1}^N p_i \cdot I(p_i) \\ &= \sum_{i=1}^N p_i \cdot \log_2 \left(\frac{1}{p_i} \right) \end{aligned} \tag{5.14}$$

Sapendo che per una funzione di distribuzione discreta uniforme, la probabilità di ogni singolo N elemento dell'alfabeto i p_i vale:

$$p_i = N \tag{5.15}$$

allora:

⁸Slide del prof — Teoria dell'informazione — pag 5
Appunti di Damiano — pag 5
Slide — Teoria dell'informazione — pag 5
Appunti — 2025-07-08 Ricevimento — pag 1.1

$$\begin{aligned}
H(X) &= \sum_{i=1}^N p_i \cdot \log_2 \left(\frac{1}{p_i} \right) \\
&= - \sum_{i=1}^N p_i \cdot \log_2 (p_i) \\
&\quad \downarrow \\
H(X) &= - \sum_{i=1}^N \frac{1}{N} \cdot \log_2 (N) \\
&= - \sum_{i=1}^N \frac{1}{N} \cdot \log_2 (N) \\
&= - \log_2 (N) \cdot \sum_{i=1}^N \frac{1}{N} \\
&= - \log_2 (N) \cdot 1 \\
&= - \log_2 (N)
\end{aligned} \tag{5.16}$$

Dal secondo al terzo passaggio (contando come prima la sostituzione con $p_i = \frac{1}{N}$), posso portare fuori dalla sommatoria il termine $\log_2 (N)$ perché è comune a tutti i termini della sommatoria stessa.

Dal terzo al quarto passaggio, $\sum_{i=1}^N \frac{1}{N} = 1$ per definizione di funzione di distribuzione discreta (la probabilità di tutti gli elementi di un alfabeto è del 100 % quindi 1).

Per una sorgente d.m.s., si avrà entropia massima se si ha una funzione di distribuzione discreta uniforme: di seguito la dimostrazione.

Consideriamo $H(X)$ una entropia generica e $H_{max}(X)$ l'entropia di una funzione di distribuzione discreta uniforme.

Possiamo scrivere che:

$$\begin{aligned}
H(X) + H_{max}(X) &= - \sum_{i=1}^N p_i \cdot \log_2 (p_i) - \log_2 (N) \\
&= \sum_{i=1}^N p_i \cdot \log_2 \left(\frac{1}{p_i} \right) + \log_2 \left(\frac{1}{N} \right) \\
&= \sum_{i=1}^N p_i \cdot \log_2 \left(\frac{1}{p_i \cdot N} \right)
\end{aligned} \tag{5.17}$$

Siccome vogliamo trovare un massimo, questa sommatoria è minore uguale a:

$$\begin{aligned}
\sum_{i=1}^N p_i \cdot \left(\frac{1}{p_i \cdot N} - 1 \right) \cdot \log_2(e) &= \left(\sum_{i=1}^N \frac{1}{N} - \sum_{i=1}^N p_i \right) \cdot \log_2(e) \\
&= \left(\sum_{i=1}^N \frac{1}{N} - \sum_{i=1}^N \frac{1}{N} \right) \cdot \log_2(e) \quad (5.18) \\
&= (1 - 1) \cdot \log_2(e) \\
&= 0 \cdot \log_2(e) \\
&= 0
\end{aligned}$$

Mettendo insieme le due equazione abbiamo che:

$$\begin{aligned}
H(X) + H_{max}(X) &\leq 0 \\
\downarrow \\
H(X) &\leq -H_{max}(X) \quad (5.19) \\
&\leq -[-\log_2(N)] \\
&\leq \log_2(N)
\end{aligned}$$

Quindi, alla fine della dimostrazione abbiamo che, data qualsiasi entropia $H(X)$ essa è sempre minore o uguale all'entropia massima, che vale $\log_2(N)$.

In formule:

$$H(X) \leq \log_2(N) \quad (5.20)$$

dove N sono gli elementi dell'alfabeto.

Un altro metodo per capire e visualizzare meglio cosa si intende per entropia, ti lascio questo video di TED-ed.

https://www.youtube.com/watch?v=YM-uykVfq_E
Cos'è l'entropia? — Jeff Phillips by TED-ed

Nel caso nostro dell'entropia collegata alla probabilità, in particolare all'entropia massima, è meglio questo video:

<https://youtu.be/2s3aJfRr9gE?si=J37vVjWY5SWied3H>
Information entropy — Journey into information theory — Computer Science — Khan Academy by Khan Academy Labs

Se vuoi andare più a fondo, c'è il mitico Josh Starmer (mitico perché l'ho scoperto quando stavo facendo un progetto sulle reti neurali ed io non sapevo nulla sull'argomento, persona molto coinvolgente):

<https://youtu.be/YtebGVx-Fxw?si=902P-pN-1X6nAi3j>
Entropy (for data science) Clearly Explained!!! by StatQuest with Josh Starmer

Una nota riguardo alle funzioni di probabilità e la differenza tra le materie.

Nella probabilità, dobbiamo considerare l'opposto rispetto al corso di misure. Se PDF uniforme è il caso peggiore in misure, qui è il caso migliore.

Se PDF gaussiano è il caso migliore nelle misure, qui è il caso peggiore perchè ci saranno dei valori più probabili rispetto agli altri dell'alfabeto.

5.4.1.3 Entropia congiunta

9

Se vuoi "divertirti" a capire meglio l'entropia congiunta, ti consiglio questo video:

<https://youtu.be/6ArSys5qHAU?si=0KeUXNtcMn2lytFE>

Neural Networks Part 6: Cross Entropy by StatQuest with Josh Starmer

L'entropia, come descritto precedentemente, descrive un comportamento medio del sistema.

Dall'entropia per una variabile X, possiamo esprimere l'entropia per due variabili (X, Y) come:

$$H(X, Y) = - \sum_{x,y} p(x, y) \cdot \log_2 [p(x, y)] \quad (5.21)$$

dove:

- $\sum_{x,y}$ si intende la somma di tutti i valori dell'alfabeto X e la somma di tutti valori dell'alfabeto Y
- $p(x, y)$ è la probabilità congiunta tra l'elemento x e y

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.2 Teoria assiomatica di Kolmogorov - pag 114 - 115

Dati due eventi A e B, la probabilità dell'evento unione $A \cup B$ è espressa dall'uguaglianza:

$$\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} + \Pr\{A \cap B\} \quad (5.22)$$

Data una coppia di eventi A e B, la probabilità dell'evento intersezione, spesso indicata semplicemente con $\Pr\{A, B\}$, è detta probabilità congiunta, mentre $\Pr\{A\}$ e $\Pr\{B\}$ hanno il significato di probabilità marginali.

■

Ma per eventi indipendenti, la definizione cambia.

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.5 Eventi statisticamente indipendenti - pag 118 - 119

Per eventi indipendenti:

$$\Pr\{A, B\} = \Pr\{A\} \cdot \Pr\{B\} \quad (5.23)$$

e cioè, che la probabilità congiunta è uguale al prodotto delle probabilità marginali (dei singoli eventi).

Considerando il caso più generale con n variabili \mathbf{X} :

$$\mathbf{X} = (X_1, X_2, \dots, X_n) \quad (5.24)$$

si può estendere la definizione di entropia alle n variabili come:

⁹ Slide del prof — Teoria dell'informazione — pag 6.1

Appunti di Damiano — pag 6.1

Slide — Teoria dell'informazione — pag 6.1

Appunti — 2025-03-21 — pag 2

$$H(\mathbf{X}) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, x_2, \dots, x_n) \cdot \log_2 [p(x_1, x_2, \dots, x_n)] \quad (5.25)$$

5.4.1.4 Entropia condizionata

[10](#)

Beccatevi il socio Josh Starmer.

https://youtu.be/eJIp_mgVLwE?si=1JD4NI1HlypPZRro

Mutual Information, Clearly Explained!!! by StatQuest with Josh Starmer

”Mutual Information” proprio perchè, come abbiamo scritto precedentemente, per noi, nella teoria dell’informazione, maggiore entropia significa maggiore informazione.

Più casino c’è, più informazione c’è.

Come nella probabilità, anche nell’entropia possiamo considerare l’entropia condizionata.

Se consideriamo solo un elemento nell’alfabeto y , cioè $Y = y$, allora possiamo formulare l’entropia condizionata tra X e Y come:

$$X(X|Y = y) = - \sum_x p(x|y) \cdot \log_2 [p(x|y)] \quad (5.26)$$

Invece, considerando tutti i possibili valori di Y , e non uno singolo come prima, possiamo scrivere, facendo alcuni passi algebrici:

$$\begin{aligned} H(X|Y) &= - \sum_x \sum_y p(y) \cdot p(x|y) \cdot \log_2 [p(x|y)] \\ &= \sum_{x,y} p(x,y) \cdot \log_2 [p(x|y)] \end{aligned} \quad (5.27)$$

dove $p(x|y)$ prende il nome di probabilità condizionata di x su y .

¹⁰Slide del prof — Teoria dell’informazione — pag 6.2
Appunti di Damiano — pag 6.2
Slide — Teoria dell’informazione — pag 6.2
Appunti — 2025-03-21 — pag 2

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.2 Teoria assiomatica di Kolmogorov - pag 114 - 115

Data una coppia di eventi A e B con $\Pr\{B\} \neq 0$, si definisce poi la probabilità condizionata:

$$\begin{aligned}\Pr\{A|B\} &= \frac{\Pr\{A \cup B\}}{\Pr\{B\}} \\ &= \frac{\Pr\{A, B\}}{\Pr\{B\}}\end{aligned}\tag{5.28}$$

la quale esprime la probabilità dell'evento A condizionata al verificarsi dell'evento B (che ha dunque il significato di evento condizionante).

■

Per eventi statisticamente indipendenti invece, dobbiamo utilizzare la formula di Bayes.

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.5 Eventi statisticamente indipendenti - pag 118 - 119

Un altro importante risultato è il teorema (o formula) di Bayes, che può essere formalizzato nel modo seguente:

$$\Pr\{A|B\} = \frac{\Pr\{B|A\} \cdot \Pr\{A\}}{\Pr\{B\}}\tag{5.29}$$

La formula di Bayes è spesso usata in combinazione con il teorema della probabilità totale.

Possiamo generalizzare l'entropia condizionata per n variabili:

$$H(X_n|X_1, X_2, \dots, X_{n-1}) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, x_2, \dots, x_n) \cdot \log_2 [p(x_1|x_2, \dots, x_n)]\tag{5.30}$$

5.4.2 Legame tra entropia condizionata ed entropia congiunta

[11](#)

Di seguito la dimostrazione del legame tra una entropia condizionata $H(X|Y)$ ed entropia congiunta $H(X, Y)$:

$$\begin{aligned}
 H(X, Y) &= - \sum_{x,y} p(x, y) \cdot \log_2 [p(x, y)] \\
 &= - \sum_{x,y} p(x, y) \cdot \log_2 [p(y) \cdot p(x|y)] \\
 &= - \sum_{x,y} p(x, y) \cdot \log_2 [p(y)] \\
 &= - \sum_{x,y} p(x, y) \cdot \log_2 [p(x|y)] \\
 &= - \sum_{x,y} p(y) \cdot \log_2 [p(y)] \\
 &= - \sum_{x,y} p(x, y) \cdot \log_2 [p(x|y)] \\
 &= H(Y) + H(X|Y)
 \end{aligned} \tag{5.31}$$

cioè:

$$H(X, Y) = H(Y) + H(X|Y) \tag{5.32}$$

Ma, alfabeti indipendenti, il legame diventa:

$$H(X, Y) = H(Y) + H(X) \tag{5.33}$$

Generalizzando a n variabili $\mathbf{X} = (X_1, X_2, \dots, X_n)$:

$$H(\mathbf{X}) = H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_1, X_2, \dots, X_{n-1}) \tag{5.34}$$

Dalla formula generale, possiamo formulare l'entropia $H(\mathbf{X})$ se ogni X_i è indipendente, come:

$$\begin{aligned}
 H(\mathbf{X}) &= H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_1, X_2, \dots, X_{n-1}) \\
 &\downarrow \\
 H(\mathbf{X}) &= \sum_{i=1}^n H(X_i)
 \end{aligned} \tag{5.35}$$

Se consideriamo X_i indipendenti e la sorgente è stazionaria, l'entropia si semplifica ulteriormente come:

¹¹Slide del prof — Teoria dell'informazione — pag 7.1
Appunti di Damiano — pag 7.1
Slide — Teoria dell'informazione — pag 7.1
Appunti — 2025-03-21 — pag 2

$$\begin{aligned} H(\mathbf{X}) &= \sum_{i=1}^n H(X_i) \\ &\downarrow \\ H(\mathbf{X}) &= n \cdot H(X) \end{aligned} \tag{5.36}$$

5.4.2.1 Entropy rate H di una sorgente stazionaria

¹²

L'entropy rate sostituisce l'entropia nel caso di una sorgente con memoria.

Per questo motivo, siccome in questo corso consideriamo solo sorgente d.m.s senza memoria, questo paragrafo è solo per approfondire un argomento, ma che realmente non andremo mai a svolgere gli esercizi.

L'entropia X_2 condizionata X_1 , cioè $H(X_2|X_1)$, rappresenta l'informazione "aggiuntiva" fornita da X_2 quando si conosce il valore di X_1 .

Considerando il caso generale, cioè $H(X_n|X_1, X_2, \dots, X_{n-1})$, rappresenta l'informazione "aggiuntiva" fornita da X_n quando si conoscono gli n-1 valori precedenti $(X_1, X_2, \dots, X_{n-1})$.

Per n che tende ad infinito, non si considera più l'entropia $H(\mathbf{X})$, bensì l'entropy rate H, che ha la seguente formula:

$$\begin{aligned} H &= \lim_{n \rightarrow \infty} H(X_n|X_1, X_2, \dots, X_{n-1}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \cdot H(X_1, X_2, \dots, X_n) \end{aligned} \quad (5.37)$$

Nel contesto pratico, non andremo a svolgere i calcoli manualmente, si utilizzeranno le formule.

Ad esempio, se consideriamo l'alfabeto di lingua inglese formato da 27 caratteri, e considerando i simboli (cioè la singola lettera dell'alfabeto) equiprobabili, possiamo calcolare l'entropia come:

$$\begin{aligned} H(X) &= \log_2(N) \\ &= \log_2(27) \\ &= 4.755 \frac{\text{bit}}{\text{simbolo}} \end{aligned} \quad (5.38)$$

¹²Slide del prof — Teoria dell'informazione — pag 7.2 - 8.1
 Appunti di Damiano — pag 7.2 - 8.1
 Slide — Teoria dell'informazione — pag 7.2 - 8.1
 Appunti — 2025-03-21 — pag 2

Invece se consideriamo l'entropy rate H , cioè l'entropia di una sorgente non d.m.s. e con memoria (cioè come succede nel linguaggio parlato), possiamo calcolarcelo "a mano", oppure possiamo utilizzare la seguente tabella:

LETTERA	ITALIANO	FRANCESE	SPAGNOLO	TEDESCO	INGLESE
-	1583	1735	1658	1442	1753
A	984	635	1078	594	724
E	945	1560	1138	1440	1042
I	863	475	652	628	641
O	844	431	672	211	551
U	255	509	343	422	195
N	684	575	595	865	554
T	670	508	319	536	842
R	539	605	527	622	557
L	512	494	512	345	320
C	395	256	408	255	231
S	376	832	563	646	668
D	303	399	464	546	319
P	202	188	236	67	148
M	193	251	197	172	145
G	174	90	105	236	126
V	120	124	97	79	91
H	93	48	65	361	419
F	82	84	51	78	205
B	72	90	108	138	75
Z	72	3	23	92	12
Q	39	62	88	5	6
J	-	21	23	28	9
K	-	-	-	71	53
W	-	-	-	113	172
X	-	18	9	8	9
Y	-	6	69	-	133
H(A) bit/simbolo	3,959	3,896	3,979	4,045	4,057

Tabella 2 - Frequenze relative dei simboli in alcune lingue. I valori riportati vanno divisi per 10.000.

dove, notiamo che, per la lingua inglese l'entropy rate H è uguale a:

$$H \approx 4.057 \frac{\text{bit}}{\text{simbolo}} \quad (5.39)$$

L'entropy rate nel linguaggio parlato è molto più basso, perchè nel linguaggio c'è ridondanza.

Questo perchè, nel linguaggio parlato, si ripete più spesso la stessa parola, quindi, nell'insieme, la singola lettera porta una bassa informazione.

Grazie a queste considerazioni, avere una sorgente non d.m.s. con memoria può essere molto utile (anche se la trattazione matematica è più complessa), perchè può ottimizzare la trasmissione quando devo ridurre al minimo i simboli da inviare per trasmettere l'informazione (si pensi al caso del wifi dove la banda è limitata e ci sono tanti dispositivi connessi alla stessa rete contemporaneamente).

5.5 Tributo a Claude Shannon

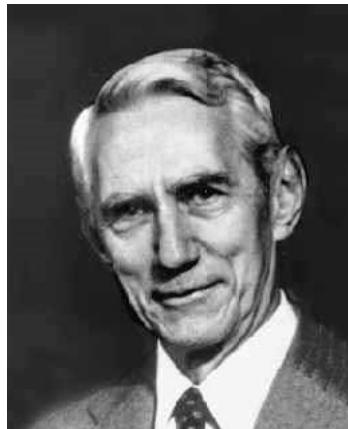
13

Grazie al socio Claude Shannon non potresti fare nulla al giorno d'oggi.

Il prof lo nomina almeno 3 miliardi di volte e fa bene.

Senza Claude non riusciresti a vedere i video su Youtube, telefonare, inviare i video su Whatsapp e neanche ricevere la foto dei piedi di taglia 36 con smalto bianco della tipa che ti piace (o la foto del tipo che ti piace in costume, qui non facciamo differenza di genere e gusti).

Ringrazia Claude Shannon per tutto questo.



Claude Shannon fu un matematico, pioniere della teoria matematica della comunicazione.

Grazie al suo paper "A Mathematical Theory of Communication":

Reprinted with corrections from *The Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656, July, October, 1948.

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set of possible messages*. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

rivoluzionò le comunicazioni e permise di introdurre il concetto di codifica di sorgente.

¹³Slide del prof — Teoria dell'informazione — pag 8.2
Slide — Teoria dell'informazione — pag 8.2
Appunti — 2025-03-21 — pag 3

Se vuoi farti una bella lettura del paper "A Mathematical Theory of Communication", ti lascio il link:

<https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>

5.5.1 Teorema della codifica di sorgente (Source Coding Theorem)

¹⁴

Grazie al suo paper "A Mathematical Theory of Communication", Claude Shannon introdusse diversi 23 teoremi, tra cui ne andremo ad analizzare uno: il teorema della codifica di sorgente, o in inglese Source Coding Theorem.

Di seguito il teorema:

Una sorgente con entropia $H(X)$ (o entropy rate H) può essere rappresentata, senza perdita di informazione, utilizzando un numero medio di simboli (ad esempio binari) per simbolo prodotto dalla sorgente che ha numero di simboli R :

$$R \geq H(X) \quad (5.40)$$

oppure:

$$R \geq H \quad (5.41)$$

Se il numero dei simboli R è:

$$R \leq H(X) \quad (5.42)$$

oppure:

$$R \leq H \quad (5.43)$$

allora, l'informazione è sicuramente ridotta (quindi distorta), per quanto complesse siano le tecniche di rappresentazione utilizzate.

La rappresentazione dei simboli prodotti da una sorgente prende il nome di codifica di sorgente.

Shannon, grazie a questo teorema, impose il limite di codifica, cioè il limite sotto al quale non bisogna scendere per una trasmissione.

¹⁴Slide del prof — Teoria dell'informazione — pag 9.1
Appunti di Damiano — pag 9.1
Slide — Teoria dell'informazione — pag 9.1
Appunti — 2025-03-21 — pag 4

5.5.2 Efficienza di codifica

¹⁵

Come ribadito dal teorema della codifica, se:

$$R \geq H(X) \quad (5.44)$$

cioè il numero di simboli R è maggiore dell'entropia $H(X)$, non si ha perdita di informazione.

Si può definire il coefficiente di efficienza di codifica η :

$$\eta = \frac{H(X)}{R} \quad (5.45)$$

dove:

- se $\eta \leq 1$ non si ha perdita di informazione perchè $R \geq H(X)$
- se $\eta > 1$ si ha perdita di informazione perchè $R < H(X)$

¹⁵Slide del prof — Teoria dell'informazione — pag 9.2
 Appunti di Damiano — pag 9.2
 Slide — Teoria dell'informazione — pag 9.2
 Appunti — 2025-03-21 — pag 4

5.5.2.1 Codifica di sorgente con sequenze codificate di lunghezza fissa

[16](#)

Consideriamo il seguente esempio numerico.

Consideriamo di avere la seguente codifica di simboli:

Simbolo	Codifica
a_1	000
a_2	001
a_3	010
a_4	011
a_5	100

Consideriamo la seguente tabella:

Codewords						
Letter	Probability	Code 1	Code 2	Code 3	Code 4	
a_1	$p_1 = \frac{1}{2}$	1	1	0	00	
a_2	$p_2 = \frac{1}{4}$	01	10	10	01	
a_3	$p_3 = \frac{1}{8}$	001	100	110	10	
a_4	$p_4 = \frac{1}{16}$	0001	1000	1110	11	
a_5	$p_5 = \frac{1}{16}$	00001	10000	1111	110	

Calcoliamoci l'entropia $H(X)$ in cui il codice di codifica ha lunghezza fissa:

$$\begin{aligned}
 H(X) &= \sum_{i=1}^N p_i \cdot \log_2 \left(\frac{1}{p_i} \right) \\
 &= \sum_{i=1}^5 p_i \cdot \log_2 \left(\frac{1}{p_i} \right) \\
 &= \frac{1}{2} \cdot \log_2(2) + \frac{1}{4} \cdot \log_2(4) + \frac{1}{8} \cdot \log_2(8) + \frac{1}{16} \cdot \log_2(16) + \frac{1}{16} \cdot \log_2(16) \\
 &= \frac{1}{2} \cdot \log_2(2) + \frac{1}{4} \cdot \log_2(4) + \frac{1}{8} \cdot \log_2(8) + 2 \cdot \left(\frac{1}{16} \cdot \log_2(16) \right) \\
 &= \frac{15}{8}
 \end{aligned} \tag{5.46}$$

¹⁶Slide del prof — Teoria dell'informazione — pag 10.1
 Appunti di Damiano — pag 10.1
 Slide — Teoria dell'informazione — pag 10.1
 Appunti — 2025-03-21 — pag 4

5.5.3 Sequenza di lunghezza variabile

¹⁷

Se abbiamo lunghezze variabili, come facciamo a capire l'inizio e la fine di un simbolo?

Come si separano le parole di codice prodotte?

Tutte queste problematiche vengono risolte dall'auto-sincronizzazione.

Per applicare l'auto-sincronizzazione, bisogna applicare questo vincolo: non bisogna introdurre simboli ai fini della rappresentazione.

Ad esempio, nel caso binario dove abbiamo un alfabeto composto da 0 o 1, se usiamo il bit 1 per l'informazione, il bit 0 sarà quello che permetterà di dividere il simbolo da un altro.

Ma siccome non avremo mai a che fare con una lunghezza di due bit per rappresentare l'i-esimo simbolo, possiamo definire R numero di simboli come:

$$R = \sum_{i=1}^N L_i \cdot p_i \quad (5.47)$$

dove:

- L_i è la lunghezza della rappresentazione dell'i-esimo simbolo
- p_i è la probabilità dell'i-esimo simbolo

Utilizzare una codifica in cui la lunghezza del simbolo è variabile porta un enorme vantaggio per la banda occupata, perché la banda va calcolata in base a quanti bit si inviano nel canale.

Più bit bisogna inviare, più banda sarà richiesta nel canale.

Continuando l'esempio della sezione precedente, consideriamo la seguente tabella:

Codewords						
Letter	Probability	Code 1	Code 2	Code 3	Code 4	
a_1	$p_1 = \frac{1}{2}$	1	1	0	00	
a_2	$p_2 = \frac{1}{4}$	01	10	10	01	
a_3	$p_3 = \frac{1}{8}$	001	100	110	10	
a_4	$p_4 = \frac{1}{16}$	0001	1000	1110	11	
a_5	$p_5 = \frac{1}{16}$	00001	10000	1111	110	

Se consideriamo la colonna del primo codice (Code 1) abbiamo che $N = 5$ e che il numero dei simboli con questo codice R_1 vale:

¹⁷Slide del prof — Teoria dell'informazione — pag 10.2 - 12.1
Appunti di Damiano — pag 10.2 - 12.1
Slide — Teoria dell'informazione — pag 10.2 - 12.1
Appunti — 2025-03-21 — pag 5 - 6

$$\begin{aligned}
R_1 &= \sum_{i=1}^N L_i \cdot p_i \\
&= \sum_{i=1}^5 L_i \cdot p_i \\
&= L_1 \cdot p_1 + L_2 \cdot p_2 + L_3 \cdot p_3 + L_4 \cdot p_4 + L_5 \cdot p_5 \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 5 \cdot \frac{1}{16} \\
&= \frac{31}{16}
\end{aligned} \tag{5.48}$$

Se consideriamo l'entropia $H(X)$ calcolata precedentemente con parola di codice di lunghezza fissa:

$$H(X) = \frac{15}{8} \tag{5.49}$$

$R_1 \geq H(X)$, quindi con il codice 1 non si ha perdita di informazione.

Inoltre, il codice 1 è auto-sincronizzante perché ogni parola di codice termina con un 1.

Ora facciamo gli stessi calcoli per il codice 2:

$$\begin{aligned}
R_2 &= \sum_{i=1}^N L_i \cdot p_i \\
&= \sum_{i=1}^5 L_i \cdot p_i \\
&= L_1 \cdot p_1 + L_2 \cdot p_2 + L_3 \cdot p_3 + L_4 \cdot p_4 + L_5 \cdot p_5 \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 5 \cdot \frac{1}{16} \\
&= \frac{31}{16}
\end{aligned} \tag{5.50}$$

In questo caso, rispetto al codice 1, ogni parola di codice inizia con un 1, quindi è un codice auto-sincronizzante.

Contrariamente al codice 1, il codice 2 non è istantaneo perché si deve attendere il primo simbolo della parola successiva per decidere che una parola di codice è terminata.

Invece, considerando il codice 3:

$$\begin{aligned}
R_3 &= \sum_{i=1}^N L_i \cdot p_i \\
&= \sum_{i=1}^5 L_i \cdot p_i \\
&= L_1 \cdot p_1 + L_2 \cdot p_2 + L_3 \cdot p_3 + L_4 \cdot p_4 + L_5 \cdot p_5 \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 4 \cdot \frac{1}{16} \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 2 \cdot \left(4 \cdot \frac{1}{16}\right) \\
&= \frac{30}{16} \\
&= \frac{15}{8}
\end{aligned} \tag{5.51}$$

$R_3 = H(X)$, perchè non si ha perdita di informazione se $R_n \geq H(X)$ e quindi si considera anche la condizione di uguaglianza tra R_n e $H(X)$.

Ma, il problema del codice 3 è quello dell'auto-sincronizzazione perchè nessuna parola di codice è prefisso di un'altra parola di codice.

Invece, considerando il codice 4:

$$\begin{aligned}
R_4 &= \sum_{i=1}^N L_i \cdot p_i \\
&= \sum_{i=1}^5 L_i \cdot p_i \\
&= L_1 \cdot p_1 + L_2 \cdot p_2 + L_3 \cdot p_3 + L_4 \cdot p_4 + L_5 \cdot p_5 \\
&= 2 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 2 \cdot \frac{1}{16} + 3 \cdot \frac{1}{16} \\
&= \frac{33}{16}
\end{aligned} \tag{5.52}$$

$R_4 \geq H(X)$, quindi R_4 è informativo, ma pone il problema dell'auto-sincronizzazione.

Consideriamo la sequenza 110110.

Questa sequenza, con il codice 4, può provenire sia da a_5a_5 che da $a_4a_2a_3$.

5.5.4 Proprietà del prefisso (prefix condition)

¹⁸

La proprietà del prefisso è una condizione necessaria e sufficiente affinché un codice (di sorgente) sia istantaneo ed univocamente decodificabile (cioè auto-sincronizzante).

Per applicare questa proprietà, nessuna parola di codice è prefisso di un'altra parola di codice.

¹⁸Slide del prof — Teoria dell'informazione — pag 12.2
Appunti di Damiano — pag 12.2
Slide — Teoria dell'informazione — pag 12.2
Appunti — 2025-03-21 — pag 6

5.5.5 Algoritmo di Huffman

¹⁹

Considerando l'esempio precedente:

Letter	Probability	Codewords			
		Code 1	Code 2	Code 3	Code 4
a_1	$p_1 = \frac{1}{2}$	1	1	0	00
a_2	$p_2 = \frac{1}{4}$	01	10	10	01
a_3	$p_3 = \frac{1}{8}$	001	100	110	10
a_4	$p_4 = \frac{1}{16}$	0001	1000	1110	11
a_5	$p_5 = \frac{1}{16}$	00001	10000	1111	110

il codice 3 implementa l'algoritmo di Huffman.

L'algoritmo di Huffman trasforma le sequenze di lunghezza fissa in sequenze di codice di lunghezza variabile.

L'idea è quella di rappresentare i simboli (o le sequenze) di informazione più probabili con le parole di codice più corte.

Per essere applicato, bisogna avere un codice ottimo in cui si ha questa relazione:

$$H(X) \leq R \leq H(X) + 1 \quad (5.53)$$

Se p_i :

$$p_i = \frac{1}{2^k} \quad (5.54)$$

dove k sono i bit, allora il codice si dice perfettamente ottimo perchè vale questa relazione:

$$R = H(X) \quad (5.55)$$

Ritornando all'esempio in figura, il codice 3 è perfettamente ottimo.

L'algoritmo di Huffman viene utilizzato negli standard JPEG e MPEG per la codifica video.

¹⁹Slide del prof — Teoria dell'informazione — pag 14.1
Appunti di Damiano — pag 14.1
Slide — Teoria dell'informazione — pag 14.1
Appunti — 2025-03-21 — pag 7

5.5.5.1 Algoritmo di Huffman: implementazione scalare e binaria

²⁰

L'algoritmo di Huffman, utilizzando lo schema a blocchi, è il seguente:

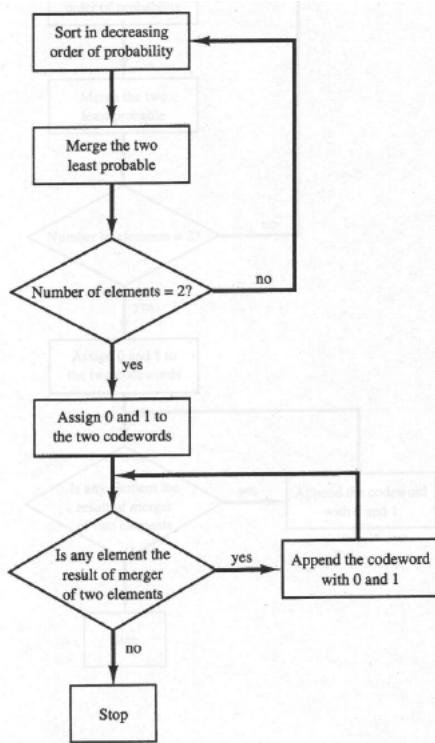


FIGURE 4.5. Huffman coding algorithm.

FIGURE 4.5. Huffman coding algorithm.

Dal punto di vista testuale, step-by-step, l'algoritmo di Huffman si implementa in questa maniera seguendo questi passaggi:

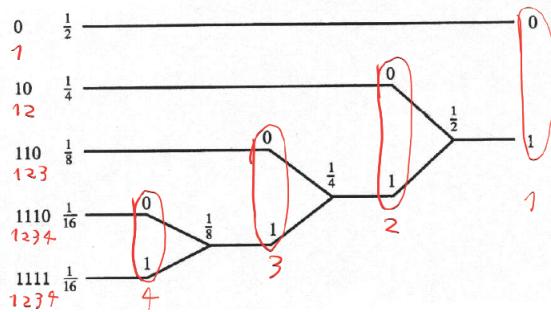
1. Si ordinano le probabilità p_i in modo decrescente (quindi dalla probabilità più alta a quella meno probabile)
2. Si uniscono i due simboli in ingresso meno probabili in un unico simbolo, la cui probabilità è la somma delle probabilità corrispondenti
3. Se il numero dei simboli si è ridotto a 2, si passa al punto successivo, altrimenti si ritorna al punto 1
4. Si assegnano (arbitrariamente) 0 e 1 ai due simboli rimanenti (da ciò deriva il famoso logaritmo in base 2)
5. Si percorrono le operazioni fatte in precedenza: ogni volta che si erano uniti due simboli, si associa 0 a uno dei due simboli e 1 all'altro
6. Le parole di codice risultano dalla concatenazione degli 0 e degli 1 così assegnati

²⁰Slide del prof — Teoria dell'informazione — pag 14.2 - 15
 Appunti di Damiano — pag 14.2 - 15
 Slide — Teoria dell'informazione — pag 14.2 - 15
 Appunti — 2025-03-24 — pag 2 - 3

Ritornando all'esempio della tabella in cui si confrontano i codici:

Letter	Probability	Codewords			
		Code 1	Code 2	Code 3	Code 4
a_1	$p_1 = \frac{1}{2}$	1	1	0	00
a_2	$p_2 = \frac{1}{4}$	01	10	10	01
a_3	$p_3 = \frac{1}{8}$	001	100	110	10
a_4	$p_4 = \frac{1}{16}$	0001	1000	1110	11
a_5	$p_5 = \frac{1}{16}$	00001	10000	1111	110

considerando il codice 3 in cui è stato implementato l'algoritmo di Huffman, possiamo visualizzare l'albero e le probabilità con lo schema seguente:



(dove, nella figura, sono segnati i percorsi che il numero ha svolto per far corrispondere a quella probabilità).

5.5.5.2 Algoritmo di Huffman: implementazione vettoriale

²¹

Si può implementare l'algoritmo di Huffman codificando n simboli alla volta.

In questo caso, R deve essere:

$$H(X) \leq R \leq H(X) + \frac{1}{n} \quad (5.56)$$

dove n sono i simboli codificati simultaneamente.

Possiamo considerare il limite in cui n tende a infinito.

In quel caso R deve essere:

$$R = H(X) \text{ per } n \rightarrow +\infty \quad (5.57)$$

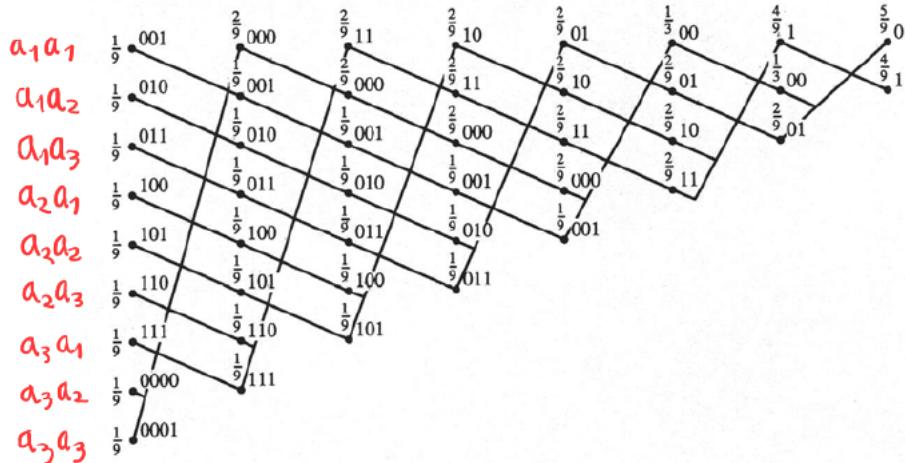
Consideriamo di avere la seguente codifica di simboli:

Simbolo	Codifica
a_1	000
a_2	001
a_3	010
a_4	011
a_5	100

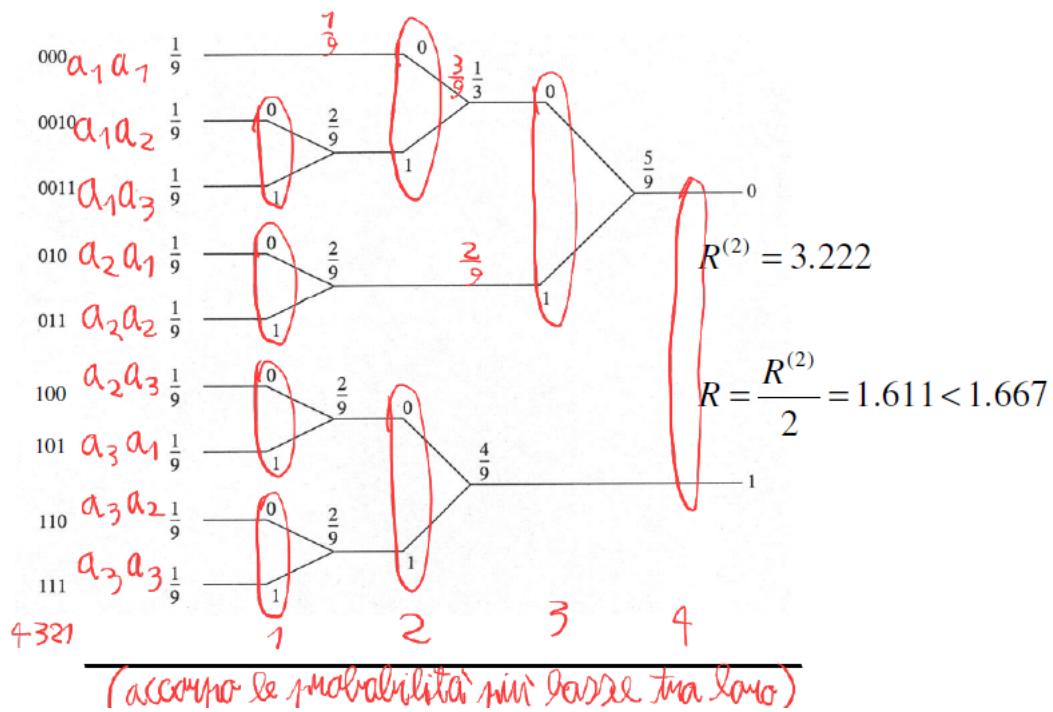
si possono implementare dei codici in cui si codificano due simboli alla volta, quindi n = 2.

²¹Slide del prof — Teoria dell'informazione — pag 16 - 18.1
 Appunti di Damiano — pag 16 - 18.1
 Slide — Teoria dell'informazione — pag 16 - 18.1
 Appunti — 2025-03-24 — pag 3 - 4

Di seguito delle rappresentazioni dell'algoritmo di Huffman con questa codifica:



oppure:



5.5.6 Algoritmo di Lempel-Ziv

²²

L'algoritmo di Lempel-Ziv trasforma sequenze di informazione di lunghezza variabile in sequenze di codice di lunghezza fissa.

Implementando una lunghezza fissa, l'algoritmo di Lempel-Ziv non ha problemi di sincronizzazione come Huffman.

Questo tipo di algoritmo è "duale" al codice di Huffman e cerca di risolvere i problemi di Huffman che sono i seguenti:

- occorre conoscere la distribuzione di probabilità
- l'implementazione vettoriale è complessa

Anche l'algoritmo di Lempel-Ziv è ottimo perchè le parole di codice verificano la proprietà del prefisso ed hanno lunghezza minima.

Si dice che l'algoritmo di Lempel-Ziv è un algoritmo di codifica di sorgente universale perchè non ha bisogno di conoscere la statistica della sorgente a priori.

Questo tipo di algoritmo è impiegato principalmente nella compressione dei file, ad esempio i file .zip .

²²Slide del prof — Teoria dell'informazione — pag 18.2
Appunti di Damiano — pag 18.2
Slide — Teoria dell'informazione — pag 18.2
Appunti — 2025-03-24 — pag 5

5.5.6.1 Algoritmo di Lempel-Ziv: esempio

²³

Si considera la sequenza:

$$0100001100001010000010100000110000010100001001001 \quad (5.58)$$

Il primo step è implementare un parsing, cioè si suddivide la stringa in piccoli frasi.

Cioè:

$$\begin{array}{l} 0100001100001010000010100000110000010100001001001 \\ \downarrow \\ 0, 1, 00, 001, 10, 000, 101, 0000, 01, 010, 00001, 100, 0001, 0100, 0010, 01001 \end{array} \quad (5.59)$$

Ogni frase è composta da una sequenza di 0 e 1 "nuova", ovvero è diversa ogni frase individuata precedentemente.

Scorrendo la sequenza in ingresso, una nuova frase è data dalla sequenza di lunghezza minima non ancora trovata in precedenza.

Ad esempio, la terza frase 00 è data dalla prima frase che è 0 più 0, la frase 001 è data dalla terza frase che è 00 più 1 e così via.

Il secondo step è memorizzare le frasi in una memoria indicizzata, come nel seguente caso:

TABLE 6.1 SUMMARY OF LEMPEL-ZIV EXAMPLE

Dictionary Location	Dictionary Contents	Codeword	
1	0001	0	0000 0
2	0010	1	0000 1
3	0011	00	0001 0
4	0100	001	0011 1
5	0101	10	0010 0
6	0110	000	0011 0
7	0111	101	0101 1
8	1000	0000	0110 0
9	1001	01	0001 1
10	1010	010	1001 0
11	1011	00001	1000 1
12	1100	100	0101 0
13	1101	0001	0110 1
14	1110	0100	1010 0
15	1111	0010	0100 0
16			1110 1

Il terzo step è quello della codifica, in cui la parola di codice associata ad ogni frase è:

- la concatenazione dell'indirizzo della sequenza d'informazione da cui la frase deriva
- la novità della frase

²³Slide del prof — Teoria dell'informazione — pag 19 - 21.1
 Appunti di Damiano — pag 19 - 21.1
 Slide — Teoria dell'informazione — pag 19 - 21.1
 Appunti — 2025-03-24 — pag 5 - 6

Devo ammettere che così spiegato l'algoritmo di Lempel-Ziv è spiegato con i piedi.

Ti lascio due video al riguardo.

<https://youtu.be/RV5aUr8sZD0?si=Fkt7DbrBK-WrSzUK>

Lossless Compression: Lempel-Ziv by Art of the Problem

<https://www.youtube.com/watch?v=Jqc418tQDkg>

Why the Lempel-Ziv algorithms are so dominant by Google for Developers

Se la sequenza di bit è molto breve, con l'algoritmo di Lempel-Ziv non si ha una compressione (anzi, si allunga la sequenza), ma diventa efficace quando si ha una sequenza molto lunga perchè le frasi si ripetono e si fanno riferimento alla frasi precedenti, fino ad arrivare a 0 e 1.

La lunghezza della parola di codice è determinata dalla scelta fatta per la dimensione della memoria.

Il processo per cui si prende nella codifica alcune locazioni che sono più richiamate e possono essere riutilizzate prende il nome di purging.

L'efficienza dell'algoritmo dipende dal purging: più viene utilizzato, e più l'algoritmo è efficiente.

5.6 Rate-distortion theory

²⁴

Da Shannon, sappiamo che se:

$$R < H(X) \quad (5.60)$$

sicuramente l'informazione sarà distorta.

Ma, nonostante ciò, possiamo minimizzarla: in alcuni casi non abbiamo memoria per avere R abbastanza grande (si pensi alla memoria di un microcontrollore).

Quindi, l'obiettivo per minimizzare la perdita di informazione è scegliere uno dei due apporci:

- per un dato R, minimizzare la distorsione D
- per una data distorsione D, minimizzare R

Come dice il detto: non puoi avere insieme la botte piena e la moglie ubriaca.

Bisogna trovare un compromesso, e molte delle volte è scegliere quanto perdere

²⁴Slide del prof — Teoria dell'informazione — pag 21.2
Appunti di Damiano — pag 21.2
Slide — Teoria dell'informazione — pag 21.2
Appunti — 2025-03-24 — pag 6 - 7

5.6.1 Misura della distorsione

²⁵

Consideriamo $x(t)$ il segnale "originale" e $\hat{x}(t)$ la rappresentazione di $x(t)$.

Si possono dare innumerevoli definizioni di distorsione D .

Di seguito alcune definizioni di D :

$$D = \max_t |x(t) - \hat{x}(t)| \quad (5.61)$$

$$D = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} |x(t) - \hat{x}(t)| dt \quad (5.62)$$

$$D = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} [x(t) - \hat{x}(t)]^2 dt \quad (5.63)$$

La distorsione D deve corrispondere alle proprietà "sensibili" in fase di rivelazione, e deve essere semplice da calcolare.

²⁵Slide del prof — Teoria dell'informazione — pag 22.1
 Appunti di Damiano — pag 22.1
 Slide — Teoria dell'informazione — pag 22.1
 Appunti — 2025-03-24 — pag 7

5.6.1.1 Distanza di Hamming

²⁶

La distorsione è una misura della distanza tra x e la sua rappresentazione \hat{x} .

Definiamo distanza di Hamming d_H tra x e \hat{x} se:

$$d_H(x, \hat{x}) = \begin{cases} 1 & \text{se } x \neq \hat{x} \\ 0 & \text{se } x = \hat{x} \end{cases} \quad (5.64)$$

Questo tipo distanza di Hamming viene definita come distanza hard perchè può assumere solo due valori.

La distanza di Hamming può avere più definizioni.

In questo corso vediamo questa qui, cioè la distanza di Hamming applicata al calcolo probabilistico.

Se vuoi farti una cultura sul perchè Hamming fu rivoluzionario, guardati questo video:

<https://www.youtube.com/watch?v=X8jsijhllIA&t=171s>

But what are Hamming codes? The origin of error correction by 3Blue1Brown

²⁶Slide del prof — Teoria dell'informazione — pag 22.2
 Appunti di Damiano — pag 22.2
 Slide — Teoria dell'informazione — pag 22.2
 Appunti — 2025-03-24 — pag 7

5.6.1.2 Distanza in media quadratica

²⁷

Possiamo definire una distanza d in media quadratica tra x e la sua rappresentazione \hat{x} come:

$$d(x, \hat{x}) = (x - \hat{x})^2 \quad (5.65)$$

In questo caso, $d(x, \hat{x})$ viene definita come distanza soft, perchè, a differenza della distanza hard dove assume solo due valori, questa distanza assume diversi valori, che, grazie alla definizione, tende a zero se la differenza tra x e \hat{x} è quasi nulla, la distanza aumenta se x e \hat{x} diventano sempre più diversi.

²⁷Slide del prof — Teoria dell'informazione — pag 23.1
 Appunti di Damiano — pag 23.1
 Slide — Teoria dell'informazione — pag 23.1
 Appunti — 2025-03-24 — pag 7

5.6.1.3 Distanza tra vettori

²⁸

Definiamo \mathbf{x}^n un vettore composto da n elementi e $\hat{\mathbf{x}}^n$ la rappresentazione di \mathbf{x}^n .

Possiamo definire la distanza d tra \mathbf{x}^n e $\hat{\mathbf{x}}^n$ come:

$$d(\mathbf{x}^n, \hat{\mathbf{x}}^n) = \frac{1}{n} \cdot \sum_{i=1}^n d(x_i, \hat{x}_i) \quad (5.66)$$

dove x_i e \hat{x}_i è il singolo elemento del vettore, rispettivamente, di \mathbf{x}^n e $\hat{\mathbf{x}}^n$.

Se non ti ricordi come è e cosa è un vettore, in informatica, si intendono questi:

index	0	1	2	3	4	5	6	7	8	9	n-1
	5	1	7	9	5	3	2	9	7	3	

$n=\text{dimensione}=10$

Nel caso probabilistico, possiamo definire la distanza D tra i vettori \mathbf{X}^n e $\hat{\mathbf{X}}^n$.

$$\begin{aligned} D &= \mathbb{E} [d(\mathbf{X}^n, \hat{\mathbf{X}}^n)] \\ &= \frac{1}{n} \cdot \sum_{i=1}^n \mathbb{E} [d(X_i, \hat{X}_i)]_{\text{stazionarietà}} \\ &= \mathbb{E} [d(X, \hat{X})] \end{aligned} \quad (5.67)$$

dove con la lettera \mathbb{E} si intende la media dei valori degli argomenti.

In questo caso, $\mathbb{E} [d(\mathbf{X}^n, \hat{\mathbf{X}}^n)]$ si intende la media delle distanze tra i vettori \mathbf{X}^n e $\hat{\mathbf{X}}^n$.

²⁸Slide del prof — Teoria dell'informazione — pag 23.2
Appunti di Damiano — pag 23.2
Slide — Teoria dell'informazione — pag 23.2
Appunti — 2025-03-24 — pag 8

5.6.1.4 Distanza nel caso di distanza d_H ²⁹

Nel caso di utilizzare la definizione di distanza di Hamming d_H :

$$d_H(x, \hat{x}) = \begin{cases} 1 & \text{se } x \neq \hat{x} \\ 0 & \text{se } x = \hat{x} \end{cases} \quad (5.68)$$

possiamo definire la distanza D tra X e \hat{X} come:

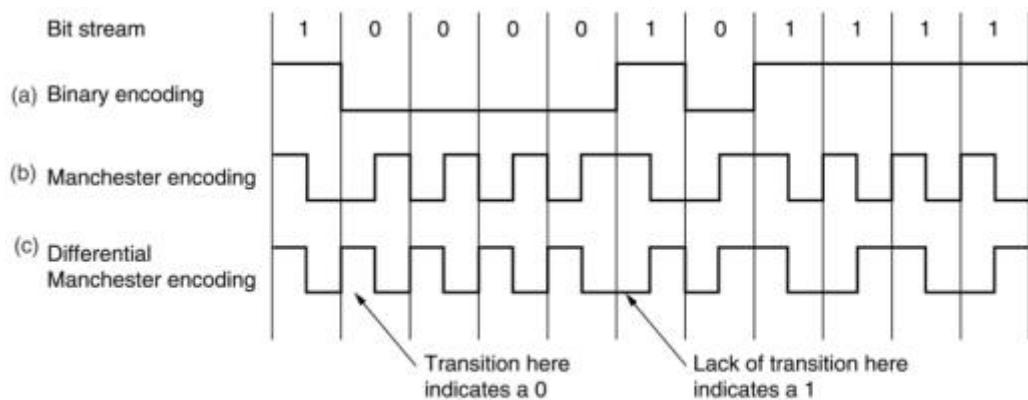
$$\begin{aligned} D &= \mathbb{E} [d(X, \hat{X})] \\ &\downarrow \\ D &= \mathbb{E} [d_H(X, \hat{X})] \\ &= 1 \cdot \Pr\{X \neq \hat{X}\} + 0 \cdot \Pr\{X = \hat{X}\} \\ &= \Pr\{X \neq \hat{X}\} + 0 \\ &= \Pr\{X \neq \hat{X}\} \\ &= \Pr\{\text{errore}\} \end{aligned} \quad (5.69)$$

Da questa dimostrazione, quando si utilizza la distanza di Hamming, la distorsione D coincide con la probabilità di errore $\Pr\{\text{errore}\}$.

²⁹Slide del prof — Teoria dell'informazione — pag 24.1
Appunti di Damiano — pag 24.1
Slide — Teoria dell'informazione — pag 24.1
Appunti — 2025-03-24 — pag 8

Capitolo 6

Codifica di linea



6.1 Cosa è la codifica di linea

1

In un mezzo trasmittivo, l'informazione è contenuta alle alte frequenze, perchè le basse frequenze, generalmente, vengono alterate dal mezzo trasmittivo stesso.

Data una sequenza di bit come la seguente:

$$0101100 \quad (6.1)$$

in cui ogni bit (ad esempio il primo bit della sequenza che è lo 0) verrà trasmesso in un tempo T_b .

Allora la frequenza di trasmissione è l'inverso di T_b , cioè f_0 :

$$f_0 = \frac{1}{T_b} \quad (6.2)$$

L'obiettivo di qualsiasi trasmissione in un mezzo trasmittivo è quello di inviare più bit utilizzando la larghezza di banda più piccola possibile (specialmente quando si tratta di inviare bit nelle comunicazioni wireless dove lo spettro è limitato e non bisogna interferire con gli altri utenti).

Inoltre, la scelta del tipo di forma d'onda da adottare per la trasmissione numerica può essere condizionata, oltre dalla necessità di controllare l'interferenza di intersimbolo (o ISI) e di minimizzare la probabilità di errore anche da diverse esigenze.

Riguardo alla probabilità di errore, è un argomento che verrà trattato in seguito in queste pagine.

Per quanto riguarda l'ISI, ne avevamo già discusso nel corso precedente di Teoria dei segnali.

Un piccolo ripasso:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 11 - Interferenza di intersimbolo - pag 101 - 108

Problemi tipici (ma non esclusivi) riguardo alla scelta del tipo di forma d'onda in un canale sono:

- necessità di minimizzare il contenuto spettrale del segnale alle basse frequenze
- necessità di semplificare l'operazione di sincronizzazione in ricezione

Il primo punto è banale perchè, essendo alle basse frequenze, lo spettro è decisamente limitato rispetto ad andare ad alta frequenza, dove il segnale, in frequenza, è molto più largo.

Il secondo punto non verrà affrontato in questo corso.

Il problema della sincronizzazione è un problema gigantesco da affrontare nella teoria dell'informazione.

L'unica cosa che dobbiamo sapere al momento è che, se non c'è una sincronizzazione tra gli apparati nelle trasmissioni digitali, gli zeri e uni che gli apparati ricevono vengono interpretati come simboli "a caso, senza senso, senza contenuto informativo" piuttosto che essere utili e informativi.

La procedura che conduce alla definizione della migliore forma d'onda per il particolare sistema o canale considerato, prende il nome di "codifica di canale".

¹Slide del prof — Codifica di linea — pag 1
Slide — Codifica di linea — pag 1
Appunti — 2025-03-25 — pag 2 - 3

6.1.1 Segnali unipolari: NRZ e RZ

²

Assumiamo come riferimento il caso di trasmissione di dati binari (0 e 1) di durata T.

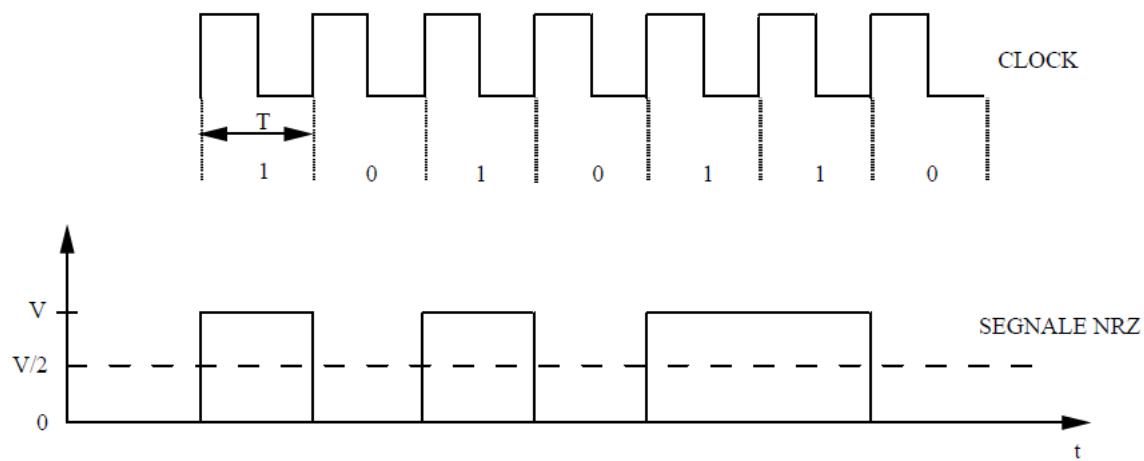
Una prima scelta che si impone è se trasmettere il segnale in forma unipolare o bipolare.

Nel caso unipolare, il segnale sarà costituito da una sequenza di livelli alti (che indicheremo con V o +V) e di livelli bassi (che indicheremo con 0), mentre nel caso bipolare comparirà, oltre al livello +V e 0, anche il livello -V.

Riferendoci al caso unipolare, sono peraltro possibili due soluzioni.

La prima soluzione di caso unipolare è denominata come NRZ (Non Return to zero) dove il singolo bit di informazione viene trasmesso utilizzando un impulso rettangolare, di durata T e di ampiezza V quando il bit vale 1; se il bit vale 0 viene trasmesso un segnale di ampiezza nulla di durata T.

Un esempio di segnale NRZ confrontato con il segnale di clock:



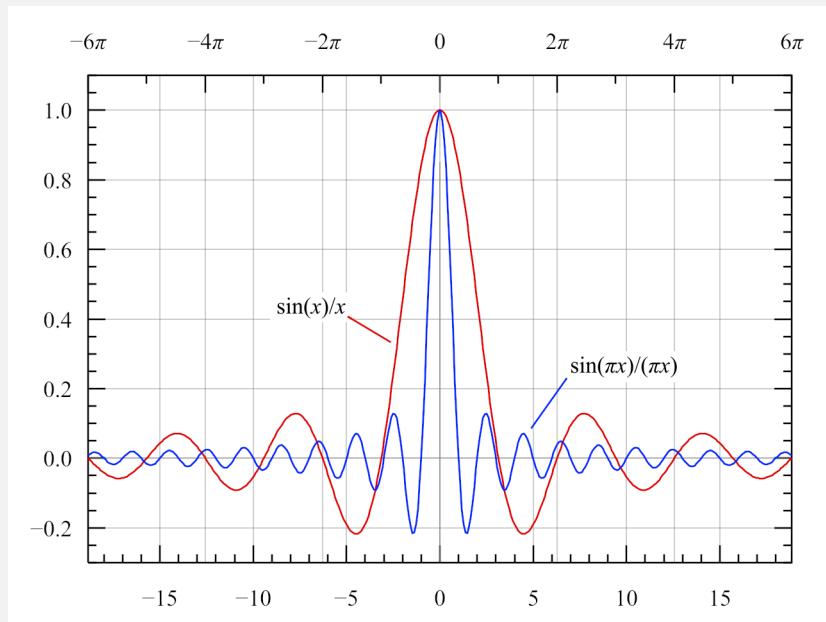
Come si può visualizzare dall'andamento del segnale NRZ e la sequenza di bit sotto al segnale di clock, si possono avere variazioni del segnale solo alla fine di un intero periodo di clock.

Questo fatto ha implicazioni sullo spettro di potenza del segnale stesso, che risulta, in questo caso, proporzionale a $\left[\frac{\sin(\pi f \cdot T)}{\pi \cdot f \cdot T}\right]^2$, oppure, detto a parole, è proporzionale a un sinc al quadrato.

²Slide del prof — Codifica di linea — pag 1 - 2
 Slide — Codifica di linea — pag 1 - 2
 Appunti — 2025-03-25 — pag 4

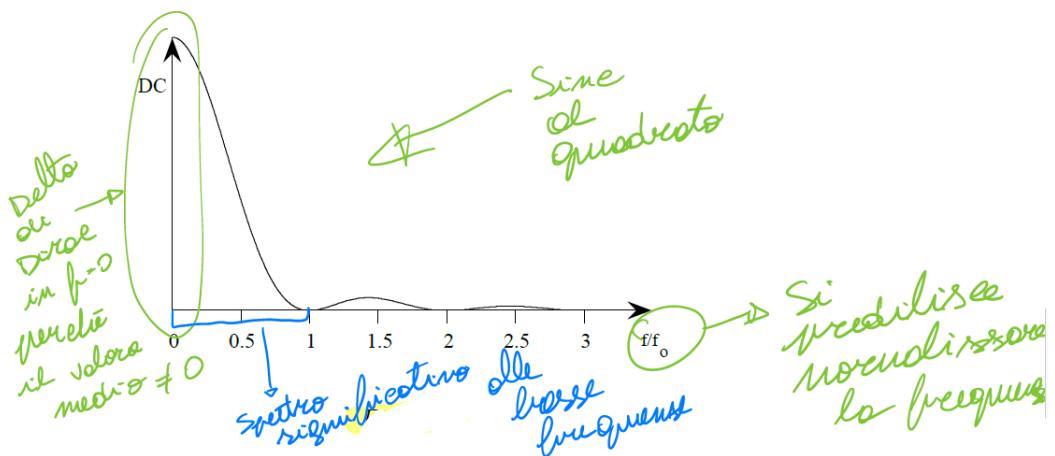
La funzione sinc la avevamo scoperta anche nel corso precedente.

Per rinfrescare le idee, questo è l'andamento di un sinc:
dove il sinc in blu è normalizzato, mentre il sinc in rosso è un sinc non normalizzato.



Per approfondire:
https://it.wikipedia.org/wiki/Funzione_sinc

Di seguito, lo spettro di potenza del segnale NRZ in frequenza normalizzato (con gli appunti presi a lezione):



Confrontando la figura del sinc e quella del sinc al quadrato, notiamo che gli andamenti negativi, grazie all'elevamento alla potenza, sono diventati positivi e, siccome si considerano andamenti con il tempo maggiore di zero, il sin al quadrato viene "tagliato" a metà, viene presa solo la metà a destra.

In figura si è posto:

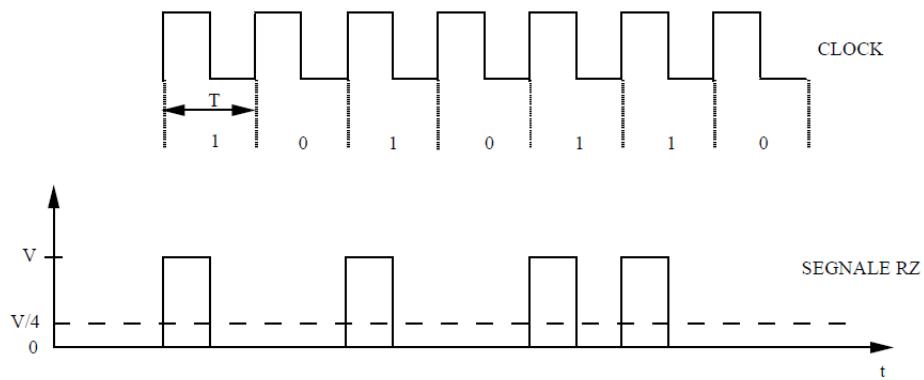
$$f_0 = \frac{1}{T} \quad (6.3)$$

Il lobo principale dello spettro in frequenza di un segnale NRZ ha spettro uguale a f_0 , e si hanno valori nulli in corrispondenza di multipli di f_0 , detta anche frequenza fondamentale.

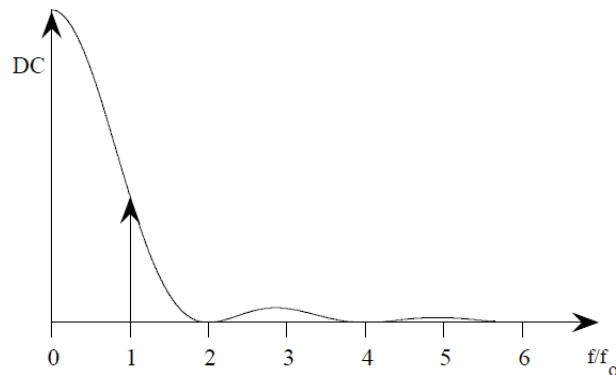
Inoltre è presente la Delta di Dirac nell'origine perché il segnale ha valore medio diverso da zero, che ha valore $\frac{V}{2}$.

Viene chiamata come NRZ (Non Return to Zero) perchè il bit 1 non diventa mai 0 [V].

Invece di seguito l'andamento nel tempo di una codifica RZ (Return to Zero) confrontato con il segnale di clock:



con il suo spettro di potenza in frequenza normalizzato:



Notando il segnale di RZ nel tempo, la trasmissione del livello V è limitata a mezzo intervallo di simbolo, cioè $\frac{T}{2}$.

In pratica, si può avere una transizione del segnale ogni $\frac{T}{2}$ secondi.

Lo spettro spettro di potenza, come il caso dell'NRZ, risultato proporzionale al sinc al quadrato, cioè $\left[\frac{\sin(\pi \cdot f \cdot T)}{\pi \cdot f \cdot T} \right]^2$, e ciò ha conseguenza un allargamento del lobo principale.

Molto importante è la relazione tra tempo e frequenza.

Sembra banale, ma ripassiamola al momento.

Se abbiamo un segnale nel tempo con transizioni lente, si avrà un andamento in frequenza molto stretto.

Se abbiamo un segnale nel tempo con transizione molto rapida, si avrà un andamento in frequenza molto largo.

Essendo la frequenza l'inverso del tempo, bisogna considerare che il comportamento del segnale nel tempo è inverso in frequenza.

Quindi, rispetto al caso NRZ, il lobo principale non si estende fino a f_0 , bensì fino a $2 \cdot f_0$.

I valori nulli sono presenti in valori multipli del lobo principale.

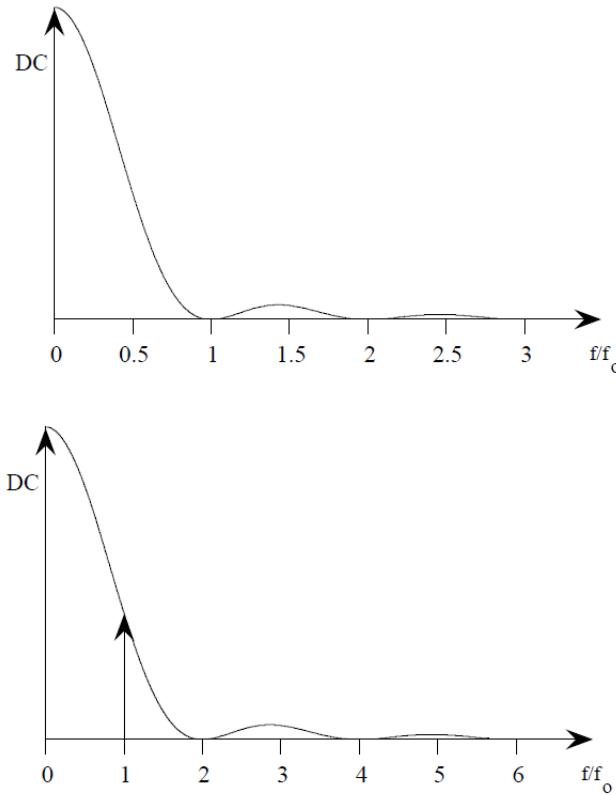
Essendo il valore medio del segnale NRZ nel tempo di $\frac{V}{4}$, avremo, in frequenza, una Delta di Dirac in continua.

Inoltre, sono presenti delle Delté di Dirac in presenza di frequenze dispare multipli di f_0 (in figura non sono mostrati, ma è presente una Delta di Dirac anche in $\frac{f}{f_0} = 3$ e in $\frac{f}{f_0} = 5$).

6.1.1.1 Confronto spettro di potenza tra NRZ e RZ

³

Dal confronto del segnale dello spettro di frequenza dell'NRZ e dell'RZ:



dove, la figura in alto è lo spettro di potenza di un segnale NRZ, e in basso è lo spettro di potenza di un segnale RZ.

Si nota che, nella soluzione NRZ, non si ha potenza alla frequenza f_0 , a differenza dello spettro dell'RZ, in cui in f_0 è presente una Delta di Dirac.

Questa osservazione è importante perché consente al ricevitore di concludere che la frequenza di clock può essere recuperata dal segnale trasmesso nel caso RZ (utilizzando un filtro passa-banda molto stretto limitato a f_0), mentre non è recuperabile nel caso NRZ.

La disponibilità della frequenza di clock (vale a dire della frequenza di trasmissione dei simboli binari) è importante in ricezione per poter sincronizzare gli apparati di rivelazione.

Per quanto detto, utilizzando lo schema NRZ, la frequenza di clock deve essere rigenerata localmente (quindi, dal punto di vista economico, bisogna avere un ricevitore con un oscillatore locale e tutti i componenti che gli stanno intorno), invece nel caso RZ può essere estratta dal segnale ricevuto (realizzare un filtro è nettamente più economico che aggiungere un oscillatore locale, anche perché quest'ultimo è instabile e va termostatato, quindi un casino).

Ricordiamo uno tra gli obiettivi della codifica di linea, che è quello di semplificare l'operazione di sincronizzazione in ricezione: per questo motivo, la codifica RZ risulta normalmente preferibile rispetto a quella NRZ.

³Slide del prof — Codifica di linea — pag 2
Slide — Codifica di linea — pag 2
Appunti — 2025-03-25 — pag 5

6.1.2 Codifica AMI

4

Sia nel caso NRZ che in quello RZ, il segnale ha un contenuto spettrale significativo alle basse frequenze (e in particolare un valore medio diverso da zero).

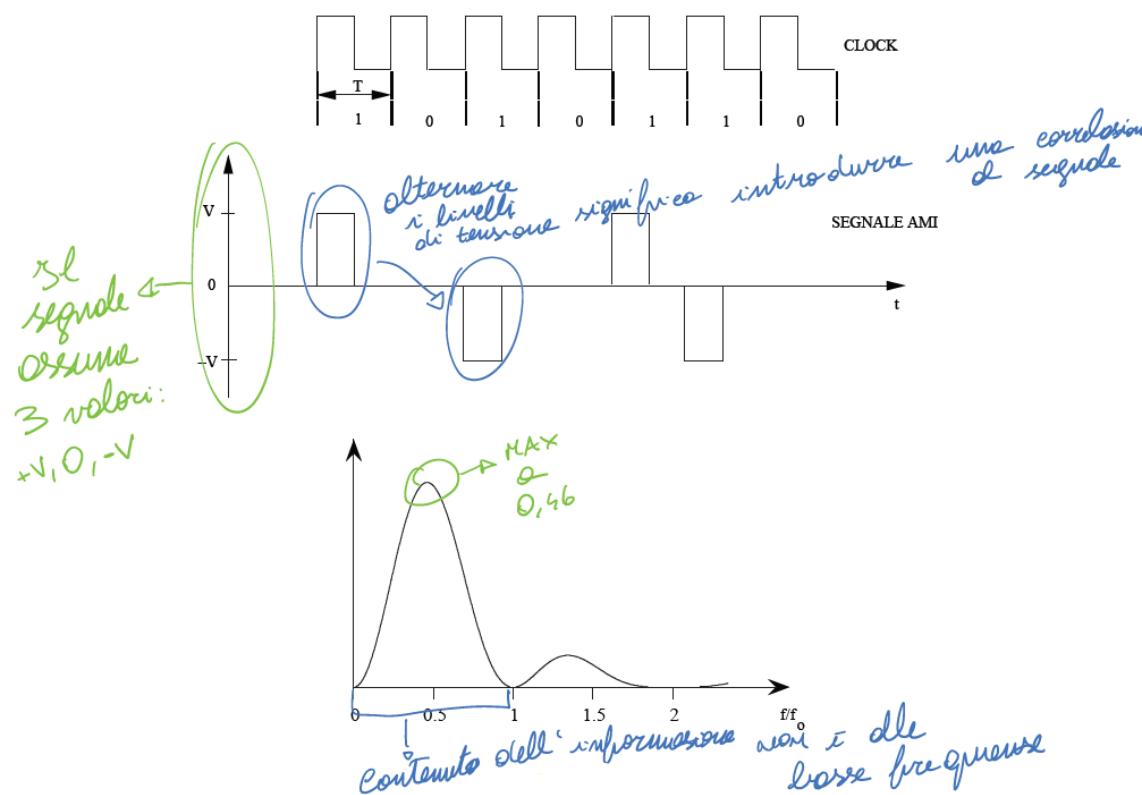
Ciò risulta svantaggioso in pratica.

Infatti, la componente continua, e più in generale le basse frequenze, vengono quasi sempre bloccate o distorte (si pensi ad un semplice condensatore presente in un circuito che, idealmente, blocca la continua).

Molte sono le tecniche di codifica utilizzabili per spostare lo spettro di potenza a frequenze più elevate.

Un esempio di tecnica senza la componente continua è la codifica AMI (Alternate Mask Inversion).

Un esempio di andamento di un segnale AMI nel tempo e del suo spettro di potenza in frequenza normalizzato a f_0 è il seguente (con le note):



Come si visualizza dalle seguenti figure, nella codifica AMI il simbolo 1 (ma la scelta è chiaramente convenzionale) viene alternativamente trasmesso il livello $+V$ ed il livello $-V$ nel tempo.

Inoltre, il segnale AMI è un segnale a 3 livelli: $+V$, $-V$ e 0.

Come la codifica RZ, il segnale AMI può avere una variazione ogni $\frac{T}{2}$ secondi.

Inoltre, lo spettro di potenza in frequenza risulta diverso rispetto alla codifica NRZ e RZ.

L'alternanza nel tempo di $+V$ e $-V$ fa sì che il valore medio del segnale sia nullo (quindi non è presente una Delta di Dirac in continua), inoltre il lobo principale è centrato su $0.46 \cdot f_0$ (cioè circa $\frac{f_0}{2}$).

⁴Slide del prof — Codifica di linea — pag 3 - 4
Slide — Codifica di linea — pag 3 - 4
Appunti — 2025-03-25 — pag 5 - 6

L'apparente mancanza della componente a frequenza f_0 può essere compensata semplicemente invertendo gli impulsi negativi nel tempo in ricezione, in tal modo ottenendo di nuovo il segnale in codifica RZ.

Così facendo, lungo il canale si trasmette il codice AMI (con i vantaggi che ciò comporta), mentre in ricezione, dove non si hanno più problemi di distorsione delle componenti a bassa frequenza, si utilizza il segnale RZ, con ciò semplificando l'operazione di sincronizzazione.

Un'altra caratteristica interessante del codice AMI è che si tratta di un codice capace di rivelare la presenza di errori.

Se il rumore sovrapposto al segnale trasforma un livello 0 in un livello $+V$ o $-V$, si verifica certamente una violazione del codice (si avranno infatti due livelli $+V$ o due livelli $-V$ consecutivi), che consente di identificare la presenza dell'errore.

La stessa considerazione vale se il rumore trasforma un livello $+V$ o $-V$ in un livello 0.

Se il ricevitore ha solo la codifica AMI, il ricevitore può capire (o rivelare) se c'è stato un errore, ma (con le conoscenze che sappiamo adesso) non sa correggerlo, quindi, l'unica cosa che può fare il ricevitore, è richiedere la ritrasmissione del segnale.

6.1.3 Codifica HBDn

5

Nello schema RZ, il recupero del sincronismo è possibile solo a condizione che non si abbiano in trasmissione lunghe sequenze di zeri.

Laddove si verifichi quest'ultima eventualità, infatti, il segnale trasmesso rimane costante per un lungo periodo di tempo e, non avendosi variazioni in ricezione, la frequenza f_0 , non può essere rivelata.

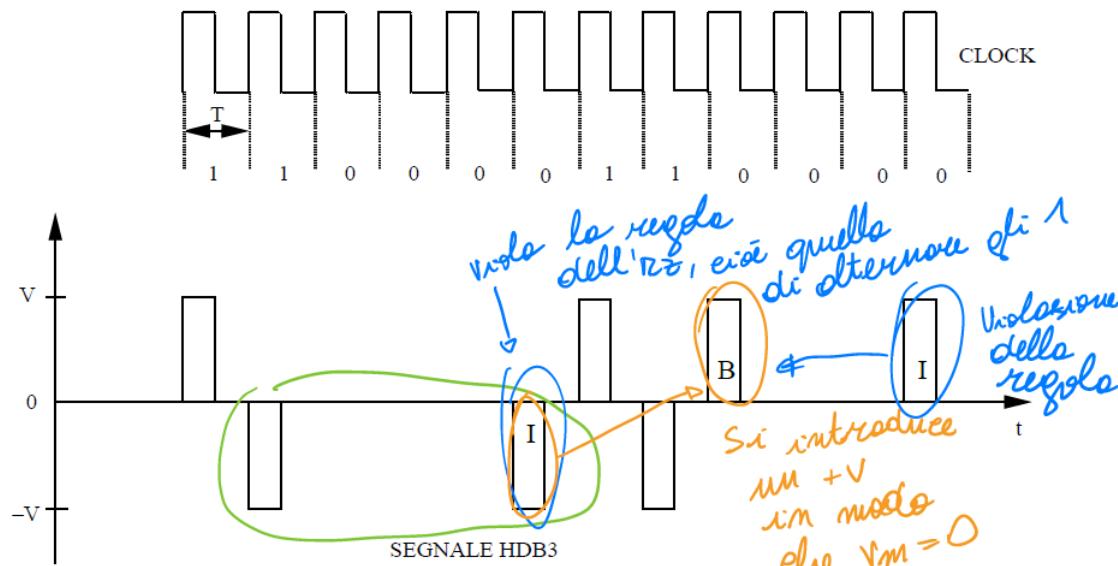
Questo importante inconveniente può essere compensato ricorrendo ad altri schemi di codifica di linea.

Il più noto tra tutti è lo schema HDBn (High Density Bipolar di ordine n), dove n è un numero intero.

Utilizzando per la trasmissione del segnale numerico questo codice di linea, non è possibile avere più di n simboli nulli consecutivi nella sequenza di segnale.

L'eventuale $(n + 1)$ -esimo livello 0, viene infatti sostituito "artificialmente" con un livello $\pm V$, ove la scelta del segno viene fatta violando intenzionalmente la regola dell'alternanza propria del codice AMI, per rendere possibile, in assenza di errori, l'identificazione.

Un esempio di un andamento nel tempo di codice HDB3 (cioè $n = 3$, quindi non si possono avere più di 3 simboli nulli) è il seguente (con gli appunti presi a lezione):



La figura della HDB3 consente anche di specificare meglio le caratteristiche del codice in cui ogni violazione (indicate con 1 in figura) deve avvenire con cambiamento di polarità.

Se ciò non è possibile "naturalmente", allora il primo zero nella sequenza di quattro simboli nulli deve, a sua volta, essere trasformato in un 1.

In figura, il livello contrassegnato con B è stato forzato a $+V$, pur corrispondendo ad un simbolo 0, proprio per avere l'inversione di polarità nell'ultimo livello della seconda sequenza.

Così facendo, il valore medio del segnale resta circa uguale a zero, e, globalmente, lo spettro in frequenza della potenza del codice HDB3 è pressoché analogo al codice AMI.

⁵Slide del prof — Codifica di linea — pag 4
Slide — Codifica di linea — pag 4
Appunti — 2025-03-25 — pag 6

6.1.4 Altri codici di linea

⁶

Codici di linea più complessi, come quelli della famiglia mBnT, trasformano m-uple di cifre binarie in ingresso in n-uple di cifre ternarie da trasmettere utilizzando tre livelli sulla linea, di cui anche qui uno a 0 e gli altri due con valore $\pm V$.

I codici della famiglia mBnT (come puri quelli HDBn) introducono, generalmente, una penalizzazione dal punto di vista della potenza necessaria per avere una certa quantità di trasmissione.

Per questo motivo, il loro uso è spesso combinato con quello di opportuni codici di canale.

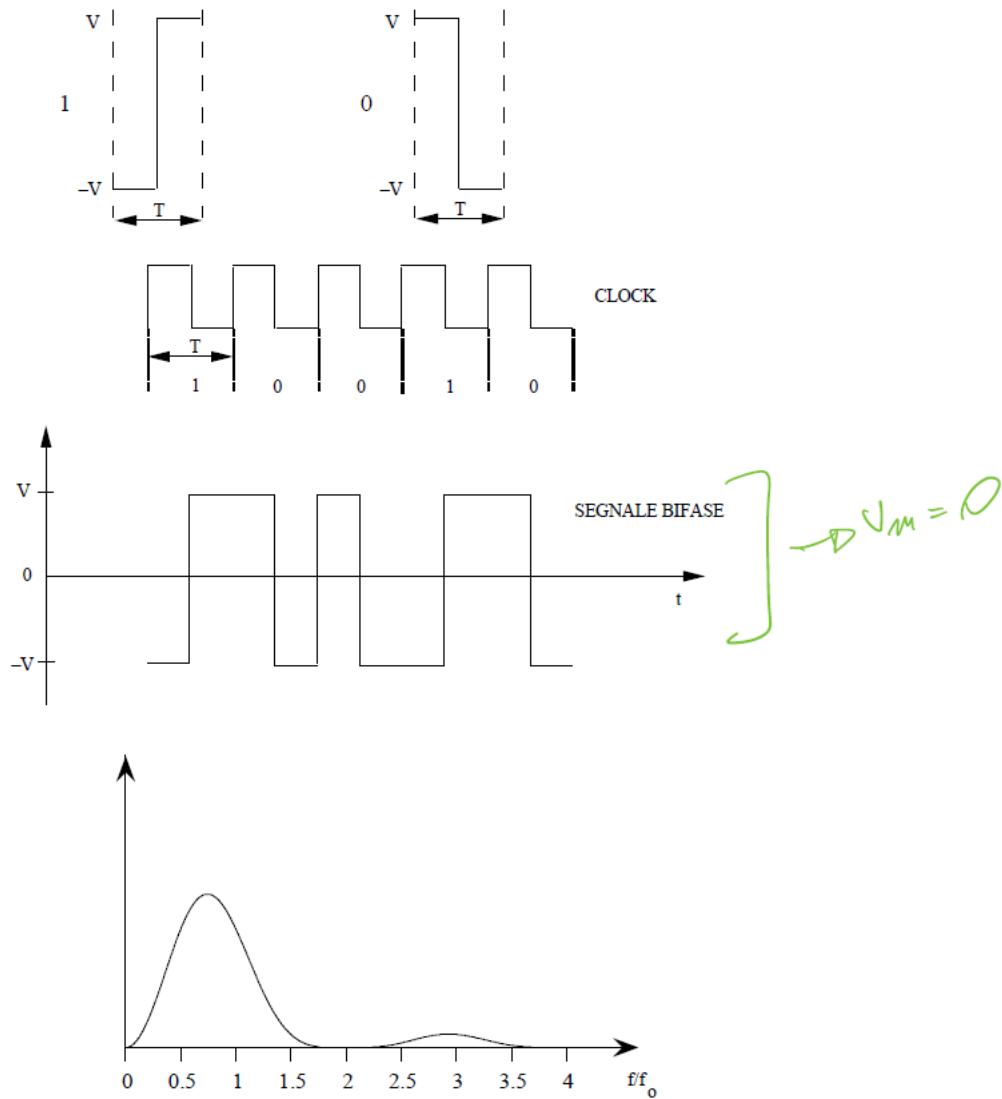
Allo stesso tempo, però, l'occupazione spettrale che essi richiedono, quando n è minore di m, è di norma minore, e quindi il loro uso è consigliato nelle applicazioni in cui è fondamentale limitare la banda occupata.

Una diversa "filosofia" nell'ottica di risolvere il problema delle lunghe sequenze di simboli nulli, consiste nell'assegnare a ciascun simbolo binario una forma d'onda che presenta comunque variazioni all'interno dell'intervallo T.

L'esempio più significativo è costituito dal "codice bifase" (comunemente noto come "codice Manchester").

⁶Slide del prof — Codifica di linea — pag 5 - 6
Slide — Codifica di linea — pag 5 - 6
Appunti — 2025-03-25 — pag 6 - 8

Come mostrato nella seguente figura:



nel codice Manchester si associa a ciascun bit una forma d'onda di valor medio nullo, di segno positivo o negativo a seconda del valore del bit.

Così facendo, il segnale presenta comunque una transizione a metà dell'intervallo di bit, e si ha transizione anche all'inizio dell'intervallo quando i due bit adiacenti sono identici.

Notando l'andamento dello spettro del codice Manchester, f_0 è certamente recuperabile, anche se, rispetto agli altri codice, l'assenza di una Delta di Dirac in f_0 rende un po' più sofisticato il recupero del sincronismo.

6.2 Cenni sul problema della sincronizzazione

⁷

Per una corretta ricostruzione dell'informazione al ricevitore, è necessaria la conoscenza della frequenza di simbolo (cioè il clock visto nelle sezioni precedenti).

L'inverso di quest'ultima, infatti, vale a dire la durata T di ciascun simbolo, determina la distanza tra due istanti di lettura e di decisione successivi.

Il problema viene qui accennato con riferimento alla trasmissione numerica in banda base, ma un problema analogo si pone ovviamente anche nelle trasmissioni in modulazione (che verranno trattate successivamente nel corso).

Come si è visto in precedenza, con alcune tecniche di codifica di linea, lo spettro del segnale trasmesso contiene una riga spettrale alla frequenza:

$$f_0 = \frac{1}{T} \quad (6.4)$$

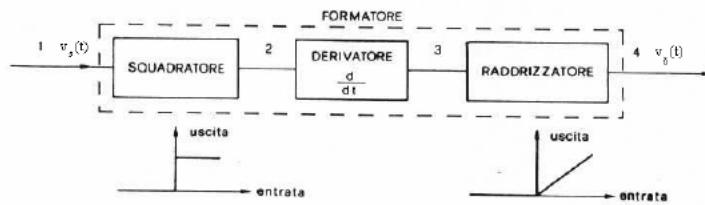
in cui f_0 è anche la frequenza di clock.

La frequenza di simbolo può essere recuperata inserendo un filtro passa-banda a banda stretta centrato sulla frequenza $\frac{1}{T}$.

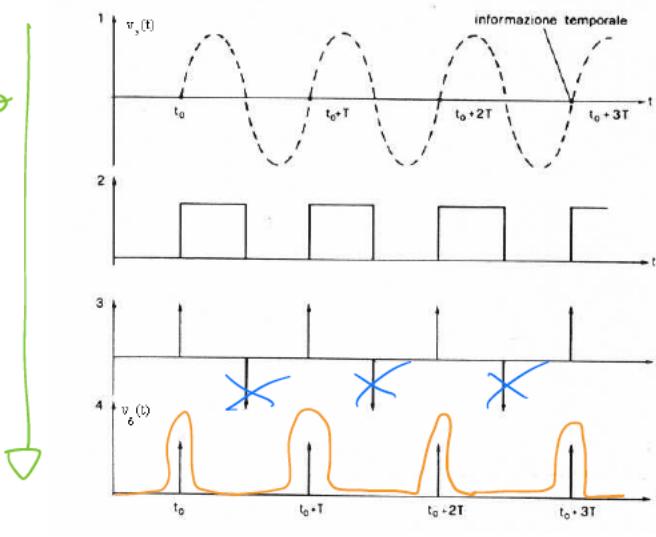
All'uscita del filtro si ottiene un'onda sinusoidale da cui è possibile ricavare gli impulsi di temporizzazione desiderati.

⁷Slide del prof — Codifica di linea — pag 7 - 8
 Slide — Codifica di linea — pag 7 - 8
 Appunti — 2025-03-25 — pag 8

Lo schema di principio utilizzabile è il seguente:



Processo IDEALE



- segnali reali, simile alle onde d'onda, ma che non sono

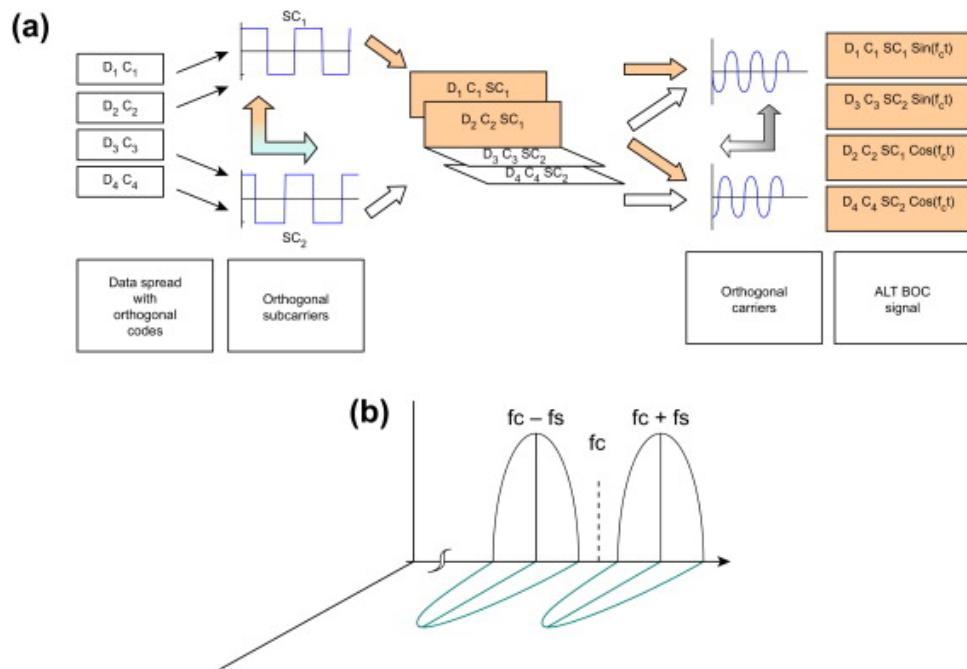
Se vuoi scoprire e andare a fondo su come è realizzato questo tipo di filtro, ti rimando al corso di Misure Elettroniche in cui è spiegato il blocco FI Formatore di impulsi.

<https://github.com/ciccio25/appunti-misure-elettroniche>
Capitolo 12.3 - Frequenzimetro - blocco FI - pag 232 - 235

Nella pratica, realizzando il circuito di estrazione del sincronismo con componenti reali, esso potrà essere soltanto approssimato.

Capitolo 7

Formati di trasmissione numerica



7.1 Procedura di ortogonalizzazione di Gram-Schmidt¹

1

Andando indietro al primo semestre del primo anno di ingegneria elettronica, precisamente al corso di Algebra lineare, abbiamo già visto la procedura di ortogonalizzazione di Gram-Schmidt per costruirsi una base vettoriale, cioè lo span dei vettori.

Un piccolo ripasso al volo:

<https://www.khanacademy.org/math/linear-algebra/alternate-bases/orthonormal-basis/v/linear-algebra-the-gram-schmidt-process>

Ma come ci siamo finiti a parlare di segnali come vettori che costituiscono una base ortogonale?

All'inizio sembra strano questo concetto, cioè associare segnali che variano nel tempo a vettori, ma poi vedrete che, come la trasformata di Fourier studiata nel corso di Teoria dei Segnali, ci renderà la vita, e specialmente i conti ed i calcoli, più semplice.

Nei vari esempi di rappresentazione di segnali che si è dovuto menzionare, in questo ed altri corsi, si è spesso sottolineata, implicitamente o esplicitamente, l'importanza di poter disporre di una base ortogonale.

Ripasso da Algebra lineare sotto tre punti di vista:

- <https://youtu.be/k7RM-ot2NWy?si=Iv-U2I81ohPvKAaX>
Linear combinations, span, and basis vectors — Chapter 2, Essence of linear algebra by 3Blue1Brown
Per capire il concetto, grazie alle animazioni, di cosa si intende per una base vettoriale, i.e. Span
- <https://www.andreaminini.org/matematica/spazio-vettoriale/como-trasformare-una-base-in-ortogonale>
Come trasformare una base in ortogonale by Andrea Minini
- <https://youtu.be/y6Z2EXB2kWQ?si=04qAQL9ZaQV7eSSp>
Base ortogonale , base ortonormale , vettori ortogonali. Ortonormalizzazione di Gram - Schmidt by Salvo Romeo (cioè quella persona che mi salvò e mi fece capire quanto è "facile" Algebra Lineare)

Alla fine stiamo parlando dello stesso argomento sotto tre punti di vista differenti.

L'esempio più noto, in questo senso, si rinviene nel classico sviluppo in serie di Fourier, in cui le funzioni espansione, esponenziali complesse e cosinusoidali, sono appunto ortogonali.

¹Slide del prof — Formati di trasmissione numerica — pag 1
Appunti di Damiano — pag 1
Slide — Formati di trasmissione numerica — pag 1
Appunti — 2025-03-25 — pag 9

Da <https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>
 Capitolo 2.1 - Sviluppo in serie di Fourier - pag 10 - 11

Se il segnale $s(t)$ è periodico con periodo T e pulsazione fondamentale ω_0 , possiamo definire una rappresentazione nota come Sviluppo in serie di Fourier.

Il dominio non sarà il tempo, bensì la frequenza (cioè l'inverso del tempo).

$s(t)$ può essere visto come, grazie alla formula di sintesi dello sviluppo in serie di Fourier:

$$s(t) = \sum_{k=-\infty}^{\infty} C_k \frac{1}{\sqrt{T}} e^{j\kappa\omega_0 t} \quad (7.1)$$

in cui:

$$\omega_0 = \frac{2\pi}{T} \quad (7.2)$$

e C_k viene calcolato dall'analisi dello sviluppo in serie di Fourier:

$$C_k = \frac{1}{\sqrt{T}} \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t) e^{-j\kappa\omega_0 t} dt \quad (7.3)$$

Ricordando, dalla matematica:

$$j^2 = -1 \quad (7.4)$$

$e^{j\kappa\omega_0 t}$ rappresenta una sinusode e ponendo:

$$\omega_k = k\omega_0 \quad (7.5)$$

possiamo dire che qualsiasi segnale $s(t)$ può essere visto come una successione di sinusoidi di diversa pulsazione.

La procedura di ortogonalizzazione di Gram-Schmidt può essere utilizzata per generare un insieme di segnali ortonormali, a partire da un set iniziale di segnali generici.

7.1.1 Due segnali ortogonali e ortonormali

²

Per chiarezza, ricordando che due segnali $s_1(t)$ e $s_2(t)$ sono ortogonali se verificano la seguenti proprietà:

$$\int s_1(t) \cdot s_2^*(t) dt = 0 \quad (7.6)$$

dove:

- per $s_2^*(t)$ si intende il complesso coniugato del segnale $s_2(t)$
- l'integrale è esteso al dominio comune di definizione tra $s_1(t)$ e $s_2(t)$
- nel caso più generale, si può assumere un intervallo di integrazione infinito, cioè con $t \in (-\infty, +\infty)$

Dal punto di vista dell'algebra lineare, per vettore complesso coniugato si intende lo stesso vettore con lo stesso modulo, ma fase opposta.

[https://www.andreaminini.org/mateematica/numeri-complessi/
numero-complesso-coniugato](https://www.andreaminini.org/mateematica/numeri-complessi/numero-complesso-coniugato)

Invece dal punto di vista di calcolo pratico, verificare l'ortogonalità tra $s_1(t)$ e $s_2(t)$ significa che dobbiamo considerare $s_2^*(t)$ con segno opposto di $s_2(t)$, poi moltiplicare $s_2^*(t)$ per $s_1(t)$ e calcolarci l'area (i.e. l'integrale).

(Un matematico è morto appena ho scritto che calcolare l'integrale significa calcolare l'area sotto ad una funzione. 1 minuto di silenzio. Riposi in pace)

Nelle dispense, l'integrale di ortogonalizzazione è scritto con $s_2(t)$ senza l'asterisco, cioè non si specifica che si intende il suo coniugato: è stato scritto senza per semplicità di notazione, ma ricordatevelo di metterlo.

I segnali $s_1(t)$ e $s_2(t)$ non solo possono essere ortogonali, ma possono essere ortonormali se ciascuno di essi è caratterizzato da un'energia E_i :

$$E_i = \int s_i^2(t) dt \text{ per } i = 1, 2 \quad (7.7)$$

che vale 1.

²Slide del prof — Formati di trasmissione numerica — pag 1
 Appunti di Damiano — pag 1
 Slide — Formati di trasmissione numerica — pag 1
 Appunti — 2025-03-25 — pag 9

7.1.2 Segnali ortogonali e ortonormali

3

Sia dunque dato un insieme M forme d'onda $s_m(t)$ con $1 \leq m \leq M$.

Ipotizzando che queste forme d'onda siano utilizzate per la trasmissione dell'informazione in un sistema di comunicazione, vogliamo costruire un insieme di $N \leq M$ forme d'onda ortonormali, che costituiranno la base per lo spazio di dimensione N , in cui potranno essere rappresentati, insieme ad altri, gli M segnali di partenza.

In accordo con la procedura di ortogonalizzazione di Gram-Schmidt, si assume la prima forma d'onda $s_1(t)$ e si costruisce il primo elemento della base ortonormale come $\Psi_1(t)$:

$$\Psi_1(t) = \frac{s_1(t)}{\sqrt{E_1}} \quad (7.8)$$

L'energia di $\Psi_1(t)$ è ovviamente unitaria.

La dimostrazione che l'energia di $\Psi_1(t)$ è unitaria, è la seguente.

Considerando $E_{\Psi_1(t)}$ l'energia di $\Psi_1(t)$, possiamo calcolarla come:

$$\begin{aligned} E_{\Psi_1(t)} &= \int [\Psi_1(t)]^2 dt \\ &= \int \left(\frac{s_1(t)}{\sqrt{E_1}} \right)^2 dt \\ &= \int \frac{s_1(t)^2}{E_1} dt \\ &= \frac{1}{E_1} \cdot \int s_1(t)^2 dt \\ &= \frac{1}{E_1} \cdot E_1 \\ &= \frac{1}{1} \\ &= 1 \end{aligned} \quad (7.9)$$

(Un altro matematico è morto quando ha visto $\frac{1}{1}$ ma è solo per evidenziare che i due E_1 si semplificano)

Per valutare il secondo elemento della base, si calcola, innanzitutto, la seguente proprietà:

$$c_{21} = \int_{-\infty}^{+\infty} s_2(t) \cdot \Psi_1(t)^* dt \quad (7.10)$$

Ipotizziamo per adesso, nella formula di c_{21} che l'integrale si estenda per $[-\infty, +\infty]$

³Slide del prof — Formati di trasmissione numerica — pag 1 - 3
 Appunti di Damiano — pag 1 - 3
 Slide — Formati di trasmissione numerica — pag 1 - 3
 Appunti — 2025-03-25 — pag 9

c_{21} può essere interpretato come la proiezione del segnale $s_2(t)$ sul primo elemento della base, cioè $\Psi_1(t)$.

In pratica, c_{21} è quella parte di $s_2(t)$ che può essere rappresentata in termini della sola $\Psi_1(t)$.

Non a caso:

$$\begin{aligned} c_{21} &= 0 \\ \uparrow \\ s_2(t) &\perp \Psi_1(t) \end{aligned} \quad (7.11)$$

$$\int_{-\infty}^{+\infty} s_2 \cdot \Psi_1(t)^* dt = 0$$

E' ovvio, perchè $\Psi_1(t)$ si può esprimere come $\frac{s_1(t)}{\sqrt{E_1}}$ e considerando solo $\frac{1}{\sqrt{E_1}}$ un coefficiente moltiplicativo, si ritorna alla formula data nella sezione precedente per due segnali $s_1(t)$ e $s_2(t)$ ortogonali.

cioè per esteso, c_{21} è nullo quando $s_2(t)$ è ortogonale a $\Psi_1(t)$, il che comporta che $\Psi_1(t)$ non ha alcuna utilità dal punto di vista della rappresentazione di $s_2(t)$.

Da quanto procede, risulta chiaro il significato dell'operazione successiva, che consiste nel sottrarre a $s_2(t)$ il prodotto tra c_{21} e $\Psi_1(t)$.

Formalmente $d_2(t)$ vale:

$$d_2(t) = s_2(t) - c_{21} \cdot \Psi_1(t) \quad (7.12)$$

Se $c_{21} \cdot \Psi_1(t) = 0$, allora $d_2(t) = s_2(t)$, cioè abbiamo una nuova funzione di base

$d_2(t)$ è certamente ortogonale a $\Psi_1(t)$, avendo infatti:

$$\begin{aligned} \int_{-\infty}^{+\infty} d_2(t) \cdot \Psi_1(t) dt &= \int_{-\infty}^{+\infty} [s_2(t) - c_{21} \cdot \Psi_1(t)] \cdot \Psi_1(t) dt \\ &= \int_{-\infty}^{+\infty} s_2(t) \cdot \Psi_1(t) dt - c_{21} \cdot \int_{-\infty}^{+\infty} \Psi_1^2(t) dt \\ &= c_{21} - c_{21} \cdot 1 \\ &= c_{21} - c_{21} \\ &= 0 \end{aligned} \quad (7.13)$$

In questa equazione non sono stati messi i complessi coniugati, ma fai finta che ci siano.

$d_2(t)$ non ha energia unitaria.

Per costruire, a partire da esso, una funzione ortonormale, ottenendo in tal modo il secondo elemento della base, valutiamo la sua energia, definita convenzionalmente come E_2 :

$$E_2 = \int_{-\infty}^{+\infty} [d_2(t)]^2 dt \quad (7.14)$$

Si potrà allora scrivere $\Psi_2(t)$:

$$\Psi_2(t) = \frac{d_2(t)}{\sqrt{E_2}} \quad (7.15)$$

e questa è una ortogonale a $\Psi_1(t)$ ed ha energia unitaria.

Se:

$$d_2(t) \neq 0 \quad (7.16)$$

$\Psi_2(t)$ può essere assunta come secondo elemento della base.

La procedura descritta può essere generalizzata, ponendo:

$$\Psi_k(t) = \frac{d_k(t)}{\sqrt{E_k}} \quad (7.17)$$

In questa espressione si ha:

$$d_k(t) = s_k(t) - \sum_{i=1}^{k-1} c_{ki} \cdot \Psi_i(t) \quad (7.18)$$

e:

$$E_k = \int_{-\infty}^{+\infty} [d_k(t)]^2 \quad (7.19)$$

con:

$$c_{ki} = \int_{-\infty}^{+\infty} s_k \cdot \Psi_i dt \quad (7.20)$$

Confrontando la formula di $d_k(t)$ e di c_{ki} afferma qualitativamente che il contributo del segnale $s_k(t)$ alla costruzione della base ortonormale è dovuto alla parte del segnale che non è già rappresentabile utilizzando una combinazione lineare delle funzioni di base già determinate.

La procedura appena descritta va ripetuta per tutti i segnali $s_m(t)$, in cui $1 \leq m \leq M$.

La cardinalità N della base (ovvero il numero degli elementi che la costituiscono) sarà uguale a M (la cardinalità dell'insieme di partenza) se gli M segnali $s_m(t)$ sono tra loro linearmente indipendenti: vale a dire che nessuno di essi può essere ottenuto per combinazione lineare degli altri.

Ciò è sicuramente vero quando gli M segnali sono tra loro ortogonali.

Se tutti i segnali sono ortogonali, si ha sempre N = M.

7.1.3 Segnali come combinazione lineare della base N-dimensione

4

Una volta costituito il set di forme d'onda ortogonali, cioè l'insieme $\{\Psi_k(t)\}$, gli M segnali $s_m(t)$ possono essere espressi come combinazione lineare delle funzioni della base, in accordo con la seguente formula generale:

$$s_m(t) = \sum_{k=1}^N s_{mk} \cdot \Psi_k(t) \quad (7.21)$$

Come nella rappresentazione in Serie di Fourier in cui si rappresentano i segnali con i loro coefficienti, anche qui si rappresentano i segnali con le loro basi $\Psi_k(t)$ finite (con k che parte da 1 e va fino a N funzioni di base)

L'orto-normalità della base rende semplice il calcolo dei coefficienti dell'espansione, s_{mk} , per i quali si avrà infatti:

$$s_{mk} = \int_{-\infty}^{+\infty} s_m(t) \cdot \Psi_k(t) dt \quad (7.22)$$

La medesima proprietà è poi alla base del calcolo dell'energia, in accordo con la formula seguente:

$$\begin{aligned} E_m &= \int_{-\infty}^{+\infty} [s_m(t)]^2 dt \\ &= \sum_{k=1}^N (s_{mk})^2 \end{aligned} \quad (7.23)$$

In accordo con la teoria generale della rappresentazione di un segnale secondo una base assegnata, le proprietà del segnale (ad esempio la sua energia), possono essere ricavate indifferentemente, dalla conoscenza diretta del segnale o dei parametri della sua rappresentazione.

Come avevamo ribadito nel corso di Teoria dei Segnali

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 2.1 - Sviluppo in serie di Fourier

Diverse volte è utile osservare lo stesso segnale in altre rappresentazioni equivalenti.

Per rappresentazione di un segnale si intende una qualsiasi modalità idonea alla sua individuazione biunivoca.

La rappresentazione può comportare un cambiamento del dominio di definizione.

Quindi, grazie ad una rappresentazione, si può passare da un dominio all'altro dello stesso segnale.

Dal punto di vista geometrico, la formula di $s_m(t)$ è evidente: le funzioni $\Psi_k(t)$ con $k = 1, 2, \dots, N$, possono essere riguardate come altrettanti versori che definiscono uno spazio N-dimensionale, sopra

⁴Slide del prof — Formati di trasmissione numerica — pag 4 - 6
 Appunti di Damiano — pag 4 - 6
 Slide — Formati di trasmissione numerica — pag 4 - 6
 Appunti — 2025-03-25 — pag 9 - 10

il quale la generica forma d'onda $s_m(t)$ è rappresentata tramite le sue componenti.

Ripasso al volo su cosa si intende per base vettoriale dal corso di Algebra Lineare e le sue proprietà:
<https://www.andreaminini.org/mathematica/spazio-vettoriale/base-algebra-lineare>

Raccolte in un vettore, possiamo scrivere le componenti della generica forma d'onda $s_m(t)$ come:

$$\vec{s}_m = (s_{m1}, s_{m2}, \dots, s_{mN}) \quad (7.24)$$

Piccola precisazione di notazione.

In tutti i corsi fatti in cui si ha a che fare con i vettori, ogni corso, quindi ogni professore, ha la propria simbologia per rappresentarli.

Il prof, nelle sue dispense, utilizza il grassetto per differenziare il vettore dalle sue componenti.

Io invece, personalmente, sono abituato a indicare il vettore con la freccia sopra, perché, quando scrivo a mano le notazioni matematiche (ad esempio durante uno scritto o durante l'orale) riesco a distinguere al volo cosa è un vettore e cosa no.

Fate come volete, l'importante è utilizzare una notazione che sia costante.

Nello spazio N-dimensionale identificato dalla base, il segnale $s_m(t)$ è rappresentato da un punto, le cui coordinate sono le componenti s_{mN} data dalla formula di \vec{s}_m .

L'utilizzo delle componenti consente anche di esprimere il prodotto intero tra due segnali come:

$$\begin{aligned} \int_{-\infty}^{+\infty} s_m(t) \cdot s_n(t) dt &= \vec{s}_m \cdot \vec{s}_n \\ &= \sum_{k=1}^N s_{mk} \cdot s_{nk} \end{aligned} \quad (7.25)$$

dove:

- per $\vec{s}_m \cdot \vec{s}_n$ si intende il prodotto scalare tra i vettori \vec{s}_m e \vec{s}_n
- per $\sum_{k=1}^N s_{mk} \cdot s_{nk}$ si intende la somma tra i singoli componenti dei due vettori \vec{s}_m e \vec{s}_n

Piccolo ripasso da Algebra lineare per cosa si intende prodotto scalare.

<https://www.youtube.com/watch?v=LyGKycYT2v0&t=30s>

Dot products and duality — Chapter 9, Essence of linear algebra by 3Blue1Brown

https://youtu.be/UqpY07EiwCU?si=Fj_7MjStUPMHjvSx

Dot Product (1 of 2: Geometric interpretation) by Eddie Woo

In definitiva, l'espansione in funzioni ortonormali è uno strumento assolutamente utile sia alla rappresentazione della classe di segnali considerata, sia ai fini della valutazioni delle loro proprietà, sia ai fini della loro elaborazione e combinazione.

In effetti, l'espansione in funzioni ortogonali è uno dei presupposti per la migliore descrizione dei sistemi di trasmissione numerici, ove i segnali che portano l'informazione sono rappresentati da forme d'onda di durata limitata T.

Infine, è opportuno osservare che, per un dato insieme di segnali $s_m(t)$, l'insieme delle forme d'onda che definisce la base ortonormale non è necessariamente unico.

L'importante è che il cambiamento di base non fa cambiare la dimensione dello spazio e l'energia di ciascun segnale o il prodotto intero tra una qualunque coppia di vettori.

In definitiva, l'algoritmo di ortogonalizzazione di Gram-Schmidt garantisce la generazione di una base ortonormale per la rappresentazione di un dato insieme di segnali.

Ma non per questo dobbiamo sempre applicare questo algoritmo per individuare una base ortonormale.

Detta in altre parole, se è la tua ultima spiaggia, utilizza l'algoritmo Gram-Schmidt, sennò, utilizza altri algoritmi.

7.2 Esempi di segnali in spazi N-dimensional

5

Assegnato un insieme di M segnali $s_m(t)$, in cui $1 \leq m \leq M$, è sempre possibile rappresentare ciascuno di essi in uno spazio di dimensione N individuato (e quindi descritto) da N forme d'onda ortonormali $\Psi_i(t)$, in cui $1 \leq i \leq N$.

In un sistema di comunicazione numerico, ciascun segnale $s_m(t)$, denominato simbolo, può essere associato alla trasmissione di k bit, dove k vale:

$$k = \log_2(M) \quad (7.26)$$

Visto che la sorgente emetterà una sequenza binaria, quindi di 0 e 1, ciascun bit avendo durata T_B , ciascun simbolo avrà durata pari a T , dove T vale:

$$T = k \cdot T_B \quad (7.27)$$

Le caratteristiche della trasmissione, dal punto di vista della qualità (cioè in presenza di rumore o altri disturbi sovrapposti), dell'occupazione spettrale e altre caratteristiche della trasmissione, dipendono, per un dato valore di M , dalla dimensione dello spazio e, per un dato N , dalla distribuzione dei segnali (loro punti rappresentativi) all'interno dello spazio.

⁵Slide del prof — Formati di trasmissione numerica — pag 8
 Appunti di Damiano — pag 8
 Slide — Formati di trasmissione numerica — pag 8
 Appunti — 2025-03-28 — pag 2

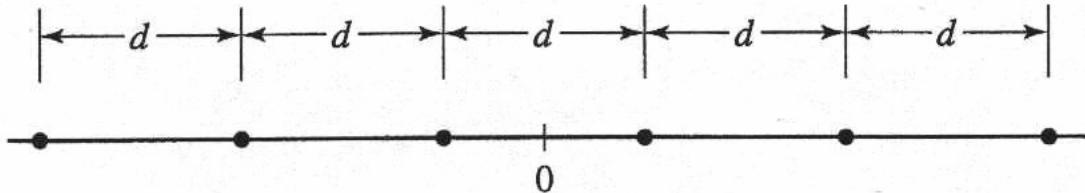
7.3 Segnali in spazi di dimensione $N = 1$

6

In questo caso, lo spazio è unidimensionale, ed è generato da una sola funzione espansione $\Psi_1(t)$.

I segnali $s_m(t)$ sono proporzionali a tale funzione.

Una possibile distribuzione dei punti (che può essere anche chiamata come costellazione) sull'asse che identifica lo spazio è la seguente:



Come si vede dalla figura, i punti sono distribuiti simmetricamente rispetto all'origine, a distanza d l'uno dall'altro.

Come funzione di base $\Psi_1(t)$ si può assumere:

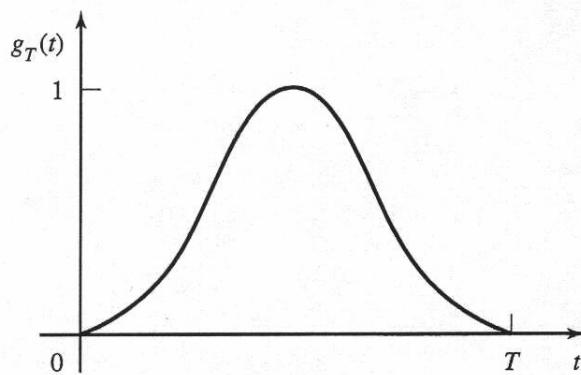
$$\Psi_1(t) = \frac{1}{\sqrt{E_g}} \cdot g_T(t) \text{ per } 0 \leq t \leq T \quad (7.28)$$

dove:

- $g_T(t)$ è una funzione generica
- E_g è l'energia di $g_T(t)$

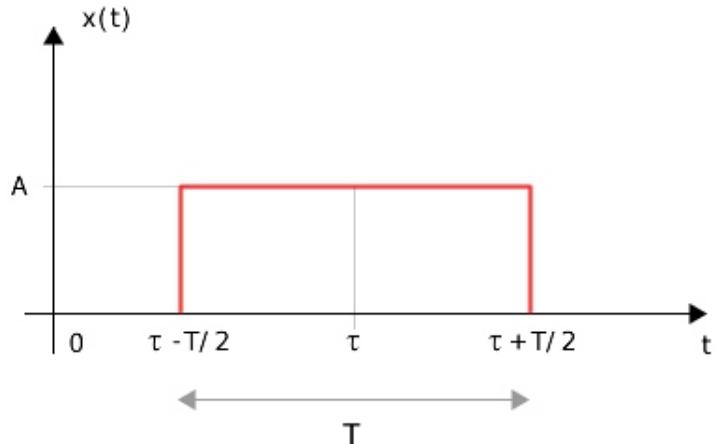
La forma di $g_T(t)$ può essere scelta con l'intento di ottenere, ad esempio, un'occupazione spettrale consona alla disponibilità di banda.

Ad esempio $g_T(t)$ può avere questo andamento:



⁶Slide del prof — Formati di trasmissione numerica — pag 8 - 10
 Appunti di Damiano — pag 8 - 10
 Slide — Formati di trasmissione numerica — pag 8 - 10
 Appunti — 2025-03-28 — pag 2 - 6
 Appunti — 2025-07-21 Ricevimento — pag 4

Ma, frequentemente, si assume per $g_T(t)$ un impulso rettangolare



WWW.ANDREAMININI.ORG

con andamento:

$$g_T(t) = \begin{cases} \sqrt{\frac{E_g}{T}} & \text{per } 0 \leq t \leq T \\ 0 & \text{altrove} \end{cases} \quad (7.29)$$

A rigore, quindi in un contesto reale di un sistema di telecomunicazioni, non si utilizzeranno gli impulsi rettangolari, perchè, dalle proprietà tra tempo e frequenza:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 3.4 - Altre caratteristiche sulla dualità tempo-frequenza - pag 24

- Un segnale $s(t)$ che è limitato nel tempo, ha spettro illimitato sull'asse delle pulsazioni
- Un segnale con spettro delle pulsazioni limitato, ha un'evoluzione temporale illimitata in t .

■

Quindi a rigore useremo $g_T(t)$ gaussiana vista precedentemente, ma siccome "telecomunicazioni" è "un corso di base", consideriamo gli impulsi rettangolari, anche perchè sono molto più semplici da trattare matematicamente con i calcoli.

Da quanto precede, si deduca che il generico segnale $s_m(t)$ può essere scritto come:

$$s_m(t) = s_m \cdot \Psi_1(t) \quad (7.30)$$

Sapendo che c'è una funzione di base $\Psi_1(t)$, i segnali $s_m(t)$ differiscono solo per l'ampiezza ed a quanto sono proporzionali da $\Psi_1(t)$, quindi significa anche essere proporzionali a $g_T(t)$.

Sapendo ciò, possiamo scrivere:

$$s_m(t) = A_m \cdot g_T(t) \text{ per } m = 1, 2, \dots, M \text{ e } 0 \leq t \leq T \quad (7.31)$$

Il coefficiente s_m si ottiene facilmente come segue:

$$\begin{aligned}
s_m &= \int_0^T s_m \cdot \Psi_1(t) dt \\
&= \int_0^T [A_m \cdot g_T(t)] \cdot \left[\frac{g_T(t)}{\sqrt{E_g}} \right] dt \\
&= \int_0^T A_m \cdot \frac{[g_T(t)]^2}{\sqrt{E_g}} dt \\
&= \frac{A_m}{\sqrt{E_g}} \cdot \int_0^T [g_T(t)]^2 dt \\
&= \frac{A_m}{\sqrt{E_g}} \cdot E_g \\
&= A_m \cdot \sqrt{E_g}
\end{aligned} \tag{7.32}$$

Per coefficiente s_m si intende un coefficiente della base ortogonale e ortonormale per rappresentare i segnali ad N-dimensione

In quanto discriminanti dall'ampiezza, i segnali $s_m(t)$ hanno diversa energia.

Risulta infatti che l'energia E_m di un segnale s_m si possa calcolare come:

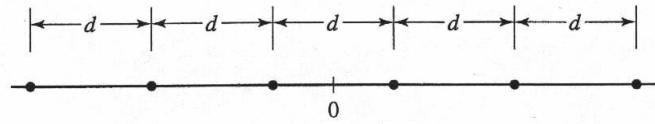
$$\begin{aligned}
E_m &= (s_m)^2 \\
&= \left[A_m \cdot \sqrt{E_g} \right]^2 \\
&= A_m^2 \cdot E_g \\
&= E_g \cdot A_m^2
\end{aligned} \tag{7.33}$$

In alcune applicazioni, in particolare per lo studio dei sistemi di telecomunicazione, ha interesse considerare l'energia media della costellazione E_{av} , definita come:

$$\begin{aligned}
E_{av} &= \frac{1}{M} \cdot \sum_{m=1}^M E_m \\
&= \frac{1}{M} \cdot \sum_{m=1}^M E_g \cdot A_m^2 \\
&= \frac{E_g}{M} \cdot \sum_{m=1}^M A_m^2
\end{aligned} \tag{7.34}$$

E_{av} , av, dall'inglese, si intende average che significa media

Nel caso di punti simmetricamente disposti rispetto all'origine come in figura:



l'ampiezza del singolo segnale è A_m :

$$A_m = (2 \cdot m - 1 - M) \quad (7.35)$$

e quindi, la potenza media della costellazione E_{av} diventa:

$$\begin{aligned} E_{av} &= \frac{E_g}{M} \cdot \sum_{m=1}^M A_m^2 \\ &= \frac{E_g}{M} \cdot \sum_{m=1}^M (2 \cdot m - 1 - M)^2 \\ &= \dots \\ &= \frac{E_g \cdot (M^2 - 1)}{3} \end{aligned} \quad (7.36)$$

Infine, la radice quadrata della distanza Euclidea tra due punti della costellazione vale:

$$\begin{aligned} d_{mn} &= \sqrt{|s_m - s_n|^2} \\ &= \sqrt{E_g \cdot (A_m - A_n)^2} \end{aligned} \quad (7.37)$$

dove:

- s_m e s_n sono due punti nella costellazione
- A_m e A_n sono, rispettivamente, le ampiezze dei punti s_m e s_n

Ricordiamo che, nella formula di d_{mn} , s_m e s_n sono scalari

Maggiore sarà la distanza tra i punti della costellazione, più semplice sarà per il ricevitore capire la differenza tra i vari simboli (ma maggiore distanza significa anche che il trasmettitore dovrà utilizzare più energia per inviare i simboli).

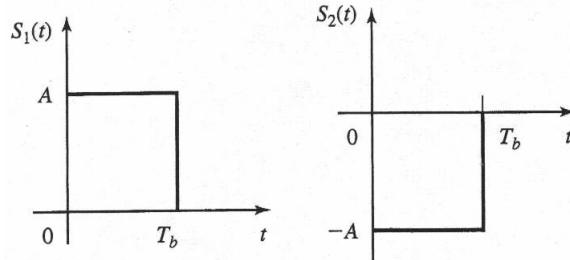
7.3.0.1 Esempi di modulazioni con $N = 1$

⁷

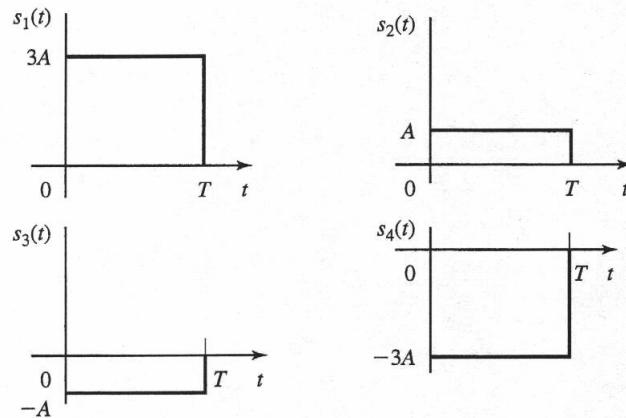
Queste considerazioni vengono svolte per una trasmissione in banda base.

Questo tipo di trasmissione prende il nome di PAM (Pulse Amplitude Modulation), che è una modulazione in banda base che, come dice il nome, ogni simbolo viene contraddistinto in base alla sua ampiezza.

Considerando $M = 2$, possiamo considerare una 2-PAM, in cui andamento dei segnali è il seguente:



Se invece si considera $M = 4$, possiamo considerare una 4-PAM, in cui andamento dei segnali è il seguente:



Invece, se moduliamo il segnale M-PAM a una sinusoide $\cos(2\pi f_c t)$, otteniamo il segnale $u_m(t)$ traslato in $2\pi f_c$, che si ottiene come:

$$u_m(t) = A_m \cdot g_T(t) \cdot \cos(2\pi f_c t) \text{ per } m = 1, 2, \dots, M \text{ e } 0 \leq t \leq T \quad (7.38)$$

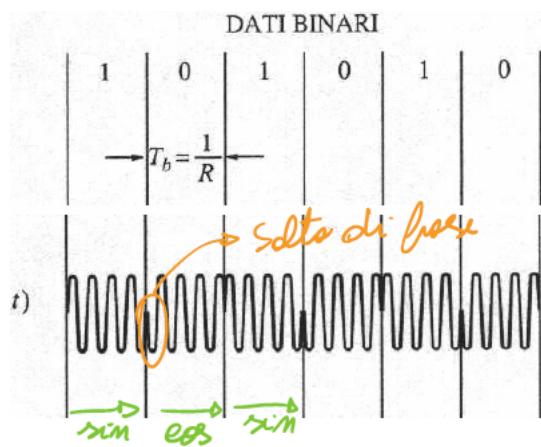
Se consideriamo il caso particolare in cui $g_T(t)$ è un impulso rettangolare, cioè:

$$g_T(t) = \begin{cases} \sqrt{\frac{E_g}{T}} & \text{per } 0 \leq t \leq T \\ 0 & \text{altrove} \end{cases} \quad (7.39)$$

$u_m(t)$ sarà un segnale ASK (Amplitude Shift Keying).

⁷Slide del prof — Formati di trasmissione numerica — pag 10 - 11
 Appunti di Damiano — pag 10 - 11
 Slide — Formati di trasmissione numerica — pag 10 - 11
 Appunti — 2025-03-28 — pag 6 - 8

Un generico andamento di un segnale ASK, nel caso di $M = 2$, è il seguente (con le note):



In effetti il caso ASK binario merita particolare menzione.

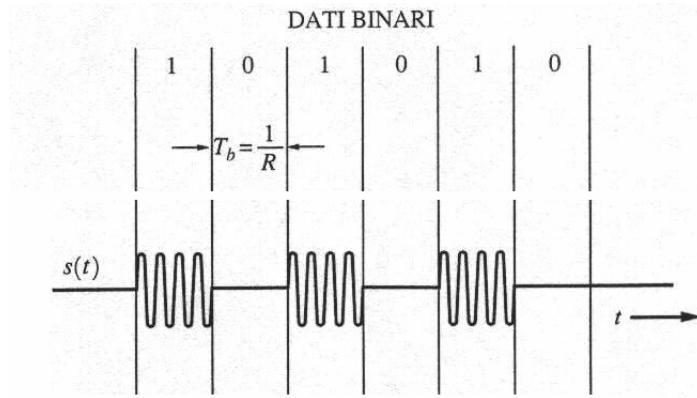
Esso è caratterizzato da forme d'onda antipodale (il coefficiente di correlazione vale -1) e si tratta di un formato di trasmissione che garantisce prestazioni estremamente elevate.

Sempre per il caso binario, in luogo di forme d'onda antipodali, è possibile assumere forme d'onda ortogonali (coefficiente di correlazione uguale a 0).

Il risultato si consegna associando alla trasmissione di uno dei due livelli logici la forma d'onda nulla.

Questo tipo di trasmissione prende il nome di OOK (On-Off Keying).

Un esempio di un andamento di un segnale in OOK:



La OOK è un formato meno efficiente della 2-ASK (a parità degli altri parametri), ma risulta più semplice da implementare.

Ad esempio la OOK è implementata nella fibra ottica, dove la presenza di luce si intende 1, l'assenza di luce si intende 0.

In generale, possiamo dire che la OOK e la ASK sono formati monodimensionali perché hanno una funzione di base.

Un altro modo per visualizzare e capire cosa è una funzione di base, pensiamo quali sono le caratteristiche che permettono di contraddistinguere i vari simboli uno dall'altro.

Ad esempio, nella OOK, l'unico parametro che permette di distinguere il simbolo da un altro è la presenza o meno della portante $\cos(2\pi f_c t)$.

Nella ASK, l'unico parametro che permette di contraddistinguere il simbolo da un altro è la sua fase.

Nella PAM, l'unico parametro che permette di contraddistinguere il simbolo da un altro è la sua ampiezza.

Sembra banale, ma adesso che andremo a visualizzare formati di trasmissione in cui N è maggiore di 1, le considerazioni da svolgere saranno diverse e più complesse.

L'importante è avere dei concetti di base, e per cosa si intende per base, solide e semplici.

7.4 Segnali in spazi di dimensione $N = 2$

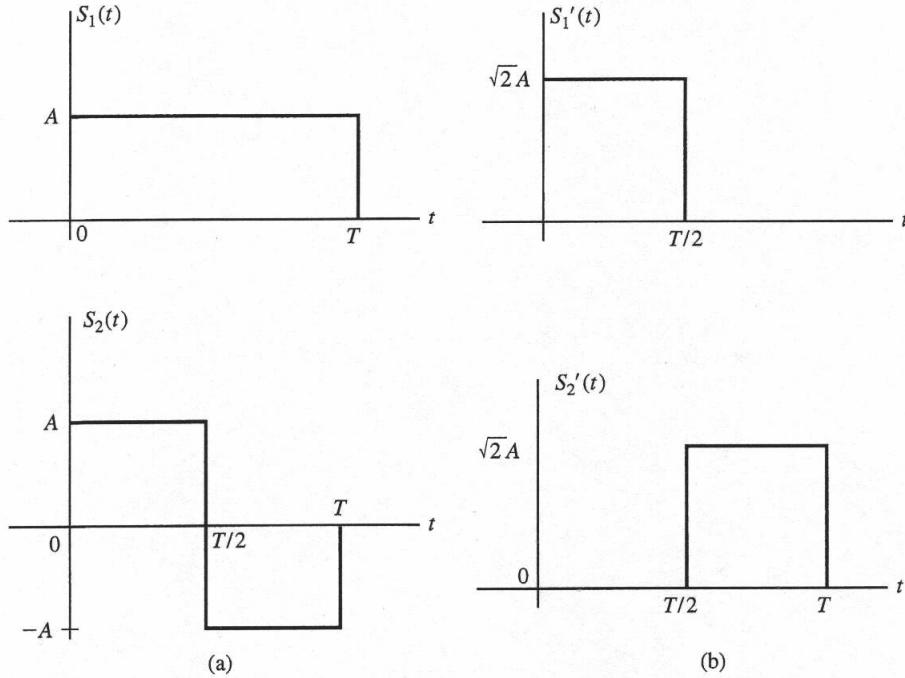
8

In questo caso, lo spazio è bidimensionale, ed è generato da due funzioni espansione $\Psi_1(t)$ e $\Psi_2(t)$.

Considerando due coppie di segnali, cioè:

- prima coppia è $s_1(t)$ e $s_2(t)$
- seconda coppia è $s'_1(t)$ e $s'_2(t)$

e visualizzando il loro andamento nelle seguenti figure:



possiamo verificare la condizione di ortogonalità:

$$\int_0^T s_1(t) \cdot s_2^*(t) dt = \int_0^T s'_1(t) \cdot [s'_2(t)]^* dt = 0 \quad (7.40)$$

Inoltre, l'energia di tutte le forme d'onda elementari considerate è la stessa e vale:

$$\begin{aligned} E &= \int_0^T s_1^2(t) dt \\ &= \int_0^T s_2^2(t) dt \\ &= \int_0^T [s'_1(t)]^2 dt \\ &= \int_0^T [s'_2(t)]^2 dt \\ &= A^2 \cdot T \end{aligned} \quad (7.41)$$

⁸Slide del prof — Formati di trasmissione numerica — pag 11 - 14

Appunti di Damiano — pag 11 - 14

Slide — Formati di trasmissione numerica — pag 11 - 14

Appunti — 2025-03-28 — pag 8

Ricordati la regola aurea dell'ingegnere: per calcolare l'energia di un segnale, cioè l'integrale in un periodo T , devi prendere l'ampiezza del segnale in quel periodo, elevarlo alla seconda perché stiamo parlando di potenze, e moltiplicarlo per il periodo T , proprio come se stessi calcolando l'area di un rettangolo

Grazie all'ortogonalità, si può presupporre che:

$$N = M = 2 \quad (7.42)$$

In altre parole, non sarebbe possibile aggiungere un ulteriore segnale ortogonale senza aumentare la dimensione della base.

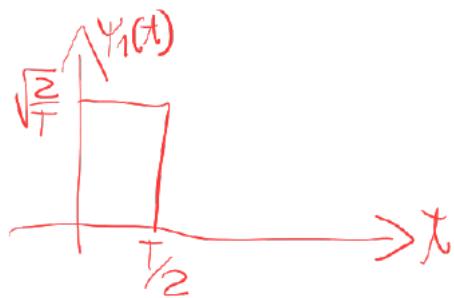
Lo spazio bidimensionale in cui rappresentare i segnali analizzati, cioè i segnali $s_1(t)$, $s_2(t)$, $s'_1(t)$ e $s'_2(t)$, può essere generato, ad esempio, dalle seguenti funzioni rettangolari:

$$\begin{cases} \Psi_1(t) = \begin{cases} \sqrt{\frac{2}{T}} & \text{per } 0 \leq t \leq \frac{T}{2} \\ 0 & \text{altrove} \end{cases} \\ \Psi_2(t) = \begin{cases} \sqrt{\frac{2}{T}} & \text{per } \frac{T}{2} \leq t \leq T \\ 0 & \text{altrove} \end{cases} \end{cases} \quad (7.43)$$

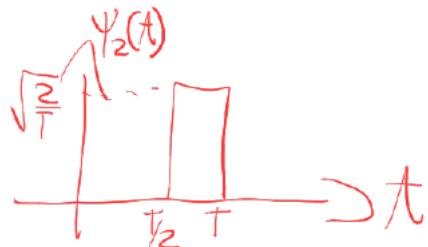
Ricordiamo sul perchè il prof ha scritto "ad esempio".

La procedura di ortogonalizzazione di Gram-Schmidt ci garantisce che va bene qualsiasi segnale come base dello spazio, basta che poi i segnali che si trovano in quello spazio siano combinazione lineare dei vettori della base stessa

Dal punto di vista grafico, possiamo visualizzare l'andamento nel tempo di $\Psi_1(t)$ come:



Invece, l'andamento nel tempo di $\Psi_2(t)$ come:



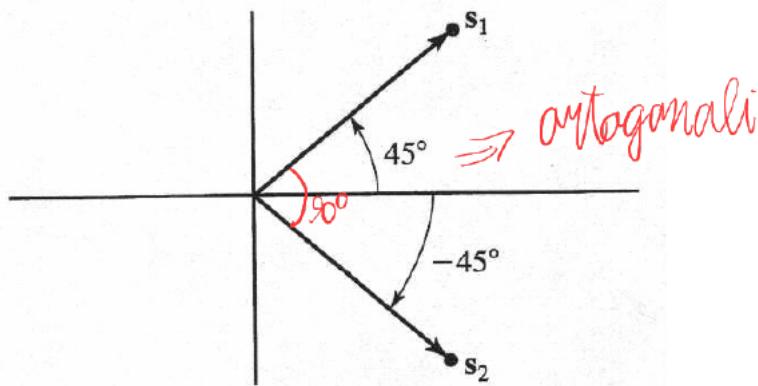
Allora, le forme d'onda $s_1(t)$ e $s_2(t)$, ad esempio, si esprimono come:

$$\begin{cases} s_1(t) = s_{11} \cdot \Psi_1(t) + s_{12} \Psi_2(t) \\ s_2(t) = s_{21} \cdot \Psi_1(t) + s_{22} \Psi_2(t) \end{cases} \quad (7.44)$$

ovvero, come si verifica immediatamente, tramite le loro componenti nello spazio bidimensionale:

$$\begin{cases} \vec{s}_1 = (s_{11}, s_{12}) = \left(A \cdot \sqrt{\frac{T}{2}}, A \cdot \sqrt{\frac{T}{2}} \right) \\ \vec{s}_2 = (s_{21}, s_{22}) = \left(A \cdot \sqrt{\frac{T}{2}}, -A \cdot \sqrt{\frac{T}{2}} \right) \end{cases} \quad (7.45)$$

Possiamo visualizzare \vec{s}_1 e \vec{s}_2 su un piano bidimensionale in questa maniera:



I due segnali \vec{s}_1 e \vec{s}_2 sono rappresentati da vettori che sono tra loro ortogonali.

La lunghezza di ogni vettore al quadrato, fornisce l'energia di ciascun segnale.

Cioè:

$$\begin{cases} E_1 = |\vec{s}_1|^2 \\ E_2 = |\vec{s}_2|^2 \end{cases} \quad (7.46)$$

dove con $|\vec{s}_1|$ e $|\vec{s}_2|$ si intendono i moduli, cioè la lunghezza dei vettori \vec{s}_1 e \vec{s}_2 .

La radice quadrata della distanza Euclidea tra i due segnali d_{12} , in questo caso, vale:

$$\begin{aligned} d_{12} &= \sqrt{|\vec{s}_1 - \vec{s}_2|^2} \\ &= \dots \\ &= A \cdot \sqrt{2 \cdot T} \end{aligned} \quad (7.47)$$

Sapendo che l'energia E vale:

$$E = A^2 \cdot T \quad (7.48)$$

possiamo riscrivere la distanza Euclidea d_{12} in funzione dell'energia E anche come:

$$\begin{aligned} d_{12} &= A \cdot \sqrt{2 \cdot T} \\ &= \sqrt{2 \cdot E} \end{aligned} \quad (7.49)$$

Comparando la distanza tra i due vettori in $N = 1$:

$$d_{12} = 2 \cdot \sqrt{E} \quad (7.50)$$

nel caso $N = 2$, la distanza tra i due vettori è diminuita perché $\sqrt{2 \cdot E}$ è minore rispetto a \sqrt{E} .

Lo si può anche notare dalle raffigurazioni nello spazio bidimensionale.

I due vettori, piuttosto che avere una differenza di 180° , differiscono solo di 90° .

Questo "avvicinamento" peggiora la probabilità di errore dei simboli, ma è più facile da implementare rispetto al caso $N = 1$.

Quindi, come ogni argomento di questo corso e di ogni corso di ingegneria, bisogna scegliere sul momento della progettazione quale criterio ci interessa di più ai nostri scopi, cioè, e lo ripeterò allo sfinito, non puoi avere la moglie ubriaca e la botta piena contemporaneamente.

La vita reale è fatta di compromessi.

La dimostrazione di d_{12} è semplice se parti da $\sqrt{2 \cdot E}$ e fai il processo inverso

Utilizzando la stessa base $\Psi_1(t)$ e $\Psi_2(t)$:

$$\left\{ \begin{array}{l} \Psi_1(t) = \begin{cases} \sqrt{\frac{2}{T}} & \text{per } 0 \leq t \leq \frac{T}{2} \\ 0 & \text{altrove} \end{cases} \\ \Psi_2(t) = \begin{cases} \sqrt{\frac{2}{T}} & \text{per } \frac{T}{2} \leq t \leq T \\ 0 & \text{altrove} \end{cases} \end{array} \right. \quad (7.51)$$

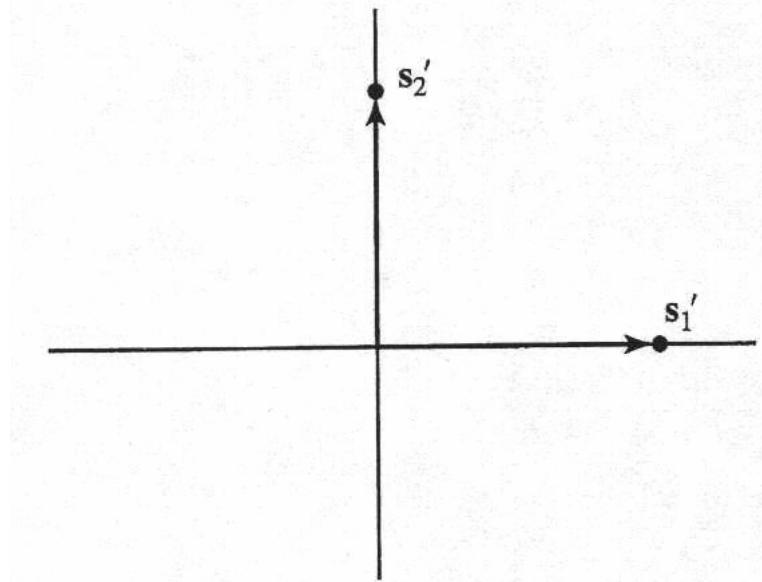
e facendo gli stessi passaggi precedenti, cioè:

$$\left\{ \begin{array}{l} s'_1(t) = s'_{11} \cdot \Psi_1(t) + s'_{12} \Psi_2(t) \\ s'_2(t) = s'_{21} \cdot \Psi_1(t) + s'_{22} \Psi_2(t) \end{array} \right. \quad (7.52)$$

possiamo esprimere $s'_1(t)$ e $s'_2(t)$ come:

$$\left\{ \begin{array}{l} \vec{s}_1 = (s'_{11}, s'_{12}) = (A \cdot \sqrt{T}, 0) = (\sqrt{E}, 0) \\ \vec{s}_2 = (s'_{21}, s'_{22}) = (0, A \cdot \sqrt{T}) = (0, \sqrt{E}) \end{array} \right. \quad (7.53)$$

La corrispondente rappresentazione grafica nello spazio bidimensionale di \vec{s}_1 e \vec{s}_2 è la seguente:



Stante la similitudine tra le due rappresentazioni, cioè quella di \vec{s}_1' con \vec{s}_2' , la distanza Euclidea tra \vec{s}_1 e \vec{s}_2 è la stessa di \vec{s}_1' e \vec{s}_2' , cioè:

$$d_{1'2'} = d_{12} \quad (7.54)$$

e inoltre, la rappresentazione dei due vettori \vec{s}_1 e \vec{s}_2 è la stessa di \vec{s}_1' e \vec{s}_2' con una semplice rotazione di 45° .

7.4.1 Esempio di modulazione con $N = 2$: la PPM

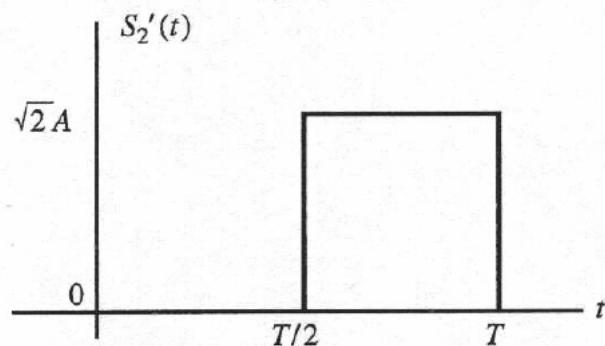
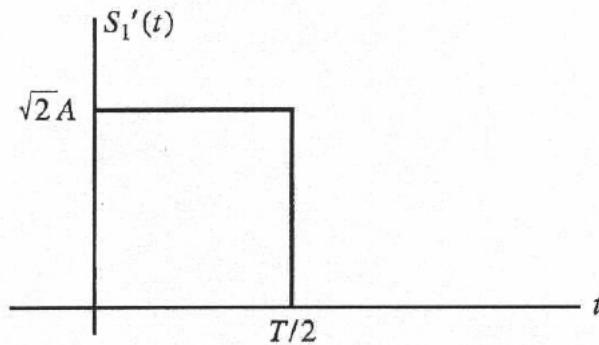
9

Un esempio di modulazione con $N = 2$, è la PPM (Pulse Position Modulation).

Dato il piano bidimensionale, i segnali differiscono per la posizione dell'impulso.

L'ortogonalità può essere conseguita facendo in maniere tale che i due impulsi non si sovrappongono.

Un esempio di due segnali che non si sovrappongono, li avevamo visti precedentemente con $s'_1(t)$ e $s'_2(t)$, in cui andamento nel tempo è il seguente:



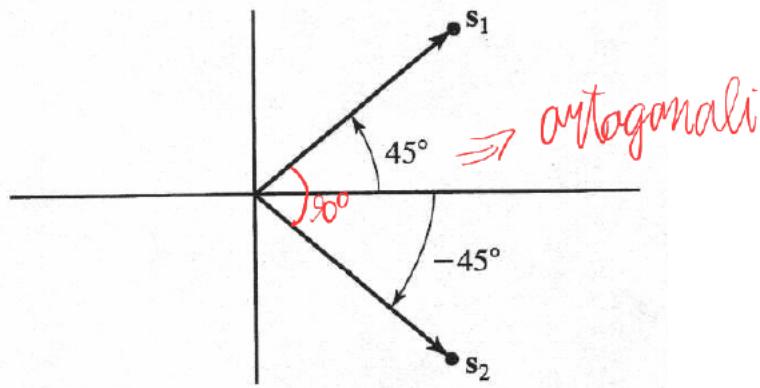
La "filosofia" della PPM può essere generalizzata ad un numero di segnali maggiore di 2, e verrà ripresa in un paragrafo successivo riguardo agli spazi multi-dimensionalali.

⁹Slide del prof — Formati di trasmissione numerica — pag 14
 Appunti di Damiano — pag 14
 Slide — Formati di trasmissione numerica — pag 14
 Appunti — 2025-03-28 — pag 8

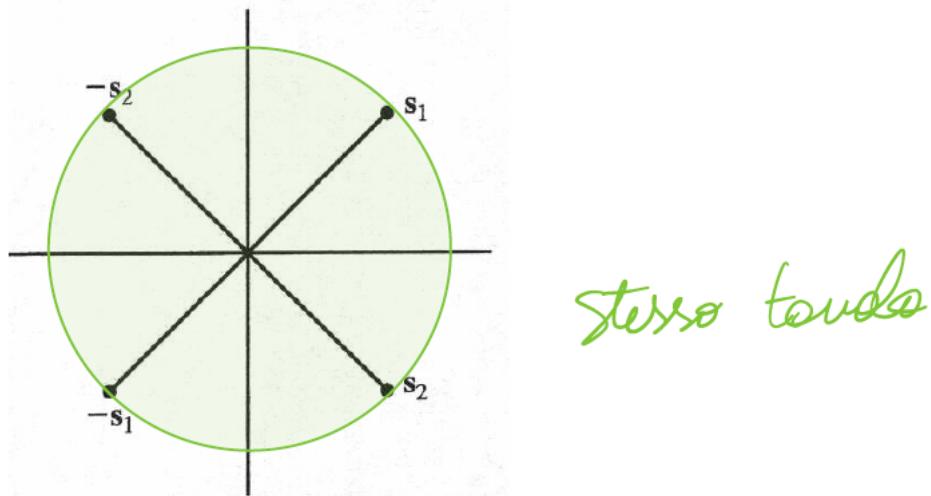
7.4.2 Segnali bi-ortogonali

¹⁰

Se ora si considera la seguente figura che abbiamo trovato precedentemente:



e si considera i loro opposti, si può ottenere la seguente costellazione:



cioè da $M = 2$, si ha una costellazione con $M = 4$.

Il concetto dei tavoli, del tavolo più grande e delle persone a tavola verrà spiegato successivamente.

Un piccolo spoiler: non è detto che dovete andare all'IKEA a comprare un nuovo tavolo, ogni tanto si dovrà stare più stretti per far entrare tutti nello stesso tavolo.

Quindi, rispetto al caso $M = 2$, nella costellazione con $M = 4$, i quattro segnali non sono più tra loro ortogonali.

Grazie alla scelta operata, questa tipo di costellazione definisce un insieme di forme d'onda decomponibile in due sottoinsieme, ciascuno di dimensione $N = 2$.

I segnali di ciascun sottoinsieme sono tra loro ortogonali, mentre i segnali di un sottoinsieme sono opposti a quelli dell'altro insieme.

¹⁰Slide del prof — Formati di trasmissione numerica — pag 14 - 15
 Appunti di Damiano — pag 14 - 15
 Slide — Formati di trasmissione numerica — pag 14 - 15
 Appunti — 2025-03-28 — pag 8

Segnali con queste caratteristiche si dicono bi-ortogonali.

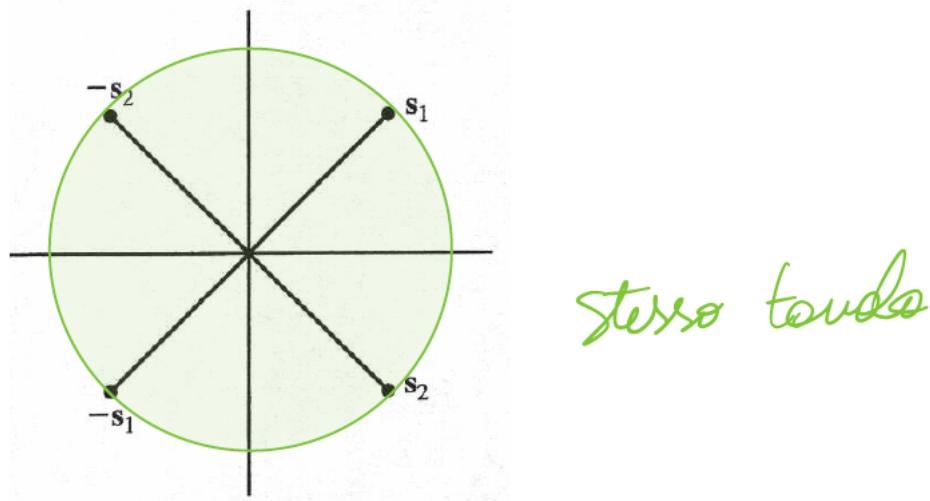
La loro generalizzazione verrà trattata successivamente.

Inoltre, in questo tipo di costellazione, con cioè $M = 4$, i segnali hanno la stessa energia e probabilità di errore di $M = 2$ perché i segnali sono distribuiti nella stessa circonferenza.

7.4.3 Segnali nella stessa circonferenza

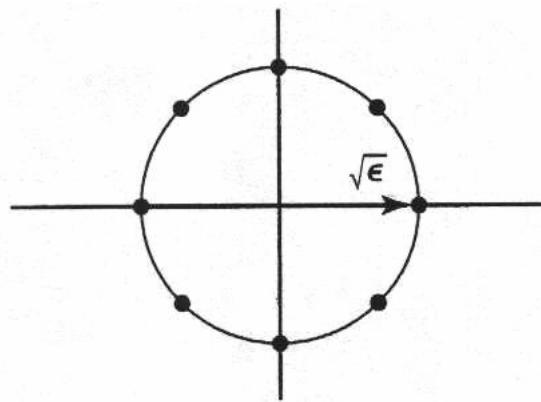
11

Ricordando che, con $M = 4$, questi segnali hanno la stessa energia perchè si trovano nella stessa circonferenza:



allora possiamo aumentare gli M segnali e porli nella stessa circonferenza.

Considerando $M = 8$, possiamo avere questo caso:



Lasciando inalterato il raggio del cerchio, e dunque l'energia E , all'aumentare di M segnali, i punti rappresentativi dei segnali diventano sempre più vicini tra loro.

Posto infatti il vettore \vec{s}_m :

$$\vec{s}_m = \left(\sqrt{E} \cdot \cos\left(\frac{2 \cdot \pi \cdot m}{M}\right), \sqrt{E} \cdot \sin\left(\frac{2 \cdot \pi \cdot m}{M}\right) \right) \quad (7.55)$$

la radice quadrata della distanza Euclidea tra due punti della costellazione vale:

¹¹Slide del prof — Formati di trasmissione numerica — pag 15
 Appunti di Damiano — pag 15
 Slide — Formati di trasmissione numerica — pag 15
 Appunti — 2025-03-28 — pag 8

$$\begin{aligned}
 d_{mn} &= \sqrt{|\vec{s}_m - \vec{s}_n|^2} \\
 &= \dots \\
 &= \sqrt{2 \cdot E \cdot \left[1 - \cos\left(\frac{2 \cdot \pi \cdot (m - n)}{M}\right) \right]}
 \end{aligned} \tag{7.56}$$

Da sottolineare, ma è anche inteso, che nella formula di d_{mn} l'energia E è la stessa perchè i due punti si trovano nella stessa circonferenza

Se consideriamo:

$$m - n = 1 \tag{7.57}$$

consideriamo due punti adiacenti, in cui la distanza minima deve valere:

$$d_{min} = \sqrt{2 \cdot E \cdot \left[1 - \cos\left(\frac{2 \cdot \pi}{M}\right) \right]} \tag{7.58}$$

La distanza minima d_{min} è il parametro più importante tra due punti perchè, se la distanza minima non viene mantenuta, il ricevitore interpreta un simbolo come un altro.

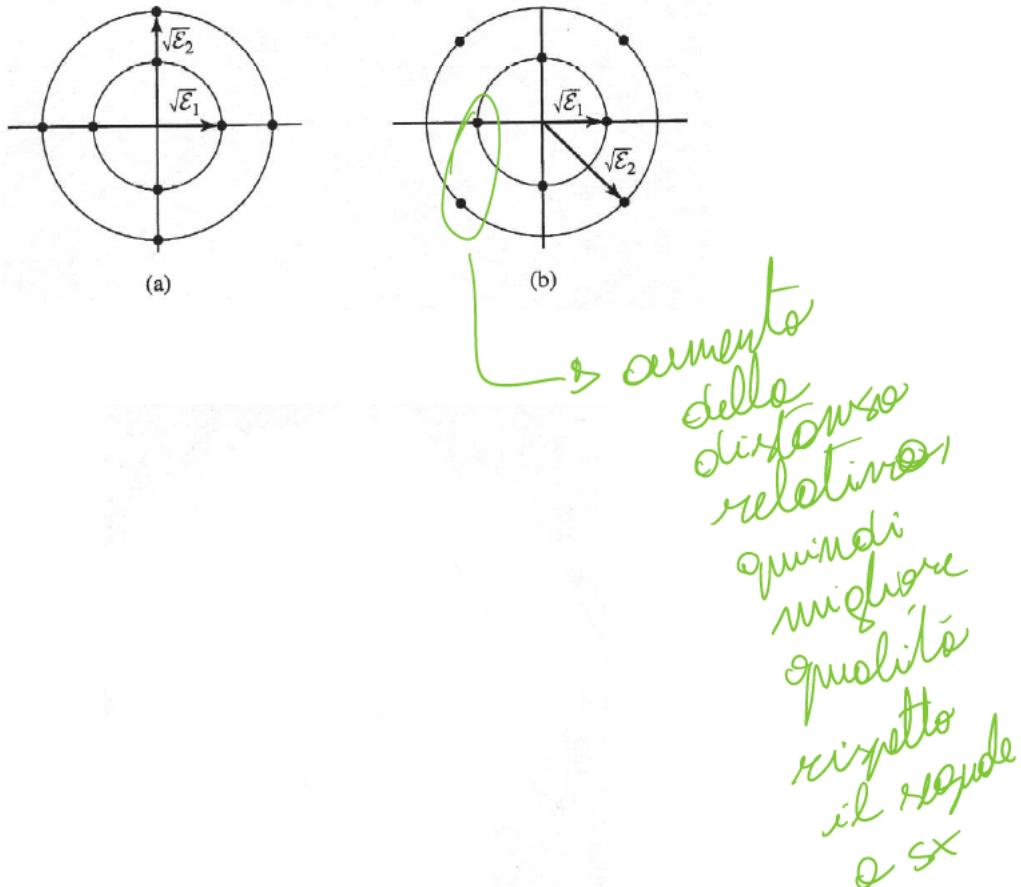
Inoltre, dalla formula di d_{min} si nota che, se si vuole aumentare d_{min} , l'unico fattore che possiamo aumentare è l'energia E dei segnali

7.4.4 Segnali in diverse circonferenze

12

Rimuovendo il vincolo sull'energia, permettendo cioè che i segnali possano avere anche energie diverse tra loro, sono possibili altre costellazioni.

Due esempi, con $M = 8$, possiamo visualizzarli di seguito (con le note):



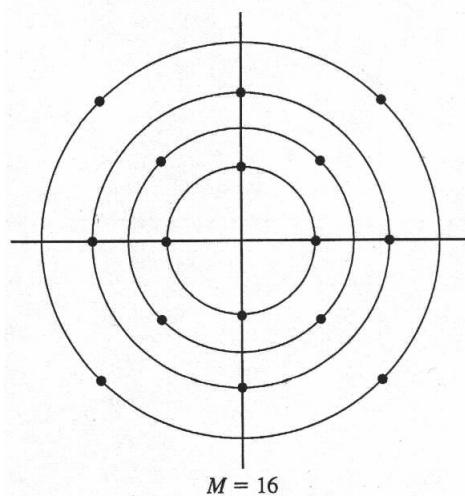
In queste due esempi, i punti sono distribuiti su due cerchi (anziché uno come in precedenza) di raggio pari a $\sqrt{E_1}$ e $\sqrt{E_2}$, rispettivamente.

Qualitativamente, si può notare che la distanza minima per la costellazione (b) è maggiore rispetto alla distanza minima della costellazione (a).

Questo fatto risulta vantaggioso quando gli 8 segnali vengono utilizzati per la trasmissione dell'informazione in un sistema di telecomunicazione.

¹²Slide del prof — Formati di trasmissione numerica — pag 15 - 16
 Appunti di Damiano — pag 15 - 16
 Slide — Formati di trasmissione numerica — pag 15 - 16
 Appunti — 2025-03-28 — pag 8
 Appunti — 2025-03-31 — pag 2

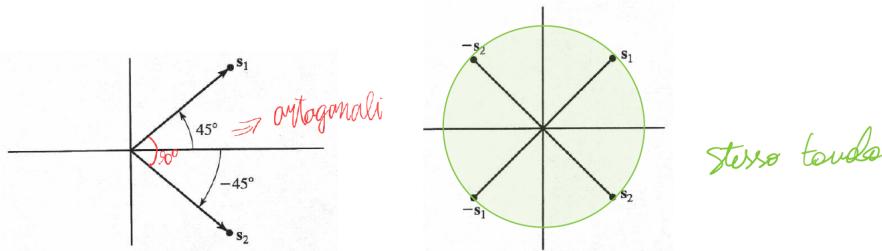
L'estensione di questa regola, cioè quella di aumentare l'energia per ogni circonferenza, la si può vedere in questo esempio in cui $M = 16$:



7.4.5 Costellazioni in segnali passa-banda

13

Le costellazioni precedentemente studiate:



riguardo ai segnali bi-ortogonal, sono segnali in banda base.

Questo tipo di costellazioni in banda base, possono essere estese al caso di rappresentazioni di segnali in passa banda utilizzando la relazione:

$$u_m(t) = s_m(t) \cdot \cos(2 \cdot \pi \cdot f_c \cdot t) \text{ per } m = 1, 2, \dots, M \text{ e } 0 \leq t \leq T \quad (7.59)$$

dove f_c è una opportuna frequenza di portante.

Se i segnali $s_m(t)$ (in cui $m = 1, 2, \dots, M$) hanno tutti la stessa energia E , cioè:

$$E = \int_0^T [s_m(t)]^2 dt \quad (7.60)$$

allora, anche i segnali $u_m(t)$ (in cui $m = 1, 2, \dots, M$) hanno, con buona approssimazione, tutti la stessa energia:

$$\begin{aligned} E_s &= \int_0^T [u_m(t)]^2 dt \\ &\approx \frac{1}{2} \cdot \int_0^T [s_m(t)]^2 dt \end{aligned} \quad (7.61)$$

Si è fatta questa approssimazione perchè, nella funzione integranda, il termine oscillante, cioè $\cos(2\pi f_c t)$, ipotizzando $f_c \gg W$, con banda W del segnale $s_m(t)$, dà un contributo trascurabile, se non, in certi casi, anche uguale a zero

D'altro canto, ove lo spazio è bidimensionale sia rappresentativo di segnali passa-banda, le funzioni di base che generano lo spazio saranno sempre due.

In questo caso:

$$\begin{cases} \Psi_1(t) = \sqrt{\frac{2}{E_g}} \cdot g_T(t) \cdot \cos(2\pi f_c t) \\ \Psi_2(t) = \sqrt{\frac{2}{E_g}} \cdot g_T(t) \cdot \sin(2\pi f_c t) \end{cases} \quad (7.62)$$

¹³Slide del prof — Formati di trasmissione numerica — pag 16 - 17
Appunti di Damiano — pag 16 - 17
Slide — Formati di trasmissione numerica — pag 16 - 17
Appunti — 2025-03-28 — pag 9
Appunti — 2025-03-31 — pag 2 - 3

dove E_g è l'energia di $g_T(t)$.

I termini sinusoidali nei termini delle funzioni di base, cioè $\cos(2\pi f_c t)$ e $\sin(2\pi f_c t)$, rimarcano che non ci troviamo più in banda base, bensì in banda traslata

Ove si assuma per $g_T(t)$ un impulso rettangolare, e cioè si pone:

$$g_T(t) = \sqrt{\frac{2 \cdot E_s}{T}} \text{ per } 0 \leq t \leq T \quad (7.63)$$

la precedente si particolarizza come segue:

$$\begin{aligned} u_m(t) &= s_m(t) \cdot \cos(2 \cdot \pi \cdot f_c \cdot t) \text{ per } m = 1, 2, \dots, M \text{ e } 0 \leq t \leq T \\ &\downarrow \\ u_m(t) &= \sqrt{\frac{2 \cdot E_s}{T}} \cdot \cos\left(2 \cdot \pi \cdot f_c \cdot t + \frac{2\pi m}{M}\right) \text{ per } m = 0, 1, \dots, M-1 \text{ e } 0 \leq t \leq T \end{aligned} \quad (7.64)$$

Analizzando la nuova formula di $u_m(t)$, la formula descrive l'andamento nel tempo di u_m .

Dal punto di vista informativo, l'ampiezza $\sqrt{\frac{2 \cdot E_s}{T}}$ rimane costante, quindi l'ampiezza non porta informazione.

Il fattore che cambia nel tempo, e che porta informazione, è $\cos(2 \cdot \pi \cdot f_c \cdot t + \frac{2\pi m}{M})$, cioè la fase di $u_m(t)$

Equivalentemente, utilizzando le funzioni di base, il generico segnale \vec{u}_m è rappresentato dalle sue componenti:

$$\vec{u}_m = \left(\sqrt{E_s} \cdot \cos\left(\frac{2\pi m}{M}\right), \sqrt{E_s} \cdot \sin\left(\frac{2\pi m}{M}\right) \right) \quad (7.65)$$

dove \vec{u}_m è un versore.

7.4.5.1 Modulazioni con costellazioni di segnali passa banda

¹⁴

Una trasmissione basata sull'invio di queste forme d'onda:

$$\begin{aligned}
 E_m &= (s_m)^2 \\
 &= \left[A_m \cdot \sqrt{E_g} \right]^2 \\
 &= A_m^2 \cdot E_g \\
 &= E_g \cdot A_m^2
 \end{aligned} \tag{7.66}$$

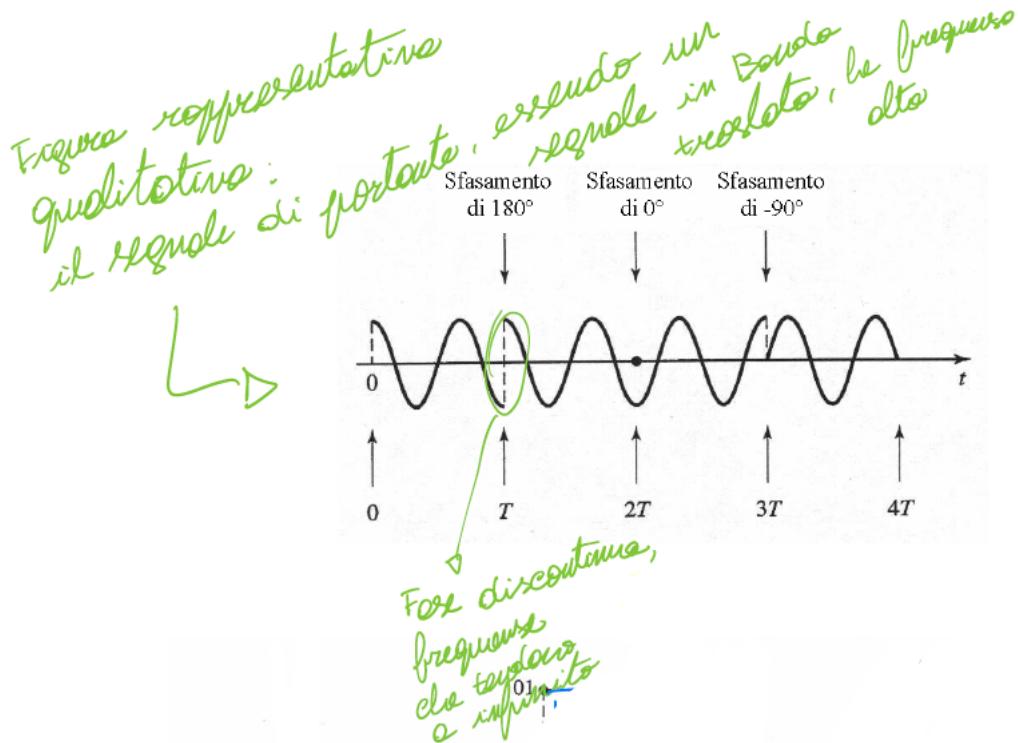
utilizza la modulazione PSK (Phase Shift Keying).

La modulazione PSK definisce, quindi, un sistema di trasmissione in banda traslata, i cui segnali differiscono per la fase e sono caratterizzate dalla medesima energia.

Due forme d'onda adiacenti, differiscono per un angolo pari a $\frac{2\pi}{M}$.

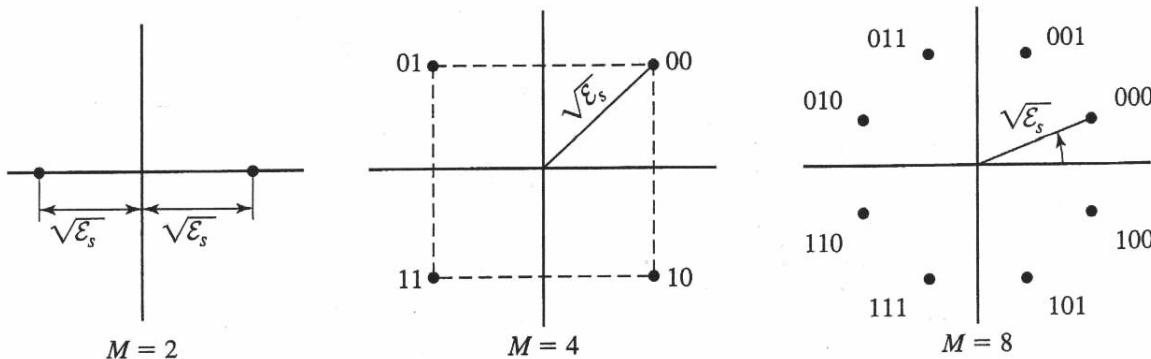
Per $M = 2$, la modulazione PSK, si riduce ad una modulazione 2-ASK introdotta in precedenza.

Un esempio di segnale 4-PSK (con le note) è mostrato di seguito:



¹⁴Slide del prof — Formati di trasmissione numerica — pag 17 - 20
 Appunti di Damiano — pag 17 - 20
 Slide — Formati di trasmissione numerica — pag 17 - 20
 Appunti — 2025-03-28 — pag 9
 Appunti — 2025-03-31 — pag 3 - 5

Un esempio di costellazioni per di PSK con $M = 2, 4, 8$ sono le seguenti figure:



Inoltre, queste figure ci permettono di mettere in evidenza, per il caso $M = 4$ e $M = 8$, la possibilità di applicare la codifica di Gray.

Per codifica di Gray, si intende l'associazione bit a simboli (cioè mapping), la cui caratteristica peculiare è che sequenze binarie relative a simboli adiacenti differiscono di un solo bit (nota la codifica del singolo punto in binario).

Questa associazione, i.e. mapping, ha un enorme vantaggio per la qualità in trasmissione, perché permette di minimizzare la probabilità di errore sul bit (o in breve BER: Bit Error Rate).

Possiamo ipotizzare che il rumore fa passare un segnale al suo punto adiacente.

Questo perchè, dal corso di Teoria dei Segnali, sappiamo che la probabilità del rumore termico è di tipo gaussiano, cioè che il rumore impatta di più sui valori più piccoli rispetto ai valori più alti.

In questo caso, il rumore termico impatta di più sull'LSB (Least Significant Bit), cioè il bit più piccolo, che, grazie alla codifica di Gray, fa al massimo muovere il segnale da punto desiderata a quello adiacente.

Inoltre, la codifica di Gray è efficace nelle modulazioni in cui i simboli sono equiprobabili.

Sempre nel caso di modulazioni in banda traslata, la rimozione del vincolo sull'energia (presente nel formato PSK), conduce alla definizione della QAM (Quadrature Amplitude Modulation).

La QAM è LA modulazione che letteralmente utilizzi tutti i giorni, perchè viene implementata dallo standard WiFi (e non solo).

Ti lascio un po' di video sulla spiegazione della QAM:

<https://www.youtube.com/watch?v=xn9zqSo0lcE&t=20s>

Wireless Communication - Eight: Quadrature Amplitude Modulation by Computer Science Lessons

La QAM viene impiegato anche nello standard 5G, e quindi anche dai modem contenuti nei chipset dei nostri cellulari:

<https://www.youtube.com/watch?v=iEWUWTfVGqI>

Snapdragon 888: la vera novità per la connettività è la carrier aggregation 5G! by Stefano Bolis

Ferma restando la rappresentazione su uno spazio bidimensionale (cioè $N = 2$), i simboli QAM sono distribuiti in modo da non trovarsi, necessariamente, alla stessa distanza dall'origine.

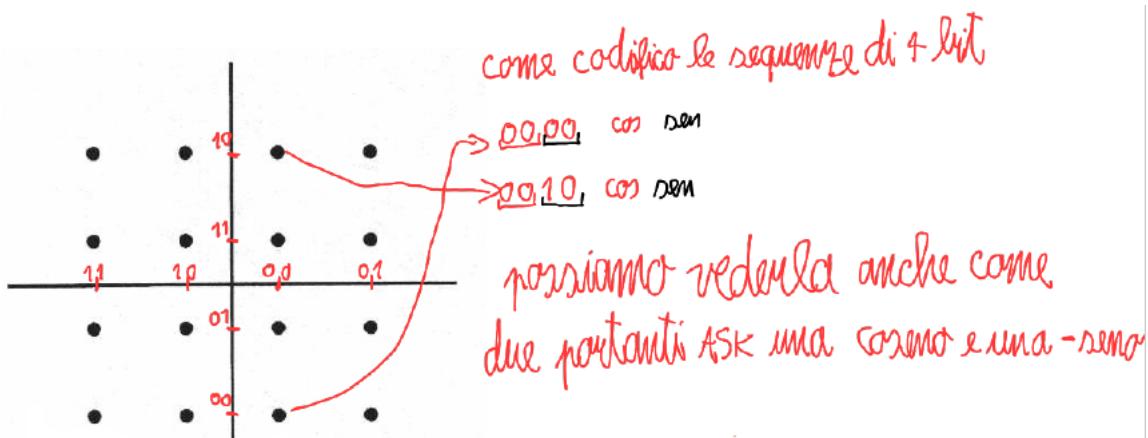
Questo corrisponde ad aumentare il numero dei gradi di libertà di sistema, per cui la generica forma d'onda trasmessa potrà scriversi:

$$u_{mn} = A_m \cdot g_T(t) \cos(2\pi f_c t + \theta_n) \text{ per } m = 1, 2, \dots, M_1, n = 1, 2, \dots, M_2 \quad (7.67)$$

I vari segnali differiscono quindi per la fase (come nel sistema PSK), ma anche per l'ampiezza e M sarà uguale:

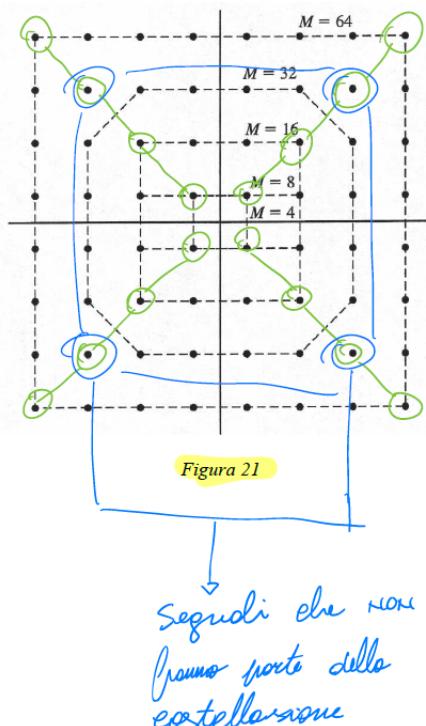
$$M = M_1 \cdot M_2 \quad (7.68)$$

Esempi classici della QAM, qui per $M = 16$, è riportata di seguito:



La 16-QAM assomiglia alla PSK, anche con la PSK hanno in comune le stesse funzioni di base

Invece, con la seguente figura, possiamo confrontare tutte le possibili costellazioni di QAM dalla $M = 4$ alla $M = 64$ (con note):



Le costellazione nel caso di:

$$M = 2^k \quad (7.69)$$

dove k è un numero pari, formano costellazioni che sono quadrati perfetti (si può notare la costellazione di M = 4, o M = 64).

Invece le costellazioni con k dispari maggiore di 3, ad esempio M = 32, formano una costellazione a croce (in inglese cross-constellation), formando una costellazione di un quadrato perfetto senza i punti degli angoli.

Proprio per questo motivo, generalmente viene utilizzando un M in cui k è pari.

Ricordiamo che, se possibile, specialmente nel campo Wireless, bisogna utilizzare tutto quello che si ha per migliorare la qualità della trasmissione e la sua efficienza.

Un detto popolare dice "Del porco, non si butta via nulla": anche nel campo delle telecomunicazioni, se possibile, non bisogna buttare via nulla.

"Buttare" dei simboli è "un privilegio" certe volte.

Lo schema di un demodulatore è il seguente:

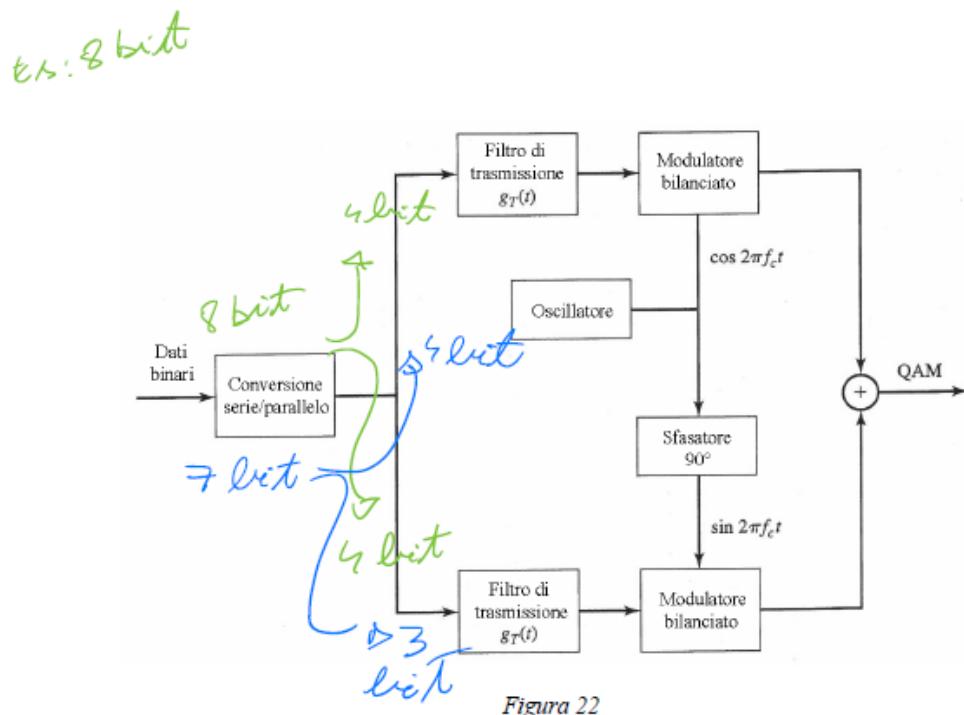


Figura 22

Ora K parci
permette di
dividere i bit
sopra e sotto

• Esempio con k dispari

Le due portanti, sfasate di 90° definiscono lo spazio bidimensionale.

La conversione serie-parallelo separa i bit che selezioneranno l'ampiezza di ciascuna delle due portanti in quadratura, in accordo con la costellazione specificata.

Ricordiamo anche l'implicazione di avere un M elevato: l'aumento della banda in frequenza.

Certe volte può essere un vantaggio (si pensi alla 4096-QAM del Wifi 7 con il quale si possono trasmettere un flusso teorici di 40 Gbps con 320 MHz di banda), ma altre volte uno svantaggio (si pensi ad un operatore di rete telefonica che paga per la banda e deve inserire più utenti contemporaneamente).

Un po' di link per avere una panoramica sulle implicazioni che esistono realmente ad oggi giorno:

<https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/Qualcomm-Wi-Fi-7-Reference-Guide.pdf>

<https://www.youtube.com/watch?v=Db4RcNPIDVU>

Il 5G ha problemi di velocità e copertura in upload: cause e soluzioni! by Stefano Bolis

<https://www.youtube.com/watch?v=mS00PbBt9hs>

Frequenze 5G in banda C: gli operatori italiani le hanno pagate il doppio! by Stefano Bolis

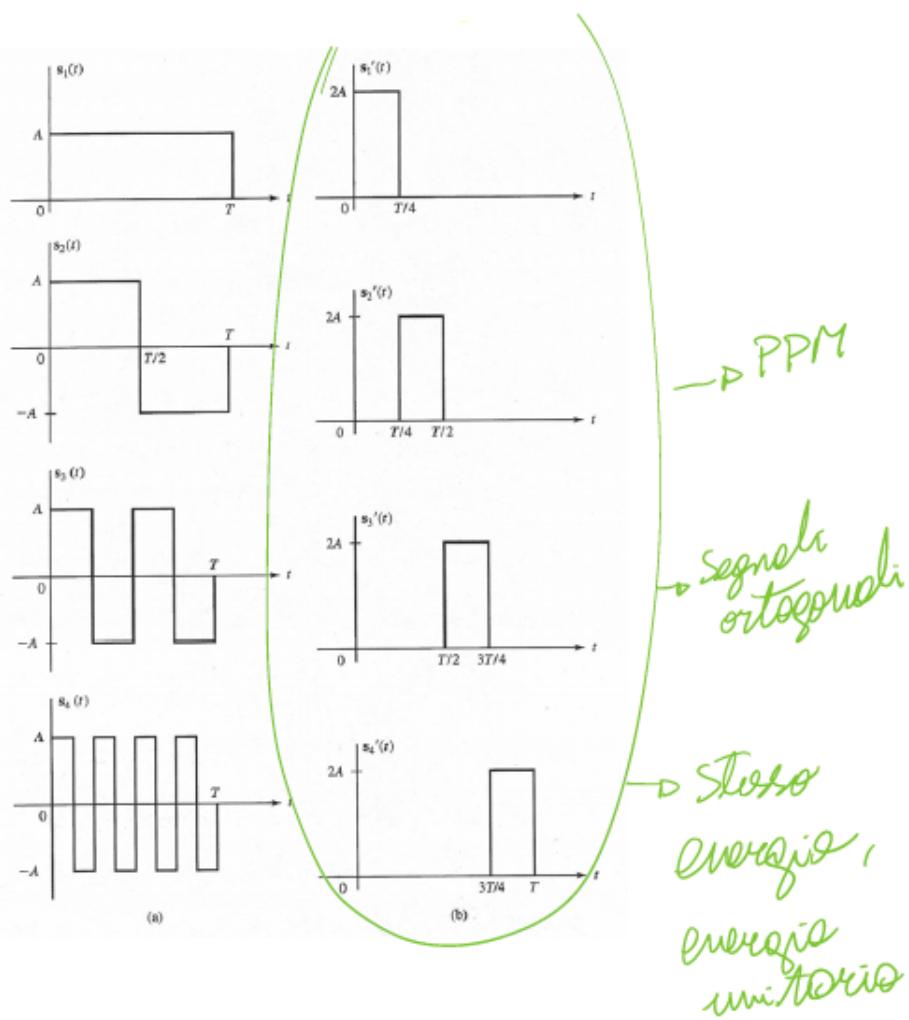
7.5 Segnali ortogonali, transagonali e bi-ortogonali

15

Già nel paragrafo precedente si sono forniti esempi di segnali ortogonali, e sono state date informazioni in merito alla loro rappresentazione, ma limitatamente al caso bi-dimensione.

In realtà, la proprietà di ortogonalità può essere estesa ad un insieme di segnali di dimensione arbitraria, ciascuno, ad esempio, di durata T .

Di seguito due esempi di segnali con $M = 4$ (con note):



A destra un sistema diverso da quello a sinistra.

L'implementazione della ortogonalità tra i vari segnali, con i due sistemi, è diversa.

Ricordiamo la definizione di ortogonalità tra due segnali:

$$\int s_1(t) \cdot s_2^*(t) dt = 0 \quad (7.70)$$

¹⁵Slide del prof — Formati di trasmissione numerica — pag 21

Appunti di Damiano — pag 21

Slide — Formati di trasmissione numerica — pag 21

Appunti — 2025-03-31 — pag 6

dove:

- si pone gli estremi di integrazione dell'integrale nei tratti in cui si vuole calcolare la ortogonalità tra $s_1(t)$ e $s_2(t)$
- per $s_2^*(t)$ si intende il coniugato di $s_2(t)$

L'ortogonalità dei segnali a sinistra, cioè i segnali $[s_1(t), s_2(t), s_3(t), s_4(t)]$, sono tutti diversi da zero nell'intervallo $[0, T]$.

In questo caso, l'ortogonalità viene conseguita attraverso una oculata assegnazione di sotto-intervalli, in cui le funzioni valgono, alternativamente $+A$ e $-A$.

Invece, nel sistema a destra, cioè i segnali $[s'_1(t), s'_2(t), s'_3(t), s'_4(t)]$, sono ottenuti con un unico impulso di durata $\frac{T}{4}$, che, nei vari segnali, viene traslato in periodi di $\frac{T}{4}$.

Ciò, garantisce una facile implementazione dell'ortogonalità tra i segnali e inoltre, nello stesso intervallo multiplo di $\frac{T}{4}$, i vari segnali non si sovrappongono: questa classe di segnali, generalizzata a un valore M qualsiasi, prende il nome di PPM (Pulse Position Modulation).

Grazie alla sua facilità di implementazione, la PPM è molto usata nel campo della fibra ottica, perchè impulso alto significa luce del laser, no impulso significa laser spento.

Delle letture sull'implementazione reale della PPM:

<https://ieeexplore.ieee.org/document/5637976>
Pulse position modulation (PPM) fiber optic architectures

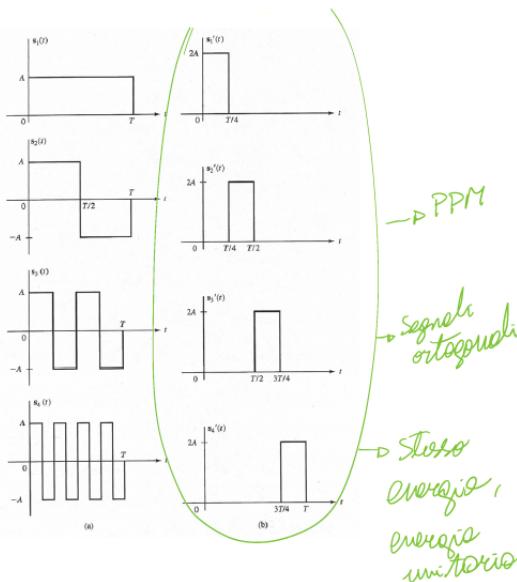
<https://ntrs.nasa.gov/api/citations/19680025702/downloads/19680025702.pdf>
The design of a Pulse Position Modulated (PPM) Optical Communication System

7.5.1 PPM (Pulse Position Modulation)

16

La PPM viene chiamata Pulse Position Modulation proprio in ragione del fatto che i vari segnali sono discriminati sulla base della posizione di un impulso (che nel caso generale avrà durata $\frac{T}{M}$).

Tutti i segnali visti nella figura dei due sistemi in cui i segnali sono ortogonali tra di loro:



hanno la stessa energia E , che vale:

$$E = A^2 \cdot T \quad (7.71)$$

Per rappresentare l'insieme di M forme d'onda ortogonali, è necessario uno spazio di dimensione:

$$N = M = 4 \quad (7.72)$$

Quindi, su questo spazio, gli M segnali sono rappresentati dai vettori con i seguenti componenti:

$$\begin{cases} \vec{s}_1 = (\sqrt{E}, 0, 0, \dots, 0) \\ \vec{s}_2 = (0, \sqrt{E}, 0, \dots, 0) \\ \dots \\ \vec{s}_M = (0, 0, 0, \dots, \sqrt{E}) \end{cases} \quad (7.73)$$

Come abbiamo studiato ad Algebra Lineare, per noi la M dimensione sarà o un vettore o una matrice

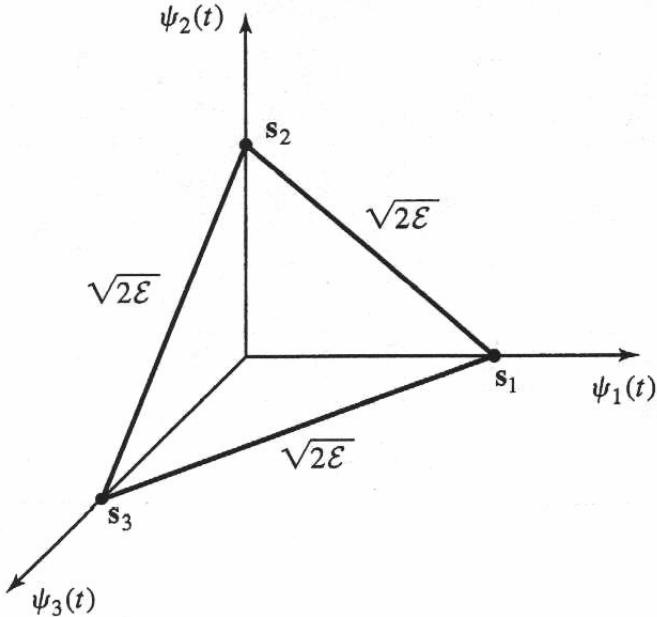
¹⁶Slide del prof — Formati di trasmissione numerica — pag 21 - 22
 Appunti di Damiano — pag 21 - 22
 Slide — Formati di trasmissione numerica — pag 21 - 22
 Appunti — 2025-03-31 — pag 6

7.5.1.1 PPM con base di impulsi rettangolari

¹⁷

Siccome l'uomo non può concepire e visualizzare la M-dimensione, cioè oltre alla terza in poi, concentriamoci al caso $M = 3$, cioè al caso tridimensionale.

Rappresentando in una figura vettoriale, possiamo rappresentare le basi di una 3-PPM con questa figura:



Ritornando al caso generale in cui M può essere un qualsiasi numero intero, le funzioni di base nel caso PPM sono:

$$\Psi_m(t) = \begin{cases} \frac{1}{\sqrt{E_g}} \cdot g \left[t - \frac{(m-1) \cdot T}{M} \right] & \text{per } \frac{(m-1) \cdot T}{M} \leq t \leq \frac{m \cdot T}{M} \\ 0 & \text{altrove} \end{cases} \quad (7.74)$$

dove $g_T(t)$ è una funzione arbitraria (nel caso della PPM, abbiamo scelto come funzione arbitraria l'impulso rettangolare).

La radice quadrata della distanza Euclidea tra due segnali è la stessa per tutte le coppie di segnali e vale:

$$\begin{aligned} d_{mn} &= \sqrt{|\vec{s}_m - \vec{s}_n|^2} \\ &= \dots \\ &= \sqrt{2 \cdot E} \end{aligned} \quad (7.75)$$

¹⁷Slide del prof — Formati di trasmissione numerica — pag 22
 Appunti di Damiano — pag 22
 Slide — Formati di trasmissione numerica — pag 22
 Appunti — 2025-03-31 — pag 6

7.5.2 FSK (Frequency Shift Keying)

¹⁸

L'ortogonalità dei segnali, ad esempio quelli che abbiamo visto nella figura della 4-PPM, viene ottenuta ragionando nel dominio del tempo.

Equivalentemente, i segnali ortogonali possono essere ottenuti anche ragionando nel dominio della frequenza.

Consideriamo, ad esempio, i seguenti segnali:

$$\begin{cases} u_1(t) = \sqrt{\frac{2 \cdot E}{T}} \cos(2\pi f_1 t) \\ u_2(t) = \sqrt{\frac{2 \cdot E}{T}} \cos(2\pi f_2 t) \end{cases} \quad (7.76)$$

queste equazioni valgono solo per $0 \leq t \leq T$.

Dal punto di vista informativo, $u_1(t)$ e $u_2(t)$ hanno la stessa ampiezza, che è $\sqrt{\frac{2 \cdot E}{T}}$, ciò che varia nel tempo è la fase perché hanno $\cos(2\pi f_1 t)$.

Inoltre, in questi segnali:

$$f_1 \neq f_2 \quad (7.77)$$

quindi, in fase di progettazione del sistema, le frequenze sono due gradi di libertà.

La frequenza di portante f_c e la differenza tra le due frequenze, sarà uguale a:

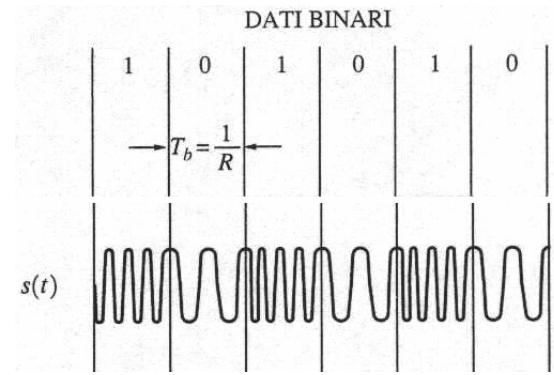
$$\begin{cases} f_c = \frac{f_1 + f_2}{2} \\ \Delta f = f_2 - f_1 \end{cases} \quad (7.78)$$

Alla fine di tutte queste osservazioni, possiamo concludere che $u_1(t)$ e $u_2(t)$ sono due funzioni oscillanti, con durata limitata all'intervallo $[0, T]$ a frequenza diversa.

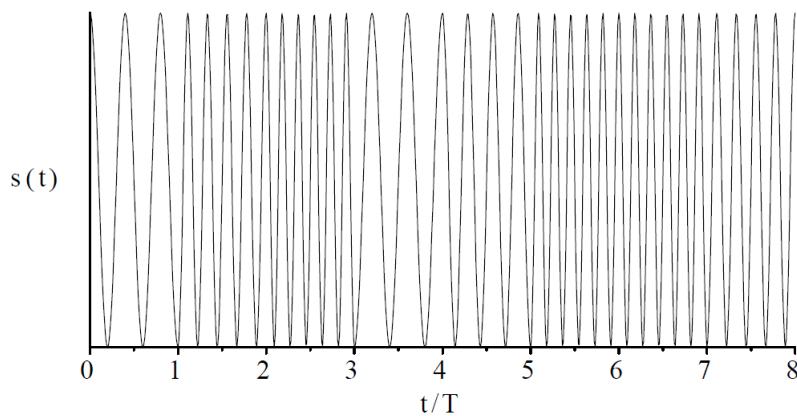
Il segnale prende il nome di FSK (Frequency Shift Keying).

¹⁸Slide del prof — Formati di trasmissione numerica — pag 22 - 24
 Appunti di Damiano — pag 22 - 24
 Slide — Formati di trasmissione numerica — pag 22 - 24
 Appunti — 2025-03-31 — pag 6 - 8
 Appunti — 2025-04-01 — pag 2
 Appunti — 2025-07-21 Ricevimento — pag 5.1

Di seguito l'andamento nel tempo di un segnale 2-FSK:



Invece, l'andamento nel tempo (normalizzato nel periodo T) di un segnale 4-FSK:



quindi nella 4-FSK ci saranno 4 frequenze differenti.

Tornando alla 2-FSK, posto:

$$f_2 = f_1 + \Delta f \quad (7.79)$$

il problema si traduce, evidentemente, nella scelta del valore di Δf .

Si può calcolare il coefficiente di correlazione che, nel caso di segnali con (approssimativamente) la stessa energia, quali quelli in esame, è così definito:

$$\gamma_{12} = \frac{\int_0^{T_b} u_1(t) \cdot u_2(t) dt}{E_b} \quad (7.80)$$

Il coefficiente di correlazione γ_{12} è un coefficiente di qualità.

I calcoli sono stati svolti considerando le due frequenze f_1 e f_2 molto vicini, così da considerare un basso su Δf dato dal rumore.

γ_{12} può assumere diversi valori, in base alle caratteristiche dei segnali $u_1(t)$ e $u_2(t)$.

Se:

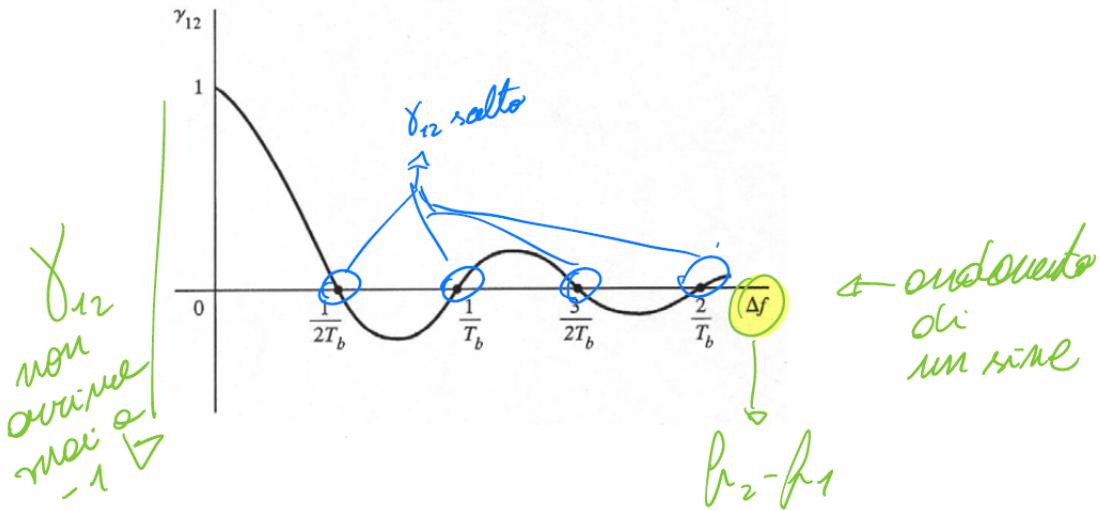
$$\gamma_{12} = -1 \quad (7.81)$$

allora:

$$u_2(t) = -u_1(t) \quad (7.82)$$

i segnali sono opposti, cioè anti-podali, e non sono ortogonali come nel caso della 2-PAM

γ_{12} rispetto a Δf ha questo tipo di andamento (con note):



Dalla figura, si nota che per:

$$\Delta f = \frac{1}{2 \cdot T_b} \quad (7.83)$$

e suoi multipli, il coefficiente di correlazione γ_{12} si annulla perchè le due forme d'onda sono tra loro ortogonali.

$u_2(t)$ si intende il suo coniugato come nel caso di calcolo dell'integrale per capire se due segnali sono ortogonali

D'altro canto, è anche possibile verificare che questo risultato, cioè la formula di $u_1(t)$ e $u_2(t)$:

$$\begin{cases} u_1(t) = \sqrt{\frac{2 \cdot E}{T}} \cos(2\pi f_1 t) \\ u_2(t) = \sqrt{\frac{2 \cdot E}{T}} \cos(2\pi f_2 t) \end{cases} \quad (7.84)$$

è valido sotto l'ipotesi che le due forme d'onda $u_1(t)$ e $u_2(t)$ abbiano identica fase iniziale (che per semplicità consideriamo nulla).

Se la fase iniziale non è nulla, e quindi possiamo scrivere $u_2(t)$ come:

$$u_2(t) = \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \cos(2\pi f_2 t + \Delta\phi) \text{ per } 0 \leq t \leq T_b \quad (7.85)$$

dove $\Delta\phi$ è la differenza di fase tra $u_1(t)$ e $u_2(t)$, allora la condizione di ortogonalità sarebbe comunque verificato, indipendentemente dal valore di $\Delta\phi$, assumendo però:

$$\Delta f = \frac{1}{T_b} \quad (7.86)$$

Controllare la fase dei segnali in un sistema di telecomunicazioni è molto complicato, perchè, come abbiamo visto nelle modulazioni angolari, il rumore può impattare tanto la ricezione.

Inoltre, il recupero della fase di un segnale va svolto con diversi componenti, tra cui il PLL, elemento complesso e difficile da implementare perchè aumenta un costo economico al ricevitore e un costo di tempo computazionale.

Più ci semplifichiamo il ricevitore, e il sistema tutto, e meglio è per tutti.

Inoltre, la figura dell'andamento di γ_{12} mette anche in evidenzia che si può conseguire un valore del coefficiente di correlazione anche più "favorevole" (dal punto di vista della differenziazione delle forme d'onda), assumendo un valore di Δf uguale a:

$$\Delta f = \frac{0.715}{T_b} \quad (7.87)$$

avremo un valore di γ_{12} uguale a:

$$\gamma_{12} = -0.217 \quad (7.88)$$

valore che è "più vicino" a $\gamma_{12} = -1$ (anche se è molto lontano), e che quindi, come abbiamo visto nella 2-PSK, garantisce una distanza tra i due punti massima.

Ma questo valore γ_{12} è difficile da gestire, quindi si sceglie di utilizzare $\gamma_{12} = 0$ perchè i segnali da generare sono ortogonali, quindi semplici da generare.

Quindi, l'ortogonalità è sempre verificata nella 2-FSK se:

$$\Delta f = \frac{k}{2 \cdot T_b} \quad (7.89)$$

con k numero intero.

Generalmente si assume:

$$k = 1 \quad (7.90)$$

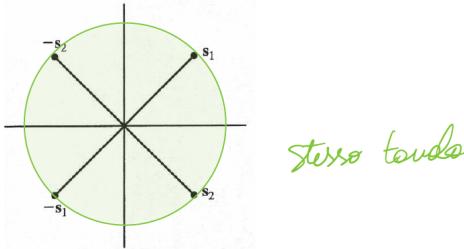
perchè permette di minimizzare l'occupazione spettrale.

7.5.3 Forme d'onda bi-ortogonali

19

A partire da un insieme di forme d'onda ortogonali, è possibile costruire un insieme di forme d'onda bi-ortogonali.

Un esempio di forme bi-ortogonali lo avevo già visto in precedenza:



Considerando M segnali, a partire dall'insieme ortogonale, si considera l'insieme delle forme d'onda opposte (sfasate cioè di 180°).

In tal modo, le M forme d'onda sono divise in due gruppi di $\frac{M}{2}$ forme d'onda.

All'interno di ciascun gruppo, le forme d'onda sono ortogonali.

Le forme d'onda di un gruppo sono, poi, opposte alle forme d'onda dell'altro gruppo.

Stanti le peculiarità dell'insieme, la dimensione dello spazio in cui sono rappresentati i segnali vale:

$$N = \frac{M}{2} \quad (7.91)$$

I vettori rappresentativi degli M segnali, avranno dunque $\frac{M}{2}$ componenti e saranno esprimibili come:

$$\left\{ \begin{array}{l} \vec{s}_1 = (\sqrt{E}, 0, 0, \dots, 0) \\ \vec{s}_2 = (0, \sqrt{E}, 0, \dots, 0) \\ \dots \\ \vec{s}_{\frac{M}{2}} = (0, 0, 0, \dots, \sqrt{E}) \\ \vec{s}_{\frac{M}{2}+1} = (-\sqrt{E}, 0, 0, \dots, 0) \\ \dots \\ \vec{s}_M = (0, 0, 0, \dots, -\sqrt{E}) \end{array} \right. \quad (7.92)$$

Quindi l'insieme dei vettori $N [\vec{s}_1, \vec{s}_2, \dots, \vec{s}_{\frac{M}{2}}]$ forma la base dello spazio M .

L'insieme dei vettori $[\vec{s}_{\frac{M}{2}+1}, \dots, \vec{s}_M]$ sono opposti ai vettori della base, cioè sono opposti ai vettori $[\vec{s}_1, \vec{s}_2, \dots, \vec{s}_{\frac{M}{2}}]$.

¹⁹Slide del prof — Formati di trasmissione numerica — pag 24 - 25
 Appunti di Damiano — pag 24 - 25
 Slide — Formati di trasmissione numerica — pag 24 - 25
 Appunti — 2025-04-01 — pag 2

7.5.4 Forme d'onda transagonali

²⁰

Accanto ai concetti di ortogonalità e bi-ortogonalità, può essere introdotto anche il concetto di transortogonalità.

Estendendo la definizione di coefficiente di correlazione γ_{mn} visto precedentemente per segnali $u_1(t)$ e $u_2(t)$ con la stessa energia E_b , possiamo estenderlo a due segnali $s_m(t)$ e $s_n(t)$ aventi rispettivamente energie E_m e E_n .

Quindi la formula del coefficiente di correlazione diventerà:

$$\begin{aligned}\gamma_{12} &= \frac{\int_0^{T_b} u_1(t) \cdot u_2(t) dt}{E_b} \\ &\downarrow \\ \gamma_{mn} &= \frac{\int_0^T s_m(t) \cdot s_n(t) dt}{\sqrt{E_m \cdot E_n}}\end{aligned}\tag{7.93}$$

Per segnali ortogonali, il coefficiente di correlazione è nullo (perchè l'integrale è nullo).

Se abbiamo delle forme d'onda che hanno la stessa energia:

$$E_m = E_n = E\tag{7.94}$$

allora, la formula del coefficiente di correlazione diventa:

$$\gamma_{mn} = \frac{\int_0^T s_m(t) \cdot s_n(t) dt}{E}\tag{7.95}$$

e si verifica che (cioè non lo dimostriamo) se lo spazio è formato da vettori bi-ortogonali:

$$\left\{ \begin{array}{l} \vec{s}_1 = (\sqrt{E}, 0, 0, \dots, 0) \\ \vec{s}_2 = (0, \sqrt{E}, 0, \dots, 0) \\ \dots \\ \vec{s}_{\frac{M}{2}} = (0, 0, 0, \dots, \sqrt{E}) \\ \vec{s}_{\frac{M}{2}+1} = (-\sqrt{E}, 0, 0, \dots, 0) \\ \dots \\ \vec{s}_M = (0, 0, 0, \dots, -\sqrt{E}) \end{array} \right. \tag{7.96}$$

il minimo valore di γ_{mn} (in modulo e segno) è:

$$\gamma_{mn} = -\frac{1}{M-1}\tag{7.97}$$

Forme d'onda con questa caratteristica si dicono trans-ortogonali o simplesse (in inglese simplex).

²⁰Slide del prof — Formati di trasmissione numerica — pag 25

Appunti di Damiano — pag 25

Slide — Formati di trasmissione numerica — pag 25

Appunti — 2025-04-01 — pag 2

7.5.4.1 Il ruolo del coefficiente di correlazione e sulla qualità della trasmissione

²¹

Il coefficiente di correlazione ha un ruolo molto importante nella determinazione della qualità della trasmissione (espressa dalla probabilità di errore).

A parità di condizioni, è preferibile il sistema caratterizzato dal minimo valore di γ_{mn} .

La situazione più favorevole, in questo caso, si ottiene per $M = 2$ quando è possibile avere $\gamma_{mn} = -1$, cioè forme d'onda anti-podali.

Chiaramente questo valore di γ_{mn} non può essere conseguibile per $M > 2$, non potendosi avere più di due forme d'onda tra loro opposte.

La rappresentazione geometrica di un insieme di segnali trans-ortogonali si ottiene dalla rappresentazione dell'insieme ortogonale da cui deriva sottraendo il valore medio dei segnali ortogonali.

In formula, e già in termini vettoriali si ha:

$$\overrightarrow{s'_m} = \overrightarrow{s_m} - \frac{1}{M} \cdot \sum_{k=1}^M \overrightarrow{s_k} \quad (7.98)$$

per $m = 1, 2, \dots, M$.

L'insieme dei vettori $\overrightarrow{s'_m}$ viene indicato i vettori simplessi.

L'effetto di sottrarre il segnale medio:

$$\overrightarrow{\bar{s}} = \frac{1}{M} \cdot \sum_{k=1}^M \overrightarrow{s_k} \quad (7.99)$$

da ogni vettore ortogonale è praticamente equivalente a traslare l'origine dello spazio M -dimensionale nel punto $\overrightarrow{\bar{s}}$ e a minimizzare l'energia dell'insieme $\overrightarrow{s'_m}$.

In effetti, l'energia dei segnali simplessi, comune per tutti i segnali, vale:

$$E' = \left(1 - \frac{1}{M}\right) \cdot E \quad (7.100)$$

dove E è l'energia sei segnali ortogonali.

Si ha quindi che:

$$E' < E \quad (7.101)$$

anche se la differenza si riduce progressivamente all'aumentare di M .

²¹Slide del prof — Formati di trasmissione numerica — pag 25 - 26
 Appunti di Damiano — pag 25 - 26
 Slide — Formati di trasmissione numerica — pag 25 - 26
 Appunti — 2025-04-01 — pag 3
 Appunti — 2025-07-21 Ricevimento — pag 5.2

Useremo i vettori simplex per rappresentare i segnali simplex.

Per sapere come si calcolano i vettori simplex:

- <https://www.youtube.com/watch?v=hQ4FKkA64z0>
Programmazione lineare strong: il metodo del simplesso #simplesso #algoritmo #programmazione by Didattica a distanza per tutti
- <https://www.youtube.com/watch?v=nabtydR4RB8>
#algoritmo del #simplesso in caso di #problemi di #minimo by Didattica a distanza per tutti

In termini di vettori rappresentativi, il coefficiente di correlazione si scrive:

$$\gamma_{mn} = \frac{\overrightarrow{s_m} \cdot \overrightarrow{s_n}}{|\overrightarrow{s_m}| \cdot |\overrightarrow{s_n}|} \quad (7.102)$$

dunque per segnali trans-ortogonali, si ha:

$$\begin{aligned} \gamma_{mn} &= \frac{\overrightarrow{s_m} \cdot \overrightarrow{s_n}}{\left| \overrightarrow{s_m} \right|^2} \\ &\dots \\ &= -\frac{1}{M-1} \end{aligned} \quad (7.103)$$

7.5.5 Dai segnali alle sequenze

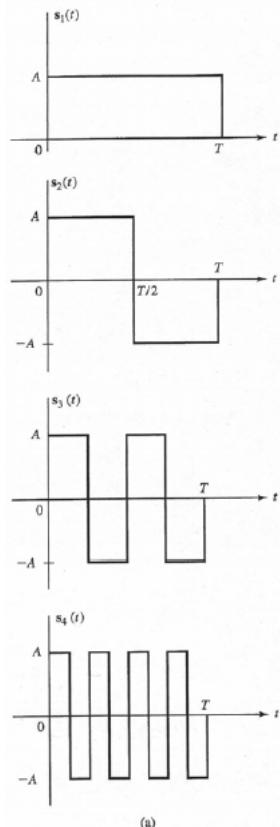
[22](#)

Questa sezione è un po' di yapping sulle M dimensioni.

Se vuoi leggila, sennò buona lettura.

Si è detto in precedenza che la sintesi di forme d'onda PPM è semplice.

Non è più, comunque, più difficile la generazione di segnali ortogonali di questo tipo:



(a)

costituite, cioè, da una successione di segmenti, in cui il segnale vale +1 o -1 (+A e -A in figura).

Segnali di questo tipo, possono essere convenzionalmente, denominati sequenze ortogonali.

Per generare un insieme di M sequenze ortogonali, possono essere utilizzate le matrici di Hadamard.

In questa sezione utilizziamo la notazione del prof per indicare le matrici.

Quindi indicheremo le matrici con il grassetto.

La matrice di Hadamard di ordine 2 ha la seguente struttura:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (7.104)$$

²²Slide del prof — Formati di trasmissione numerica — pag 27 - 29

Appunti di Damiano — pag 27 - 29

Slide — Formati di trasmissione numerica — pag 27 - 29

Appunti — 2025-04-01 — pag 3

Leggendo le colonne di \mathbf{H}_2 i conti portano perchè la prima colonna ha valori $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ che equivalgono i valori della prima sequenza nella figura $s_1(t)$, cioè: $\begin{bmatrix} A \\ A \end{bmatrix}$, invece per la seconda colonna di \mathbf{H}_2 ha valori $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, nella figura $s_2(t)$ ha valori $\begin{bmatrix} A \\ -A \end{bmatrix}$

A partire da essa, supposto che n sia una potenza di 2, si utilizza la seguente procedura iterativa:

$$\mathbf{H}_{2n} = \begin{bmatrix} \mathbf{H}_n & \mathbf{H}_n \\ \mathbf{H}_n & -\mathbf{H}_n \end{bmatrix} \quad (7.105)$$

Così, ad esempio, la matrice di Hadamard di ordine 4 si ottiene come:

$$\mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (7.106)$$

In questa maniera, per ogni colonna e dividendo ogni cella per un multiplo di T , otteniamo delle sequenze ortogonali tra loro.

A partire dalla matrice che identifica l'insieme ortogonale, è immediato ottenere la matrice che identifica l'insieme bi-ortogonale.

Indicato con $\mathbf{H}_{\frac{M}{2}}$, una matrice di Hadamard di ordine $\frac{M}{2}$ si avrà infatti la matrice dell'insieme bi-ortogonale:

$$\mathbf{H}_2|_{biortogonale} = \begin{bmatrix} \mathbf{H}_{\frac{M}{2}} \\ -\mathbf{H}_{\frac{M}{2}} \end{bmatrix} \quad (7.107)$$

Considerando $M = 4$, la matrice di Hadamard di ordine $\frac{M}{2}$ sarà:

$$\mathbf{H}_4|_{biortogonale} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (7.108)$$

Per ottenere M forme d'onda trans-ortogonali, si ottiene eliminando la prima colonna della matrice ortogonale da cui deriva.

Quindi, per $M = 4$, da:

$$\mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (7.109)$$

diventa:

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \quad (7.110)$$

7.6 Alcune considerazioni sullo spettro e la banda occupata

²³

In questo paragrafo ci limitiamo a riportare, senza dimostrazione, l'effetto sull'occupazione spettrale (banda) di un incremento del valore di M (cardinalità del formato di trasmissione).

Nelle trasmissione M -arie, ogni simbolo sostituisce k :

$$k = \log_2(M) \quad (7.111)$$

dove k sono i bit per ogni simbolo.

La durata di ciascun simbolo M -aria è k volte maggiore della durata del singolo bit.

Nella trasmissione binaria, si possono avere variazioni ogni T_b secondi, mentre nella trasmissione M -aria, il segnale può variare ogni T :

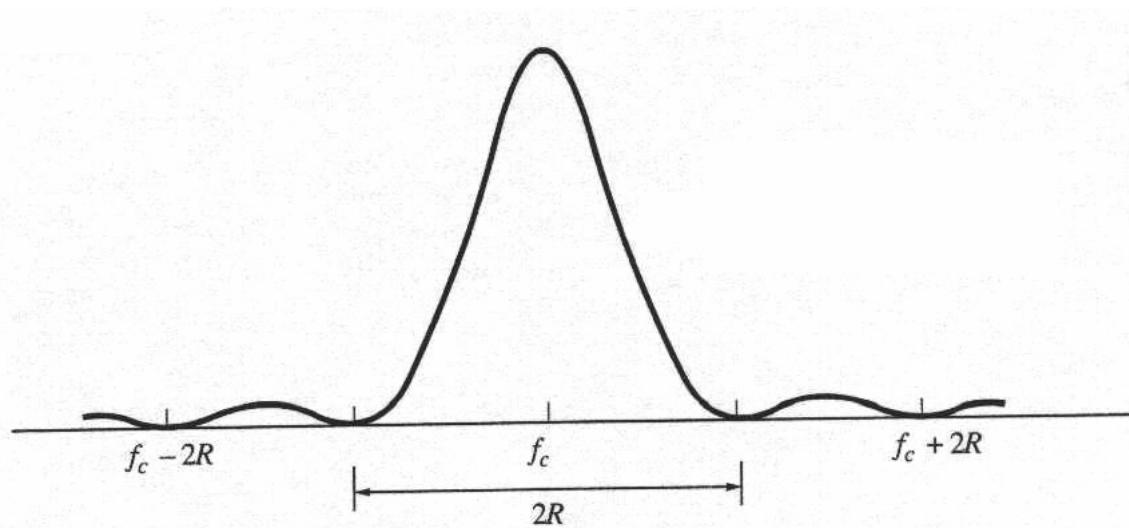
$$T = T_b \cdot \log_2(M) \quad (7.112)$$

dove con T è in secondi.

k è minore di T , quindi, nella trasmissione M -aria ci vorrà più tempo per inviare un simbolo.

Il simbolo M -aria ha un'occupazione spettrale significativa $\log_2(M)$ volte più piccola della corrispondente occupazione spettrale del segnale binario.

A titolo di esempio si riporta l'andamento dello spettro di potenza per il formato 2-ASK avendo posto $R = \frac{1}{T_b}$:



Se si considera un valore di M maggiore di 2, la forma dello spettro rimane inalterata, ma si riduce la banda intorno alla frequenza f_c .

Ma, l'aumento M significa anche ridurre la qualità della trasmissione.

Generalmente questo principio vale nella maggior parte dei formati, tranne i formati ortogonali o derivati dagli ortogonali.

Ad esempio nella PPM, all'aumentare di M si allarga la banda.

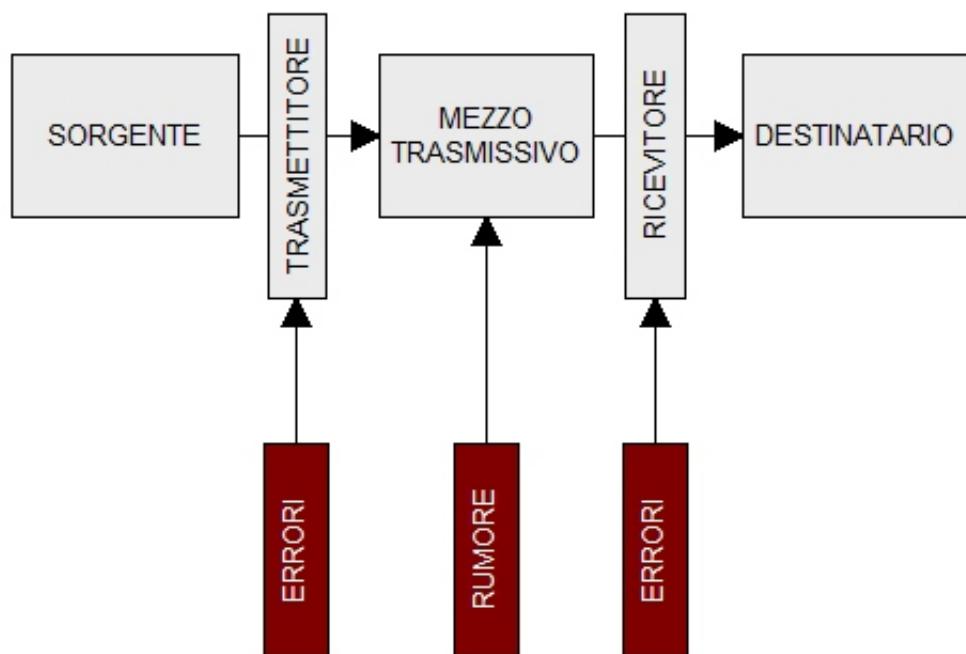
²³Slide del prof — Formati di trasmissione numerica — pag 30 - 32
 Appunti di Damiano — pag 30 - 32
 Slide — Formati di trasmissione numerica — pag 30 - 32
 Appunti — 2025-04-01 — pag 4

Inoltre, all'aumentare dalla banda, la qualità migliora e si abbassa la probabilità di errore.

Il vincolo della banda è un problema nelle trasmissione senza fili, ma nella fibra ottica abbiamo banda "illimitata", quindi non è un problema e possiamo implementare una PPM con M molto elevato.

Capitolo 8

Ricevitore numerico



WWW.ANDREAMININI.ORG

8.1 Ricevitore ottimo

1

Consideriamo questo modello per la trasmissione di un segnale ricevuto da un ricevitore:

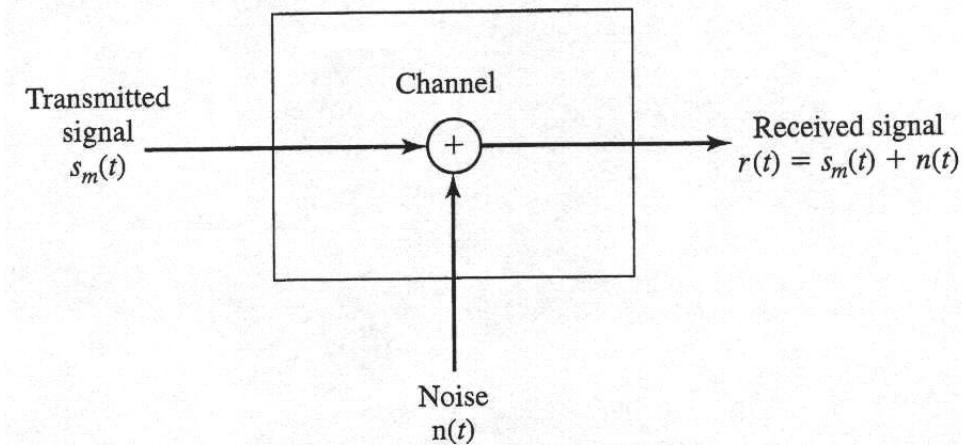
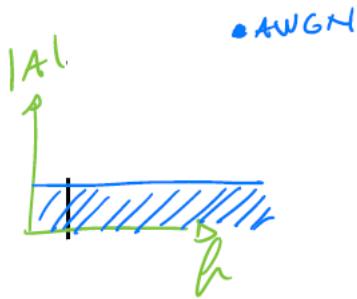


Figure 7.30 Model for received signal passed through an AWGN channel.

Questo tipo di modello vale sotto queste ipotesi:

- Non si considera l'attenuazione nel canale, quindi il segnale trasmesso $s_m(t)$ a sinistra della figura è lo stesso di quello a destra
- il rumore $n(t)$ è AWGN (Additive White Gaussian Noise), quindi è solo additivo, e non si considerano altri tipi di rumore

Essendo il rumore $n(t)$ AWGN, nello spettro il rumore sarà additivo e costante in frequenza:



Dato questo modello, l'obiettivo è progettare un ricevitore che sia ottimo, nel senso che minimizzi la probabilità di errore.

Rimarchiamo una caratteristica rispetto al caso analogico.

Siccome adesso abbiamo a che fare con sistemi numerici, il rumore impatta la probabilità di errore del simbolo, non il segnale analogico tutto perciò, nei ricevitori numerici, i simboli fanno parte di un alfabeto discreto, e non infinito come nelle modulazioni analogiche.

Per i nostri obiettivi, è meglio suddividere il ricevitore in due parti (partendo dal segnale ricevuto $r(t)$):

¹Slide del prof — Ricevitore numerico — pag 1
Appunti di Damiano — pag 1
Slide — Ricevitore numerico — pag 1
Appunti — 2025-04-01 — pag 5 - 6

- Demodulatore del segnale ricevuto
- Decisore

Il compito del demodulatore è quello di convertire la forma d'onda ricevuta $r(t)$ in un vettore N-dimensione \vec{r} con le seguenti componenti:

$$\vec{r} = (r_1, r_2, \dots, r_N) \quad (8.1)$$

Il vettore \vec{r} è la rappresentazione geometrica nella N-dimensione in cui i segnali, cioè simboli, dell'alfabeto di trasmissione.

Nel caso ideale in cui il rumore $n(t)$ è nullo:

$$\vec{r} = \vec{s}_m \quad (8.2)$$

cioè, il simbolo ricevuto \vec{r} è uguale a il simbolo trasmesso \vec{s}_m .

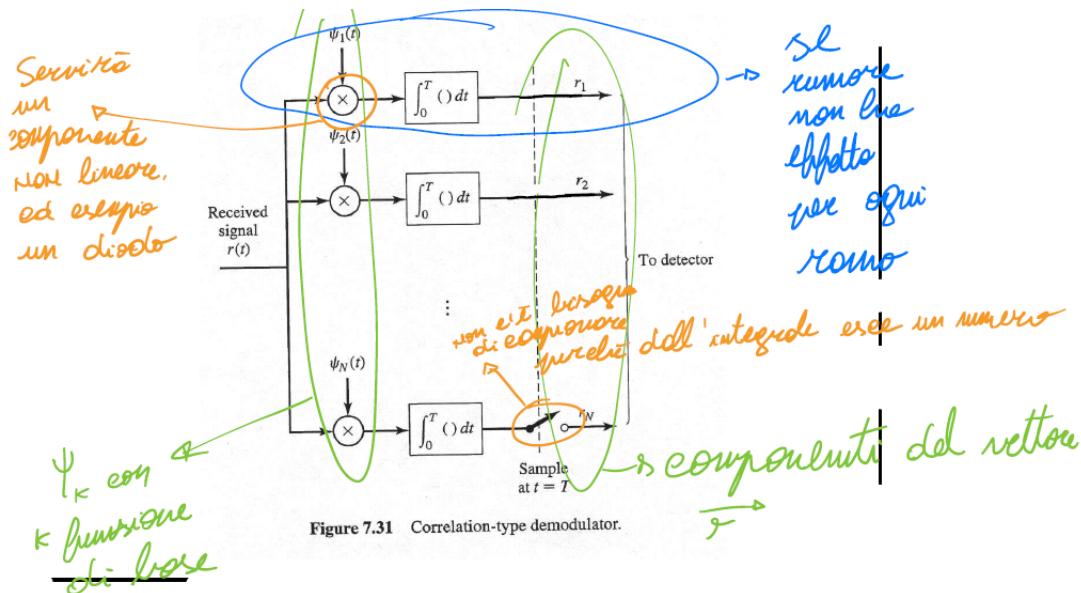
Ma, siccome il rumore $n(t)$ non è nullo e inciderà sul segnale trasmesso $s_m(t)$, allora sarà il compito del decisore decidere, dall'osservazione di \vec{r} , quale degli M simboli possibili è stato trasmesso facendo una stima, perchè il rumore è un processo stocastico, cioè probabilistico.

Dalla stima del simbolo trasmesso, è poi possibile recuperare l'informazione originale.

8.2 Demodulatore (ricevitore) a correlatore

2

Di seguito lo schema del demodulatore (ricevitore) a correlatore (con le note):



Alcune note che non ho evidenziato nello schema:

- Il blocco integratore di ogni ramo è visto, in frequenza, come un filtro passa-basso
- $\Psi_k(t)$ è un versore della base N-dimensione

Sapendo che quello che riceviamo dal demodulatore è il vettore \vec{r} :

$$\vec{r} = \vec{s}_m + \vec{n} \quad (8.3)$$

cioè il vettore ricevuto \vec{r} è la somma del segnale trasmesso \vec{s}_m e il vettore rumore \vec{n} .

Possiamo scomporre i vettori nei loro componenti.

Se consideriamo k:

$$k = 1, 2, \dots, N \quad (8.4)$$

dove N è la dimensione dello spazio, possiamo considerare ogni ramo del demodulatore un singolo componente del vettore \vec{r} , quindi anche di \vec{s}_m e di \vec{r} .

Se consideriamo il singolo ramo del demodulatore, possiamo esprimere matematicamente il segnale ricevuto come:

$$\begin{aligned} \int_0^T r(t) \cdot \Psi_k(t) dt &= \int_0^T [s_m(t) + n(t)] \cdot \Psi_k(t) dt \\ &= \int_0^T s_m(t) \cdot \Psi_k(t) dt + \int_0^T n(t) \cdot \Psi_k(t) dt \end{aligned} \quad (8.5)$$

²Slide del prof — Ricevitore numerico — pag 2
 Appunti di Damiano — pag 2
 Slide — Ricevitore numerico — pag 2
 Appunti — 2025-04-01 — pag 6
 Appunti — 2025-04-04 — pag 2 - 3

Sapendo che il singolo elemento del vettore del segnale ricevuto \vec{r}' è composto da:

$$r_k = s_{mk} + n_k \text{ per } k = 1, 2, \dots, N \quad (8.6)$$

allora possiamo scrivere la equazione precedente nei vari componenti.

Per quanto riguarda il segnale trasmesso:

$$\begin{aligned} & \int_0^T s_m(t) \cdot \Psi_k(t) dt \\ & \downarrow \\ s_{mk} &= \int_0^T s_m(t) \cdot \Psi_k(t) dt \text{ per } k = 1, 2, \dots, N \end{aligned} \quad (8.7)$$

Per quanto riguarda il rumore:

$$\begin{aligned} & \int_0^T n(t) \cdot \Psi_k(t) dt \\ & \downarrow \\ n_k &= \int_0^T n(t) \cdot \Psi_k(t) dt \text{ per } k = 1, 2, \dots, N \end{aligned} \quad (8.8)$$

Considerando il segnale trasmesso s deterministico (anche se in realtà non lo è, ma per i nostri scopi lo è), il rumore n è aleatorio perchè è un rumore termico: ciò comporta dell'errore sul singolo elemento della base $\Psi_k(t)$.

8.2.1 Caratteristiche del rumore additivo

³

Un ripasso da Teoria dei Segnali per quanto riguarda il rumore additivo AWGN ed i calcoli che andremo a svolgere.

Non dobbiamo sapere tutti i calcoli matematici qui riportati, però dobbiamo giustificare perché traiamo queste conclusioni alla fine.

Ripassa o tieni sotto, mentre leggi questo capitolo:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 13 - Processi stocastici - pag 151 - 158

Ritorniamo alla formula del rumore nel singolo ramo del demodulatore a correlatore:

$$n_k = \int_0^T n(t) \cdot \Psi_k(t) dt \text{ per } k = 1, 2, \dots, N \quad (8.9)$$

Per le considerazioni svolte nel precedente corso di Teoria dei Segnali, formalmente non possiamo indicare il rumore come un processo deterministico, bensì come processo probabilistico perché il rumore è un fenomeno stocastico.

Il rumore non è un fenomeno stoccafisso, si dice stocastico.

Lo stoccafisso lo mangi dopo, adesso pensa allo studio

Sapendo le caratteristiche del rumore AWGN, il rumore nel singolo ramo n_k è una variabile gaussiana.

Ripassa o tieni sotto, mentre leggi questo capitolo:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 14 - Rumore termico - pag 161 - 163

Se si è scelti un intervallo di tempo appropriato in base al disturbo, cioè si è scelti un intervallo $[0, T]$ adeguato al disturbo, sappiamo che la media del rumore AWGN in un periodo vale:

$$\mathbb{E}[n(t)] = 0 \quad (8.10)$$

dove con \mathbb{E} si intende il valore medio del rumore $n(t)$ nel periodo T .

Inoltre, considerando la formula di n_k :

$$n_k = \int_0^T n(t) \cdot \Psi_k(t) dt \text{ per } k = 1, 2, \dots, N \quad (8.11)$$

possiamo visualizzare l'integrale \int_0^T una somma di tempi infinitesimi dt .

³Slide del prof — Ricevitore numerico — pag 3
Appunti di Damiano — pag 3
Slide — Ricevitore numerico — pag 3
Appunti — 2025-04-04 — pag 2 - 3
Appunti — 2025-07-21 Ricevimento — pag 5.3

Un altro minuto di silenzio per aver ucciso un altro matematico.



Quindi, calcolando la media di n_k nel periodo T vale, per qualsiasi valore di $\Psi_k(t)$:

$$\begin{aligned}\mathbb{E}[n_k] &= \int_0^T \mathbb{E}[n(t)] \cdot \Psi_k(t) dt \\ &= \int_0^T 0 \cdot \Psi_k(t) dt \\ &= 0\end{aligned}\tag{8.12}$$

Invece, se si vuole calcolare il valore medio tra due rami, cioè calcolare la loro correlazione, i calcoli diventano un pochino più complicati.

Considerando il rumore di due rami, cioè n e k, calcoliamo la correlazione tra i due rami:

$$\mathbb{E}[n_k \cdot n_m] = \int_0^T \int_0^T \mathbb{E}[n(t) \cdot n(t + \tau)] \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau\tag{8.13}$$

Dalle proprietà del rumore termico, possiamo scrivere:

$$\mathbb{E}[n(t) \cdot n(t + \tau)] = \mathbb{R}(\tau)\tag{8.14}$$

dove con $\mathbb{R}(\tau)$ si indica l'auto-correlazione del rumore termico.

Quindi, la correlazione diventa:

$$\begin{aligned}\mathbb{E}[n_k \cdot n_m] &= \int_0^T \int_0^T \mathbb{E}[n(t) \cdot n(t + \tau)] \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau \\ &\downarrow\end{aligned}\tag{8.15}$$

$$\mathbb{E}[n_k \cdot n_m] = \int_0^T \int_0^T \mathbb{R}(\tau) \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau$$

Sempre dal corso di Teoria segnali, sappiamo che:

$$\mathbb{R}(\tau) = \frac{N_0}{2} \delta(\tau) \quad (8.16)$$

dove con N_0 si intende la densità spettrale del rumore.

Quindi:

$$\begin{aligned} \mathbb{E}[n_k \cdot n_m] &= \int_0^T \int_0^T \mathbb{R}(\tau) \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau \\ &\downarrow \\ \mathbb{E}[n_k \cdot n_m] &= \int_0^T \int_0^T \left[\frac{N_0}{2} \cdot \delta(\tau) \right] \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau \end{aligned} \quad (8.17)$$

Andando a fare delle semplici sostituzioni algebriche:

$$\begin{aligned} \mathbb{E}[n_k \cdot n_m] &= \int_0^T \int_0^T \left[\frac{N_0}{2} \cdot \delta(\tau) \right] \cdot \Psi_k(t) \cdot \Psi_m(t + \tau) dt d\tau \\ &= \int_0^T \frac{N_0}{2} \cdot \left[\int_0^T \delta(\tau) \cdot \Psi_m(t + \tau) d\tau \right] \cdot \Psi_k(t) dt \\ &= \int_0^T \frac{N_0}{2} \cdot \Psi_m(t) \cdot \Psi_k(t) dt \\ &= \frac{N_0}{2} \cdot \int_0^T \Psi_m(t) \cdot \Psi_k(t) dt \\ &= \frac{N_0}{2} \cdot \delta_{mk} \end{aligned} \quad (8.18)$$

dove δ_{mk} prende il nome di delta di Kronecker e vale:

$$\delta_{mk} = \begin{cases} 1 & \text{per } m = k \\ 0 & \text{per } m \neq k \end{cases} \quad (8.19)$$

Tutte queste formule per dire che il rumore del singolo ramo è indipendente dal rumore degli altri rami.

Quindi, da tutte queste osservazioni, possiamo dire che, in un ricevitore numerico a correlatore, le N componenti di rumore n_k sono variabili aleatorie gaussiane incorrelate, con valore medio nullo e varianza:

$$\sigma_n^2 = \frac{N_0}{2} \quad (8.20)$$

Ma questo lo sapevamo già dal corso precedente

Scrivendo un'unica formula f per mettere insieme tutte le densità congiunte del rumore del ricevitore di ogni singolo ramo i, possiamo scriverla con:

$$\begin{aligned}
 f(\vec{n}) &= \sum_{i=1}^N f(n_i) \\
 &= \dots \\
 &= \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp\left(-\frac{\sum_i^N n_i^2}{N_0}\right)
 \end{aligned} \tag{8.21}$$

dove:

- n_i indica il rumore del singolo ramo
- $f(\vec{n})$ è una funzione di variabili gaussiane miste

Considerando la ricezione del simbolo ricevuto come un fenomeno probabilistico, è necessario capire la sua varianza

8.2.2 Caratteristiche del vettore in uscita dal correlatore

4

Recuperando lo schema della trasmissione del segnale:

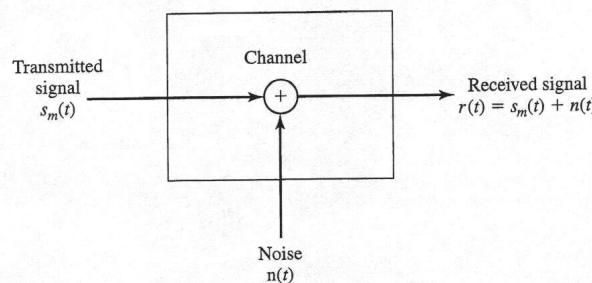
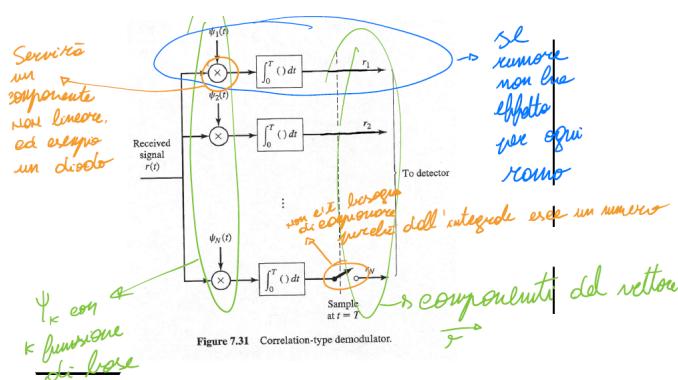


Figure 7.30 Model for received signal passed through an AWGN channel.

e lo schema del demodulatore (ricevitore) a correlatore (con le note):



possiamo scrivere che la media nel singolo ramo del demodulatore a correlatore vale:

$$\mathbb{E}[r_k] = \mathbb{E}[s_{mk} + n_k] = \mathbb{E}[s_{mk}] = s_{mk} \quad (8.22)$$

$\mathbb{E}[s_{mk}] = s_{mk}$ all'inizio sembra strano, ma guardati l'esempio della sezione successiva sull'esercizio della M-PAM.

Perché, per esteso dovrebbe essere $\mathbb{E}[s_{mk}(t)] = s_{mk}$, cioè la media nel tempo del segnale s_{mk}

Inoltre poniamo che la varianza del rumore è uguale alla varianza del segnale, quindi:

$$\sigma_r^2 = \sigma_n^2 = \frac{N_0}{2} \quad (8.23)$$

In questo caso, consideriamo il caso più semplice, cioè che il simbolo sia deterministico, e che quindi ha varianza nulla, ma su cui è sommato il rumore termico, che ha varianza $\frac{N_0}{2}$.

Quindi, la varianza totale da considerare del segnale ricevuto è uguale a quella del rumore termico

Quindi, consideriamo la densità congiunta tra il vettore ricevuto e il vettore segnale trasmesso:

⁴Slide del prof — Ricevitore numerico — pag 4
Appunti di Damiano — pag 4
Slide — Ricevitore numerico — pag 4
Appunti — 2025-04-04 — pag 3
Appunti — 2025-07-21 Ricevimento — pag 6.1

$$f(\vec{r} | \vec{s}_m) = \sum_{k=1}^N f(r_k | s_{mk}) \text{ per } m = 1, 2, \dots, M \quad (8.24)$$

Calcoliamo la densità condizionata perché presupponiamo che sia stato trasmesso un determinato \vec{s}_m (alche \vec{s}_m lo consideriamo deterministico).

Considerano i singoli rami dal correlatore, possiamo scrivere:

$$f(r_k | s_{mk}) = \frac{1}{\sqrt{\pi \cdot N_0}} \cdot \exp \left[-\frac{(r_k - s_{mk})^2}{N_0} \right] \text{ per } k = 1, 2, \dots, N \quad (8.25)$$

e quindi possiamo riscrivere la densità congiunta tra il vettore ricevuto e il vettore segnale trasmesso come:

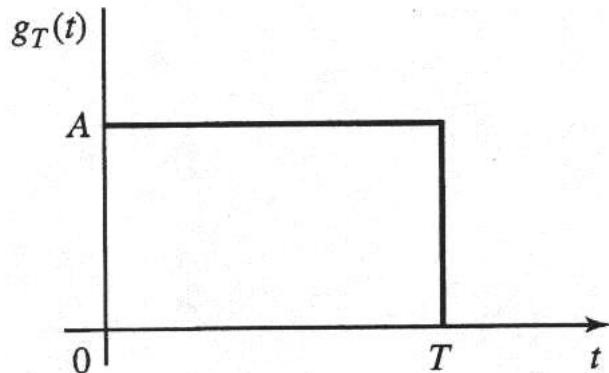
$$\begin{aligned} f(\vec{r} | \vec{s}_m) &= \sum_{k=1}^N f(r_k | s_{mk}) \text{ per } m = 1, 2, \dots, M \\ &\downarrow \\ f(\vec{r} | \vec{s}_m) &= \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp \left[-\frac{\sum_{k=1}^N (r_k - s_{mk})^2}{N_0} \right] \text{ per } m = 1, 2, \dots, M \\ &= \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp \left[-\frac{|\vec{r} - \vec{s}_m|^2}{N_0} \right] \text{ per } m = 1, 2, \dots, M \end{aligned} \quad (8.26)$$

dove con $|\vec{r} - \vec{s}_m|$ si indica il modulo del vettore differenza che si ricava sottraendo il vettore \vec{r} da \vec{s}_m .

8.2.2.1 M-PAM con demodulatore a correlatore

5

Ricordando che la M-PAM è formata da una sola funzione di base (vedi capitolo 7.3.0.1 Esempi di modulazioni con $N = 1$ pag 206), come ad esempio quella rettangolare:



Calcoliamo l'uscita del segnale dal demodulatore a correlatore :

$$\begin{aligned}
 r &= \int_0^T r(t) \cdot \Psi(t) dt \\
 &= \int_0^T r(t) \cdot \frac{1}{\sqrt{T}} dt \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T r(t) dt \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T [s_m(t) + n(t)] dt \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T s_m(t) dt + \frac{1}{\sqrt{T}} \cdot \int_0^T n(t) dt \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T s_m(t) dt + n \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T s_m \cdot \Psi(t) dt + n \\
 &= \frac{1}{\sqrt{T}} \cdot \int_0^T s_m \cdot \frac{1}{\sqrt{T}} dt + n \\
 &= \frac{1}{\sqrt{T}} \cdot \frac{1}{\sqrt{T}} \cdot s_m \int_0^T dt + n \\
 &= \frac{1}{T} \cdot s_m \cdot T + n \\
 &= s_m + n
 \end{aligned} \tag{8.27}$$

dove:

⁵Slide del prof — Ricevitore numerico — pag 5.1
Slide — Ricevitore numerico — pag 5.1
Appunti — 2025-04-04 — pag 3

$$s_m(t) = s_m \cdot \Psi(t) \quad (8.28)$$

e con n si indicano i contributi di rumore.

Sembra un piccolo dettaglio, ma semplifica di tanto i calcoli

Quindi, utilizzando la formula precedente:

$$f(\vec{r} | \vec{s}_m) = \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp \left[-\frac{|\vec{r} - \vec{s}_m|^2}{N_0} \right] \quad (8.29)$$

e adattandola al caso $N = 1$ per la M-PAM, abbiamo che:

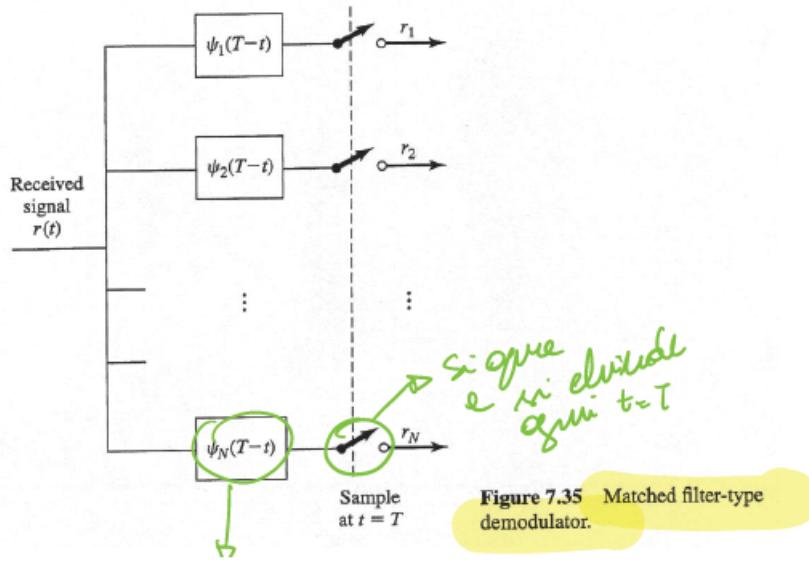
$$f(r | s_m) = \frac{1}{\sqrt{\pi \cdot N_0}} \cdot \exp \left[-\frac{(r - s_m)^2}{N_0} \right] \quad (8.30)$$

8.3 Demodulatore (ricevitore) a filtri adattati (Matched-Filter)

6

Un altro tipo di demodulatore, oltre a quello a correlatore, è quello a filtri adattati (in inglese Matched-Filter).

Di seguito uno schema dell'architettura (con note):



sistema lineare
con una
risposta
impulsiva
nel tempo

Rispetto al demodulatore a correlatore visto precedentemente:

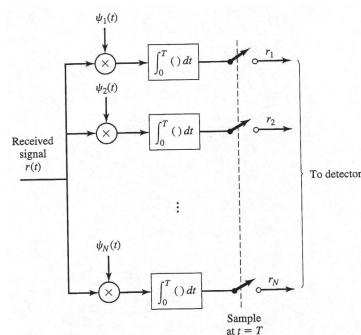


Figure 7.31 Correlation-type demodulator.

⁶Slide del prof — Ricevitore numerico — pag 5.2 - 6.1
Appunti di Damiano — pag 5.2 - 6.1
Slide — Ricevitore numerico — pag 5.2 - 6.1
Appunti — 2025-04-04 — pag 4
Appunti — 2025-07-21 Ricevimento — pag 6.2

La formula che esce dal demodulatore a filtri adattati è la stessa del demodulatore a correlatore, ma, in questo caso, non ha bisogno di sistemi non lineari.

Come si vede dalla figura, non bisogna moltiplicare il singolo rame per la base (moltiplicare significa avere un componente non lineare come il diodo), bensì ci sono dei filtri, quindi sistemi e componenti lineari, con le rispettivi componenti della base.

Per quanto riguarda gli interruttori, qui vanno alzati e abbassati ogni tempo T.

La risposta impulsiva del sistema lineare nel singolo ramo k del demodulatore è il seguente:

$$h_k(t) = \Psi_k(T - t) \text{ per } 0 \leq t \leq T \quad (8.31)$$

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 5 - Sistemi lineari - pag 41

Dal capitolo dei sistemi lineari, svolto nel precedente corso, sappiamo che l'uscita da un filtro $y_k(t)$ è la convoluzione tra il segnale di ingresso $r(\tau)$ e la risposta impulsiva del filtro $h_k(t)$.

In formule:

$$y_k(t) = \int_{-\infty}^{+\infty} r(\tau) \cdot h_k(t - \tau) d\tau \quad (8.32)$$

Si vuole progettare un filtro che abbiamo, oltre l'intervallo che non ci serve:

$$h_k(t) = 0 \text{ per } t < 0 \text{ e } t > T \quad (8.33)$$

quindi, riprendendo la formula impulsiva nel filtro in cui si fa una convoluzione, significa che:

$$\begin{aligned} h_k(t - \tau) &= 0 \text{ per } t - \tau < 0 \text{ e } t - \tau > T \\ &\downarrow \\ h_k(t - \tau) &= 0 \text{ per } t < \tau \text{ e } t - T > \tau \end{aligned} \quad (8.34)$$

oppure:

$$h_k(t - \tau) = 0 \text{ per } \tau > t \text{ e } \tau < t - T \quad (8.35)$$

Quindi possiamo limitare da $[-\infty, +\infty]$ all'intervallo $[t - T, t]$, e possiamo riscrivere l'equazione della convoluzione come:

$$\begin{aligned} y_k(t) &= \int_{-\infty}^{+\infty} r(\tau) \cdot h_k(t - \tau) d\tau \\ &\downarrow \\ y_k(t) &= \int_{t-T}^t r(\tau) \cdot h_k(t - \tau) d\tau \end{aligned} \quad (8.36)$$

Siccome sappiamo che:

$$h_k(t) = \Psi_k(T - t) \text{ per } 0 \leq t \leq T \quad (8.37)$$

possiamo cambiare ulteriormente la risposta impulsiva come:

$$\begin{aligned} y_k(t) &= \int_{t-T}^t r(\tau) \cdot h_k(t - \tau) d\tau \\ &\downarrow \\ y_k(t) &= \int_0^t r(\tau) \cdot \Psi_k(T - t + \tau) d\tau \end{aligned} \quad (8.38)$$

dove $k = 1, 2, \dots, N$.

Ponendo:

$$t = T \quad (8.39)$$

dove T è il periodo del segnale, allora possiamo considerare la risposta impulsiva nel periodo T del filtro:

$$\begin{aligned} y_k(t) &= \int_0^t r(\tau) \cdot \Psi_k(T - t + \tau) d\tau \\ &\downarrow \\ y_k(T) &= \int_0^T r(\tau) \cdot \Psi_k(T - T + \tau) d\tau \\ &= \int_0^T r(\tau) \cdot \Psi_k(\tau) d\tau \\ &= r_k \end{aligned} \quad (8.40)$$

dove:

- $k = 1, 2, \dots, N$
- r_k è un elemento del vettore \vec{r} ricevuto

La scelta di utilizzare dei componenti lineari o non lineari in fase di progetto dipende da "quello che si ha a scaffale".

Per la teoria, il prof preferisce utilizzare il demodulatore a filtri adattati perchè si fa una moltiplicazione in frequenza, e che, grazie alla relazione tra tempo e frequenza studiata nel vecchio corso, implica una correlazione nel tempo.

8.3.1 Filtro adattato

7

Spiegazione di cosa è un filtro adattato in generale, che non per forza deve essere collegato ad una trasmissione numerica.

Un filtro la cui risposta impulsiva è:

$$h(t) = s(T - t) \text{ per } 0 \leq t \leq T \quad (8.41)$$

si dice adattato al segnale $s(t)$ (dove $s(t)$ è il segnale di ingresso del filtro).

L'uscita del filtro vale:

$$\begin{aligned} y(t) &= \int_0^t s(\tau) \cdot h(t - \tau) d\tau \\ &= \int_0^t s(\tau) \cdot h(T - t + \tau) d\tau \end{aligned} \quad (8.42)$$

Consideriamo come esempio l'andamento nel tempo del segnale $s(t)$ in ingresso al filtro adattato (figura a sinistra) e l'andamento del tempo della risposta impulsiva $h(t)$ del filtro adattato:

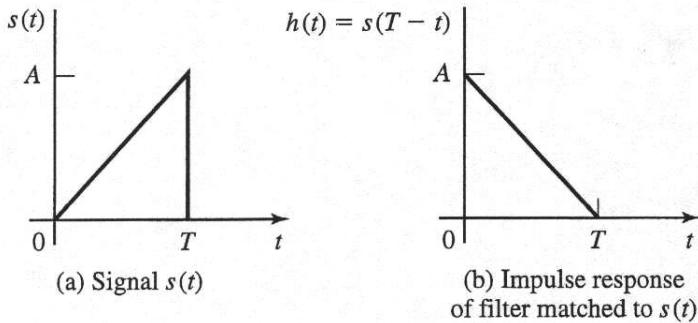


Figure 7.33 Signal $s(t)$ and filter matched to $s(t)$.

L'uscita dal filtro adattato $y(t)$ è la seguente:

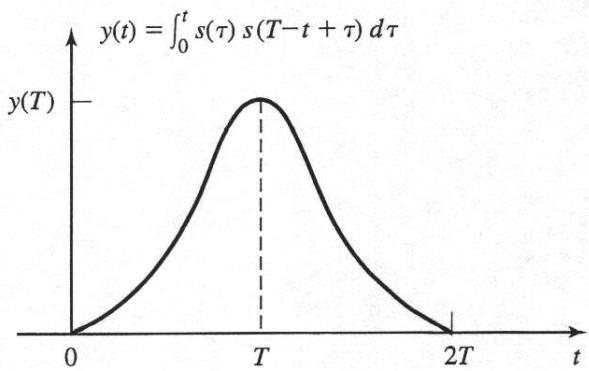


Figure 7.34 Matched filter output is the autocorrelation function of $s(t)$.

Generalmente si prende l'uscita $y(t)$ al tempo $t = T$, perché, come si nota dalla figura, si ha la potenza massima del filtro.

⁷Slide del prof — Ricevitore numerico — pag 6.2 - 7.1
 Appunti di Damiano — pag 6.2 - 7.1
 Slide — Ricevitore numerico — pag 6.2 - 7.1
 Appunti — 2025-04-04 — pag 4

8.3.1.1 Proprietà del filtro adattato

⁸

Analizzando in frequenza, cioè in Fourier, la risposta impulsiva del filtro adattato:

$$\begin{aligned}
 H(f) &= \int_0^T s(T-t) \cdot e^{-j2\pi ft} dt \\
 &= \left[\int_0^T s(\tau) \cdot e^{-j2\pi f\tau} d\tau \right] \cdot e^{-j2\pi fT} \\
 &= S^*(f) \cdot e^{-j2\pi fT}
 \end{aligned} \tag{8.43}$$

Questi passaggi matematici dimostrano che la trasformata di Fourier di un filtro adattato $h(t)$, cioè $H(f)$, è uguale al complesso coniugato della trasformata di Fourier di $s(t)$, cioè $S^*(f)$, ritardata di un tempo T : questo significa adattarsi al segnale di ingresso $s(t)$.

Per quanto riguarda il modulo del filtro adattato:

$$|H(f)| = |S(f)| \tag{8.44}$$

cioè il modulo del filtro adattato è uguale al modulo del segnale in ingresso.

⁸Slide del prof — Ricevitore numerico — pag 7.2
 Appunti di Damiano — pag 7.2
 Slide — Ricevitore numerico — pag 7.2
 Appunti — 2025-04-04 — pag 4

8.3.1.2 SNR in uscita dal filtro adattato

9

Ricordo che i calcoli matematici non ci servono al fine dell'esame orale.

Servono solo per la dimostrazione.

Le cose importanti sono le conclusioni e quello che scrivo a parole.

Consideriamo l'uscita di un filtro adattato utilizzando la frequenza.

Avremo:

$$\begin{aligned}
 y_s(t) &= \int_{-\infty}^{+\infty} Y(f) \cdot e^{j2\pi ft} df \\
 &= \int_{-\infty}^{+\infty} [S(f) \cdot H(f)] \cdot e^{j2\pi ft} df \\
 &= \int_{-\infty}^{+\infty} S(f) \cdot [S^*(f) \cdot e^{-j2\pi T}] \cdot e^{j2\pi ft} df \\
 &= \int_{-\infty}^{+\infty} |S(f)|^2 \cdot e^{-j2\pi T} \cdot e^{j2\pi ft} df
 \end{aligned} \tag{8.45}$$

Siccome consideriamo l'uscita al tempo t:

$$t = T \tag{8.46}$$

perchè abbiamo la potenza massima del filtro, allora l'uscita $y_s(t)$ diventa:

$$\begin{aligned}
 y_s(t) &= \int_{-\infty}^{+\infty} |S(f)|^2 \cdot e^{-j2\pi T} \cdot e^{j2\pi ft} df \\
 &\downarrow \\
 y_s(T) &= \int_{-\infty}^{+\infty} |S(f)|^2 df
 \end{aligned} \tag{8.47}$$

Grazie al teorema di Parseval:

$$|S(f)|^2 = |s(t)|^2 \tag{8.48}$$

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 4.2.1 - Teorema di Parseval - pag 32

Quindi, possiamo scrivere:

⁹Slide del prof — Ricevitore numerico — pag 8.1
Appunti di Damiano — pag 8.1
Slide — Ricevitore numerico — pag 8.1
Appunti — 2025-04-04 — pag 5

$$\begin{aligned}
 y_s(T) &= \int_{-\infty}^{+\infty} |S(f)|^2 df \\
 &\downarrow \\
 y_s(T) &= \int_{-\infty}^{+\infty} |s(t)|^2 df \\
 &= E_s
 \end{aligned} \tag{8.49}$$

dove E_s si intende l'energia del segnale.

Sapendo che la potenza del rumore S_{n_0} è:

$$S_{n_0}(f) = |H(f)|^2 \cdot \frac{N_0}{2} \tag{8.50}$$

La potenza del rumore P_n vale:

$$\begin{aligned}
 P_n &= \int_{-\infty}^{+\infty} S_{n_0}(f) df \\
 &= \int_{-\infty}^{+\infty} \frac{N_0}{2} \cdot |H(f)|^2 df \\
 &= \frac{N_0}{2} \cdot \int_{-\infty}^{+\infty} |H(f)|^2 df \\
 &= \frac{N_0}{2} \cdot \int_{-\infty}^{+\infty} |S(f)|^2 df \\
 &= \frac{N_0}{2} \cdot E_s
 \end{aligned} \tag{8.51}$$

Mettendo in relazione tutti i dati, calcoliamo il rapporto segnale rumore del filtro adattato:

$$\begin{aligned}
 \left(\frac{S}{N}\right)_o &= \frac{[y_s(T)]^2}{P_n} \\
 &= \frac{(E_s)^2}{\frac{N_0}{2} \cdot E_s} \\
 &= \frac{2 \cdot E_s}{N_0}
 \end{aligned} \tag{8.52}$$

Se ci mettiamo a $t = T$, con un filtro adattato, questo è il rapporto segnale rumore massimo in uscita.

8.4 Decisore

¹⁰

Come scritto precedentemente, in un ricevitore numerico, dopo il demodulatore è necessario un decisore.

Consideriamo come esempio grafico il caso di una trasmissione, dopo il demodulatore, in $N = 3$, con $M = 4$, nella costellazione:

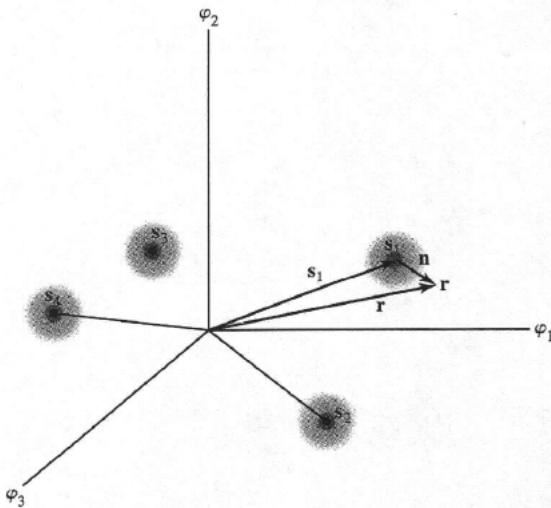


Figure 7.37 Signal constellation, noise cloud and received vector for $N = 3$ and $M = 4$. It is assumed that s_1 is transmitted.

Si considerano sistemi cooperanti.

Per sistemi cooperanti si intende che il decisore collabora con il demodulatore, cioè il demodulatore "passa" l'informazione del segnale ricevuto al decisore.

Se invece si considera un sistema non cooperante, il decisore dovrebbe prendere una decisione "alla cieca", da solo

Spiegando graficamente la costellazione, se non consideriamo il rumore che riceviamo in trasmissione, in ricezioni dovremmo ricevere il vettore originariamente trasmesso.

Cioè, in formule, se non esistesse il rumore termico:

$$\vec{r} = \vec{s}_m \quad (8.53)$$

MA, siccome esiste il rumore termico:

$$\vec{r} \neq \vec{s}_m \quad (8.54)$$

Il rumore varia la posizione di \vec{s}_m nella costellazione e, come indicato nella figura, \vec{r} si troverà in un intorno di \vec{s}_m .

Inoltre, essendo il rumore termico AWGN un processo stocastico, quindi probabilistico, dobbiamo utilizzare la probabilità.

In questo caso, si possono calcolare due tipi di probabilità:

¹⁰Slide del prof — Ricevitore numerico — pag 8.2 - 9.1
Appunti di Damiano — pag 8.2 - 9.1
Slide — Ricevitore numerico — pag 8.2 - 9.1
Appunti — 2025-04-04 — pag 5 - 6
Appunti — 2025-07-21 Ricevimento — pag 7

- probabilità massima, che è una probabilità che possiamo calcolare prima della trasmissione, cioè a priori
- probabilità massima condizionata dal segnale ricevuto, che è una probabilità che possiamo solo dopo aver ricevuto il segnale \vec{r} , cioè a posteriori

Ovviamente, a noi interessa di più la seconda, quindi vogliamo calcolare la seguente probabilità:

$$\Pr\{\vec{s}_m | \vec{r}\} \quad (8.55)$$

8.4.1 Criterio MAP: Maximum A Posteriori probability

¹¹

Il decisore sarà comparativo: cioè confronterà la probabilità a priori di \vec{s}_m e la confronterà con il segnale ricevuto \vec{r} .

In particolare, con il criterio MAP, il decisore seleziona, come segnale trasmesso, quello che massimizza la $\Pr\{\vec{s}_m | \vec{r}\}$:

$$\Pr\{\vec{s}_m | \vec{r}\} = \frac{f(\vec{r} | \vec{s}_m) \cdot \Pr\{\vec{s}_m\}}{f(\vec{r})} \quad (8.56)$$

Questa formula è l'applicazione della formula di Bayes applicata alla costellazione, quindi in ambito vettoriale.

La avevamo già vista nel corso di Teoria dei Segnali.

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.5 - Eventi statisticamente indipendenti - pag 118 - 119

$$\Pr\{A|B\} = \frac{\Pr\{B|A\} \cdot \Pr\{A\}}{\Pr\{B\}} \quad (8.57)$$

Se invece sei un iPad kid (o iPad girl, qui non facciamo discriminazioni), puoi vedere i seguenti video per capire meglio la formula di Bayes:

- <https://www.youtube.com/watch?v=HZGCoVF3YvM&t=249s>
Bayes theorem, the geometry of changing beliefs by 3Blue1Brown (per visualizzare l'applicazione della formula di Bayes)
- <https://www.youtube.com/watch?v=0Byl4RJxnKA>
Bayes' Theorem of Probability With Tree Diagrams & Venn Diagrams by The Organic Chemistry Tutor (per visualizzarla e fare dei semplici esercizi di calcolo)

Quindi per massimizzare la formula di $\Pr\{\vec{s}_m | \vec{r}\}$, bisogna massimizzare il numeratore, quindi $f(\vec{r} | \vec{s}_m) \cdot \Pr\{\vec{s}_m\}$ deve essere la più grande possibile.

Calcolare $f(\vec{r} | \vec{s}_m)$ significa confrontare il vettore \vec{r} ricevuto con tutti i possibili \vec{s}_m della costellazione.

Possiamo ignorare il denominatore perché $f(\vec{r})$ è un fattore moltiplicativo, lo stesso vale per $\Pr\{\vec{s}_m\}$.

Quindi alla fine, ciò che dovremmo massimizzare è $f(\vec{r} | \vec{s}_m)$.

Inoltre, il criterio MAP è il miglior criterio di decisione a posteriori, ma lo si può applicare solo se i vettori \vec{s}_m sono equiprobabili.

¹¹Slide del prof — Ricevitore numerico — pag 9.2
Appunti di Damiano — pag 9.2
Slide — Ricevitore numerico — pag 9.2
Appunti — 2025-04-04 — pag 6

8.4.2 Criterio ML: Maximum Likelihood

¹²

Ipotizzando che i simboli dell'alfabeto sono equiprobabili, possiamo scrivere:

$$\Pr\{\vec{s}_m\} = \frac{1}{M} \quad (8.58)$$

Siccome dobbiamo massimizzare $f(\vec{r}|\vec{s}_m)$, dalla sezione 8.2.2 Caratteristiche del vettore in uscita dal correlatore a pag 259, abbiamo visto che:

$$f(\vec{r}|\vec{s}_m) = \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp \left[-\frac{|\vec{r} - \vec{s}_m|^2}{N_0} \right] \text{ per } m = 1, 2, \dots, M \quad (8.59)$$

Siccome nella formula abbiamo un esponenziale, quindi con una semplice manipolazione matematica, possiamo scrivere:

$$\begin{aligned} f(\vec{r}|\vec{s}_m) &= \frac{1}{(\pi \cdot N_0)^{\frac{N}{2}}} \cdot \exp \left[-\frac{|\vec{r} - \vec{s}_m|^2}{N_0} \right] \\ &\downarrow \\ \ln [f(\vec{r}|\vec{s}_m)] &= -\frac{N}{2} \cdot \ln(\pi \cdot N_0) - \frac{|\vec{r} - \vec{s}_m|^2}{N_0} \\ &= -\frac{N}{2} \cdot \ln(\pi \cdot N_0) - \frac{D(\vec{r}, \vec{s}_m)}{N_0} \end{aligned} \quad (8.60)$$

dove con $D(\vec{r}, \vec{s}_m)$ si intende la distanza tra i vettori \vec{r} e \vec{s}_m .

Quindi, adesso dobbiamo calcolare la minima distanza $D(\vec{r}, \vec{s}_m)$ perchè tutti gli altri parametri sono costanti.

¹²Slide del prof — Ricevitore numerico — pag 10.1
 Appunti di Damiano — pag 10.1
 Slide — Ricevitore numerico — pag 10.1
 Appunti — 2025-04-04 — pag 6

8.4.3 Distanza Euclidea minima

¹³

Esplodendo la formula di $D(\vec{r}, \vec{s}_m)$ nei vari componenti della base N del vettore, possiamo sviluppare il quadrato come:

$$\begin{aligned} D(\vec{r}, \vec{s}_m) &= |\vec{r} - \vec{s}_m|^2 \\ &= |\vec{r}|^2 - 2 \cdot \vec{r} \cdot \vec{s}_m + |\vec{s}_m|^2 \end{aligned} \quad (8.61)$$

per $m = 1, 2, \dots, M$.

Quindi per minimizzare la distanza $D(\vec{r}, \vec{s}_m)$, dobbiamo minimizzare i termini $-2 \cdot \vec{r} \cdot \vec{s}_m + |\vec{s}_m|^2$.

Il decisore individua il vettore \vec{s}_m che minimizza:

$$D'(\vec{r}, \vec{s}_m) = -2 \cdot \vec{r} \cdot \vec{s}_m + |\vec{s}_m|^2 \quad (8.62)$$

cioè applicando la metrica di distanza.

Oppure equivalentemente, il vettore \vec{s}_m che massimizza:

$$C(\vec{r}, \vec{s}_m) = 2 \cdot \vec{r} \cdot \vec{s}_m - |\vec{s}_m|^2 \quad (8.63)$$

cioè applicando la metrica di correlazione.

Per i formati in cui i simboli hanno tutti la stessa energia, nel calcolo di $D'(\vec{r}, \vec{s}_m)$ o di $C(\vec{r}, \vec{s}_m)$, il termine $|\vec{s}_m|^2$ può essere ignorato.

La distanza minima $D'(\vec{r}, \vec{s}_m)$ è l'opposto della correlazione $C(\vec{r}, \vec{s}_m)$.

Quindi bisogna avere sia \vec{r} e \vec{s}_m per calcolare questi due parametri: ciò è possibile solo in sistemi cooperanti

¹³Slide del prof — Ricevitore numerico — pag 10.2
Appunti di Damiano — pag 10.2
Appunti — 2025-07-21 Ricevimento — pag 8.1

Capitolo 9

Prestazioni dei sistemi numerici

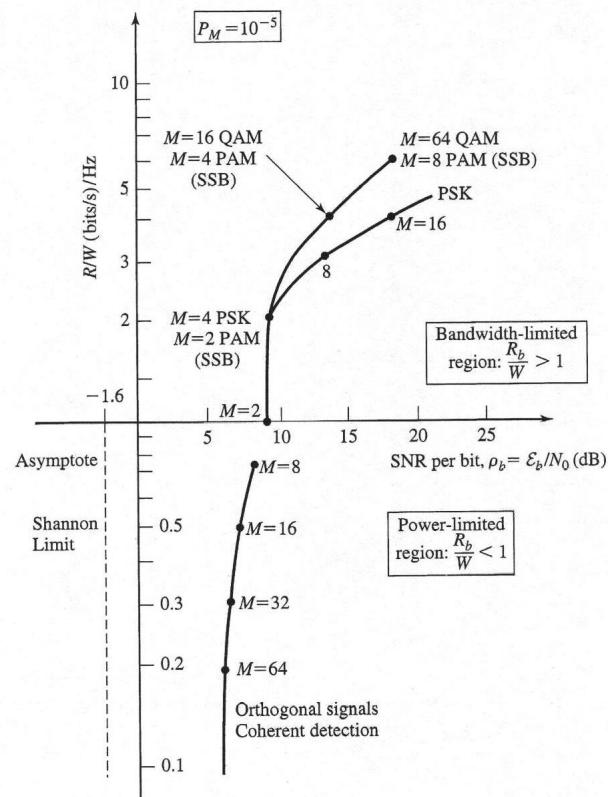


Figure 7.66 Comparison of several modulation methods at 10^{-5} symbol error probability.

Questo capitolo sarà un capito pieno di formule che non dovete imparare a memoria (in caso userete queste formule per l'esercizio scritto).

Per l'orale, l'importante sono le considerazioni finali e sul perchè abbiamo tratto quelle conclusioni.

9.1 Probabilità di errore per formati antipodali

¹

Consideriamo la costellazione di un formato anti-podale, quindi come nella 2-PAM, 2-ASK e 2-PSK:

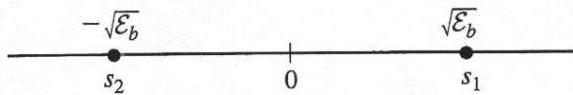


Figure 7.50 Signal points for antipodal signals.

Siccome in questi formati la dimensione dello spazio è $N = 1$, non ci sarà la notazione vettoriale.

La possiamo utilizzare, perchè alla fine è un vettore con $N = 1$.

Per semplicità di notazione, per adesso, lasciamo tutto senza notazione vettoriale

Ipotizzando che è stato trasmesso il simbolo s_1 , allora possiamo dire che il segnale ricevuto è uguale a:

$$\begin{aligned} r &= s_1 + n \\ &= \sqrt{E_b} + n \end{aligned} \tag{9.1}$$

dove:

- s_1 è il segnale che dovremmo ricevere
- n è il contributo di rumore
- per segnali anti-podali, dai capitoli precedenti, sapevamo che vale $\sqrt{E_b}$
- E_b è l'energia del bit

¹Slide del prof — Prestazioni sistemi numerici — pag 1 - 4
Appunti di Damiano — pag 4
Slide — Prestazioni sistemi numerici — pag 1 - 4
Appunti — 2025-04-08 — pag 2 - 4

Grazie alle considerazioni svolte nel capitolo precedente, cioè che:

- ogni segnale trasmesso s_m ha una PDF gaussiana
- il segnale r dovrà essere confrontato con ogni simbolo della costellazione

allora, possiamo intanto graficare l'andamento delle PDF dei due simboli:

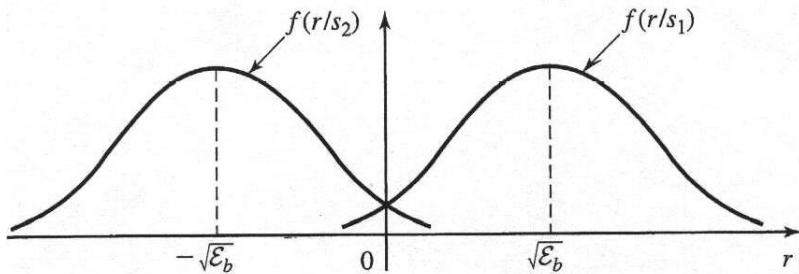


Figure 7.51 Conditional PDF's of two signals.

e scrivere le due formule che comparano r a simboli s_m della costellazione, che in questo caso sono due:

$$f(r|s_1) = \frac{1}{\sqrt{\pi \cdot N_0}} \cdot \exp \left[-\frac{(r - \sqrt{E_b})^2}{N_0} \right] \quad (9.2)$$

$$f(r|s_2) = \frac{1}{\sqrt{\pi \cdot N_0}} \cdot \exp \left[-\frac{(r + \sqrt{E_b})^2}{N_0} \right] \quad (9.3)$$

Calcoliamo adesso la probabilità di errore, sapendo di aver trasmesso s_1 piuttosto che s_2 , cioè ci ritroviamo r minore di 0, quindi nell'altra sezione di s_2 .

Cioè in formule:

$$\Pr\{e|s_1\} = \int_{-\infty}^0 f(r|s_1) dr \quad (9.4)$$

Alche impostiamo l'integrale di $f(r|s_1)$ nell'intervallo $[-\infty, 0]$

Facendo un po' di passi algebrici (che puoi tranquillamente ignorare perché sono conti):

$$\begin{aligned}
\Pr\{e|s_1\} &= \int_{-\infty}^0 f(r|s_1) dr \\
&= \frac{1}{\sqrt{\pi \cdot N_0}} \cdot \int_{-\infty}^0 \exp\left[-\frac{(r - \sqrt{E_b})^2}{N_0}\right] dr \\
&= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{-\sqrt{\frac{2E_b}{N_0}}} \exp\left(-\frac{x^2}{2}\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \cdot \int_{\sqrt{\frac{2E_b}{N_0}}}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx \\
&= Q\left(\sqrt{\frac{2 \cdot E_b}{N_0}}\right) \\
&= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)
\end{aligned} \tag{9.5}$$

Molto utile per lo scritto sono le tabelle erfc: le puoi visualizzare qui:

https://www.univpm.it/Entra/Engine/RAServeFile.php/f/P001926/allegati_ins/Erfc.pdf

L'operatore erfc lo avevamo già studiato nel corso di Teoria dei segnali quando, appunto, stavamo parlando della variabile gaussiana.

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.9.4 - Gaussiana - pag 127 - 129

Lo stesso risultato varrà anche per s_2 :

$$\Pr\{e|s_2\} = \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{9.6}$$

perchè le PDF sono uguali.

Ipotizzando che i simboli sono equiprobabili, possiamo scrivere la probabilità di errore sul bit (cioè il bit o può valere 1 o 0, quindi o può valere s_1 o può valere s_2) e facendo qualche semplice passo algebrico:

$$\begin{aligned}
P_b &= \frac{1}{2} \cdot \Pr(e|s_1) + \frac{1}{2} \cdot \Pr(e|s_2) \\
&= Q\left(\sqrt{\frac{2 \cdot E_b}{N_0}}\right) \\
&= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)
\end{aligned} \tag{9.7}$$

Possiamo sommare le probabilità di errore di s_1 e s_2 perchè le due probabilità sono incompatibili, quindi indipendenti.

Nella formula di P_b è presente il fattore $\frac{E_b}{N_0}$ e può essere visto come un rapporto segnale-rumore per bit.

Da ora in poi, consideriamo sempre questo rapporto cruciale per le varie trasmissioni perchè in digitale, rispetto all'analogico, ci interessa quanta energia E inviare dato una certa probabilità di errore.

L'argomento lo avevamo accennato nel primo capitolo di queste dispense, ma non ci erano noti a priori tutti gli argomenti e approfondimenti fatti fino ad adesso.

Questa espressione vale per tutti i formati binari anti-podali, quindi: 2-PAM, 2-ASK e 2-PSK.

Si considerano dei formati antipodali con canale di trasmissione di tipo BSC (in italiano Canale Binario Simmetrico), canale che può far variare il bit trasmesso in questa maniera:

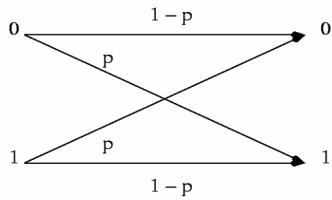


Figura 1.5.1 Schema di un canale binario simmetrico

cioè, se si trasmette uno 0, alla fine del canale, o può rimanere 0 o cambiare e diventare 1; stessa cosa vale se è stato trasmesso un 1.

Questa formula di P_b è la migliore probabilità di errore, i.e. il gold standard, per un formato binario perchè è quella più piccola.

Inoltre, si considera un ricevitore coerente (e che quindi si de-modula moltiplicando il segnale modulato con una funzione sinusoidale $\cos(2\pi f_c t)$ e filtrando nell'intorno di f_c) perchè i bit vengono inviati con la stessa energia.

All'aumentare di N, la probabilità della trasmissione peggiorerà.

Oppure visto in un'altra maniera:



<https://x.com/scottecs/status/1403686697938866177>

La formula:

$$P_b = \frac{1}{2} \cdot \operatorname{erfc} \left(\sqrt{\frac{E_b}{N_0}} \right) \quad (9.8)$$

ce la dobbiamo ricordare a memoria, proprio perchè è il gold standard.

Non faremo i conti "carta e penna" per i prossimi formati.

Ci dobbiamo fidare che siano veri, come i dogmi

9.2 Probabilità di errore per un generico formato binario con simboli equiprobabili

²

La probabilità di errore sul bit per un generico formato binario con simboli equiprobabili vale:

$$\begin{aligned} P_b &= Q\left(\sqrt{\frac{(d_{12})^2}{2 \cdot N_0}}\right) \\ &= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{(d_{12})^2}{4 \cdot N_0}}\right) \end{aligned} \quad (9.9)$$

dove con d_{12} si indica la distanza tra due punti.

Considerando come esempio i segnali ortogonali:

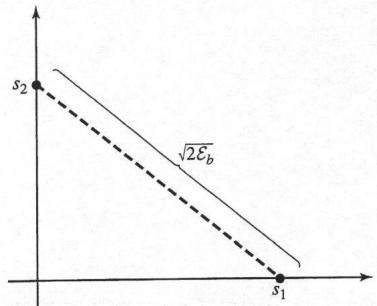


Figure 7.52 Signal points for orthogonal signals.

Allora, la formula precedente diventa:

$$\begin{aligned} P_b &= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{(d_{12})^2}{4 \cdot N_0}}\right) \\ &= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{(\sqrt{2 \cdot E_b})^2}{4 \cdot N_0}}\right) \\ &= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{2 \cdot E_b}{4 \cdot N_0}}\right) \\ &= \frac{1}{2} \cdot \operatorname{erfc}\left(\sqrt{\frac{E_b}{2 \cdot N_0}}\right) \end{aligned} \quad (9.10)$$

²Slide del prof — Prestazioni sistemi numerici — pag 5 - 6
Appunti di Damiano — pag 5 - 6
Slide — Prestazioni sistemi numerici — pag 5 - 6
Appunti — 2025-04-08 — pag 4 - 5

Ricordiamo ancora una volta che la qualità, in una trasmissione numerica digitale, non si definisce con un rapporto segnale-rumore classico, bensì tra definendo la probabilità di errore media su una finestra di misura molto ampia

A parità di errore, il formato ortogonale richiede una energia doppia rispetto a quella del formato antipodale.

In dB, dovremmo aumentare la potenza di 3 dB nel formato ortogonale per avere la stessa probabilità di errore del formato antipodale.

Questa affermazione si può dimostrare perchè se moltiplichiamo per 2 il termine sotto la radice quadra dell'erfc, quindi:

$$2 \cdot \frac{E_b}{2 \cdot N_0} = \frac{E_b}{N_0} \quad (9.11)$$

troviamo lo stesso argomento dell'erfc, quindi la stessa probabilità di errore P_b del formato antipodale.

Generalmente, piuttosto che scrivere la formula di P_b e metterla in relazione con il rapporto $\frac{E_b}{N_0}$ si prediligono le curve come la seguente (con note):

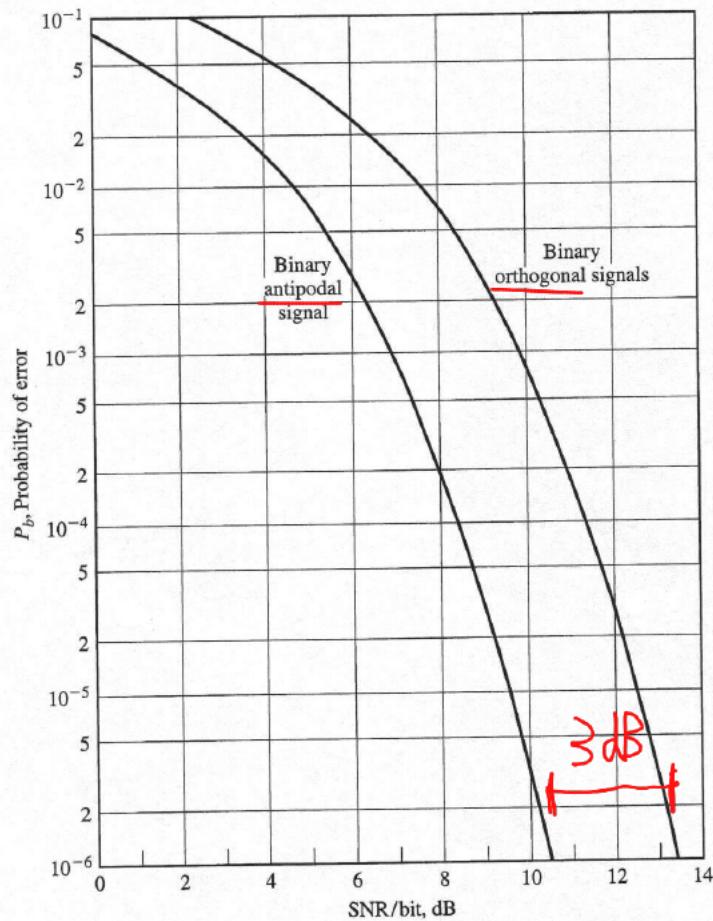


Figure 7.53 Probability of error for binary signals.

Come sottolineato precedentemente, la differenza tra le curve dei due formati è proprio di 3 dB.

Generalmente queste curve vengono utilizzate per fissare la probabilità di errore P_b (quindi fissando il valore delle ordinata), sapendo quale valore di $\frac{E_b}{N_0}$ utilizzare (quindi scegliendo di conseguenza il valore delle ascissa: in questa figura questo rapporto è indicato in dB).

9.3 Probabilità di errore per M-PAM

3

Consideriamo la M-PAM, dove, come i casi precedenti, c'è un'unica funzione di base, quindi $N = 1$, ma rappresenteremo più segnali, quindi M sarà maggiore di 2.

Dal punto di vista grafico, possiamo visualizzare i simboli della M-PAM in una ascissa:

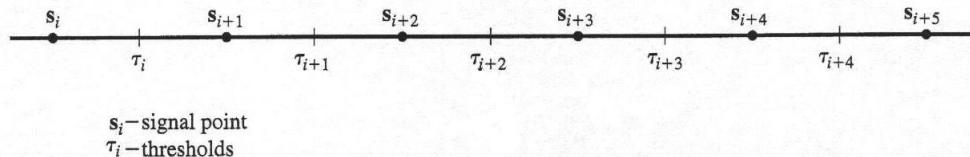


Figure 7.54 Placement of thresholds at midpoints of successive amplitude levels.

Come i casi precedenti, anche qui le PDF per ogni simbolo è di tipo gaussiano.

Se consideriamo le PDF di ogni simboli della M-PAM, potremmo vedere la figura precedente in questa maniera:

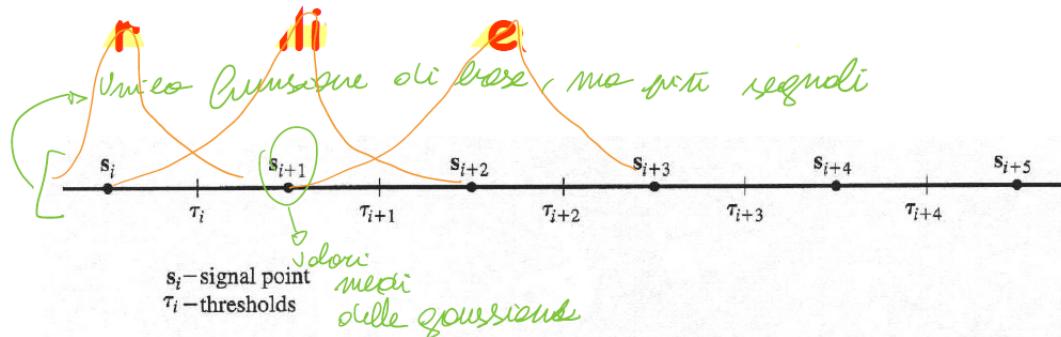


Figure 7.54 Placement of thresholds at midpoints of successive amplitude levels.

Possiamo calcolare l'energia media E_{ave} di tutti i segnali della M-PAM come:

$$E_{ave} = \frac{1}{M} \cdot \sum_{m=1}^M E_m \quad (9.12)$$

dove E_m è l'energia media di ogni simbolo di ogni gaussiana.

Facendo l'ipotesi che nella M-PAM vengano utilizzati degli impulsi rettangoli di ampiezza $\pm 1, \pm 3, \dots$, la potenza media della costellazione diventa:

Siccome partiamo da segnali con $M = 1$ che vanno da $[-1, +1]$, in cui la distanza minima è 2, all'aumentare di M , si vuole mantenere la stessa distanza minima.

Per esempio, per $M = 2$, cioè per segnali che vanno da $[-3, +3]$, la distanza minima rimane 2.

Quindi, in modulo, per segnali M-PAM, si considerano impulsi con ampiezza dispari.

³Slide del prof — Prestazioni sistemi numerici — pag 7 - 9
Appunti di Damiano — pag 8 - 9
Slide — Prestazioni sistemi numerici — pag 7 - 9
Appunti — 2025-04-08 — pag 5 - 6
Appunti — 2025-07-21 Ricevimento — pag 8.2

$$E_{ave} = \frac{M^2 - 1}{3} \cdot T \quad (9.13)$$

dove:

- M sono i simboli della M-PAM
- T è il periodo del simbolo

La potenza media per bit $E_{b,ave}$ è:

$$E_{b,ave} = \frac{E_{ave}}{\log_2(M)} \quad (9.14)$$

La probabilità di errore sull'intero segnale vale P_M :

$$\begin{aligned} P_M &= \frac{2 \cdot (M - 1)}{M} \cdot Q\left(\frac{6 \cdot \log_2(M) \cdot E_{b,ave}}{(M^2 - 1) \cdot N_0}\right) \\ &= \frac{M - 1}{M} \cdot \operatorname{erfc}\left(\sqrt{\frac{3 \cdot \log_2(M)}{(M^2 - 1)} \cdot \rho_{b,ave}}\right) \end{aligned} \quad (9.15)$$

dove:

$$\rho_{p,ave} = \frac{E_{b,ave}}{N_0} \quad (9.16)$$

Per noi, $\rho_{p,ave}$ sarà il nostro SNR, rapporto segnale-rumore

Applicando il Mapping di Gray (vedi 7.4.5.1 Modulazioni con costellazioni di segnali passa-banda pag 223 - 228) quindi che il vettore adiacente varia solo, in codifica, di un bit, allora la probabilità di errore sul bit P_b vale:

$$P_b = \frac{P_M}{\log_2(M)} \quad (9.17)$$

Visualizzando graficamente la probabilità di errore del simbolo P_M rispetto all'SNR sul bit in dB, per un qualsiasi M, P_M vale (con note):

Come scritto nelle note della figura, all'aumentare di M, peggiora $\rho_{p,ave}$ (che per noi, in digitale è il rapporto segnale rumore).

Come scritto in precedenza, idealmente si vorrebbe aumentare M per stringere la banda, ma ciò comporta una maggiore probabilità di errore.

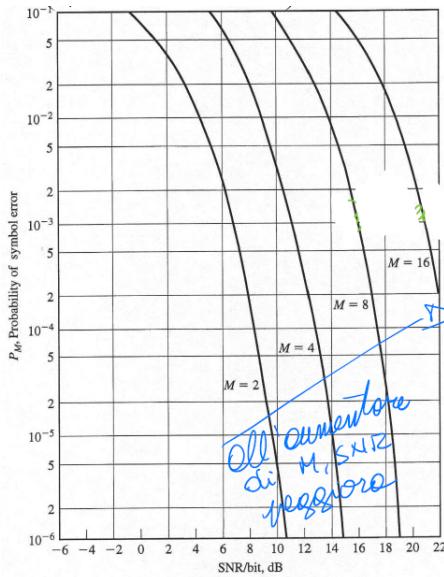


Figure 7.55 Probability of a symbol error for PAM.

Di seguito un confronto tra le varie M-PAM e di quanti dB aumentare per avere la stessa probabilità di errore:

$$M = 2 \text{ vs. } M = 4 : \frac{3\log_2 2}{(4-1)} \rho_b^{(2)} = \frac{3\log_2 4}{(16-1)} \rho_b^{(4)} \rightarrow \rho_b^{(4)} = 2.5 \rho_b^{(2)} \rightarrow +4 \text{ dB}$$

$$M = 4 \text{ vs. } M = 8 : \frac{3\log_2 4}{(16-1)} \rho_b^{(4)} = \frac{3\log_2 8}{(64-1)} \rho_b^{(8)} \rightarrow \rho_b^{(8)} = 2.8 \rho_b^{(4)} \rightarrow +4.47 \text{ dB}$$

$$M = 8 \text{ vs. } M = 16 : \frac{3\log_2 8}{(64-1)} \rho_b^{(8)} = \frac{3\log_2 16}{(256-1)} \rho_b^{(16)} \rightarrow \rho_b^{(16)} = 3.04 \rho_b^{(8)} \rightarrow +4.82 \text{ dB}$$

...

$$M = 2^k \text{ vs. } M = 2^{k+1} : \frac{3\log_2 2^k}{(2^{2k}-1)} \rho_b^{(2^k)} = \frac{3\log_2 2^{k+1}}{(2^{2k+2}-1)} \rho_b^{(2^{k+1})} \rightarrow \rho_b^{(2^{k+1})} = \frac{2^{2k+2}-1}{2^{2k}-1} \frac{k}{k+1} \rho_b^{(2^k)}$$

$$\text{per } k \rightarrow \infty : \rho_b^{(2^{k+1})} = 4 \rho_b^{(2^k)} \rightarrow +6 \text{ dB}$$

Siccome consideriamo il dato peggiore, per M che tende ad infinito, possiamo considerare che aumentando una M-PAM tra una M che è una potenza di 2^k e la successiva 2^{k+1} , si considera un aumento di +6 dB per mantenere lo stesso SNR, cioè bisogna aumentare in trasmissione di 4 volte l'energia.

9.4 Probabilità di errore per M-ASK

⁴

Le prestazioni del formato M-ASK (probabilità di errore sul simbolo e probabilità di errore sul bit) sono identiche a quelle del formato M-PAM.

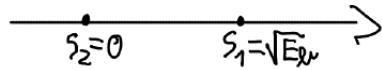
Questa frase è ovvia pensando che la M-ASK è una M-PAM traslata in frequenza.

⁴Slide del prof — Prestazioni sistemi numerici — pag 10
Slide — Prestazioni sistemi numerici — pag 10
Appunti — 2025-04-08 — pag 6

9.5 Probabilità di errore per OOK

5

Visualizzando nel piano vettoriale i simboli dell'OOK:

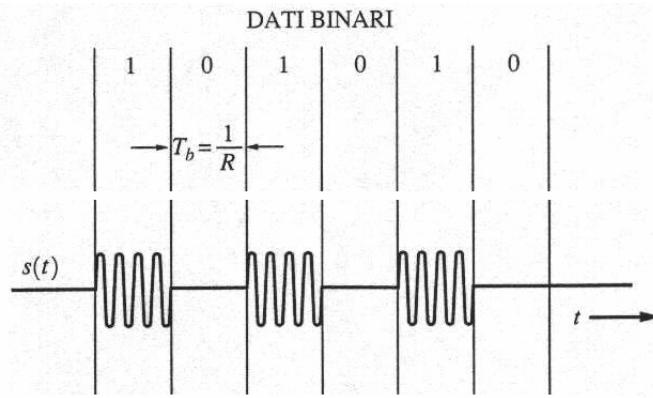


Considerando che la distanza d_{12} tra due punti nella OOK vale:

$$d_{12} = \sqrt{E_b} \quad (9.18)$$

dove con E_b si considera l'energia del livello alto, cioè l'energia di piccola (perchè nella OOK si modula il segnale quando c'è il livello alto, invece non si manda nulla se il livello logico è basso).

Per completezza, riporto l'andamento temporale di un segnale OOK (On-Off Keying):



La probabilità del bit con energia di picco P_b vale:

$$\begin{aligned} P_b &= \frac{1}{2} \cdot \operatorname{erfc} \left(\sqrt{\frac{(d_{12})^2}{4 \cdot N_0}} \right) \\ &= \frac{1}{2} \cdot \operatorname{erfc} \left(\sqrt{\frac{E_b}{4 \cdot N_0}} \right) \end{aligned} \quad (9.19)$$

A probabilità di errore, il formato OOK richiede una energia di picco 4 volte maggiore rispetto al "gold standard" della 2-ASK (in dB moltiplicare per 4 è aggiungere +6 dB).

Invece, se non si considera l'energia di picco, ma la sua media, quindi:

$$E_{b,ave} = \frac{E_b}{2} \quad (9.20)$$

e considerando $E_{b,ave}$ l'energia media per bit, allora si può calcolare la probabilità di errore per bit P_b :

$$P_b = \frac{1}{2} \cdot \operatorname{erfc} \left(\sqrt{\frac{E_{b,ave}}{2 \cdot N_0}} \right) \quad (9.21)$$

⁵Slide del prof — Prestazioni sistemi numerici — pag 11

Appunti di Damiano — pag 11

Slide — Prestazioni sistemi numerici — pag 11

Appunti — 2025-04-08 — pag 6 - 7

Come prima, possiamo fare la seguente considerazione.

A parità di probabilità di errore, il formato OOK richiede una energia media doppia di quella del formato 2-ASK (cioè +3 dB).

Generalmente il rumore è più difficile gestirlo e/o diminuirlo (basti pensare, come esempio, un amplificatore rumoroso, un canale rumoroso, etc.), quindi, per semplicità, si aumenta l'energia in trasmissione, in modo da avere e mantenere un SNR costante.

9.6 Probabilità di errore per M-PSK

6

La probabilità di errore sull'intero segnale in M-PSK vale circa:

$$\begin{aligned} P_M &\approx 2 \cdot Q \left(\sqrt{2 \cdot \log_2(M) \cdot \sin^2\left(\frac{\pi}{M}\right) \cdot \rho_b} \right) \\ &= \operatorname{erfc} \left(\sqrt{\log_2(M) \cdot \sin^2\left(\frac{\pi}{M}\right) \cdot \rho_b} \right) \end{aligned} \quad (9.22)$$

Si considera circa perchè bisognerebbe svolgere un integrale doppio, quindi si predilisce un'approssimazione di P_M

dove:

$$\rho_b = \frac{E_b}{N_0} \quad (9.23)$$

Per quanto riguarda l'energia sul bit E_b della M-PSK, esso vale:

$$E_b = \frac{E_m}{\log_2(M)} \quad (9.24)$$

dove E_m è l'energia media del segnale nel simbolo.

Inoltre, la probabilità di errore sul bit P_b vale:

$$P_b = \frac{P_M}{\log_2(M)} \quad (9.25)$$

Considerando il caso particolare della 2-PSK, quindi $M = 2$, P_b vale:

$$P_b = \frac{1}{2} \cdot \operatorname{erfc}(\sqrt{\rho_b}) \quad (9.26)$$

⁶Slide del prof — Prestazioni sistemi numerici — pag 12

Slide — Prestazioni sistemi numerici — pag 12

Appunti — 2025-04-08 — pag 7

9.7 Probabilità di errore per DPSK

7

Consideriamo per adesso una rivelazione coerente per una DPSK.

L'obbiettivo è quello di eliminare il problema dell'ambiguità di fase attraverso la codifica differenziale.

La demodulazione è identica a quella della PSK, quindi l'errore sul bit P_b vale:

$$P_b = \operatorname{erfc}(\sqrt{\rho_b}) \cdot \left(1 - \frac{1}{2} \cdot \operatorname{erfc}(\sqrt{\rho_b})\right) \approx 2 \cdot (P_b)_{PSK} \quad (9.27)$$

Il fattore 2 è dovuto al fatto che, un errore nella DPSK produce di solito 2 bit errati in uscita.

Invece, se si svolge una rivelazione non coerente, P_b vale:

$$P_b = \frac{1}{2} \cdot \exp(-\rho_b) \quad (9.28)$$

Facendo i grafici di P_b rispetto all'SNR sul bit e tracciando le curve della DPSK con rivelazione non coerente e la PSK, abbiamo il seguente andamento:

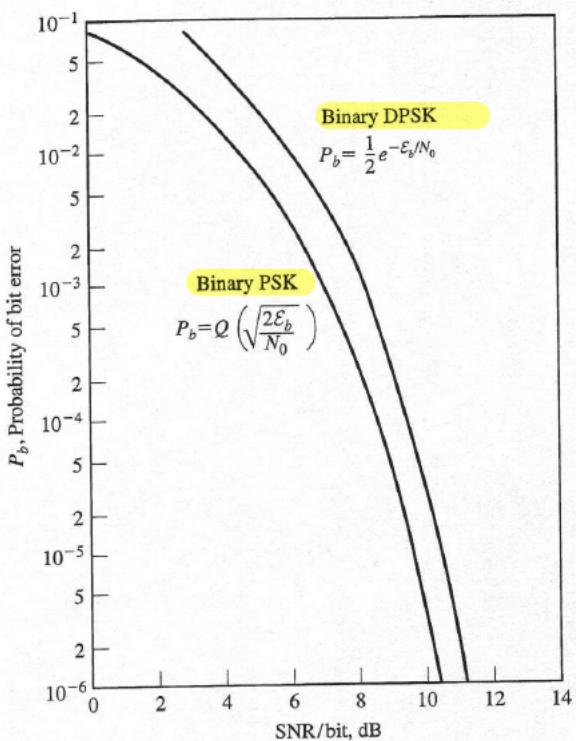


Figure 7.58 Probability of error for binary PSK and DPSK.

⁷Slide del prof — Prestazioni sistemi numerici — pag 13
Slide — Prestazioni sistemi numerici — pag 13
Appunti — 2025-04-08 — pag 7
Appunti — 2025-07-21 Ricevimento — pag 9

Quindi utilizzare la DPSK con rivelazione non coerente rispetto alla PSK si ha solo una penalizzazione di 1 dB.

Nella pagina 15 della slide ci si riferisce a, sinistra, una modulazione classica, invece a destra, ad una modulazione mista.

Tutte e due le modulazione hanno la stessa distanza minima.

Invece a pag 16, si considerano delle modulazioni con la stessa distanza minima, ma hanno energia diversa per ogni simbolo

9.8 Probabilità di errore per M-QAM

8

La probabilità di errore P_M sull'intero segnale in M-QAM, vale:

$$\begin{aligned} P_M &\approx 4 \cdot Q \left(\sqrt{\frac{3 \cdot \log_2(M)}{M - 1} \cdot \rho_{b,ave}} \right) \\ &= 2 \cdot \operatorname{erfc} \left(\sqrt{\frac{3 \cdot \log_2(M)}{2 \cdot (M - 1)} \cdot \rho_{b,ave}} \right) \end{aligned} \quad (9.29)$$

Dalla formula di P_M si può notare che, all'aumentare di M , P_M peggiora di un fattore molto alto.

Di seguito l'andamento della probabilità di errore P_M rispetto all'SNR per bit:

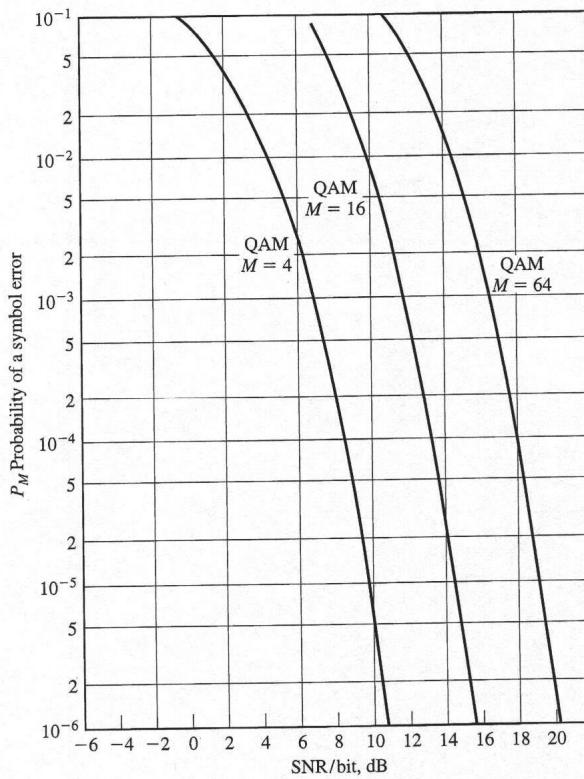


Figure 7.62 Probability of a symbol error for QAM.

Si considera che nella M-QAM si è applicato un mapping di Gray, quindi con k pari (in cui si avrà un quadrato perfetto della costellazione), invece se k è dispari, si dice che si applica un mapping quasi Gray.

Se $M = 4$, avremo che la 4-QAM è uguale alle 4-PSK.

⁸Slide del prof — Prestazioni sistemi numerici — pag 17
Slide — Prestazioni sistemi numerici — pag 17
Appunti — 2025-04-14 — pag 2

9.9 QAM vs. PSK

9

Confrontiamo le due probabilità di errore della M-QAM e della M-PSK:

$$(P_M)_{QAM} \approx 2 \cdot \operatorname{erfc} \left(\sqrt{\frac{3 \cdot \log_2(M)}{2 \cdot (M - 1)} \cdot \rho_{b,ave}} \right) \quad (9.30)$$

$$(P_M)_{PSK} \approx \operatorname{erfc} \left(\sqrt{\log_2(M) \cdot \sin^2\left(\frac{\pi}{M}\right) \cdot \rho_b} \right) \quad (9.31)$$

A livello energetico, la PSK ha il vincolo che tutti i segnali M devono avere la stessa energia, a differenza della QAM che l'energia degli M segnali può essere diversa in base alla circonferenza in cui si trova il simbolo.

QAM e PSK possono avere le stesse prestazioni, ma, all'aumentare di M, si predilisce utilizzare la M-QAM.

Di seguito il vantaggio analitico di usare una M-QAM rispetto a una M-PSK:

TABLE 7.1 SNR ADVANTAGE OF M-ARY QAM OVER M-ARY PSK

M	$P_{(dB)} = 10 \log_{10} \mathcal{R}_M$
8	1.65
16	4.20
32	7.02
64	9.95

M
aumento
↓

Risparmio
di energia
↓

Invece la seguente tabella:

TABLE 7.2 QAM SIGNAL CONSTELLATIONS

Number of signal points M	Increase in average power (dB) relative to M = 2
4	3
8	6.7
16	10.0
32	13.2
64	16.2
128	19.2

⁹Slide del prof — Prestazioni sistemi numerici — pag 18 - 20
Slide — Prestazioni sistemi numerici — pag 18 - 20
Appunti — 2025-04-14 — pag 3 - 4

dimostra che, per mantenere lo stesso SNR, bisogna aumentare la potenza del segnale in trasmissione.

Quindi, siccome vogliamo aumentare M per stringere la banda, dobbiamo per forza aumentare la potenza in trasmissione per mantenere lo stesso SNR.

9.10 Probabilità di errore formati ortogonali (ricevitore a correlatore)

10

Considerando dei segnali equiprobabili, possiamo calcolare la probabilità di errore sul segnale P_M utilizzando la seguente formula:

$$P_M = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{-\infty}^{+\infty} \left\{ 1 - [1 - Q(x)]^{M-1} \right\} \cdot e^{-\frac{(x-\sqrt{2 \cdot \log_2(M) \cdot \rho_b})^2}{2}} dx \quad (9.32)$$

Invece la probabilità di errore sul bit P_b vale:

$$P_b = \frac{2^{k-1}}{2^k - 1} \cdot P_M \quad (9.33)$$

Con il simbolo k si intende $M = 2^k$

Se la P_b fosse uguale a 1, sarebbe il caso favorevole perchè basterebbe solo cambiare il bit: il problema si pone quando $P_b = \frac{1}{2}$ perchè la probabilità sta in mezzo.

Come sempre, grafichiamo l'andamento della probabilità dell'errore sul bit P_b rispetto all'SNR sul bit:

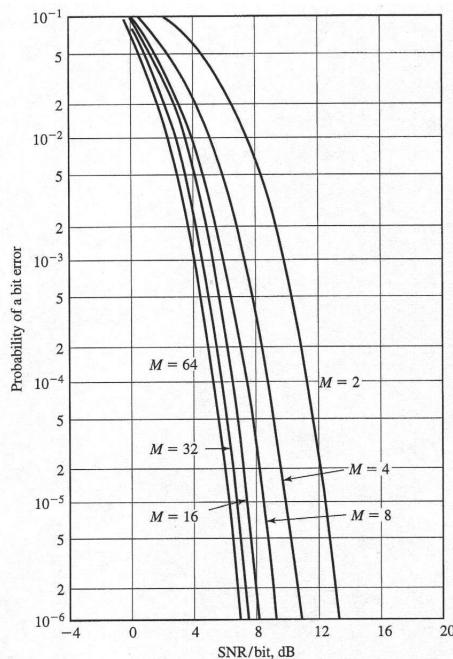


Figure 7.63 Probability of bit error for coherent detection of orthogonal signals.

Confrontando la figura e la formula di P_b , in particolare quella di P_M , notiamo che, all'aumentare di M , l'esponenziale diminuisce, è un decadimento molto veloce che fa tendere il fattore moltiplicativo a 1.

All'aumentare di M , l'SNR migliora per formati ortogonali con ricevitore a correlatore.

¹⁰Slide del prof — Prestazioni sistemi numerici — pag 21
 Slide — Prestazioni sistemi numerici — pag 21
 Appunti — 2025-04-14 — pag 4 - 5
 Appunti — 2025-07-21 Ricevimento — pag 10

9.10.1 Limite di Shannon

¹¹

Il matematico Shannon ha dimostrare che, se M tende ad infinito:

$$P_M = 0 \text{ per } \frac{E_b}{N_0} > -1.6 \text{ dB} \quad (9.34)$$

Questa formula prende il nome di Limite di Shannon.

I formati ortogonali tendono al limite di Shannon per M che tende ad infinito.

Dal punto di vista grafico:

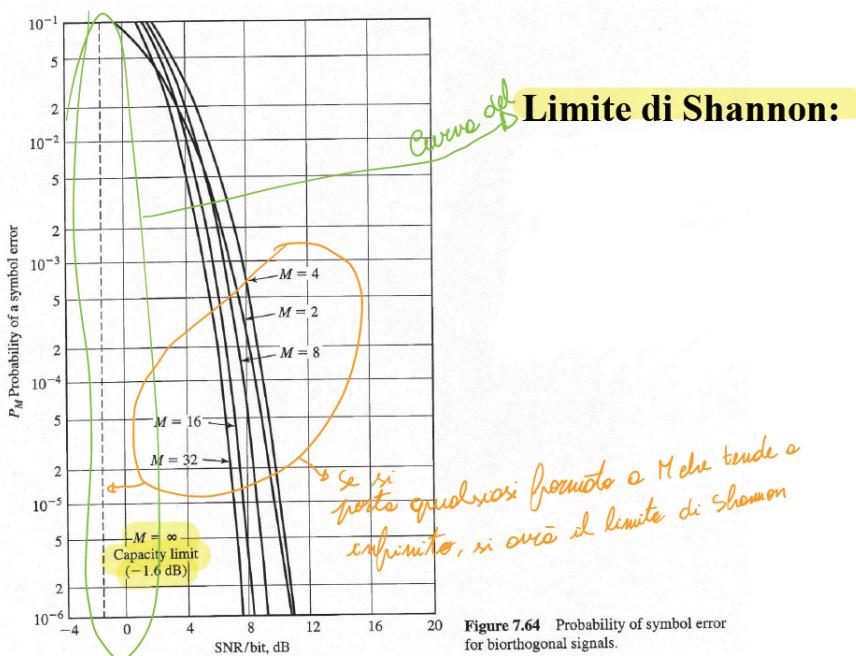


Figure 7.64 Probability of symbol error for biorthogonal signals.

¹¹Slide del prof — Prestazioni sistemi numerici — pag 22
Slide — Prestazioni sistemi numerici — pag 22
Appunti — 2025-04-14 — pag 5

9.11 Probabilità di errore formati bi-ortogonali e trans-ortogonali

12

A livello qualitativo, le probabilità di errore dei formati bi-ortogonali e trans-ortogonali sono le stesse delle ortogonali.

Di seguito le caratteristiche dei seguenti formati:

Biortogonali

$$P_M = 1 - \int_{-\sqrt{2E_m/N_0}}^{+\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-(v+\sqrt{2E_m/N_0})}^{(v+\sqrt{2E_m/N_0})} e^{-x^2/2} dx \right]^{\frac{M}{2}-1} e^{-v^2/2} dv \quad E_m = \log_2 M \cdot E_b$$

Transortogonali

$$P_b(\rho_b, \gamma_{mm}) = P_b[\rho_b(1 - \gamma_{mm}), 0] = P_b\left(\frac{M}{M-1}, 0\right) \quad \gamma_{mm} = -\frac{1}{M-1}$$

esprimere i segnali transortogonali utilizzando i biortogonali

Per M che tende ad infinito, anche la banda del segnale tende ad infinito.

¹²Slide del prof — Prestazioni sistemi numerici — pag 23
Slide — Prestazioni sistemi numerici — pag 23
Appunti — 2025-04-14 — pag 5 - 6

9.12 Probabilità di errore M-FSK ortogonali (ricevitore non coerente)

13

Di seguito i dettagli della M-FSK:

$$P_M = \sum_{n=1}^{M-1} (-1)^{n+1} \binom{M-1}{n} \frac{1}{n+1} \exp\left(-\frac{n}{n+1} \log_2 M \cdot \rho_b\right)$$

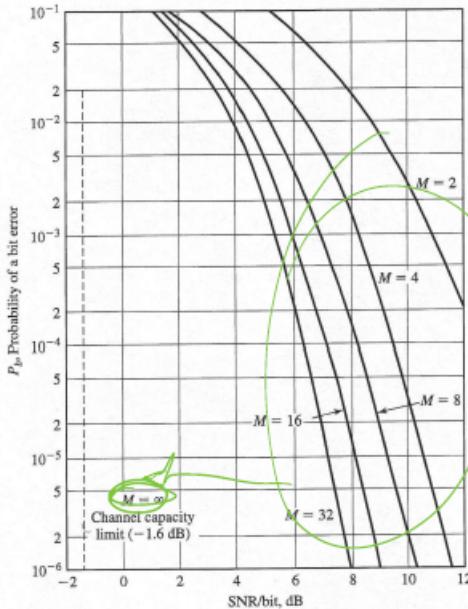


Figure 7.65 Probability of a bit error for noncoherent detection of orthogonal FSK signals

Ponendo:

M = 2:

Oltre

$$P_b = \frac{1}{2} \exp\left(-\frac{\rho_b}{2}\right)$$

✓ La 2-FSK perde 3 dB rispetto alla DPSK

Curve peggiori
rispetto pag 22

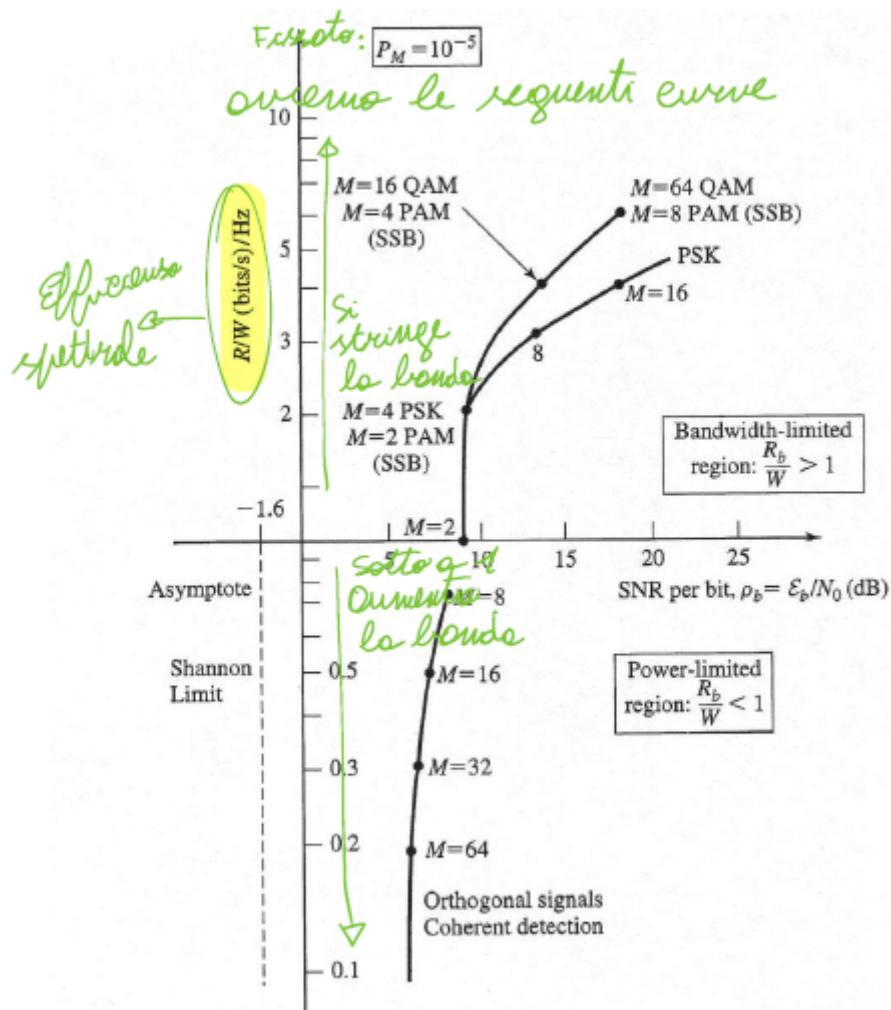
$$P_{b,1} = \frac{1}{2} e^{-\beta_b}$$

¹³Slide del prof — Prestazioni sistemi numerici — pag 24
Slide — Prestazioni sistemi numerici — pag 24
Appunti — 2025-04-14 — pag 6

9.13 Confronto tra le modulazioni fissato una probabilità di errore sul simbolo

14

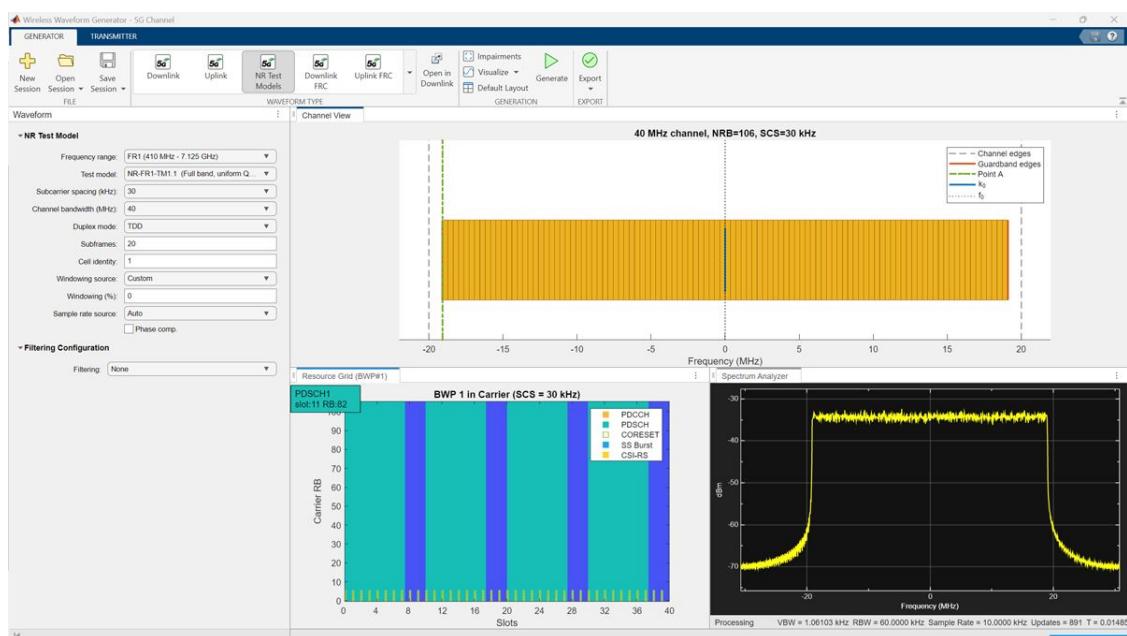
Fissato un P_M , con la seguente figura possiamo confrontare tutte le modulazioni studiate:



¹⁴Slide del prof — Prestazioni sistemi numerici — pag 25
Slide — Prestazioni sistemi numerici — pag 25
Appunti — 2025-04-14 — pag 6

Capitolo 10

Codifica di canale



10.1 Cosa è una codifica di canale

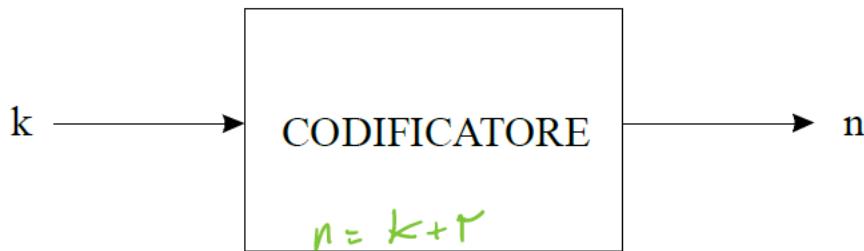
1

La codifica di canale è l'operazione mediante cui una trasmissione, in forma numerica, può essere protetta dagli effetti dei disturbi (soprattutto il rumore termico) introdotti dal canale e dagli apparati che elaborano il segnale.

Più precisamente, attraverso una opportuna strutturazione della sequenza dei simboli trasmessi, è possibile, in reazione, rivelare la presenza di errori o, meglio ancora, correggere fino a un numero prefissato, dipendente dalle caratteristiche del codice.

Si possono implementare diverse caratteristiche.

Faremo riferimento, per questo corso, il seguente schema più classico, definito "codifica a blocco":



In questo corso, consideriamo solo gli errori dovuti al canale e non tratteremo su come è realizzato il codificatore e decodificatore.

Il codificatore opera su blocchi di k simboli in ingresso, fornendone in uscita n , in cui:

$$n > k \quad (10.1)$$

Questo significa che, l'operazione di codifica introduce una certa ridondanza, cioè aggiungere r simboli, in particolare:

$$r = n - k \quad (10.2)$$

che non portano informazione.

Benché l'operazione di codifica possa essere applicata tanto a canali binari quanto a canale M-ari, di seguito ci riferiremo esclusivamente al caso binario.

La possibilità di rivelare o correggere errori in una sequenza di simboli è legata alla capacità di individuare, in ricezione, sequenze che non potevano essere trasmesse.

Se la sorgente emette simboli binari (cioè bit) equiprobabili ed incorrelati, un blocco di k simboli può assumere qualunque delle 2^k sequenze possibili.

L'aggiunta degli r simboli di ridondanza, al contrario, consente di differenziare maggiormente le sequenze trasmissibili (alfabeto di trasmissione).

Avendo a disposizione n simboli, si possono infatti "sintetizzare" 2^n sequenze, ma essendo necessarie solo 2^k sequenze (cioè la parola di codice) la ridondanza potrà essere utilizzata proprio per differenziare maggiormente ciò che può essere trasmesso, riducendo il rischio di equivocazione.

¹Slide del prof — Codifica di canale — pag 1
Slide — Codifica di canale — pag 1
Appunti — 2025-04-14 — pag 7 - 8

10.1.1 Codice a bit di parità

²

Consideriamo l'esempio di "bit di parità": a partire da una sequenza di k di bit (con k arbitrariamente scelto) il codificatore aggiunge un altro simbolo binario, in modo che il numero totale di simboli "1" nella sequenza trasmessa sia pari.

Così facendo, solo metà delle 2^n sequenze vengono riconosciute come "ammissibili" in ricezione, dove:

$$2^n = 2^{k+1} \quad (10.3)$$

Se la sequenza ricevuta contiene un numero dispari di "1", il decodificatore riconosce che si è verificato almeno un errore.

Dunque, il sistema è in grado di ricevere la presenza di 1 errore senza peraltro essere in grado di correggerlo.

Ma, se il numero di errori è pari, il numero totale di simboli "1" torna ad essere accettabile.

Quindi, il decodificatore è inefficiente.

Visto in un'altra maniera, più elettronica e che viene implementata negli FPGA e negli LDPC, è che il codificatore del bit di parità è uno XOR di n ingressi, sia codificatore che al decodificatore.

Se vuoi farti una cultura in merito:

FPGA

<https://www.ibm.com/it-it/think/topics/field-programmable-gate-arrays>

LDPC

https://en.wikipedia.org/wiki/Low-density_parity-check_code

XOR

<https://www.elemania.altervista.org/digitale/circuiti/circ2.html>

Siccome con questo codice non si può rivelare l'errore, cioè capire dove c'è stato l'errore, o si può dire che è presente un errore o il ricevitore chiede al trasmettitore di inviare di nuovo la sequenza.

Oppure, si implementa un altro codice, come ad esempio il codice a ripetizione.

²Slide del prof — Codifica di canale — pag 1 - 2
 Slide — Codifica di canale — pag 1 - 2
 Appunti — 2025-04-14 — pag 8

10.1.2 Codice a ripetizione

3

Nel codice a ripetizione il valore k è prefissato e poniamo come esempio:

$$k = 1 \quad (10.4)$$

Nel codice a ripetizione, il codice replica n volte il simbolo emesso dalla sorgente.

Considerando un n uguale a:

$$n = 5 \quad (10.5)$$

il simbolo 0 generato dalla sorgente diventa:

$$0 \rightarrow 00000 \quad (10.6)$$

Continuando l'esempio, si supponga che, a seguito di errori di trasmissione, la sequenza generata dal codice a ripetizioni diventa:

$$00000 \rightarrow 10010 \quad (10.7)$$

Il decodificatore riconoscerà che la sequenza ricevuta non potrà essere stata trasmessa.

Dovendo prendere una decisione sul simbolo inviato, il decodificatore riterrà più plausibile che sia stato trasmesso il simbolo 0 piuttosto che il simbolo 1.

Con questo criterio, il decodificatore sceglie "a maggioranza".

In questo esempio citato, il decodificatore sarà in grado di correggere fino a 2 errori.

Per una sequenza lunga n , è più probabile che avvenga un errore in un bit che in più bit.

In generale, per un codice a ripetizione di lunghezza n , il numero di errori correggibili è pari a:

$$t = \left[\frac{n-1}{2} \right] \quad (10.8)$$

dove con le parentesi quadre di $\frac{n-1}{2}$ si indica la parte intera di quel rapporto.

Ai fini delle capacità di correzione, non si ha alcun vantaggio ad assumere un n pari.

Portando un esempio con $n = 6$, si possono correggere 2 errori, che è lo stesso numero di errori che si possono correggere con $n = 5$.

Quindi, in generale, meglio utilizzare un numero n più basso possibile per correggere gli errori.

³Slide del prof — Codifica di canale — pag 2
Slide — Codifica di canale — pag 2
Appunti — 2025-04-14 — pag 8 - 9

10.2 Caratteristiche di una codifica di canale

⁴

Gli esempi descritti del codice a ripetizione e del codice a bit di parità, mettono in evidenza due aspetti fondamentali e di validità generale:

- indipendentemente dal modo con cui viene implementata la decodifica, la stima in ricezione della sequenza trasmessa è generalmente basata su un criterio di massima verosimiglianza, chiamato anche criterio della minima distanza.
Il decodificatore sceglie la parola di codice \vec{v} più vicina dalla sequenza ricevuta \vec{r}
- la capacità di correzione è tanto maggiore quanto più le parole di codice sono diverse tra loro

In questa sequenza, codice e codifica verranno utilizzati come sinonimi.

Inoltre, per bit si intende sequenza di zeri e uni, e non si considerano bit informativi (si specificherà quando lo sono).

L'abilità di chi progetta il codice sta nell'individuare, a parità di altre condizioni, la parola di codice che meglio soddisfa la seconda proprietà.

Vi è un modo esplicito per quantificare la diversità tra le due sequenze di n simboli, che consiste nel misurare la relativa "distanza".

Praticamente stiamo ritornando al concetto fatto sugli M vettori con N vettori di base fatto nelle varie modulazioni numeriche.

Soltanto che, in questo caso, viene applicato questo concetto ad una sequenza di bit e non a dei vettori N -dimensionali.

⁴Slide del prof — Codifica di canale — pag 2
 Slide — Codifica di canale — pag 2
 Appunti — 2025-04-14 — pag 9
 Appunti — 2025-04-15 — pag 2 - 4

10.2.1 Distanza di Hamming

5

Esistono vari tipi di definizioni di distanza: per adesso ci soffermiamo sulla definizione di distanza come distanza di Hamming.

La distanza di Hamming d_{ij} tra due parole di codice \vec{v}_i e \vec{v}_j rappresenta il numero di posizioni per cui le due parole differiscono.

Ad esempio:

$$\begin{cases} \vec{v}_i = 0001011 \\ \vec{v}_j = 0010110 \end{cases} \rightarrow d_{ji} = 4 \quad (10.9)$$

Le due parole di un codice a ripetizione di lunghezza n hanno distanza di Hamming n; e così via.

Dato un codice, si può definire la distanza minima, o semplicemente distanza di Hamming, tra due o più sequenze.

Per un codice generico, si dimostra che il numero di errori correggibili t è legato a d_{min} dalle seguenti espressioni:

$$d_{min} = \begin{cases} 2 \cdot +1 \text{ se } d_{min} \text{ è dispari} \\ 2 \cdot +2 \text{ se } d_{min} \text{ è pari} \end{cases} \quad (10.10)$$

Come si vede dalla formula di d_{min} , non compare il numero n.

Questo significa che, all'aumentare di n, non è detto che si può correggere le sequenze.

Un esempio concreto è quello che accade ogni giorno nell'ambito delle telecomunicazioni wireless: la banda è stretta, quindi è sempre meglio inviare meno n bit possibili e quindi una d_{min} minima per risolvere gli errori.

Il codice a bit di parità risulta un caso particolare di d_{min} perchè la capacità di rivelazione degli errori è data da $d_{min} - 1$.

⁵Slide del prof — Codifica di canale — pag 2 -3
Slide — Codifica di canale — pag 2 - 3
Appunti — 2025-04-15 — pag 4 - 6

10.2.2 Sottospazio di una decodifica

⁶

Quando si genera un codice, dal punto di vista matematico, si genera un sottospazio rispetto a tutte le possibili configurazioni del bit.

I codici hanno diverse proprietà, di seguito ne elenchiamo due.

Se il codice è lineare, qualunque coppia di codice è presente solo la d_{min} .

Per maggiori informazioni riguardo ai codici lineari:

https://en.wikipedia.org/wiki/Linear_code

Se il codice è sistematico, possono essere progettati con:

$$n = d_{min} \quad (10.11)$$

fissa.

Ad esempio, il codice di Hamming può essere implementato con $d_{min} = 3$.

Se d_{min} è bassa, può essere valutata e non stimata.

Per le codifiche "vecchie" si poteva valutare, quindi quantificare la distanza minima tra un simbolo rispetto all'altro perché i simboli erano pochi.

Con le codifiche attuali, ciò non è possibile farlo perché i simboli sono tanti e si utilizzano delle tecniche di codifica molto complicate: in questo caso, si dice che si può stimare la distanza minima, e non quantificarla come nelle codifiche "vecchie"

Il codice opera quando il numero di errori non eccede le capacità correttive.

Cosa succede quando il numero di errori è maggiore di t ?

Più frequentemente, nel tentativo di correggere la sequenza ricevuta, il decodificatore potrebbe aggiungere altri errori, con ciò "peggiorando" la situazione.

Consideriamo il seguente esempio:

A rigore i vari sottospazi dovrebbero essere di dimensione n , ma nell'esempio in figura $n = 1$, quindi è un serie di punti lungo un segmento.

Le parole di codice sono contrassegnate dai cerchietti e separate da sequenze (contrassegnate con rettangoli) che possono essere ricevute, ma non corrispondono a parole di codice.

In figura, le due parole di codice \vec{v}_i e \vec{v}_j sono a distanza 5.

Ciascuna circonferenza, centrata su \vec{v}_i e \vec{v}_j ha raggio t , in cui t sono il numero di errori correggibili.

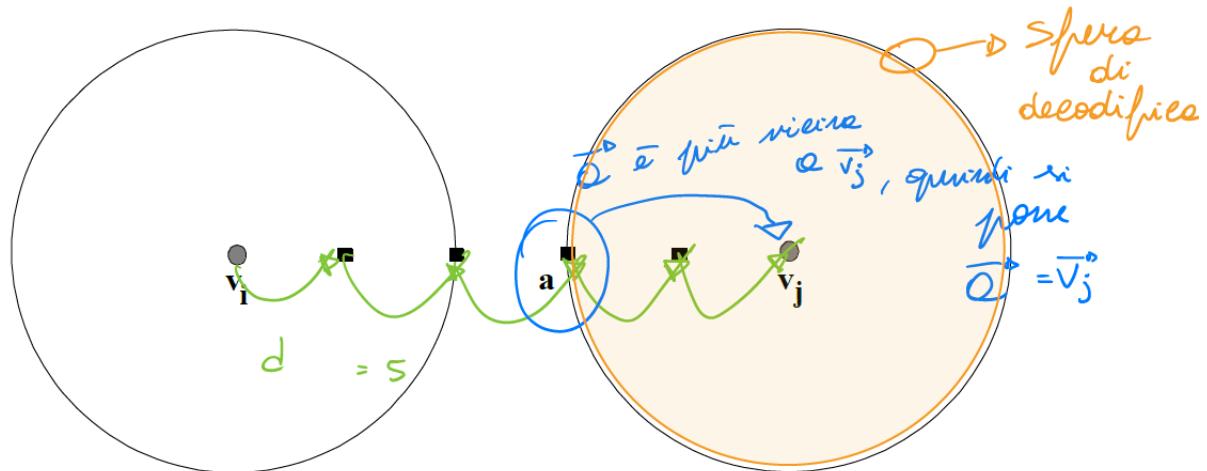
Se trasmessa la parola \vec{v}_i e si riceve la sequenza \vec{a} (in ricezione 3 simboli sono errati), il decodificatore stimerà, grazie al principio della vero-somiglianza (o anche chiamato della minima distanza), che la parola di codice più plausibile non è quella trasmessa che sia \vec{v}_i , bensì stimerà che la parola di codice trasmessa sia \vec{v}_j .

⁶Slide del prof — Codifica di canale — pag 3

Slide — Codifica di canale — pag 3

Appunti — 2025-04-15 — pag 4 - 6

Appunti — 2025-07-21 Ricevimento — pag 11 - 12.1



Ciò introduce 2 errori, pari alla distanza tra \vec{d} e \vec{v}_j .

Per l'esempio preso in figura:

$$t = 2 \quad (10.12)$$

cioè, al numero massimo di errori certamente correggibili.

Nel caso più sfortunato, il decodificatore introduce, nel tentativo di correggere una situazione che eccede le sue capacità correttive, un numero uguale a t di errori addizionali.

Questo perchè, ritornando all'esempio, se io mando \vec{v}_i , ma ricevo $\vec{d} = \vec{v}_j$, in quel caso il codice non può risolvere l'errore.

10.3 Probabilità di un decodificatore di correggere un errore

7

Prima di implementare la codifica di canale, bisogna valutare l'efficienza del codice.

Tenendo conto anche delle osservazioni, si può concludere che la probabilità di decodificare correttamente una parola di codice soddisfa la seguente relazione:

$$P_{rc} \geq \sum_{i=0}^t \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \quad (10.13)$$

dove p è la probabilità di transizione errata del singolo bit binario (da 0 a 1 o viceversa, in accordo con la definizione di canale binario simmetrico o BSC in inglese).

Ricordiamo al volo come è fatto un canale BSC (che già avevamo introdotto nei capitoli precedenti).

Il canale BSC causa questo cambiamento:

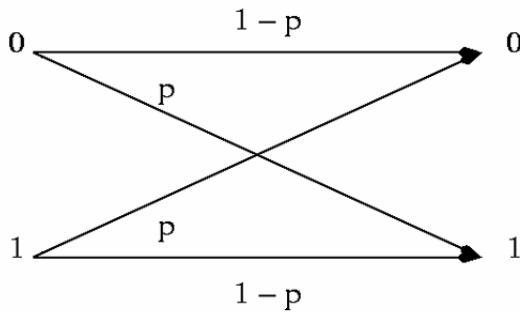


Figura 1.5.1 Schema di un canale binario simmetrico

cioè, se si trasmette uno 0, alla fine del canale, o può rimanere 0 o cambiare e diventare 1; stessa cosa vale se è stato trasmesso un 1.

La formula che esprime P_{rc} è l'applicazione dell'esperimento di Bernoulli.

Ti consiglio di ripassarlo dagli appunti del precedente corso.

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 12.6 - Esperimento di Bernoulli - pag 120

L'applicazione di Bernoulli lo possiamo applicare perché alla fine il bit se è 0 o 1, e la sua probabilità, è uguale a calcolare l'esperimento del lancio della moneta (o è testa o è croce).

Possiamo visualizzarlo anche graficamente con questi video:

- <https://youtu.be/6YzrVUV09M0?si=Qae7VPG994RL1N9L>
Overexplaining the binomial distribution by Primer
- <https://www.youtube.com/watch?v=8idr1WZ1A7Q>
Binomial distributions — Probabilities of probabilities, part 1 by 3Blue1Brown
- <https://youtu.be/J8jNoF-K8E8?si=ieyBa7voY9DxjTLN>
The Binomial Distribution and Test, Clearly Explained!!! by StatQuest with Josh Starmer

⁷Slide del prof — Codifica di canale — pag 3 - 5
Slide — Codifica di canale — pag 3 - 5
Appunti — 2025-04-15 — pag 6 - 7
Appunti — 2025-07-21 Ricevimento — pag 12.2

Analizzando la formula di P_{rc} si giustifica osservando che sarà certamente corretta la sequenza che non è affetta da errori, e la probabilità, quindi, sarà uguale a $(1 - p)^n$, ma, saranno ricostruite correttamente anche tutte le sequenze che contengono fino a t errori.

Per ciascun valore di i in cui $1 \leq i \leq t$, sono esattamente $\binom{n}{i}$ le sequenze contenenti gli errori.

Il fatto di considerare una disuguaglianza è conseguenza del fatto che il decodificatore, sporadicamente, potrebbe correggere un errore maggiore di t .

La formula di P_{rc} può essere utilizzata se non si abbia dipendenza statistica tra la ricezione di un simbolo e quella degli altri.

Ciò accade perchè si considerano canali in cui i simboli sono statisticamente indipendenti

Questa ipotesi è certamente vera in alcuni casi, come ad esempio nei canali con rumore additivo gaussiano bianco (o AWGN), dove gli errori sono dispersi in modo del tutto casuale nella sequenza ricevuta.

Facendo il complemento a 1 di P_{rc} , sappiamo la probabilità di errore del codificatore.

Indicando il complemento a 1 di P_{rc} come P_{Eg} , questo ultimo vale:

$$\begin{aligned} P_{Eg} &\leq 1 - P_{rc} \\ &= 1 - \sum_{i=0}^t \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \\ &= \sum_{i=0}^n \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \end{aligned} \tag{10.14}$$

Siccome si considera la probabilità di errore su n , P_{Eg} è la probabilità di errore globale sulla sequenza ricevuta.

A noi interessa di più la probabilità di errore del codificatore sul singolo bit, che chiameremo P_{Eb} .

Considerando un calcolo pessimistico, possiamo scrivere che:

$$\begin{aligned} P_{Eb} &\approx \frac{1}{2} \cdot P_{Eg} \\ &= \frac{1}{2} \cdot \sum_{i=0}^n \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \end{aligned} \tag{10.15}$$

La formula di P_{Eb} si giustifica sulla base che, scegliendo "a caso", o si prende la sequenza giusta o il suo complementare.

Siccome ci troviamo nel caso dei bit, o il bit è corretto o il bit è sbagliato.

Considerando invece una P_{Eb} più accurata:

$$P_{Eb} \approx \sum_{i=0}^n \frac{i+t}{n} \cdot \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \tag{10.16}$$

dove p è diverso dal codice che viene implementato.

Questa formula è giustificata dal fatto che il canale introduce i errori, invece il decodificatore introduce t errori, che sono già presenti.

Tutto questo va valutato per gli n simboli, quindi $\frac{i+t}{n}$.

Ma, generalmente, per il calcolo di P_{Eb} si considera la formula più pessimistica.

10.4 Espansione della banda con un codice

8

Adesso consideriamo brevemente la valutazione delle prestazioni conseguibili con un codice.

L'aggiunta del codice, per come ne abbiamo discusso adesso, sembra portare solo benefici perché permettono di correggere gli errori.

Ma, l'aggiunta di simboli di ridondanza (in numero di r) vengono aggiunti a quelli significativi (in numero di k) e non possono modificare la quantità di informazione trasmessa per unità di tempo.

Ma, non possiamo aggiungere infiniti simboli di ridondanza r perché il segnale deve essere trasmesso entro un tempo pre-fissato.

Siccome un segnale numerico è ottenuto a partire da un segnale analogico, la frequenza di cifra F_c non può essere minore di $2B$, cioè:

$$F_c \geq 2 \cdot B \quad (10.17)$$

dove:

- B è la massima frequenza significativa del segnale campionato
- F_c è la frequenza di cifra senza la codifica, quindi il segnale senza i simboli di ridondanza r

Aggiungendo i k simboli introdotti dalla codifica, si considera un nuovo segnale in cui ha frequenza F_s pari a:

$$F_s = \frac{n}{k} \cdot F_c \quad (10.18)$$

A lezione abbiamo visto invece F_s , ma sotto l'aspetto energetico.

Se consideriamo E_b l'energia del singolo bit, ed E_s l'energia del segnale, allora possiamo esprimere E_s l'energia del segnale con la codifica come:

$$\begin{aligned} n \cdot E_s &= k \cdot E_b \\ &\downarrow \\ E_s &= \frac{k}{n} \cdot E_b \end{aligned} \quad (10.19)$$

dove:

- n sono i bit del segnale
- k sono i bit della parola di codice con l'aggiunta della codifica

Quindi, nel tempo riservato alla trasmissione dei k bit, devono essere contenuti gli n simboli.

All'aumentare di F_c frequenza di simbolo, aumenta anche F_s .

In assenza di codifica, si può calcolare la potenza di rumore N come:

$$N = K \cdot T \cdot \frac{F_c}{2} \quad (10.20)$$

⁸Slide del prof — Codifica di canale — pag 5 - 6
Slide — Codifica di canale — pag 5 - 6
Appunti — 2025-04-15 — pag 7 - 8

dove:

- K è la costanza di Boltzmann
- T è la temperatura in kelvin
- F_c è la frequenza del segnale senza codifica

Invece, in presenza di rumore, la potenza di rumore N' diventa:

$$\begin{aligned} N' &= K \cdot T \cdot \frac{F_s}{2} \\ &= K \cdot T \cdot \left(\frac{n}{k} \cdot \frac{F_c}{2} \right) \end{aligned} \quad (10.21)$$

Ovviamente:

$$\begin{aligned} F_s &> F_c \\ \downarrow \\ N' &> N \end{aligned} \quad (10.22)$$

Il fattore $\frac{n}{k}$ è un parametro caratteristico e prende il nome di espansione di banda dovuto al codice.

Questo ci dice che, se dobbiamo aggiungere un codice ad un segnale, abbiamo bisogno di più banda da quella iniziale.

L'inverso di $\frac{n}{k}$, cioè $\frac{k}{n}$, prende il nome del "rate" (in inglese tasso) del codice.

A parità di capacità di correzione, sono evidentemente preferibili i codici che presentano il massimo rate perché hanno la minima espansione di banda.

Però, ciò, significa avere un codice nettamente più complesso da elaborare.

10.5 Effetti del codice sull'SNR

9

A parità di S potenza trasmessa dal trasmettitore, il rapporto segnale-rumore in presenza di codifica peggiora.

Il rapporto segnale rumore $\frac{S}{N}$, con l'aggiunta della codifica, diventa $\frac{S}{N'}$:

$$\frac{S}{N'} = \frac{k}{n} \cdot \frac{S}{N} \quad (10.23)$$

dove $\frac{k}{n}$ è il rate del codice.

Visto che la probabilità di transizione errata P_{Eb} dipende dal rapporto segnale-rumore, in cui P_{Eb} aumenta al diminuire del rapporto (perchè ci sarà più rumore e quindi si avrà più probabilità di sbagliare i bit), in presenza di codifica:

$$p' > p \quad (10.24)$$

anche perchè, dato un set di n bit da inviare, il simbolo codificato contiene i bit del codice di errore.

Un altro modo di vederla è quando si è in macchina.

Se la macchina ha 5 posti e siete in 3, se ti trovi nei posti dietro, stai più comodo perchè hai più spazio e quindi puoi riempire lo spazio in più con le borse della spesa (che sono i bit del codice).

Se la macchina ha 5 posti e siete in 5, se ti trovi nei posti dietro, ti trovi più scomodo e non hai nessun posto dove mettere le buste della spesa.

Quindi, la probabilità di correzione del bit, con l'introduzione della codifica, diventa, nel caso pessimistico:

$$\begin{aligned} P_{Eb} &\approx \sum_{i=0}^n \frac{i+t}{n} \cdot \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \\ &\downarrow \\ P_{Eb} &\approx \sum_{i=0}^n \frac{i+t}{n} \cdot \binom{n}{i} \cdot p'^i \cdot (1-p')^{n-i} \end{aligned} \quad (10.25)$$

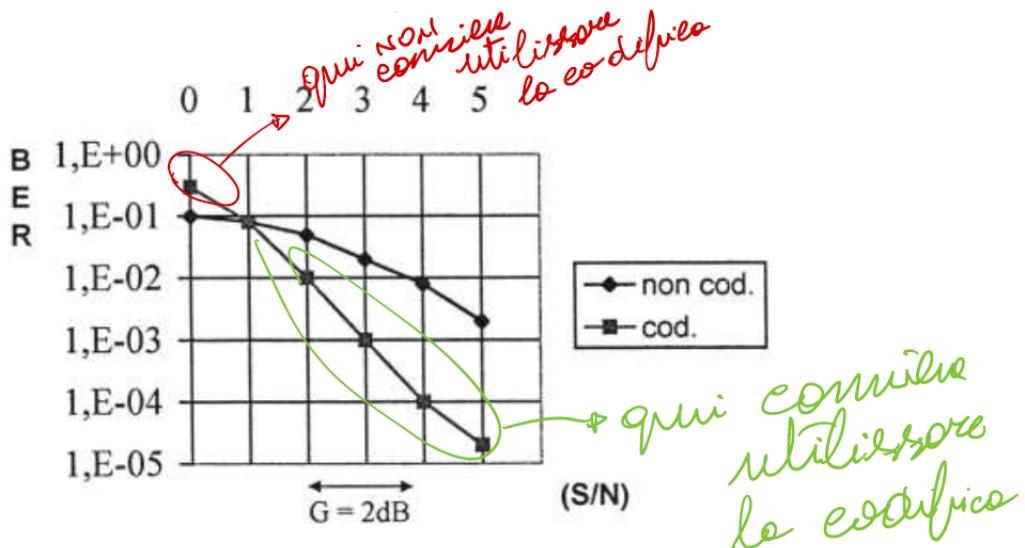
Se invece consideriamo il caso peggiorativo:

$$\begin{aligned} P_{Eb} &\approx \frac{1}{2} \cdot \sum_{i=0}^n \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \\ &\downarrow \\ P_{Eb} &\approx \frac{1}{2} \cdot \sum_{i=0}^n \binom{n}{i} \cdot p'^i \cdot (1-p')^{n-i} \end{aligned} \quad (10.26)$$

⁹Slide del prof — Codifica di canale — pag 6 - 8
Appunti di Damiano — pag 6 - 8
Slide — Codifica di canale — pag 6 - 8
Appunti — 2025-04-15 — pag 8 - 9

Dove l'applicazione del codice risulta vantaggiosa, definiamo "guadagno di codifica" la riduzione del rapporto $\frac{S}{N}$, appunto conseguente alla codifica, a parità di BER (Bit Error Rate).

Un esempio di andamento del BER in funzione dell'SNR:



La scelta del codice va ottimizzata per la specifica applicazione.

In pratica, fissata la probabilità di errore di interessa, occorre cercare un codice il cui punto di incrocio con la curva non codificata sia al di sopra della regione di interesse, e che quindi garantisca un guadagno di SNR.

10.6 Codici BCH

10

Accanto ai codici semplici menzionati, esiste una famiglia di blocchi chiamati BCH (Bose-Chaudhuri-Hocquenghem, cioè i nomi dei ricercatori che per primi li proposero).

Grazie a questi codici possiamo scegliere dalla tabella del codice i possibili valori di n, k e t, come in questa tabella:

*codici sopra
la capacità di
quanto può
correggere
il codice*

n	k	t
7	4	1
15	11	1
	7	2
	5	3
31	26	1
	21	2
	16	3
	11	5
	6	7
63	57	1
	51	2
	45	3
	39	4
	36	5
	30	6
	24	7
	18	10
	16	11
	10	13
	7	15
127	120	1
	113	2
	106	3
	99	4
	92	5
	85	6
	78	7
	71	9
	64	10
	57	11
	50	13

Se p' è sufficientemente piccolo, dell'ordine di 10^{-2} o meno, la P_{Eb} diventa, considerando quella peggiorativa:

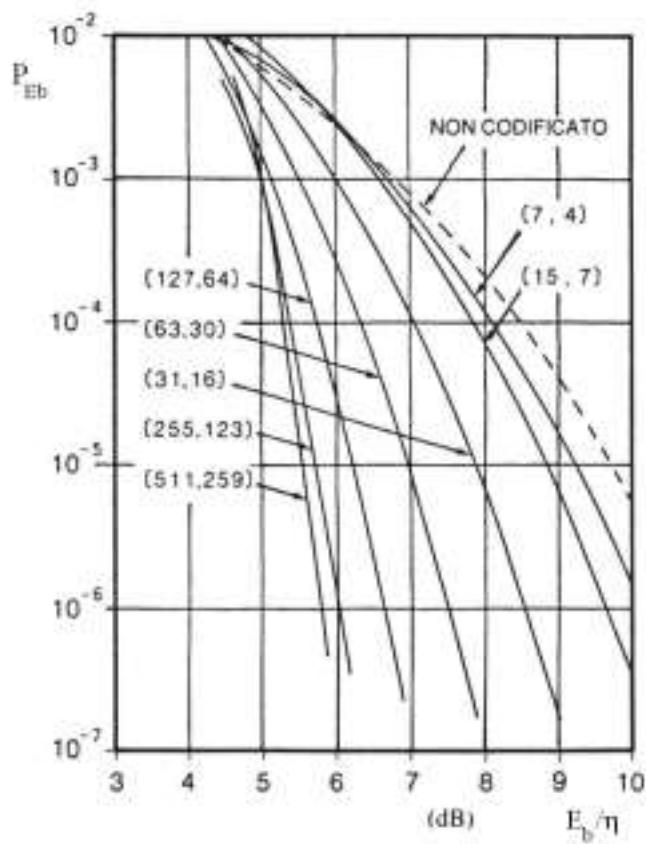
$$P_{Eb} \approx \frac{1}{2} \cdot \binom{n}{t+1} \cdot p^{t+1} \quad (10.27)$$

Considerando invece quella approssimata non peggiorativa:

$$P_{Eb} \approx \frac{2 \cdot t + 1}{n} \cdot \binom{n}{t+1} \cdot p^{t+1} \quad (10.28)$$

¹⁰Slide del prof — Codifica di canale — pag 8 - 9
 Appunti di Damiano — pag 8 - 9
 Slide — Codifica di canale — pag 8 - 9
 Appunti — 2025-04-15 — pag 9

Utilizzando questa ultima formula, si può tracciare il grafico come il seguente:



In particolare, questo grafico mostra l'andamento del BER di diversi codici BCH.

10.7 Codici MDS

¹¹

Considerando i codici in cui abbiamo una base M-aria, possono essere utili i codici Reed-Solomon (in breve RS).

I codici RS operano su un alfabeto di simboli M:

$$M = 2^m \quad (10.29)$$

in cui:

$$m > 1 \quad (10.30)$$

Il codice RS fa parte della famiglia dei codici MDS (Maximum Distance Separable).

I codici RS vengono impiegati nella PPM per la registrazione ottica dei CD.

¹¹Slide del prof — Codifica di canale — pag 9
 Appunti di Damiano — pag 9
 Slide — Codifica di canale — pag 9
 Appunti — 2025-04-15 — pag 9

Capitolo 11

Generazione di variabili aleatorie



11.1 Definizioni utili per le sequenze casuali

¹

Come scritto nel titolo della sezione, di seguito daremo delle definizioni per capire questo capitolo.

¹Slide del prof — Generazione di variabili aleatorie — pag 1
Appunti di Damiano — pag 1
Slide — Generazione di variabili aleatorie — pag 1
Appunti — 2025-04-15 — pag 10

11.1.1 Definizioni di sequenze

²

Sequenza binaria

Una sequenza \underline{u} composta da u_k elementi:

$$\underline{u} = (u_k) \quad (11.1)$$

dove il singolo elemento u_k appartiene ai valori dei numeri interi \mathbb{Z} e può valere solo 0 o 1:

$$u_k = \{0, 1\} \text{ per } u_k \in \mathbb{Z}_2 \quad (11.2)$$

Talvolta la sequenza \underline{u} può essere una sequenza binaria bipolare.

Sequenza binaria bipolare

Una sequenza \underline{u} composta da b_k elementi:

$$\underline{u} = (b_k) \quad (11.3)$$

dove il singolo elemento b_k può valere:

$$b_k = \{-1, +1\} \quad (11.4)$$

dove se:

$$\begin{cases} u_k = 0 \leftrightarrow b_k = -1 \\ u_k = 1 \leftrightarrow b_k = +1 \end{cases} \quad (11.5)$$

Sequenza binaria random

Sequenza \underline{u} dove ogni simbolo u_k è equiprobabile:

$$P(u_k = 0) = P(u_k = 1) = 0.5 \quad (11.6)$$

e indipendente:

$$P(u_k | u_i) = P(u_k) \text{ con } i \neq k \quad (11.7)$$

cioè u_k non dipende da u_i .

Nel seguito, salvo diversa indicazione, le sequenze si assumeranno di lunghezza infinita.

Se una sequenza ha lunghezza finita N , la interpreteremo come una sequenza di lunghezza infinita e periodo N , ottenuta mediante ripetizione.

In alte parole, all'interno di N i bit sono casuali, ma siccome si ripetono, la sequenza non è proprio random, cioè casuale.

²Slide del prof — Generazione di variabili aleatorie — pag 1
 Appunti di Damiano — pag 1
 Slide — Generazione di variabili aleatorie — pag 1
 Appunti — 2025-04-15 — pag 10

11.1.2 Definizioni di grandezze per le sequenze casuali

³

Definiamo ora le seguenti grandezze.

Numero di simboli

$$\begin{cases} d_0 = \text{numero di simboli } u_k = 0 \text{ presenti in un periodo} \\ d_1 = \text{numero di simboli } u_k = 1 \text{ presenti in un periodo} \end{cases} \quad (11.8)$$

Funzione di autocorrelazione

Per funzione di autocorrelazione si intende la seguente formula $R(i)$:

$$R(i) = \sum_{k=1}^N b_k \cdot b_{k+i} \text{ per } i = 0, 1, \dots, N-1 \quad (11.9)$$

Per $i = 0$, R vale:

$$R(0) = N \quad (11.10)$$

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 13.2 - Medie di insieme - pag 153

Per le variabili aleatorie estratte dal processo sono calcolabili le medie di insieme.

Consideriamo il valore medio in x_1 , che sarà dato da:

$$\langle x_1 \rangle = \int_{-\infty}^{+\infty} x_1 f_1(x_1; t_1) dx_1 \quad (11.11)$$

Il momento congiunto di ordine (1, 1) delle variabili x_1 e x_2 si otterrà come:

$$\langle x_1 x_2 \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 f_{12}(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (11.12)$$

Il momento congiunto di ordine (1, 1), che rappresenta la correlazione tra le variabili aleatorie estratte x_1 e x_2 ; prende il nome di autocorrelazione statistica del processo e si indica con $R(t_1, t_2)$.

■

Invece, qui, siccome siamo nell'ambito dei numeri discreti, non useremo l'integrale, bensì una sommatoria.

Corsa

Una corsa, in inglese run, C_L di lunghezza L è una sotto-sequenza di \underline{u} costituita da L simboli consecutivi uguali.

In particolare, indicheremo con C_L^0 una corsa di L simboli consecutivi uguali a 0.

³Slide del prof — Generazione di variabili aleatorie — pag 1
Appunti di Damiano — pag 1
Slide — Generazione di variabili aleatorie — pag 1
Appunti — 2025-04-15 — pag 10

Indicheremo con C_L^0 una corsa di L simboli consecutivi uguali a 1.

Nel caso di sequenza di lunghezza finita (e dunque trattata come sequenza periodica), per attribuire correttamente il bit u_0 ad una corsa, si deve guardare U_{N-1} , cioè U_{-1} .

Si indica con N_T il numero totale di corse nel periodo.

Si indica con N_L il numero di corse C_L nel periodo.

Si indica con N_L^0 il numeri corse C_L^0 nel periodo.

Si indica con N_L^1 il numeri corse C_L^1 nel periodo.

11.1.3 Proprietà delle sequenze random

⁴

Asintoticamente, cioè per N molto grande, le sequenze random tendono ad avere queste tre proprietà:

- $d_0 = d_1$ (cioè il numero di simboli 0 ed 1 tende ad essere uguale)
- $R(i) = 0$ per $i \neq 0$ (cioè il numero di coincidenze tende ad essere uguale al numero di discordanze, per ogni shift dato alla sequenza di partenza)
- $N_L = \frac{N_T}{2^L}$ (cioè metà delle corse sono di lunghezza 1, un quarto 2, etc.) e $N_L^1 = N_L^0$ (cioè le corse di una data lunghezza sono per metà corse di 1 e per metà corse di 0)

⁴Slide del prof — Generazione di variabili aleatorie — pag 2
 Appunti di Damiano — pag 2
 Slide — Generazione di variabili aleatorie — pag 2
 Appunti — 2025-04-15 — pag 10

11.2 Generazione di sequenze pseudo-random mediante LFSR

5

Siccome abbiamo a che fare con strumenti, come i microcontrollori, che sono degli strumenti progettati per svolgere dei compiti deterministici, purtroppo i dispositivi che abbiamo non riescono a generare delle sequenze veramente casuali.

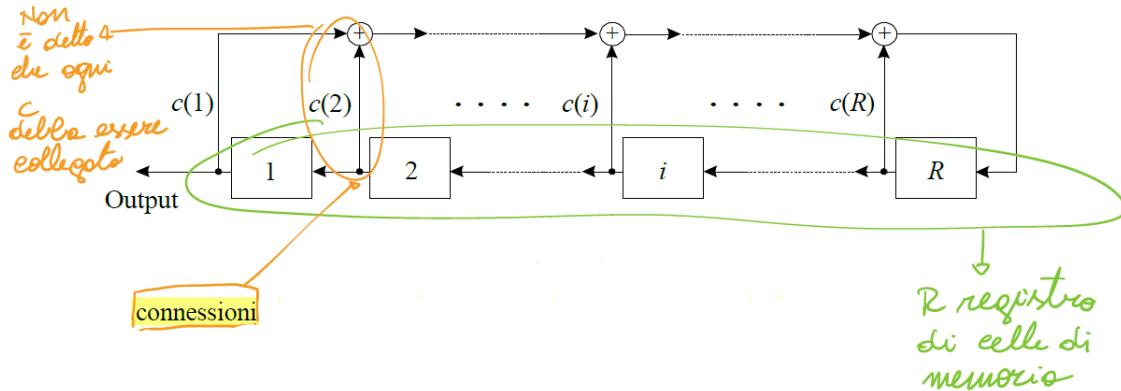
Quindi, in pratica, ci si accontenta di generare sequenze pseudo-random, cioè con proprietà il più possibile simili a quelle delle sequenze random.

La generazione di buone sequenze pseudo-random può avvenire mediante una struttura che prende il nome di Linear Feedback Shift Register (oppure in acronimo LFSR).

Perché, alla fine, il computer o qualsiasi microprocessore a basso livello, è formato da registri.

Un LFSR è un sistema a tempo discreto (generalmente temporizzato da un clock) costituito da R celle e R possibili connessioni.

Lo schema generale di un LFSR è il seguente:



Come scritto nelle note della figura, le connessioni sono numeri binari.

Ogni connessione i-esima si può scrivere come $c(i)$ e può essere:

$$c(i) = \{0, 1\} \text{ per } c(i) \in \mathbb{Z}_2 \quad (11.13)$$

Dal punto di vista circuitale:

- Se $c(i) = 1$, il collegamento è presente
- Se $c(i) = 0$, il collegamento è assente

Le connessioni $c(i)$ rimangono costanti nel tempo.

Il contenuto delle celle è binario e varia in corrispondenza di istanti discreti t_k :

$$t_k = (k \cdot T) \quad (11.14)$$

(identificabili, ad esempio con il fronte di salita di un clock con periodo T).

⁵Slide del prof — Generazione di variabili aleatorie — pag 2 - 3
 Appunti di Damiano — pag 2 - 3
 Slide — Generazione di variabili aleatorie — pag 2 - 3
 Appunti — 2025-04-15 — pag 10

Circuitalmente un temporizzatore lo possiamo implementare in diverse maniere.

Ti lascio un esempio sui timer 555:

https://youtu.be/oZzjmAbyyIQ?si=DRkYR_8KggpFJ2OU

How 555 timers Work - The Learning Circuit by element14presents

Inoltre, qui spiegherò di seguito come funziona un LFSR per generare una sequenza pseudo-random in forma scritta, ma grazie a questo video lo puoi visualizzare step-by-step anche graficamente e con un esempio in Python:

<https://www.youtube.com/watch?v=Ks1pw1X22y4>

Random Numbers with LFSR (Linear Feedback Shift Register) - Computerphile by Computerphile

Indicheremo il contenuto delle celle all'istante (discreto) k con i simboli:

$$s(1)_k, \dots, s(R)_k \text{ con } s(i)_k \in \mathbb{Z}_2 \quad (11.15)$$

All'istante $k = 0$, si assegna il valore iniziale alle R celle, ovvero si scelgono:

$$[s(1)_0, \dots, s(R)_0] \quad (11.16)$$

Ho messo le parentesi quadre come se la sequenza fosse un vettore nel linguaggio di programmazione C

Al generico istante k , il sistema si evolve nel seguente modo:

1. bit generato, si genera il bit, quindi:

$$u_k = s(1)_k \quad (11.17)$$

2. calcolo reazione, si calcola il bit di reazione:

$$x_k = \sum_{i=1}^R s(i)_k \cdot c(i) \quad (11.18)$$

Le somme ed i prodotti sono in \mathbb{Z}_2 (cioè possono essere solo 0 o 1): in altre parole si fanno le somme modulo 2 utilizzando degli XOR

3. calcolo stato futuro, avanzamento dello shift register e inserimento della reazione nell'ultima cella.

Si ha:

$$s(R)_{k+1} = x_k \quad \forall j : 1 \leq j \leq R-1, s(j)_{k+1} = s(j+1)_k \quad (11.19)$$

Per fare le somme modulo 2 si usano gli XOR.

Un po' di materiale che ti può aiutare:

<https://www.settorezero.com/wordpress/operazioni-algebriche-booleane-come-si-utilizzano-gli-operatori-di-logica/>

<https://www.youtube.com/watch?v=VPw9vPN-3ac>

XOR & the Half Adder - Computerphile by Computerphile

Si dimostra, cioè non lo dimostriamo, che, scegliendo opportunamente le connessioni, si ottiene una sequenza \underline{u} :

$$\underline{u} = (u_k) \quad (11.20)$$

con periodo massimo N:

$$N = 2^R - 1 \quad (11.21)$$

dove R sono il numero delle celle binarie.

Inoltre, si vuole ribadire che la scelta di avere un LFSR in cui l'evoluzione avviene da destra verso sinistra è puramente convenzionale.

11.2.1 Proprietà di una sequenza con N massimo da LFSR

⁶

Una sequenza \underline{u} di periodo massimo $N = 2^R - 1$, generata mediante LFSR, possiede le seguenti proprietà:

- $d_1 = d_0 + 1$
- $R(i) = -1$ per ogni $i \neq 0$
- A partire dal valore di N_T , per qualunque lunghezza L si ha:
 - Per L compreso tra 1 e R , N_L vale: $N_L = \frac{N_T}{2^L}$ e $N_L^1 = N_L^0$
 - Per $L = R$ N_L vale: $N_L = N_L^0 = 1$
 - Per $L = R$, N_L vale: $N_L = N_L^1 = 1$
 - Per $L > R$, N_L vale: $N_L = 0$

Confrontando le proprietà delle sequenze random con quelle random LFSR, notiamo che:

- $d_1 = d_0 + 1$, che è simile a $d_0 = d_1$ delle sequenze random
- $R(i) = -1$ per ogni $i \neq 0$, invece nelle sequenze random non è vera che la relazione deve valere per ogni valore diverso da zero
- Per L compreso tra 1 e R , N_L vale: $N_L = \frac{N_T}{2^L}$ e $N_L^1 = N_L^0$, che è la stessa proprietà della sequenza random

⁶Slide del prof — Generazione di variabili aleatorie — pag 3 - 4
 Appunti di Damiano — pag 3 - 4
 Slide — Generazione di variabili aleatorie — pag 3 - 4
 Appunti — 2025-04-15 — pag 10

11.2.2 Collegamenti in un LFSR per avere periodo massimo

7

Il contenuto iniziale delle R celle determina univocamente la sequenza \underline{u} generata di periodo N.

Partire da una diversa configurazione delle R celle, corrisponde a generare una versione shiftata, nel tempo, della stessa sequenza.

Le connessioni che corrispondono ad un periodo massimo sono note e tabulate.

Dato un R, si possono utilizzare la seguente tabella per ricavare un periodo massimo:

$R = 2$	(1,2)
$R = 3$	(1,3)
$R = 4$	(1,4)
$R = 5$	(1,4)
$R = 6$	(1,6)
$R = 7$	(1,7)
$R = 8$	(1,5,6,7)
$R = 9$	(1,6)
$R = 10$	(1,8)
$R = 11$	(1,10)
$R = 12$	(1,7,9,12)
$R = 13$	(1,10,11,13)
$R = 14$	(1,5,9,14)
$R = 15$	(1,15)
$R = 16$	(1,5,14,16)
$R = 17$	(1,15)
$R = 18$	(1,12)
$R = 19$	(1,15,18,19)
$R = 20$	(1,18)
$R = 21$	(1,20)
$R = 22$	(1,22)
$R = 23$	(1,19)
$R = 24$	(1,18,23,24)
$R = 25$	(1,23)
$R = 26$	(1,21,25,26)
$R = 27$	(1,23,26,27)
$R = 28$	(1,26)
$R = 29$	(1,28)
$R = 30$	(1,8,29,30)
$R = 31$	(1,29)
$R = 32$	(1,11,31,32)
$R = 33$	(1,21)
$R = 34$	(1,8,33,34)
$R = 61$	(1,2,3,6)

Va ribadito che le configurazioni sopra riportate non sono le uniche possibili e, a parità di valore di R, ne possono essere trovate altre equivalenti.

⁷Slide del prof — Generazione di variabili aleatorie — pag 4 - 7
Appunti di Damiano — pag 4 - 7
Slide — Generazione di variabili aleatorie — pag 4 - 7

11.3 Rappresentazione polinomiale di un LFSR

⁸

Un aspetto molto interessante riguarda le modalità di rappresentazione dell'LFSR.

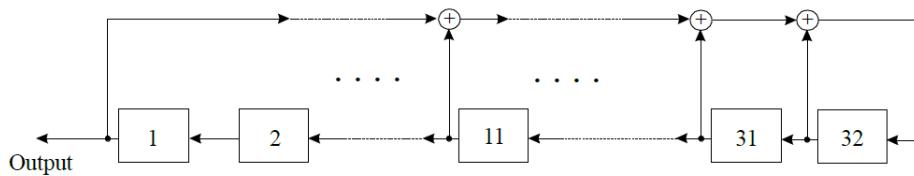
Se può rappresentare il LFSR con un polinomio $h(x)$ di grado R , in cui saranno presenti soltanto le potenze della variabile formale x , rappresentative delle connessioni non nulle.

In particolare, visto che la potenza di grado R sarà sempre presente (sta a rappresentare la retroazione che determina l'aggiornamento dell'ultima cella dell'LFSR), se:

$$c(i) = 1 \quad (11.22)$$

comparirà, nella rappresentazione polinomiale, la potenza di x^{i-1} .

A titolo di esempio, si passerà dall'LFSR con i suoi elementi in vista grafica:



a questa sua rappresentazione polinomiale:

$$h(x) = x^{32} + x^{31} + x^{30} + x^{10} + 1 \quad (11.23)$$

Nel polinomio, è presente il termine $+1$, proprio perchè, come si vede in figura, la cella dell'LFSR [1] ha un ramo di retroazione prima della scritta "Output".

Essendo ora un polinomio, quindi una rappresentazione matematica, si può scegliere la base del polinomio, molto utile se abbiamo a che fare con polinomi molto grandi.

A titolo di esempio, nell'esempio c'è scritto come convertire il polinomio $h(x)$ da binario a ottale, quindi un polinomio in base 8.

Ricordando che la rappresentazione è una funzione biunivoca, anche il polinomio ha una sequenza di periodo massimo:

$$N_{max} = 2^R - 1 \quad (11.24)$$

⁸Slide del prof — Generazione di variabili aleatorie — pag 8 - 9
Appunti di Damiano — pag 8 - 9

11.4 PSL di un LFSR

⁹

Può essere interessante osservare che, le proprietà delle sequenze di periodo massimo discusse in precedenza, vengono significativamente modificate nel caso in cui sia necessario troncare la sequenza.

Questa operazione va effettuato, ad esempio, nel caso in cui la sequenza generata debba adattarsi ad una struttura di frame pre-fissata.

Si pensi ad avere una lunga stringa e la memoria in un microcontrollore è piccolissima.

Dal punto di vista non tecnico, è come avere un trasloco e mettere tutto quello che hai della vecchia casa in un'auto: sarebbe impossibile.

Ma se smonti ogni singolo componente e fai diversi viaggi con l'auto, poi "traslocare" facilmente.

Emblematico, sotto questo aspetto, è il comportamento di un parametro fondamentale (che è molto utile nelle applicazioni che utilizzano sequenze pseudo-random) denominato Maximum Peak Slide Lobe, ed indicato con PSL.

Esso è definito come segue:

$$PSL = \max_{i \neq 0} |R(i)| \quad (11.25)$$

Nel caso di una sequenza di periodo massimo, con cioè:

$$N_{max} = 2^R - 1 \quad (11.26)$$

allora possiamo dire che la sequenza ha PLS:

$$PSL = 1 \quad (11.27)$$

⁹Slide del prof — Generazione di variabili aleatorie — pag 8 - 9
Appunti di Damiano — pag 8 - 9
Slide — Generazione di variabili aleatorie — pag 8 - 9

11.5 Generazione di variabili casuali

¹⁰

Un problema che spesso si pone nella simulazione di un sistema di comunicazione è quello della generazione di campioni secondo una prefissata statistica.

L'esempio più importante, in questo caso, è quello dato dal rumore termico, i cui campioni, come noto, seguono una statistica gaussiana.

Per generare una variabile gaussiana, è preliminare necessario disporre di un algoritmo capace di generare campioni indipendenti di una variabile casuale uniformemente distribuita tra 0 e 1.

¹⁰Slide del prof — Generazione di variabili aleatorie — pag 10
Appunti di Damiano — pag 10
Slide — Generazione di variabili aleatorie — pag 10
Appunti — 2025-05-06 — pag 2

11.5.1 Generazione di una variabile causale con distribuzione uniforme

¹¹

Un buon algoritmo per la generazione della variabile uniforme è il seguente:

1. si fissa un valore iniziale v_0 :

$$v_0 \neq 0 \quad (11.28)$$

2. per generare un k-esimo campione w_k si calcola prima v_k :

$$v_k = a \cdot v_{k-1} \bmod m \quad (11.29)$$

e poi w_k :

$$w_k = \frac{v_k}{m} \quad (11.30)$$

Ricordo al volo che l'operatore mod restituisce il modulo tra $a \cdot v_{k-1}$ e m , dove per modulo si può intendere anche il resto intero della divisione.

Per ricavare w_k si divide v_k per m , in modo tale che w_k venga normalizzato, cioè rimanga tra un valore intero.

Se i valori di m e a sono scelti opportunamente, la sequenza v_k ha periodo $(m-1)$, cioè assume $(m-1)$ valore diversi (tutti i numeri interi tra 1 e $(m-1)$ compresi) in ordine pseudo-random, prima di ritornare al valore v_0 .

La divisione per m , restituisce poi per un valore reale w_k compreso tra 0 e 1 (esclusi).

Il resto della divisione non deve essere mai 0, altrimenti $v_k = 0$ e rimarrà 0 da quel momento in poi.

¹¹Slide del prof — Generazione di variabili aleatorie — pag 10
 Appunti di Damiano — pag 10
 Slide — Generazione di variabili aleatorie — pag 10
 Appunti — 2025-05-06 — pag 2

11.5.2 Generazione di una variabile aleatoria secondo una assegnata statistica

¹²

Una volta che si dispone un generatore di variabile casuale uniforme, è possibile costruire semplici algoritmi per la generazione di una variabile casuale con distribuzione generica.

Indichiamo con V la variabile aleatoria uniformemente distribuita da 0 e 1.

La sua funzione di ripartizione (cioè la distribuzione di probabilità cumulativa), che esprime la probabilità che sia:

$$V \leq x \quad (11.31)$$

risulta:

$$F_V(x) = \begin{cases} 0 & \text{per } x < 0 \\ x & \text{per } 0 \leq x \leq 1 \\ 1 & \text{per } x > 1 \end{cases} \quad (11.32)$$

Consideriamo ora la distribuzione di probabilità cumulativa di un'altra variabile, la X , la cui densità di probabilità siamo interessati a simulare (cioè ad estrarre i relativi campioni).

Per definizione, sarà:

$$F_X(a) = \Pr [X \leq A] \quad (11.33)$$

avendo indicato con $\Pr [X \leq A]$ la probabilità dell'evento $X \leq A$.

Poniamo ora la seguente trasformazione di variabile aleatoria:

$$X = F_X^{-1}(V) \quad (11.34)$$

Questa formula, è stata presentata nel corso senza dimostrazione

che corrisponde a considerare l'inversa della funzione di ripartizione assegnata ad X e nel sostituire alla variabile indipendente la V (che, per ipotesi, è distribuita uniformemente).

¹²Slide del prof — Generazione di variabili aleatorie — pag 14
 Appunti di Damiano — pag 14
 Slide — Generazione di variabili aleatorie — pag 14
 Appunti — 2025-05-06 — pag 2 - 3

11.5.2.1 Dimostrazione dell'ammissibilità della trasformazione

¹³

Dimostrazione utile solo se ti interessa, sennò saltala

Di seguito la dimostrazione dell'ammissibilità della formula:

$$X = F_X^{-1}(V) \quad (11.35)$$

Verifichiamo che, con il cambiamento di variabile proposta, la $\Pr[X \leq A]$ restituisce proprio $F_X(a)$, e quindi il cambiamento di variabile è ammissibile e può essere utilizzato per generare i campioni di X a partire dai campioni di V.

In altre parole, si vuole fare la dimostrazione che $F_X^{-1}(V)$ ha le stesse proprietà di X.

Infatti:

$$\Pr[X \leq A] = \Pr[F_X^{-1}(V) \leq a] \quad (11.36)$$

Ora è evidente che:

$$\begin{aligned} \Pr[F_X^{-1}(V) \leq a] &= \Pr\{F_X[F_X^{-1}(V)] \leq F_X(a)\} \\ &= \Pr[V \leq F_X(a)] \end{aligned} \quad (11.37)$$

Per definizione risulta:

$$\Pr[V \leq F_X(a)] = F_V[F_X(a)] \quad (11.38)$$

e poiché:

$$F_V(x) = x \text{ per } 0 \leq x \leq 1 \quad (11.39)$$

ed effettivamente risulta:

$$0 \leq F_X(a) \leq 1 \quad (11.40)$$

si ha infine:

$$\Pr[V \leq F_X(a)] = F_X(a) \quad (11.41)$$

il che dimostra la tesi.

¹³Slide del prof — Generazione di variabili aleatorie — pag 14 - 15
 Appunti di Damiano — pag 14 - 15
 Slide — Generazione di variabili aleatorie — pag 14 - 15
 Appunti — 2025-05-06 — pag 2 - 3

11.5.2.2 Metodo di trasformazione

¹⁴

Sulla base della dimostrazione precedente, l'algoritmo per generare i campioni di una generica variabile aleatoria X , di cui si conosca in forma esplicita la distribuzione di probabilità cumulativa o la funzione di densità di probabilità (visto che una è la derivata dell'altra), può essere sintetizzato con il "metodo di trasformazione".

Di seguito l'algoritmo step-by-step:

1. Si inverte la funzione di ripartizione assegnata trovando $F_X^{-1}(x)$
2. Si genera una variabile aleatoria uniforme $V \in [0, 1]$
3. Si genera un campione (in inglese prende il nome di variate) della variabile X attraverso la trasformazione $X = F_X^{-1}(V)$

¹⁴Slide del prof — Generazione di variabili aleatorie — pag 15
 Appunti di Damiano — pag 15
 Slide — Generazione di variabili aleatorie — pag 15
 Appunti — 2025-05-06 — pag 3

11.5.2.3 Algoritmo Box-Muller

¹⁵

Quando si hanno variates della variabile gaussiana va sotto il nome di Algoritmo (o Metodo) di Box-Muller.

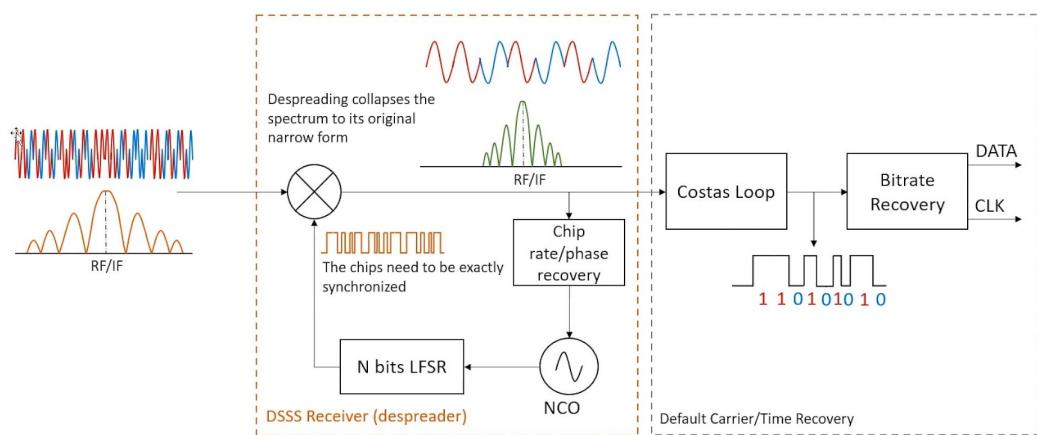
Da un punto di vista operativo, questi sono gli step dell'algoritmo:

1. Si generano le variates di due variabili uniformi tra -1 e +1
2. Si costituisce le variate della variabile uniforme tra 0 e 1, sommando i quadrati delle variates trovate al passo precedente
3. Si verifica che la variate generata non sia maggiore di 1, sennò si ritorna al primo step, altrimenti si procede
4. Si calcolano le variates della variabile gaussiana utilizzando le varietes del primo e del secondo passaggio

¹⁵Slide del prof — Generazione di variabili aleatorie — pag 20
Appunti di Damiano — pag 20
Slide — Generazione di variabili aleatorie — pag 20
Appunti — 2025-05-06 — pag 4

Capitolo 12

Spettro espanso



12.1 Tecniche a spettro espanso

¹

Un sistema a spettro espanso è realizzato sulla base di tecniche per cui il segnale trasmesso ha una occupazione spettrale maggiore (tra i 10^3 ed i 10^6 volte) di quella che avrebbe il convenzionale segnale modulato.

Di seguito le tre tecniche SS (Spread Spectrum) in cui si possono catalogare queste tecniche:

- FH-SS, cioè Frequency Hopping Spread Spectrum (principalmente utilizzato in ambito militare)
- TH-SS, cioè Time Hopping Spread Spectrum
- DS-SS, cioè Direct Sequence Spread Spectrum

¹Slide del prof — Spettro espanso — pag 1
Appunti di Damiano — pag 1
Slide — Spettro espanso — pag 1
Appunti — 2025-05-06 — pag 5

12.1.1 Teorema di Shannon

²

Lo Spread Spectrum si basa sul teorema di Shannon.

Di seguita il teorema.

La capacità C, espressa in bps (cioè bit al secondo), di un canale AWGN è data da:

$$\begin{aligned} C &= W \cdot \log_2 \left(1 + \frac{S}{N_0 \cdot W} \right) \\ &= W \cdot \log_2 \left(1 + \frac{S}{N} \right) \end{aligned} \quad (12.1)$$

La capacità C fornisce la massima quantità di informazione che può essere trasmessa lungo il canale con probabilità di errore arbitrariamente piccola.

La capacità C è una funzione di W, cioè la banda in frequenza, e il rapporto segnale-rumore $\frac{S}{N}$: quindi per aumentare C si deve aumentare la W, o aumentare il rapporto segnale rumore $\frac{S}{N}$, che a sua volta significa o aumentare il segnale S o diminuire il rumore N (cosa che è generalmente quasi impossibile fare).

Grazie al teorema di Shannon, possiamo esprimere questa formula:

$$\begin{aligned} C &= W \cdot \log_2 \left(1 + \frac{S}{N} \right) \\ &\downarrow \\ \frac{C}{W} &= \log_2 \left(1 + \frac{S}{N} \right) \\ &= 1.44 \ln \left(1 + \frac{S}{N} \right) \end{aligned} \quad (12.2)$$

Esprimere $\frac{C}{W}$ da una base 2 a una base e (che è appunto ln), è molto comodo per i calcoli.

Quando:

$$\frac{S}{N} \rightarrow 0 \quad (12.3)$$

allora:

²Slide del prof — Spettro espanso — pag 2 - 4
 Appunti di Damiano — pag 2 - 4
 Slide — Spettro espanso — pag 2 - 4
 Appunti — 2025-05-06 — pag 5 - 6

$$\begin{aligned}
 \frac{C}{W} &= \log_2 \left(1 + \frac{S}{N} \right) \\
 &\downarrow \\
 \frac{C}{W} &\approx 1.44 \cdot \frac{S}{N} \\
 &\downarrow \\
 W &\approx \frac{1}{1.44} \cdot \frac{N}{S} \cdot C
 \end{aligned} \tag{12.4}$$

Se C è finito, W tende a:

$$W \rightarrow +\infty \tag{12.5}$$

Se W tende ad infinito significa che l'informazione è minore del rumore termico, quindi è più difficile capire se qualcuno sta trasmettendo perché l'informazione sarà "nascosta" nel rumore.

12.2 Vantaggi e modello base di un sistema SS

3

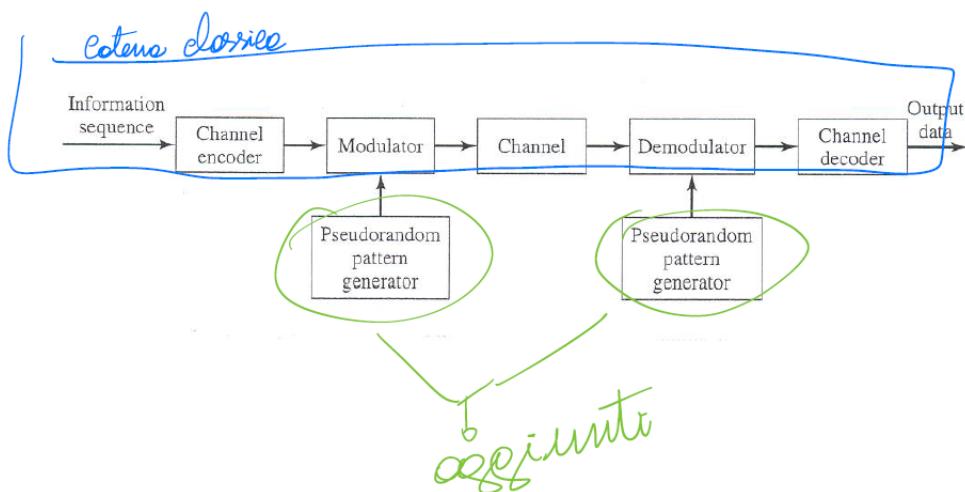
”Spalmare” lo spettro del segnale originale comporta numerosi vantaggi.

Di seguito ne elenchiamo i principali:

- Protezione dall’interferenza
- Spettri a bassa densità di potenza
- Sicurezza nelle comunicazioni (negli anni c’è stato un cambio di paradigma riguardo all’uso dell’SS come metodo di sicurezza delle comunicazioni)
- Messaggi ”schermo” per chi cerca di spiare
- Capacità di opporsi ad interferenze intenzionali (o il cosiddetto jamming)
- Il CDMA (Code Division Multiple Access) è una tecnica che sfrutta l’SS nella sua implementazione

Per messaggi schermo si intende chi intercetta il messaggio ha difficoltà a decifrato perchè è protetto da una lunga sequenza PN

Di seguito un modello base di un sistema SS:



12.3 Direct-Sequence Spread-Spectrum (DS-SS)

⁴

Per applicare la DS-SS, viene applicata una modulazione in 2-PSK, la quale, come abbiamo studiato, ha una occupazione spettrale di banda di R:

$$R = \frac{1}{T_b} \quad (12.6)$$

dove T_b è il tempo di bit.

Siccome il modulatore, in un sistema SS, viene moltiplicato per una sequenza pseudo-casuale, la banda occupata con la tecnica DS-SS è di B_c :

$$B_c \gg R \quad (12.7)$$

cioè, scritto per esteso, si intende che, il segnale DS-SS ha un'occupazione di banda molto larga.

Questo accade perchè, al trasmettitore, la banda del segnale viene espansa fino a W:

$$W = B_c \quad (12.8)$$

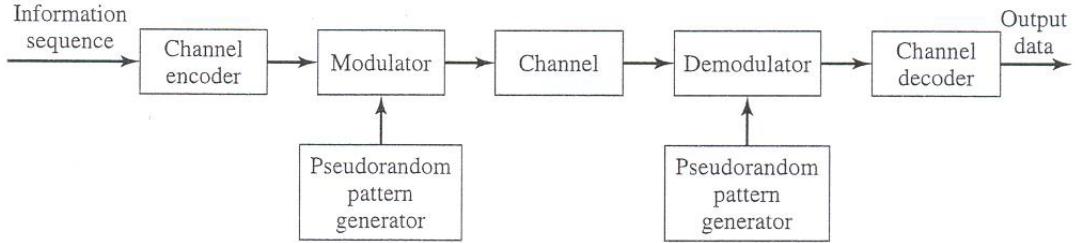
utilizzando le sequenze PN (Pseudo-Noise).

⁴Slide del prof — Spettro espanso — pag 7
Slide — Spettro espanso — pag 7
Appunti — 2025-05-09 — pag 2

12.3.1 DS-SS: da bit rate a chip rate, formulazione matematica, e spreading

5

Riportando il modello base di un sistema SS:



notiamo che il modulatore moltiplica la sequenza informativa per la sequenza PN.

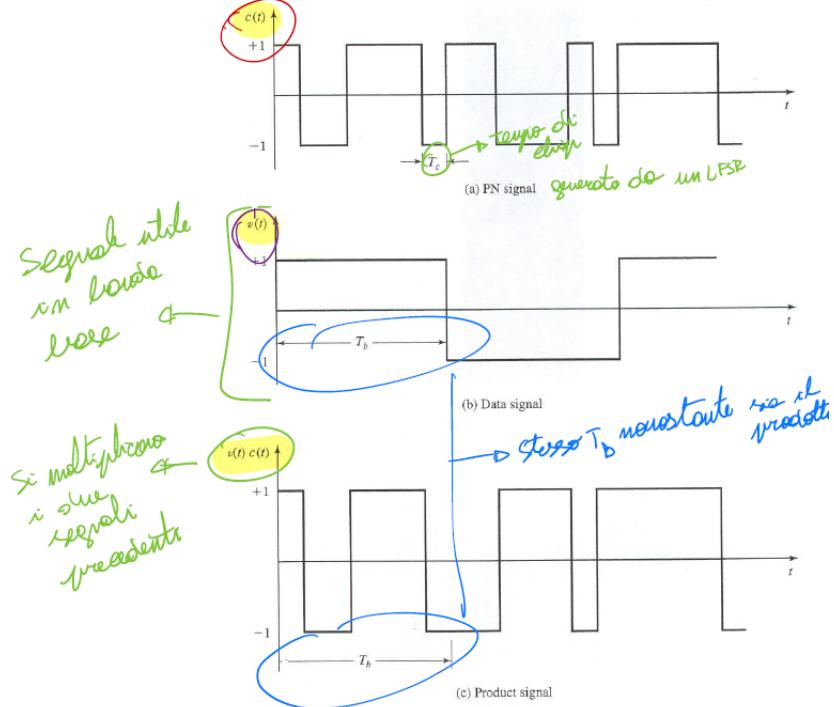
Per le formule matematiche che scriveremo di seguito, consideriamo:

$$T_c \ll T_b \quad (12.9)$$

dove:

- T_c viene definito come tempo di chip, cioè il tempo tra un'onda quadra e quella successiva nel segnale PN
- T_b viene definito come tempo di bit, che è il tempo del bit della sequenza informativa

Di seguito una figura dell'andamento dei segnali nel trasmettitore (con note):



Possiamo esprimere il segnale dell'informazione $v(t)$ come:

$$v(t) = \sum_{n=-\infty}^{+\infty} a_n \cdot g_T(t - n \cdot T_b) \quad (12.10)$$

⁵Slide del prof — Spettro espanso — pag 8 - 11

Appunti di Damiano — pag 9, 11

Slide — Spettro espanso — pag 8 - 11

Appunti — 2025-05-09 — pag 2 - 4

dove:

$$a_n = \pm 1 \quad (12.11)$$

e g_T è la funzione di base, dove, nei sistemi digitali, consideriamo degli impulsi rettangoli di periodo T_b .

Matematicamente, possiamo scrivere che l'andamento della sequenza PN va da $-\infty$ a $+\infty$, anche se, come abbiamo studiato nello scorso capitolo, un sistema reale elettronico non riesce ad elaborare infiniti elementi e verranno divisi in N elementi, in cui N è un numero molto elevato

Invece, possiamo esprimere il segnale PN come:

$$c(t) = \sum_{n=-\infty}^{+\infty} c_n \cdot g_T(t - n \cdot T_b) \quad (12.12)$$

dove:

$$c_n = \pm 1 \quad (12.13)$$

e p è la funzione di PN da un LFSR.

Il modulatore moltiplica la portante con il segnale $v(t)$ e il segnale $c(t)$, e forma il segnale $u(t)$, che si esprime come:

$$\begin{aligned} u(t) &= [v(t) \cdot c(t)] \cdot A_c \cdot \cos(2\pi f_c t) \\ &= A_c \cdot \cos(2\pi f_c t + \theta(t)) \end{aligned} \quad (12.14)$$

dove:

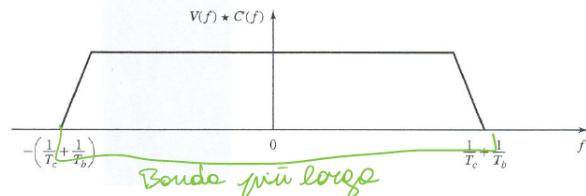
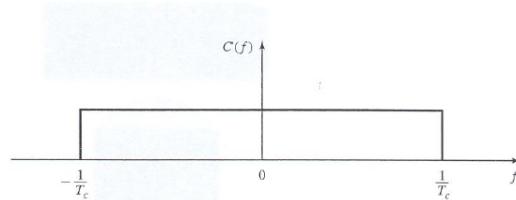
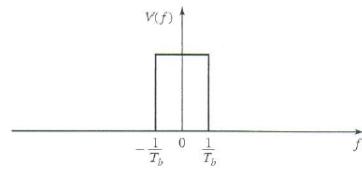
$$\theta(t) = v(t) \cdot c(t) \quad (12.15)$$

Alcune osservazioni riguardo l'espressione di $u(t)$:

- si fa una modulazione per traslare $v(t) \cdot c(t)$ alla frequenza f_c
- se $v(t)$ e $c(t)$ sono impulsi rettangolari, quindi il prodotto può valere solo 1 o 0

In frequenza possiamo confrontare i segnali $v(t)$, $c(f)$ e $u(t)$ con questo figura:

la Banda Base



Altre osservazione osservando questo grafico:

- non per forza i segnali devono essere rettangolari: vengono usati per comodità grafiche
- l'espansione della banda da $\frac{1}{T_b}$ diventa $\frac{1}{T_c} + \frac{1}{T_b}$, quindi la banda viene aumentata dalle 10^3 alle 10^6 volte
- il canale, per non avere distorsione, deve avere la banda disponibile
- Siccome si fa una moltiplicazione nel tempo, grazie alle proprietà tra tempo e frequenza, in frequenza avviene una convoluzione tra $V(f)$ e $C(f)$, che sono rispettivamente $v(t)$ e $c(t)$ in frequenza

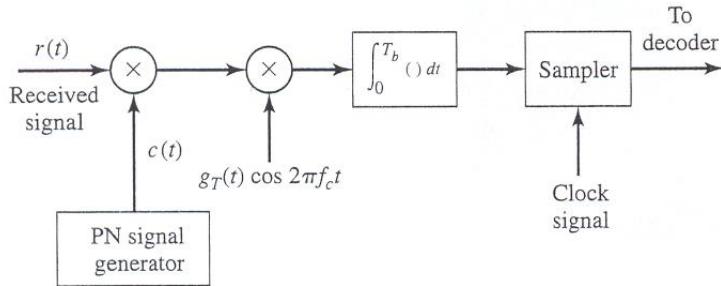
12.3.2 DS-SS: despreading, demodulazione e interferenze a banda stretta

6

Quando si de-modula in digitale, il primo step è la sincronizzazione, perché in digitale, se non si ha una sincronizzazione, si ricevono sequenze che non hanno senso.

Il rumore, in particolare, cambia i valori.

Di seguito lo schema di un demodulatore che fa anche da despreading:



Ricordando la formula di $r(t)$, per adesso senza rumore, cioè il segnale trasmesso:

$$r(t) = [v(t) \cdot c(t)] \cdot A_c \cdot \cos(2\pi f_c t) \quad (12.16)$$

Se c'è perfetta sincronizzazione, la fase in ricezione e in trasmissione è la stessa.

Quindi possiamo porre in ricezione:

$$\theta(t) = 0 \quad (12.17)$$

Se moltiplichiamo $r(t)$ per una sequenza PN $c(t)$, possiamo esprimere la formula del segnale in despreading:

$$\begin{aligned} r(t) \cdot c(t) &= \{[v(t) \cdot c(t)] \cdot A_c \cdot \cos(2\pi f_c t)\} \cdot c(t) \\ &= A_c \cdot v(t) \cdot c^2(t) \cos(2\pi f_c t) \\ &= A_c \cdot v(t) \cdot \cos(2\pi f_c t) \end{aligned} \quad (12.18)$$

Ma, se consideriamo l'interferenza $i(t)$, il segnale $r(t)$ ricevuto non è quello scritto precedente, ma diventa:

$$r(t) = [v(t) \cdot c(t)] \cdot A_c \cdot \cos(2\pi f_c t) + i(t) \quad (12.19)$$

Quindi, se consideriamo il despreading, se si moltiplica $r(t)$ per $c(t)$ non è quello precedente del caso ideale, bensì:

$$r(t) \cdot c(t) = [A_c \cdot \cos(2\pi f_c t)] + [i(t) \cdot c(t)] \quad (12.20)$$

La prima parte della equazione tra parentesi quadre, cioè $A_c \cdot \cos(2\pi f_c t)$ è voluta, invece $i(t) \cdot c(t)$ non è voluta, è indesiderata.

L'operazione di despreading trasforma il segnale a banda stretta del disturbo $i(t)$, in segnale banda larga perché $i(t)$ è banda stretta e viene moltiplicato per $c(t)$, che è a banda larga.

⁶Slide del prof — Spettro espanso — pag 12 - 13
Slide — Spettro espanso — pag 12 - 13
Appunti — 2025-05-09 — pag 5

Solo una frazione di $\frac{R}{W}$ (che per noi, in digitale è il rapporto segnale-rumore) della potenza interferente disturba il segnale utile.

Considerando l'interferenza ideale $i(t)$:

$$i(t) = A_I \cdot \cos(2\pi f_I t) \quad (12.21)$$

dove:

- A_I è l'ampiezza del disturbo $i(t)$
- f_I è la frequenza in cui si trova il disturbo a banda stretta $i(t)$

Sapendo che, il disturbo $i(t)$ è ergodico, stazionario e bilatero, possiamo scrivere che, la potenza del rumore $i(t)$ vale:

$$P_I = \frac{(A_I)^2}{2} \quad (12.22)$$

La densità spettrale di potenziale dell'interferente, dopo il despread, vale:

$$I_0 = \frac{P_I}{W} \quad (12.23)$$

dove W è la banda del canale.

La potenza interferente all'uscita del demodulatore vale:

$$\begin{aligned} I_0 \cdot R &= P_I \cdot \frac{R}{W} \\ &= \frac{P_I}{W/R} \end{aligned} \quad (12.24)$$

Quindi, la potenza interferente all'uscita del demodulatore dipende dal rapporto $\frac{W}{R}$: più $\frac{W}{R}$ è grande, e più $\frac{P_I}{W/R}$ diminuisce e viceversa.

12.4 Guadagno di processo (Process Gain)

7

Per guadagno di processo si intende di quando aumenta la banda in SS dalla banda originale.

Se consideriamo L_C il guadagno di processo, possiamo esprimere come:

$$L_C = \frac{W}{R} = \frac{T_b}{T_c} \quad (12.25)$$

dove:

- T_b è il tempo di bit
- T_c è il tempo di chip, cioè il periodo di un bit della sequenza Pseudo-Noise

Esprimere il rapporto tra i rapporti dei segnali rumori tra il segnale SS e quello con interferenza dentro all'SS, possiamo scriverla come:

$$\frac{\left(\frac{S}{N}\right)_o}{\left(\frac{S}{N}\right)_i} \quad (12.26)$$

dove:

- $\left(\frac{S}{N}\right)_o$ è il rapporto segnale rumore con SS
- $\left(\frac{S}{N}\right)_i$ è il rapporto segnale rumore con rumore con SS

applicando le proprietà dei dB, la frazione diventa una sottrazione, cioè:

$$\frac{\left(\frac{S}{N}\right)_o}{\left(\frac{S}{N}\right)_i} \rightarrow \left(\frac{S}{N}\right)_o \Big|_{dB} - \left(\frac{S}{N}\right)_i \Big|_{dB} \quad (12.27)$$

Nelle SS, generalmente $\left(\frac{S}{N}\right)_i \Big|_{dB}$ è molto basso: ecco perchè si preferisce utilizzare le tecniche SS.

La riduzione della potenza interferente, storicamente, è il primo motivo che ha condotto all'introduzione dei sistemi SS, migliorando le prestazioni in canali affetti da interferenza.

Un esempio pratico di come è molto utile una SS è proprio quella impiegata nella trasmissione Wi-Fi.

Ti consiglio questo video sul Wi-Fi 7 e di come venga gestito lo spettro in caso di interferenze:

https://www.youtube.com/watch?v=D_1qfsd0mJU

Arriva il Wi-Fi 7 in Italia: come funziona il nuovo standard portato da Iliad by Stefano Bolis

Per un maggior approfondimento, in particolare riguardo alla tecnica Puncturing per evitare le bande occupate da altri utenti o dovute alle interferenze, ti lascio il sito della intel dove ne discute in modo approfondito:

<https://www.intel.com/content/www/us/en/products/docs/wireless/wi-fi-7.html>

12.5 Ruolo delle sequenze PN

⁸

La sequenza Pseudo-Noise (o chiamata anche sequenza dei coefficienti $\{C_n\}$) è nota solo al ricevitore autorizzato: questo è un primo livello di protezione dato dall'SS con LFSR.

Un ricevitore che non conosce la sequenza PN, non può de-modulare correttamente l'informazione.

L'uso della sequenza PN fornisce un primo livello di sicurezza, che non è possibile con le modulazioni convenzionali.

La sequenza di coefficienti $\{C_n\}$ è come la chiave di una porta: se ce l'hai, puoi entrare nella stanza, se non ce l'hai, non entri.

Per adesso consideriamo il caso ideale, anche se sai già bene che questa caratteristica non è del tutto vero perché abbiamo a che fare con LFSR, i quali implementano degli algoritmi, e quindi deterministicici e non del tutto realmente casuali

Ma (in ingegneria c'è sempre un ma) il costo del miglioramento ottenuto è, oltre all'allagamento della banda, un incremento della complessità del sistema.

Da https://www.youtube.com/watch?v=EkX_00AaS10
La spada nella roccia: "Questo il mondo far girar"

Questo il mondo fa girar
per ogni qua c'è sempre un là
per ogni se "c'è sempre un ma"
per ogni su "c'è sempre un giù"
per ogni meh "c'è sempre un più"

⁸Slide del prof — Spettro espanso — pag 15
Appunti di Damiano — pag 15
Slide — Spettro espanso — pag 15
Appunti — 2025-05-09 — pag 5

12.6 Probabilità di errore in ricezione

9

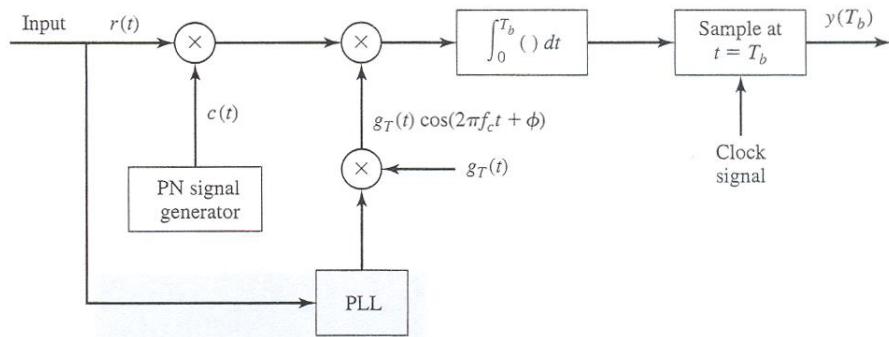
Come sempre, l'importante sono le considerazioni e non le formule per la teoria

In questo caso, nella SS, non consideriamo il rumore termico.

Consideriamo solo un segnale indesiderato interferente sovrapposto, che abbrevieremo con la dicitura interferente.

L'interferente influisce e aumenta la probabilità di errore del segnale in ricezione $r(t)$.

Consideriamo questo schema per un ricevitore SS:



Consideriamo il segnale il segnale prima dell'integratore $s(t)$ come:

$$s(t) = a_0 \cdot g_T(t) \cdot c(t) \cdot \cos(2\pi f_c t) \quad (12.28)$$

dove:

$$a_0 = \pm 1 \quad (12.29)$$

perchè è una sequenza digitale, si pone non solo $a_0 = \pm 1$, ma anche la funzione di base ha la seguente funzione di base:

$$g_T(t) = \begin{cases} \sqrt{\frac{2 \cdot E_b}{T_b}} & \text{per } 0 \leq t \leq T_b \\ 0 & \text{altrove} \end{cases} \quad (12.30)$$

e inoltre manca anche la formulazione matematica della sequenza Pseudo-Noise $c(t)$, che vale:

$$c(t) = \sum_{n=0}^{L_c-1} c_n \cdot p(t - n \cdot T_c) \quad (12.31)$$

In quanto la sequenza $\{c_n\}$ si può assumere incorrelata, allora l'energia tra il prodotto dei $\{c_n\}$ in ricezione e il prodotto dei $\{c_m\}$ coefficienti generati dal ricevitore vale:

$$E(c_n \cdot c_m) \approx E(c_n) \cdot E(c_m) \quad (12.32)$$

Sapendo che il singolo valore c_n appartiene a questo intervallo:

⁹Slide del prof — Spettro espanso — pag 16 - 19
 Appunti di Damiano — pag 17 - 18
 Slide — Spettro espanso — pag 16 - 19
 Appunti — 2025-05-09 — pag 5 - 6

$$c_n \in [-1, +1] \quad (12.33)$$

e ogni c_n della sequenza $\{c_n\}$ è equiprobabile, allora l'energia di ogni c_n vale:

$$E(c_n) = 0 \quad (12.34)$$

In altre parole, il valore medio di c_n nel periodo T_b è uguale a zero.

Invece l'energia del valore quadratico medio di ogni elemento di c_n , vale:

$$E(c_n^2) = 1 \quad (12.35)$$

Considerano il segnale in uscita, quindi dopo l'integratore e il campionatore, avremo $y_I(T_b)$ che vale:

$$\begin{aligned} y_I(T_b) &= \int_0^{T_b} c(t) \cdot g_T(t) \cos(2\pi f_c t + \theta) \cdot i(t) dt \\ &= \int_0^{T_b} \sum_{n=0}^{L_c-1} c_n \cdot p(t - n \cdot T_c) \cdot \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \cos(2\pi f_c t + \theta) \cdot i(t) dt \\ &= \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \sum_{n=0}^{L_c-1} c_n \cdot \int_0^{T_b} p(t - n \cdot T_c) \cdot i(t) \cdot \cos(2\pi f_c t + \theta) dt \end{aligned} \quad (12.36)$$

A questo punto, sapendo che:

$$T_c \ll T_b \quad (12.37)$$

cioè il tempo di chip T_c della PN è molto minore del tempo di bit T_b , allora possiamo svolgere l'integrale non nell'intervallo $[0, T_b]$, bensì nell'intervallo $[n \cdot T_c, (n+1) \cdot T_c]$, cioè tra due periodi di T_c .

Inoltre, sappiamo che, dalle sequenze PN con LFSR:

$$p(t - n \cdot T_c) \neq 0 \quad (12.38)$$

in un periodo multiplo di T_c .

Con queste osservazioni, possiamo riscrivere la formula di $y_I(T_b)$ come:

$$\begin{aligned} y_I(T_b) &= \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \sum_{n=0}^{L_c-1} c_n \cdot \int_0^{T_b} p(t - n \cdot T_c) \cdot i(t) \cdot \cos(2\pi f_c t + \theta) dt \\ &\downarrow \\ y_I(T_b) &= \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \sum_{n=0}^{L_c-1} c_n \cdot \int_{n \cdot T_c}^{(n+1) \cdot T_c} i(t) \cdot \cos(2\pi f_c t + \theta) dt \\ &= \sqrt{\frac{2 \cdot E_b}{T_b}} \cdot \sum_{n=0}^{L_c-1} c_n \cdot v_n \end{aligned} \quad (12.39)$$

Possiamo scrivere:

$$v_n = \int_{n \cdot T_c}^{(n+1) \cdot T_c} i(t) \cdot \cos(2\pi f_c t + \theta) dt \quad (12.40)$$

Grazie al teorema del limite centrale per L_c molto elevato, si può approssimare la sommatoria ad una gaussiana .

12.6.1 Interferente monocromatico

¹⁰

Se consideriamo interferenza monocromatica del tipo:

$$i(t) = \sqrt{2 \cdot P_I} \cdot \cos(2\pi f_c t + \Theta_I) \quad (12.41)$$

e consideriamo Θ_I una variabile aleatoria uniformemente distribuita tra:

$$\Theta_I \in [0, 2\pi] \quad (12.42)$$

Allora possiamo esprimere la probabilità di errore sul bit in una DS-SS in ricezione come:

$$P_{Eb} = \left(\sqrt{\frac{2 \cdot E_b}{I_0}} \right) \quad (12.43)$$

dove con la funzione Q si intende la funzione erfc della gaussiana.

Ricordati di utilizzare e di portare all'esame le tabelle dell'erfc

¹⁰Slide del prof — Spettro espanso — pag 20
Slide — Spettro espanso — pag 20

12.6.2 Interference margin

¹¹

Consideriamo che il rapporto tra l'energia del bit E_b e la potenza del rumore I_0 sul bit vale:

$$\frac{E_b}{I_0} = \frac{P_S \cdot T_b}{\frac{P_I}{W}} = \frac{\frac{P_S}{R}}{\frac{P_I}{W}} = \frac{\frac{W}{R}}{\frac{P_I}{P_S}} = \frac{L_C}{\frac{P_I}{P_S}} \quad (12.44)$$

e quindi:

$$\frac{E_b}{I_0} = \frac{L_C}{\frac{P_I}{P_S}} \quad (12.45)$$

dove:

- E_b è l'energia sul bit
- I_0 è l'energia dell'interferente sul bit
- P_S è la potenza del segnale con SS
- T_b è il tempo di bit
- P_I è la potenza dell'interferente
- W è la banda del segnale
- R è il rate
- L_C è il guadagno di processo dovuto alla SS

Allora in dB possiamo scrivere:

$$\left(\frac{P_I}{P_S} \right)_{dB} = (L_C)_{dB} - \left(\frac{E_b}{I_0} \right)_{dB} \quad (12.46)$$

$\frac{P_I}{P_S}$ prende il nome di Interference margin.

¹¹Slide del prof — Spettro espanso — pag 21
Appunti di Damiano — pag 21
Slide — Spettro espanso — pag 21

12.6.2.1 Interference margin: esempio numerico

¹²

Considerando i seguenti dati:

$$\begin{cases} \left(\frac{E_b}{I_0}\right)_{dB} = 10dB \\ \left(\frac{P_I}{P_S}\right)_{dB} = 20dB \end{cases} \quad (12.47)$$

dove $\frac{E_b}{I_0}$ è il rapporto segnale-rumore senza SS, allora abbiamo bisogno di un L_C in dB che vale:

$$(L_C)_{dB} = 30 \text{ dB} \quad (12.48)$$

che in numeri non logaritmici vale:

$$\begin{aligned} (L_C)_{dB} &= 30 \text{ dB} \\ &\downarrow \\ \frac{W}{R} &= 1000 \end{aligned} \quad (12.49)$$

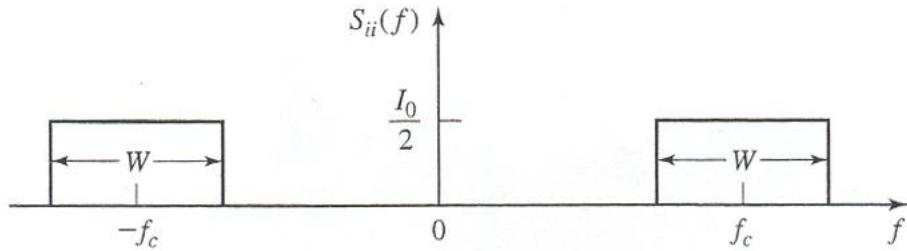
Questo significa che la potenza dell'interferente può essere 1000 volte maggiore della potenza del segnale utile, ma, grazie all'SS, si può mantenere la qualità desiderata.

¹²Slide del prof — Spettro espanso — pag 21
Appunti di Damiano — pag 21
Slide — Spettro espanso — pag 21

12.6.3 Interferente a banda larga

¹³

Rispetto al caso monocromatico precedente, consideriamo adesso un interferente a banda larga come nella seguente figura:



Senza dimostrazione, formule che ci dovete credere come un dogma.

La probabilità di errore sul bit considerando una interferente a banda larga come in figura, vale:

$$P_{Eb} = Q \left(\frac{2 \cdot E_b}{I_0 \cdot J(\alpha)} \right) \quad (12.50)$$

dove:

$$J(\alpha) = 2 \cdot \int_0^{+\infty} \left(1 - \frac{x}{\alpha} \right) \cdot \frac{\sin(pi \cdot x)}{pi \cdot x} dx \quad (12.51)$$

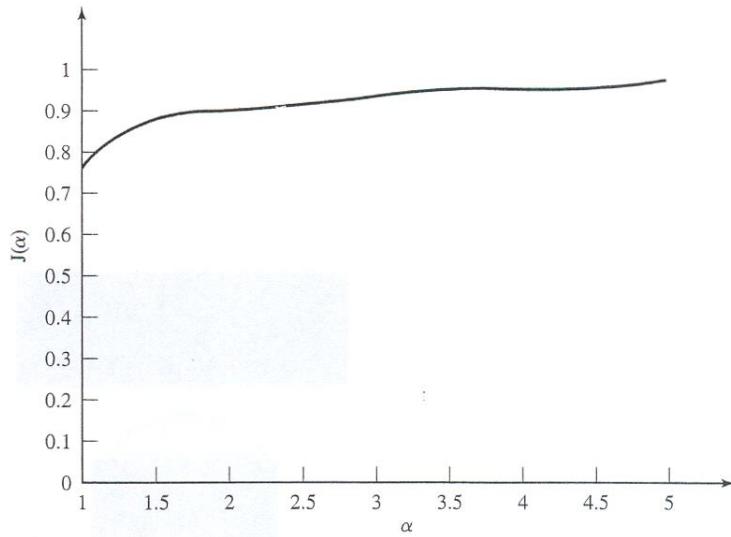
dove a sua volta:

$$\alpha = W \cdot T_c \quad (12.52)$$

¹³Slide del prof — Spettro espanso — pag 22 - 23
Slide — Spettro espanso — pag 22 - 23
Appunti — 2025-07-21 Ricevimento — pag 13

Generalmente non calcoliamo questo integrale, ma, come spesso accade, utilizziamo i grafici.

Possiamo tracciare l'andamento di $J(\alpha)$ come:



La probabilità di errore con la interferente a banda larga è minore rispetto all'esempio della banda mono-cromatica.

12.7 DS-SS: applicazioni

14

Il prof scrive le applicazioni del DS-SS in Inglese (cosa buona e giusta), però qui proverò a dare una traduzione in italiano perchè so che l'inglese non è il vostro forte.

Utilizzo questa occasione per ricordare di studiare e parlare l'inglese bene, non a livello madrelingua, ma almeno bene.

Sulla rete ci sono taaaanti materiali: da video a podcast, a dirette su Twitch a libri. Tutto.

Per piacere, usufruitene

Di seguito delle applicazioni del DS-SS:

- Trasmissioni di segnale a basso rilevabilità (Low-detectability signal transmission)
- Accesso multiplo a divisione di codice (Code Division Multiple Access o CDMA)
- Comunicazioni in canali con percorso multipli (Communication over channel with multipath)

I percorsi multipli li avevamo già visti nel precedente corso.

Un breve ripasso da Tds:

<https://github.com/ciccio25/appunti-teoria-dei-segnali/blob/main/Appunti%20Teoria%20dei%20segnali.pdf>

Capitolo 5.3.3 - Distorsione causata da percorsi multipli - pag 50 - 51

I multipath sono alla base delle comunicazioni moderne: basti pensare alla tecnica MI-MO (Multiple Input, Multiple Output) del Wi-Fi o del 5G.

Un esempio in questo video:

<https://youtu.be/e9x5PANDAuU?si=FdVu5UgL0YWPbdBY>

MU-MIMO nelle reti Wi-Fi: come funziona, perchè è importante e potenzialità con il Wi-Fi 6 by Stefano Bolis

Invece per visualizzare il concetto da un punto di vista generale:

- <https://www.youtube.com/watch?v=B2h8FohnwUs>
MU-MIMO Explained by PowerCert Animated Videos
- <https://youtu.be/0ncIWlhsu1A?si=2txuVFDsNfOhYnXA>
What is Multi-User MIMO Communications (MU MIMO)? by Iain Explains Signals, Systems, and Digital Comms

¹⁴Slide del prof — Spettro espanso — pag 24
Slide — Spettro espanso — pag 24
Appunti — 2025-05-13 — pag 2

12.7.1 Low detectability signal transmission

¹⁵

Il segnale utile è trasmesso con un livello di potenza molto basso (confrontabile con il rumore).

Sapendo che la potenza del rumore P_N vale:

$$P_N = W \cdot N_0 \quad (12.53)$$

dove:

- W è banda che consideriamo
- N_0 è il rumore

Se consideriamo P_R la potenza del segnale in ricezione, possiamo compararla con la potenza del rumore P_N :

$$\frac{P_R}{P_N} \ll 1 \quad (12.54)$$

cioè, come scritto precedentemente, il segnale utile ha un livello di potenza molto basso.

Un ricevitore autorizzato, e che ha quindi i famosi coefficienti $\{c_n\}$ dell'LFSR Pseudo-Noise, può recuperare l'informazione, con qualità soddisfacente, utilizzando il processo di despreading.

Invece, il ricevitore non autorizzato non può conseguire il necessario miglioramento del rapporto segnale-rumore.

Quindi, con il DS-SS si ha una bassa probabilità di intercettazione.

Possiamo considerare il segnale con DS-SS, grazie a questa proprietà, come segnale LPI (Low Probability of Intercept).

¹⁵Slide del prof — Spettro espanso — pag 25
Slide — Spettro espanso — pag 25
Appunti — 2025-05-13 — pag 2

12.7.2 CDMA

¹⁶

Il guadagno di processo L_C può essere utilizzato per consentire la trasmissione simultanea di segnali provenienti da utenti diversi (e sovrapposti nel tempo e in frequenza).

Il caso delle comunicazioni moderne, specialmente quelle senza filo, è come entrare in una stanza piena di persone, dove si sta stretti, ognuno parla ad alta voce e tu sei lì che devi capire la persona che ti sta davanti o dall'altra parte della stanza.

Nella realtà o scappi dalla stanza o cerchi il possibile di farti capire con le gesta (come un vero italiano stereotipato).

Nelle tecnologie moderne SS, questo "casino" è intenzione e inoltre è possibile, se si è autorizzati, capire qualsiasi persona nella stanza

Ciascun utente deve essere identificato da una diversa sequenza Pseudo-Noise, chiamata firma (o signature in inglese).

Le sequenze PN devono essere scelte in modo da avere buone proprietà di correlazione (quindi auto-correlazione e cross-correlazione).

Ricordo al volo che la correlazione è diversa dalla auto-correlazione, che a sua volta è diversa dalla correlazione mutua

Al ricevitore, i segnali con sequenza di codice diversa da quella di interesse appaiono come segnali interferenti.

L'entità dell'interferenza dipende principalmente dalle seguenti caratteristiche:

- dal numero di utenti
- dalle proprietà di cross-correlazione delle sequenze scelte

Il grosso vantaggio del CDMA rispetto agli altri schemi di multiplazione è la flessibilità.

¹⁶Slide del prof — Spettro espanso — pag 28 - 29
Appunti di Damiano — pag 28 - 29
Appunti — 2025-05-13 — pag 3

In altri casi, ad esempio la FDM (Frequency Division Multiplexing), per aggiungere un altro trasmettitore, bisogna ridividere in parti uguali la banda assegnata W.

Invece nella CDMA ho un numero massimo di firme PN.

Finché un trasmettitore trasmette, avrà la sua firma, quando non trasmetterà quella stessa firma verrà assegnata ad un altro utente, così da non aumentare il disturbo dovuto agli altri segnali e/o utenti nello spettro.

È ovvio che, nelle tecniche di trasmissione moderne, si utilizza il CDMA insieme ad altre tecniche di modulazione, in questo caso di esempio CDMA-FDM, oppure CDMA-TDM e così via
Piccola parentesi storica: l'azienda Qualcomm ha spinto tanto, quando introdusse l'UMTS/GSM (o anche chiamata dagli operatori "Great Source of Money"), quello che comunemente chiamavamo 3G, sul CDMA.

Un po' di storia di quello che era successo ai tempi:

<https://www.youtube.com/watch?v=Zi0SjDSa8Dc>

The iPhone was 3G's Killer App by Asianonmetry

Parlando del nostro paese nostrano, ti lascio qualche spot della Tim sull'UMTS del tempo:

- <https://youtu.be/9WqjGfQwd2E?si=9HVaFiqpRTfgeiAM>
- <https://www.youtube.com/watch?v=cIRuGydsN84>

Invece nel 5G, il CDMA viene poco utilizzato.

Video molto utile riguardo al 5G:

<https://www.youtube.com/watch?v=cLEfKpsSAEU>

Whatever Happened to Millimeter-Wave 5G? by Asianonmetry

Per informazione personale, e come parentesi storica, ti lascio anche questi video:

<https://www.youtube.com/watch?v=9e0Jat-sKLl>

How'd we get to 5G? The history of cell networks — Upscaled by Engadget

12.7.2.1 CDMA: valutazione di massima delle prestazioni (worst case scenario)

¹⁷

Ipotizzando che tutti i segnali hanno identica potenza media (o come scritto nei documenti power control)

Il rapporto segnale-interferente complessivo vale:

$$\begin{aligned} \frac{P_S}{P_N} &= \frac{P_S}{(N_u - 1) \cdot P_S} \\ &= \frac{1}{N_u - 1} \end{aligned} \tag{12.55}$$

dove N_u sono il numero degli utenti simultanei.

Il rapporto $\frac{1}{N_u - 1}$ sarà sempre minore o uguale a 1.

Possiamo esprimere la potenza di rumore P_N :

$$P_N = (N_u - 1) \cdot P_S \tag{12.56}$$

perchè, per il singolo utente, nel CDMA, gli altri utenti sono di disturbo.

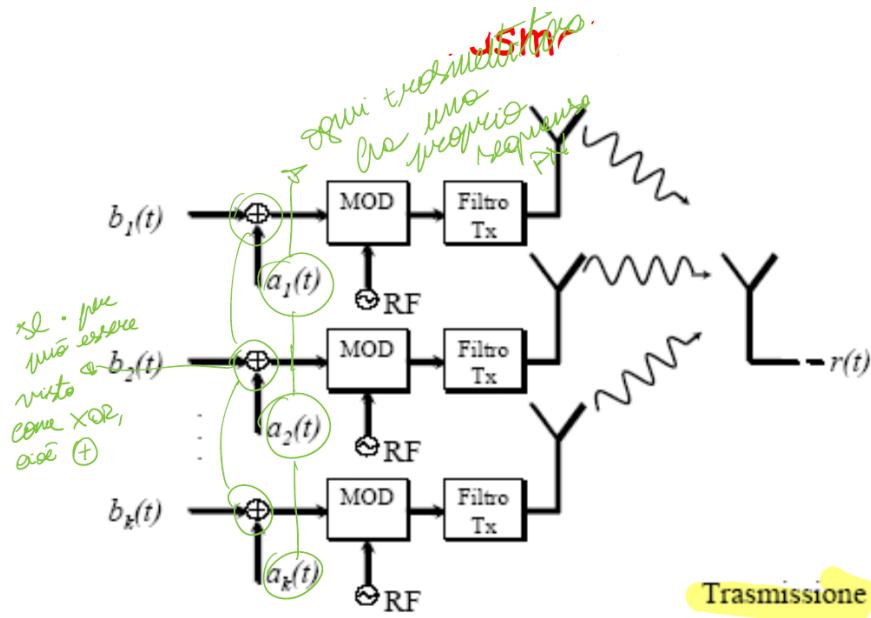
A parità degli altri parametri, N_u può essere aumentato utilizzando un opportuno codice di canale.

¹⁷Slide del prof — Spettro espanso — pag 30
Appunti di Damiano — pag 30
Slide — Spettro espanso — pag 30
Appunti — 2025-05-13 — pag 3

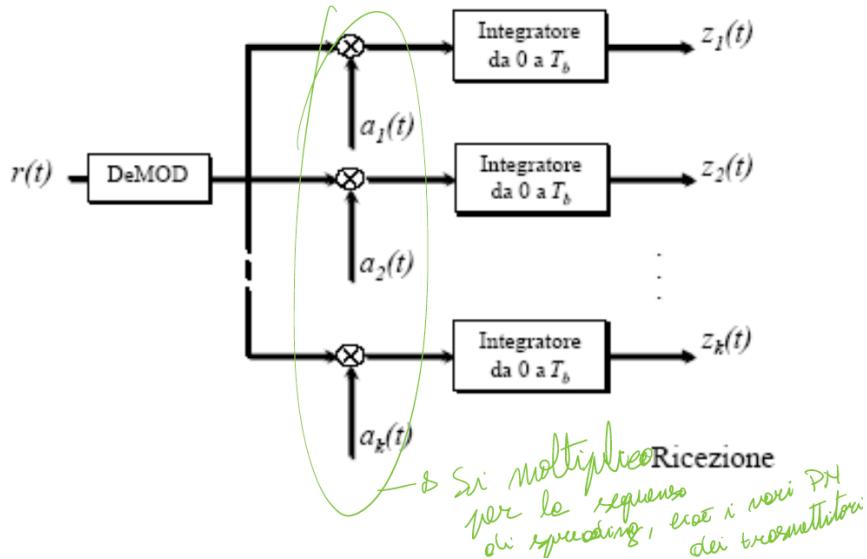
12.7.2.2 CDMA: Trasmettitore e ricevitore

18

Di seguito uno schema a blocchi un trasmettitori CDMA (con note):



Di seguito uno schema a blocchi di un ricevitore in cui i trasmettitori applicano il CDMA (con note):



¹⁸ Slide del prof — Spettro espanso — pag 32 - 33

Appunti di Damiano — pag 33

Slide — Spettro espanso — pag 32 - 33

Appunti — 2025-05-13 — pag 4

12.7.3 CDMA: valutazione più accurata

¹⁹

Facendo una valutazione più accurata del segnale in ingresso in CDMA, avremo che il segnale $r(t)$ vale:

$$r(t) = \sum_{k=1}^{N_u} \sqrt{2 \cdot P_k} \cdot a_k(t - \tau_k) \cdot b_k(t - \tau_k) \cdot \cos[\gamma_k(t)] + n(t) \quad (12.57)$$

dove:

- $n(t)$ sono termini di rumore
- nelle parentesi $(t - \tau_k)$, τ_k indica il tempo di ritardo del k -esimo trasmettitore perché i trasmettitori, nel caso reale, non sono sincronizzati allo stesso tempo, e quindi si considera un ritardo

Inoltre, nella formula di $r(t)$ comprare il termine $\gamma_k(t)$ dove:

$$\gamma_k(t) = \omega_c \cdot (t - \tau_k) + \phi_k \quad (12.58)$$

dove ϕ_k è la fase del singolo trasmettitore k -esimo.

Se consideriamo l' i -esimo utente, abbiamo che:

$$\begin{cases} \tau_i = 0 \\ \phi_i = 0 \end{cases} \quad (12.59)$$

La demodulazione si realizza moltiplicando $r(t)$ per $d(t)$:

$$d(t) = \sqrt{2} \cdot a_i(t) \cdot \cos(\omega_c \cdot t) \quad (12.60)$$

Allora possiamo esprimere questa formula:

$$\begin{aligned} z_i &= r_i \cdot d(t) \\ &= \sqrt{P_i} \cdot b_i \cdot \int_0^{T_b} a_i^2(t) dt \\ &+ \sum_{k=1, k \neq i}^{N_u} \sqrt{P_k} \cdot \int_0^{T_b} a_k(t - \tau_k) \cdot b_k(t - \tau_k) \cdot a_i(t) \cdot \cos(\phi_k - \omega_c \cdot \tau_k) dt \\ &+ \eta_i(t) \end{aligned} \quad (12.61)$$

Nella formula di z_i , il termine:

$$\sqrt{P_i} \cdot b_i \cdot \int_0^{T_b} a_i^2(t) dt \quad (12.62)$$

¹⁹Slide del prof — Spettro espanso — pag 34 - 36
Appunti di Damiano — pag 34 - 36

Slide — Spettro espanso — pag 34 - 36
Appunti — 2025-05-13 — pag 4 - 5

è il termine utile perchè il termine di auto-correlazione.

Invece, l'altra parte dell'equazione è considerata disturbo.

In particolare η_i è il contributo del rumore termico che può espresso come segue:

$$\eta_i = \int_0^{T_b} n'(t) \cdot a_i(t) dt \quad (12.63)$$

Se:

$$\tau_k = 0 \quad (12.64)$$

allora il termine:

$$\int_0^{T_b} a_k(t - \tau_k) \cdot b_k(t - \tau_k) \cdot a_i(t) \quad (12.65)$$

diventa la correlazione mutua tra a_k e a_i .

Nel CDMA possiamo considerare due scenari.

Il primo scenario, quello ideale, ha le seguenti caratteristiche:

- ritardi uguali per tutti i segnali, quindi:

$$\tau_k = 0 \text{ per } \forall k \quad (12.66)$$

- fase iniziale nulla per tutti le portanti, quindi:

$$\phi_k = 0 \quad (12.67)$$

In questo primo scenario, z_i diventa:

$$\begin{aligned} z_i &= \sqrt{P_i} \cdot b_i \cdot R_{ii}(0) \\ &+ \sum_{k=1}^{N_u} \sqrt{P_k} \cdot b_k \cdot R_{ki}(0) \\ &+ \eta_i \end{aligned} \quad (12.68)$$

dove:

- $R_{ii}(0)$ è l'auto-correlazione di $a_i(t)$
- $R_{ki}(0)$ è la mutua-correlazione tra $a_k(t)$ e $a_i(t)$

Se ci troviamo nel caso ideale:

$$R_{ki}(0) = 0 \quad (12.69)$$

e di conseguenza z_i diventa:

$$z_i = \sqrt{P_i} \cdot b_i \cdot R_{ii}(0) + \eta_i \quad (12.70)$$

cioè, in parole, se la mutua-correlazione tra $a_k(t)$ e $a_i(t)$ è nulla, allora avremo z_i minimo e potremo ricostruire il messaggio senza problemi.

Se invece consideriamo il secondo scenario, dove si hanno ritardi diversi per i segnali per ogni utente, cioè:

$$\tau_i \neq \tau_k \quad (12.71)$$

avremo una trasmissione asincrona, ad esempio in un caricamento di un radiomobile, ad esempio un cellulare, verso una stazione radio-base.

Un caso storico di tecnologia mobile, al suo fallimento dovuto al ritardo di sincronizzazione e di fase tra la cella radio-base e il dispositivo radiomobile è lo standard WiMAX, il quale ha perso il successo tecnologico mobile grazie alla migliore implementazione dello standard 4G.

Il tema è discusso più a fondo in questo video:

<https://youtu.be/bq7iUap1YbQ?si=X8UFFf18bmVDRNs0C>
How 4G killed Sprint by Speedify

Cioè:

$$R_{ki}(\tau_k) \rightarrow R_{ki}(0) \quad (12.72)$$

o in inglese, out-of-phase cross-correlation.

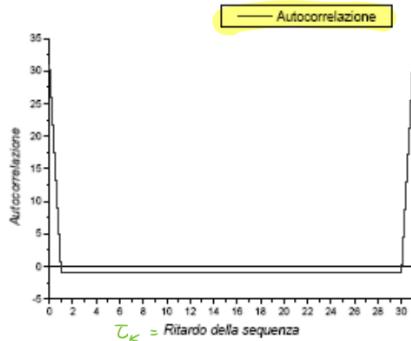
Anche se a priori non conosciamo quanto vale $R_{ki}(\tau_k)$, vogliamo che sia la più bassa possibile.

12.7.4 Auto-correlazione, cross-correlazione, sequenza PN, prestazioni CD-MA

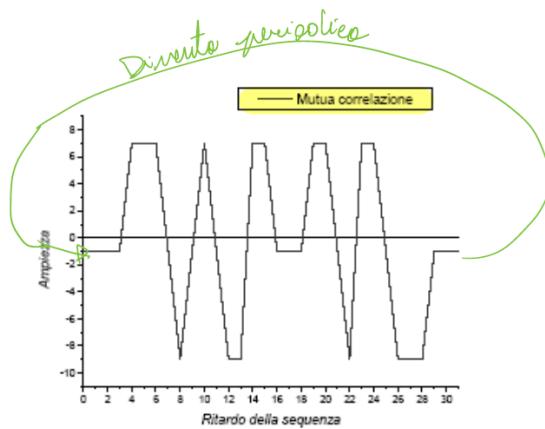
20

Come al solito, ci sono molto utili i grafici piuttosto che le formule.

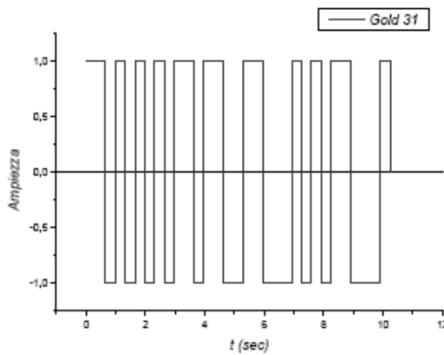
Per esempio una auto-correlazione di una sequenza di lunghezza 31 è la seguente:



Una gold-sequence (ne approfondiremo dopo) di lunghezza 31 ha l'andamento di una mutua correlazione di questo tipo:

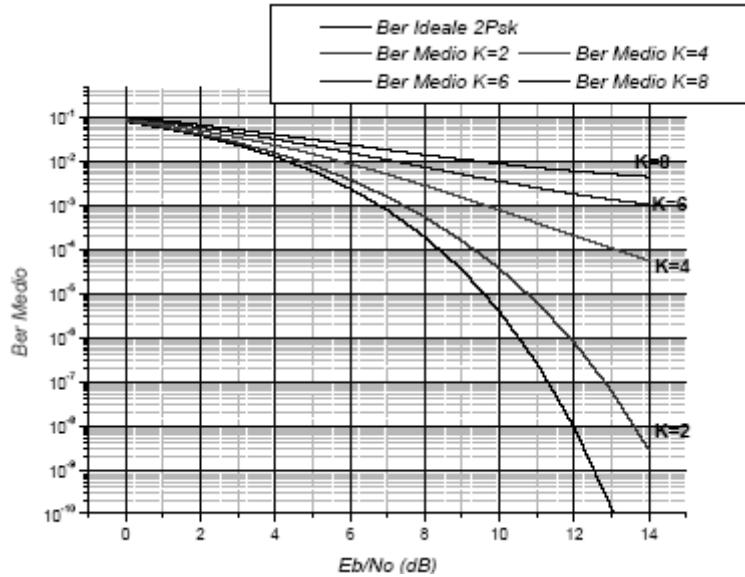


Una gold sequence PN ha questo andamento:



²⁰Slide del prof — Spettro espanso — pag 37 - 40
 Appunti di Damiano — pag 37 - 38, 40
 Slide — Spettro espanso — pag 37 - 40
 Appunti — 2025-05-13 — pag 5

Un andamento della prestazione del BER rispetto al numero di utenti K, e l'SNR, che nel caso digitale è $\frac{E_b}{N_0}$, è il seguente:



Come si può notare, all'aumentare degli utenti K, il BER peggiora (la curva si trova più in alto).

12.7.5 Communications over multipath channels

²¹

Come il caso precedente della valutazione della CDMA, nella DS-SS bisogna considerare anche le comunicazioni lungo percorsi multipli, dall'inglese multipath channels.

Possiamo considerare due scenari su questi percorsi multipli per una DS-SS:

1. utente singolo
2. multi-utente

Considerando il primo caso, cioè il singolo utente, possiamo modellare il segnale ricevuto $r(t)$ in DS-SS come:

$$r(t) = a(t) \cdot b(t) \cos(\omega_c \cdot t) + \sum_{i=1}^L A_i \cdot a(t - \tau_i) \cdot b(t - \tau_i) \cos(\gamma_i(t)) + n(t) \quad (12.73)$$

Nella formula di $r(t)$, il termine:

$$\sum_{i=1}^L A_i \cdot a(t - \tau_i) \cdot b(t - \tau_i) \cos(\gamma_i(t)) + n(t) \quad (12.74)$$

dove con L si indicano le L repliche del segnale $a(t) \cdot b(t) \cos(\omega_c \cdot t)$ dovute dai percorsi multipli.

Al ricevitore è presente un integratore e quindi riceveremo il segnale z :

$$z = b \cdot R_{SS}(0) + \sum_{i=1}^L A_i \cdot \int_0^{T_b} a(t - \tau_i) \cdot b(t - \tau_i) \cos(\gamma_i(t)) + \eta(t) \quad (12.75)$$

Invece, nel secondo caso, cioè in caso di comunicazione in percorsi multipli con diversi utenti, occorre tener conto degli altri segnali interferenti, ciascuno con il proprio insieme di cammini multipli.

La modellazione matematica sarebbe più complicata da svolgere rispetto al caso del singolo utente.

²¹Slide del prof — Spettro espanso — pag 41
Slide — Spettro espanso — pag 41
Appunti — 2025-05-13 — pag 6

12.7.6 Pulsed interference

²²

Si definisce pulsed interference un segnale interferente a burst (nel tempo), con spettro di potenza piatto nell'intera banda di trasmissione (in frequenza).

Ricordiamo bene la relazione tra tempo e frequenza dal vecchio corso.

La differenza tra un disturbo a banda larga generico, e un disturbo burst sta nel suo tempo, che ha un carattere "impulsato".

L'interferenza di tipo impulsato, realmente, disturba un pacchetto di bit.

Andando più a fondo nella trattazione matematica, se indichiamo con il carattere β la percentuale di tempo per cui l'interferente disturba la trasmissione, se P_I è la potenza media trasmessa, negli intervalli di disturbo "impulsato", la potenza dell'interferente diventa P'_I :

$$P'_I = \frac{P_I}{\beta} \quad (12.76)$$

In presenza di solo disturbo impulsato, quindi non considerando rumore termico o altre cause di disturbo, possiamo calcolarci la probabilità di errore sul bit P_{Eb} che vale:

$$P_{Eb} = Q \left(\sqrt{\frac{2 \cdot E_b}{\frac{I_0}{\beta}}} \right) \quad (12.77)$$

Questa formula di P_{Eb} è valida quando l'interferente è attivo.

Come visto nei precedenti capitoli, con Q si intende di calcolarci l'erfc dell'argomento di Q.

²²Slide del prof — Spettro espanso — pag 42 - 43
 Appunti di Damiano — pag 42 - 43
 Slide — Spettro espanso — pag 42 - 43
 Appunti — 2025-05-13 — pag 6 - 7

12.7.6.1 Pulsed interference: worst case scenario

²³

Dalla formula di probabilità di errore media è:

$$P_{Eb} = \beta \cdot Q\left(\sqrt{\frac{2 \cdot E_b}{\frac{I_0}{\beta}}}\right) \quad (12.78)$$

Se poniamo $\beta = 1$ nella formula di P_{Eb} , ricaviamo la formula di trasmissione in DS-SS con solo rumore termico

Derivando questa formula (senza fare la seconda derivata), cioè:

$$\frac{\partial P_{Eb}}{\partial \beta} = 0 \quad (12.79)$$

possiamo trovare il valore β^* che massimizza la funzione P_{Eb} media.

Quindi, senza fare i calcoli, possiamo scrivere che β^* vale:

$$\beta^* = \begin{cases} \frac{0.71}{\frac{E_b}{I_0}} & \text{per } \frac{E_b}{I_0} \geq 0.71 \\ 1 & \text{per } \frac{E_b}{I_0} \leq 0.71 \end{cases} \quad (12.80)$$

Mettendo il valore di β^* in P_{Eb} , possiamo ricavarci $P_{Eb}|_{max}$ come:

$$P_{Eb}|_{max} = \begin{cases} \frac{0.082}{\frac{E_b}{I_0}} & \text{per } \frac{E_b}{I_0} \geq 0.71 \\ Q\left(\sqrt{\frac{2 \cdot E_b}{I_0}}\right) & \text{per } \frac{E_b}{I_0} \leq 0.71 \end{cases} \quad (12.81)$$

²³Slide del prof — Spettro espanso — pag 44
Appunti di Damiano — pag 44
Slide — Spettro espanso — pag 44
Appunti — 2025-05-13 — pag 7

12.7.6.2 Pulsed interference: prestazioni

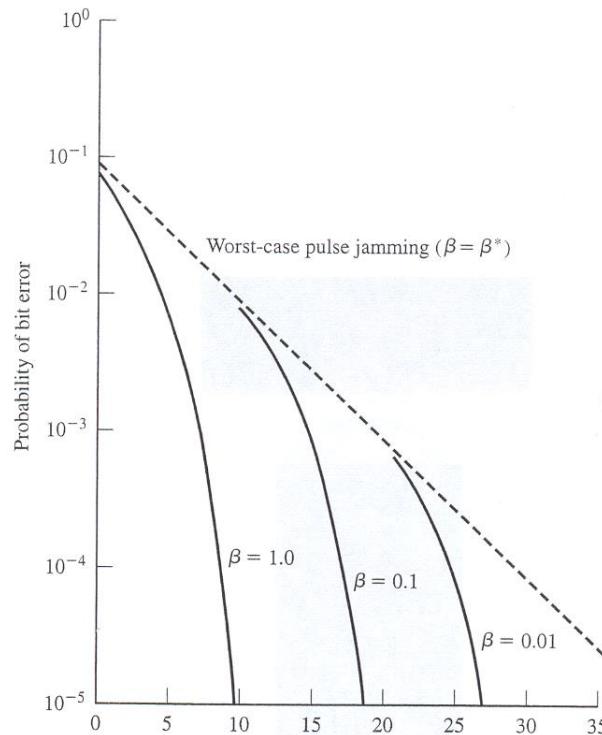
²⁴

Sapendo la formula media di P_{Eb} è:

$$P_{Eb} = \beta \cdot Q\left(\sqrt{\frac{2 \cdot E_b}{I_0}}\right) \quad (12.82)$$

possiamo metterla in relazione con il rapporto segnale-rumore, cioè in digitale $\frac{E_b}{I_0}$.

In base al valore di β , possiamo tracciare il seguente grafico per confrontare le prestazioni delle pulsed intereference:



La retta tratteggiata in figura è proprio la formula che avevamo trovato precedentemente, cioè:

$$P_{Eb}|_{max} = \frac{0.082}{\frac{E_b}{I_0}} \text{ per } \frac{E_b}{I_0} \geq 0.71 \quad (12.83)$$

Dal seguente grafico possiamo svolgere due considerazioni riguardo al caso peggiore:

- la penalizzazione può essere molto rilevante
- l'uso di un FEC (Forward Error Correction, cioè un algoritmo di correzione in cui non si chiede la ritrasmissione del segnale) classico, non risolve il problema

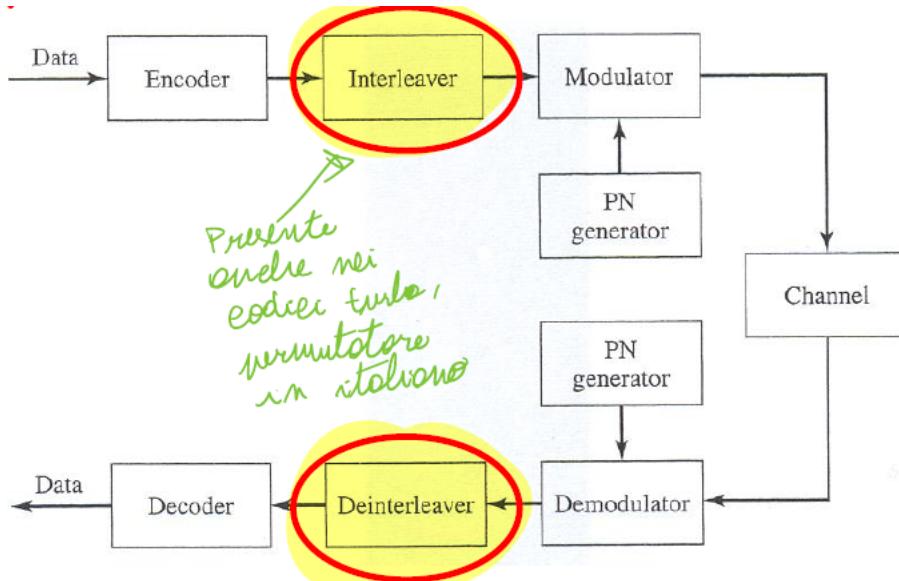
²⁴Slide del prof — Spettro espanso — pag 45
Appunti di Damiano — pag 45
Slide — Spettro espanso — pag 45
Appunti — 2025-05-13 — pag 7

12.7.6.3 Pulsed interference: miglioramento delle prestazioni tramite interleaver

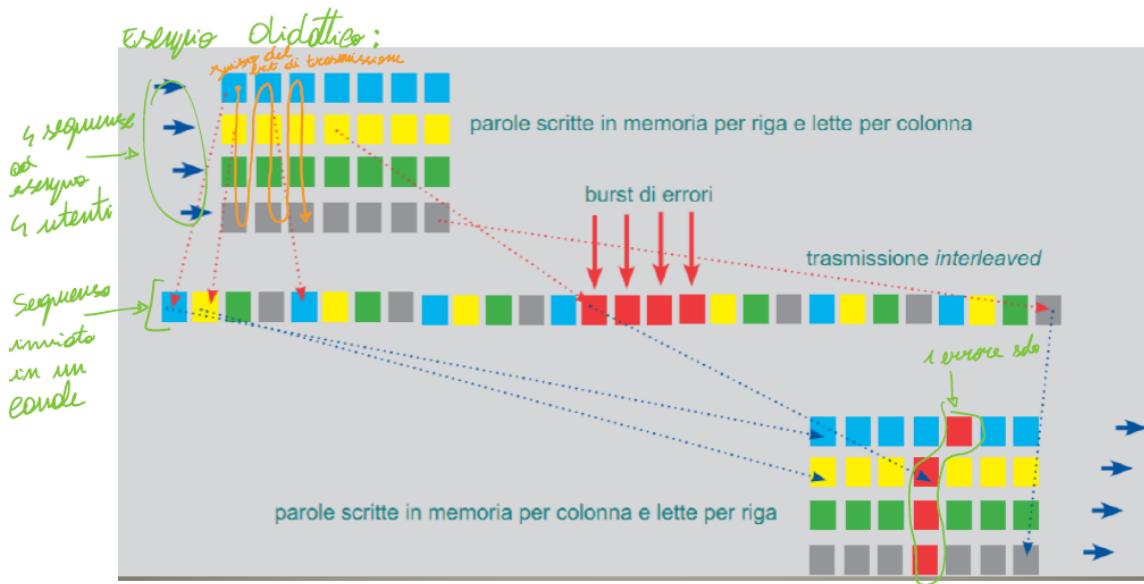
25

Siccome, in questi casi peggiori, non possiamo usare i FEC, allora si è pensato di usare l'interleaver (in italiano permutatore) per migliore la trasmissione.

Uno schema di una trasmissione DS-SS con interleaver al trasmettitore e de-interleaver al ricevitore è il seguente (con note):



Da un punto di vista grafico, con un esempio di 4 utenti, possiamo descrivere il funzionamento dell'interleaver, che opera righe per colonne, con il seguente schema (con note):



Come si vede dalla figura, a causa del rumore burst, avremo solo un errore di bit per ogni utente, quindi l'errore viene "distribuito".

Ma, il problema è che questo "mischiamento" ha bisogno di tempo e aggiunge ritardo al sistema, specialmente se abbiamo a che fare con delle matrici molto grandi.

²⁵Slide del prof — Spettro espanso — pag 46 - 47
Appunti di Damiano — pag 46 - 47
Slide — Spettro espanso — pag 46 - 47
Appunti — 2025-05-13 — pag 7 - 8

12.8 Generazione di sequenze PN

26

Un piccolo ripasso del PN con LFSR, ma approfondiremo un aspetto molto importante che vedrete poi

Una sequenza pseudo-random o pseudo-noise (PN) è una sequenza binaria in cui la proprietà di correlazione è simile (per quanto possibile) a quella del rumore bianco.

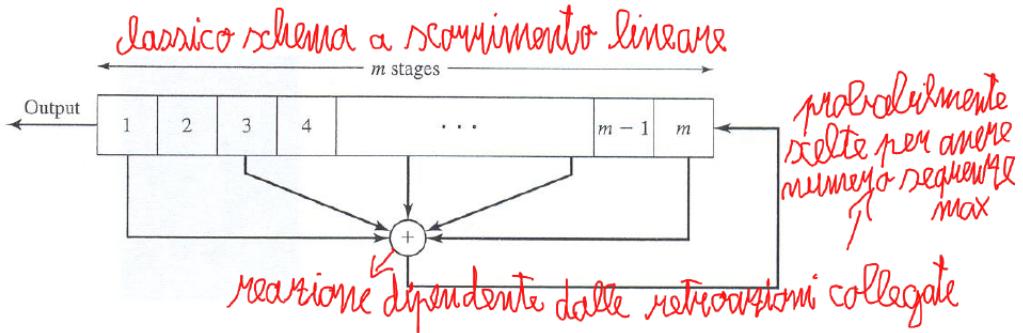
12.8.1 Maximum-length sequences (m-sequences)

27

Dal precedente capitolo, sappiamo che una generazione di una sequenza può essere svolta con un LFSR (Linear Feedback Shift Register), il quale ha una lunghezza L massima pari a:

$$L = 2^m - 1 \quad (12.84)$$

Uno schema di un LFSR per la generazione di PN è la sequenza (con note):



Il LFSR è inizializzato con un valore qualsiasi che prende il nome di seme (o seed in inglese) purché diverso da zero.

Per ogni m ci sono più possibilità di collegamento retro-azionati.

Di seguito un esempio per ogni m :

SHIFT-REGISTER CONNECTIONS FOR GENERATING ML SEQUENCES

m	Stages connected to modulo-2 adder	m	Stages connected to modulo-2 adder	m	States connected to modulo-2 adder
2	1, 2	13	1, 10, 11, 13	24	1, 18, 23, 24
3	1, 3	14	1, 5, 9, 14	25	1, 23
4	1, 4	15	1, 15	26	1, 21, 25, 26
5	1, 4	16	1, 5, 14, 16	27	1, 23, 26, 27
6	1, 6	17	1, 15	28	1, 26
7	1, 7	18	1, 12	29	1, 28
8	1, 5, 6, 7	19	1, 15, 18, 19	30	1, 8, 29, 30
9	1, 6	20	1, 18	31	1, 29
10	1, 8	21	1, 20	32	1, 11, 31, 32
11	1, 10	22	1, 22	33	1, 21
12	1, 7, 9, 12	23	1, 19	34	1, 8, 33, 34

La sequenza prodotta dall'LFSR è, originariamente, una sequenza di 0 e 1.

Ai fini dell'utilizzo nel sistema DS-SS, la sequenza viene convertita in una sequenza di coefficiente $\{c_n\}$ in cui il singolo elemento c_n diventa:

$$\begin{cases} c_n = 0 \rightarrow c_n = -1 \\ c_n = 1 \rightarrow c_n = 1 \end{cases} \quad (12.85)$$

Come studiato nel capitolo precedente, siccome siamo nel valore dei discreti \mathbb{Z}_2 , la funzione di autocorrelazione $R_c(m)$ non sarà un integrale, bensì una sommatoria esprimibile come:

²⁷ Slide del prof — Spettro espanso — pag 49 - 51
Appunti di Damiano — pag 49 - 50
Appunti — 2025-05-16 — pag 1

$$R_c(m) = \sum_{n=1}^L c_n \cdot c_{n+m} \text{ per } 0 \leq m \leq L-1 \quad (12.86)$$

Anche la funzione di auto-correlazione $R_c(m)$ è periodica di periodo L.

12.8.1.1 Proprietà delle m-sequences

²⁸

La sequenza PN ideale dovrebbe avere:

$$R_C(m) = \begin{cases} L & \text{per } m = 0 \\ 0 & \text{per } 1 \leq m \leq L - 1 \end{cases} \quad (12.87)$$

Invece nelle m-sequenze l'auto-correlazione $R_C(m)$ è uguale a:

$$R_C(m) = \begin{cases} L & \text{per } m = 0 \\ -1 & \text{per } 1 \leq m \leq L - 1 \end{cases} \quad (12.88)$$

Se facciamo il rapporto tra l'autocorrelazione e la correlazione, calcoliamo la cross-correlazione, che vale:

$$\frac{R_C(m)}{R_C(0)} = -\frac{1}{L} \quad (12.89)$$

Se L tende ad infinito, il rapporto diventa:

$$\frac{R_C(m)}{R_C(0)} \rightarrow 0 \text{ per } L \rightarrow +\infty \quad (12.90)$$

²⁸Slide del prof — Spettro espanso — pag 52
Appunti di Damiano — pag 52
Appunti — 2025-05-16 — pag 1

12.8.2 Cross-correlazione

29

Come sempre detto, generalmente si useranno le seguenti tabelle (con note):

PEAK CROSSCORRELATIONS OF m SEQUENCES AND GOLD SEQUENCES							
m	$L = 2^m - 1$	Number	m sequences		Gold sequences		
			$R(0)$	R_{max}	$R_{max}/R(0)$	R_{max}	$L = 2^m R_{max}/R(0)$
				Peak cross-correlation R_{max}			
3	7	2	5	$\frac{5}{7} = 0.71$	5	9	0.71
4	15	2	9	0.60	9	17	0.60
5	31	6	11	0.35	9	33	0.29
6	63	6	23	$\frac{23}{63} = 0.36$	17	:	0.27
7	127	18	41	$\frac{41}{127} = 0.32$	17	:	0.13
8	255	16	95	0.37	33	:	0.13
9	511	48	113	0.22	33	:	0.06
10	1023	60	383	0.37	65	:	0.06
11	2047	176	287	0.14	65	:	0.03
12	4095	144	1407	$\frac{1407}{4095} = 0.34$	129	$4097 \downarrow$ chiam disponibili qualità alta	$0.03 = \frac{129}{4097}$

(11 è meglio di 12)

numero sequenze possibili (chiavi disponibili)

molte di più con sequenze gold

All'aumentare di m , il rapporto $\frac{R_{max}}{R(0)}$ non decresce: ha un andamento altalenante.

Nella tabella qui in figura, $\frac{R_{max}}{R(0)}$ è troppo alto.

Nel DS-SS si vuole avere un $\frac{R_{max}}{R(0)}$ più bassa possibile: per risolvere questo problema vengono usate le gold-sequence.

Se $\frac{R_{max}}{R(0)}$, si possono servire più utenti nel CDMA.

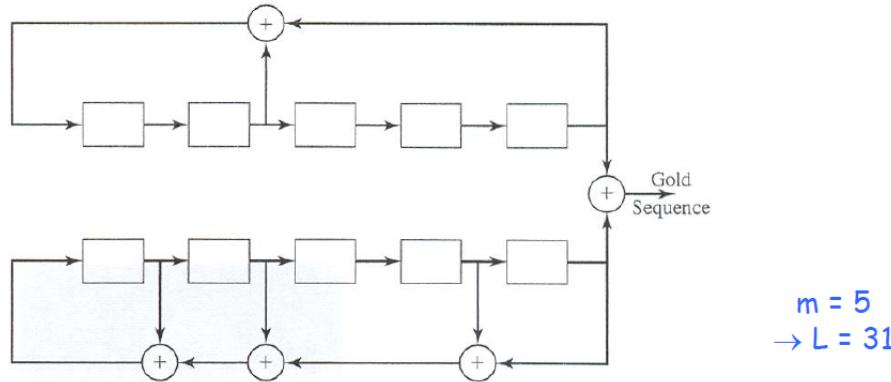
²⁹Slide del prof — Spettro espanso — pag 53
Appunti di Damiano — pag 53
Appunti — 2025-05-16 — pag 1 - 2

12.9 Gold-sequences

30

Le gold-sequences sono quelle sequenze che sono ottenute a partire da una coppia selezionata di m-sequenze, e formando le somme modulo 2 per ognuna delle L versioni shiftate di una sequenza rispetto all'altra.

Di seguito un esempio di una sequenza gold (con note):



Per generare una sequenza gold, bisogna partire da una condizione iniziale, diversa da zero, nel primo generatore.

Si ottengono 2^m sequenze di gold cambiando il seme del secondo generatore, da 0 a L, in cui:

$$L = 2^m - 1 \quad (12.91)$$

Questo è equivalente a dire che si fa partire il secondo generatore da una di 2^m diverse condizioni iniziali.

Un'ultima sequenza di Gold è quella fornita dal secondo generatore da solo, ottenuta ponendo a zero il seme del primo generatore.

In totale si hanno $2^m + 1$ sequenze di Gold.

Nelle sequenze di Gold, si possono avere 3 valori possibili di cross-correlazione:

- -1
- $-t(m)$
- $t(m) - 2$

dove $t(m)$ è la massima cross-correlazione: può essere indicata anche con R_{max} .

$t(m)$ può avere due tipi di valori in base alla scelta di m:

$$t(m) = \begin{cases} 2^{\frac{m+1}{2}} + 1 & \text{per } m \text{ dispari} \\ 2^{\frac{m+2}{2}} + 1 & \text{per } m \text{ pari} \end{cases} \quad (12.92)$$

Siccome bisogna sempre progettare il caso peggiore, bisogna prendere il caso peggiore, che è $t(m)$ con m pari.

Ricordiamo che il numero massimo delle sequenze è $L + 2$.

³⁰Slide del prof — Spettro espanso — pag 57 - 61
Appunti di Damiano — pag 57 - 60
Appunti — 2025-05-16 — pag 3 - 4

Per m dispari e sufficientemente grande, R_{max} vale:

$$\begin{aligned} R_{max} &= \sqrt{2^{m+1}} + 1 \\ &\approx \sqrt{2 \cdot 2^m} \\ &\approx \sqrt{2 \cdot L} \end{aligned} \tag{12.93}$$

Se invece m è pari e sufficientemente grande, R_{max} vale:

$$\begin{aligned} R_{max} &= \sqrt{2^{m+2}} + 1 \\ &\approx \sqrt{2^2 \cdot 2^m} \\ &\approx 2 \cdot \sqrt{L} \end{aligned} \tag{12.94}$$

12.10 Lower bound

³¹

Per aumentare gli utenti, ad esempio nel CDMA DS-SS, oltre alle sequenze gold, sono presenti altri tipi di sequenze per massimizzare $\frac{R_{max}}{R(0)}$.

Per un insieme di N sequenze di periodo L, il matematico Welch definì il Welch bound, cioè un valore minimo possibile teorico di R_{max} , che vale:

$$R_{max} \geq L \cdot \sqrt{\frac{N-1}{N \cdot L - 1}} \quad (12.95)$$

Quando N è sufficientemente grande, la formula precedente diventa:

$$\begin{aligned} R_{max} &\geq L \cdot \sqrt{\frac{N-1}{N \cdot L - 1}} \\ &\downarrow \\ R_{max} &\geq \sqrt{L} \text{ per } N \gg 1 \end{aligned} \quad (12.96)$$

Confrontando l' R_{max} con quello che si può ottenere con le sequenze gold:

$$R_{max} \geq \sqrt{2 \cdot L} \quad (12.97)$$

se si vuole massimizzare i clienti e si considera come parametro solo R_{max} , meglio prendere le sequenze di Welch che hanno un R_{max} minore.

Le sequenze di Gold si definiscono sub-ottime (significa che si può fare di meglio).

Invece altri sequenze, come quella di Kasami, sono definite ottime perché hanno:

$$R_{max} \approx \sqrt{L} \quad (12.98)$$

³¹Slide del prof — Spettro espanso — pag 62
Appunti di Damiano — pag 62
Appunti — 2025-05-16 — pag 5

12.11 Kasami sequences (small set)

32

Le sequenze Kasami sono definite ottime perché hanno:

$$R_{max} \approx \sqrt{L} \quad (12.99)$$

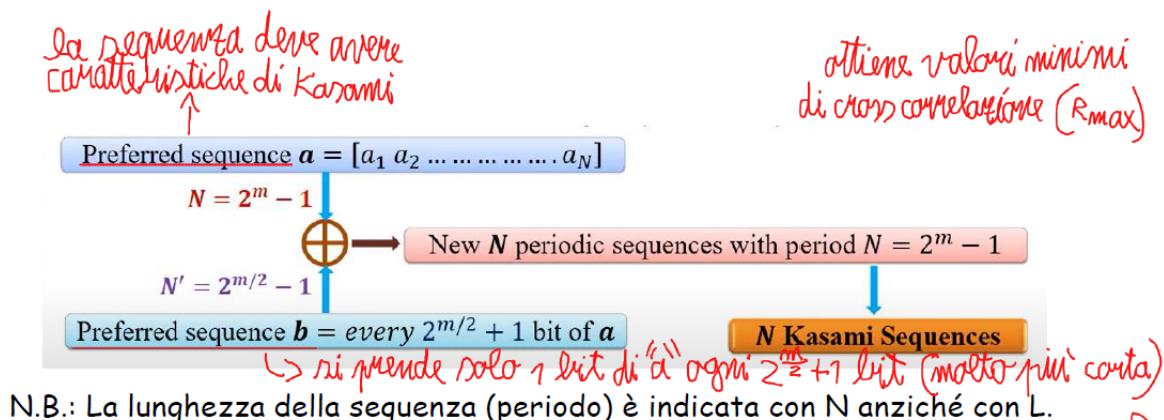
In altri termini, si utilizzano le sequenze Kasami per ottenere valori minimi di cross-correlazione, cioè R_{max} .

Possiamo definire due tipi di sequenze Kasami:

- small set
- Large set

Adesso andiamo ad analizzare le sequenze Kasami small set.

Uno schema su come si ottengono le sequenze Kasami (con note):



Dalle sequenze Kasami si hanno N sequenze, dove N vale:

$$N = 2^{\frac{m}{2}} \quad (12.100)$$

A parità di m , come le sequenze gold, le sequenze Kasami sono minori rispetto a quelle gold.

Cioè si passa dalle $2^m - 1$ sequenze delle gold alle $2^{\frac{m}{2}}$ delle Kasami.

Come illustrato dallo schema, le sequenze Kasami sono ottenute considerando due vettori \vec{a} e \vec{b} e tutte le versioni traslate di \vec{b} , che ha periodo $2^{\frac{m}{2}} - 1$, quindi $2^{\frac{m}{2}-2}$.

Vettori e sequenze vengono utilizzati come sinonimi, perchè, come ben sai dai corsi di programmazione, un computer o un microcontrollore non riescono ad interpretare un vettore come un qualcosa di fisico, bensì il vettore viene interpretato come una sequenza di bit.

Andando a fondo, la sequenza \vec{b} viene ottenuta per "decimazione" delle sequenza \vec{a} , selezionando un bit ogni $2^{\frac{m}{2}} + 1$ bit di \vec{a} .

\vec{b} è periodica e di periodo $2^{\frac{m}{2}} - 1$.

Ogni periodo di \vec{a} include $2^{\frac{m}{2}} + 1$ periodi di \vec{b} .

Inoltre, per costruire le sequenze Kasami, m deve essere pari.

La funzione di cross-correlazione assume 3 valori:

- -t
- -1
- t - 2

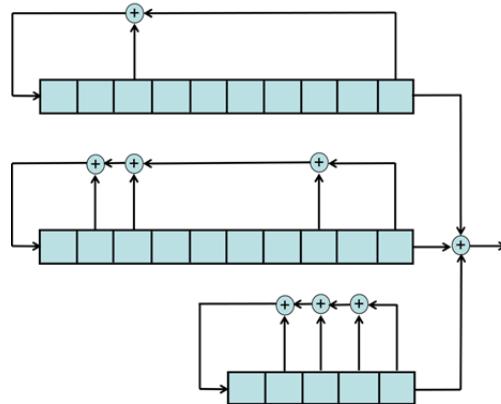
dove t vale:

$$t = 2^{\frac{m}{2}} + 1 \quad (12.101)$$

12.12 Kasami sequences (large set)

33

Un esempio di schema di sequenze Kasami large-set è il seguente:



In questo caso, al posto delle small set, si utilizzano due sequenze preferite, e la versione decimata di una di queste.

La versione decimata deve essere di lunghezza:

$$\text{mod } (m, 4) = 2 \quad (12.102)$$

Ad esempio, $m = 2$ è un valore possibile da utilizzare.

Il periodo L è come quello delle small set:

$$L = 2^m - 1 \quad (12.103)$$

Invece il numero di sequenze N è diverso, cioè:

$$N = 2^{\frac{3 \cdot m}{2}} + 2^{\frac{m}{2}} \quad (12.104)$$

La funzione di cross-correlazione assume questi 5 valori:

- -t
- -s
- -1
- s - 2
- t - 2

dove:

$$\begin{cases} t = 2^{\frac{m}{2+1}} + 1 \\ s = \frac{t+1}{2} \end{cases} \quad (12.105)$$

R_{max} che si ottiene con le sequenze Kasami large set è un valore maggiore rispetto alle small set.

³³Slide del prof — Spettro espanso — pag 66 - 68
Appunti di Damiano — pag 66 - 68
Appunti — 2025-05-16 — pag 6

12.13 Sequenze di Barker

³⁴

Una caratteristica principale delle sequenze utilizzate fino adesso è stata l'auto-correlazione periodica.

Invece le sequenze di Barker sono sequenze che hanno un'auto-correlazione aperiodica.

Nello specifico, definita la funzione di auto-correlazione aperiodica di una sequenza bipolare $\{b_j\}$, dove $j = 1, \dots, L$ come:

$$C_k = \sum_{j=1}^{L-K} b_j \cdot b_{j+k} \text{ per } k = 0, \dots, L-1 \quad (12.106)$$

Come scritto precedente, l'auto-correlazione periodica è diversa rispetto all'auto-correlazione aperiodica; questo lo si può notare anche dalle sommatorie impiegate.

Useremo $\sum_{j=1}^{L-K}$ per l'auto-correlazione aperiodica, invece $\sum_{j=1}^L$ per l'auto-correlazione periodica.

Una sequenza, o codice, di Barker verifica la proprietà:

$$|C_k| \leq 1 \text{ per } \forall k \neq 0 \quad (12.107)$$

Inoltre, non esistono codici di Barker con $L > 13$.

Esempi di utilizzi di sequenze di Barker sono:

- come prefisso per il recupero del sincronismo di una trama
- come sequenza di speading nell'IEEE 802.11b, cioè il Wi-Fi 1 adottato nel 1999, se $L = 11$

³⁴Slide del prof — Spettro espanso — pag 69 - 71
Appunti di Damiano — pag 69
Appunti — 2025-05-16 — pag 7

12.14 Frequency-Hopping Spread-Spectrum (FH-SS)

35

La FH-SS è una tecnologia usata principalmente nell'ambito militare.

La banda disponibile W è divisa in un certo numero elevato di slot di frequenze distinte.

Nella trasmissione dell' n -esimo simbolo, si utilizza uno o più di questi slot.

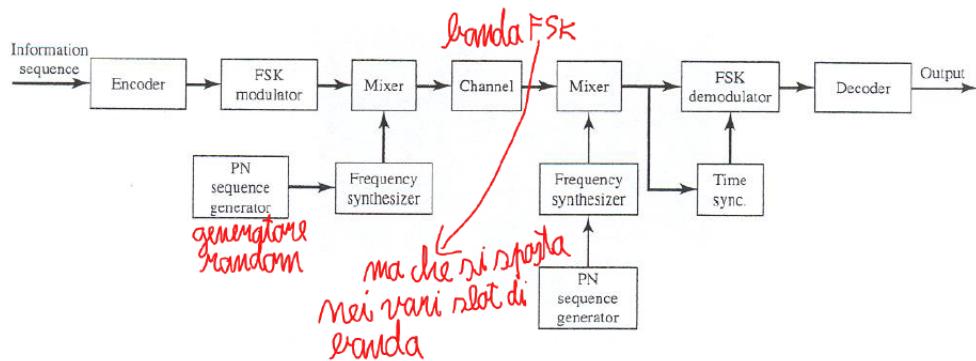
Gli slot vengono selezionati in modo pseudo-random, sulla base dell'uscita di un generatore PN.

Oppure, spiegato in altri termini, il trasmettitore salta da una banda all'altra secondo regole pseudo-casuali.

Solo chi è a conoscenza della sequenza pseudo-causale tra una frequenza all'altra può ricevere il segnale.

Chi intercetta una porzione del segnale poi non sa dove si troverà lo stesso segnale in un secondo momento.

Di seguito uno schema di trasmissione e ricezione in FH-SS (con note):

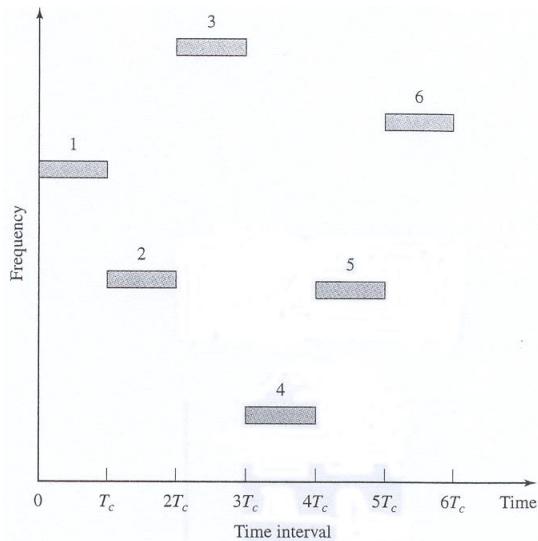


Generalmente la modulazione utilizzata nell'FH-SS è M-FSK.

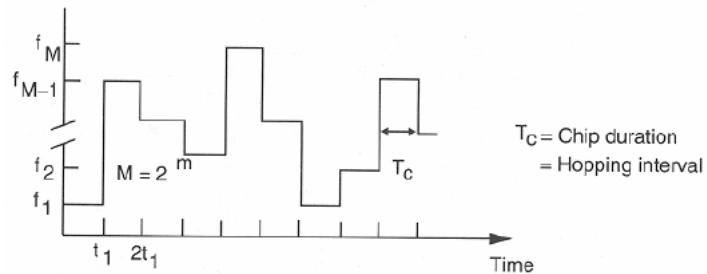
La frequenza in uscita dal modulatore FSK viene traslata in ragione dell'uscita del sintetizzatore di frequenza.

³⁵Slide del prof — Spettro espanso — pag 72 - 75
Appunti di Damiano — pag 72 - 73
Appunti — 2025-05-16 — pag 8

Di seguito un grafico di una trasmissione FH-SS in cui viene indicata la frequenza utilizzata lungo il tempo:



oppure un altro grafico di FH-SS:



Rispetto alle figure, l'intervallo di Hopping non verrà indicato come T_c , bensì con T_H .

I vantaggi della FH-SS sono le seguenti:

- minori problemi di sincronizzazioni
- guadagni di processo più elevati rispetto alla DS-SS
- notevole immunità al jamming intenzionale

12.14.1 Classificazione dei sistemi FH-SS

³⁶

I sistemi FH-SS si possono classificare in:

- SFH, o Slow Frequency Hopping
- FFH, o Fast Frequency Hopping
- IFH, o Intermediate Frequency Hopping

La loro classificazione dipende dal periodi di Hopping T_H rispetto al periodo di bit T_b .

Si definisce una SF quando:

$$T_H > T_b \quad (12.108)$$

o, visto in dal punto di vista del tasso (o rate in inglese):

$$R_H < R_b \quad (12.109)$$

Andando più nel dettaglio, nel periodo di tempo fra due portanti vengono generati più bit, cioè, sulla stessa portante vengono trasmesse sequenze di alcuni o molti bit (in certe implementazioni vengono trasmessi più di 1000 bit nella stessa frequenza).

Si definisce una FFH quando:

$$\begin{array}{c} T_H < T_b \\ \Downarrow \\ R_H > R_b \end{array} \quad (12.110)$$

La frequenza di portante cambia più volte durante la trasmissione di un bit, anche la FFH è largamente diffusa per scopi militari.

In altri termini, per trasmettere un bit viene diviso in più più frequenze.

Alla fine, possiamo definire l'ultima FH-SS che è la IFH, in cui:

$$\begin{array}{c} T_H \approx T_b \\ \Downarrow \\ R_H \approx R_b \end{array} \quad (12.111)$$

³⁶Slide del prof — Spettro espanso — pag 76
Appunti di Damiano — pag 76
Appunti — 2025-05-16 — pag 8

FINE PROGRAMMA A.A. 2024/2025

BUON ESAME !!!

